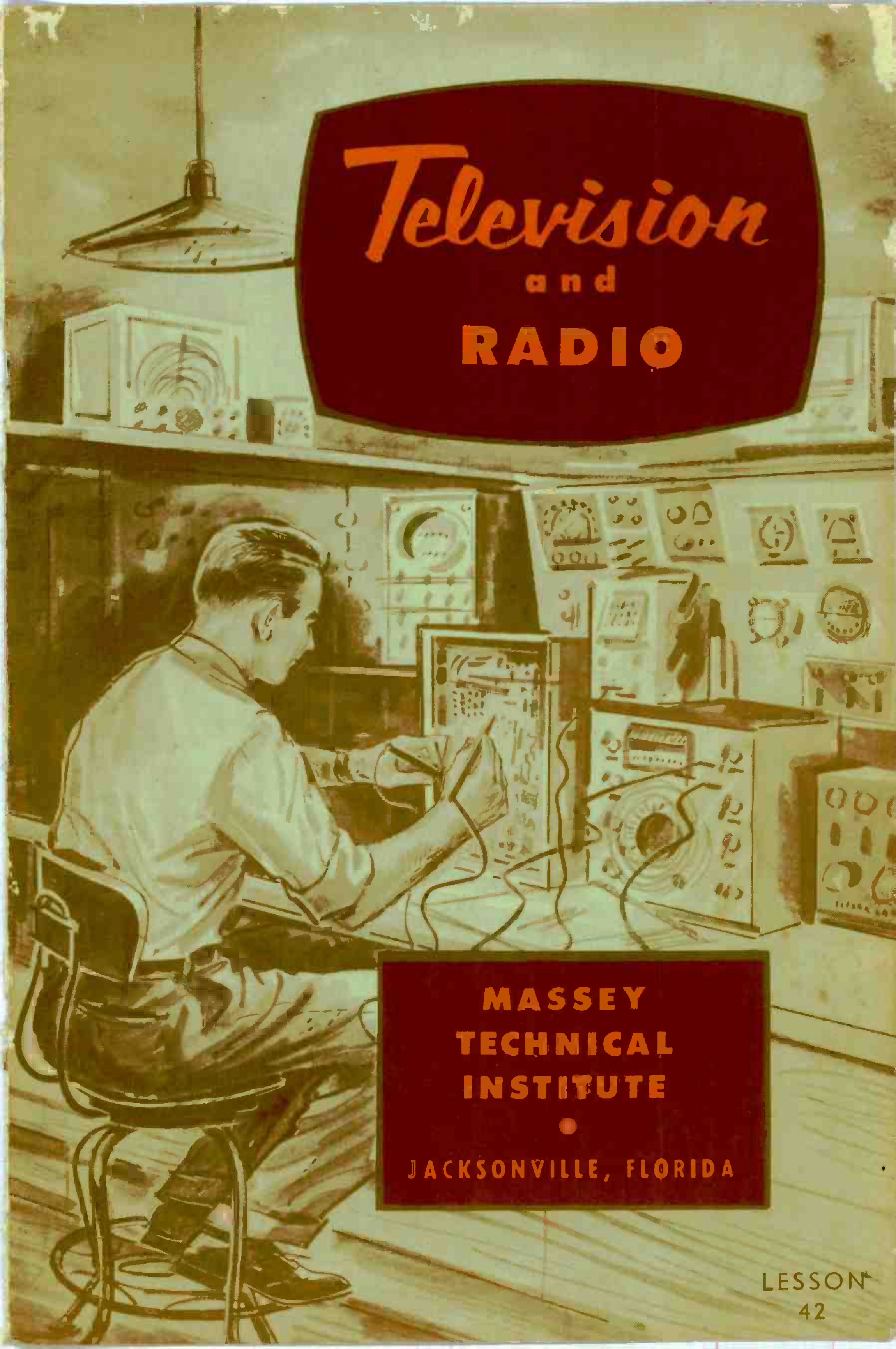


Television and RADIO

MASSEY TECHNICAL INSTITUTE

JACKSONVILLE, FLORIDA

LESSON 42

MOST THINGS WE WORRY ABOUT NEVER HAPPEN.

TRAINING SERVICE

SEMI-CONDUCTORS AND TRANSISTORS

1. INTRODUCTION

The vacuum tube has long served the electronic industry. In fact, it is the very heart of electronics, and when we think of electronics we automatically think of vacuum tubes.

It is natural that this should be so, since each is such an intimate part of the other.

During recent years, however, we have been hearing with increasing frequency about a new electronic component which is taking over some of the functions formerly performed solely by vacuum tubes. We refer to *transistors*.

Due to the fact transistors appear to have certain points of superiority over vacuum tubes for a variety of purposes, some persons have leaped to the conclusion they are likely to supplant, and replace, vacuum tubes. This is by no means correct.

Transistors should be looked upon as supplementing vacuum tubes. Doing some things better than vacuum tubes, and even doing a few things which vacuum tubes cannot do.

On the other hand, there are many fields in which a vacuum tube is vastly superior to transistors, and there are many things vacuum tubes do with ease which are beyond the range of transistors.

Working together they supplement each other nicely. In a few cases, although not many, both

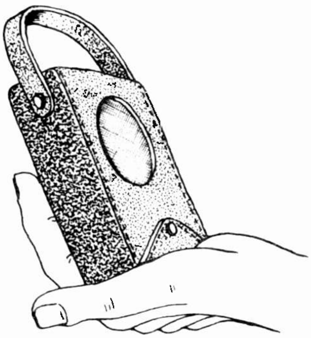


FIG. 1.—*Transistor Radio.*

transistors and vacuum tubes are used in the same electronic device. There are a few cases where it is desirable to utilize the unique qualities of each. The most outstanding example of this is in some of the newer model radios for use in automobiles.

2. BACKGROUND OF SEMI-CONDUCTORS

Semi-conductors are interwoven into the pattern of radio since the very earliest days. Back in the days when many of us were still experimenting with what we then called "wireless," one of the most important components in receiving equipment was a little device which we called a "cat whisker," or "crystal detector."

In our lesson on DETECTORS we touched briefly on the part which crystal detectors played in the early days of radio. It may seem a little strange for us to return to that subject at this time, but there is a good reason for doing so.

A crystal detector is a simple device. It consists merely of a tiny piece of semi-conductive crystal substance which possesses the property of permitting electrical current to flow rather freely in one direction, but which blocks its flow in the opposite direction. This, as we have explained in earlier lessons, is essentially what we call rectification.

The manner in which such a rectifier is placed in a radio circuit to perform the duties of a detector are indicated in Figure 3. It serves to rectify the high frequency R-F carrier signal so the audio component can be extracted.

It is interesting to note that a crystal detector—unlike a vacuum tube—is not a perfect rectifier. The forward resistance is much lower than that of the reverse direction, yet the back resistance does not completely cut off current during the reverse half-cycles.

A vacuum tube is a perfect rectifier, and thus a perfect detector. This means that it does not permit any "back current" to flow during the interval in which the voltage and current are trying to move in the backward, or reverse, direction.

Since a crystal is unable to completely cut off the flow of current in the reverse direction, it cannot be considered a perfect detector, but it is fully capable of reducing the back current to such a low level that it makes a satisfactory detector. For the most part, the minute amount of back current is not a matter of importance, and can be ignored.

However, there are a few situations where even a small amount of back current cannot be tolerated. In such case, a crystal detector could not be used.

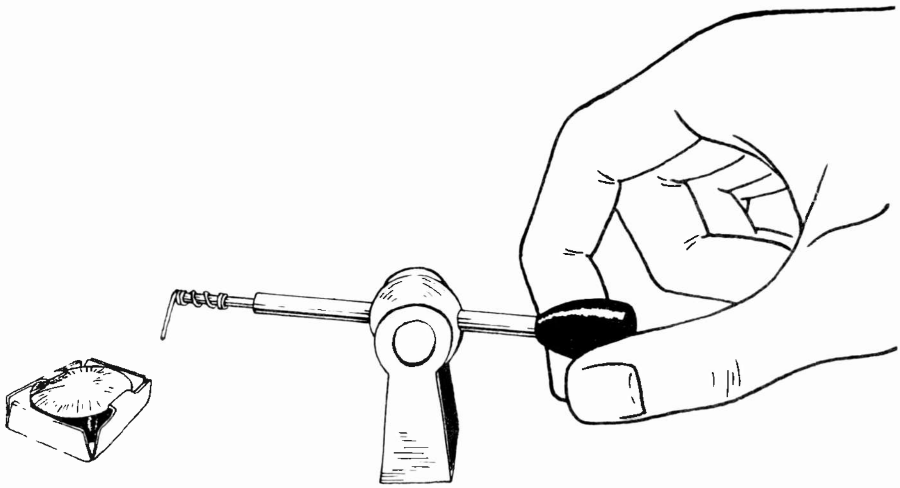


FIG. 2.—*Crystal Detector.*

To obtain a better idea for comparison of the forward and reverse currents in crystal types of rectifiers, we present the series of graphs in Figure 4. The top graph in Figure 4 shows the wave-form shapes of a modulated R-F carrier. This is the way the graph would appear on the screen of an oscilloscope if the modulating signal recurred at regular intervals.

The center graph shows what happens when the signal is applied through a crystal detector, and is rectified. Most of the lower half of the modulated envelope is cut-off, but if you examine the drawing carefully you will see it is not quite obliterated completely.

The bottom graph is intended to show the audio component of the signal after it has been recovered during the detection process.

3. CRYSTAL DETECTORS

Back in the days of "wireless telegraphy" crystal detectors gave good and faithful service. They fell into disuse, however, when vacuum tubes were invented. Vacuum tubes could do more things than the simple crystal, and for the most part were able to do them better.

The early model vacuum tube detector served the dual purpose of demodulating the R-F carrier signal while providing some degree of amplification to the audio signal. This matter of amplification was important at that time because all the amplification it was possible to attain was needed. Because of the added gain, vacuum tubes soon replaced crystal detectors except for a few limited applications.

For many years about the only places where crystal detectors remained in use were in connection with experimental radios which were hand-built by amateur experimenters.

For many years diode detection, such as that provided by crystals, was virtually abandoned. It was

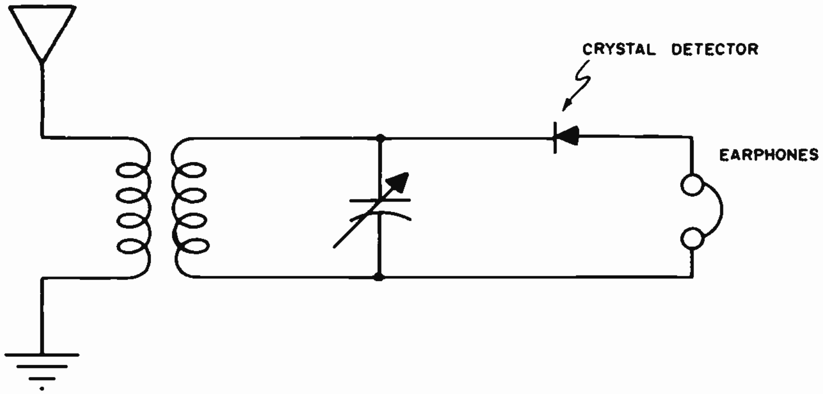


FIG. 3.—*How Crystal Detector is Placed in Circuit.*

only after the development of pentode voltage amplifier tubes that diode detection again became popular. By this time, however, the new diode detectors were nearly always diode *tubes*, not diode crystals.

4. PUTTING CRYSTALS TO WORK IN RADAR

Oddly enough, crystal detectors gradually came back into use. This occurred during the second World War. But this time they were used in Radar equipment, not radio.

It is even more odd that crystals were selected, since radar operates on frequencies far higher than those normally used with radio. For the most part Radar frequencies exceed even those used with UHF television. Such frequencies have extremely short wave-lengths.

In some cases the wave-length of a radar signal is so short that the distance between the cathode and the anode of a vacuum tube becomes an appreciable part of the length of each wave. This means the time required for an electron to travel from the cathode to the anode is a substantial part of each cycle. The result of all this is that the increasingly higher frequency makes it increasingly difficult for a vacuum tube to handle the signals.

When this situation was reached, research men remembered something they had almost forgotten about crystal detectors. Having no anode or cathode, as such, their rectifying action was largely independent of frequency.

Radar engineers faced a peculiar situation. They were working with frequencies so high that the "transient time" of the electrons between tube elements constituted a substantial part of the time involved in each wave-length and each cycle. Before they could begin amplifying the received signal it was necessary to convert it to some lower frequency which could be handled by the available amplifier tubes.

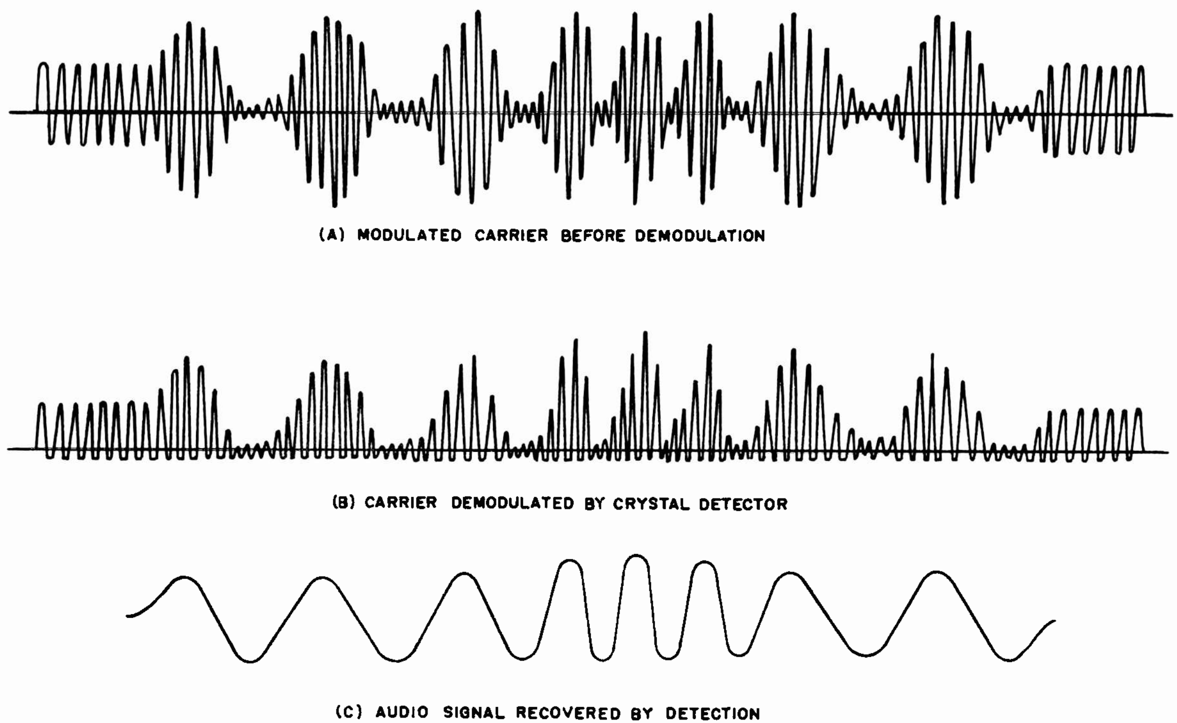


FIG. 4.—*Graphs Showing Detector Action of Crystal.*

Engineers worked out a scheme which is roughly described in Figure 5. They introduced the incoming signal directly from the receiving antenna to a crystal diode. Simultaneously, they introduced to the same diode another signal from a local oscillator. The signal from the local oscillator was usually much lower than that of the incoming signal picked up by the antenna.

Introducing the two signals to the diode crystal in this manner brought about a "heterodyning," or mixing, action. This resulted in the creation of a new *difference* frequency, a frequency low enough so it could be amplified by an ordinary amplifier tube.

This use of crystal diodes gave them a new lease on life, from which has stemmed an incredible number of innovations. One of the most outstanding has been the development of *transistors*.

5. SEMI-CONDUCTORS

In the earlier lessons of this course we introduced you to certain types of materials which are classed as insulators. They are so classed because they do not support the conduction of electrical current. It is customary to say that electrical current cannot pass through them.

Similarly, we introduced you to another group of materials which are classified as conductors. They

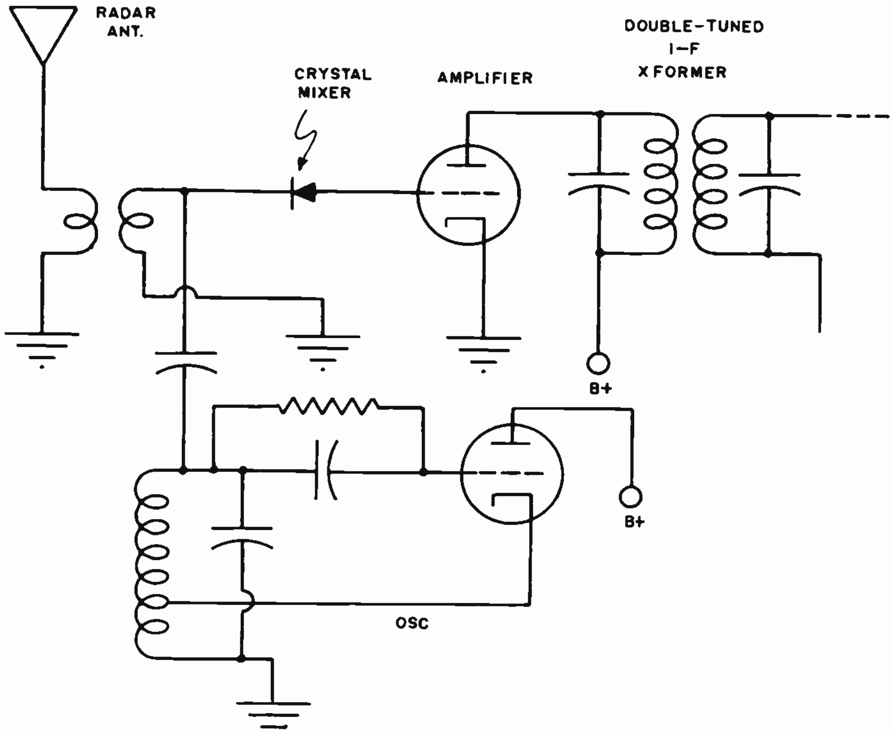


FIG. 5.—*High-frequency Receiver Using Crystal Mixer.*

are so classified because they are capable of permitting electrical current to pass through them. In short, they conduct electrical current.

Up to this point in our lessons we have acted as though all materials can be arbitrarily classified as either conductors or insulators. For the most part, and for practical considerations, this is the way most materials are classified.

However, there is a zone in which we find a few materials which are neither good conductors nor absolute insulators. We refer specifically to those which permit electrical current to flow through in one direction, yet inhibit its flow in the opposite direction.

For the purpose of convenience, scientific men have chosen to classify this group of materials as "semi-conductors." It is from this group that the original "crystal detectors" of early day radio came.

It is also from this same group of materials that modern transistors have been developed. This is a field of electronics which is growing with incredible rapidity, especially so since the urgent need for lightweight electronic components in modern guided missiles is so critical.

Launching of the "Sputnik" satellites by the Russians has given an impetus to this growth which promises still more spectacular things in the future.

In order to understand the functional operation of transistors it is first necessary to know something more about atomic structure, and something of what causes certain materials to exhibit the peculiar symptoms associated with semi-conductors. In short, it is necessary to understand atomic structures, to understand semi-conductors and transistors.

6. SHELLS IN ATOMIC STRUCTURES

What we are about to say is to belabor the obvious, yet it must be said. Crystals which are capable of serving as radio detectors owe their peculiar ability to their atomic construction.

A little thought about this situation makes it clear that such a conclusion is little short of inevitable.

In an earlier lesson we explained the make-up of an atom. We explained how an atom is composed of one or more heavy protons and neutrons which go to make up the nucleus of the atom. We can look upon this as being the heart of all atoms.

Around this relatively heavy nucleus, following fixed orbits, are one or more electrons. For the most part there are approximately the same number of electrons in the orbits around the nucleus as there are protons in the atom.

We say "approximately" since there are occasions when there may be one more electron than proton, or one more proton than there are electrons. Nevertheless, the balance between protons and electrons is maintained fairly constant.

It is the arrangement of the electrons around the nucleus of each atom which determines whether a given crystal of such atoms is a conductor, an insulator, or a semi-conductor.

We are not going to attempt to delve deeply into atomic construction at this time. We believe we can make the action of semi-conductive crystal reasonably understandable without doing so.

At the same time, it is necessary to acquaint you with a few additional facts over and above those we have already given you on this subject. This is necessary in order for you to obtain a reasonably good idea of how transistors work. It is solely for the purpose of explaining transistors that we are going into this subject at all.

Atoms of various *elements* are arranged in a variety of ways. Some have a single "shell" of orbits for the electrons, others have two such shells, but the vast majority of elements have a larger number of shells.

It is the arrangement of the electrons in the *outer* shell of each atom which determines whether the element, of which the atom is a part, is a conductor, insulator, or semi-conductor. Because of this fact most of our attention must be directed toward this outer shell.

Because of the nature of the subject much of our explanation in this lesson may tend to become a bit dry. If it does so, we ask that you bear with us. It is necessary to clearly understand the fundamentals we are trying to explain in this lesson before you can acquire a clear insight into the manner in which transistors work.

We have mentioned previously this matter of the orbital arrangement of electrons into the so-called "shells." Thus, it is not entirely new to you.

Before going any further on this subject we would like to point out that these shells are also referred to in certain quarters as "rings." So, if you find us using these terms interchangeably we want you to know that we are doing so intentionally. It is not done to confuse you; quite the contrary. Since the terms are used regularly in scientific work, we are using them to make you familiar with them.

The more familiar you become with the terms at this time the less trouble you will have with them later. Furthermore, when you pick up other scientific literature, and find these terms being used, you will not be confused by them.

7. ATOMIC STRUCTURES

To further add to any possible confusion you may feel concerning these terms, we would like to inject our own personal comment that none of these names are so precisely accurate as most scientific men would like. However, they serve to convey the general idea, and that seems the best we can hope.

In dealing with atomic structures we are dealing with things which it is hard for our ordinary human senses to grasp and understand. There is so little in our daily experience which prepares us to deal with them.

Further than this, no human has ever actually seen an atomic structure. This means that all our descriptions of them must be based upon indirect observation.

Despite this, we can rest reasonably assured that the theories concerning atomic structures are accurate within a close degree. Everything dealing with atomic power, missile propulsion, electronic action, and many other things are based on these theories.

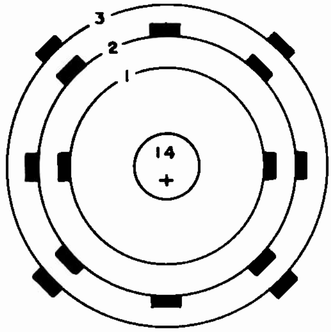


FIG. 6.—*Structure of Silicon Atom.*

The fact these things have been brought into existence as a result of relying on the accuracy of our theories comes pretty close to guaranteeing that they are reasonably accurate.

It has come to be the common practice in scientific circles to describe atoms by using diagrams for that purpose. The diagrammatic drawing in Figure 6 is typical of the drawings used for this purpose. The drawing in Figure 6 is intended to show the atomic structure of an atom of *Silicon.*

Such a drawing or diagram has come to be known as a "conventionalized diagram." It is regularly used by scientists and physicists because it provides a convenient method of accounting for, and showing the relationship among, each of the electrons and protons in the atom.

The drawing of the Silicon atom shows two electrons in Shell No. 1. It shows eight electrons in Shell No. 2. Finally, it shows 4 electrons in Shell No. 3.

It is a peculiarity of atomic structure that the inner shell, commonly called shell No. 1, usually has only two electrons. This is true of all electrons except that of hydrogen. Shell No. 2 has eight electrons. All the other shells have eighteen electrons each, except the outermost shell.

The outermost shell has a varying number of electrons, depending on the atomic weight and structure. In the case of Lithium, the second shell is the outermost, thus does not have the full quota of eight electrons.

Referring back to Figure 6, we find the shells contain a total of 14 electrons. These 14 electrons are electrically balanced by 14 protons in the nucleus. This is indicated by the Figure "14" in the center of the drawing.

While we intend to use the same system of diagramming atoms which other scientific men regularly employ, we want to make it clear that this method of diagramming the structure should not be understood as being a precise drawing of the actual action. You should not acquire the idea that one electron chases another around the shells like a bunch of puppies chasing each other's tails. Such a conception would be neither accurate nor true.

The shells should be thought of as being layers around the nucleus of an atom. The various shells surround the nucleus in a manner something like the various rings, or layers, of an onion surround the center.

If the atomic shells are then considered as being hollow areas

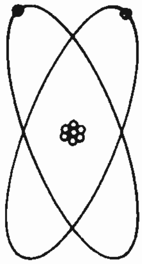


FIG. 7.—*Electrons in Shell No. 1.*

of space within which electrons are free to move, you will be acquiring the proper conception of just how they are.

Within each of these shells, except the innermost and the outermost, electrons sweep around and around in their orbits in a manner closely resembling the manner in which a Sputnik satellite swings around the Earth in its orbit.

In fact, the manner in which a Sputnik continues to follow its orbital path in space is closely related to the manner in which an electron follows its orbit within the shell of an atomic structure. The first Sputnik was placed in its orbital path because Russian scientists followed the same principles which occur in atomic structures.

The drawing in Figure 7 is typical of the manner in which an atom is commonly diagrammed when it is desirable to show the orbits of the electrons as well as the nucleus. The drawing shown is that of an atom with a single shell, and only two electrons.

A diagram of a slightly more complicated atom is shown in Figure 8. The diagram in Figure 8 is intended to represent an atom with two shells, although the diagram is obviously incomplete. All the electrons are not shown.

Note that electrons in the inner shell are relatively close to the nucleus of the atom. Note further, although electrons in the inner shell are moving within what is a definitely fixed *layer* of space, the orbits of each electrons does not follow the same path as that of the other electron in the same shell.

Much the same is true of the electrons in the outer shell. Those electrons follow orbits which are located within the same shell space surrounding the nucleus; nevertheless, the orbits of the individual electrons do not follow the same path.

These points we are emphasizing are not critically important. At the same time we would like for you to get them clearly in your mind, since they are items which cause some degree of confusion among students to whom they are not clearly explained.

If you think of an atom as resembling an onion, with each ring of the onion corresponding to a shell of an atom—a shell of empty space—then can imagine designated numbers of tiny objects whirling around the center of the onion, but within the limitations of each ring, you will have a pretty good idea of an atom. We agree that this is a most rough comparison, yet it comes close to providing a working idea of the make-up of an atom. After all, that is the most we can hope to do, acquire a working idea.

Figure 9 is an illustration of a segmented onion, which helps convey the idea we are trying to explain. Each layer of the onion corresponds to a shell in an atomic structure.

Although electrons are arranged in their respective shells around the nucleus of an atom in the manner described in Figures 7 and 8, for the sake of convenience scientific men have found it more practical to utilize diagrams on the order of the one shown in Figure 6. Such a diagram enables them to describe what occurs within the atomic structure, and do so with great clarity.

We can carry the idea presented in Figure 6 somewhat further by considering an atom of Germanium. An atom of Germanium is somewhat more complicated than one of Silicon, it has more shells.

The diagram in Figure 10 shows how a scientist would describe an atom of Germanium. Note that it has four shells. No. 1 has two electrons, No. 2 has eight electrons, shell No. 3 has its full quota of eighteen, and the outer orbit has only four electrons.

It is that outer shell, and its complement of electrons, which provides a Germanium atom with its pecular electrical properties which result in what we call a semi-conductor. It is the peculiar action which revolves around the electrons in that outer shell of the Germanium toward which we have

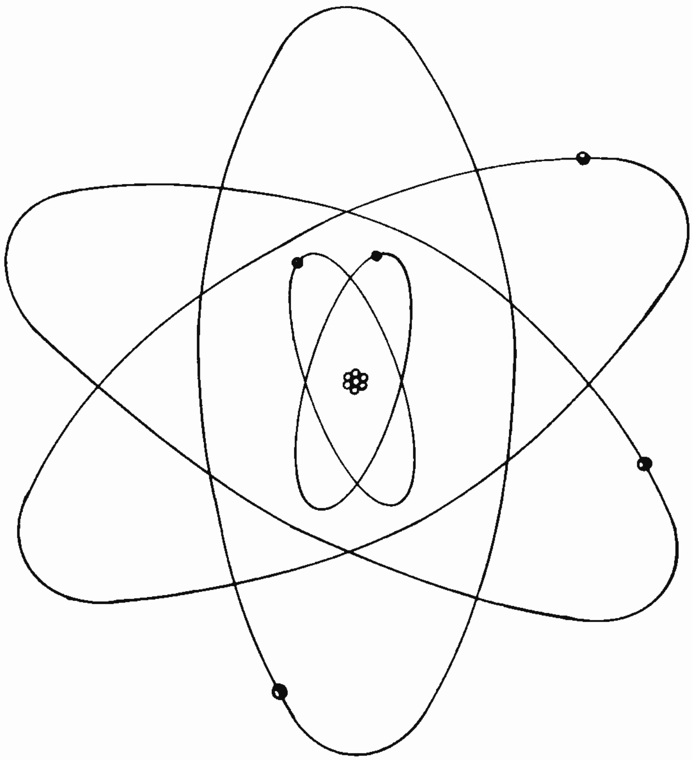


FIG. 8.—*Electrons in Two Different Shells.*

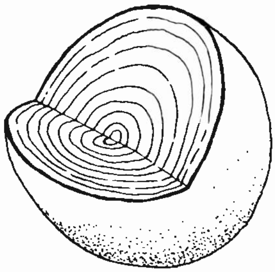


FIG. 9.—*Segmented Onion Showing Layers Which Correspond to Shells of an Atom.*

been aiming during this rather involved discussion of atomic structures.

Scientific men have found it rather difficult and cumbersome to draw illustrations like those in Figures 8 or 9 when trying to describe some action with respect to an atomic structure. It is for that reason they have chosen to use diagrams similar to those in Figures 6 and 10. The latter method of diagramming is more simple, yet conveys to a trained mind the same amount of information.

We are going into detail concerning this matter of diagramming so you will understand the reason behind the differing methods of describing atoms and their structures. Without an explanation of the reasoning it would be impossible for you to become a bit confused by the varying methods of presentation.

Before going any further we want to pass along another bit of advice concerning your attempt to visualize the actions of electrons in an atomic structure. If these things were not pointed out to you it would be easy to fall into error in your attempts to visualize what is going on inside them.

Drawings and diagrams prepared in the manner of those in Figures 6 and 10 tend to suggest the electrons within the structures are stationary in the position in which the drawings locates them. Such, of course, is not true.

Electrons are constantly moving within their orbits, the orbits of each electron being fixed within one of the shells surrounding the nucleus. It is only for the purpose of explanation of the arrangement that we pretend that all movement is momentarily frozen and suspended during the interval necessary to make our explanation.

While on this subject we would like to point out that it is not likely you will pursue the study of atomic structures beyond the elementary level. At least it will not be necessary to do so to support your knowledge of electronics.

This being true it will not be necessary for you to understand all the fine details of the structures, nor to delve into the subject as deeply as is necessary when studying some of the other branches of science.

It is still true, however, that you will undoubtedly want to read additional technical books and magazines touching on the subject of semi-conductors, especially as they concern transistors. This requires that you acquire a reasonably good *general* knowledge of atomic structure.

8. EARLY HISTORY OF CRYSTAL DETECTORS

We are directing our purpose toward acquiring a basic knowledge of transistors. To help us

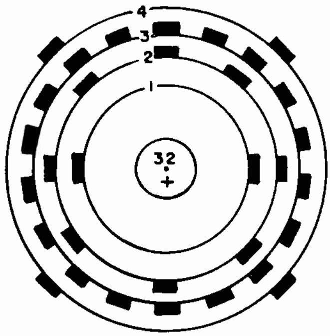


FIG. 10.—*Structure of Germanium Atom.*

understand something of the background of these intriguing little electronic components, it is desirable that we go back to the beginning of radio and direct a little of our attention toward natural crystals which were used as one of the original radio detectors.

There are four types of crystals which can function as a radio detector. The two most widely used for this purpose were *Silicon* and *Germanium.* It is a reasonable assertion to say that they are the only two which were ever used to any great extent for that purpose.

Although the element Silicon is relatively unfamiliar to many persons, it is, strangely enough, one of the most plentiful elements on Earth. It is the second most plentiful element; only oxygen being more plentiful. It has been known to scientists, chemists and research men for more than a century.

Despite the fact it is so plentiful, it is never found in a pure state. It is so active it is always found combined chemically with some other element.

The best known forms of Silicon include combinations of Silicon Dioxide. These are certainly the most plentiful forms. For the most part it takes the form of sand, flint and quartz.

The first known record of Silicon being separated into a pure form was accomplished in 1823. A Swedish scientist named Jons Jakob Berselius was the man who first succeeded in reducing Silicon to a pure form.

The element Silicon had long been suspected of existing, but it was not until Berselius isolated it in pure form that its existence was actually proven.

Back in Figure 6 we gave you a diagram of a Silicon atom. The atom has three shells within which the 14 electrons of the atom are free to revolve in their orbits.

As we pointed out a little earlier in this lesson, the inner shell contains the orbits of two electrons. The second shell contains the orbits of eight electrons. This is the pattern of the atomic arrangement in all other atoms having more electrons than those of Silicon.

It is the outer shell of Silicon which comes in for special attention, and is the factor that sets it apart from other elements insofar as its electrical action is concerned.

The normal arrangement of atomic structures is for the third concentric shell to contain the orbits of 18 electrons. Such arrangement provides what scientists call a *stable arrangement.*

They mean by the term "stable arrangement" that such an atom maintains its electrons in fixed and orderly positions. This arrangement prevents electrons drifting from one atom to another.

Remember—in the case of silicon, each atom has only four electrons in the outer, or third, shell.

Electrons in the outer shell of silicon do not have the freedom to move from atom to atom as is true of metallic elements like silver, copper and iron. On the other hand, the electrons are not so firmly fixed in their orbits within the outer shell as is the case with true insulators.

This puts Silicon in a sort of twilight zone, existing between conductors and insulators.

There are certain conditions under which it is possible for an atom of Silicon to accept an extra electron from an adjacent atom of the same element. Likewise, there are other conditions which cause a Silicon atom to give up one of its electrons.

This peculiar ability is modified when Silicon is adulterated, or contaminated, by certain other types of substances or materials. We find this is especially true if the adulterating substance is arsenic, antimony or boron.

Contamination can also occur from Gallium or Indium. In this case, however, the resulting action is different from that which occurs with arsenic or boron.

Both Silicon and Germanium were used as early day radio detectors. Germanium is much more scarce than Silicon, but is more suitable for certain uses in electronic work.

If you study the diagrams in Figures 6 and 10 you will find that the atoms of the two elements are apparently quite different, yet they have one thing in common. Both have exactly 4 electron orbits in the outer shell.

Existence of Germanium as an element was suspected for a long time, but it was not until 1886 that it was proven to exist. It was in that year that a German scientist succeeded in separating it into a pure form.

The German scientist, Clemens Alexandre Winkler, derived the pure Germanium from a complex ore called Argyrodite. This ore was obtained from a mine dump in Saxony in Germany. He gave the Latin name for Germany to the new element he had isolated.

Since that time Germanium has been found among other metallic ores. There are several sources of the ore in this country, one of the most important being in Arkansas.

Early wireless experimenters constantly sought to find new and better "detection devices" for use with their equipment. Among those who kept looking for better methods were all the oldtimers, Hertz, Marconi, Dunwoody, DeForest and Pickard.

In fact, it was DeForest's efforts to secure a better detecting device that led to the invention of his three-element *Audion* tube. He was not trying to invent an amplifier, he was trying to develop a better detector.

Due to the excessive gaseous conditions inside his imperfectly evacuated glass tube, his first audions were much better detectors than they were amplifiers. It was not until other scientists provided more highly evacuated tubes that the three-element tube's ability to amplify was discovered.

Marconi's method of "detecting" radio signals was to use a *coherer,* a device which is now all but forgotten. It was little more than a glass tube filled with iron filings. The coherer was severely limited because it was unable to receive any kind of signals except telegraphic codes.

History seems to credit Henry Harrison Dunwoody, who was then working with the DeForest Wireless Company, with being the first person to use a metallic crystal as a radio detector. His method was to place a piece of carborundum in contact with brass holders.

But it was Greenleaf Whittier Pickard who first used the more practical method of contacting a crystal held in a mount with a tiny "cat whisker." His patent covering that method was filed in 1906, almost the same time that DeForest was perfecting his Audion.

Pickard used Silicon crystals, then placed an extremely finely pointed wire under stress to press against the surface of the crystal.

To demonstrate the odd patterns followed by history—and the development of science—it was this same process of pressing a fine point of wire against a crystal which was used some forty years later when the first Transistor came into existence.

Getting back to Germanium, it was a good many years after Winkler discovered the element in a German mine dump before it was discovered to possess the peculiar properties needed to make a radio detector. It was a Swedish inventor who discovered Germanium's rectifying ability which made it useful as a radio detector.

The Swedish inventor, Carl Axel Benedicks, made the discovery in 1915 during the height of the First World War. The discovery came more or less by accident.

Prior to that time very little was really known about Germanium. Very little of the mineral was in existence, and it was so expensive to refine that little was done to find uses for it. With the discovery that it made an excellent radio detector, however, it soon came to have wide use for that purpose.

9. FROM CRYSTAL DETECTOR TO TRANSISTOR

During the period following adoption of vacuum tubes for radio work, crystals fell into disfavor. Most scientists, experimenters, and even radio engineers lost interest in them.

Not everybody forgot them, however. A few experimenters who liked to build radios which could work without outside power, continued to play with them. But such interest was extremely minor.

It was the peculiar needs of radar which rekindled the interest of radio men in crystal diodes, or rectifiers. As mentioned earlier in this lesson, crystals were put to use in the front end of radar receivers to convert the extremely high radar frequencies to lower frequency levels which could be handled by vacuum tubes.

It was from that renewed interest, spurred on by the needs of radar, which resulted in the invention of transistors.

The first basic patent on transistors was issued by the United States patent office on June 17, 1948. It was issued in the names of John Bardeen and Walter H. Brattain. These two men headed the research teams at Bell Telephone Laboratories which developed the transistor.

It would be well for us to point out that the patent papers were taken out in the names of these two men for technical and administrative reasons. Both men played important parts in the development and invention of the transistor.

Yet, both men would be among the first to assert that they were not solely responsible for the invention, nor that they should be given full credit for it. The transistor came into existence as the result of team effort on the part of a vast number of highly trained scientists, engineers and technicians.

It would be well to point out that most scientific research and invention at this time are largely a matter of teamwork. Few present day advances in the field of electronics result from the starving garret inventor, as was so true in the earlier days.

With teams of highly skilled and trained scientists and technicians working together, advances in electronics are being made in leaps and bounds rather than the slow groping which existed during the early days. This makes things better for everybody, and opens tremendous opportunities which might not otherwise exist.

The action of a transistor consists very largely of varying the surface charges on a high-back-voltage rectifier by controlling the density of the surface charges through the addition of a cat-whisker contact. In effect, what is occurring is a matter of transferring the resistance.

These crystal devices were first termed *transfer resistors*. What had been TRANSfer resISTORS became TRANSISTORS. Thus the present name was coined from parts of the two words in the original name.

It is by the name *Transistor* that they are now known everywhere.

