ACHIEVEMENT THROUGH ELECTRONICS

CLASS C RF POWER AMPLIFIERS

C205

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

CLASS C RF POWER AMPLIFIERS

C205

STUDY SCHEDULE

KL3M774 **1974 EDITION** **Litho in U.S.A.**

CLASS C RF POWER AMPLIFIERS

A radio transmitter is a device for converting some form of intelligence into electrical impulses suitable for transmission through space. In its simplest form, a transmitter consists of a source of rf energy, called the master oscillator, and one or more stages of rf power amplification.

In practical transmitters, such as the one shown in block diagram form in Fig. 1, there are a number of stages between the master oscillator and the antenna. Since each stage is a form of rf power amplifier, let's briefly discuss the particular role each one plays in the overall operation of the transmitter.

Immediately following the master oscillator is a stage called the buffer amplifier. Its purpose is to present a light constant load to the oscillator, which helps maintain a stable oscillator frequency. The FCC requires that very close control of output frequency be maintained on all radio transmitters under its jurisdiction.

The next stage, called a frequency multiplier, produces an output whose frequency is some multiple of the input frequency. The presence of this stage permits the master oscillator to be operated at a frequency lower than the transmitted frequency. It is much easier to design highly stable oscillator circuits at the lower radio frequencies; therefore, one or more frequency multiplier stages are an essential part of most radio transmitters.

The driver and power output stages provide the remaining amplification necessary to supply power to the antenna. This output power may range from less than 100 watts to 1 million watts, depending on the transmitter type and the purpose for which it is to be used. In recent years low power circuits, which formerly used vacuum tubes, have been

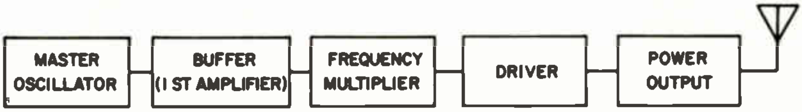


Fig. 1. Block diagram of a basic transmitter. All the stages are operated class C.

redesigned to use economical, efficient transistors. In many low to medium power mobile transmitters, transistors are used in all stages delivering as much as 75 watts of rf output power at 175 MHz. In other transmitters, all but the driver and final amplifier (power output) stages have been replaced by transistors. There is every reason to suppose that this trend will continue as the high-frequency power handling capability of transistors is improved.

The vacuum tube, however, is still used in high power stages of transmitters which are employed in the AM, TV and FM broadcast fields. It will no doubt remain so for quite some time to come. Very large vacuum tubes are required to handle the enormous power outputs of these transmitters. There are three characteristics of all rf power amplifiers, transistors or vacuum tubes which are of concern to us. They are linearity, power gain, and efficiency.

The linearity of an amplifier is a measure of how closely the amplified output follows the input; in other words, a measure of how much distortion is introduced into the output signal by the amplifier. Linear amplifiers, which introduce very little distortion into the signals they amplify, are a subject in themselves and will be considered in a later lesson.

The power gain of an amplifier, usually expressed in db, tells us how much the power level of the input signal is increased by the amplifier. Power gain depends on circuit design and the tube or transistor type used in the circuit. Beam power tetrodes have the highest power gain of any other conventional vacuum tube type. For this, as well as other reasons, the beam power tetrode is the most commonly used tube in modern transmitters. The power gain of transistors does not compare favorably with that of vacuum tubes at higher radio frequencies. This limitation may be partially overcome by adding more stages or using more than one transistor in each stage.

The efficiency of an amplifier, expressed as a percentage, is the amount of dc input power to the stage actually converted to rf energy at the output. In a vacuum tube stage, the power input is the product of the plate supply voltage times the average current. For example, suppose the plate supply voltage is 3000 volts, the plate current 450 milliamps, and the power output of the stage 1000 watts. The dc input power to the stage is:

$$P = E \times I$$

$$P = 3000 \times .45 = 1350 \text{ watts}$$

The efficiency of the stage can then be found by using the following formula:

$$\%\ \text{Efficiency} = \frac{\text{Power Out}}{\text{Power In}} \times 100$$

$$\%\ \text{Efficiency} = \frac{1000}{1350} \times 100 = 74\%$$

In a previous lesson, you learned that class C amplifiers give the highest practical efficiency, ranging up to 75%. This is compared to efficiencies of 35% to 50% for class B and as little as 30% to 35% for class A. However, the high efficiency of class C amplifiers is obtained at the expense of linearity. As you'll remember, output current flows for less than half the input cycle in a class C amplifier. This output current pulse bears little resemblance to the input signal which produced it and is therefore highly distorted.