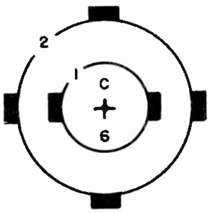
10. HOW ATOMIC STRUCTURE AFFECTS CONDUCTIVE ACTION

We are now about to embark on our technical discussion of transistors. It seems only fair to warn that some of the terms to which we will be introducing you will be totally new. So will many of the words.

In fact, some of the terms used in connection with transistors were actually invented in an effort to describe the action which takes place in them. In other cases new meanings—new *technical* meanings—have been given old words.

Insofar as we are concerned, it is entirely probable that we could describe the general actions of transistors without employing these terms. However, that would not be entirely fair to you.

It is probable that you will prefer to continue your studies of transistors by reading and studying magazine articles, and will want to read and study bulletins covering technical information relating to transistors. If you do, you will find that most articles on



RING NO. 1–2 ELECTRONS
RING NO. 2–4 ELECTRONS

FIG. 11.—*Atom of Carbon.*

this subject are written in such manner that these technical terms and words are used.

In order for you to acquire as much benefit from such articles and bulletins as is desirable, it seems best that we follow the same pattern and practice. We will try to explain each of the new terms as we go along.

One of the first terms to which you are introduced when working with transistors is what physicists call a *valence bond.*

Rather than attempt to define a valence bond in a formal definition, we think it more desirable to provide an example. We think you can arrive at the meaning of the term in that manner more easily than through a mere definition alone.

We will start our explanation with an atom of carbon. A diagram of an atom of carbon is shown in Figure 11.

Carbon takes several forms. The atom of carbon shown in Figure 11 is the form taken by carbon when it is a diamond.

Diamonds are not used in transistor work, of course; they are too expensive. Technically, however, they possess the possibilities which could be used in transistor work.

The atomic structure of an atom of carbon when in the form of a diamond is closely similar to, and related to, atoms of Silicon and Germanium. Like Silicon and Germanium, an atom of carbon in the form of a diamond has four electrons in its outer shell.

A diamond's atomic structure is extremely simple. It is for that purpose that we are using it to demonstrate a transistor principle. This should help make the underlying principles of transistor operation easier to understand.

An atom of carbon in the form of a diamond has only two shells. The inner shell has two electrons, the outer shell has four electrons. There are no other shells.

We have provided an imaginative arrangement of five carbon atoms in Figure 12. We want to base an explanation of the lattice structure on these atoms. We are using an atom of diamond because of its simplicity, yet, which is closely similar to the lattice arrangement in the outer shells of the atoms of Silicon and Germanium.

Atom No. 1 is shown in the center of the diagram. The interesting thing about diamond atoms is that each of the four electrons in the outer shell of each atom is shared by another nearby atom.

Note this carefully. It is the heart of the transistor action. It is what we mean by *valence bonds.*

Each atom in Figure 12 has four electrons in the outer shell, but in each case one of the electrons is shared by a nearby atom.

Each of the four electrons of atom No. 1 is shared with a near-by atom. This means that each atom has the correct number of electrons in its outer shell, but in order to have the correct number it is necessary for the atom to share electrons with other close-by atoms.

Getting down to details, note that one of the electrons in the outer shell of atom No. 1 is shared with atom No. 2. While this is going on, one of the electrons in the outer shell of atom No. 2 is shared with atom No. 1.

The same thing is occurring between atom No. 1 and atom No. 3. These two atoms also share another pair of electrons.

Going another step further, you will note the same arrangement exists between atom No. 1 and atom No. 4, as well as between atom No. 1 and atom No. 5.

When this arrangement exists we find atom No. 1 behaving as though its outer shell possessed its full complement of eight electrons. Although it has only four electrons of its own, we find a total of eight electrons entering and leaving the outer shell.

Keep in mind that electrons in their orbits move in elliptical paths. At one point in the orbit they swing in closer to the nucleus. At the other side of the orbit they swing farther away.

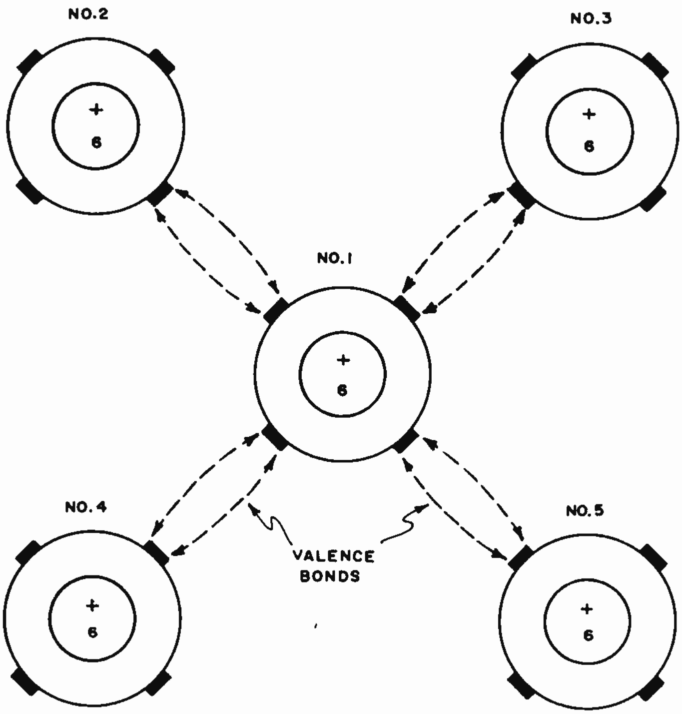


FIG. 12.—*Valence Bond in a Tetravalent Atom Lattice.*

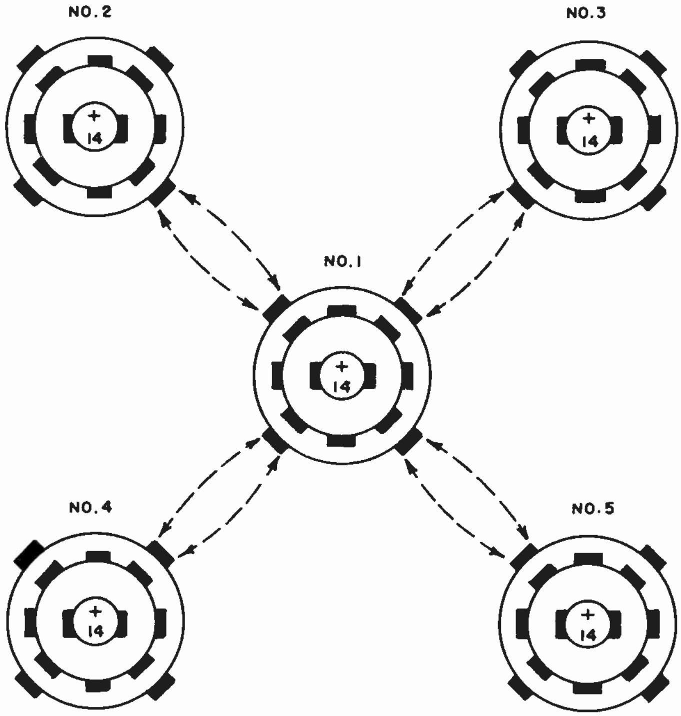


FIG. 13.—*Tetravalent Atom Lattice of Silicon.*

It is this elliptical shape of their orbits which permit atoms to behave as though they had more electrons in their outer orbits than they actually have. Two electrons follow an elliptical path which causes it to encircle two atomic nuclei, rather than a single nucleus. One pair encircles the nuclei of atoms No. 1 and 2. Another pair encircle atoms No. 1 and 3.

And so it goes.

Keep in mind that we are describing the action in a very small portion of a particle of carbon. We are directing our attention to only a very few atoms and their electrons.

While the pattern described in this extremely small area is occuring as we have explained, similar actions are occurring during the same instants of time in all other parts of the diamond particle.

While electrons move in their own orbits within the shells of two atoms, they involve only those two atoms. Neither of the electrons move into the orbits or shells of any other electrons or any other atoms.

Since the electrons are not free to move from one atom to another —except the two atoms to which they are bound through their valence bond—such atoms cannot be used as conductors of electricity. They are, in fact, insulators.

The term *valence bond* is used in chemistry to describe a peculiar action which is constantly occurring in and among atoms and molecules of an element. The term is used similarly in physics.

The term, actually, is used to describe this action of one or more pairs of electrons binding together a pair of atoms into a bond.

It is this binding together of two atoms which is referred to as a *valence bond.*

By referring back to Figure 12 we find that this binding action also serves to bind the central atom to each of the other atoms.

By giving our imagination a little free rein it is not hard to visualize that each of the four outer atoms, Nos. 2, 3, 4 and 5, in Figure 12, are also bound in a similar manner to still other atoms which are not shown on the diagram. And these other atoms, in turn, are bound to still others. Thus the bond action continues in all directions without limit.

So long as the electrons and atoms of a substance are arranged as shown in Figure 12, that substance is an insulator. There are no free electrons to move from atom to atom.

11. SEMI-CONDUCTOR VALENCE BOND STRUCTURES

We used an atom of diamond carbon in Figure 12 to describe the valence bond which exists among atoms in certain types of materials. The diamond atom was used solely because it is the most simple type of atom which has four electrons in the outer shell.

Now let us turn our attention to another type of atom, one which possesses the same characteristics but is more practical for electronic work. This is the Silicon atom. A group of five Silicon atoms are shown in Figure 13, with the five atoms being held together through valence bonds in a manner similar to that which existed among the five diamond atoms in Figure 12.

Silicon in its pure form is an electrical insulator. It is an insulator because it does not possess free electrons in its outer shell which are needed to make it a conductor of electricity.

Despite the fact that in its pure state Silicon is not a conductor, it can be changed into a semi-conductor by adding small quantities of impurities. By controlling the kinds and amounts of the impurities added to Silicon it is possible to control the degree of its conductivity.

If the new and different atoms of the adulterating material has an extra "free" electron in their outer shells the adulterating atoms are referred to as *donor* atoms. Materials such as arsenic, antimony and boron are classified as being *donor* atoms.

Keep this word "donor" clear in your mind. You can think of donor atoms as *donating* electrons to the atoms of silicon in order to make the Silicon a semi-conductor.

Since Silicon has what we have already explained as being called

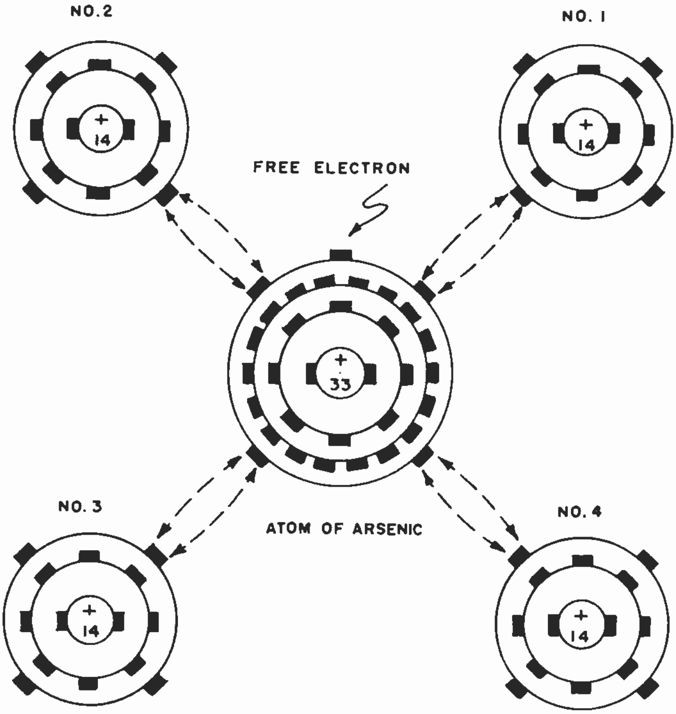


FIG. 14.—*Four Silicon Atoms Combined with a Penta-valent Atom of Arsenic.*

tetra-valent lattices in its atomic structure, each outer shell of each atom has *four* electrons.

Whereas Silicon has four electrons in its outer shell, arsenic has *five* electrons. This makes arsenic a *penta-valent* element.

Keep it clearly in mind that these actions all occur because of the *number* of electrons in the outer shells of each of the atoms of the various elements. Four electrons create a *tetra-valent* lattice arrangement; five electrons create a penta-valent arrangement.

When a small amount of arsenic is mixed with Silicon there is a resulting combination of their atoms in a manner which can be diagrammed as shown in Figure 14.

Note what happens. Four of the electrons in the outer shell of each atom of arsenic create valence bonds with adjacent atoms of Silicon. So far as this goes the action is closely similar to that which would occur if all the atoms were tetra-valent in nature.

But look at that fifth electron in the outer shell of the arsenic atom. It does not form a valence bond with any of the other atoms, therefore remains free. Under certain electrical conditions it is possible for this free electron to move from that atom to another. When this occurs we have what

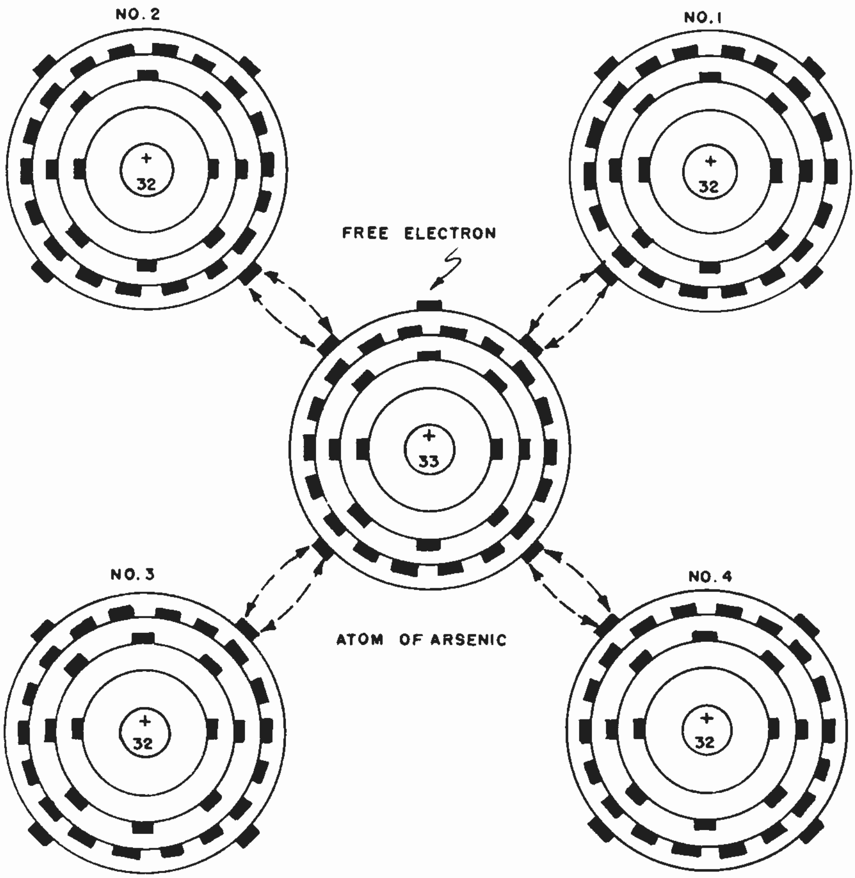


FIG. 15.—*Four Germanium Atoms Combined With an Atom of Arsenic.*

amounts to partial electrical conduction.

You will note that the normally insulative Silicon has been changed into a semi-conductor by the addition of a small quantity of arsenic atoms.

This is the first step in making a semi-conductor from Silicon. It is also the first step in the creation of a transistor.

12. SEMI-CONDUCTOR ELECTRICAL PROPERTIES

A semi-conductor is distinguished from an ordinary electrical conductor or insulator by its peculiar ability to act as a conductor in one direction while acting as an insulator in the opposite direction. You will recognize this property as being the essential ingredient of a rectifier.

Although it was not known at the time they were being used for that purpose, we now know that crystals owed their ability to "detect" radio signals to this fact. The early day crystals were accidentally adulterated by nature, and radio men took advantage of their peculiar abilities without knowing what contributed to that ability.

Although we have discussed Silicon as though it were the principal element in which this condition occurs, we should not overlook Germanium. In point of fact, Germanium is used even more widely than Silicon, although late research is being conducted with Silicon which is making it constantly more useful.

Germanium atoms have a somewhat more complex atomic structure than Silicon. It has more shells, and a more complex atomic structure. In Figure 15 we show a single atom of arsenic combined with four atoms of Germanium to bring about the same condition which we have already described in connection with four atoms of Silicon. Note the similarity of the action in Figure 15 with that which we described in Figure 14.

Note the many additional electrons in a Germanium atom. Whereas an atom of Silicon had only 14 electrons, we find an atom of Germanium has 32 electrons. This gives Germanium a much greater atomic weight than Silicon.

However, the atom of arsenic has 33 electrons, which again provides the Germanium with a free electron when adulterated with the arsenic atom.

We could carry this discussion into Selenium, which possesses properties similar to those of Silicon and Germanium. However, the action is so closely similar to that we have described in connection with Silicon we doubt there is any need to go into it—at least, not right now.

All three elements are used in electrical work. Their exact characteristics differ somewhat from each other, which causes each to be used in special ways. Silicon and Germanium have found their greatest usefulness in connection with so-called "crystal diodes," whereas Selenium is used most frequently to make large capacity rectifiers.

However, Selenium is also used in connection with certain types of photoelectric work. In fact, layers of Selenium were the first "electric eyes," so-called.

Selenium is used most often where the frequencies are relatively low. Silicon, on the other hand, is used more widely on the higher frequencies. Nearly all the crystal diodes found in tuner sections of UHF television receivers, for example, are made from Silicon.

Germanium possesses still different characteristics. It is also used for high-frequency crystal diodes, but it possesses qualities which make it especially suitable for transistors, and phototransistors.

We have directed most of our attention to the effect of impurities in the atomic structure makeup in giving to certain crystals the peculiar property of uni-direction conductivity, or making of them semi-conductors. But there are other factors in addition to

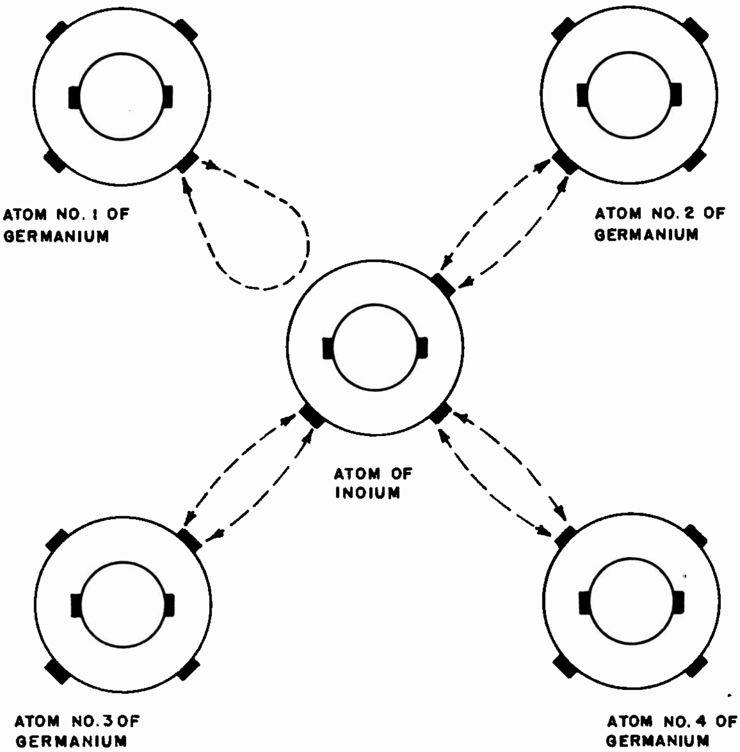


FIG. 16.—*Unsatisfied Valence Bond Which Exists When Adulterant Atom Has Few Electrons in Outer Shell.*

that of impurities which affects conductivity.

One of these, which we have already mentioned, is light. Light has a very strong effect on the conductivity of certain types of crystals.

Another is heat. Some crystals are so sensitive that they must be kept within carefully controlled temperature ranges if they are to function properly. This sensitivity to heat is one of the things which tends to limit the usefulness of transistors. Some of them quit working when the temperature exceeds certain limits. In fact, some are ruined when the temperature rises to levels which are too high.

Sensitivity to light of certain types of these semi-conductive crystals has made them useful in photoelectric work. This has served to open entirely new areas in which they can be used.

Another thing which affects their conductivity is the presence of electrical fields on their surface.

It is this ability which led to the development of transistors.

In the case of transistors an electrical field is deliberately placed on the surface of the crystal, and the presence of this electrical field affects the conductivity of the crystal. More will be explained about this property as we go along.

13. ACCEPTOR ATOMS IN SEMI-CONDUCTORS

We have already pointed out that adulterating Silicon and Germanium with minor amounts of arsenic, antimony or boron provides what are termed *donor* atoms. These are atoms which possess an extra electron in the outer shell—one more electron than is possessed in corresponding atoms of Silicon and Germanium.

Now we find ourselves considering the opposite condition. We now turn our attention to the situation in which an adulterating atom has *one less* electron in the outer shell than the atoms of Silicon and Germanium. The diagram in Figure 16 shows such a situation.

One of the most common adulterants is Indium. An atom of Indium is called an *acceptor* atom, because it has one less electron, and is capable of accepting electrons from the atoms of Silicon or Germanium with which they are mixed.

The diagram in Figure 16 shows the action. Whereas Germanium has four electrons in the outer shell, Indium has only three electrons. This creates an interesting situation. It is one that has a direct bearing on the operation of transistors.

Note that valence bonds have been created between the atom of Indium in Figure 16 and atoms Nos. 2 and 3 and 4 of the Germanium. But there is no valence bond between atom No. 1 and the atom of Indium.

This leaves one of the valence bonds unsatisfied.

Now we come to a peculiar situation with respect to transistors. Since an electron is missing, and nothing but empty space exists where the electron would normally be, physicists reason that a "hole" exists.

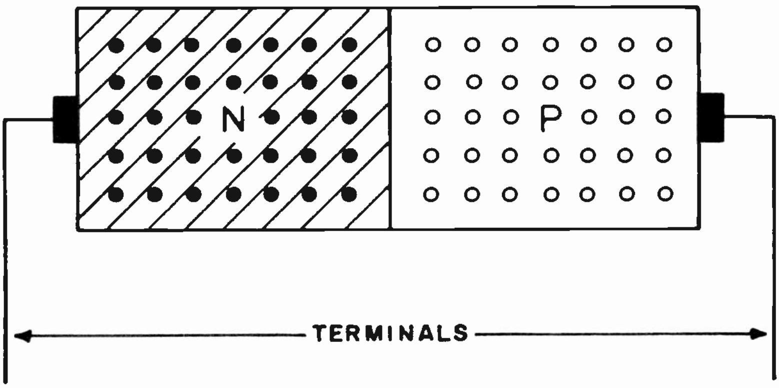


FIG. 17.—*Positive and Negative Crystal Joined Together.*

This matter of a "hole" has come to assume an extremely important place in transistor work. Physicists and scientists treat that "hole" as though it had a physical existence. They talk of these "holes" moving from one place to another, and constituting a flow of electrical current.

This means you will have to become accustomed to thinking of them as having physical presence. It is a bit difficult to imagine an empty space—a "hole"—as having physical existence, yet learned men treat these "holes" in this manner.

Under many conditions it is the movement of these "holes" from point to point which constitute an electrical current flow.

Researchers have proven to their own satisfaction that, under the influence of electrical voltages, holes in semi-conductor crystals are capable of moving from one atom to another.

14. POSITIVE AND NEGATIVE CRYSTALS

Semi-conductive crystals can be endowed with *positive* characteristics, or with *negative* characteristics. Whether a crystal is positive or negative depends on the type of adulterant used with it.

If the crystal is adulterated with a *donor* atom, it has an excess of electrons, thus becomes a *negative* crystal.

On the other hand, if it is adulterated with an *acceptor* atom, it becomes a *positive* crystal.

This means that a Germanium crystal which has been adulterated with Arsenic is a negative crystal. But a Germanium crystal which has been adulterated with Indium is a positive crystal.

You may feel we are placing just a bit too much emphasis on these matters right at this time, and that we are delving too deeply into atomic structures. However, as you progress a little further with your studies you will find that these things assume real importance. Much of the manner of action of a transistor depends on the manner in which positive crystals are combined with negative crystals.

The important thing you should remember, is that these crystals have two basic types of semi-conductive characteristics. One type has positive characteristics. The other has negative characteristics.

15. RECTIFICATION

It would be interesting to go into a full explanation of all the things which occur in a semi-conductor during the process of rectification. However, for our present purpose that seems neither necessary nor desirable. There is too much chance that concentration on that point could divert your attention from the one thing which is important right now.

At the same time, there are certain things we want to make clear to you.

Rectification occurs when two crystals, with different polarities, are joined together. The joining together can be in the manner diagrammed in Figure 17.

We have previously explained that a Silicon or Germanium crystal can be caused to have a positive polarity when adulterated with a material which has *acceptor* atoms. This means that a crystal adulterated with a small amount of Indium, for example, would be a *positive* crystal.

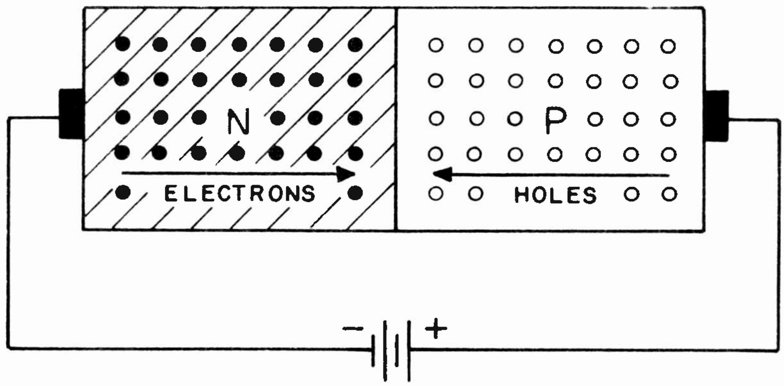


FIG. 18.—*Applying a Voltage to Joined Dissimilar Crystals Starts an Action.*

We also explained that when a Silicon or Germanium crystal, and this includes Selenium crystals, is adulterated with arsenic, boron or other similar adulterants having *donor* atoms we wind up with *negative* crystals. It is when crystals having opposite polarities, such as shown in Figure 17, are joined together that we have a set-up for rectification.

The crystal at the left side of the diagram in Figure 17 is marked with the capital letter *N*. This designates it as having a negative polarity.

The crystal at the right side is marked with the letter *P*, to indicate that it has *positive* polarity.

The mere act of joining together two such opposite crystals does not automatically produce any special activity. It is only when an external voltage is applied to them that action occurs.

From our many previous discussions you know that a crystal, or any other substance, which has a negative polarity is one which has a surplus of electrons. The mere matter of negative polarity is closely linked with a surplus of electrons.

When it comes to the positive crystal we must change our attitude of mind slightly.

We know, of course, that a positive polarity refers to something which is *deficient* in electrons. It means the normal supply of electrons is missing. There are fewer electrons in each atom than there are protons.

In transistor work, however, technical men tend to view the situation in a slightly different manner.

If an electron is missing it must mean there is an empty space in the atomic structures of the material. An empty space is a "hole."

All of which means that a positive crystal is one which has within itself a quantity of "holes." It is this concept of "holes" which sets apart semi-conductors, such as transistors, from other types of electrical conductors.

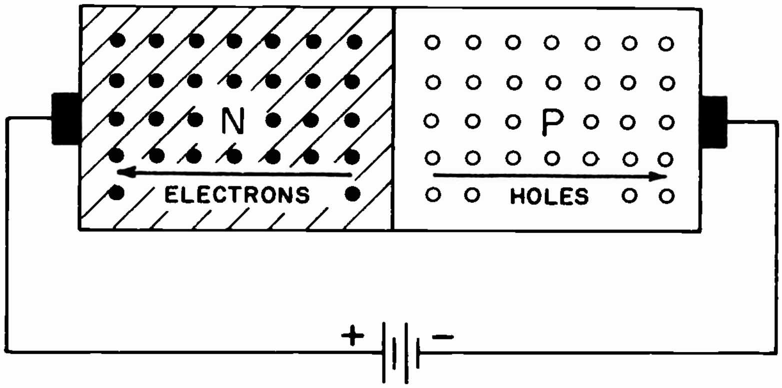


FIG. 19.—*No Current Flows When Voltage is Applied with This Polarity.*

Yet, to keep our thinking straight, it is generally deemed essential that this seemingly peculiar concept be applied to transistors.

We have taken the pair of crystals diagrammed in Figure 17 and used them in a slightly different form in Figure 18. In Figure 18 we show one crystal with an excess of electrons and the other with an excess of "holes." The two crystals are then connected to the terminals of a battery, which is capable of applying an electromotive force to the joined crystals.

Note that the positive terminal of the battery is connected to the positive crystal, while the negative terminal of the battery is connected to the negative crystal. When voltages with these polarities are applied to the joined crystals the movement of electrons and holes in the respective crystals is toward the junction between the crystals.

The positive voltage applied to the positive crystal causes holes to move from the connection to the battery toward the junction with the other crystal. At the same time, the negative voltage on the negative crystal causes electrons in the negative crystal to move from the electrical connection toward the junction.

In essence, this produces what we call electrical current flow.

At the junction we find a peculiar situation occurring. As the surplus electrons in the negative crystal move toward the junction they join with the holes moving toward the junction in the positive crystal. The electrons and the holes neutralize each other, and electrical conditions are temporarily satisfied.

For all practical purposes, an electrical current flows through the connecting wires—and the crystal—from one terminal of the battery to the other. In short, the crystal permits electrical current to flow through the circuit.

If the battery were connected to the crystals in the opposite direction, no current would flow.

When connected in that manner, as shown in Figure 19, both holes and electrons tend to try to move *away from* the junction. Since they no longer balance, or neutralize, each other, no current is able to flow in the external circuit.

When all this action is summed up it means that electrical current can flow through such a pair of crystals when the polarity is in one direction, but it is prohibited from flowing when the polarity is in the opposite direction. In short, if an AC voltage were to be applied to the respective terminals of the joined crystals, the action of the crystals would *rectify* the AC current into uni-directional current.

16. HOW CRYSTALS WORK IN PRACTICE

So far in our discussion of crystals we have pretended that a crystal is either wholly positive or wholly negative. By this we mean that we have acted as though a crystal was a pure positive crystal, or a pure negative crystal.

In short, we have handled our explanation as though a *negative* crystal had no "holes," and that a *positive* crystal had no excess electrons. It would be well if crystals could be grown or constructed in this manner, but as a matter of practical fact they do not follow this pattern precisely.

Scientist and other technical men know that negative characteristic crystals are never perfectly negative crystals. Such crystals have predominantly negative characteristics, but they also have a few "holes."

Much the same thing occurs in the case of positive crystals. While they possess predominantly positive characteristics, the fact still remains that they have some excess electrons.

It is because of this deviation from pure characteristics that perfect rectification by means of crystals is not possible.

It is in this feature that crystals tend to differ considerably from vacuum tubes. Vacuum tubes are capable of providing perfect rectification, as contrasted with the crystals.

For many purposes this inability to perform as perfect rectifiers is a matter of little concern. Crystals can perform an entirely adequate job.

On the other hand, there are some situations where the back-current permitted by crystals is a matter of considerable concern, and crystals cannot be used. While the matter of back-current is frequently of no importance, this is a fact which should always be kept clearly in mind, and looked upon as a feature which presents a limitation on their usefulness.

ITEMS OF SPECIAL INTEREST

Transistors have taken over many jobs formerly handled exclusively by vacuum tubes.

Vacuum tubes, however, are capable of handling many jobs which transistors are not fitted to perform.

Transistors should be looked upon as a tremendously useful supplement to vacuum tubes, but scarcely as something to completely replace them.

Transistors are made from crystals, and count among their ancestors the crystal detectors used in the very earliest days of wireless telegraphy and radio.

A crystal detector is not a perfect rectifier, whereas a vacuum tube does not permit any back-current to flow during the reverse cycle.

Crystals were used as detectors in the early days of radio due to their ability to rectify high-frequency signals.

Crystal detectors went out of style for many years, but came back into a period of usefulness during the development of radar receivers.

Crystal diodes are frequently used in the front end of radar receivers to act as a mixer to convert the extremely high-frequency radar signals into a lower frequency which can be handled by ordinary vacuum tube amplifiers.

It is customary, in ordinary electrical work, to classify all materials as though they were either electrical conductors, or electrical insulators. However, there is a twilight zone in which certain materials act as semi-conductors in one direction, and as an insulator in the other. Such materials are classed as *semi-conductors.*

Crystals possessing the ability to rectify electrical AC currents can all be classified as *semi-conductors.*

The ability of a material to act as a semi-conductor depends on its atomic structure.

Semi-conductors are usually created when some normally insulator material is adulterated with some other substance which changes its atomic structure to create "free" electrons, or to provide "holes" within the atomic structure.

Adulterant substances which provide "free" electrons in an atomic structure are commonly classified as *donor* substances, and as possessing *donor* atoms.

Substances which possess *donor* atoms are arsenic, boron, and antimony.

Substances which lack sufficient electrons in their outer shell to fill the full complement are said to be *acceptor* substances, and to possess *acceptor* atoms.

The most common substance used in electronic work which possesses *acceptor* atoms is *Indium.*

Crystals which possess the atomic structure most desirable for semi-conductor work are Silicon, Germanium and Selenium. These are not the only elements which can be used for this work, but are the ones most frequently used.

Each kind of element possesses individual characteristics which make it especially suitable for specific types of electronic work. Silicon is used in high-frequency work, but very frequently for middle-frequency transistors.

Germanium is used in diodes for extremely high frequencies, and is also used in photo-transistors which are sensitive to light.

Selenium is also sensitive to light, and is used frequently for that type of work. It is also used for the construction of heavy duty, high current carrying rectifiers.

Most rectifying crystals are sensitive to heat, and are frequently rendered inoperative when subjected to high temperatures.

A negative crystal is one whose atomic structure possesses a surplus of electrons.

A positive crystal is one which atomic structure is deficient in electrons. A deficiency of electrons is looked upon as being an atomic structure which has an unfilled "hole."

In transistor work there is a concept of current flow consisting of "holes" moving from the positive terminal of the source, around the circuit to the negative terminal. This is the reverse of the direction taken in electron current flow.

In transistor theory there is a concept of the electron current flowing in one direction while the "hole" current flows in the opposite direction.

Crystals are not perfect rectifiers, which limits their usefulness for some purposes.

Transistors derive their name from a partial joining of the words TRANSfer and resISTORS. Joining the capitalized letters of these two words results in the coined name TRANSISTORS.

A *valence bond* in an atomic structure consists of a linkage between two adjacent atoms in which a pair of electrons have orbits around both atoms.

TRAINING SERVICE

Massey Technical Institute, Inc. Jacksonville, Florida

Print or Use Stamp:

Name__________ Date__________

Student No.__________

Address__________ Zone______ Grade__________

City__________ State__________

Lesson 42

SEMI-CONDUCTORS AND TRANSISTORS
Part 1

Underline or fill in the ONE correct answer, just as in previous lessons.

1. When a pair of crystals with opposing polarities are joined together they posses the property to:
   a. multiply. b. rectify. c. amplify. d. oscillate.

2. When a crystal has predominant donor crystals its polarity is:
   a. positive. b. neutral. c. negative. d. reversed.

3. Boron, and arsenic possess what are called:
   a. neutral atoms. b. acceptor atoms. c. donor atoms. d. neutrons.

4. In semi-conductor theory it is customary to think of moving holes as being:
   a. negative particles of electricity. b. an electrical current.
   c. impossible. d. donor atoms.

5. A semi-conductor is distinguished from an ordinary conductor by its peculiar ability to:
   a. act as a conductor in one direction while acting as insulator in other direction.
   b. amplify. c. oscillate. d. modulate. e. agitate.

6. When two or more electrons link two atoms together with encircling orbits, it is called:
   a. free electrons. b. a valence bond. c. a donor atom. d. acceptors.

7. Diamonds, Silicon and Geranium have the same property of possessing:
   a. penta-valent lattice atoms. b. the ability to conduct electricity.
   c. abundant acceptor atoms. d. tetra-valent atom lattices.

8. A Silicon or Germanium crystal contaminated with Indium becomes:
   a. a positive crystal. b. saturated with donor atoms.
   c. a perfect insulator. d. negative crystal.

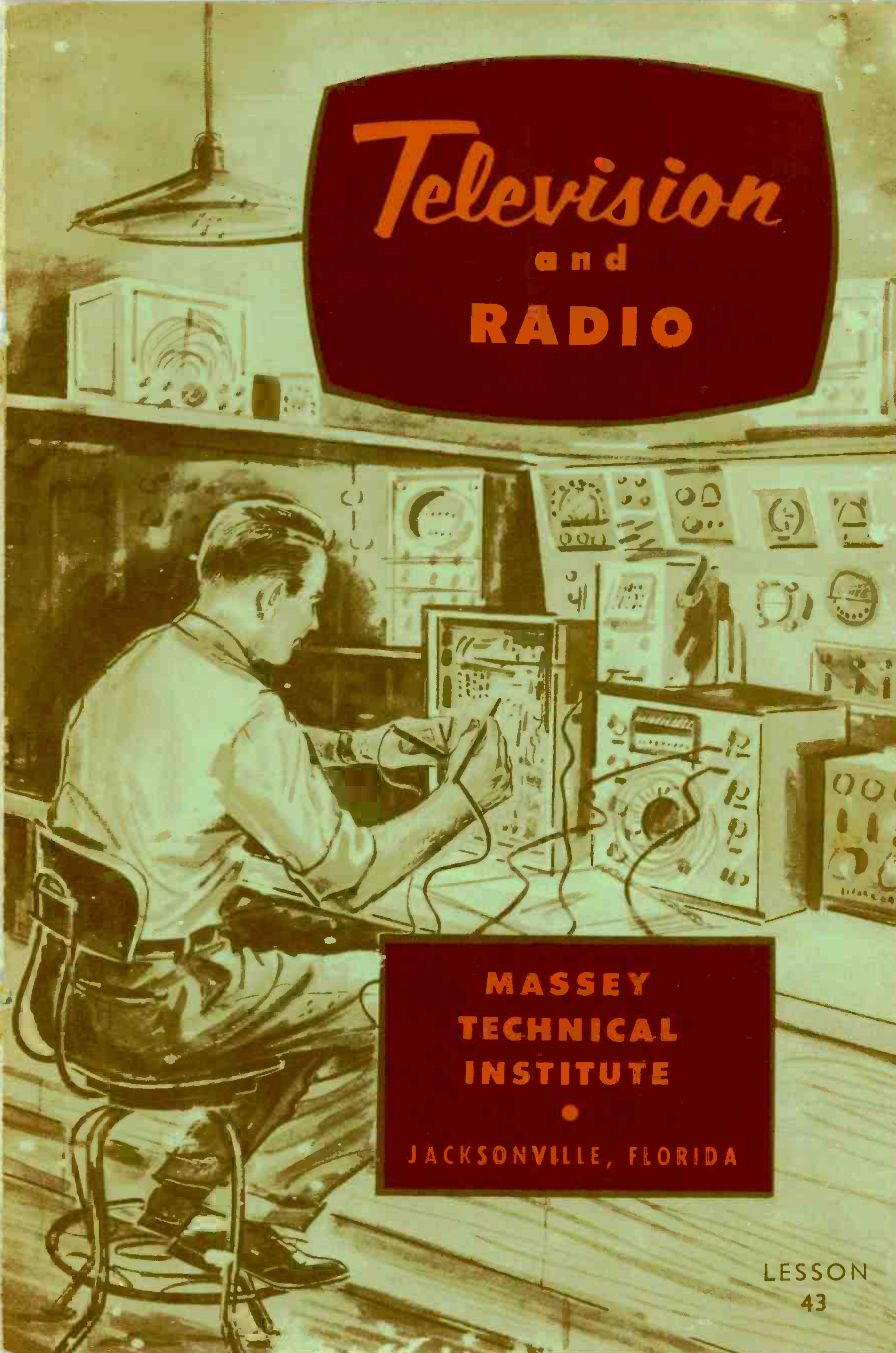
9. Crystals used in connection with transistors are closely related to those used:
   a. to maintain frequency on transmitters. b. as early day detectors.
   c. in phonograph pick-ups. d. in crystal microphones.

MASSEY TECHNICAL INSTITUTE, Inc.

MTI

Jacksonville, Florida

Television and RADIO



MASSEY TECHNICAL INSTITUTE

JACKSONVILLE, FLORIDA

LESSON 43

ALIBIS PAY NO DIVIDENDS

TRAINING SERVICE

TRANSISTOR ACTION

1. INTRODUCTION

Transistors are beginning to play an important part in many branches of Electronics. Their importance is bringing about a change in the thinking of persons whose attention has been focused too closely on the action which occurs in vacuum tubes.

While transistors are taking over some functions, under certain conditions, which previously were limited exclusively to vacuum tubes, they are doing so in an entirely different manner. Transistors work differently from vacuum tubes.

We have already described something of the action of the crystals as they perform the function of rectification. You will recall from our explanation that they perform that function differently from the way that it is performed by vacuum tubes.

Transistors have the ability to act as amplifiers. Again, the manner in which they go about the performance of that function is different from the action of a vacuum tube.

During these preliminaries we want you to keep in mind that both vacuum tubes and transistors are fully capable of acting as amplifiers. At the same time, however, we want to make it emphatically clear they do so in entirely different ways.

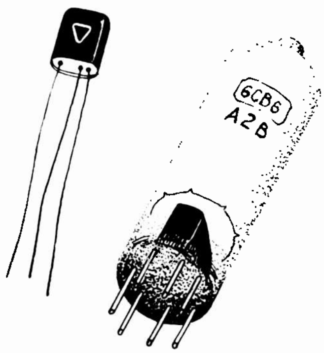


FIG. 1.—*Comparison between transistor and small tube.*

Since transistors are capable of rectifying AC voltages, and of amplfying AC voltage, it would be easy for a person new to this branch of electronics to jump to the conclusion that transistors can do everything that a vacuum tube can do. Since this supposition seems so reasonable—at first glance—one has some justification for arriving at a conclusion which soon is shown to be a fallacy.

By this we merely mean that while transistors are capable of duplicating many vacuum tube actions, they should not be considered a direct substitute for them. They are not.

2. COMPARISON OF TRANSISTORS WITH VACUUM TUBES

When it comes to comparison between transistors and vacuum tubes, the matter of size immediately attracts attention. A transistor is so much smaller than the smallest miniature vacuum tube, that a transistor walks away with all honors when miniaturization is a matter of consideration.

The illustration in Figure 1 provides an excellent means for comparing size of a transistor with that of a smaller vacuum tube. The vacuum tube is a type 6CB6, one of the smallest R-F pentode amplifiers. So far as vacuum tubes are concerned, this is one of the smallest types.

Yet, a glance at the illustration in Figure 1 quickly shows that even this small vacuum tube is much larger than a transistor.

One of the big advantages which a transistor holds over a vacuum tube is this matter of size. There are an increasing number of places where the size of the operating components has considerable importance.

Electronic equipment for use in aircraft is an illustration directly in point. Some of the modern transport aircraft seem incredibly large, yet the innumerable component parts which enable the craft to fly must be extremely small, and light in weight, while still performing their assigned tasks.

Weight saved by using transistors instead of vacuum tubes for many jobs in aircraft, such as the one shown in Figure 2, give such giant craft greater cargo capacity and wider cruising range.

3. DIRECT ADVANTAGES OF TRANSISTORS OVER TUBES

We have already pointed out one important advantage which transistors have over vacuum tubes. That is the item of size.

It was the extremely small size of transistors which first provided them with an area of usefulness. The first place where they

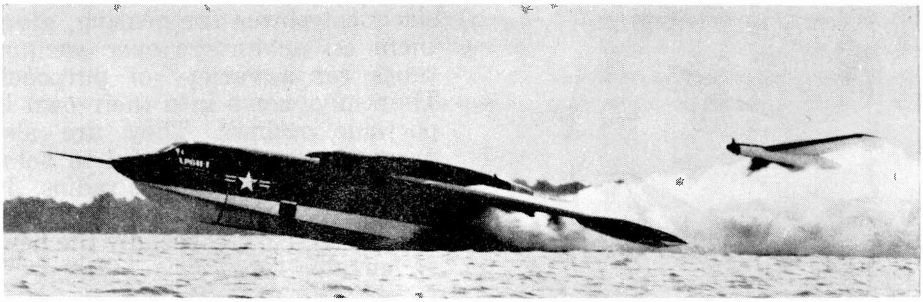


FIG. 2.—*Many transistors are used in the electronic gear for this giant seaplane.*

were put to practical use was in hearing aids for the deaf.

At first, transistors were directly substituted for vacuum tubes in regular hearing aids. The difference in size between them and the tubes, which had been previously used, enabled manufacturers of hearing aids to make them radically smaller.

Later, the extremely small size of transistors, together with even newer types which were still smaller than the originals, made it possible for hearing aid manufacturers to build the aids so small that a complete hearing aid could be built into the frames of a pair of glasses. This made it possible for persons who already wore glasses to also use the same frames to conceal inconspicuous hearing aids. The illustration in Figure 3 shows a person wearing a pair of such combination glasses and hearing aid. The hearing aid cannot be seen.

But size alone is not the only advantage possessed by transistors. For some purposes, the lack of any need for filament heating current is even more important.

Since a transistor works on an entirely different principle from vacuum tubes, it does not employ thermionic emission. This means that it does not have any filament to be heated. Therefore, no filament heating current is needed.

Because a transistor does not need filament heating power it has become widely used in electronic gadgets of a portable nature. This includes such things as hearing aids, which we have just mentioned, portable radios, some types of automobile radios, aircraft electronic equipment, and other similar things.

Probably most important of all, transistors have become extremely important in the construction and launching of guided missiles. Every ounce which can be removed from the operating equipment inside a guided or ballistic missile means that 1000 ounces can be removed from the overall weight of the missile as it sits on the ground.

All vacuum tubes, even the smallest, require some form of filament heating power. Filament power for portable electronic equipment usually means some form of electric cell, or a battery of cells.

This further means extra weight. All electric cells are heavy, even the smallest.

FIG. 3.—*Transistor hearing aid concealed in frames of glasses.*

Eliminating filament heating power requirements automatically means a radical reduction in weight. In the case of either guided or ballistic missiles, weight reduction is critically important, and if enough weight can be removed and still make it possible to launch the missile it helps provide the missile with added range. See Figure 4.

So it is, that in the field of missiles transistors are playing an important part. They are certain to play equally important parts in the construction and launching of manned space satellites, and space ships.

Small size and light weight of transistors, combined with the fact they have no filament requiring electrical power for heating, gives them an advantage over vacuum tubes for a variety of purposes. They have come into their own in portable radios. They are also proving extremely useful in some of the newer automobile radios. In a few cases they are used in television receivers, especially the newer portable types.

4. DISADVANTAGES OF TRANSISTORS

If one were to read only as far as we have gone so far in this lesson it would be possible to obtain the idea that transistors are about to push vacuum tubes aside, and make them as obsolete as the fabled Dodo bird.

The truth is that such is not likely to happen.

Useful as transistors have proven themselves to be, they are not without their limitations. They have definite, and valuable, advantages, but they also have some disadvantages.

For one thing, transistors are quite sensitive to heat. If the temperature is permitted to rise above certain definite limits, for each type, they are damaged. In some cases transistors lose all their useful properties, and in others their characteristics are changed to such extent they can no longer be used for their original purpose.

This means that wherever they are used they must be provided with protection against excessive heat. In some cases it is decidedly inconvenient to provide such protection, although in others it is not a matter of difficulty.

In their present state of development, transistors are unable to compare in amplifying ability with comparable vacuum tubes. This is something which doubtless will be improved with passage of time, but as of this writing transistors labor under a distinct disadvantage on this point.

There are some cases where it is necessary to employ at least one vacuum tube to deliver the power needed to perform some given task, or job. If so much as one vacuum tube is needed in an electronic device it means that a power supply must be provided so the tube will have the voltage and current needed for it to operate.

What this adds up to is that if a power supply must be provided for one tube, the same power supply might just as well be used to activate other tubes to provide whatever other amplification is needed.

This means there is not usually any advantage in using part transistors and part tubes in an electronic device. If one or more tubes are to be used, it is usually just as well to go ahead and use all tubes. Tubes are better amplifiers than transistors, and if a power supply must be provided for one tube it might as well be used for others.

There are certain exceptions to this general statement. An outstanding one is provided by some of the newer automobile radios.

Oddly enough, tubes are used in the voltage amplifier circuits of such automobile radios, the anodes being supplied directly from the 12-volt electrical systems. Such low voltage is not normally practical for power output tubes, yet is sufficient for transistors. So, such radios use vacuum tube voltage amplifiers operating on low anode

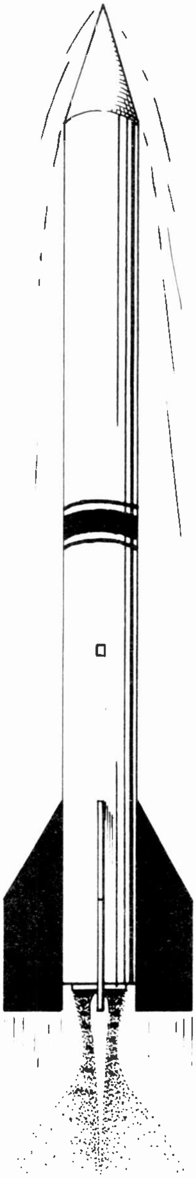


FIG. 4.—*Guided missiles depend heavily on transistors.*

voltages, but use transistors for the power output stage.

This is a complete switch from the normal practice, but in this isolated instance has been found to be highly practical.

There is another difficulty with transistors, at least such is true at this stage of their development. It is difficult to maintain their characteristics in a stable condition. This is an inherent difficulty which will become more apparent to you as we progress with our explanations.

This factor has made it difficult to use transistors in regular production models. Electronic equipment produced in mass production factories depend on close adherence to tolerances. In the case of transistors, there are such wide degrees of differences in the characteristics of similar types, it has been difficult to set up manufacturing tolerances.

However, this problem has been partially solved. Transistors are being used on production line manufactured products.

5. HOW VACUUM TUBES AND TRANSISTORS COMPARE

As we have just indicated, transistors possess certain definite advantages which provide them with a usefulness superior to that possessed by vacuum tubes. At the same time, they do not possess other advantages typical of vacuum tubes.

It is difficult to draw an arbitrary line, and make the flat statement that things on one side of the line can be done better by transistors, and those on the other side can be handled better by vacuum tubes.

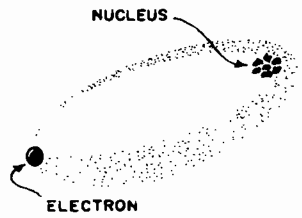


FIG. 5.—*How electron swings around the nucleus in its orbit.*

It should be kept clearly in mind that we have not had nearly so much experience working with transistors. Technical men have worked many more years with vacuum tubes than with transistors, and have thus become more familiar with their capabilities and limitations.

As time passes, and we accumulate additional experience with transistors, this situation could change. Undoubtedly, it will be found that many things which cannot be done with transistors at this time can be handled by other types of transistors which possess improved characteristics.

To understand the comparative advantages and disadvantages of both vacuum tubes and transistors it is usually convenient to describe each in a tabular form. This makes it easier to compare them, and thus acquire a better idea of where each is most useful. Table I makes this quite clear.

When the comparative characteristics tabulated in Table I are studied it becomes evident that vacuum tubes possess superior characteristics for certain applications, while transistors are more adaptable for use in other situations.

Where an adequate supply of electrical power is available, and the matters of size and space are not important, vacuum tubes possess advantages. Tubes are able to provide much higher voltage gain per stage, and deliver more power, than transistors. Furthermore, they are more stable in their operation, and are not so greatly affected by surrounding temperature changes.

On the other hand, where the electronic device must be portable, and the matter of size, space and weight are important, transistors have numerous points of superiority. The power supply for a portable transistor device may be very simple; it is common to use a tiny flashlight dry cell as the sole source of power.

Vacuum tubes still possess an imposing advantage in the matter of cost. A portable radio employing transistors, for example, still costs about twice as much as a vacuum tube radio having comparative sensitivity and quality. However, it seems probable that passage of time, and improvement in transistor techniques, will eventually narrow this matter of cost.

TABLE I.

HOW TRANSISTORS COMPARE WITH VACUUM TUBES

Advantages:

1. Require low operating voltages.
2. No need for filament heating power.
3. Have longer life, if not abused.
4. Are much smaller than tubes.
5. Drain very little power.

Disadvantages:

1. Hard to find exact replacement when one is defective.
2. Relatively low voltage gain.
3. Operating characteristics not so uniform as tubes.
4. Easily damaged by heat, or electrical current.
5. Unstable when ambient temperature varies.
6. Unstable in ambient electrical fields.
7. Not so reliable at high operating frequencies.
8. Limited power gain per stage.

HOW VACUUM TUBES COMPARE WITH TRANSISTORS

Advantages:

1. Fewer tubes do more jobs.
2. Tubes have greater amplifying ability.
3. Tubes have more uniform characteristics.
4. Tube action is more predictable.
5. Tubes can deliver more power.

Disadvantages:

1. Filaments require power for heating.
2. Must be replaced more frequently.
3. Require higher operating voltage.
4. More susceptible to mechanical damage.
5. Require special power supplies in portable equipment.
6. Are larger, and occupy more space.

6. ELEMENTARY ATOMIC PHYSICS

In previous lessons we have touched lightly on atomic physics, and the basic elements involved in the structure of atoms. We have explained in a very brief manner some of the factors which relate atomic structure to electrical current flow.

On several occasions we have pointed out that the ability of an element, or compound, to conduct electricity depends on its atomic structure.

We do not believe it necessary to penetrate very deeply into the science of atomics, or go into a lengthy or involved discussion of nucleonics, in order for us to teach you about transistors and electronics. At the same time, there are certain basic fundamentals about atomic structure which you should understand if you want to acquire a clear picture of how transistors work.

Magnetic forces are present within an atom—so are electrical and gravitational forces. They all go to make up what we usually group together into what we call *nuclear forces.*

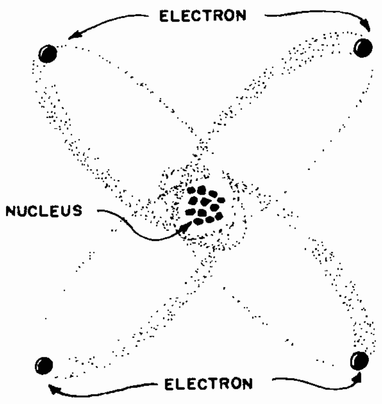


FIG. 6.—*Two electrons follow different orbits, but the orbits both occupy the same shell of space.*

These various forces interact within, and upon, each other. The result of this interaction is that varying groups of electrons keep revolving in their orbits around the central nucleus.

Never lose sight of the fact that electrons, traveling within their respective orbits located within specific shells of space surrounding the nucleus, move at incredible speed.

Within the tiny space occupied by a single atom is a degree of power so tremendous it is hard to appreciate. Before the first atomic bomb exploded above the New Mexican desert in the darkness of that early July dawn, few persons outside a handful of scientists had any real inkling of that incredible power.

We are better able to appreciate that power than was true a few years ago. Even so, it still remains hard for many of us to clearly understand how the power is locked up, or to understand how it can be released.

At the heart of each atom is a nucleus. The nucleus consists of neutrons and protons. Neutrons possess the massive weight of protons, but are electrical neutral. Thus, in most electrical and electronic work it is practical to ignore neutrons.

Each electron of the atom moves around and around the nucleus within the limitations of a fixed orbit. The drawing in Figure 5 shows how the electron swings around the nucleus in an elliptical path, swinging in close to the nucleus at one point in the orbit and away from the nucleus at another point.

The drawing in Figure 5 shows only a single electron swinging around its nucleus, following repeatedly the same path of an orbit. In Figure 6 we can see two such electrons following their orbits around the nucleus.

It will be noted that the two electrons in Figure 6 follow different paths, or different orbits, in their movements around the nucleus. However, each of them occupy approximately the same area of space around the nucleus.

In the language of the atomic scientist, it would be said that the two electrons in Figure 6 occupy the same *shell*. Their orbits are positioned in the same layer of space whose center is the nucleus.

To acquire a better idea of this matter of the shells, and the manners in which the electrons follow

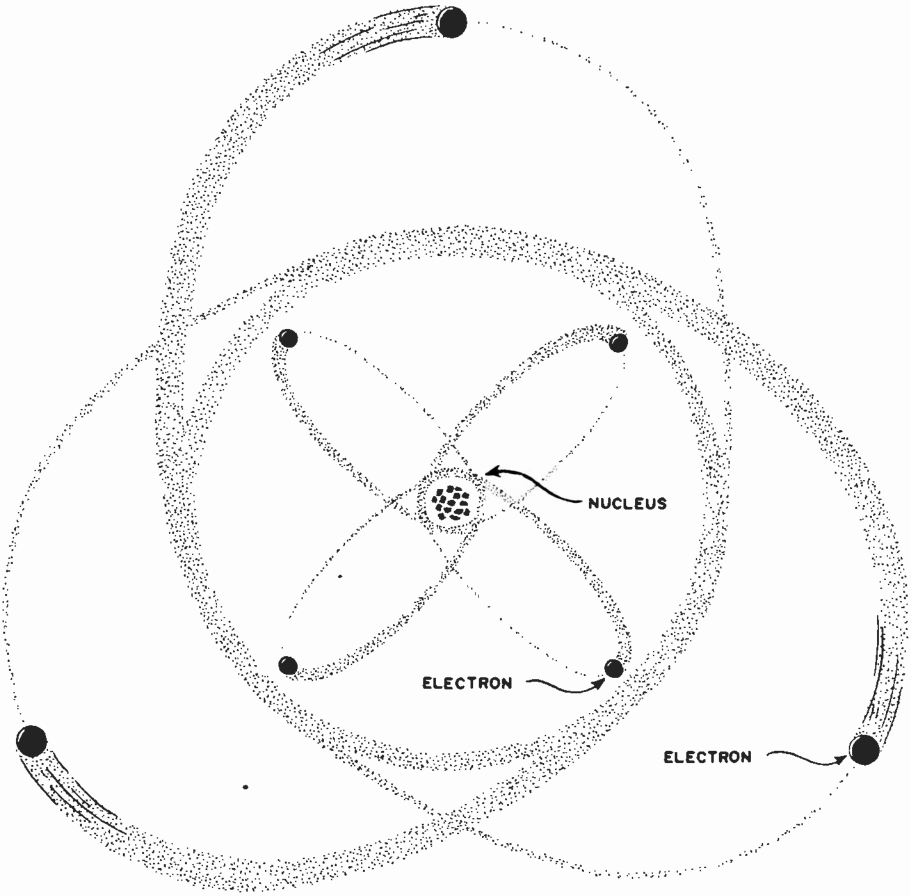


FIG. 7.—*Different groups of electrons following orbits in different layers of space, or shells.*

their respective orbits within those shell layers, it would be well to study the drawing in Figure 7.

A close study of Figure 7 shows the two electrons in the inner shell, represented by the shaded ball-like area around the nucleus, as forming a center around which the electrons in the second shell can move in their own orbits. Electrons in the second shell move within the limits of that shell, and not out into outer space, nor do they penetrate within the inner shell.

Atoms which have additional shells have other electrons moving in another layer of space which surrounds the space occupied by the second shell.

This brings us to the subject mentioned briefly in a preceding lesson in which we introduced you to the matter of *valence bonds.*

In the case of atoms of certain kinds of material the electrons in the outer shells act to form orbits which encircle the nuclei of two atoms. Some idea of how a pair of electrons have their orbits in this manner is shown in Figure 8.

The drawing in Figure 8 fails to completely illustrate the action, since some of the details have been omitted in the interest of clarity. Nevertheless, we see one atom, which we have designated as Atom No. 1, in one location. At a little distance from it is a second atom nucleus which we have designated as Atom No. 2.

Each of these two atoms have their own groups of electrons, following orbits within definitely fixed shells.

Linking both atoms, however, is another pair of electrons. These are what we know as *valence bond* electrons, since they link with their orbits two separate atoms.

It is this linking together by means of these valence bond electrons which hold the atoms together in a permanent structural arrangement.

Although we have shown only two atoms linked in this manner, the fact is that each of the atoms shown are also linked with other nearby atoms in the same manner with other pairs of valence bond electrons.

In some types of substances the electrons are all tightly bound to the nuclei in such manner that none are free to move from one atom to another. Such substances are known as electrical insulators.

On the other hand, the valence bond electrons in some substances are so arranged that it is possible for them to move from atom to atom under the influence of external electrical pressures. Such substances are what we know as electrical conductors.

Most substances can be classified as either conductors or insulators. However, there is the twilight group of substances we have mentioned in a previous lesson which are now classified as semi-conductors. They are the substances which conduct electrical current with a reasonable degree of facility in one direction, but act as a high resistance in the opposite direction.

It is this group of semi-conductors which are of most interest to us, since it is from them that we build our transistors.

It is often useful to compare the relative conductivities of various

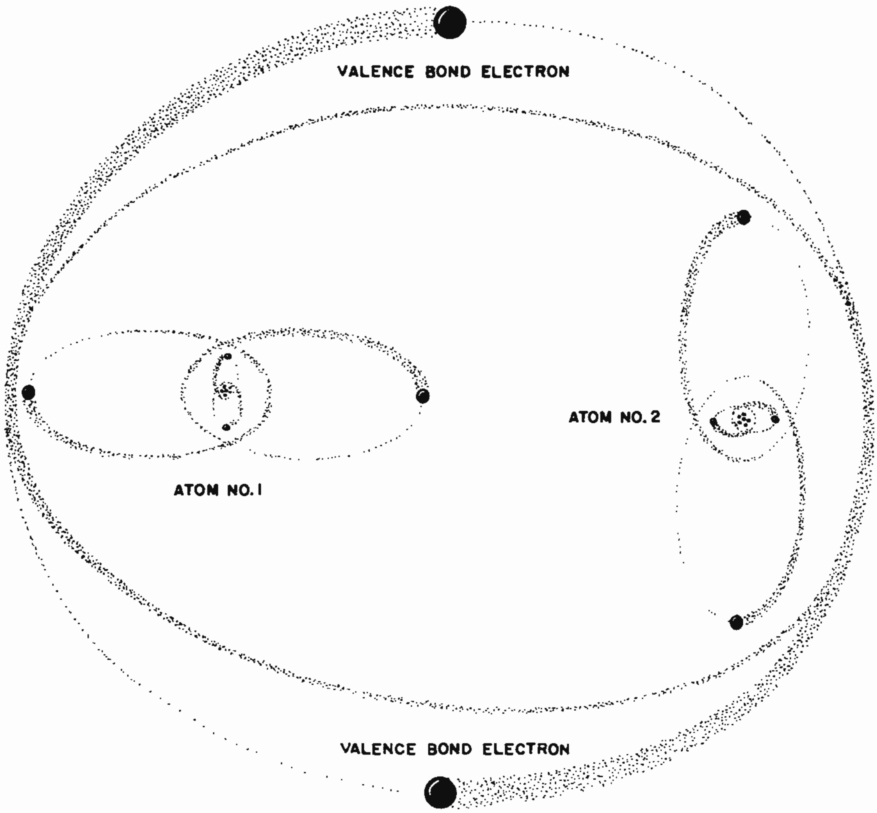


FIG. 8.—*How Valence Bond electrons link two atoms.*

kinds of substances. Such comparison sometimes help make the relationships among them a little more meaningful.

Some scientist has made the estimate that a crystal of Germanium which has been specially prepared for use as a transistor has a conductivity about a trillion times greater than glass. Glass, you know, is an insulator.

Carrying this matter of comparison a bit further, it has also been estimated that copper has a conductivity about 30,000,000 times greater than the same Germanium crystal. These two methods of comparing the relative conductivities of conductors, semi-conductors, and insulators is a bit dramatic, yet it helps convey the thinking we are trying to get across to you.

7. GERMANIUM AND SILICON AS SEMI-CONDUCTORS

From the things we explained in a previous lesson, we believe it is quite clear to you that neither Germanium nor Silicon, when in a

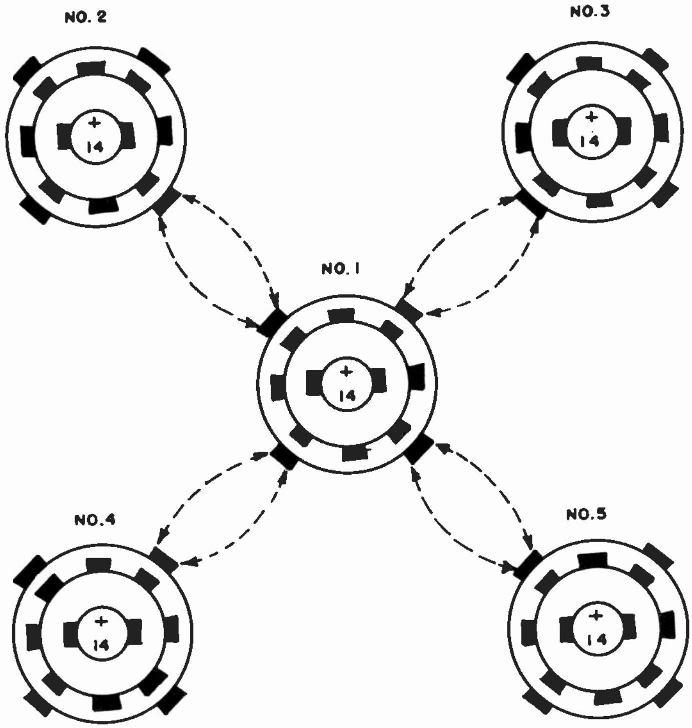


FIG. 9.—*Tetra-valent Lattice of Silicon.*

pure form, is an electrical conductor. Both must be classified as insulators.

At the same time, the atomic structure of both Germanium and Silicon is such that either can be changed into a semi-conductor quite simply. All that is necessary is to add small quantities of impurities of a specific nature. When the pure crystal is adulterated in this manner it becomes a semiconductor.

One of the difficulties experienced with transistors stems from the fact that it is hard to control the amount of adulterant added to the pure crystal, yet the percentage of such adulterant has a highly significant effect on the manner in which the crystal works. It is to perfect the degree of control during manufacture that scientists are directing much of their attention in transistor work.

We have gone into some detail in the preceding lesson to explain how the electrons in a Silicon atom are arranged into what is called a *tetra-valent lattice*. The arrangement is shown diagrammatically in Figure 9. There are four electrons in the outer shell of each atom.

When a crystal of Silicon is adulterated with a penta-valent atom, as shown in Figure 10, an electron is left free in the outer shell, and does not form a valence bond with any other atom. Such electron is then free to move from atom to atom, thus making possible a modified conduction of electricity.

A penta-valent atom, one having an extra electron in the outer shell as shown in Figure 10, is classified as a *donor* atom, thus capable of providing, or *donating*, an electron when mixed with tetra-valent atoms. Typical *donor* atoms used in transistor work are Boron, Arsenic and an occasional other substance. Arsenic and Boron seem to be used for this purpose more often than any other kind.

For the purpose of keeping your thinking clear you should remember that adulteration of Silicon or Germanium atoms with a *donor* atom like Arsenic tends to endow the crystal with *negative* characteristics. The negative characteristics stem from the excess electron provided by the penta-valent atoms.

On the other hand, the Silicon or Germanium crystals can be given *positive* characteristics. This can be done by adulterating the Silicon or Germanium crystals with a substance containing *acceptor* atoms. An acceptor atom

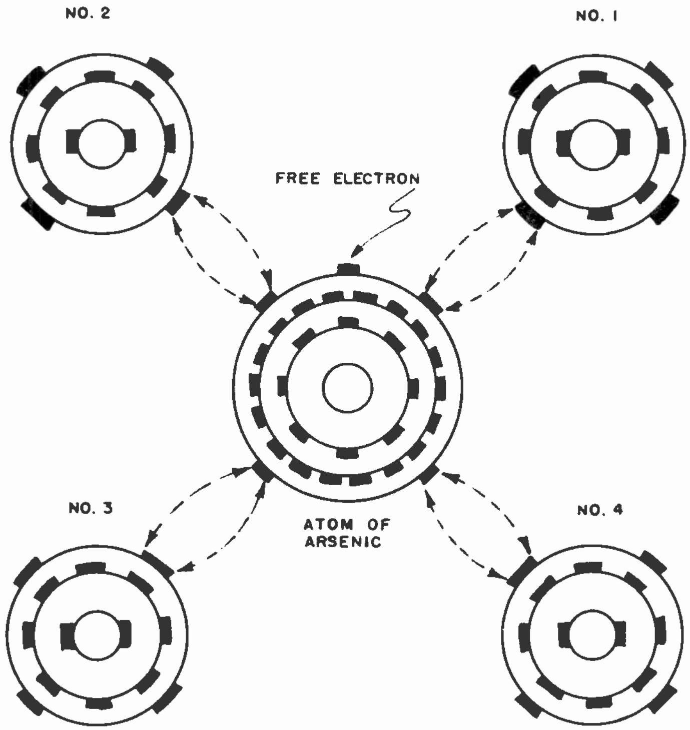


FIG. 10.—*Tetra-valent Lattice adulterated by penta-valent atom.*

is one which has less than four atoms in its outer shell.

An acceptor atom has been given that name because it is capable of accepting an electron from an atom of the Silicon or Germanium.

To carry this thought a step further, an acceptor atom is one which is missing an electron in its outer shell. The outer shell, in other words, has an empty space. In the technical terminology of scientists who work with semi-conductors, the empty space is referred to as a "hole."

Because a Silicon or Germanium crystal which has been adulterated by substances containing acceptor atoms—atoms with "holes"—is lacking in electrons, such a crystal is understood to possess positive electrical characteristics.

In the terminology of semi-conductor technicians, a crystal which has negative characteristics is commonly referred to as an N-type crystal. This is a designation which you will run into repeatedly in transistor work, so it would be wise to keep it clearly in mind.

A crystal which has positive characteristics is referred to, logically enough, as a P-type crystal.

8. PRODUCTION PROBLEMS FACING TRANSISTOR MANUFACTURERS

It is doubtful that you will have occasion to work in a manufacturing plant which is making transistors. This may lead you to feel there is no need to know these details. However, your understanding of transistors, and the problems connected with their manufacture and use, will be much more clear if you know something about the underlying structure of transistor crystals.

One of the most serious problems connected with the manufacture of transistors revolves around the difficulty of maintaining precise control over the percent of impure material used to adulterate a crystal. If the amount of adulterant material is too little or too much the crystal is useless for transistor work.

Something of the magnitude of this problem can be better imagined if we use some actual figures for the purpose of demonstration.

Suppose we have the basic element of Germanium in absolutely pure state. This means a crystal which is completely free from contaminating atoms. This is possible, but extremely difficult to achieve as a practical matter.

Now, suppose that to every 100,000,000 Germanium atoms we add 1 atom of some adulterant substance, such as Indium or Arsenic or Boron.

When we add an adulterant at this rate—at the rate of 1 part in 100-million—the conductivity of the crystal is increased 16 times. Note the small amount of adulterating material, and the effect it has on the conductivity of the crystal.

When a crystal is adulterated at the rate of 1 part to 100 million the crystal is satisfactory for transistor work.

However, instead of the adulterant being added at this rate, suppose 1 part of adulterant is added to 10,000,000 parts of Germanium. When this is done the conductivity of the crystal is increased by 160 times.

A conductivity this high is too great for satisfactory transistor operation. Therefore, such a crystal cannot be used for transistor work.

If you give this a little thought you soon begin to realize the magnitude of the problem facing the manufacturer. Adding 1 adulterating atom to 100 million Germanium atoms is about on the order of adding one additional person to the population of the United States east of the Mississippi River. It is about on the order of adding 1 drop of liquid to a municipal water tank.

On the other hand, adding 1 part to 10 million, is about the same as adding one additional person to the people in New York State. Remember, when the adulterant is added at this rate—it is too much.

If the manufacturing processes could be reduced to the problem of first obtaining a pure Germanium crystal, then adding the proper proportion of adulterant, it would be one thing. But the problem is even more difficult than that.

It is very rare that a pure form of Germanium can be obtained. This means that the manufacturer starts with a Germanium crystal which already possesses some degree of contamination mixed with the Germanium. In most cases, there are impurities present which are neither penta-valent nor tri-valent in character.

Such impurities do not always affect the conductivity of the crystal. However, they create imperfections in the crystal's atomic structure, and often set up conditions which make it unfit for transistor work.

9. HOW IMPURITIES AFFECT CRYSTALS

In transistor work it is desirable to have crystals with distinctive N-type characteristics, or distinctive P-type characteristics. When such crystals can be obtained they can be joined together in a definite manner, and thus create the conditions which cause them to exhibit the characteristics which we know as transistors. When so joined they are known as junction transistors.

Presence of P-type impurities in a Germanium N-type crystal exerts a strong effect on the conductivity of the crystal. Much the same situation prevails when N-type impurities are present in a P-type crystal.

In those situations where the N-type impurities are present in equal amounts with the P-type impurities, the two types tend to cancel each other's effectiveness. In such cases the crystal acts very much as though no impurities of any kind were present.

One of the major problems in manufacturing transistors is to reduce the degree of contamination so that less than 1 contaminating atom is present for each 100,000,000 Germanium atoms. Once that degree of purity is achieved the manufacturer is faced with the problem of then adding the correct percentage of adulterant.

When speaking of extremely pure Germanium we run into a situation which is called *intrinsic*. This term is used to describe Germanium crystals in which the donor and acceptor atoms are evenly balanced, or where the crystal is virtually completely free from contaminating atoms.

When a Germanium crystal has been purified to intrinsic value the valence bond electrons are affected by heat and light applied externally. This means that either heat or light is able to bring about a condition of conductivity.

There are cases where this condition is desirable. There are others where it is not desirable.

When heat is applied to a Germanium crystal of this type, the heat creates a high degree of excitation, with electrons being released from their valence bonds. This permits conduction to occur.

This situation raises a point of importance in connection with Germanium crystals. They have what is technically known as a *negative coefficient of resistance.*

In less technical language this means that increasing the temperature reduces the resistance.

This matter of temperature has an important bearing on the use of Germanium crystals in transistor work. When the temperature of a Germanium crystal transistor rises to 80° C, the thermal excitation of the electrons in the valence bond permits them such a degree of freedom as to turn the crystal into a conductor.

This ruins the transistor's usefulness for the purpose for which it was originally designed.

If you are accustomed to thinking of temperatures in terms of Centigrade, we will point out that 80° C is approximately equivalent to 175° Fahrenheit. This is about the temperature at which hot water emerges from normal hot water faucets.

Do not permit these statements to mislead you. Many transistors are ruined at temperatures much lower than the one we mentioned. In practice, it is seldom considered wise to subject transistor crystals to temperatures exceeding about 125° F, if that can be avoided.

10. EFFECT OF LIGHT ON GERMANIUM CRYSTALS

Germanium crystals, in common with several types of other crystals, are sensitive to light. In transsistor work, however, we need not worry much about light damaging transistors. It is usually possible to shield them against the effects of light.

Sensitivity of the crystals to light is put to good use in certain types of transistors. Special types of photo-transistors have been developed. These are types which become conductive when subjected to light.

Photo-transistors are assuming a high degree of importance for a number of uses. However, this is a subject which we must return to later. We do not have either the time nor space to explain them in this lesson.

11. DIRECTION OF CURRENT FLOW THROUGH CRYSTALS

We have previously pointed out that N-type crystals possess the ability to conduct electrical current. At least, they are reasonably good conductors.

We want to make it equally clear that P-type crystals are equally good conductors. This means that either kind of crystal is capably of conducting electricity.

More than this, either type of crystal is capable of conducting

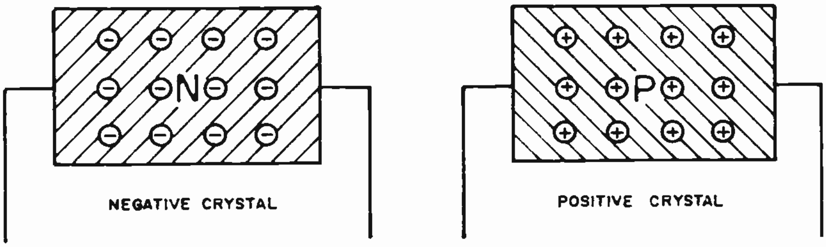


FIG. 11.—*Either type crystal conducts current in either direction.*

current in either direction. This means that neither a P-type nor an N-type crystal can be looked upon as possessing rectifying characteristics.

To illustrate these conditions we show diagrams of each type of crystal in Figure 11.

We carry this illustration to a step further in Figure 12, where we show that applying an electrical potential to a P-type crystal causes current to flow. It makes no difference which way the polarity is applied, current will flow.

In the case of the P-type crystal in Figure 12, applying an electrical potential to the two sides of the crystal causes the "holes" to move from the positive terminal toward the negative terminal. As a matter of fact, this is technically the same as having negative electrons moving from the negative terminal toward the positive terminal, but in crystal and transistor work, current flow in a P-type crystal is looked upon as the movement of "holes."

Much the same thing could occur had the crystals in Figure 12 been N-type crystals. In fact, we show that such could happen in Figure 13.

The point of what we are trying to make clear is that electrical current will flow in either direction through a crystal which has either polarity characteristic predominating. Current will flow in any direction dictated by the polarity of the applied voltage.

However, we run into an entirely different situation when a pair of crystals, with opposite polarities, are joined together.

When a pair of crystals with opposite polarity are joined as shown in Figure 14 current will flow when the applied voltage has the proper polarity. However, the polarity of the applied voltage is critically important.

For current to flow through the joined crystals, it is necessary for the negative terminal of the voltage source (battery) to be connected to the N-type crystal. It is equally necessary for the positive terminal of the voltage source to be connected to the P-type crystal.

Conduction inside the joined crystals occurs in the manner outlined in Figure 15. The negative voltage from the battery forces electrons away from the point where connection is made to the

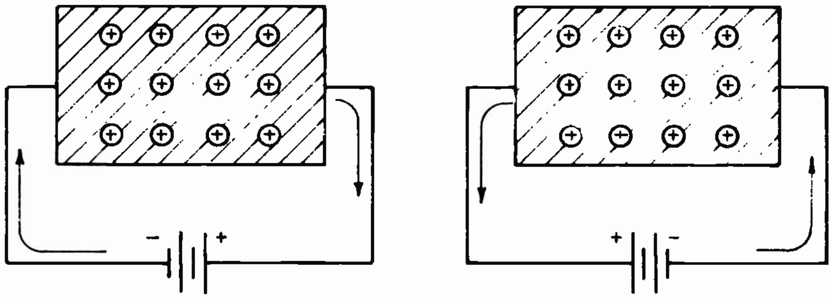


FIG. 12.—*Applying polarity in either direction permits conduction to occur.*

N-type crystal. Simultaneously, the positive voltage on the P-type crystal is causing "holes" to move toward the junction of the two crystals.