If the output circuit of the class C amplifier is a resonant tank, this current pulse shock-excites the tank so that a complete sine wave is produced. Thus the nonlinearity of the class C amplifier is effectively eliminated when a single rf frequency (a sine wave) is to be amplified. This, along with the class C amplifier's high efficiency, makes it suitable to many rf power amplifier applications.

In the next section we'll discuss class C rf amplifier fundamentals. The information presented in this section applies to both vacuum tubes and transistors. Later, we'll discuss specific applications of vacuum tube and transistor amplifiers. In the final section, we'll talk about the various adjustments which may be made to both types of amplifier circuits.

RF Power Amplifier Fundamentals

In any amplifier, heat is generated by the current flow through the internal resistance of the stage. The power used to generate this heat represents wasted energy and subtracts from the power that could go to the output. The high efficiency of a class C amplifier is due largely to the fact that current flows for a relatively short portion of the input cycle. It is only during this short conducting period that power-wasting heat is generated within the amplifier. To begin our discussion of class C amplifiers, we'll consider the relationships between the current conducting time and the signal voltage waveforms in an operating class C amplifier.

CURRENT AND VOLTAGE RELATIONSHIPS

The graph in Fig. 2 shows the output current pulse produced by an input signal at various dc bias levels. Look first at the signal at bias level 1. This signal is below the cutoff value of the amplifier for all of the negative half-cycle and nearly all of the positive half-cycle. Output current flows only for the time the input voltage goes above cutoff. Now look at the signal at bias level 2. It has the same amplitude as the first signal but, because we've increased the bias, this signal exceeds the cutoff level for a shorter period. As a result, the output current pulse produced

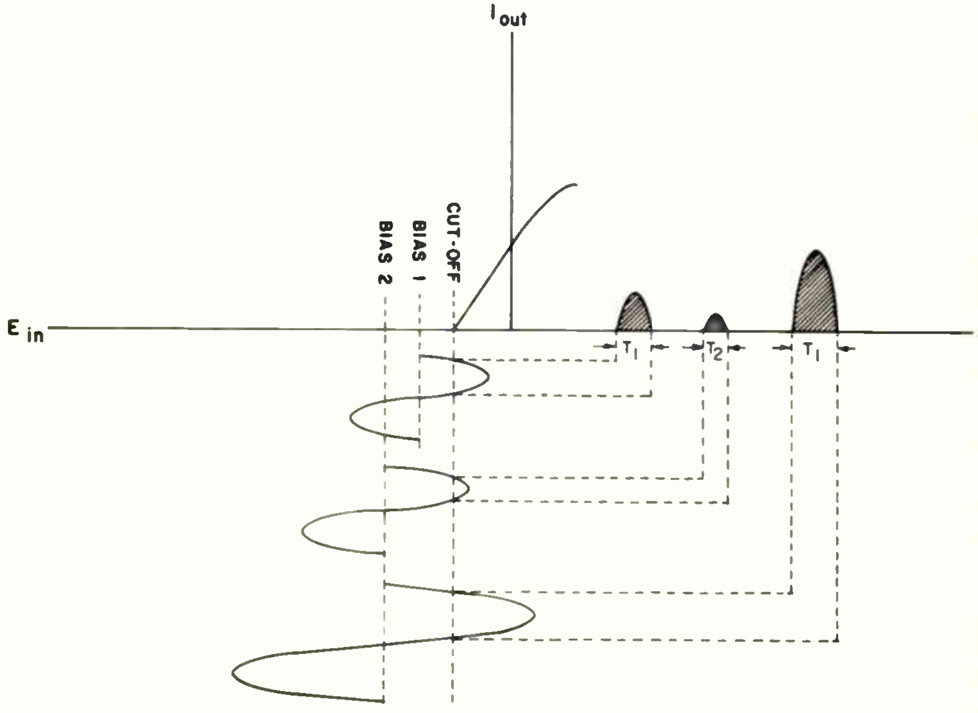


Fig. 2. Output current pulse produced by input signals at various amplitudes and bias levels.

also flows for a shorter period and is lower in amplitude. Increasing the amplitude of the input voltage has the same effect as decreasing the bias. That is, the output current flows for a longer time.