Thus we have a condition in which electrons are moving through the N-type crystal toward the junction, while "holes" are moving in the opposite direction through the P-type crystal toward the junction.

Remember, a "hole" in one of these crystals is merely a condition in which an electron is missing. Thus we have excessive electrons moving in one crystal toward the junction where such electrons are needed to fill the "holes."

In short, the excess electrons and the "holes" neutralize each other at the junction, thus balancing out, and permitting current to flow through the joined crystals.

Now let us take a look at what happens when the polarity of the voltage source is reversed. This condition is shown in Figure 17.

When the voltage is applied in this direction no current can flow. The explanation is that a negative voltage applied by the battery to the P-type crystals causes the "holes" in the crystal to move toward the electrical connection

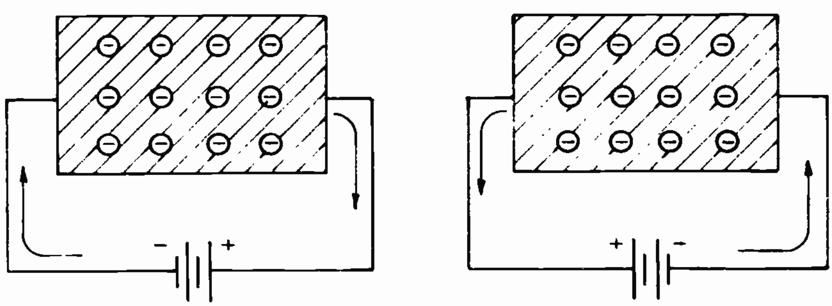


FIG. 13.—*Current flows through N-type crystal as readily as through P-type.*

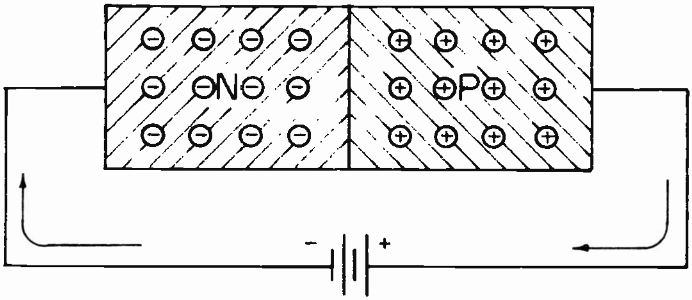


FIG. 14.—*Voltage polarity must be in correct direction when opposing crystals are joined.*

from the battery. This is a direction away from the junction of the two crystals.

Simultaneously, the positive terminal of the battery is placing a positive voltage on the N-type crystal. This causes electrons in the N-type crystal to move in a direction away from the junction and toward the connection to the battery.

These respective movements quickly cease. Only a limited number of electrons can be caused to move through the N-type crystal, and only a limited number of "holes" can be caused to move through the P-type crystal. After a relatively few of each are moved, stresses are set up inside the respective crystals which prevent further movement.

Thus, no conduction occurs.

The net effect of all this is that when a pair of crystals with opposite polarity characteristics are joined as shown in Figures 14 through 16, conduction can occur when the polarities are correct. On the other hand, conduction is inhibited when the polarities are reversed.

The result is to permit conduction in one direction, but to prevent it occurring in the opposite direction. This is the essential ingredient for rectification.

12. ELECTRICAL ACTION INSIDE A CRYSTAL DETECTOR

It is interesting to note how this action of crystals is applied to those once widely used as radio detectors.

The drawing in Figure 17 shows a typical detector crystal locked into a holder, then the "cat whisker" used to pinpoint a "good" position on the crystal. It will be noted from the drawing that the crystal is composed of several crystals with differing polarity characteristics.

Early day radio men did not know all this at the time they were using crystals to "detect" radio signals. Actually what happened is that the crystals they used were either predominantly N-type, with small areas of P-type crystals, or were predominantly P-type with small areas of N-type crystals. Either way would work.

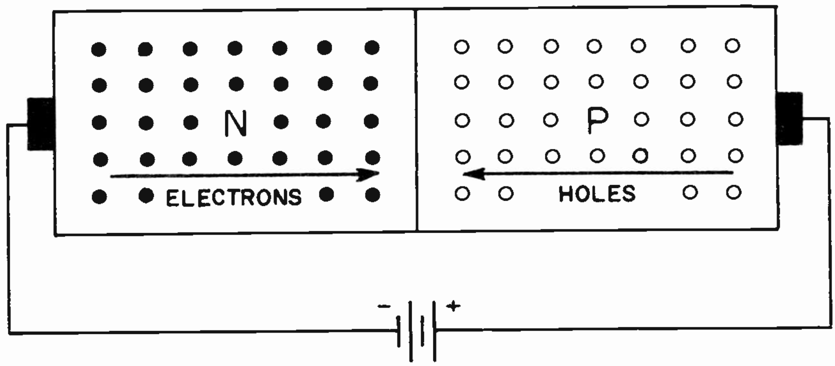


FIG. 15.—*Conduction inside joined crystals when polarity is correct.*

What the radio experimenters would do is feel around over the surface of the crystal until they found a sensitive spot. What they actually did was to find a spot which differed in polarity from that which predominated in the crystal. When they did so, they established all the essentials for rectification, which amounts to detection of high-frequency radio signals.

13. POTENTIAL HILL AT CRYSTAL JUNCTIONS

The illustration in Figure 17 shows a crystal which is predominately negative. However, there are several spots on the crystal which are predominately positive. It is on one of the positive spots that the "cat-whisker" is resting.

A thought which may come to you is this: why do not the free electrons in the N-type crystal neutralize themselves by filling in the holes in the P-type crystal when a junction is formed between the two crystals?

This is a good question. It is one which puzzled scientists for quite a while. They think they now know the answer. At least, they have come up with a theory—or hypothesis—which serves to explain the condition, and most scientists accept it as being accurate.

For our purposes, anyway, the theory can be considered as being reasonable and acceptable.

To help understand this theory we have prepared the drawing in Figure 18 as an aid to understanding it. At the left side of the drawing is a P-type crystal. Being a P-type crystal, it has a significant number of acceptor atoms.

Because the acceptor atoms are deficient in electrons, this means that such a crystal has a number of holes.

On the right side of the drawing is an N-type crystal. This crystal has a number of donor type atoms. These, as has been previously explained, are atoms which possess an excess electron. This means that among the atoms making up the N-type crystal are a number of free electrons.

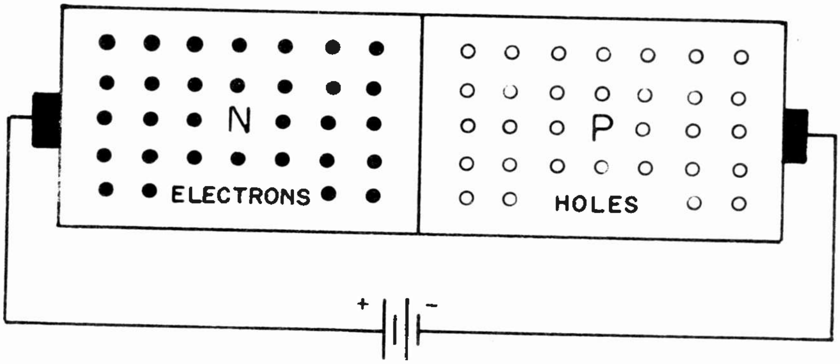


FIG. 16.—*When polarity of voltage source is reversed no current flows.*

Scientists and researchers have come to the conclusion that the acceptor atoms in the P-type crystals are fixed in their positions. They are not free to move around. Certainly they are not free to move around as are the electrons.

Much the same is true of the donor atoms in the N-type crystals. The donor atoms are no more free to move around than are the acceptor type atoms. This means that both the donor and the acceptor atoms in the respective crystals are fixed in their normal positions, and do not move from them.

Scientists accept the fact that a few free electrons from donor atoms closest to the junction between the two types of crystals may neutralize a few of the holes in the P-type crystal on the other side of the junction. But this is

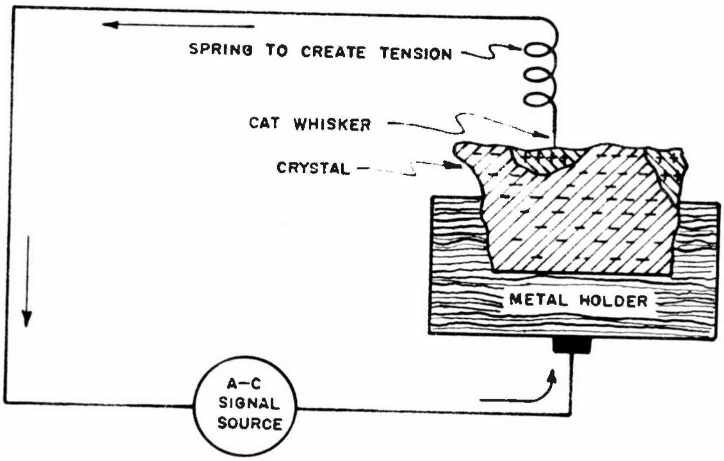


FIG. 17.—*Cat-whisker touching sensitive spot on crystal to bring about radio detection.*

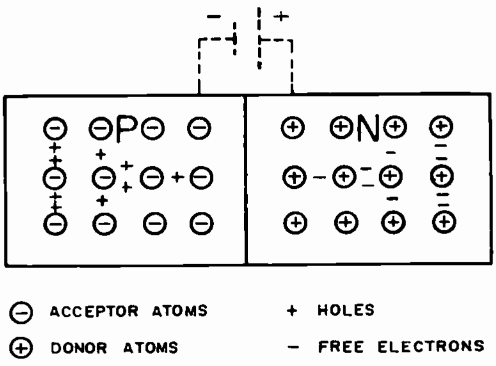


FIG. 18.—*Creation of barrier at junction which prevents neutralization of acceptor and donor atoms.*

a minor matter, and the neutralizing action does not extend back more than an infinitesimal distance from the actual junction surface.

However, once this neutralization does occur at the immediate vicinity of the junction, a barrier is created at the junction surface.

What happens is that a barrier of acceptor atoms line the surface of the N-type crystal nearest the junction, as shown in Figure 18, and a barrier of donor atoms line the surface of the P-type crystal on its side of the junction. This barrier effectively bars the neutralization of any additional atoms on either side of the junction until, and unless, a voltage is applied to the respective crystals.

If you study the drawing in Figure 18 carefully, you will note that the surface of the N-type crystal nearest the junction has no free electrons. By shifting your attention to the P-type crystal, you will note that the surface at the junction has no holes, or acceptor atoms.

When this condition exists it creates a tiny voltage between the two surfaces. This tiny voltage, which exists directly on the surface between the two crystals, is opposite in polarity to that of the main crystals themselves.

In Figure 18 we have shown this tiny voltage as existing in the form of an imaginary battery, a battery whose polarity is opposite that of the crystals of which the battery is a part.

To put this in other words, we can say that the N-type crystal is predominantly negative with an excess of free electrons. Yet, on the exact surface where the N-type crystal joins the other crystal, its surface is positive.

The same situation exists with respect to the P-type crystal. The P-type crystal is predominantly positive, with a surplus of acceptor atoms. Yet, at the surface where the P-type crystal joins the other the surface is negative.

This opposing voltage which exists at the surface where the two crystals are joined has been given

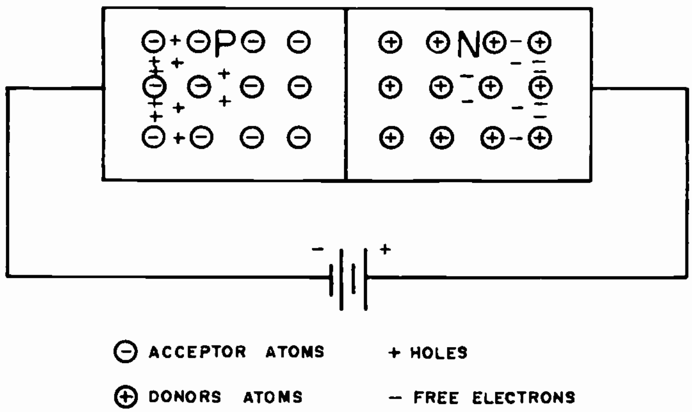


FIG. 19.—*Reverse bias.*

a technical name. It is called a *potential hill.*

The potential hill voltage is enough to prevent a progressive neutralization of the acceptor and donor atoms in the two crystals. Yet, the voltage is so low it is easy for a stronger voltage applied by an external source to overcome it.

What we are trying to say is this: The *potential hill* voltage is sufficiently large to prevent neutralization among the atoms in the two types of crystals. However, if an outside voltage is applied by a battery or other voltage source, it is easily able to overcome the voltage of the potential hill.

14. BIAS OF APPLIED VOLTAGE

When an external voltage is applied to a pair of crystals joined as shown in Figure 18 it is customary to refer to such voltage as a *bias.* In a limited way this voltage is similar to the biasing voltage used on control grids of vacuum tubes, yet the technical sense in which it is used with crystals and transistors is different, and one should not confuse the two usages.

If the external voltage from a battery is applied to the joined crystals in such manner that the positive voltage from the battery is applied to the negative crystal and the negative voltage is applied to the positive crystal, we have a condition which is called *reverse bias.* This condition is shown in the diagram in Figure 19.

When a reverse bias is applied to the joined crystals, the atoms, holes and electrons will align themselves in positions approximately as shown in Figure 19. Free electrons in the N-type crystal move away from the junction. At the same time, holes in the P-type crystal move away from the junction.

Under these conditions no current flows from the battery through the joined crystals.

We have a different situation when the voltage polarity of the battery is reversed. The reversed condition is shown in Figure 20.

When the positive terminal of the voltage source is connected to the P-type crystal, and the negative terminal of the voltage source is connected to the N-type crystal, as shown in Figure 20, it is possible for current to flow through the circuit including the joined crystals. The arrangement of the electrons, holes and atoms inside the two crystal structures is much as shown in Figure 20.

Free electrons in the P-type crystal move toward the junction. At the same time, "holes" in the P-type crystal also move toward the junction.

At the junction, the free electrons and "holes" neutralize each other, and a flow of current occurs in all parts of the circuit.

This brings us back to our explanation of the *potential hill* in the preceding section of this lesson. At that time we explained that the potential hill consisted of a low voltage existing on the surface of the joined crystals.

Since this potential hill exists, you may wonder just how much voltage must be applied to the two crystals in order to overcome that potential hill, and thus permit current to flow.

The potential hill has a voltage on the order of only a few tenths of a volt. However, to cause current to flow it is necessary to apply a forward voltage to the joined crystals somewhat larger than that of the potential hill before current can flow.

In the case of transistors—and that is what we are primarily interested in at the moment—an external voltage of about two volts is needed for current to flow through the joined crystals.

15. A NEW LOOK AT THE ELECTRON THEORY

In one of the earliest lessons of this course we introduced you to what is known as the "Electron theory." This is the basis for all work in the electronic field, and the fact that the theory is sound is demonstrated by the miraculous

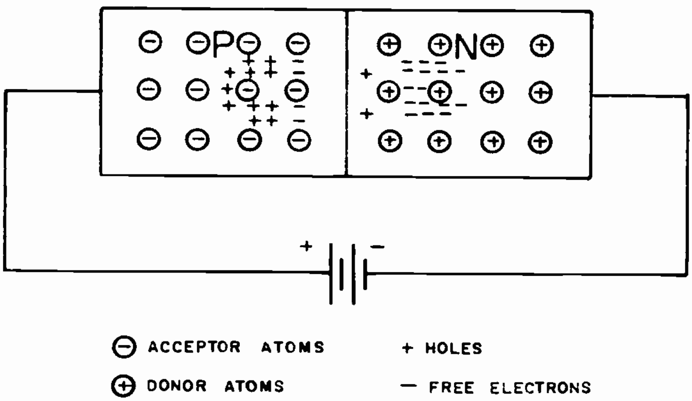


FIG. 20.—*Forward bias.*

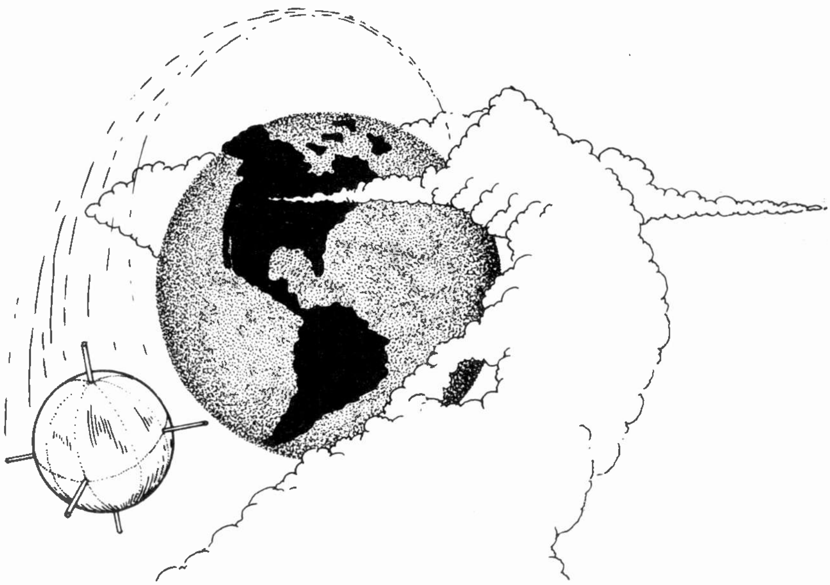


FIG. 21.—*Orbiting of a Sputnik helps prove the soundness of electron theory.*

wonders accomplished by adhering to it.

A more or less extensive explanation of the electron theory can be found in virtually every textbook dealing with the subject of electronics. An understanding of the theory is considered prerequisite to any study of electronics.

However, the Electron Theory is exactly what it has always been proclaimed to be—a *theory*.

It is a theory, scientifically arrived at by learned men, to explain natural physical phenomena which are so ephemeral that human senses are too slow and too dull to study them directly.

That the theory is basically sound cannot be denied. Far too many sciences and industries have been built upon the acceptance of its soundness. This includes radio and television, of course, but it also extends to the newer fields of photoelectronics, guided missile, and the Sputniks and space ships already traveling through empty space outside the Earth's atmosphere.

The point we are trying to emphasize is that despite what we are going to tell you, we do not want you to acquire the idea there is something unsound about the Electron Theory. There is not. The theory has been proven sound so repeatedly it would be folly to suggest a suspicion of its unsoundness.

Yet, continued research into the electron since the theory was first announced has brought to light additional information which suggests that minor amendments could

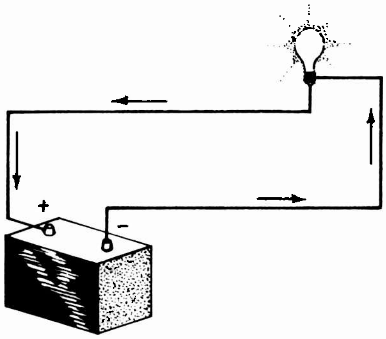


FIG. 22.—*Electrical current flow according to electron theory.*

be made to the theory to include new data not known at the time the theory was given to the scientific world. This applies especially to the concept of "holes" in crystal structure.

Experiments which revolve around semi-conductors, and their use in the construction of transistors, causes us to face up to the necessity of making a slight addition to the Electron Theory. A full explanation of the action inside transistors cannot be achieved by leaning on the Electron Theory without modification.

Suppose we focus our attention for a moment on the matter of current flow. This simple matter has been a bone of contention between two schools of thought within the electrical industry for many years.

Electronic men use what they describe as the "left hand rule" to describe electrical current flow. This theory is simple enough, it merely asserts that electrons move from the negative terminal of a voltage source, flow through the external circuit, then return to the positive terminal of the voltage source. The diagram in Figure 22 illustrates this, although the information in the diagram is not new to you after having studied our preceding lessons.

Men who have worked exclusively with industrial and commercial electricity look upon the flow of electrical current in what they always call the "conventional" manner. The conventional concept of electrical current flow is that it flows from the positive terminal of the source, through the load, then back to the negative terminal.

Despite these discrepancies in concept of a basic electrical action, no really serious conflict has developed between the two schools. Even the better educated electrical men are willing to admit that the older, so-called "conventional" view of current flow is wrong.

16. ROWLAND'S EXPERIMENTS

Although electrical men, as well as those working directly with electronics, have long accepted the soundness of the Electron Theory, there have been certain occurrences which cannot be accounted for by adhering solely to the theory. An outstanding item of this nature is what is commonly known as "Rowland's Experiment."

A scientist by the name of Rowland conducted a series of experiments as far back as 1889 which, when viewed objectively, seem to cast a strong doubt on the full accuracy of the Electron Theory. Certainly, the results of his experiment are such that they cannot be accounted for by the Electron Theory alone and without modification.

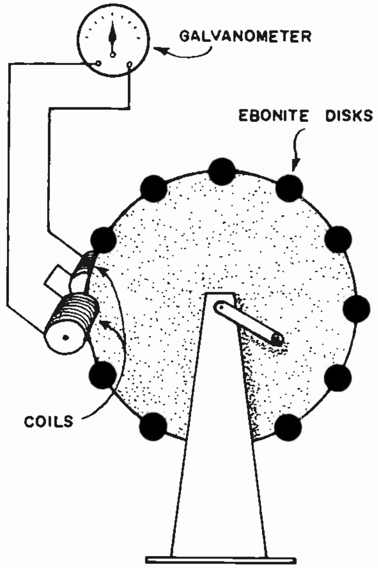


FIG. 23.—*Rowland's Experiment.*

What makes Rowland's experiment so impressive is the fact that they were conducted before the Electron Theory was proposed. Nevertheless, the Electron Theory was advanced, and accepted, despite this known discrepancy, or apparent contradiction.

The basis of Rowland's experiment was a group of ebonite disks mounted on a large disk, and equally spaced around the circumference of the larger disk. The experiment was arranged something like that shown in Figure 23.

Rowland placed an electrical charge on each of the ebonite disks by the time-honored method of rubbing the disks with cat's fur, and by rubbing glass with cat's fur, then placing the electrical charge on the ebonite disks. When so charged, the disks possessed negative electrical charges.

Rowland then caused the large disk to rotate so the charged ebonite disks passed between a pair of coils of wire. The negative charges of electricity on the ebonite disks passing through the coils of wire clearly possessed magnetic fields. This was proven by the galvanometer connected to the coils of wire. The galvanometer proved the direction of the current flow.

There was nothing unusual about this portion of the experiment, since it merely served to illustrate basic principles of electricity which were then known, and which are accepted at this time by all physicists and scientific researchers. It is a commonly accepted fact that movement of electrons constitutes a moving electrical current, and that a moving electrical current can create a magnetic field which can, in turn, induce other electrical voltages.

The negative electrical charges on the moving disks represented moving electrons when the disks on which the negative charges existed were caused to move through the coils of wire.

However, Rowland was not satisfied with stopping at this point in his experiments. He next proceeded to place *positive* electrical charges on the Ebonite disks, and then rotated the large disk in the *opposite* direction. When he did so, he found that the galvanometer registered the same results as before.

His deductions were that: a *positive current flow in one direction is exactly the same as a negative current flow in the opposite direction.*

When we think of this experiment we should be reminded that our knowledge of electrical current in a conductor must be based on the magnetic field produced by a current. We should also keep in mind there are definite limits to our precise knowledge. Actually, we have no uncontroverted methor for proving the direction of current flow; our knowledge is derived from indirect observations.

For the most part it is satisfactory to accept the fact that an electrical current consists of electrons moving through a conductor. We can also accept as a reasonable assumption that the current is also moving from *left to right.*

At the same time, Rowland's experiment forces us to accept the fact that the same results can be attained by causing *positive* electrical charges to move from *right to left.*

Rowland's experiments force us to accept the fact that external evidences around a conductor are such that we cannot tell whether a negative current is moving in one direction or a positive current moving in the opposite direction. Once we bring ourselves to accept this line of reasoning we are faced with what seem to be puzzling things.

Yet, acceptance of this reasoning makes what were puzzling features of transistor action considerably less puzzling.

Researchers are tending to orient their thinking to the concept that electrical current flow may consist of simultaneous movement of electrical charges in opposite directions. They tend to suspect that negative electrical charges may be moving in one direction while positive charges are moving in the opposite.

If we think of positive charges as being "holes" of the character which we have already explained, the entire situation begins to make considerably better sense.

The fact that this is exactly what does occur is fundamental to the theory of transistor action.

17. ELECTRICAL CONDUCTION BY MEANS OF "HOLES"

Scientific circles commonly accept as a fact that an atom consists of a nucleus of positive protons, and—in most elements—a definite number of neutrons. Both protons and neutrons are relatively heavy particles. It is not thought that they play any direct part in electrical current flow.

As we have explained previously, electrons are constantly revolving around the central nucleus. The electrons in their orbits tend to electrically balance the protons in the nucleus.

There are circumstances where an electron in the outer shell of an atom may be free to move from atom to atom, thus bringing about what we know as electrical conduction. This leads to a condition in which an atom may temporarily have an extra electron, or be missing one, thus creating an unbalanced electrical condition.

All this goes to the very heart of the Electron Theory.

While the Electron Theory is accepted without exception, and nothing said in this lesson should be interpreted as casting doubt on its validity, we must also accept the fact that there are conditions in which an atom has one electron less than normal to a balanced condition. This leads us to the "hole" theory.

Although it is normal for us to think of a hole as being empty space, and without physical existence, in transistor work we look upon "holes" as though they had a physical presence.

To emphasize this a little more strongly, we will point out that scientists have come to look upon these holes as going through physical movements and exercises as though they had an actual physical presence, or physical entity.

It is proper that we regard them in this manner. This concept of the physical presence of "holes" goes to the heart of transistor theory.

Research scientists act as though holes were endowed with a definite mass, and possessed a definite electrical charge. They treat them as though they move with a definite velocity, and with associated energy.

In short, they treat a hole exactly as though it were a particle of matter.

They have proven to their own satisfaction that holes are acted upon by magnetic fields and electric charges in a manner equivalent to that in which electrons are acted upon by similar charges.

Holes are attracted—in transistor work—by negative electrical charges. This being so, it is acceptable to look upon holes as being positive electrical charges.

18. POINT TRANSISTOR

After laying the groundwork with the foregoing explanation we will now turn our attention directly to the action which takes place in a transistor.

There are two principal types of transistors. One is called a *point transistor*. The other is called a *junction transistor*.

The junction transistor has such superiority over a point transistor that the latter seems to be gradually disappearing from use. Nevertheless, the point transistor was the first type developed, and it is logical to explain how it works before getting into the junction types.

A point transistor uses a Germanium crystal. It is usually an N-type crystal. This, of course, is a Germanium crystal contaminated with donor type atoms.

The actual crystal for use as a point type transistor is usually quite small. It is about .05 inch in length and something like .02 inch thick. In size, this is about that of a very thin piece of pencil lead.

The pellet is mounted on some type of base material. The base material makes electrical contact to the crystal.

The arrangement of the electrical circuits to the crystal is pretty much as shown in Figure 24. The base itself is not shown clearly in Figure 24, but the electrical connections can be visualized.

In addition to the base connections to the crystal, there are connections to two other locations. One is called the *collector*, and the other is called the *emitter*.

These are two terms you should memorize. You will be running into them repeatedly while working with transistor circuits.

The diagram in Figure 24 shows electrical connections to the collector and the emitter, but no contact has been made by either of them to the crystal.

Before we make contact with the crystal by either of these elements we want to point out something so it will hold your attention. On the surface of the N-type crystal in Figure 24 are shown a number of electrons. These are represented by a group of minus signs.

These are called *surface-bound electrons* by physicists specializing in semi-conductor work.

They have found through their experimentation that some electrons diffuse to the surface of a Germanium pellet, then lose their way. They cannot find their way back to the interior of the crystal.

In effect, the electrons form a skin-like covering for the crystal, covering the outer surface.

Through their research they have found that the surface-bound electrons tend to combine with a layer of donor atoms just below the surface in a manner which tends to form a "potential hill."

If the electrode which is called the collector is then touched to the surface of the crystal, it causes an electrical potential to be applied to the crystal. However, no conduction occurs. The potential hill prevents any current flow. It prevents current flowing despite the fact a relatively high voltage may be applied by some external source.

The nature of this condition is shown in Figure 25.

The diagram in Figure 25 shows the collector electrode touching the surface of the crystal, but no current is flowing. Part of the reason is that the collector electrode has a negative potential, and that is not conducive to conduction under these conditions.

Now let us take another step forward. Let us turn our attention to Figure 26.

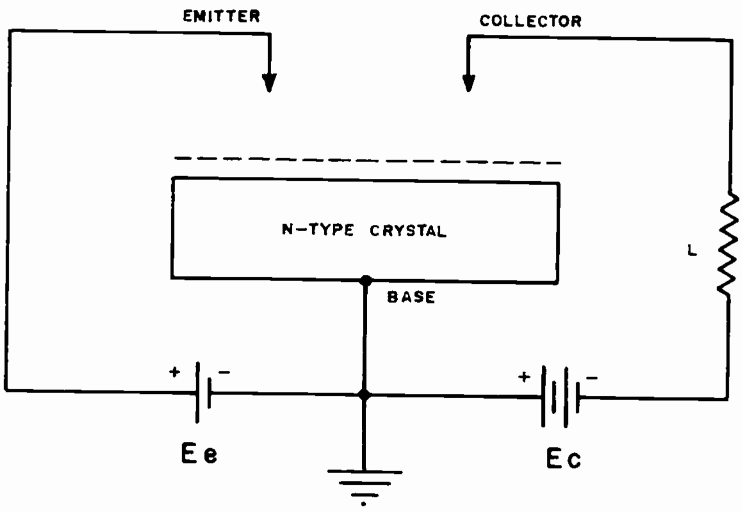


FIG. 24.—*Electrical connections to point transistors.*

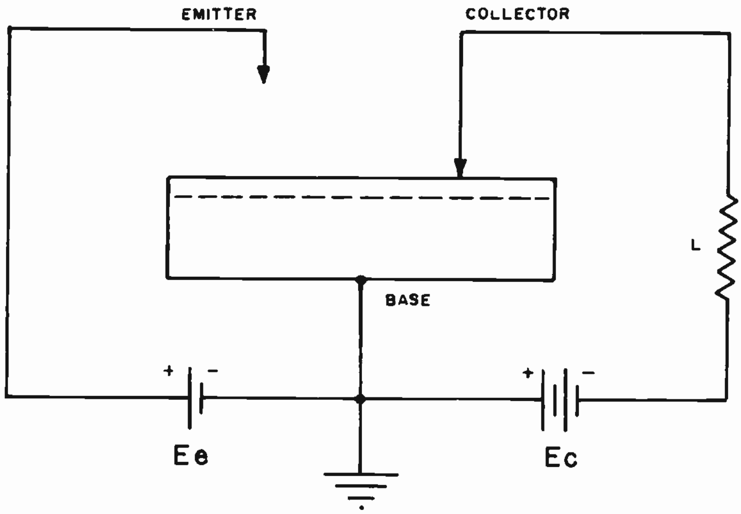


FIG. 25.—*No current can flow in collection circuit under these conditions.*

The action in the drawing in Figure 26 has been deliberately exaggerated. This has been necessary in order to make the action clear.

On the drawing the distance between the contacts of the emitter and the collector electrodes appears to be relatively large. Actually, it is quite small. The two contacts must be very close together.

The points of the fine contacts are very small. They should not be more than 5 thousandths (.005) of an inch in diameter. Nor should the spacing between the two electrodes exceed 2 thousandths (.002) of an inch at points of contact.

It can be quickly recognized that from a mechanical viewpoint these dimensions are quite small. On the other hand, in terms of electron size, they are very large.

Now for the electrical action.

When the emitter touches the surface of the crystal it immediately begins applying a positive voltage. The positive voltage tends to inject "holes" into the crystal, which has the effect of removing electrons from the surface of the crystal.

Note the theory of this hole injection. This action is important in transistor action.

The holes tend to diffuse through the crystal. They move toward the collector, as is indicated by the diagram in Figure 26. The movement toward the collector electrode stems from the negative voltage on the collector electrode.

This action is the reason why it is necessary to space the collector and emitter electrodes so closely together. As the holes tend to move toward the collector they meet free electrons in the crystal surface structure. When this occurs the holes and electrons tend to cancel each other.

When the distance between the two electrodes is kept small the recombination of holes and electrons is held to a minimum.

When holes from the emitter reach the potential hill which surrounds the collector they recombine with the surface-bound electrons. This action tends to reduce the potential hill.

When the potential hill around the collector is reduced it becomes possible for the collector to inject free electrons into the crystal surface. These electrons then make their way—for the most part—toward the positive base. This action tends to greatly increase the collector circuit current.

Note this carefully: when the potential hill around the collector is broken down by action of the holes from the emitter, some of the electrons injected into the crystal by the collector make their way to the emitter. Such action is unavoidable, and is not normally a matter of major concern.

However, a much larger number of them make their way to the base.

The result of this action is that a small current in the emitter circuit acts to control a much larger current in the collector circuit.

We can compare this action with that which takes place in a vacuum tube. In the case of a vacuum tube a small AC voltage on the control grid sets up conditions which cause a major variation in the anode current, and thus bring about major variations in the anode voltage.

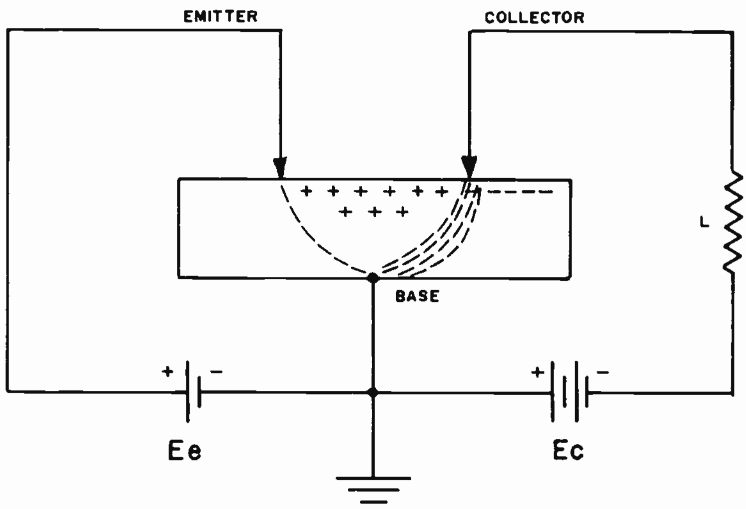


FIG. 26.—*Holes from emitter act to neutralize the potential hill and permit conduction in collector circuit.*

However, in the case of a transistor, a small varying current in the emitter circuit acts upon the crystal in such manner as to cause wide variations in the current which flows in the collector circuit. It is logical to make a rough comparison between the emitter circuit of a transistor and the grid circuit of a vacuum tube. In like manner it is practical to make a rough comparison between the collector circuit of a transistor and the anode circuit of a vacuum tube.

However, we want to warn you against jumping to conclusions too quickly. There is justification for *rough* comparisons between the elements of a transistor and those of a vacuum tube. But you should not fall into the error of thinking they are *directly* comparable. That is not exactly true.

19. GAIN IN A POINT TRANSISTOR CIRCUIT

When working with transistors we are normally more interested in the *changes* which occur in the currents in the various circuits. We are commonly more interested in such changes than in the total volume of current flowing. This is similar to the situation we find in vacuum tube work.

The best way to study such varying currents is to apply some actual figures to them, then see what happens.

Stage *gain* in a transistor circuit is represented by the Greek letter Alpha, much as is true in a vacuum tube amplifier circuit. However, for purposes of our convenience we can use the letter *A*.

In studying a typical point-contact transistor we will find that an increase in the emitter current amounting to 1 milliammeter will bring about a change in the collector circuit amounting to approximately 2.5 milliammeters.

At first glance this would indicate a very small gain in the transistor circuits. The collector current increases only 2½ times as much as the emitter, which suggests a very small gain.

However, first appearances are deceiving, just as they are in so many other cases.

The emitter circuit is a low resistance circuit. Its resistance is on the order of some 300 ohms.

The collector circuit, however, has a much higher resistance. A typical resistance is on the order of 20,000 ohms.

This begins to open up possibilities. If we can use a circuit with an impedance of only 300 ohms to bring about current changes two and a half times greater in a circuit where the resistance is some 60 times greater, we have the possibilities of considerable gain. Actually, the impedance of the collect circuit is 67 times greater than that of the emitter circuit.

Getting back to the fundamentals of Ohm's Law, we know that the voltage in a circuit—any circuit—is equal to the current multiplied by the resistance.

In the emitter circuit we have a voltage which amounts to the emitter current multiplied by the emitter resistance. Since the emitter current amounts to 1 milliampere (equal to .001 ampere) and the emitter circuit resistance is 300 ohms, it works out that the voltage is equal to .001 x 300, which is equal to .3 volts.

There are 2.5 milliamperes (equal to .0025 amperes) in the collector circuit. There are also 20,000 ohms of resistance. Multiplied together, this works out to .0025 x 20,000, which is equal to 50 volts.

Thus we have 50 volts in the collector circuit as compared with .3 volts in the emitter circuit, a theoretical gain of approximately 165 times.

This, of course, is a theoretical gain. Part is necessarily lost in handling the currents in the actual circuits, but the gain is substantial.

Gain in a transistor circuit can also be calculated by comparing the power gain. This is accomplished by calculating the power in the emitter and then the power in the collector circuit, and finally finding the ratio between them.

In describing this action we want you to keep it clearly in mind that we are describing *theoretical* actions, and theoretical gains. Actual gain is necessarily always considerably less.

You will also be interested in the technical manner in which gain is shown in formula form for a transistor. The formula is:

$$A = \frac{i_c R_c}{i_e R_e}$$

What this means is that the gain of a circuit is merely the ratio of the collector current and resistance, and that of the emitter current and resistance. Again, this is theoretical, and is subject to modifications.

We will go into this subject a little more deeply when we get into *junction* transistors.

ITEMS OF SPECIAL INTEREST

Transistors are important in electronic work.

Transistors are capable of performing many of the jobs formerly handled exclusively by vacuum tubes.

Despite the fact that transistors are capable of doing many things which only vacuum tubes were formerly able to do, they should not be thought of as replacing vacuum tubes.

Transistors should be looked upon as supplementing vacuum tubes, rather than replacing them.

Transistors, and other crystals, are able to rectify AC currents and voltages into uni-directional DC.

Transistors are also capable of acting as amplifiers under proper conditions.

Transistors do not require filament heating current, therefore do not need such an extensive power supply as vacuum tubes.

Transistors operate on much lower anode voltages than vacuum tubes, thus reducing again their need for a power supply.

TRAINING SERVICE

Massey Technical Institute, Inc. MTI Jacksonville, Florida

Print or Use Stamp:

Date__________

Name__________ Student No.__________

Address__________ Zone_____ Grade__________

City__________ State__________

Lesson 43

TRANSISTOR ACTION

Part 1.

Underline or fill in the ONE correct answer, just as in previous lessons.

1. If a point transistor has 1 milliampere and 350 ohms in the emitter circuit and 2.2 milliamperes and 18,000 ohms in the collector circuit, what is its theoretical gain?
   a. 73. b. 113. c. 153. d. 165. e. 179. f. ________.

2. How much current flows in the collector circuit of a transistor when none is flowing in the emitter circuit?
   a. more than normal. b. slightly less than normal. c. none at all.

7. The term "A negative coefficient of Resistance" means:
   a. resistance increases linearly with temperature.
   b. increasing temperature means less resistance.
   c. there is a geometric progression of temperature and resistance.
   d. the temperature is equal to the coefficient of the resistance.

8. If the temperature of a Germanium crystal rises above $80^{\circ}C$:
   a. it becomes an insulator.
   b. its efficiency is an excellent transistor.
   c. it is ruined as a transistor.
   d. it is unchanged.

9. Germanium crystals are:
   a. excellent conductors.
   b. sensitive to light.
   c. perfect insulators.
   d. used in Camera tubes.

10. For transistor work approximately what amount of adulteration is permitted?
   a. 1 part in a million.
   b. two parts in a million.
   c. 1 part in ten million.
   d. 1 part in 100 million.

11. An acceptor atom is one that:
   a. has an extra electron in the outer shell.
   b. is negative.
   c. is missing an electron in the outer shell.
   d. has changed elements.

12. Adulterating Silicon or Germanium with Arsenic endows it with:
   a. positive characteristics.
   b. negative characteristics.
   c. properties of an insulator.
   d. doubled atomic weight.

13. Atoms are linked to each other by:

a. electrons which provide valence bonds. b. magnetic attraction.
c. magnetic lines of force. d. attraction between protons.

14. Atoms which have many electrons arrange them in:
a. layers, with a single electron in each layer. b. clustered groups.
c. orbits, with the orbits in shells. d. a single shell.

15. Two big advantages of transistors over vacuum tubes are:
a. cheaper and smaller. b. smaller, and require less power.
c. better amplifiers and less expensive. d. fewer do more work.

Part 2.

16. If an inductance coil has a reactance of 25 ohms at 60 cycles, and a resistance of 7 ohms, what is its total impedance? ______________.

17. Small insulating tubing used to provide additional insulation for wires and pigtails in electronic circuits is called ______________.

18. The mixer tube on most broadcast receivers is a ______________.

19. If a 470 mmf capacitor is needed, which of the following would be the best substitute:
a. .047 mfd. b. .0047 mfd. c. .0005 mfd. d. .0000005 mfd.

20. Give the formula for a pentode resistance coupled amplifier ______________.

d. total equals the amount normally carried in both circuits.

3. The empty space normally occupied by an electron which is missing from an atom is referred to in semi-conductor work as a:
   a. vacancy. b. well. c. negative charge. d. hole.

4. Rowland's experiment proved that:
   a. the electron theory was immutable.
   b. electrons always moved through a circuit from positive voltage to negative voltage.
   c. positive charges moving in one direction produced effects similar to those of negative charges moving in the opposite direction.
   d. electrons are sometimes positive.

5. What happens to the two crystals when a positive crystal is joined to a negative crystal?
   a. they are immediately neutralized.
   b. surplus electrons from the N-type crystal fill the holes in P-type.
   c. a potential hill is built up which prevents further neutralization.
   d. N-type changes to P-type, and P-type changes to N-type.

6. When a P-type crystal is joined to an N-type crystal, current flows through the junction when:
   a. a negative voltage is applied to the N-type crystal and a positive voltage to the P-type crystal.
   b. N-type crystal is connected to positive voltage and P-type to negative voltage.
   c. only AC voltage is applied.
   d. the bias is reverse.

One of the first practical uses for transistors was in the field of hearing aids.

Eliminating the filament batteries for hearing aid tubes by replacing the tubes with transistors made it possible to radically reduce the size of the hearing aids.

Hearing aids can be made so small that they can be fitted into the frames of eye-glasses.

Transistors have found ready use where the demands of space and weight are such that they must be reduced to the minimum.

Transistors are not so superior, in comparison with vacuum tubes, where there is an abundant supply of electrical power.

Transistors are not capable of providing the high degree of amplification that can be handled by a vacuum tube.

Transistors cannot withstand high temperatures so easily nor readily as vacuum tubes.

Unless abused, transistors commonly have longer lives than vacuum tubes.

Transistors depend for their usefulness on the peculiar properties of valence bonds in certain types of crystal structures.

Semi-conductor crystals are the starting point for the manufacture of transistors.

Germanium and Silicon crystals are the types most frequently used for transistors.

Germanium and Silicon crystals must be refined to a high degree of purity before they are suitable for transistors.

One of the problems encountered in transistor manufacture, and one of the things contributing to their high cost, is the difficulty of attaining the proper degree of purity in the crystals.

Point transistors have three basic electrodes. They are the emitter, the collector and the base.

The emitter circuit of a transistor is roughly comparable to the control grid in a vacuum tube, although they are not directly comparable.

The collector circuit around a transistor is roughly comparable to the anode circuit of a vacuum tube.

The gain of a transistor is figured as the ratio of the collector voltage to the emitter voltage.

The gain of a transistor can be figured from this formula:

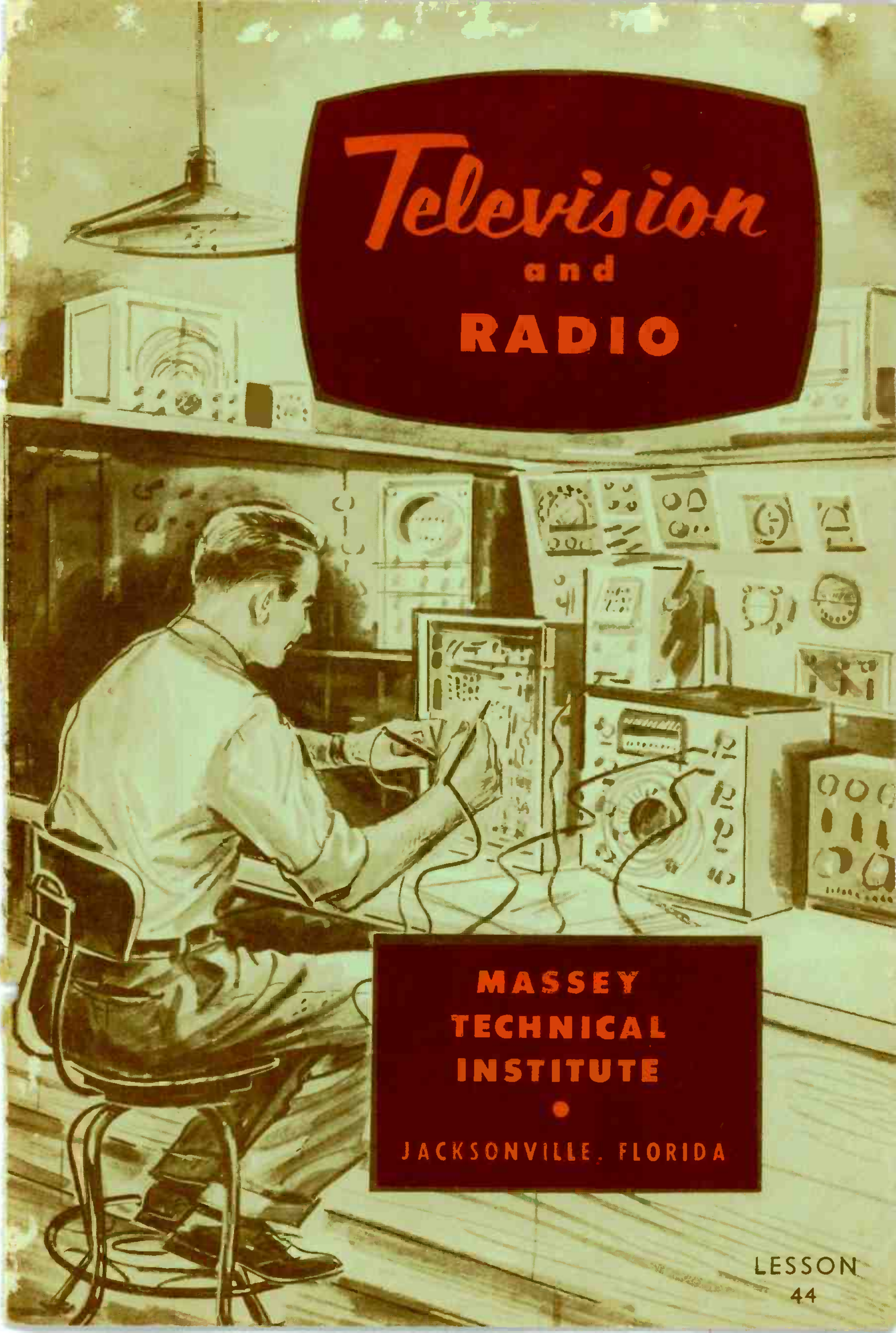
$$A = \frac{i_c R_c}{i_e R_e}$$

MASSEY TECHNICAL INSTITUTE, Inc.

MTI

Jacksonville, Florida

Television and RADIO



MASSEY TECHNICAL INSTITUTE

JACKSONVILLE, FLORIDA

LESSON 44

PEOPLE DON'T PLAN TO BE FAILURES
- - THEY JUST FAIL TO MAKE PLANS.

TRAINING SERVICE

TRANSISTOR CIRCUITS

CONTENTS:—Introduction—Junction Transistors—Electrical Action in a Junction Transistor—P-N-P Transistors—Junction Transistor Gain—Power Gain—P-N-P-N Transistors—Tetrode Transistors—Electronics of Transistors—Transistor Symbols—Circuit Connections—Grounded-Emitter Connections—Frequency Response of Grounded-Emitter Transistor—Grounded-Collector Transistor—Mechanical Standards for Transistor Construction—Transistors in Radio Receivers—Items of Special Interest.

1. INTRODUCTION

In our preceding lesson we succeeded in introducing you to what are known as *point-contact* transistors. This was the first type of transistors.

Point-contact transistors have certain inherent limitations which caused research engineers to seek different and better types.

For one thing, a point-contact transistor requires an intense electrical field if it is to work properly. Yet, if the battery voltage applied to the transistor is raised to a level high enough to produce what are considered optimum results, the high potential tends to damage the transistor.

Such is not, of course, desirable.

If a high voltage is applied to the emitter circuit it causes an excessive current to flow in that circuit. The reason it does so results from the comparatively low resistance which normally exists in that circuit.

When the current rises above a definite level it acts to burn out the transistor, thus making it worthless.

Even in the collector circuit, the voltage must be held to relatively low levels. Compared to voltages used in vacuum tube work, the voltages used on transistors are extremely low. Voltages on transistor elements rarely exceed 24 volts, and usually are lower than that.

If the voltage in the collector circuit is permitted to rise above some definite limit it causes a breakdown of the potential hill in the transistor. When it does so it ruins the usefulness of the transistor for the purpose for which it was intended.

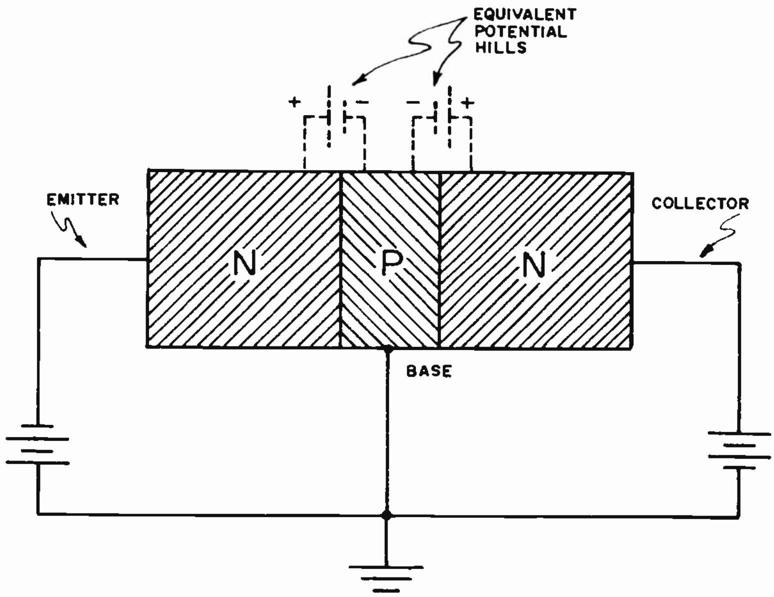


FIG. 1.—*How crystals are joined in junction Transistor.*

In the case of point-contact transistors, the act of placing the contacts of the cat-whisker points extremely close together serves to partially overcome the limitations of transistors with respect to voltage. Even so, efficiency of point-contact transistors is quite low.

However, there is minimum beyond which it is not practical to go. When a pair of contact points are carefully placed so they are no more than two thousandths (.002) of an inch apart, that is just about as close as it is practical to go. Thickness of this paper is on the order of about .002 inch. It is not practical to place the points much closer together.

When the contact points are placed much farther apart than .002 inch it has an effect on the frequency response. The farther apart the points are spaced the lower is the top frequency the transistor is capable of handling.

2. JUNCTION TRANSISTORS

Some of the limitations of point-contact transistors have been overcome by using a different principle in transistor construction. Instead of permitting a couple of metal contact points to rest on the surface of the crystal, three or more crystals are joined together so no two crystals with the same polarity are in contact with each other.

Such types have been named *junction transistors.*

Junction transistors are not plagued by the close spacing of the contact points which creates so

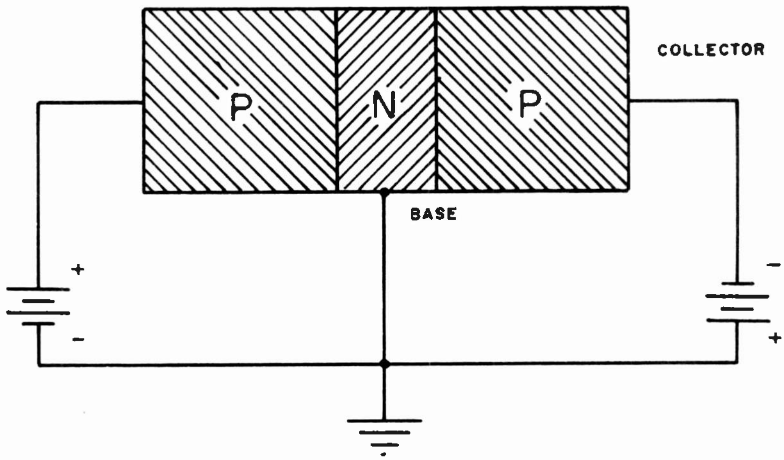


FIG. 2.—*P-N-P Transistor.*

much trouble in point transistors. Instead of the points, crystals with different polarities are joined together. The drawing in Figure 1 helps show how the crystals are arranged.

The transistor in Figure 1 is what is known as an N-P-N Junction Transistor. This name comes from the arrangement of the different crystals. The name indicates that a P-type crystal is sandwiched between a pair of N-type crystals.

In the case of the transistor shown in Figure 2, we find an N-type crystal sandwiched between a pair of P-type crystals. That kind of transistor is known as a P-N-P junction transistor.

Returning our attention to the transistor in Figure 1, we should first understand that the drawing is many times larger than an actual transistor. The bits of crystal are quite small, and the total overall size of the transistor is much less than that occupied by the drawing.

When the crystals are joined together, as shown in Figures 1 and 2, *potential hills* are created at the exact location where the junction between adjacent crystals occurs. This is indicated by the potential hills shown in the drawing, the manner of indicating taking the form of tiny imaginary batteries.

As a matter of fact, a small fractional voltage does exist between the adjacent crystals, the polarity of which is the reverse of the normal polarity of the two crystals. As indicated by the battery symbols in Figure 1, the positive terminal of the *potential hill* battery is located on the negative crystal, while the negative terminal of the potential hill battery is located on the positive crystal.

If this action is not entirely clear, we suggest you go back to your earlier lesson where we first explained this matter of *potential*

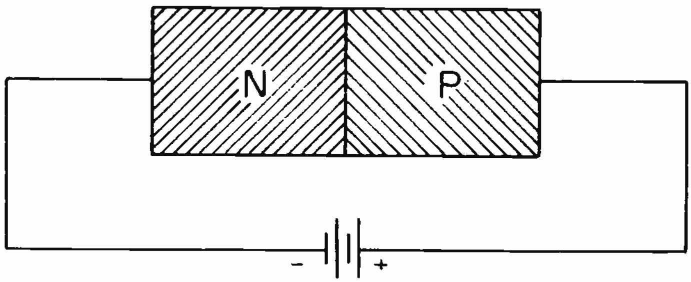


FIG. 3.—*Simple diode crystal.*

hills. If you study the explanation given there, in conjunction with the one we are giving now, we believe the action will become somewhat more clear to you.

However, this is an action which is new to most electrical men, and the minds of many are not entirely ready to accept the explanation. However, as you pursue your studies more deeply into transistor theory and action we think you will find the explanations more easily acceptable.

3. ELECTRICAL ACTION IN A JUNCTION TRANSISTOR

When you get right down to it, a junction transistor is little more than a double crystal diode.

You will recall from what we told you in a preceding lesson, that a P-type and an N-type crystal joined together in the manner shown in Figure 3 are capable of acting as a rectifier. This means that when two dissimilar crystals are joined as shown in Figure 3 electrical current is able to flow through the crystal in one direction, but it encounters high resistance when attempting to flow in the opposite direction.

Now, suppose we take a pair of such crystal diodes, then connect the two positive crystals together as shown in Figure 4. In that case we have all the essentials of a junction transistor.

If you study the diagram in Figure 4 quite closely you will see there is a close similarity between it and the diagram in Figure 1. The main difference being that the transistor in Figure 1 has only a single P-type crystal, whereas Figure 4 has two of them connected together.

There is another practical difference. The P-type crystals in the circuit in Figure 4 would probably be much thicker than the single P-type crystal in Figure 1.

In an actual transistor it is essential to good operation that the center crystal of a junction transistor be extremely thin. Therefore, despite its electrical similarity, the circuit in Figure 4 is scarcely a practical one.

Let us take a look at the construction of a junction transistor, and note one special improvement which makes it decidedly superior to a point transistor.

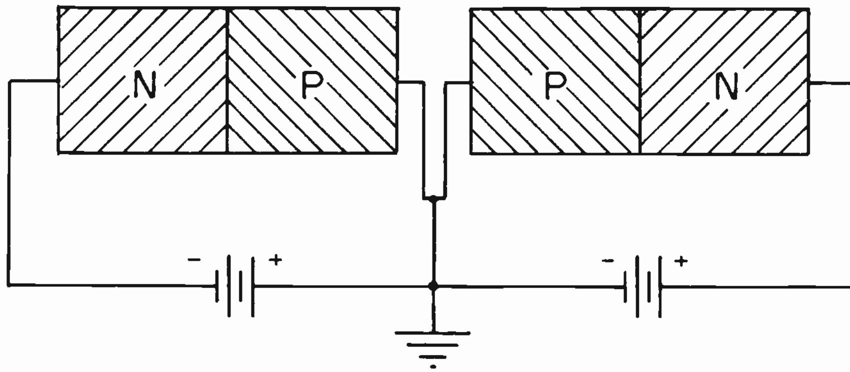


FIG. 4.—*Two diodes connected together.*

You will remember from our description of point transistors that two of the three contacts consisted merely of fine cat-whiskers pressing against the crystal. Necessities of the situation being what they are, those two electrodes must consist of extremely fine wire. This limits their current carrying ability.

Further, it is difficult to anchor the fine wires permanently to the crystal. In most cases the connection depends on spring tension continuing to hold the contacts in place.

We find a different situation with respect to junction transistors. For one thing we start with a much broader area of contact between the various elements of the transistor crystals. Also, it becomes possible to make soldered electrical contacts directly to the crystal. This works to make greater electrical stability.

It makes a junction transistor capable of handling larger voltages and currents. It also increases their power handling ability. All of this acts to increase the usefulness of transistors.

Getting back to our comparison of a junction transistor with a pair of crystal diodes. While the analogy is apt, the internal conduction and action is different from that which occurs in ordinary crystal diodes.

Going another step further, the action is also considerably different from that which takes place in a point-contact transistor.

Let us first take a look at a junction transistor when no external voltage is applied to any of the terminals. This is the condition which exists when no connection is being made to an electrical battery.

In such a condition we find potential hills existing at the junction between each pair of crystals. These potential hills are represented in Figure 1 by the "equivalent" battery signs shown in the form of dashed lines.

Polarities of the potential hills are shown by signs adjacent the simulated batteries.

It is possible to bring about electrical conduction by placing a negative voltage on the emitter electrode. This causes conduction through the left hand portion of the circuit, represented by the left N-type crystal, the center P-type crystal, and the connection through the base back to the battery.

It is important to remember that *the number of electrons which pass through the barrier junction is proportional to the magnitude of the emitter voltage.*

What this works out to is that the act of varying the emitter voltage causes a variation in the number of electrons *entering* the P-type crystal.

In short, varying the emitter voltage causes a variation in the number of electrons entering the P-type crystal.

An alternating voltage placed on the emitter electrode causes a varying number of electrons to pass through the junction barrier into the P-type crystal. This being true, it opens wide possibilities for using an alternating current and voltage on the emitter to control conditions in the P-type center crystal.

Here are the reasons:

Under the influence of the voltage and current applied to the emitter electrode, electrons pass through the barrier from the N-type crystal at the emitter into the P-type crystal in the center of the transistor. When they do so, some of the electrons combine with the holes in the P-type crystal, and both are neutralized.

Even more of the electrons, however, pass through the P-type crystal and through the second barrier between the P-type crystal and the N-type crystal on the right in Figure 1. This permits them to reach the collector electrode, and emerge into the collector circuit in the form of electrical current.

You may wonder why so few electrons are neutralized in passing through the P-type crystal in the center.

It is because the P-type crystal in the center is kept quite small—or to be more accurate, is made quite thin. The thinner the P-crystal in the center, the fewer electrons are neutralized during passage through it.

Another thing that aids passage of electrons through the P-type crystal, and acts against their neutralization, is their accelerated movement. Electrons are accelerated by the potential hill which exists between the center P-type crystal and the N-type crystal at the collector electrode.

Let us take a good look at this action to be sure we have it clear.

Some external voltage is needed on the emitter electrode to enable the electrons to overcome the potential hill between the emitter electrode crystal and the base crystal. Polarities of the two crystals—and the opposing polarities of the potential hill which exists at their junction—prevent normal movement of electrons through the junction. However, when an external voltage is applied to the emitter, electrons can move readily from the N-type emitter crystal into the P-type base crystal.

This external voltage can be thought of as being in the nature of a bias voltage.

Once electrons reach the P-type base crystal, however, their movement is actually accelerated toward the N-type collector crystal.

This is especially true if the base crystal is very thin, since this makes the potential hill between the base crystal and the collector crystal even more effective in its influence on the movement of the electrons.

Once electrons reach the N-type collector crystal they instantly release other electrons to flow through the collector electrode's external circuit to its voltage source.

4. P-N-P TRANSISTORS

Our foregoing explanation revolved around the action which occurs in an N-P-N junction transistor. This is a transistor of the type shown in Figure 1. As the name suggests, it consists of a P-type crystal sandwiched between a pair of N-type crystals.

A transistor constructed in this manner, possessing larger cross-sectional contact between the surfaces of the several crystals with their differing polarities, is capable of handling larger currents than the earlier type point-contact transistors. Higher voltages can also be applied to them.

Further than this, they are able to handle larger volumes of power.

Without getting into an involved technical discussion at this point in our discussions, we can point out that a slightly different type of junction transistor has been found to possess certain technical advantages. This is one which has an N-type crystal sandwiched between a pair of P-type crystals.

Such a transistor is shown in Figure 2.

A comparison of the voltage polarities applied to the transistor in Figure 2 discloses the fact that almost everything about the transistor and its circuits is reversed from the manner in which it was arranged in the transistor circuit in Figure 1.

A transistor of the type shown in Figure 2 is known as a P-N-P junction transistor.

To better understand the reasons for the differences in the manner of applying electrical voltages to the two types of transistors, suppose we take an imagined look into the inside of each. We will first look at an N-P-N transistor, such as the one in Figure 1. The arrangement of atoms, electrons and holes inside the transistor can be imagined as being as shown in Figure 5.

We have shown the atoms, electrons and holes in the several crystals by means of symbols. There is nothing sacred nor significant in our selection of these particular symbols to represent the things we are discussing. We have merely chosen them because they seem convenient. We could have represented them in some other manner, and would have conveyed the same meaning.

Note that the emitter electrode and crystal are biased in the *forward direction.* This is in keeping with normal transistor practice.

The voltage applied to the collector electrode and crystal by its battery, however, acts to bias them in the *reverse direction.* You can see this by studying the diagrams in Figure 5.

Now let us turn our attention to the action inside a P-N-P junction transistor. This is one like that shown in Figure 6.

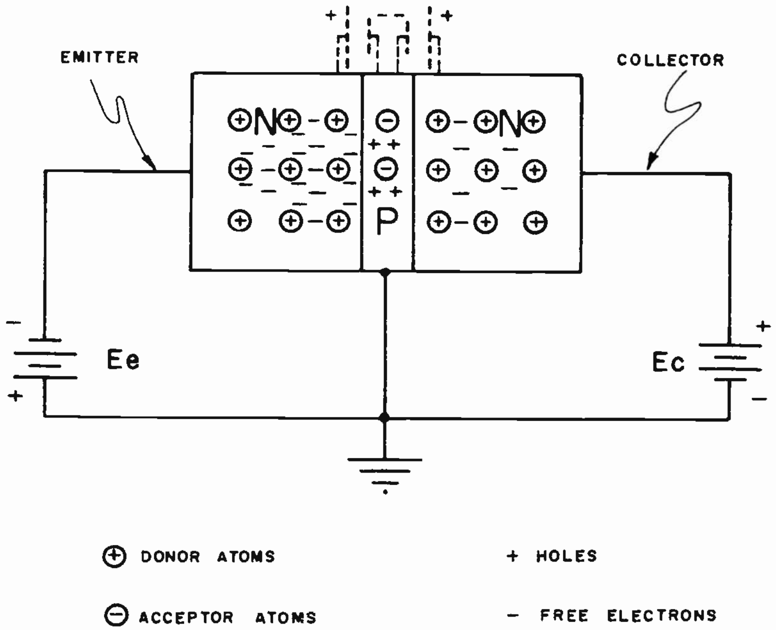


FIG. 5.—*Electrical arrangement in N-P-N Transistor.*

The first thing which catches our attention in connection with a P-N-P junction transistor is that conduction occurs through the movement of "holes" rather than through movement of electrons. The emitter injects "holes" into the P-type crystal to which it is connected, and it is the movement of "holes" rather than electrons which accounts for current flow.

Applying a positive voltage to the emitter electrode causes it to inject holes into the P-type emitter crystal. These holes move toward the junction barrier between the P-type emitter crystal and the N-type base crystal.

Note that this action is similar to that which we described in connection with the N-P-N transistor, except that it is holes moving instead of electrons.

The emitter voltage should be sufficient to cause the holes to move toward the barrier junction. The emitter voltage acts to overcome the opposing voltage of the potential hill which exists between the two crystals. When it does so, some of the holes from the emitter crystal pass into the N-type base crystal.

This emitter voltage is DC in character, and can be looked upon as a bias voltage. The AC voltage of the signal is then superimposed upon the bias voltage.

If you have followed our explanations carefully with respect to the two types of junction transistors you have arrived at the reason for the "emitter" being so named. It is because in one case it acts to "emit" electrons into the

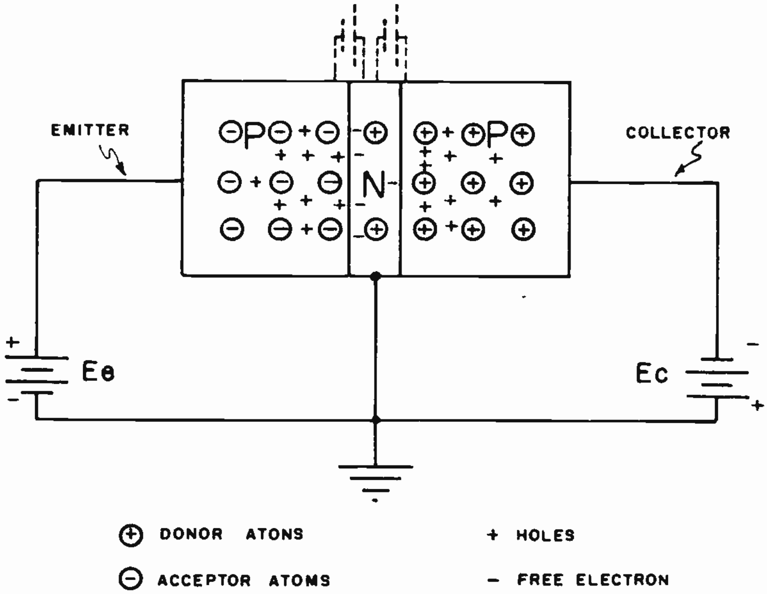


FIG. 6.—*P-N-P Transistor.*

base crystal, and in the other case it acts to emit holes into the base crystal.

In many ways, the electrons or holes emitted into the base crystal serve a purpose closely similar to that of a control grid in a vacuum tube. The emitter, itself, being biased, is closely analogous to a tube grid.

Getting back to our discussion of the action which occurs in a P-N-P transistor. When the emitter injects some holes into the N-type base crystal, they act to combine with free electrons in the base crystal, and thus neutralize each other. This is similar to the action we previously described with respect to the N-P-N transistor, with the exception that holes are now involved rather than electrons.

During the course of the passage of holes through the base crystal, some combine with free electrons and are neutralized. We have already mentioned this.

But the percentage of holes which are neutralized in this manner is relatively small—percentage wise. Possibly five percent of the total holes which move from the emitter crystal toward the collector crystal are neutralized in this manner. The rest reach the collector crystal to perform the duties expected of them.

Once the holes reach the potential hill at the junction between the base crystal and the collector crystal, the polarity of the potential hill is such that movement of the electrons is accelerated.

The action in the collector crystal then becomes something like this:

As a hole crosses the potential hill into the collector crystal, it sets up an electrical action which causes the collector electrode to emit an electron into the collector crystal. It does so to neutralize the holes arriving through the base crystal from the emitter crystal.

For every hole that is neutralized through recombination with an electron, there is a new electron detached from its co-valent bond near the emitter electrode. Such detached electron moves through the emitter crystal to the emitter electrode, then leaves the crystal. In doing so, it creates a new hole.

The action then continues. The new holes created in this manner begin moving immediately toward the junction with the base crystal. The resulting action is such that a constant flow of holes is maintained from the vicinity of the emitter electrode to the collector crystal.

When first described, all this action seems a bit involved, and a bit complicated. As we progress with our studies these things will fall into place, and you will understand them better.

Because of this we suggest that after you have finished a few additional lessons it will be advisable for you to return to this lesson for a review. We think you will learn fully as much from such review as you are learning this first time through the lesson, possibly more. Some of the terms will not be so new.

5. JUNCTION TRANSISTOR GAIN

By this time it should begin to be reasonably clear that junction transistors are capable of handling more current and voltage than point-contact transistors. We have pointed this out earlier in this lesson.

This being true it also seems reasonable that junction transistors are more efficient, and capable of being applied to more practical jobs, than the earlier model

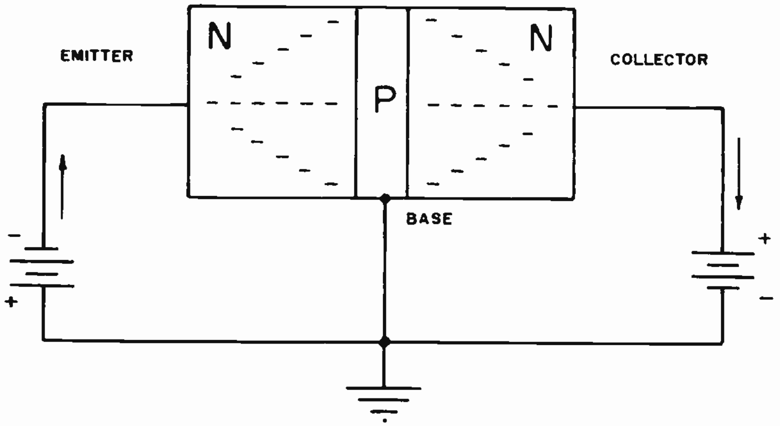


FIG. 7.—*Electron movement in N-P-N Transistor.*

point-contact transistors. That is completely true. It is true of both kinds of junction transistors.

When the action of a junction transistor is first explained it might be thought that a loss is actually occurring in the circuits, rather than a gain. This results from the fact that some loss through neutralization does occur within the transistor circuits.

A little closer analysis of the action shows that a junction transistor is a true amplifier. Furthermore, it does a pretty good job of amplifying.

Let us pause for a moment and make a matter clear which might otherwise become a trifle hazy.

In scientific circles the major current carriers are given considerable attention. In the case of N-P-N junction transistors it is customary to look upon *electrons* as being the major current carriers.

Electrons are looked upon as being the current carriers, largely because in this type transistors electrons are emitted by the emitter electrode, and make their way through the transistor to the collector circuit. This is illustrated by the diagram in Figure 7.

In the case of a P-N-P transistor, the emitter still continues to emit something, and thus set up the requirements of a current carrier. In this case, however, the current carrier consists of *holes* rather than electrons.

Note the drawing in Figure 8. This is intended to show the emitter emitting holes into the emitter crystal, and setting up a current carrier movement toward the collector crystal and electrode. In contrast with the supposed movement of electrons in Figure 7, we see that a movement of holes is occurring in Figure 8.

In both cases, however, the emitter is acting to control the action in the transistor. In each case electrons or holes are moving toward the base crystal, and from there to the collector crystal.

To repeat, in the case of the N-P-N transistor, *electrons* are

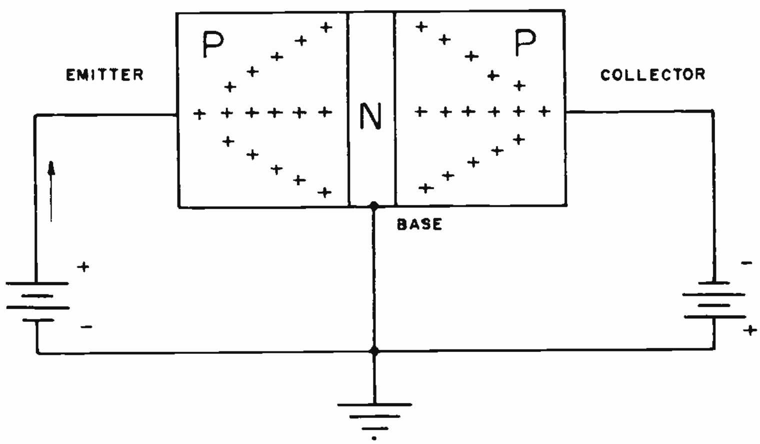


FIG. 8.—*Hole movement in P-N-P Transistor.*

the major current carrier; in the P-N-P transistor, *holes* are the major current carrier.

In a junction transistor the collector current is always smaller than that in the emitter circuit. It is smaller by the factor of the base current, which amounts to about 5%.

As we will explain later, the base current in a junction transistor is actually the difference between the emitter current and collector current.

This works out to a condition in which the collector current is about 95% of the emitter current. The base current is the other 5%.

There is no need to fall into the error of thinking that because the collector current is smaller than the emitter current that a loss is occurring. We showed in our discussion of point-contact transistors that circuit impedance has more to do with gain than anything else.

At that time we learned that voltage gain between the input and output circuits depends more on the resistances in the two circuits than upon the current ratios.

We would like to point out that it is this ability of transistors to transfer resistances from one circuit to another that they owe their name. You will recall that we have already explained that transistors derive their coined name by taking parts of the two words, TRANSfer resISTORS, and combining them into their technical name.

With this brief introduction you are probably prepared for the fact that the respective resistances in input and output circuits of a junction transistor are quite different.

The input, or emitter circuit, resistance of a typical junction transistor is on the order of approximately 500 ohms. The actual resistance may vary to some extent above or below that typical value, but this can still be looked upon as being typical.

This input resistance, you will recall, is somewhat higher than that of a typical point-contact transistor.

Now let us take a look at the resistance in the output circuit of a junction transistor. We find that is on the order of *1 megohm.*

It is easy to see—even at a glance—that the output resistance of the collector circuit of a junction transistor is many, many times higher than the input resistance of the emitter circuit. Further than this, it is many times greater than the output resistance of a point-contact transistor.

In fact, the megohm output resistance of a junction transistor is some 50 times greater than the 20,000 ohms of resistance of a typical point-contact transistor.

From this it can be quickly seen that the voltage gain of a junction transistor can be much greater than with a point-contact transistor. The theoretical gain of a junction transistor approaches that of a vacuum tube.

Suppose we look at an example to better understand just how much gain is involved.

We will act on the assumption that the loss of current to the base crystal reduces the collector current so it is only 95% of the emitter current. This makes its ratio .95.

Since the input resistance of a typical junction transistor is on the order of 500 ohms and the output resistance is on the order of 1,000,000 ohms, the ratio amounts to the larger number divided by the smaller. This means that 1,000,000 divided by 500 works out to 2000.

Note this:

Ratio of the currents of the input and output circuits is .95.

Ratio of the resistances of the two circuits is 2000.

Since these are in inverse ratio, we can multiply 2000 by .95, which gives us a total of 1900. This is the theoretical gain of such a transistor.

Keep in mind that this is the *theoretical gain*. Actual gain, in most cases, is usually considerably less, since full theoretical gain is rarely attained in actual circuits.

However, you will recall from your previous studies that full theoretical gain is not normally attained in vacuum tube amplifier circuits, either.

6. POWER GAIN

Power gain is calculated by squaring the ratio of the currents, then multiplying that by the ratio of the resistances. This is approximately equivalent to applying the basic power formula $P = I^2R$. But it is applied to ratios, not actual values.

Returning to our typical transistor, we square the ratio of the currents. This is done by multiplying .95 by .95. This works out to .9025.

The next step is to multiply by the ratio of the resistances. As we have already learned, the ratio of the resistances is 2000.

This means that to find the power gain we must multiply .9025 by 2000. This works out to:

$$.9025 \times 2000 = 1805.$$

This means that the power gain in a circuit using such a typical transistor would amount to 1805. However, this is, again, a theoretical gain, not necessarily an actual gain. In most cases the actual gain would be considerably less.

You should not fall into the error of believing that all junction transistors have current ratios, resistance ratios and power ratios identical with those we have been using as an example. As a matter of fact, we have simply selected typical values which are easy to use for the purpose of explanation. Actual values in commercial transistors can be considerably different.

In the transistor we have selected as typical, the output resistance of the collector circuit was said to be 1,000,000. It could have been considerably greater.

Many of the newer transistors have output resistances on the order of 3,000,000 ohms, and more.

7. P-N-P-N TRANSISTORS

A four-crystal transistor has been created for the purpose of improving the operating characteristics of earlier models.

The manner in which the various crystals are arranged is shown in the diagram in Figure 9.

This is a type of transistor in which the major current carriers

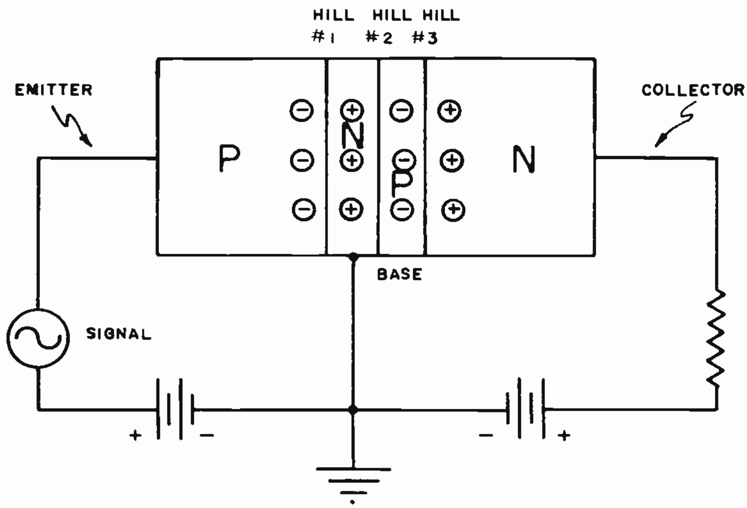


FIG. 9.—*P-N-P-N Transistor.*

are holes. The holes move from the emitter electrode through the two central crystals to the collector electrode. In this action there is not much to distinguish this type transistor from some of the others.