Another way of looking at these basic relationships is shown in Fig. 3. Fig. 3A shows two basic class C amplifiers; one uses a vacuum tube, the other a transistor. Fig. 3B shows the voltage and current waveforms appearing at the inputs and outputs of the amplifiers. Again notice that output current flows only during the period when the input signal exceeds the cutoff level of the amplifier. The output voltage waveform E_{out} is produced by the flywheel action of the resonant output circuit.

Conduction Angle. There are 360 electrical degrees in one complete cycle of a sine wave. The number of electrical degrees the output current flows in a class C amplifier is called the conduction angle or operating angle of the stage. As you've seen, the operating angle depends on both the dc bias and the amplitude of the driving signal. Although amplifier efficiency is higher at the smaller operating angles, the power output is less because the output current pulse is reduced in amplitude and flows for a shorter period. Therefore, the operating angle must be a compromise between maximum efficiency and the highest power output. In making this compromise, the driving signal is maintained at a level sufficient to drive the stage into saturation while the bias is adjusted for the correct operating angle.

Driving Power. To drive a vacuum tube amplifier into saturation requires that the grid be driven positive. The positive grid

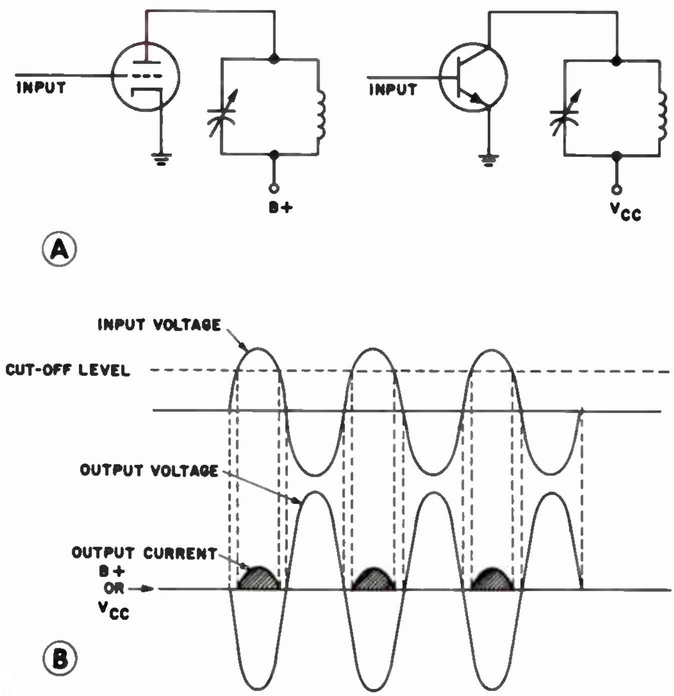


Fig. 3. Basic relationships between current and signal voltages in a class C amplifier.

draws current, causing power to be consumed in the grid circuit. Likewise, in a transistor amplifier, base current flows during the time the driving signal forward-biases the emitter-base junction. The result of this base current flow is that power is consumed in the base circuit.

The power consumed in the input circuit of a class C amplifier, called the driving power, must be supplied by the previous stage. Thus the input circuit of one class C amplifier represents the load on the stage which comes before it. Furthermore, this load presented by the input circuit varies over the period of an operating cycle, reaching a maximum when the input signal causes maximum current to be drawn. As we'll see later, we can use this current flow in the input circuit to develop bias for the stage.