However, the holes, in their movement, are stopped by presence of the No. 3 potential hill as shown in Figure 9. It is the potential hill which exists at the junction between the second P-type crystal and the collector N-type crystal.

The holes pile up at the edge of the potential hill. When they do so they serve to partially neutralize the effect of the potential hill.

When the holes pile up at the edge of the potential hill they act to attract electrons from the vicinity of the collector electrode, and flow from it toward the junction with the adjacent P-type crystal.

When this action occurs some of the electrons are neutralized by recombination with the holes. Most of them, however, come under the influence of the voltage of the second potential hill (Hill No. 2), and move into the central N-type crystal. When they do so they move toward the base.

There is a purpose behind this arrangement, of course. The purpose is to secure a gain in the collector current over that in the emitter circuit.

In a transistor constructed in this manner, and with the electrical connections made as they are in the diagram in Figure 9, the current in the collector circuit is materially greater than in an ordinary junction transistor without the extra crystal. While studying the circuit it will be well to note that the polarity of the collector voltage is opposite that in an ordinary junction transistor circuit, such as the one in Figure 6.

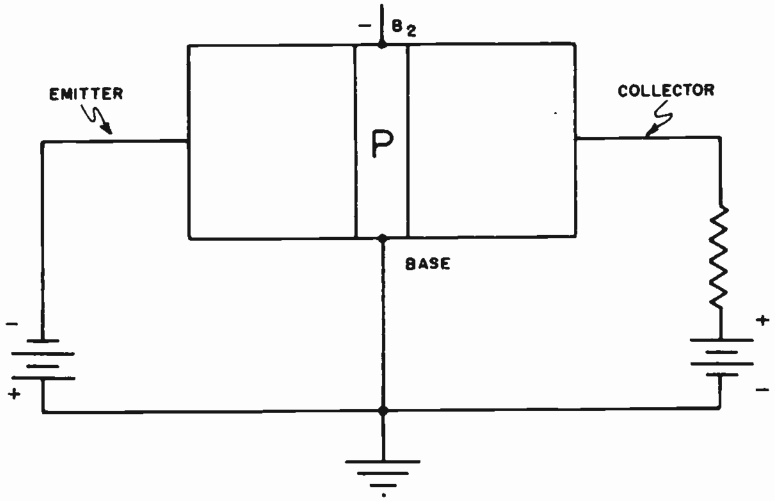


FIG. 10.—*Tetrode Transistor.*

8. TETRODE TRANSISTORS

At first glance one might fall into the error of believing that the four-crystal transistor, such as the one described in connection with Figure 9, is a tetrode transistor. That, however, would be wrong.

Nevertheless, there are such things as tetrode transistors. One of that kind is shown in Figure 10.

Just as tetrode vacuum tubes were designed to enable them to handle frequencies higher than those which could be readily handled by ordinary triode vacuum tubes, so were tetrode transistors developed so they could be used on higher frequencies.

One of the things which limits the frequency at which junction transistors can normally be used is the thickness of the base crystal. We mentioned earlier that the base crystal is made as thin as possible. One of the reasons for doing so is to enable the crystal to operate on higher frequencies.

However, the top operating frequencies of a junction transistor are still much lower than for vacuum tubes. Radically lower.

There is a direct relationship between the resistance of the base crystal and the frequency response of the transistor. The greater the resistance the lower the frequency response. The reverse is also true.

The obvious conclusion concerning this state of facts is that anything which can be done to reduce the resistance of the base, or reduce its thickness, acts to increase the frequency response of the transistor.

Since there are practical limits below which it is not practical to go in reducing the thickness of transistor crystals, the next best thing is to approach the problem from another angle. If anything

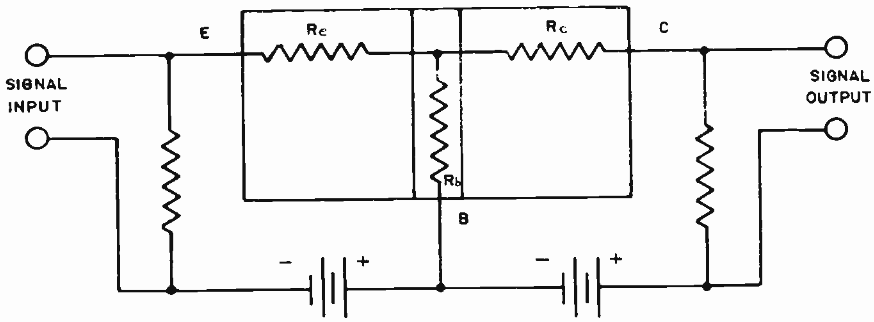


FIG. 11.—*Showing internal resistances in Transistors.*

can be done to reduce the resistance of the base crystal it has the same electrical effect of reducing the thickness.

One method being used to reduce the resistance of the base crystal in junction transistors is to make a second connection to the base crystal, and inject a new voltage on the crystal. The manner in which this is done is shown in Figure 10.

A new voltage is injected into the base crystal as shown in Figure 10. The voltage in Figure 10 is shown to be a negative voltage applied to the positive crystal base.

The new tetrode connection is made to the junction layer of the base in the same manner as the normal base connection, but it is connected on the opposite side of the crystal. When the negative voltage is applied to the P-type crystal, it acts to materially reduce the resistance of the base crystal.

Electrons injected onto the crystal by the tetrode connection serve another purpose. The electrons injected into the base crystal prevents electrons from the collector crystal passing through the junction near the location of the connection for the negative voltage. This forces a concentration of electrons near the connection to the regular base terminal.

The negative bias voltage applied to the tetrode base connection is commonly on the order of about 6 negative volts. When so applied, the voltage causes a current of about 1 milliampere to flow.

With respect to the resistance of the base crystal, this resistance is reduced from about 1000 ohms to about 40 ohms. The arrangement has no noticeable effect on the resistance of the emitter crystal or circuit.

The current gain is reduced. Where the current ratio in the collector circuit of an ordinary junction transistor is about .95 to that in the emitter circuit, we find the ratio dropping to about .75 in a tetrode transistor. This action has the effect of reducing the gain.

Further than that, there is also a reduction in the collector circuit resistance. It is commonly reduced about half. For example, a collector circuit which would have a

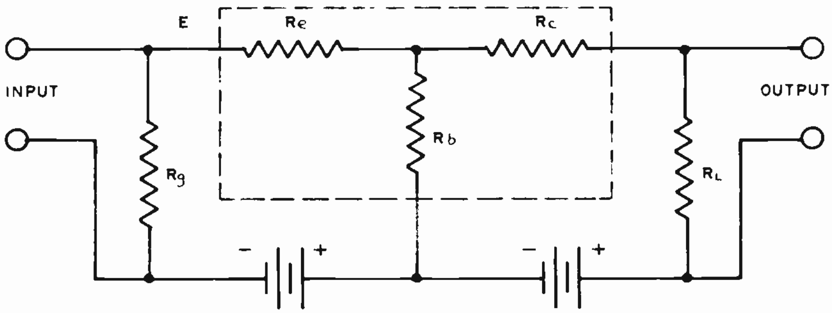


FIG. 12.—*Equivalent Circuit of Transistor.*

normal resistance of about 3,000,000 ohms without the second base connection would have a collector circuit resistance of only about 1,500,000 ohms with the second base connection.

This brings about an even greater loss in gain for the transistor when used as an amplifier.

However, sacrificing the gain of a transistor is a relatively small penalty to pay if the new arrangement brings about a higher frequency response. This the new circuit does. There is a very sharp increase in the frequency response of the transistor with the second base connection.

As a matter of fact the frequency response may be increased as much as ten times. This means that an ordinary transistor which has a top frequency response of some 0.5 megacycle has its frequency response upped to the vicinity of 5 megacycles by use of the second base connection on the transistor.

9. ELECTRONICS OF TRANSISTORS

To understand how transistors work in regular circuits it is first necessary to reduce the internal action in them to some form which is easier to understand. One way of doing that is to represent the resistances of the various parts of a transistor by the normal symbol for a resistor.

Such a method is shown in Figure 11.

However, the manner shown is sometimes a bit confusing because of the mixing of circuit symbols with the mechanical outline of the transistor. There is no real need to include the transistor outline if we intend to direct our attention only to the electrical action which is occurring.

The manner of arranging the circuits shown in Figure 12 is easier to understand.

The internal resistances of the transistor elements are shown inside the dashed box in Figure 12. One is designated as R_e. This is the resistance of the emitter circuit. Another is designated as R_c. This resistor symbol represents the resistance of the collector circuit.

Finally, we have the resistance designated as R_b. That, of course, is the resistance of the base circuit.

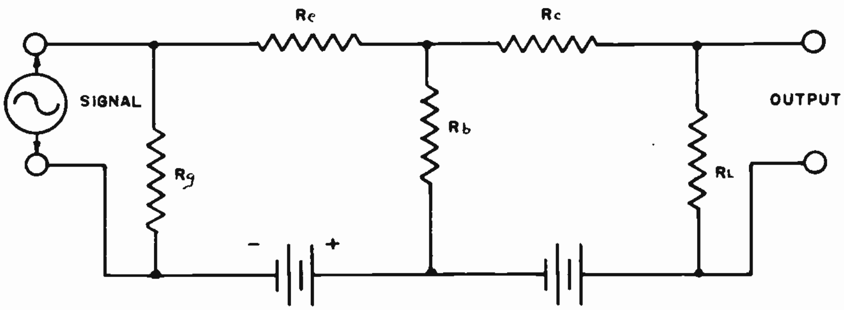


FIG. 13.—*AC signal in transistor circuit.*

Transistors, like vacuum tubes, require external circuit elements to make them work properly. In this case we have a load resistor in the output circuit designated as R_L, which is equivalent in all respects to the load resistances commonly used in the output circuits of vacuum tube amplifiers.

Note also, the resistor marked as R_g in the input circuit. This resistor corresponds in most respects to the grid-return resistor used in the input circuits of vacuum tubes used as amplifiers.

With these points of similarity established, let's take another look at the transistor's circuits.

We can look upon the battery in the emitter circuit as corresponding to the "C" battery formerly used in connection with vacuum tubes. It is, essentially, a biasing battery.

The other battery shown in Figure 12 can be looked upon as performing a duty similar to that performed by the "B" battery in vacuum tube work.

Once these things are established, we now come to the matter of placing a signal on the input to the transistor. It could be injected by using an AC generator of some kind, and connected to the emitter circuit, as shown in Figure 13.

The AC signal voltage placed across the emitter circuit, as shown in Figure 13, affects the movement of electrons and holes through the internal resistances R_e, R_c, and R_b in a manner similar to that already described. This means the steady DC voltage and current previously used on the elements of the transistor have been modified by adding an AC voltage and current from the signal source.

The AC signal causes the current in the emitter circuit to vary in accordance with the changes in the AC voltage. The changing emitter current, in turn, affects the collector current, bringing about variations in the collector's external circuit in a manner resembling the action of a vacuum tube.

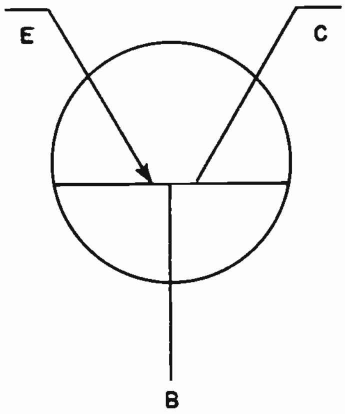


FIG. 14.—*Symbol for Transistor.*

The important point is that small changes in the voltage from the AC source bring about changes in the collector current so the AC voltage which is reproduced across the load resistor R_L in the output circuit is much greater than the original signal voltage. This, of course, is what we know as amplification.

10. TRANSISTOR SYMBOLS

One of the first things you learned when you began studying this course was the manner in which a vacuum tube was expressed symbolically.

It should occasion no suprise that a transistor is indicated in a similar manner; and so it is. The symbol for a transistor is somewhat different from that of a vacuum tube, but the differences are not material.

The drawing in Figure 14 shows what a typical transistor symbol looks like. Note the tiny arrowhead on the line for the emitter. That serves to identify the emitter, thus distinguishing it from the collector.

The letter *E* serves to further identify the emitter. The collector is identified by the letter *C*, if further identification of that element is needed. The base, of course, is the additional element. It is identified by the letter *B*. In most circuit diagrams the letters are omitted.

11. CIRCUIT CONNECTIONS

In the preceding pages we have drawn a comparison between the emitter in a transistor and the control grid in a vacuum tube. There are many points of similarity between them, and it is entirely practical to make such comparison.

In like manner, we drew a comparison between the collector circuit of a transistor and the anode circuit of a vacuum tube. Again such manner of comparison is entirely practical and logical.

Carrying this matter of comparison a step further, we have compared the base of a transistor with the cathode of a vacuum tube. Again, the comparison is sufficiently apt so it is logical.

At the same time we want to warn you against drawing inaccurate and erroneous conclusions from our attempt to make these comparisons. The emitter of a transistor is definitely not *identical* with the control grid in a vacuum tube, and if you fall into the error of so considering it you will be letting yourself in for unnecessary trouble.

The same is true of the other two elements in a transistor. The base is comparable in a number of ways with the cathode of a vacuum tube, but it certainly is not identical. The same is true of the collector. It is similar to the anode in a vacuum tube, and performs comparable functions; but by no stretch of the imagination should it be considered identical.

The drawing in Figure 15 serves to show what is often thought of as being a typical transistor connection. It shows the emitter serving as the input circuit, corresponding in many ways with the control grid in a vacuum tube. It also shows the base being grounded, in a manner similar to that often used with vacuum tube cathodes. The collector is serving as the output from the transistor circuit.

The circuit shown in Figure 15 shows a transistor being used as an amplifier so that it is performing its function in a manner closely similar to that followed by a vacuum tube. In fact, from all appearances, the transistor could almost have been substituted for a vacuum tube triode in the circuit, except for the difference in the values of voltage used on the two types of amplifiers.

The circuit shown in Figure 15 is what is known in transistor work as a *grounded-base circuit*.

The circuit is all right, but it has certain technical drawbacks. Among them is the fact that the transistor requires two separate sources of voltage, one for the emitter circuit, the other for the collector circuit.

Another drawback is that a transistor connected in this manner is not capable of developing the gain it can achieve if connected in one of the other ways.

Therefore, despite the fact we have used the grounded-grid method of connecting the circuits to a transistor so as to provide a method of comparison, the fact remains that transistors are not always connected in this manner in practice. An even more common method is known as the grounded-emitter method. There is another known as the grounded-collector method. Both of these latter methods have certain features of superiority.

12. GROUNDED-EMITTER CONNECTIONS

For testing and experimental purposes it is the general practice to connect transistors with the base grounded. This is the typical grounded-base connections as shown in Figure 15. This method of connecting a transistor into an

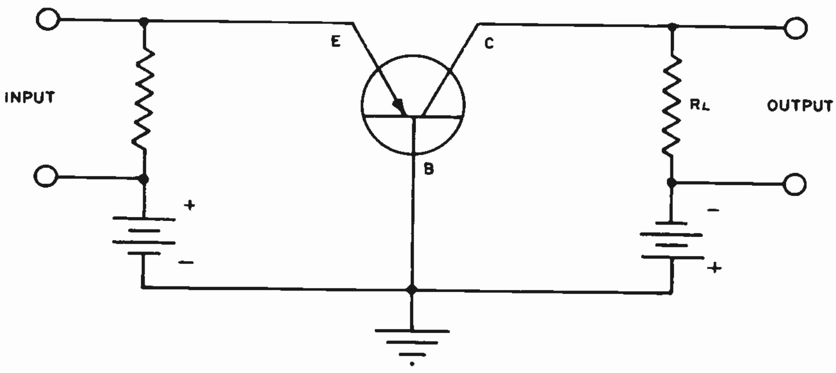


FIG. 15.—*Grounded-Base transistor Circuit.*

amplifier circuit is occasionally referred to as the "conventional" method, although this latter designation is falling into disuse.

The two other systems used in connection with transistor amplifier circuits are known as the "grounded-emitter" method and the "grounded-collector" method. Both are widely used; each possessing special kinds of advantages.

The circuit diagram in Figure 16 shows, in simplified form, the wiring arrangement of a grounded-emitter transistor. Note, that when a transistor is connected in this manner only one battery, or voltage source, is needed.

Contrast the circuit in Figure 16 with that in Figure 15, where two voltage sources are needed.

Reduction in the need for additional voltage source is one big advantage of a grounded-emitter transistor circuit. But that is not the only one. A transistor connected with a grounded-emitter has a much higher gain than one with a grounded base.

Let us pause a moment to compare this matter of differences in the gain for the two methods of grounding.

As a rough basis for comparison, and that is all we are attempting to show at this time, let us say that a typical point-contact transistor has been connected with a grounded base, and the circuit has a voltage gain of 30. If the same point-contact transistor were connected with grounded-emitter, it would have a gain on the order of 130, or approximately 5 times greater than for grounded-base, when used in audio circuits.

Now let us take a look at a junction transistor. For one that has a normal gain of approximately 150 when connected as a grounded-base audio amplifier, the same transistor would have a gain of about 575 when connected as a grounded-emitter transistor. Thus, when connected as a grounded-emitter amplifier the transistor has almost 4 times as much gain as when connected as a grounded-base amplifier.

There is another difference between the two types of operation. When a transistor is connected as a grounded-base amplifier the output signal is in phase with the input signal.

On the other hand, when a transistor is connected as a grounded-emitter the output signal is 180° out of phase with the input signal. This is a situation similar to that which we have previously discussed with respect to vacuum tube amplifiers.

With respect to power gain, a transistor connected as a grounded-emitter amplifier has about four or five times as much gain as when connected as a grounded-base amplifier.

Thus, we have two important advantages when a transistor is connected as a grounded emitter amplifier rather than a grounded-base amplifier. For one thing, only one power source is required. For another, greater gain is obtained.

A disadvantage—provided it can be called that—is the fact the output signal is 180° out of phase with the input signal. Since this is the same situation with which we have been working in connection with vacuum tubes for many years, it is doubtful if this can

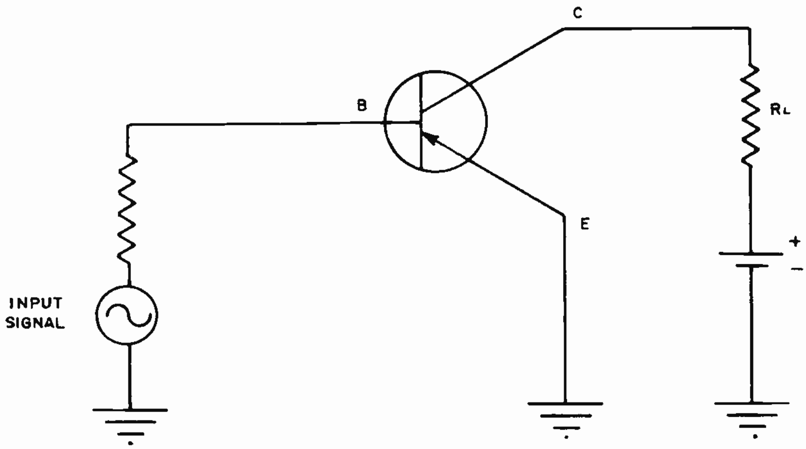


FIG. 16.—*Ground-emitter transistor circuit.*

actually be called a disadvantage. Another disadvantage is that gain falls off at higher frequencies.

Perhaps we should point out that a grounded-emitter transistor is sometimes a little unstable when the transistor is a point-contact type. However, in the case of junction transistors, the circuit is always stable.

Perhaps you will be interested in learning why two sets of power sources are needed with a grounded-base transistor, while only a single source is required for a grounded-emitter transistor. Beginning with Figure 17, we can make this situation a little more clear.

Note the manner in which the currents are flowing. In the emitter circuit the current is flowing from the emitter electrode through its external circuit to its battery, E_1. From there it flows to the ground, and back to the base of the transistor.

Current flows in a similar manner in the collector circuit. It flows from the battery, E_2, to the collector electrode, through the transistor, out the base, and finally through the ground to the battery again.

Since current in the collector circuit is flowing through the base circuit in a direction opposite that of the emitter current which is also flowing in the base circuit, it follows that the only actual current which flows in the base circuit is the difference between that in the emitter circuit and that in the collector circuit.

Now look at the circuit in Figure 18, and note how the current flows. Current flows from the battery, through the load resistor, R_L, to the collector electrode. It then flows through the transistor and out the emitter electrode, from whence it goes to ground.

Current can still flow in either direction in the base circuit, except that it is also subject to the

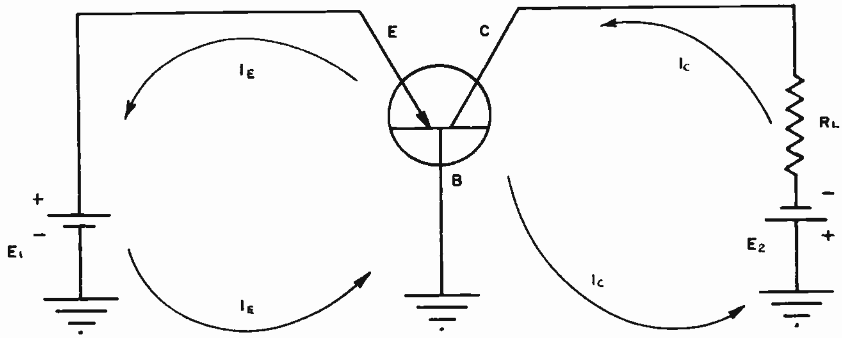


FIG. 17.—*Current flow in Emitter-collector circuits.*

control of any externally applied signal voltage. The signal voltage, usually AC in character, is represented by the small sine-wave generator in the base circuit of Figure 18.

Variations in the AC signal voltage, as applied to the base crystal of the transistor by the base connection, acts to affect the current in both the collector and the emitter circuits.

The transistor in Figure 18 is a P-N-P junction type. However, an N-P-N junction transistor could be used in a similar manner, merely by reversing the connection to the battery voltage supply.

The connections and direction of current flow, in a transistor circuit involving a grounded-emitter N-P-N type, are shown in Figure 19.

13. FREQUENCY RESPONSE OF GROUNDED-EMITTER TRANSISTOR

We have already pointed out that a grounded-emitter transistor has much greater gain than a grounded-base transistor. This makes the grounded-emitter much more useful for a number of purposes. This is particularly true when the transistor is used in an audio circuit.

However, the gain which can be attained by a grounded-emitter transistor is strongly affected by the frequency. The gain can be readily attained at frequencies below 0.5 megacycles, which makes the transistor operate satisfactorily at audio and intermediate frequencies. However, at frequencies above .5 megacycles the gain falls off quite rapidly as the frequency is increased.

In the case of junction transistors the gain falls off almost to zero at 1.5 megacycles, and is quite low at 1 megacycle. There is a falling off in frequency response in the case of point-contact transistors, but not quite so drastically as with junction transistors.

What this means is that the advantage of connecting a transistor as a grounded-emitter amplifier depends on the frequency of the signal to be amplified. If the frequency is relatively low, such as audio signals or low-frequency radio signals, the gain of a grounded-emitter transistor is high as compared with a grounded-base transistor. On the other hand, if the signal frequency to be handled is higher than .5 megacycles, wisdom of using a grounded-emitter circuit is highly debatable.

Much of the loss in gain at the higher frequencies stems from stray capacity in the transistor circuits. Stray capacities in the transistor circuits act to shunt the higher frequency signals around the transistor, thus they are not affected by the normal transistor action. If the signal is not affected by the transistor action, no gain results.

14. GROUNDED-COLLECTOR TRANSISTOR

The third principal method for connecting transistors is by using what is called the grounded-collector method. This is shown in diagrammatic form in Figure 20.

When connected in this manner the input resistance into the transistor circuit is quite high, while the output resistance is quite low. This is just the reverse condition needed for good gain in an amplifier.

A grounded-collector transistor can be roughly compared with a cathode-follower circuit used with vacuum tubes. It serves to couple a high-impedance input circuit into a low-impedance output circuit.

The output signal is taken from the emitter circuit in a grounded-collector transistor circuit. This means the signal swing must be kept within relatively narrow limits.

In practice, a grounded-collector circuit is rarely used as an amplifier. Its principal use is as a coupler where it is necessary to match impedances for some reason. This manner of connecting is definitely not recommended as either a power amplifier nor as a voltage amplifier.

15. MECHANICAL STANDARDS FOR TRANSISTOR CONSTRUCTION

When transistors first came into existence little effort was made to standardize their physical construction so one could be replaced by another. As they have become

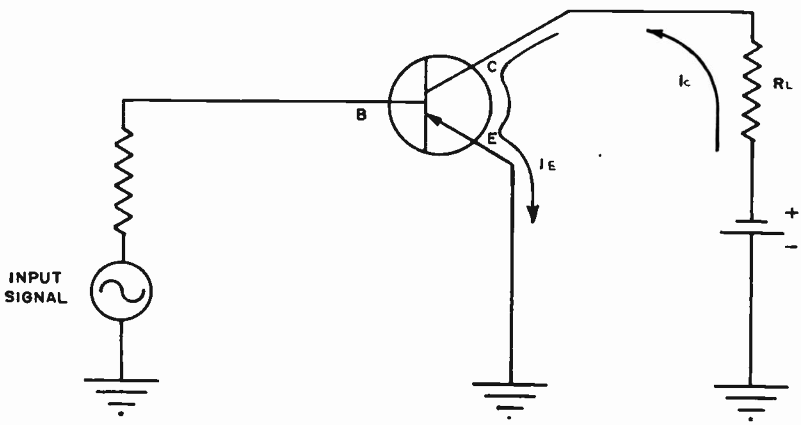


FIG. 18.—*Current flow in Emitter-collector circuit of Grounded-emitter transistor.*

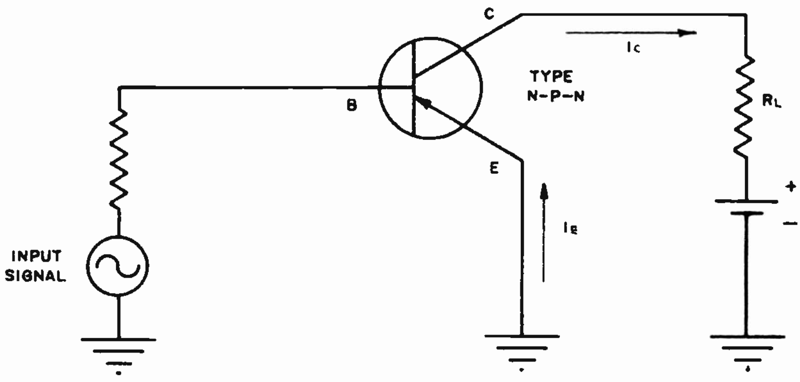


FIG. 19.—*Grounded-emitter N-P-N transistor.*

more common, it has been recognized that some system of standards should be set up.

Manufacturers of transistors have developed certain standards which most of them are trying to follow. There are exceptions in the case of special types, but most follow these standards fairly closely.

The basic standards are shown in Figure 21. This drawing shows the dimensions of the various parts of a transistor.

Since a transistor is much smaller than the drawing, it follows that all the dimensions must of necessity, be smaller than is indicated by the drawing. Nevertheless, we have tried to indicate the actual measurements of most of the dimensions.

The three principal elements of a transistor have connecting leads brought out in the form of wire pigtails. These are shown under the body of the transistor in Figure 21. The collector connection is marked with the letter *C*. The

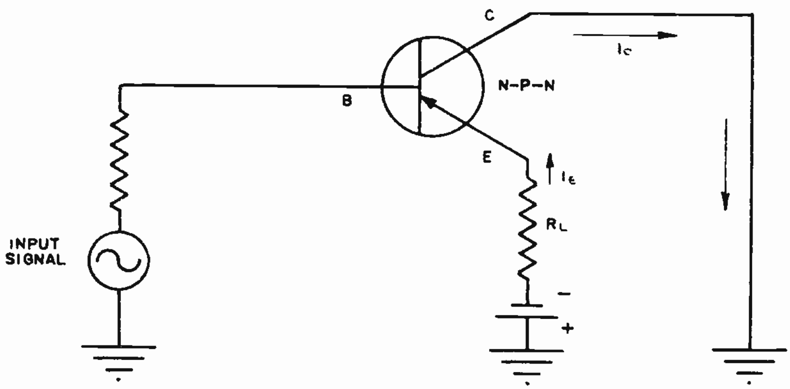


FIG. 20.—*Grounded-collector Transistor.*

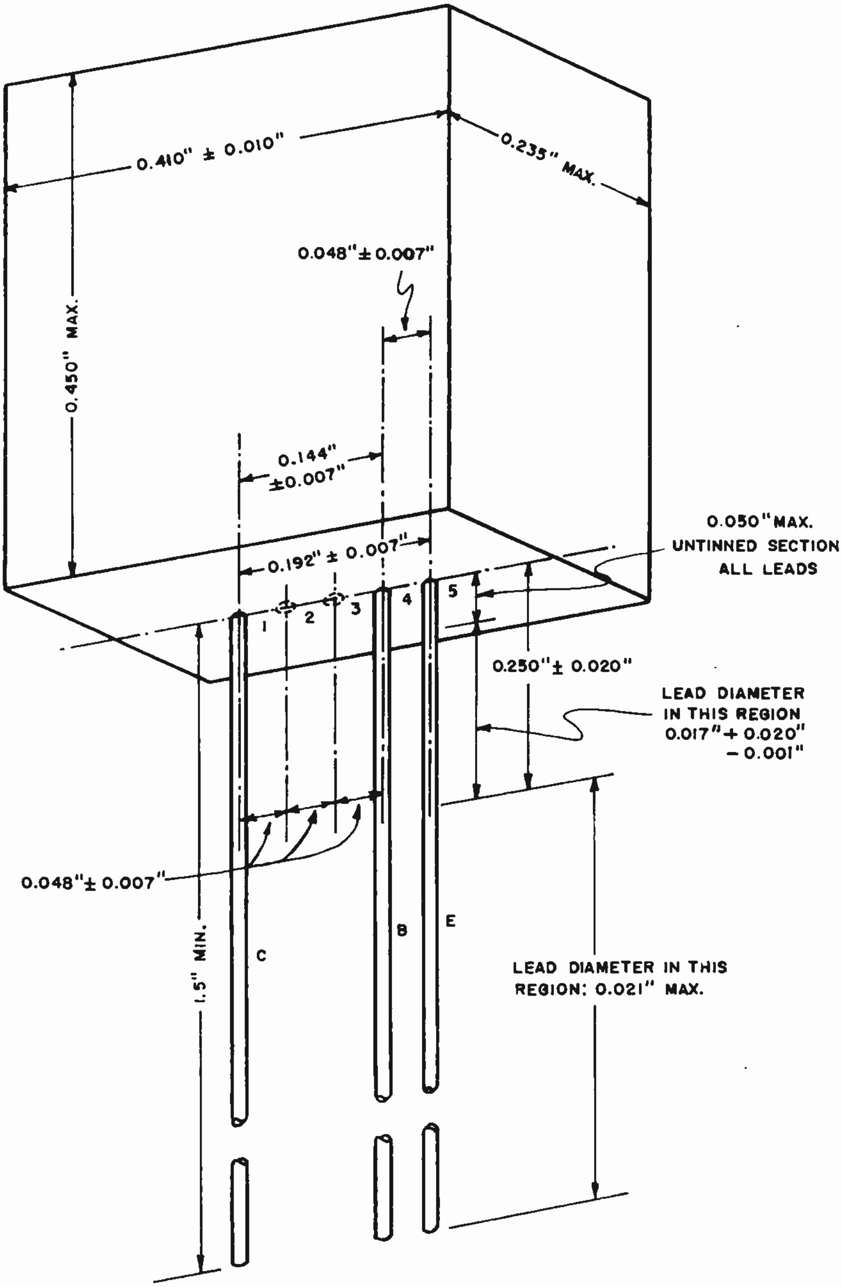


FIG. 21.—*Proposed standards for transistor construction.*

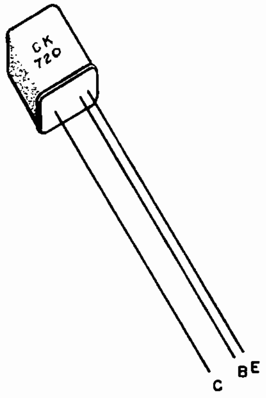


FIG. 22.—*Typical Transistor with connecting leads.*

base is marked with a *B*, while the emitter is marked with an *E*.

It is not required that the body of the transistor follow the shape of the one shown in the drawing. In fact, most transistors do not follow this shape.

However, manufacturers try to hold their finished product within the limits shown on the drawing in Figure 21. Frequently, the finished product is considerably smaller than the size shown.

The actual appearance of a transistor more nearly resembles the drawing in Figure 22. The letters at the ends of the connecting leads indicate the internal elements to which they are connected. The collector and emitter electrode connections are the ones the farthest apart. The base connection is closest to the emitter connection.

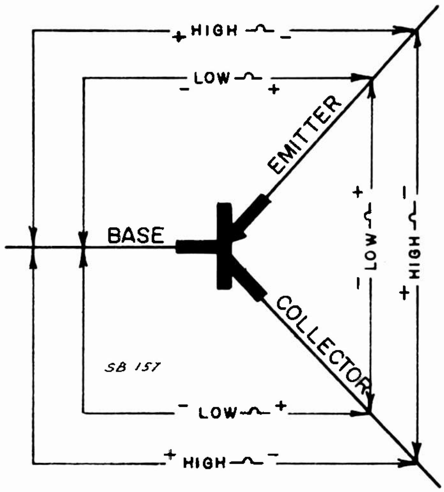
16. TRANSISTORS IN RADIO RECEIVERS

There is not sufficient space, nor time within the limits of this lesson, to go into details connected with the use of transistors in actual radio circuits. Nevertheless, we can introduce you to the manner in which they are so used, then return in later lessons to go into the subject more deeply.

The Motorola Company has been a leader in the use of transistors in radio equipment. They have set up some specific rules concerning the manner in which their service technicians should handle the servicing of transistor radio equipment.

First, they set standards by means of which service technicians can test a transistor for defects.

One of the tests they suggest concerns the matter of checking the resistance between elements in a transistor. They recommend use of an ohmmeter for making the



PNP TYPES 2N186, 2N187, 2N188, 2N189, 2N190, 2N191, 2N192, 2N241

FIG. 23.—*Testing forward and reverse resistances of P-N-P Transistor.*

tests, but they warn that the ohmmeter must have an internal battery voltage which does not exceed 7.5 volts. It is desirable to keep the ohmmeter battery voltage below that level to avoid damaging the transistors, some of which cannot withstand voltages greater than that value.

The ohmmeter is used to measure the forward and the reverse resistance between each pair of the transistor terminals. The drawing in Figure 23 shows how the forward and reverse resistances of a P-N-P type transistor are tested. No attempt is made to show the exact values of resistance in either forward or reverse directions, since these are frequently subject to some degree of change.

The manner in which the forward and reverse resistances of an N-P-N transistor are checked is shown in Figure 24.

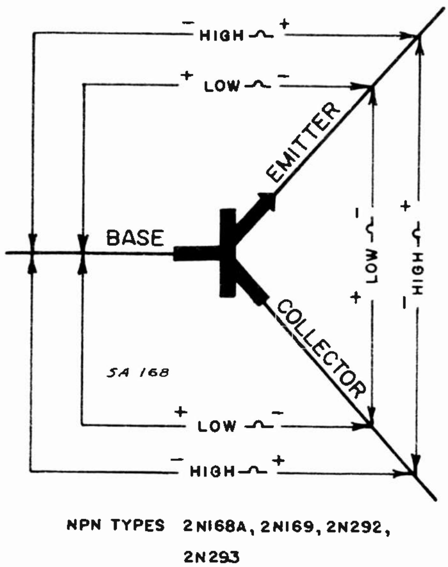


FIG. 24.—*Testing resistances of N-P-N Transistor.*

If a transistor is to be replaced, Motorola warns its service technicians that the new transistor must be protected against overheating during the soldering process. A pair of long nose pliers should be used to grasp the pigtail being soldered. This helps conduct the heat away from the transistor during the soldering operation.

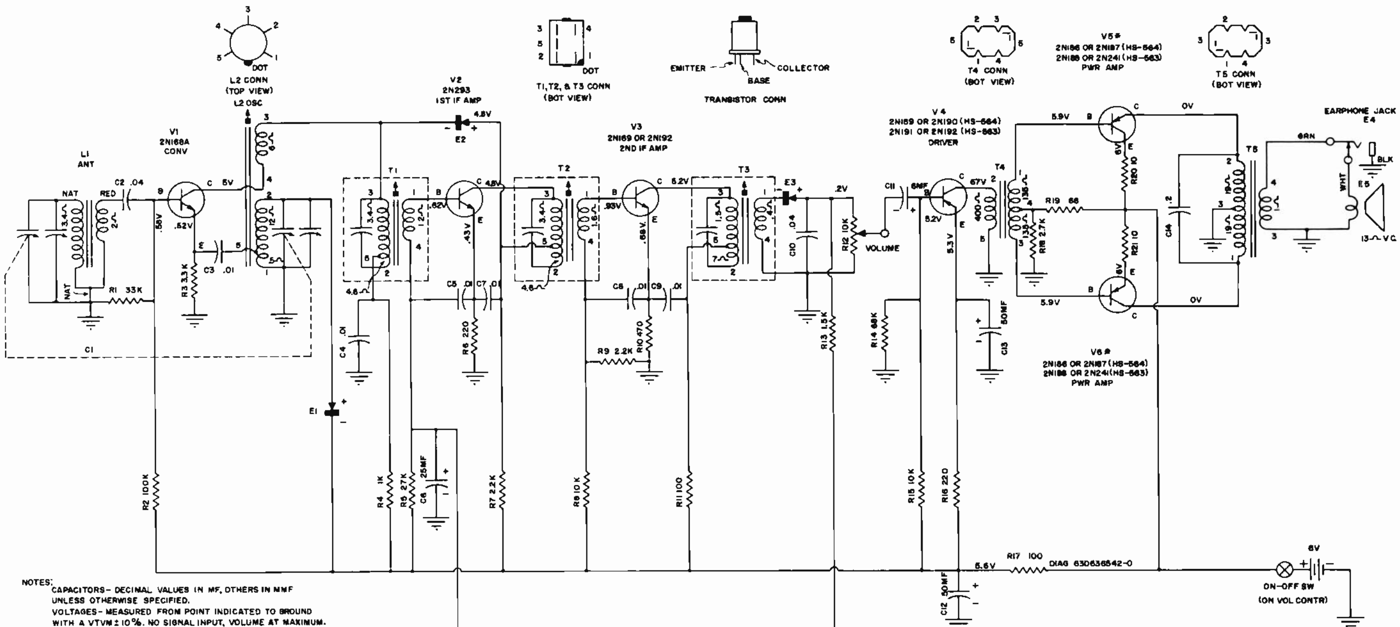
No more heat should be applied to the pigtail than is absolutely necessary, and the heat should not be applied any longer than is essential. Applying too much heat, or for too long a period, will ruin the new transistor.

The secret is to apply heat quickly, and then cool the soldered connection quickly by blowing, or applying a little moisture.

We have provided a complete circuit diagram of a transistor radio in Figure 25. This has not been done with the thought that you should already understand everything there is to know about such a radio. On the contrary, we are including this circuit diagram in this lesson for the purpose of providing you with a circuit you can study, and thus begin to familiarize yourself with many of the circuit elements.