TANK CIRCUITS

The resonant circuit in the output of a class C amplifier has several important jobs to do. We've already mentioned the most basic of these, that of changing the output current pulse into a complete sine wave. This resonant circuit is also required to present the proper load impedance to the stage, and to suppress the undesired harmonics generated within the stage. Let's discuss these last two in detail.

Load Impedance. In order to obtain the maximum power gain from a class C amplifier, or any other amplifier for that matter, the impedance of the load must match the internal impedance of the amplifier. However, do not confuse power gain with power output. It is quite possible that an amplifier operating with a matched load, for maximum power gain, is not delivering its maximum output power.

This is especially true for transistor rf power amplifiers. These amplifiers are very often designed to operate from automotive type battery supplies, thus limiting collector supply voltages to the 12 to 28 volt range. With the load matched to the internal impedance of the amplifier, there may be insufficient collector current flow to give the required power output. Using a value of load resistance lower than the input impedance of the stage results in a greater collector current flow and higher power output. Therefore, in some cases power gain must be sacrificed for power output.

Factors Affecting Impedance. Since the output tank circuit must offer the correct load impedance for the class C amplifier, let's look at some of the factors which affect this impedance. We know that to be resonant, the L and the C of the tank circuit must offer equal reactances at the operating frequency. If we increase the value of L, the value of X_L will increase. To maintain resonance, we must increase X_C by decreasing C. Having increased the value of X_L and X_C by equal amounts, we have increased the total impedance of the circuit without affecting the resonant frequency.

Any resistance present in the tank acts to decrease the total impedance of the circuit. The values of C, L, and R are related to total impedance by the following formula:

$$Z = \frac{L}{CR}$$

From the formula, you can see that increasing the ratio of L to C in the tank causes the impedance to increase. Increasing the resistance in the tank causes impedance to decrease. This leads us to a discussion of tank circuit Q.

Circuit Q. The Q of a coil, as you know, is the ratio of its reactance to its resistance or:

$$Q = \frac{X_L}{R}$$

A capacitor also has a Q, but its value is very large due to the capacitor's low internal resistance. The Q of a tank circuit, therefore, is equal to the Q of the coil.

Fig. 4 shows an amplifier with a parallel-tuned tank circuit in its output. At the operating frequency of this amplifier, let's assume the X_L of the coil is 6000 ohms and its resistance (R_S) is 20 ohms. The Q of this unloaded tank circuit then is:

$$Q = \frac{X_L}{R_S} = \frac{6000}{20} = 300$$

The Q of unloaded tank circuits in practical transmitters may range from 200 to 800. Fig. 5A shows the same amplifier of Fig. 4 inductively coupled to a load. This load might be a transmission line, an antenna, or another rf amplifier. The effect of this load is to reflect an additional resistance into the tank circuit. The equivalent circuit, shown in Fig. 5B, contains this reflected resistance (R_L') in

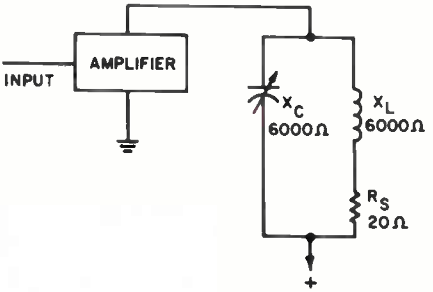


Fig. 4. Amplifier with a parallel tuned output tank showing impedances at resonance.

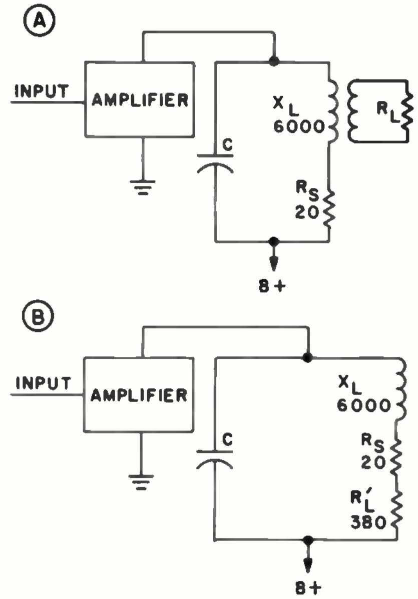


Fig. 5. Tank circuit coupled to a load and its equivalent circuit.