You will note many points of similarity between the transistor radio and those which use tubes. You will also find a number of differences.

One of the things which will probably catch your attention most quickly is the absence of filament circuits, and filament power supply. Since transistors have no filaments, this is one electrical circuit which can be eliminated from a transistor radio.



NOTES:
CAPACITORS- DECIMAL VALUES IN MF, OTHERS IN MMF UNLESS OTHERWISE SPECIFIED.
VOLTAGES- MEASURED FROM POINT INDICATED TO GROUND WITH A VTVM ± 10%. NO SIGNAL INPUT, VOLUME AT MAXIMUM.
TUNING RANGE- 530 KC TO 1620KC.
IF- 455 KC
DO NOT INTERMIX POWER AMPLIFIER TYPES, USE TWO OF THE SAME TYPE TRANSISTORS IN THE OUTPUT STAGE.
RESISTANCES- MEASURED WITH THE TRANSISTORS OUT OF ASSOCIATED CIRCUITS.

FIG. 25.—*Transistor Radio Circuit.*

Another thing which will probably catch your attention is the extremely low voltages needed by a radio of this character. A simple 6-volt battery, which could be a small dry cell battery, is all that is required to operate the radio. This small battery will provide sufficient power to keep the radio operating for months.

Some of the transistors used in the radio are P-N-P types; others are N-P-N transistors. This will serve to familiarize you with the manner in which manufacturers provide this information, since this diagram is a copy of one prepared by a manufacturer.

ITEMS OF SPECIAL INTEREST

Point-contact transistors were the first types developed, but junction transistors are taking over many of the functions formerly handled by them exclusively.

Point-contact transistors had many limitations, yet they are still used in limited applications because of certain areas of superiority for certain purposes.

Voltage used with transistors must be kept low to prevent ruining the transistors.

Transistors are susceptible to over-current and over-voltage.

Some of the limitations of point-contact transistors were overcome by using the wider contact surfaces of junction transistors.

Junction transistors consist of sandwiching a crystal having one polarity between two other crystals which have opposite polarity, then making appropriate connections to each of the crystals.

A potential hill appears at each junction between crystals with differing polarities. Polarity of the potential on the potential hills is opposite that on the joining crystals.

The thinner the base, or center, crystal of a junction transistor the better its operating characteristics.

The base crystal is made as thin as manufacturing facilities permit. In some cases the base crystal is applied by diffusing a thin layer of molecules having the proper polarity so they separate the two end crystals with opposite polarity.

A junction transistor is, basically, a double crystal diode.

There is much broader area of contact between the surfaces of adjoining crystals in a junction transistor than is true at the contacts of a point-contact transistor.

A potential hill exists at the junction between crystals with opposite polarities.

Varying the emitter circuit voltage causes a variation in the current which enters the emitter crystal.

Some junction transistors consist of a pair of P-type crystals with a thin N-type crystal between them. Another kind consists of a pair of N-type crystals with a thin P-type crystal between.

The major current carrier in an N-P-N transistor is considered to be electrons.

In the case of a P-N-P junction transistor, the major current carrier is considered to be holes.

The emitter circuit is biased in the *forward direction.* The collector circuit is biased in the *reverse direction.*

Voltage gain in a transistor amplifier depends more heavily on the ratio of resistances in the emitter and collector circuits than on the ratio of currents.

Theoretical gain in a transistor circuit can be calculated by multiplying the square of the ratio of currents in the emitter and collector circuits by the ratio of resistances in the two circuits.

Gain of a transistor can be increased by using four crystals instead of three, but this affects the frequency response.

Frequency response in an ordinary three-crystal transistor can be improved by applying added voltage to the base crystal. The so-called *tetrode* voltage should have a polarity opposite that of the base crystal.

The bias voltage applied to the base crystal of a tetrode transistor is on the order of about 6 volts.

Adding the bias voltage to a tetrode transistor reduces the gain, but it greatly increases its frequency response.

A transistor can be connected into an amplifier circuit so the base is connected to the ground. When so connected it is commonly referred to as a grounded-base transistor.

A grounded-base transistor is similar in many respects to a vacuum tube amplifier. The base corresponds roughly to the cathode of a vacuum tube, the emitter to the control grid of a vacuum tube, and the collector to the anode circuit of a vacuum tube.

A grounded-base transistor amplifier does not have the gain that some of the other methods of connecting, but it has a good frequency response.

A grounded-base transistor is used where good frequency response is more important than high gain.

Another method of connecting transistor amplifiers is what is known as grounded-emitter.

A grounded-emitter transistor has several times the gain of a grounded-base transistor.

Where high gain is a matter of importance it is customary to connect a transistor amplifier as a grounded-emitter.

A grounded-emitter transistor does not have so good frequency response as one with grounded-base.

In practice, transistors intended to amplify audio frequencies and low-frequency radio signals are connected as grounded-emitter amplifiers, while high-frequency amplifiers are connected grounded-base.

A third method of connecting transistors into amplifier circuits is known as grounded-collector.

Grounded-collector transistor amplifiers are comparable to cathode-coupler vacuum tube circuits.

The input impedance of a grounded-collector transistor is quite high, but the output resistance is quite low.

Grounded-collector transistors are used to couple circuits where the impedances differ, but they are not considered good amplifiers.

The two most comon methods for connecting transistors as amplifiers are grounded-grid and grounded-emitter.

Where high gain is important grounded-emitter is used; where high-frequency response is important, grounded-base is used.

At low frequencies a grounded-emitter transistor has about four to six times the gain of a grounded-base transistor.

At frequencies above .5 megacycles, grounded-emitter transistors lose their advantages of gain.

At frequencies much above .5 megacycles, grounded-emitter transistors have little, if any, gain advantage over grounded-base transistors.

Lack of gain at high frequencies in grounded-emitter transistors is due to shunt capacity in the transistor circuit by-passing the signal voltages around the transistor circuits.

Manufacturers have set mechanical standards for transistor construction.

The center lead from a transistor body is normally the base connection.

The outside lead closest to the base lead on a transistor is the emitter connection.

TRAINING SERVICE

Print or Use Stamp:

Date__________
Name__________ Student No.__________
Address__________ Zone_____ Grade__________
City__________ State__________

Lesson 44

TRANSISTORS.
Part 1.

Underline or fill in the ONE correct answer, just as in previous lessons.

1. A circuit used with electronic devices employing tubes, but omitted when transistors are used, is the:
   a. detector circuit. b. power supply circuits.
   c. filament circuit. d. ground circuit.

2. During the process of replacing a transistor it is important:
   a. to protect the new transistor against overheating.

9. At audio frequencies the highest gain transistor circuit is one connected:
   a. as a grounded-emitter amplifier.
   b. in cascade.
   c. directly to the voice coil.
   d. as a grounded-base amplifier.

10. The major current carriers in an N-P-N transistor are:
   a. holes.
   b. ions.
   c. electrons.
   d. protons.
   e. arsenic atoms.

11. The center crystal in a junction transistor:
   a. is thicker than the other two crystals.
   b. is kept quite thin.
   c. breaks easily.
   d. can withstand much heat.

12. Connection to a transistor base is made through the:
   a. isolated outside pigtail.
   b. center pigtail.
   c. emitter circuit.
   d. collector circuit.

13. Electronic devices using transistors are distinguished by:
   a. the extremely low voltages needed from the power supply.
   b. their ability to operate at extremely high frequencies.
   c. their elaborate and intricate circuitry.
   d. their ability to operate at extremely high temperatures.

14. Voltages used on transistors:
   a. are approximately the same as used with vacuum tubes.
   b. should be slightly higher than commonly used with vacuum tubes.
   c. must be kept low to prevent ruining the transistors.
   d. must be AC to prevent developing reverse bias.

15. a. The emitter circuit is biased in the ____________ direction.

b. The collector circuit is biased in the ____________ direction.

Part 2.

16. The reactance (at 60 cycles) of the primary winding in a typical power transformer for an electronic device is commonly on the order of:
a. 10 ohms. b. 100 ohms. c. 750 ohms. d. 1500 ohms. e. 2500 ohms.

17. The relationship of electrical power, voltage and current to each other is commonly known as:
a. Ohm's Law. b. Boyle's Law. c. Watt's Law. d. Joule's Law.

18. There are how many watts in a horsepower? ____________

19. Secondary emission is avoided in vacuum amplifier tubes by using what kind of tube element? ____________________________

20. If you have a triode with a mu of 40 and a plate resistance of 20,000 ohms, then use a 30,000-ohm load resistance, how much gain will it have as a resistance-coupled amplifier?
a. 10. b. 15. c. 24. d. 30. e. 33. f. 37. g. ______.

b. to prevent the transistor freezing.
c. for the base connection to be soldered first.
d. to ground the emitter.

3. Principal advantage of a grounded-collector transistor circuit is:
   a. to couple different impedances.
   b. greater amplification.
   c. to operate at higher frequencies.
   d. greater power.

4. Principal advantage of tetrode transistors is their ability to:
   a. develop greater power.
   b. operate at higher frequencies.
   c. couple different impedances.
   d. operate at higher temperatures.

5. One big advantage of junction transistors over point-contact transistors is:
   a. its ability to handle larger currents.
   b. higher voltage ratings.
   c. greater delicacy.
   d. longer experience in use.

6. In a P-N-P transistor:
   a. electrons are the major current carrier.
   b. an indium crystal is in center.
   c. holes are the major current carriers.
   d. there are no holes.

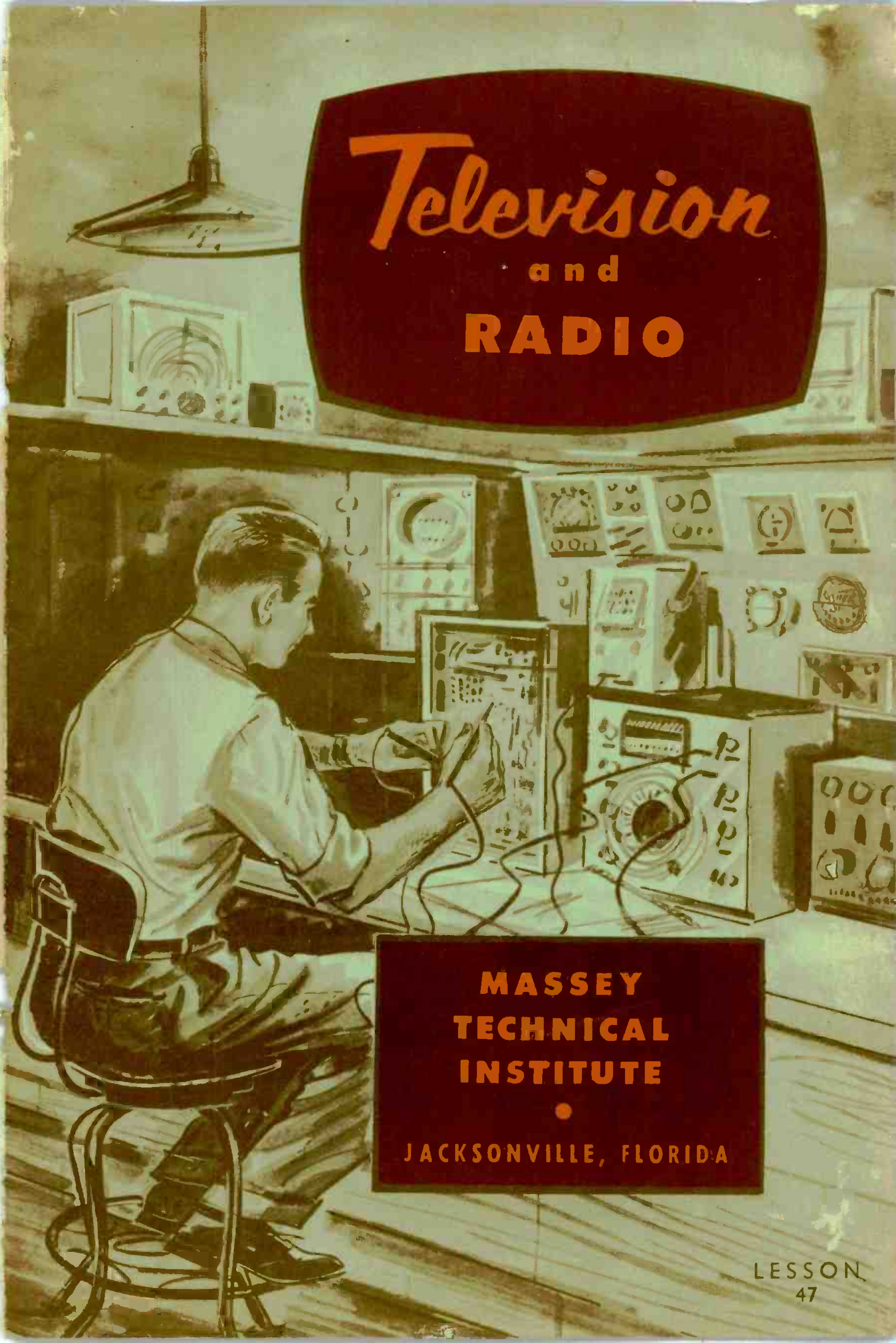
7. The emitter in a transistor is:
   a. similar to the anode in a tube.
   b. always grounded.
   c. same as cathode in a tube.
   d. closely analogous to a tube grid.

8. The DC voltage normally applied to a transistor emitter can be looked upon as:
   a. the B-plus voltage.
   b. the bias voltage.
   c. part of the signal.
   d. a reverse bias.

MTI

MASSEY TECHNICAL INSTITUTE, Inc.

Jacksonville, Florida



Television and RADIO

MASSEY TECHNICAL INSTITUTE

JACKSONVILLE, FLORIDA

LESSON 47

NOTHING DETERS
THE DETERMINED.

TRAINING SERVICE

FREQUENCY MODULATION

CONTENTS: Introduction—Modulation—Frequency Modulation—How an FM Signal Compares With an AM Signal—How Amplitude Modulation Occurs—Purposes for Which Each Type of Modulation is Used—Advantages in FM Radio—Disadvantages of FM Radio—Carrier Modulation—Modulation in FM Radio—Frequency Deviation—Limitations on Frequency Deviation—Degree of Deviation—Loudness and Pitch of the Audio Signal—Deviation Ratio—Range of FM Signals—FM in Television—Items of Special Interest.

1. INTRODUCTION

There is scarcely a person who has not heard the terms "FM Radio" and "FM Stations" and similar expressions. Most of our larger cities have FM broadcast radio stations, and where we find such stations providing programs there are FM receivers.

Despite the fact that so many persons have heard of FM radio and FM receivers, the truth is that few know what the term means.

The name comes from the manner in which the R-F carrier signal transmitted by the station is modulated. The audio component is carried from the transmitter to the receiver by the R-F carrier through the process of modulating the carrier's *frequency* rather than modulating its *amplitude*.

Hence the name, which is actually *Frequency Modulation*.

So, the distinction between the two principal kinds of radio transmission consists of the difference in which the carrier is modulated. To make certain that you fully appreciate this difference we think it desirable to review briefly the matter of modulation, then go into the details which causes FM radio to differ from the more conventional "amplitude modulated" radio.

2. MODULATION

In a number of previous lessons we have explained that the process of transmitting intelligence from one location to another by means of radio signals it is first necessary to create a "carrier" signal which possesses the capability of radiating through space. This is the high-frequency R-F carrier we have mentioned so many times. It possesses a frequency sufficiently high as to permit it to radiate.

In most cases the frequency of the R-F carrier is too high to carry the desired intelligence directly. In such case it has become the general practice to modify or change—or to *modulate*—the R-F carrier in some manner. It is modulated by the electrical signal representing the intelligence it is desired to transmit.

The most common method of modulating the R-F carrier is to cause the intelligence signal to bring about a variation in the amplitude of the carrier signal. This is accomplished in a manner similar to that diagrammed in Figure 1.

In the upper diagram note that we have a repetitive sine-wave signal. This represents the high-frequency R-F carrier, whose signal, as it comes from the oscillator, has an even amplitude for all successive cycles. Each half-cycle rises just as high above the zero reference line, and each half-cycle drops just as far below it. All successive cycles have the same amplitude.

The second graph represents the modulating signal. We commonly think of the modulating signal as being an audio signal, and in most cases it is. However, the modulating signal may carry some other forms of intelligence, and in the case of television we find both the video information and the audio information being used to modulate carrier signals. The video signal has a frequency much higher than the audio.

Finally, the bottom graph in Figure 1 shows the carrier signal after it has been amplitude modulated by the audio signal. Note that the effect of the modulating signal has been such as to vary the *amplitude* of the individual R-F cycles without changing the *frequency* of the carrier.

When the *amplitude* of the R-F carrier signal is caused to vary in accordance with variations in the

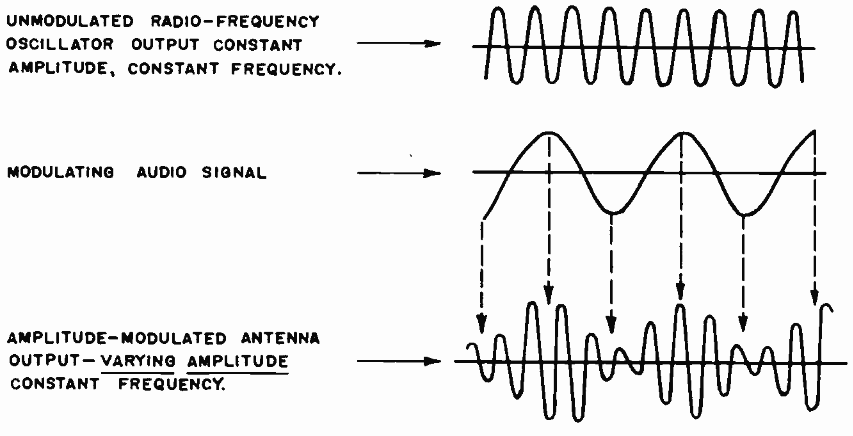


FIG. 1.—*How Carrier's Amplitude is Modulated.*

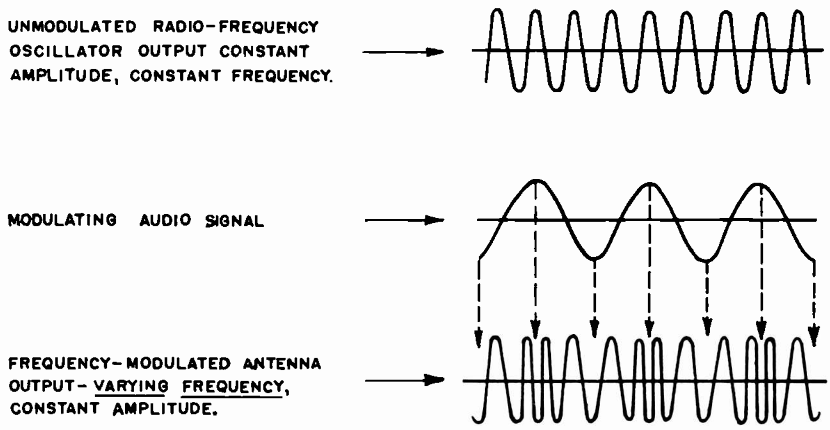


FIG. 2.—*Varying Carrier's Frequency through Modulation.*

strength or frequency of the modulating signal we call the act *amplitude modulation.* For many years this was the only form of modulation employed in radio work. It continues to be the form of modulation used by standard broadcast stations. This is the kind of signal picked up by typical radio receivers, or what are so frequently referred to as "standard radio receivers," or "AM receivers."

However, it is possible to cause the R-F carrier signal to vary in other ways, or to change its form, in response to the effect of the modulating signal. The carrier signal can be varied with respect to its frequency, with respect to the difference in phasing between the carrier's voltage and current nodes, by assuming a sequence of pulses, and in other manners.

For the moment we will confine our attention to the matter of modulating the carrier by causing its frequency to vary.

This is what we call *frequency modulation.*

3. FREQUENCY MODULATION

In order for television programs to be transmitted from a broadcast station to the millions of receivers tuned to it, it is necessary to provide means to transmit both the video signal and the audio signal from transmitter to the receivers. To prevent interaction between the visual intelligence and the aural, or sound, intelligence, standards were set up to use different methods of modulation for handling the two types of intelligence.

The video portion of the television broadcast is carried by the R-F carrier in the form of amplitude modulation. Thus, the video portion of the television programs is handled between transmitter and receiver in the same manner as sound is carried in ordinary radios.

The sound portion of the broadcast, however, is carried by the R-F carrier in the form of *frequency modulation.* Actually, in television broadcasts, one R-F carrier is modulated by varying the

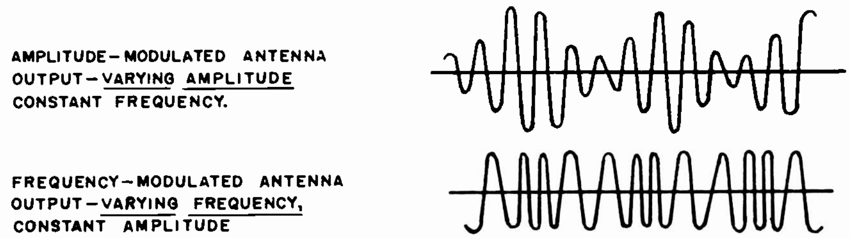


FIG. 3.—*Comparison of Frequency Modulated Signal with an Amplitude Modulated Signal.*

amplitude and acts to handle the video intelligence; a second R-F carrier is modulated by varying the frequency in accordance with the sound intelligence.

Thus, for each TV channel we find two separate R-F carriers, each modulated in a different manner, and each carrying a different form of intelligence.

In the case of a frequency modulated signal we find the signal being generated in a manner similar to that for any other kind of broadcast. The original signal is some form of unvarying sine-wave voltage and current going through their cycles at very high frequencies.

However, the intelligence to be carried by the carrier signal is used to bring about a variation in the frequency of the R-F carrier. Sometimes the modulating signal voltage is applied directly to the oscillator creating the R-F carrier, and in other cases it is applied in a different manner. At the present moment we need not concern ourselves greatly as to the precise form in which the modulation is brought about.

In any event, the R-F carrier can be thought of as being created so it assumes the shape approximating that shown in the top graph of Figure 1, and for practical purposes we can look upon them as being identical.

Nor is there any significant difference in the shape of the modulating signal shown in the center graph in Figure 2. It can be thought of as being the same as the modulating signal in Figure 1.

However, when you look at the bottom graph in Figure 2, the shape is shown to be considerably different.

The amplitude of the signal represented by the bottom graph in Figure 2 does not vary. It has the same amplitude for all successive cycles.

However, when you study the frequency you note that it varies. At times the frequency is slightly greater than that of the original R-F carrier, and at other times the frequency is slightly lower. We have what amounts to frequency *variation.*

Such a carrier signal is said to be *frequency modulated.*

To better understand the difference between the two carriers after they have been modulated by each of the two methods, suppose you take a look at the two graphs in Figure 3. The top graph

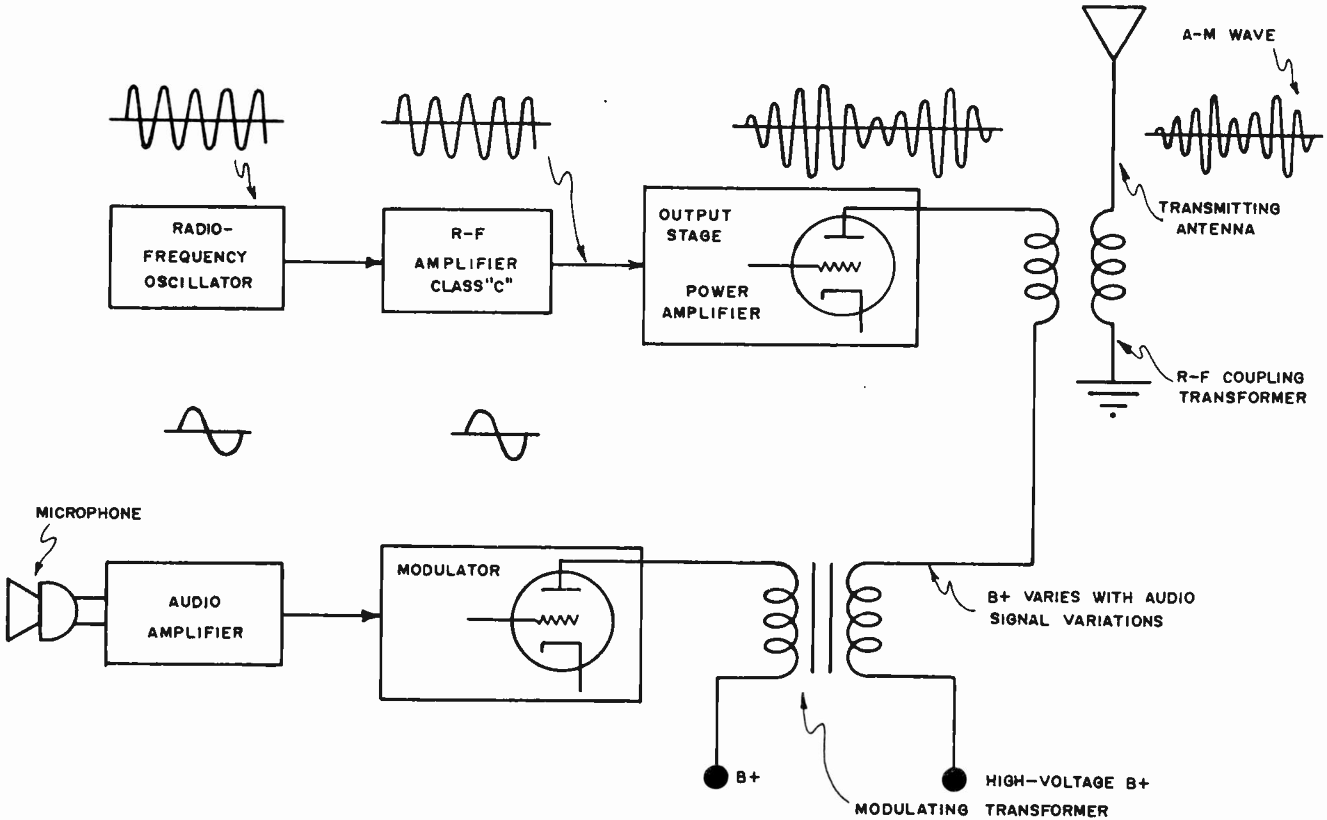


FIG. 4.—*How Amplitude Modulation Occurs.*

in Figure 3 is the same, or approximately the same, as that of the modulated signal in Figure 1. The bottom graph in Figure 3 is the same as that of the frequency-modulated signal in Figure 2.

Both carrier signals are—presumably—carrying the same intelligence. In one case the intelligence is being carried in the form of an amplitude modulated signal, in the other it is being carried in the form of a frequency modulated signal.

In one case the amplitude of the carrier signal is varied in conformance with the intelligence being transmitted. In the other case, the frequency of the carrier signal is varied to conform with the intelligence being transmitted.

4. HOW AN FM SIGNAL COMPARES WITH AN AM SIGNAL

Theoretically, there is no reason why an AM carrier and an FM carrier cannot operate on similar frequencies. In practice, we frequently find great differences.

In strictly radio work, amplitude modulation may be employed with either low-frequency carriers or with high-frequency carriers. In practice, we find it being employed more widely on the lower carrier frequencies than on the higher ones.

In the case of frequency-modulation, much the same thing holds true in theory, but practice is somewhat different. Theoretically, a low-frequency carrier can be frequency modulated as well as a high-frequency carrier; in practice, frequency modulation is used only at the higher carrier frequencies.

An exception to this broad statement is the case of radio amateurs. They commonly use FM on frequencies as low as 14 megacycles, and some go even lower than that. Some go even further and chop off one-half of the carrier to attain what they call "Narrow Band FM."

However, this is getting just a little outside the scope of our present discussion, and is scarcely a matter of interest until after we have explained what FM is, and how it works.

Right now our interest centers around the differences between the two principal types of modulation, and the places where each is most widely used. Also a brief discussion of the similarities and differences which exist between signals modulated by the two differing methods.

5. HOW AMPLITUDE MODULATION OCCURS

To begin with, the methods used to modulate carrier signals are different for the two types of modulation. In one case the R-F carrier is generated in a typical oscillator, often crystal controlled. That signal is amplified and sometimes the frequency is doubled or tripled.

Finally, usually in the final stage, the modulation is applied to the R-F carrier. The diagram in Figure 4 provides a rough idea of how this act of modulation is brought about.

Although we have deliberately refrained from going into a full description of the act of modulation, this method is commonly called "plate modulation." It is also called the *Heising method of modulation* after the designer who first used the system.

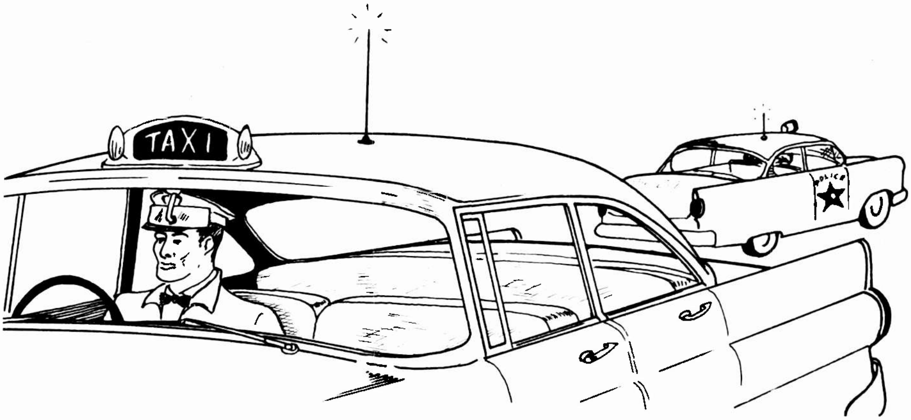


FIG. 5.—*Most Taxicabs and many Police Cars use FM Radio for Communications.*

It will be noted that the audio signal is superimposed upon the R-F carrier by causing the plate voltage of the output R-F amplifier tube to vary from instant to instant as the audio voltage goes through its successive cycles. The audio voltage is superimposed on the R-F carrier through the medium of the *modulating transformer*.

Note that the audio signal from the modulator amplifier is fed into the primary of the modulating transformer. The plate voltage to the R-F output amplifier must pass through that same secondary winding to reach the plate of the R-F tube. Thus, any voltage induced in the secondary winding of the transformer by the varying current in the primary must necessarily act upon the plate voltage of the R-F amplifier tube.

The net result of this action is that the voltage applied to the plate of the R-F amplifier tube varies in accordance with the varying voltage on the audio modulating signal. This varying voltage on the plate of the R-F amplifier causes the amplitude of the R-F carrier signal to vary in accordance with the audio modulating signal.

Note carefully that the action we have been describing is that which occurs during modulation when the R-F carrier signal is being amplitude modulated. The process for modulating the carrier in the case of Frequency Modulation is different.

6. PURPOSES FOR WHICH EACH TYPE OF MODULATION IS USED

As we indicated a little earlier in this lesson, it is hard to make a definite statement and say that Amplitude modulation is used for certain definite purposes, and no others. It is equally hard to say that frequency modulation is used for certain other purposes, and no others.

At the same time, it is possible to outline the general purposes for which each type of transmission is used most commonly, and from that we can obtain a reasonably good idea of how each is used.

Amplitude modulation is used for ordinary radio broadcast systems. Those radio broadcast stations which can be picked up on the average home radio, and the types of radios used in most automobiles, almost always use amplitude modulation.

Much communication between aircraft and ground stations is handled by means of amplitude modulation. Communications between ships at sea is usually in the form of amplitude modulation, although not always.

The R-F carrier frequencies used in these types of operations are nearly always relatively low. By the term "relatively low" we mean to imply that the frequencies are under 30 megacycles. This is an exceedingly broad statement and should be understood to be a most general one. We know full well there are plenty of exceptions.

Still, right now, we are trying to point out the purposes for which each type of radio is used predominantly. In this sense, amplitude modulation is used predominantly for the purposes mentioned.

There is a broadcast band reserved primarily for frequency modulation. It lies at 100 megacycles, ranging from several megacycles below 100 megacycles to several megacycles above 100 megacycles.

Actually, the FM broadcast band occupies 20 megacycles. It ranges from 88 megacycles to 108 megacycles. There is a movement on foot to persuade the FCC to reduce the FM broadcast band by a few megacycles, and reassign those frequencies to TV broadcasting, or for other purposes. But, as of this writing, the band is between 88 and 108 megacycles.

This band is used for commercial broadcasting. It is the band of frequencies which can be picked up on a home FM radio receiver. It is the band over which high quality music is broadcast in so many localities.

But FM radio is also used for other purposes.

Much of the communication between moving vehicles is by means of FM radio. Taxicabs and police cars, such as those in Figure 5, usually use FM radio. A tip-off is the small vertical antenna sticking upward from the roof of the vehicle. Because the carrier frequency is high the antenna is short.

There is a direct relationship between the wave-length of the signal being handled and the length of the antenna. The higher the frequency the shorter the wavelength, and thus the shorter the antenna. This means that a very high frequency signal would use a very short antenna.

When the signal used for communication has a very high frequency the signal is nearly always frequency modulated.

Frequency modulated radio signals are used for other purposes. They are also used for some communications between airplanes and the ground, and between one plane and another.

7. ADVANTAGES OF FM RADIO

FM radio possesses two outstanding advantages over AM radio. One of these deals with the

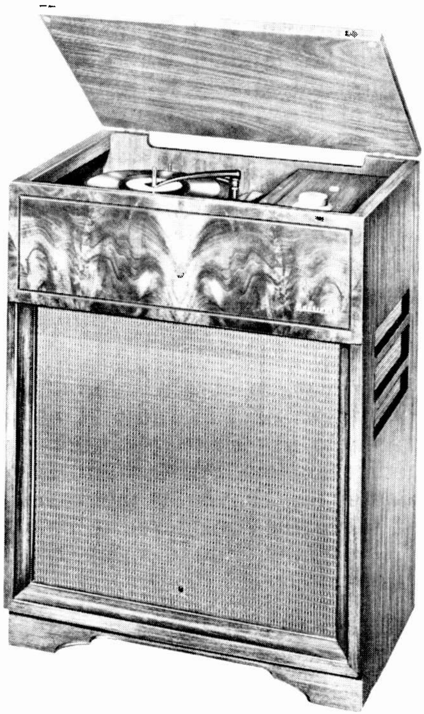


FIG. 6.—*High-Fidelity Receiver Using FM Tuner.*

matter of "electrical noise," and the other with the matter of wide-band, high fidelity audio signal transmission.

So called "electrical noise" consists of radiated electrical signals having random frequencies and random amplitude. For the most part they are created by sparking at the brushes of electrical motors, by sparking in Neon electrical sign transformers, high-tension electrical power lines, and similar things.

These electrical signals radiate through space, and frequently reach the signal circuits of radio receivers. Those signal circuits, being extremely sensitive, and being fully capable of amplifying any signal which reaches the grid of any tube, accept such "electrical noise" signals and pass them along to the audio circuits, and eventually to the loudspeaker.

Signals of that kind, which originate in a small motor at its brushes, act to cause a harsh and unpleasant "buzzing" sound in the loudspeaker.

Distant lightning creates electrical signals which also produce noises in the loudspeaker.

For the most part such signals have widely varying amplitudes, and frequently have disconnected breaks, thus producing irritating noise in the loudspeaker. It is the varying amplitude which produces the audio signal which affects the loudspeaker.

For the most part, FM receivers are not affected by variations in amplitude level, therefore are not sensitive to electrical noises caused by variations in the amplitude of the random signal. In plain words this means that FM receivers do not reproduce such unwanted electrical noises.

So far as the listener is concerned, electrical noises are not heard on FM receivers.

FM receivers have another advantage. This lies in their ability to reproduce a wide range of audio frequencies during the demodulating process.

To explain this a little more clearly we should point out that the AM broadcast channels are so positioned and located that they are only 10 kilocycles apart. Since one-half the frequency range is eliminated during the demodulating process this means that 5 kilocycles is just about the maximum audio signal frequency which can be reproduced. Five kilocycles, as you know, is equivalent to 5000 cycles.

To put this in still other words, it means that audio frequencies up to about 5000 cycles can be handled and reproduced in an ordinary AM radio. However, in the case of fine music, the higher frequency notes are lost during the transmission, and they are not reproduced.

For persons who like fine music the loss of these higher frequency audio notes is disturbing.

FM radio, on the other hand, is fully capable of reproducing audio frequencies much wider than the human ear can hear. This means that any audio frequency which can be handled by the mechanical parts of the radio transmitting and receiving equipment can be reproduced by the electrical circuits of the radio. Largely because of this factor, FM radio music is greatly preferred by persons who know, and appreciate, good music.

The radio receiving equipment used in most of the high-fidelity audio amplifying systems is almost universally FM equipment. Only FM radio tuners are capable of reproducing music requiring the wide frequency handling ability of high-fidelity audio amplifiers. The receiver in Figure 6 is an example of high-fidelity amplifying equipment using an FM tuner.

8. DISADVANTAGES OF FM RADIO

The principal disadvantages of FM radio transmission is that it is normally used only for short range communication. It is considered to be a line-of-sight radio system, meaning that it can be used only for such transmissions as can be handled where the receiving antenna is within line-of-sight of the transmitting antenna.

This is not strictly true.

Technically, FM can be used over longer distances. To be so used it is necessary to use lower carrier frequencies than those normally used with FM transmissions.

However, to obtain some of the advantages of FM radio it is desirable to use relatively high frequencies. Most FM is handled on frequencies above 30 megacycles, and these are considered to be useful only for line-of-sight transmissions.

If the carrier signal is reduced to a frequency lower than 30 megacycles most of the advantages of using FM are lost. So, to obtain the advantages of FM radio it is necessary to use higher frequencies, meaning that it is necessary to use frequencies which are capable of being transmitted for only relatively short distances.

Therefore, it is normal to look upon FM radio as being handicapped by having a relatively short range.

9. CARRIER MODULATION

We have no intention of delving into the problems normally connected with transmission of radio signals. At least, not with this lesson.

At the same time, to explain some of the differences between AM and FM radio systems it is necessary to give some thought to

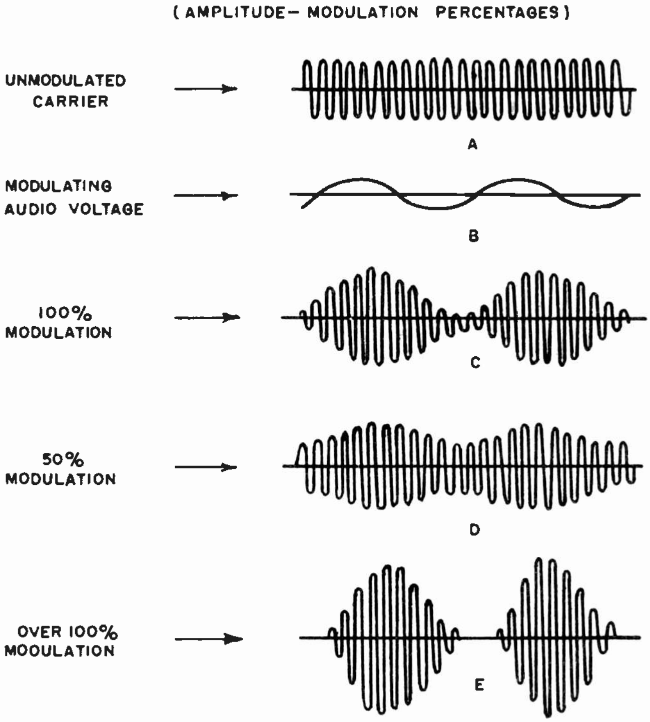


FIG. 7.—*Differing "Percentages of Modulation."*

the matter of modulation so the distinction between them can be better understood.