series with the resistance of the coil. The exact value of R_L' depends on the value of the load resistance as well as the coupling to the load. We'll assume a value of 380 ohms for our discussion. The Q of the tank now becomes:

$$Q = \frac{X_L}{R_L' + R_S} = \frac{6000}{400} = 15$$

Thus, the Q of the tank circuit went from an unloaded value of 300 to the loaded value of 15 due to the resistance reflected into the tank circuit by the load. From the previous discussion of tank impedance, you know that this additional resistance in the tank also decreases the impedance of the tank. Tank Q and tank impedance are closely related quantities. Factors which change one will also change the other in the same direction.

Let's see now why this is important.

You know that only the resistance in a circuit consumes power. Inductive and capacitive reactances, under conditions of resonance, merely transfer energy back and forth between themselves. Therefore, when the tank is loaded, all of the power in the circuit is consumed by the resistances R_S and R_L'. The power consumed in R_L' represents power consumed by the load, while that consumed by R_S is dissipated as heat in the tank circuit. From Ohm's Law we derive that $P = I^2 R$, so the power consumed by the load far exceeds that lost as heat in the tank. This is because of a larger value of R_L'. We can actually calculate the efficiency of the tank circuit by the formula:

$$E_{ff} = \left(1 - \frac{Q_L}{Q_U}\right) \times 100$$

In our example, the unloaded Q (Q_U) was 300, and the loaded Q (Q_L) was 15. Therefore:

$$E_{ff} = \left(1 - \frac{15}{300}\right) \times 100$$

$$E_{ff} = (1 - .05) \times 100 = 95\%$$

Suppose we increased the coupling to the load and reflected a larger value of resistance into the tank. This would decrease Q_L without affecting Q_U, resulting in a higher tank circuit efficiency. But remember, tank impedance is dependent on the resistance in the tank, so changing the coupling to the load changes the tank impedance. Since the stage is designed for best operation at a particular tank impedance, there is only one correct value of loading on the tank.

Reducing Harmonics. The output pulse of a class C amplifier contains, in addition to the fundamental, numerous harmonic frequency components. As you'll learn later in this lesson, this fact enables us to operate a class C stage as a frequency multiplier. The output tank circuit offers maximum impedance at the frequency to which it is tuned. Harmonic frequencies, seeing a relatively lower impedance, are not developed across the tank circuit in any great magnitude. The circuit which couples the tank to its load is usually designed with harmonic suppression in mind. Sometimes, special traps must be used in output coupling networks which either shunt the harmonics to ground or block their passage to the antenna.

An additional precaution against harmonic radiation is to use an electrostatic shield between two inductively coupled circuits. A shield of this type, called a Faraday screen, is shown in Fig. 6.

The Faraday screen consists of a number of wires fastened together at one end and open at the other. The ends of the wires that are connected together are grounded. Capacitively coupled harmonic currents will flow to the screen wires rather than to the pickup coil. At the

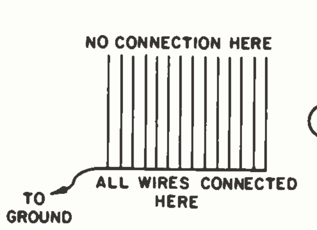


Fig. 6. An electrostatic shield or Faraday screen between the output tank and the antenna pickup coil is used to prevent harmonic currents from flowing through the capacity between the coils.

same time, because the wires do not form closed circuits, there can be no voltage induced in them by the magnetic field. Therefore, they do not interfere with the inductive coupling between the output tank and the link coil. This method is very effective in reducing the transmission of harmonics from an output tank circuit to an antenna or transmission line.

COUPLING METHODS

The resonant tank in the output of a class C amplifier forms the basis of the coupling circuit to the amplifier's load. We know that the impedance presented to the stage by the output tank circuit depends, to a large measure, on the equivalent resistance in the tank. We also know that the value of this equivalent resistance is primarily that reflected into the tank by the load. To obtain the correct tank impedance for the amplifier, the coupling must reflect a certain value of resistance into the tank. In most cases, the actual value of the load resistance connected directly across the tank would not reflect the correct resistance into the tank. Therefore, the coupling method must give an impedance transformation. The simplest way to accomplish this impedance transformation is to use a transformer as a method of inductive coupling.

Inductive Coupling. Fig. 7 shows two amplifier stages inductively coupled together. In this circuit, the resistance reflected into the output tank for amplifier 1 is adjusted by varying the spacing between the coils. Varying this spacing also adjusts the drive to amplifier 2. The circuit is designed to reflect the correct resistance and provide the proper drive at the same setting.

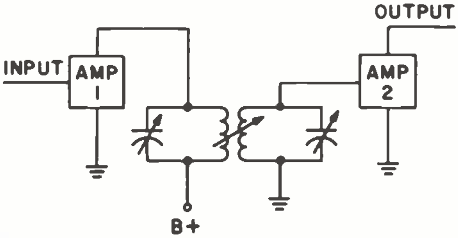


Fig. 7. Inductive coupled amplifiers.

A variation of inductive coupling is shown in Fig. 8. This method is called link coupling. It consists of a coil with only a few turns of wire inductively coupled to an output tank. A similar coil is inductively coupled to the load. The connection between the two coils is usually by means of shielded coaxial cable, so it may run some distance with very little loss. Link coupling may also be used between the final power amplifier in a transmitter and a low impedance transmission line. As in the conventional inductive coupling already discussed, the coupling is adjusted by varying the spacing between one of the link coils and the tank. Sometimes the link itself is tuned by a variable reactance. When this is done, the tuned link provides additional suppression of harmonics generated in the previous stages.

Notice that the method of applying B+ to the stage in Fig. 8 differs from that of Fig. 7. In Fig. 8, this voltage is applied through a radio frequency choke (rfc). The rfc offers a very high impedance at the operating frequency, so it keeps the signal voltage out of the power supply. When the power supply, the tank circuit, and the stage are connected in series, as in Fig. 7, the amplifier is said to be series-fed. When the power supply, tank circuit, and stage are in parallel (or shunt), as in Fig. 8, the amplifier is said to be shunt-fed.

Tapped Tank Circuits. Another

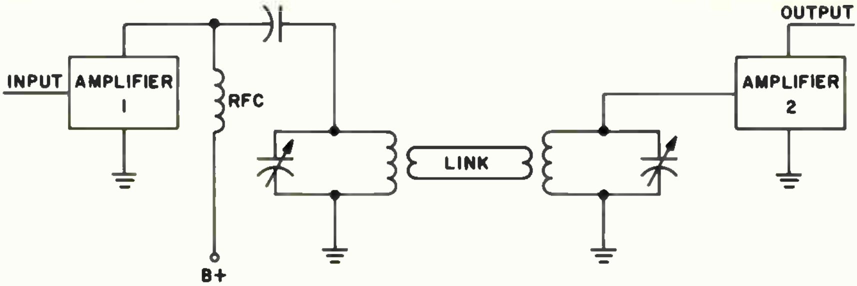


Fig. 8. Link coupled amplifiers.

method of obtaining an impedance transformation is to connect the load across only a portion of the tank coil. Such a method is shown in Fig. 9A. The value of resistance reflected into the tank is dependent on the position of the tap. In Fig. 9B, the tank capacitor is split to provide the impedance transformation. The values of C_1 and C_2 determine the value of load resistance seen by the tank (reflected resistance).

The methods shown in Figs. 9A and 9B may be combined as shown in Fig. 9C. With this circuit arrangement, the internal impedance of the stage, as seen by the tank, is transformed to a higher value by the tapped coil. Using this method the required value of loaded Q in the tank may be maintained in spite of low values of internal impedance (such as found in transistor stages). The values of C_1 and C_2, as before, determine the value of load resistance seen by the tank.