Back in Figure 4 we gave you a rough outline of the bare essentials of an AM transmitter system. We described briefly some of the things leading up to the act of modulating an R-F carrier.

You can understand that the greater the magnitude of the audio signal the more effect it has on the R-F carrier. To make this emphatically clear we can point out that a weak audio signal has relatively little effect on the amplitude of the carrier signal. On the other hand, a strong audio signal has a high degree of effect on the carrier signal.

The relative effectiveness of the audio signal in varying the amplitude of the carrier is referred to as the "percentage of modulation."

To understand this more clearly suppose you turn your attention to the series of graphs in Figure 7.

The top graph in Figure 7, the one marked *A*, is that of a continuous sine-wave. It represents an unmodulated R-F carrier. It is a carrier which is not being affected, or modulated, by an audio signal in the plate circuit. During the transmission of an R-F carrier signal represented by a graph of this kind the normal plate voltage is not being varied in any manner.

The second graph, the one shown at *B* in Figure 7, represents an audio signal. It is the modulating voltage. We can look upon it as being an audio voltage signal, and usually it is such. However, it could be almost any kind of signal with a frequency substantially below that of the carrier signal.

In AM transmission the audio signal is impressed upon the carrier signal in such a way that the carrier's amplitude is varied in accordance with the frequency and strength of the modulating signal. This has been pointed out previously.

However, the modulating signal does not always have the same strength. Sometimes it is weaker than the carrier signal. In other cases its maximum strength is comparable to that of the carrier. In extreme cases the strength of the modulating signal may be greater than that of the carrier.

The relative strengths of the two signals give rise to a condition which is described as "percentage of modulation." What this means is just how much the modulating signal is capable of modulating the carrier.

In the case of the carrier signal shown at *C* in Figure 7, the modulating signal at its maximum amplitude has a strength virtually equal to that of the carrier. Such a modulating signal has the ability to increase the amplitude of the individual cycles of the carrier signal when the audio modulating signal half-cycle is adding to the plate voltage of the carrier's output amplifier tube. On the other hand, during the half-cycles of the modulating voltage which oppose the plate voltage on the output amplifier, the modulating voltage signal has the power to virtually cut-off the output of the carrier.

Such a condition is called 100% modulation.

This is the type of modulation most frequently striven for in AM radio work. It is considered the most desirable condition. It is the ideal.

When a carrier is modulated at 100% modulation the system is operating at top efficiency.

However, 100% modulation is not always attained in practice. In fact, it is rarely ever attained, although efforts are made to approach 100% modulation as closely as possible.

A more common degree of modulation is called "under-modulation." The graph at *D* in Figure 7 shows the carrier modulated at 50%. It is quite common for the carrier to be modulated somewhere between 50% and 100%.

Still a third degree of modulation is called "over-modulation." This is shown by graph at *E* in Figure 7.

"Over-modulation" is a condition in which the carrier is actually interrupted during the peak of the negative half-cycles of the modulating voltage. This means the modulating voltage is so strong that its negative half-cycles are stronger that the plate voltage supplying the R-F output amplifier.

When the modulating voltage goes into its negative half-cycles it over-balances the plate voltage on the R-F output amplifier to such an extent that the plate is driven negative. This cuts off the tube and it ceases to work until the negative modulating voltage is removed from the tube.

This is not a desirable condition.

When a carrier is over-modulated it has a tendency to create unwanted interference of various kinds. The FCC frowns on over-modulation, and the condition should be avoided so far as possible.

It is largely because of the fear of causing over-modulation that most AM modulation is handled at less than the optimum 100%.

10. MODULATION IN FM RADIO

The amplitude of the carrier signal is not affected by the modulating signal in FM radio. Instead, the audio modulating signal is applied to the carrier in such manner that it causes the frequency to vary, rather than causing the amplitude to vary.

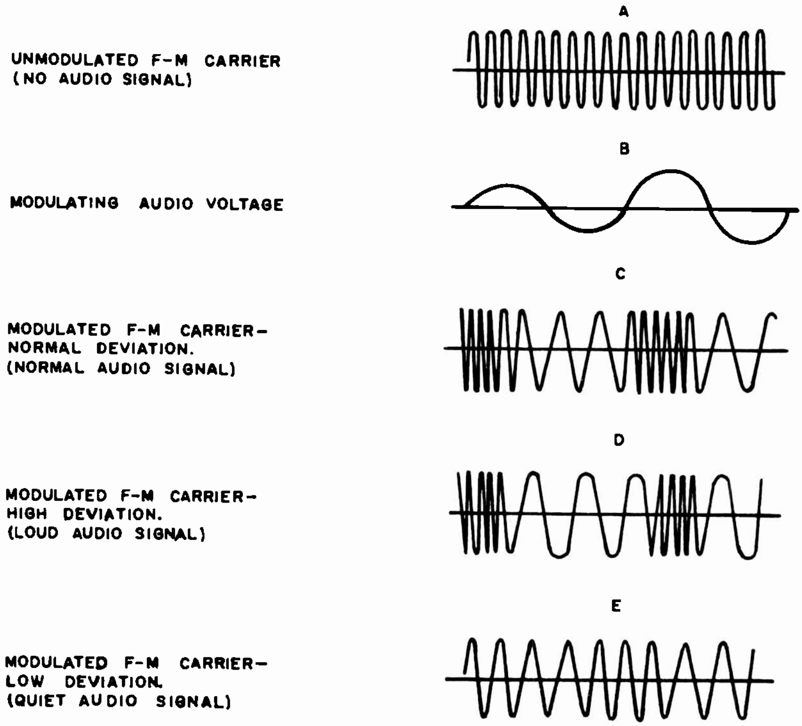


FIG. 8.—*Frequency "Deviation" caused by Modulating Signal.*

The series of graphs in Figure 8 helps to make this point clear.

The top graph in Figure 8 is another unmodulated carrier signal, similar in all respects to that in the upper graph in Figure 7. It represents a continuous sine-wave which is not varied, or "modulated," in any way.

The second graph in Figure 8 is also similar to the second graph in Figure 7. It is a typical audio signal, being used as the modulating signal.

It is in the three lower graphs in Figure 8 where we find the signal differing from those shown in Figure 7. This shows the degree of "deviation."

"Deviation" in FM radio is roughly comparable to "Percentage of modulation" in the AM radio. However, the two meanings are by no means identical, nor synonomous.

The amplitude of the carrier signal in FM radio remains constant. At least it is supposed to remain constant. If the amplitude varies it has no significance insofar as handling the intelligence of the transmitted message is concerned.

In FM radio the unmodulated carrier has some specific "normal operating frequency." This can be considered as being equivalent to the normal frequency of the carrier used in AM radio.

In FM radio this is called the *mid-frequency.*

It is also called the *center frequency.* In either case it means the same thing. It means the frequency of the unmodulated carrier.

When modulation is applied to a carrier to achieve frequency modulation the frequency of the carrier *deviates* from the mid-frequency. During one-half of the modulating signal cycle the frequency of the carrier is increased above the mid-frequency. During the other half of the modulating signal cycle the frequency of the carrier is reduced slightly below that of the mid-frequency.

For all practical purposes we can think of the *frequency deviation* in FM radio as corresponding to *percentage of modulation* in AM radio.

11. FREQUENCY DEVIATION

To acquire a concrete understanding of frequency deviation it is desirable for us to apply some definite and specific figures and values to the subject under discussion.

For the purpose of providing an example let us consider the case of an FM carrier whose mid-frequency is 100 megacycles. This means that the carrier without modulation has a frequency of 100 megacycles. This is also sometimes called the "resting frequency."

Let us now suppose that the 100-megacycle carrier signal is to be modulated through the form of frequency changes by an audio signal voltage. During the process of modulation the frequency of the carrier will vary from some frequency slightly below 100 megacycles to some other frequency slightly above 100 megacycles.

Let us suppose the transmitter is so designed that full modulation by the audio signal will cause the frequency of the carrier to vary by as much as 75 kilocycles on each side of its normal mid-frequency. In terms of actual cycles this means that the frequency of the carrier will then vary from a value of 75,000 cycles below 100 megacycles to a value of 75,000 cycles above 100 megacycles.

Again, these figures can be changed into actual values.

The lowest frequency to which the carrier would swing would be 75,000 cycles below 100 megacycles. This means that the lowest frequency of the signal would be 100,000,000 cycles minus 75,000 cycles. This works out to be a minimum frequency of 99.925 megacycles.

The highest frequency to which the carrier would swing would be 75,000 cycles above 100 megacycles. This means it would swing upwards to 100.075 cycles.

Now let's look closely at these figures.

After modulation, the frequency of the carrier will vary between 99.925 megacycles and 100.075 megacycles. This is a maximum swing in the frequency over a range of 150,000 cycles, or 150 kc.

The *frequency deviation* of the signal is the difference between the mid-frequency and either the upper or the lower frequency limit. In the case of the signal being described the frequency deviation would amount to 75,000 cycles.

12. LIMITATIONS ON FREQUENCY DEVIATION

The limits within which a frequency modulated carrier signal is permitted to swing are controlled by the FCC. They are further controlled by practical considerations.

The FCC is interested in the frequency deviations because they want no interference between the signal from one station and that of a station operating on some adjacent frequency either above or below. Unless the frequency deviation were limited it would be possible for the signal from one station to swing so low as to interfere with the signal of another station.

There are equally compelling practical considerations.

For one thing, there is no real reason why the signal frequency needs to deviate any further. On the other hand, devising circuits to accomodate an FM carrier signal which is deviating more than 75,000 cycles presents problems.

It is necessary that the tuned circuits in the receiver be so designed that they can accommodate the full deviation of the signal without introducing distortion into it. Otherwise, there is danger that the audio signal will be so distorted as to cause trouble.

Receivers for receiving FM signals are designed in a manner similar to those designed for receiving AM signals. The only basic difference is in the manner in which the carrier is demodulated.

The R-F stage of the receiver picks the signal from space, and provides it some degree of amplification. This is indicated by the block diagram in Figure 9. From the R-F stage the signal is passed along to the mixer stage.

From the details supplied by the block diagram there is little so far which distinguishes an FM receiver from one intended to receive AM signals. In point of fact there is no difference.

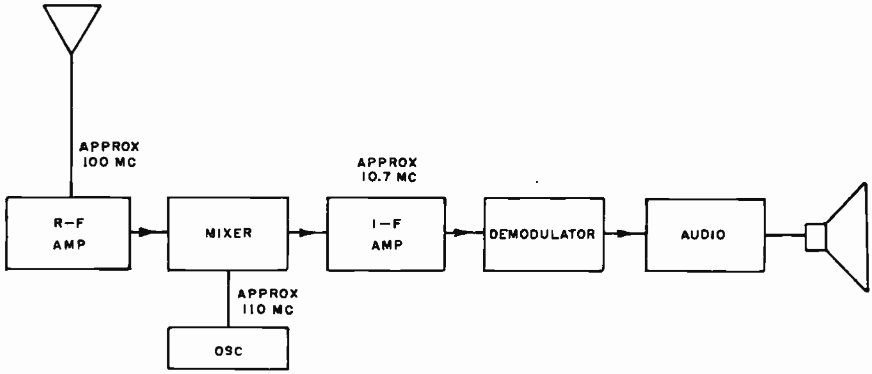


FIG. 9.—*Block Diagram of FM Receiver.*

However, most FM receivers have their circuits designed to handle only signals which have a high frequency. If you study the block diagram carefully you will note that the R-F signal handled by the R-F amplifier has a frequency of approximately 100 megacycles.

This would be a typical frequency for an FM receiver.

Note, also, that the frequency of the oscillator is approximately 110 megacycles. This is also typical. The frequency of the oscillator in an FM receiver is approximately 10 megacycles higher than that of the incoming R-F signal.

Actually, in many commercial FM receivers the I-F frequency is 10.7 MC, which means that the oscillator signal would have a frequency of 110.7 MC, if the R-F signal is exactly 100 megacycles.

Going a step further, you will note that the I-F frequency in the block diagram in Figure 9 is 10.7 MC.

In each of these stages of the receiver the frequencies are much higher than those normally encountered in most AM radio receivers. However, there would be no other specific difference between the circuits for any of these stages, regardless of whether the receiver was intended to receive AM or FM signals. So far as capability is concerned, the circuits of such a receiver could handle either AM or FM signals. At least, up to the detector stage.

From a practical point of view, such a receiver would handle only FM signals, since that is the only kind of signals operating within that frequency range.

It is in the *demodulating* stage that we find the only practical difference between an AM and FM receiver. An AM radio has what we call a "detector" stage. The corresponding stage in an FM receiver is more commonly called a "demodulator."

However, special types of demodulator circuits have been given special names. There is the "discriminator" circuit, invented by

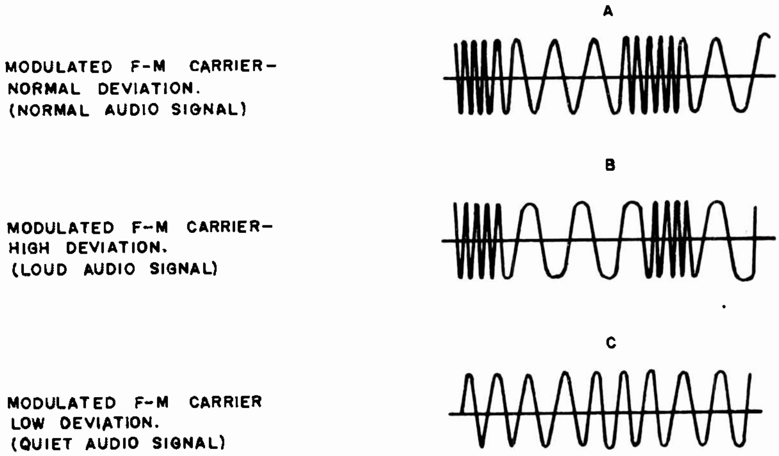


FIG. 10.—*Degrees of Modulation of FM Signal.*

Major Armstrong, the inventor of FM radio. There is also the "Quadrature detector," or "Gated Beam" detector, first used by Zenith, and later copied by many other manufacturers.

Finally, there is the so-called "ratio detector," devised by RCA to avoid patents on the original FM discriminator. *Ratio detectors* are found on millions of inexpensive FM receivers, especially in table models.

The audio section of an FM receiver is essentially the same as that for any other kind of radio. However, since an FM receiver is inherently capable of reproducing a much wider range of audio signals than an AM receiver, it is frequently found that on the better quality FM radios the audio section is more carefully designed, and is capable of handling the higher frequency audio signals passed along by the FM demodulator.

13. DEGREE OF DEVIATION

If we accept 75,000 cycles of deviation as being the maximum permitted by regulations and in practice, we can associate that limitation with 100% modulation as in the case of amplitude modulation. You will recall we explained earlier that 100% is the maximum amount of modulation permitted in AM radio.

In FM radio we can look upon a deviation of 75,000 cycles from the mid-frequency of the carrier as being the maximum amount of modulation permitted.

We can show this action in the form of graphs. In Figure 10, for example, the graph at the top at *A* can be looked upon as being *normal* modulation.

This does not necessarily mean this is the maximum degree of modulation. Rather, we can think of it as being about midway between no modulation and full modulation.

On the second graph, at *B* we have an "over-modulated" carrier. We see the distance between some cycles is too little, that the cycles occur too frequently. At the other extreme, the distance between them is too great, meaning that the cycles occur too far apart.

Finally, at *C*, the bottom graph, we have a case of "under-modulation." In that case there is not enough modulation to do the job properly.

In the case of "under-modulation" the highest frequency attained by the carrier is not very high. On the other hand, the lowest frequency is not very low. In both cases the frequency comparison is made against the mid-frequency, or "resting" frequency, of the carrier.

When a carrier is frequency modulated in this manner there is no danger the carrier will be over modulated. Nor will there be distortion.

The signal that is received—the audio signal—will be quiet.

Deviation of the carrier is approximately 10 kc rather than the full 75 kc which is permissible.

14. LOUDNESS AND PITCH OF THE AUDIO SIGNAL

In any kind of radio transmission the final objective is to carry some form of intelligence to the receiver, and reproduce it with some degree of faithfulness.

This means that the original audio signal should be so reproduced that the pitch of each signal cycle is reproduced in full faithfulness, and amplitude of the various audio cycles is reproduced in true tone. The question probably filling your mind is how this is accomplished in the case of FM radio transmission.

To understand this better, suppose we return our attention to Figure 10.

The upper graph in Figure 10 represents a frequency-modulated signal whose carrier frequency is varying between an upper limit and a lower limit. The *number of times that frequency varies up and down represents the number of* times the audio component goes through its cycles.

For example, suppose that the carrier frequency varies from its upper limit to its lower limit 155 times within the period of a second. That means that the audio component being carried by the carrier signal has a frequency of 155 cycles.

In terms of music, the frequency with which the carrier signal varies between its upper and lower frequencies represents the "pitch" of the audio signal. This is the same as the *frequency* of the audio signal.

Note this: the frequency of the *rise and fall* of the carrier signal represents the exact frequency of the *audio signal.*

Then there is the loudness—the amplitude of the audio signal.

This is represented by the swing —or deviation—of the carrier signal frequency.

This can be said another way. If the carrier swings from 100 mc. to 100.065 mc. the audio is louder than if the carrier had swung from 100 mc. to 100.045 mc.

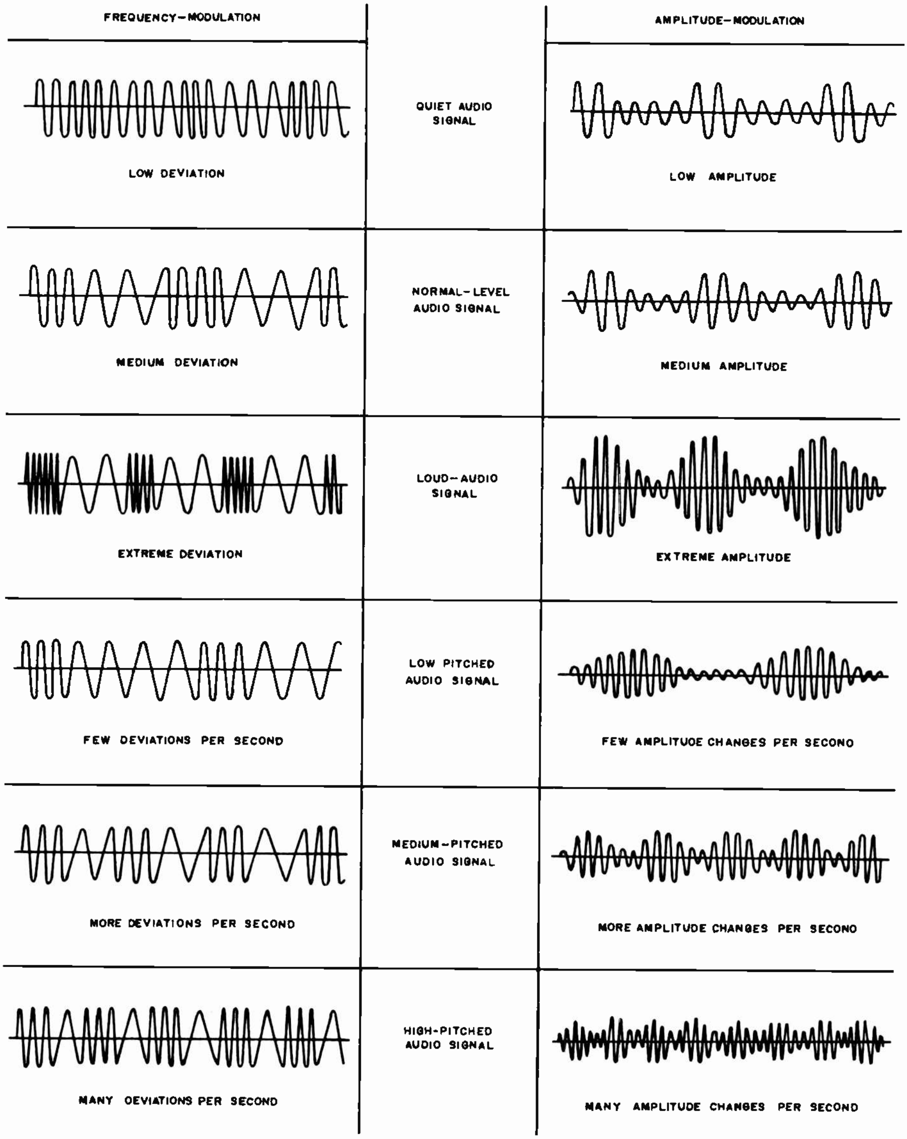


FIG. 11.—*Comparison of Modulated Signals in FM and AM Radio.*

FIG. 12.—*The Earth is a Huge Ball.*

The farther the carrier frequency swings from its mid-frequency the louder is the audio signal.

In FM radio, the loudness, or amplitude, of the audio signal is directly proportional to the degree by which the carrier deviates from its normal or "resting" frequency. Compare this with AM radio. In AM radio the loudness of the audio is comparable and proportional to the amplitude of the carrier.

The graphs in Figure 11 provide a good comparison between AM and FM signals.

15. DEVIATION RATIO

In FM radio we have a term which is not used in connection with AM radio. It is *deviation ratio.*

This is a term which is often used erroneously, even by men with considerable experience in radio work. The term is used to compare the frequency of the highest audio note with the maximum deviation of the carrier.

Note this carefully. It is easy to be confused.

The ratio is between the *deviation* of the carrier—its maximum deviation—and the highest *frequency of the audio signal* being handled.

To understand this more clearly, suppose we reduce it to practical figures. We will suppose that the maximum deviation of the carrier is 75,000 cycles, a common and typical condition.

We will also suppose that the highest audio frequency to be transmitted has a frequency of 15,000 cycles.

In this case the maximum deviation of the carrier frequency is 5 times greater than the maximum audio frequency to be transmitted. This is a ratio of 5-to-1, or just plain 5. This is the deviation ratio.

Significance of the deviation ratio is that the higher the deviation ratio the better the noise level of the receiver. A high deviation ratio makes for a low noise level.

16. RANGE OF FM SIGNALS

FM radio is commonly looked upon as being a short range system of radio communication. There are several factors which cause this.

One of the peculiarities of FM radio which makes it useful is the fact that it possesses the ability to reproduce original sounds with a high degree of true fidelity. To achieve that true fidelity it is necessary to have a relatively high deviation ratio.

To obtain this wide band-width it is necessary that the carrier frequency be quite high. There are a number of technical reasons for this situation.

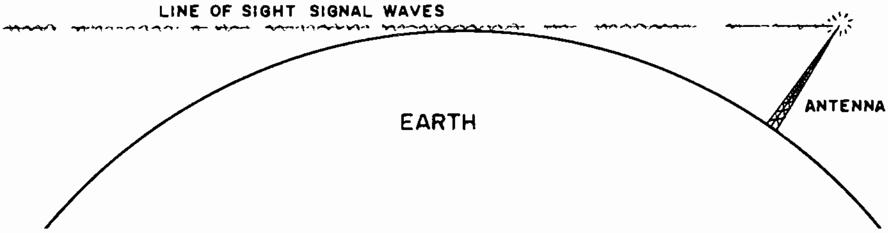


FIG. 13.—*High-Frequency Signal Waves Travel in Straight Line and Leave the Earth's Surface.*

The important point is that most FM transmitters work on frequencies well above 30 megacycles.

We have not stressed this point previously, but as carrier frequencies increase we find them acquiring many of the properties of light. Instead of being refracted, and reflected, they tend to travel in straight lines.

Since the Earth is a huge ball, and the surface of the Earth is a constant curve, this means that no portion of the surface is a true level. See Figure 12.

Carrying this thought a step further, we can see that an antenna erected at any location on the Earth's surface can radiate signals, but when those signals have a sufficiently high frequency they tend to travel in a straight line. That action, in turn, forces the signal waves to leave the surface of the Earth and move outward into space. This is indicated in Figure 13.

You may wonder why this happens to high-frequency signal waves, but not to signal waves which have lower frequencies. The answer lies in a peculiar condition involving what are known as "sky waves," and reflected signals.

The drawing in Figure 14 helps to make this clear.

High above the ground of the Earth is a layer of ionized atmosphere. Just what causes the ionization has never been clearly determined, but it is generally believed that cosmic rays from outer space act upon the atoms in the atmospheric layer, and ionize some of them.

There is also some evidence that unusual activity on the sun, such as active "sun spots," has an effect on the ionized layer of atmosphere.

In any event, presence of the layer of ionized atmosphere acts as a reflector for the lower frequency radio waves, and causes a major portion of the signal strength to be reflected back to the earth. This makes it possible to transmit low-frequency radio waves over long distances by taking advantage of what is called "skip," or reflection.

Low frequency radio waves have relatively long waves. The wavelength of such low-frequency signals is so much greater than the thickness of the surface of the ionosphere that reflection occurs in much the same manner that

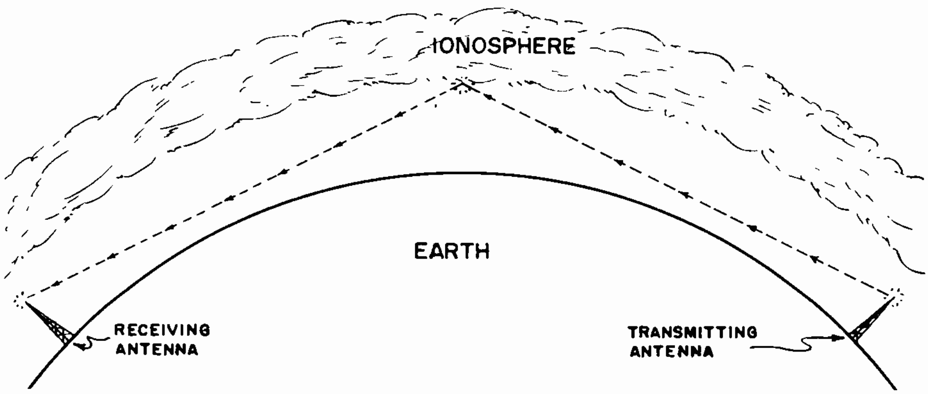


FIG. 14.—*Low-Frequency Signals are Reflected from Ionosphere.*

reflection of light occurs from the surface of a smooth mirror. This causes low-frequency radio waves to be reflected back to Earth in far distant places.

In the case of the high-frequency signals, those having frequencies above 30 megacycles, the wave-length is so short that very little of the signal energy is reflected. Therefore, such signals continue moving outward through space, and does not return to the Earth.

FM radio signals usually have frequencies above the level which are normally reflected back to earth. This means that transmission of such signals is limited to line-of-sight.

In practice, the transmitting antenna is raised to considerable height. This makes it possible to radiate FM signals for distances up to 75 and 100 miles, and in some cases to greater distances. For many purposes this is sufficient distance to accomplish the purpose for which that kind of radio is used.

In the case of FM broadcast stations, the antenna is usually mounted on the top of a high building in the center of a metropolitan area. Having a range of 75 miles, or somewhat more, in all directions, permits the station to serve a relatively large listening audience.

FM radio has proven itself especially useful for a number of purposes in the field of communications. It is used very widely between base stations operated at the main offices of taxicab companies to communicate with moving taxicabs on the street. FM radio is ideal for this purpose.

For the most part such taxicabs rarely have reason to travel beyond the normal operating range of the FM transmitter. Therefore, the limitations with respect to distance is not a matter of great importance. In fact, the short distance such signals travel is a distinct advantage; it avoids interference with stations in other cities.

Due to the ability of an FM demodulator circuit to pass audio signal frequencies, without permitting noise signals to pass, enables an FM receiver to reproduce spoken messages with a minimum of background noises. This is a matter of supreme importance in the case of a moving vehicle where normal electrical noise levels are always high.

When so used, FM radios provide excellent radio communications under conditions where ordinary AM radio would be "swamped" with electrical noise, and the desired voice signal blanked out. Furthermore, the carrier signal requires very little power, because there is no need to use power to blank out noise.

Much the same thing holds true with respect to police radios. Also for radio equipment used in fire department vehicles, public utility vehicles, and other kinds of moving vehicles.

Several companies have specialized in building FM radio systems for mobile service of this kind. Outstanding among them is the Motorola Company, which builds more than half of all the mobile radio equipment now in use in this country.

Another major builder is the Hallicrafters Company. At one time it was the leader in building this type of equipment, but during

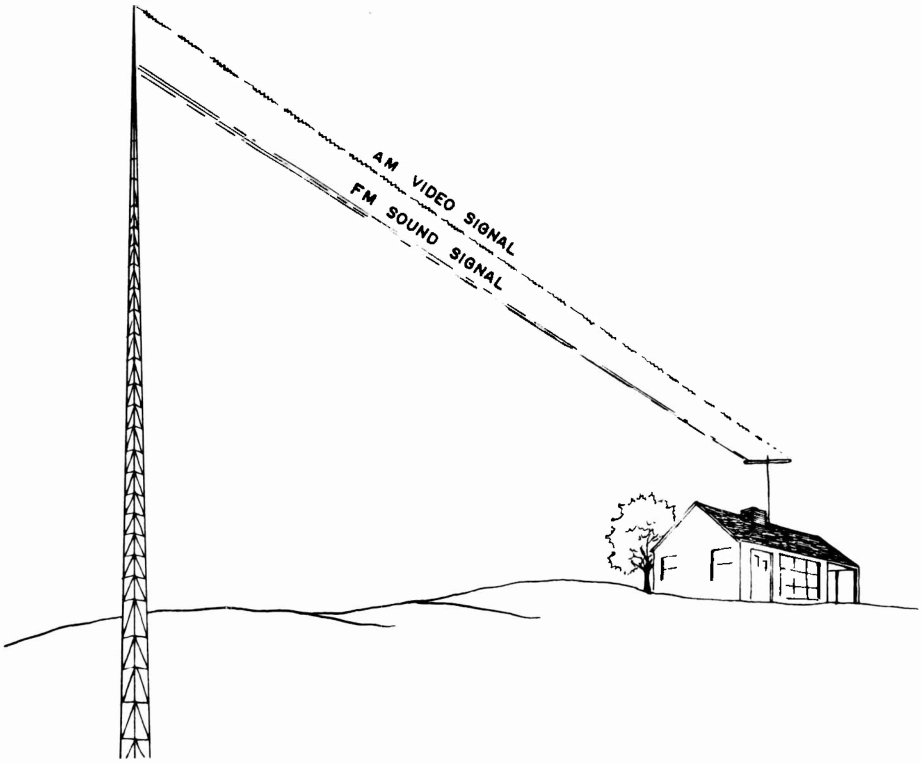


FIG. 15.—*Two Carriers, One FM and the Other AM, are Needed to Transmit a TV Program.*

recent years its attention has been directed toward other types of radio gear, so that Motorola has taken over undisputed leadership.

Still a third company which builds a sizeable amount of communications radio equipment for this class of service is the Collins Radio Co., of Cedar Rapids, Iowa. This company is also a major builder of radio gear for the armed services.

One of the earliest uses of FM radio equipment was in armored tanks during the Second World War, and for communications between jeeps and various headquarters units. Usefulness of FM radio for that purpose was that its range was so short there was little likelihood of the signals reaching the more distant enemy. Further, due to its noise-free characteristics, the signals could be handled with low power, yet be readily understood.

17. FM IN TELEVISION

The most widely spread use to which FM is put is in connection with television.

We have already hinted that two types of carrier signals are used to transmit a television program from the studio to the listener. One of these signals is Amplitude modulated. The other signal is frequency modulated.

The video portion of the program—the portion carrying the picture intelligence— is amplitude modulated.

The sound portion of the program, however, is frequency modulated.

This means that picture information is carried by means of one carrier signal with the intelligence needed to reproduce the picture being carried in the form of amplitude modulation. The sound portion intelligence is transmitted in the form of frequency modulation on another carrier signal.

In the TV receiver one set of circuits are designed to amplify and detect the AM signals. Another set of circuits amplify and demodulate the FM sound signals.

We will make no attempt to describe such circuits in this lesson. We merely mention this fact to illustrate the importance of understanding as much as possible about FM radio systems.

The block diagram in Figure 15 provides some idea of the arrangement of the circuits in a TV receiver. It does little more than outline the arrangement, however, and it should not be thought that the diagram does much more than provide a most general idea.

ITEMS OF SPECIAL INTEREST

Almost everybody has heard of FM radio, even though they may not know exactly what it is.

FM radio is merely the abbreviation for "frequency modulation" referring to the manner in which the signal intelligence is impressed on the carrier signal.

In FM radio the intelligence of the signal is carried in the form of frequency variations rather than in that of amplitude variations.

During the process of modulation the carrier signal is changed or varied in some manner so that such variation at the receiving end can be used to reproduce the original intelligence.

In the process of amplitude modulation the amplitude of the carrier signal is varied from instant to instant in accordance with the varying cycles of the modulating audio signal.

In the process known as frequency modulation the amplitude of the carrier remains constant, but the frequency of the carrier is varied from instant to instant.

The carrier signal in FM radio normally has a very high frequency. In most cases the carrier frequency exceeds 30 megacycles, although there are certain exceptions to this general rule.

One of the most common methods for achieving amplitude modulation of a carrier signal is to vary the plate voltage on the final power amplifier of the transmitter. When the plate voltage is varied in accordance with the audio modulating signal it causes the amplitude of the carrier to vary in a like manner.

Frequency modulation has certain advantages over ordinary AM, but it also has disadvantages.

Since FM radio requires relatively high carrier frequencies it means that transmission of the signal is restricted to what is generally referred to as line-of-sight. This is its biggest disadvantage, although it is looked upon as an advantage for certain purposes.

FM radio receivers, if properly constructed, are almost completely free from electrical noise. This is a big advantage for a number of purposes.

FM radio is also capable of reproducing a much wider range of audio frequencies than ordinary AM radio. FM radio is widely used in high-fidelity sound systems.

FM radio systems are widely used for communications purposes between moving vehicles and base control stations. Its noise-free characteristics make FM radio ideal for this purpose.

FM signals can be transmitted at low power for communications purposes, yet the received signal is not drowned in electrical noises as would be the case of the AM signals.

In AM radio the optimum modulation is what is known as 100% modulation.

100% modulation is rarely achieved in practice for fear of over-modulating the signal. Over-modulation is undesirable since it causes interference on nearby frequency channels.

In FM radio the unmodulated carrier signal has what is commonly known as a *mid-frequency*. It is the frequency at which the signal is generated when no modulation is placed upon it.

The mid-frequency is also known as the "resting frequency," and sometimes as the "center frequency."

When a high-frequency carrier is frequency modulated the modulating signal causes the carrier frequency to vary somewhat above and below the normal mid-frequency.

Tendency for the carrier frequency to vary at frequencies above and below the mid-frequency is called *frequency deviation.*

Frequency deviation in FM radio is roughly comparable to percentage of modulation in AM radio.

Frequency deviation is restricted in FM radio transmissions to 75,000 cycles each side of the mid-frequency.

The term "deviation ratio" refers to the ratio between frequency deviation of the carrier and the maximum frequency of the modulating signal.

Loudness in an FM signal is controlled by controlling the deviation of the carrier signal. The greater the deviation the louder the signal. Conversely, the less the deviation the softer the signal.

Pitch of the audio signal in FM radio is controlled by the number of times each second the carrier signal swings from maximum frequency to minimum frequency and returns.

Each time the carrier signal varies from its maximum frequency to its minimum frequency, then returns to the maximum frequency, it is equivalent to one cycle of the modulating signal.

Range of FM signals is restricted by the fact that the high carrier frequencies which are used in FM radio work are not readily reflected by the ionosphere which surrounds the Earth. FM signals are commonly restricted to line-of-sight distances.

The sound signals in television broadcasts are handled by frequency modulation.

Television receivers must provide means for amplifying and demodulating FM signals as well as AM signals.

a. unmodulated.
b. audio modulated.
c. amplitude modulated.
d. frequency modulated.

3. Carrier frequencies above approximately 30 megacycles:
   a. are able to travel long distances around the earth.
   b. are readily reflected by the ionosphere.
   c. cannot be used as TV carrier frequencies.
   d. have such short waves very little signal energy is reflected to Earth.

4. The ratio between the deviation of an FM carrier and the highest frequency of the modulating signal is called:
   a. deviation constant.
   b. frequency deviation.
   c. deviation ratio.
   d. percentage ratio.

5. The loudness of an FM signal is represented by:
   a. the swing, or deviation, of a carrier frequency.
   b. the frequency at which the signal varies.
   c. percentage of modulation.
   d. deviation ratio.

6. "Percentage of modulation" in AM carrier signals is comparable to what in FM carrier signals?
   a. degree of modulation.
   b. deviation ratio.
   c. deviation.
   d. percent of deviation.

7. The "normal operating frequency" used in describing AM carriers is equivalent to what in FM carriers?
   a. modulating frequency.
   b. demodulated frequency.
   c. mid-frequency.
   d. frequency deviation.

14. What gives R-F signals below 30 megacycles the ability to travel long distances?

a. reflection from the ionosphere. b. refraction by stratosphere.
c. sun spots. d. curvature of the Earth.

15. The sound portion of a television broadcast is transmitted by:

a. amplitude modulation. b. frequency modulation.
c. the video transmitter. d. composite video signal.

Part 2.

16. When the filament of a vacuum tube is used as the cathode, the tube is said to have:

a. a directly heated cathode. b. an indirectly heated cathode.
c. a transistor action. d. uncoupled.

17. Four things which affect the resistance of a conductor are:

a. size of conductor, weight, temperature, and length.
b. size of conductor, length, material of which made, temperature.
c. material from which made, weight, temperature, and length.
d. temperature, length, size, and weight.

18. The fourth band of color on a carbon resistor is used to indicate ______________.

19. The cross-sectional area of a conductor is measured in ____________________.

20. An alloyed metal used to make good quality permanent magnets is called:

a. soft iron. b. stainless steel. c. copper. d. alnico. e. permalloy.

8. When the modulating signal is applied to an AM carrier over-balances the normal plate voltage, it results in:
   a. under-modulation
   b. over-modulation, and distortion.
   c. 100% modulation.
   d. 50% modulation.

9. One disadvantage of ordinary FM radio is:
   a. the signal carries too far.
   b. the difficulty of achieving modulation.
   c. useful only over short distances
   d. signals are not clear.

10. Audio frequencies above 5000 cycles are:
   a. handled easily by AM radio carriers.
   b. never heard by human ears.
   c. transmitted by FM, but not AM radio.
   d. cannot be handled by FM.

11. "Electrical Noise" originating in lightning discharges, electro-mechanical devices, and similar sources:
   a. are troublesome in FM radio.
   b. cannot be heard on AM receivers.
   c. are not normally reproduced by FM receivers.
   d. can be ignored.

12. What is the greatest frequency deviation normally permitted in connection with frequency modulation?
   a. 2500 cycles. b. 5000 cycles. c. 25 kc. d. 75 kc. e. 150 kc.

13. At what percentage of modulation is an AM carrier operating when it is operating at top efficiency?
   a. 22 1/2% b. 45% c. 63% d. 85% e. 100% f. ________

TRAINING SERVICE

Print or Use Stamp:

Date__________

Name__________ Student No.__________

Address__________ Zone_____ Grade__________

City__________ State__________

Lesson 47

FREQUENCY MODULATION
Part 1.

Underline or fill in the ONE correct answer, just as in previous lessons.

1. Tendency for a carrier frequency to vary at frequencies above or below the normal mid-frequency is called:
   a. deviation ratio.
   b. frequency deviation.
   c. deviation constant.
   d. frequency constant.

2. The video portion of a television broadcast is:

MASSEY TECHNICAL INSTITUTE, Inc.

MTI

Jacksonville, Florida