Network Coupling. Fig. 10 shows three types of networks frequently used to couple class C amplifiers to their loads. The various arms of each are shown as impedances in the figure. In practical networks of this type these impedances will be combinations of L and C components. Later on in your course you'll learn to calculate the reactance values for the arms of these networks to give a required impedance transformation. For now, it is enough for you to know

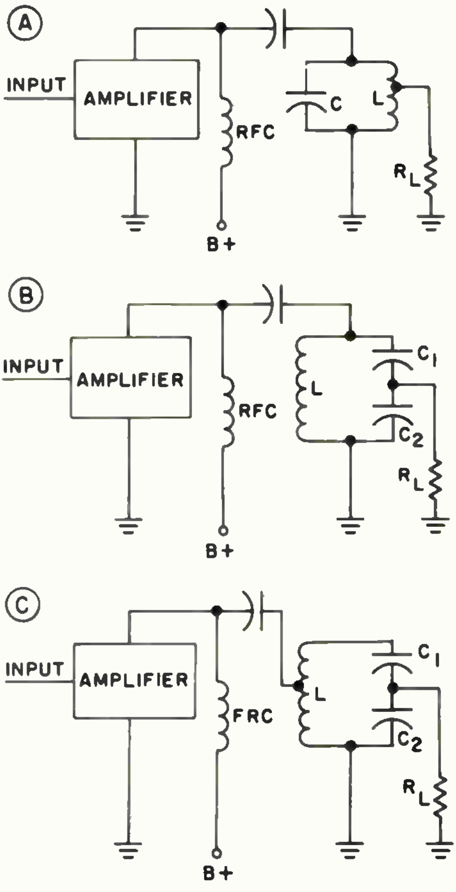


Fig. 9. Tapped tank circuits used for impedance transformation.

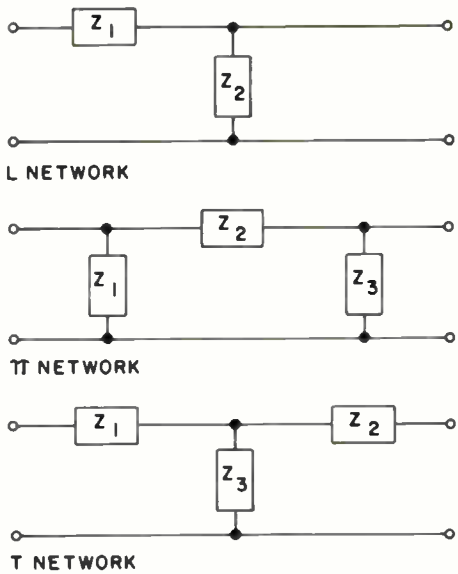


Fig. 10. Networks used to couple an amplifier to its load.

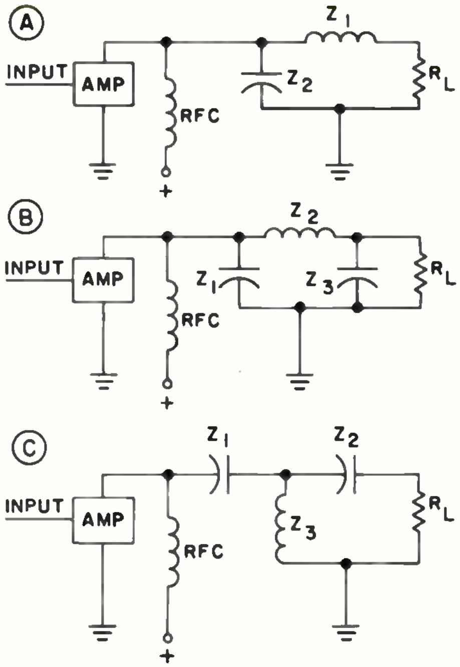


Fig. 11. Amplifiers using L, π, and T network coupling to load.

that they fulfill all the requirements of a tank circuit and can provide impedance transformations over a very wide range of values. In addition, these circuits can be easily designed to attenuate undesired harmonic frequencies.

Fig. 11 shows an example of how each of these networks (L, π, T) is used to couple an amplifier to its load. The network used in any particular application depends on the magnitude of the impedance levels to be transformed. The networks themselves may be coupled together in a variety of combinations to provide the proper load to the amplifier and greater harmonic attenuation. It is important to remember in the examples of this section that R_L may represent the input of another amplifier, a transmission line, or an antenna. The only difference between these three types of loads, as far as the amplifier is concerned, is the impedance level. An antenna or transmission line usually offers a lower impedance, and therefore a greater load, to an amplifier than the input circuit of another amplifier.

SELF-TEST QUESTIONS

(a) What is the primary reason class C amplifiers operate at higher efficiencies than class A or class B amplifiers?

(b) What two factors affect the operating angle of a class C amplifier?

(c) Is impedance matching between an amplifier and its load always desirable? Why?

(d) Suppose we wanted to increase the impedance of a parallel-tuned tank without changing the coupling or the resonant frequency. What components in the tank should be changed and in what direction?

(e) If the coupling to a tank is adjusted to increase the load on the tank, what would happen to the loaded Q? The unloaded Q? The efficiency of the tank?

(f) If the loaded Q of a tank decreases, what would happen to the tank impedance?

(g) How does the output tank circuit reduce the harmonics present in the output of a class C amplifier?

(h) How would you increase the coupling between two inductively coupled amplifiers?

(i) Is amplifier I (shown in Fig. 6) series-fed or shunt-fed?

(j) In a transistor rf power amplifier, the internal resistance of the stage is found to load the output tank so heavily that a high enough loaded Q cannot be obtained. If the amplifier is connected to the tank as shown in Fig. 6, what change could be made in the circuit to increase the loaded Q of the tank?

(k) Normally, which would more heavily load a class C amplifier: an antenna or another class C amplifier?

Vacuum Tube RF Power Amplifiers

Now that you have a basic understanding of rf power amplifiers, let's take a detailed look at some applications which use vacuum tubes. In the first section we'll discuss methods of obtaining the class C bias necessary to get the proper conduction angle from the amplifier. Then we'll look at some of the methods employed to insure stable amplifier operation. Finally, we'll take a look at some practical rf amplifier circuits, including frequency multipliers.

BIAS METHODS

Some typical class C bias methods are shown in Fig. 12. The three most common biasing methods are shown at A, B, and C.

External Bias. In Fig. 12A, the bias is obtained from an external bias supply and is coupled through an isolating rf choke to the grid of the stage. The rf choke acts as a high impedance and prevents the power supply circuit from acting as a shunt for the radio-frequency energy.

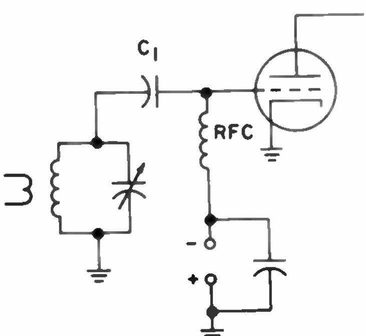


Fig. 12A. External bias.

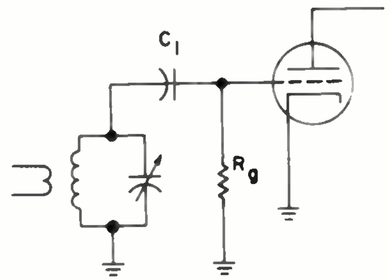


Fig. 12B. Grid-leak bias.

Grid-Leak Bias. In Fig. 12B, grid-leak bias is used. With this method of biasing, the grid current that flows during the crest of the positive half of the cycle of the incoming signal charges capacitor C_1 to a high negative value. During the interval between grid current pulses, the capacitor discharges through grid-leak resistor R_g. The discharge current develops a steady negative voltage across R_g. The value of this voltage depends upon the value of R_g and on the current through it. One advantage of this circuit is that the bias adjusts itself when the driving power is changed. Increasing the driving power increases the grid current and therefore increases the voltage drop across R_g. Thus, with grid-leak bias, changing the driving power does not appreciably change the operating angle.

A disadvantage of using grid-leak bias alone is that if there is a failure in the preceding stage, so that no excitation is supplied to the grid of the amplifier, there will be no bias developed. Excessive plate current will then flow, and if the circuit is not suitably protected, the tube and its associated components will be damaged.

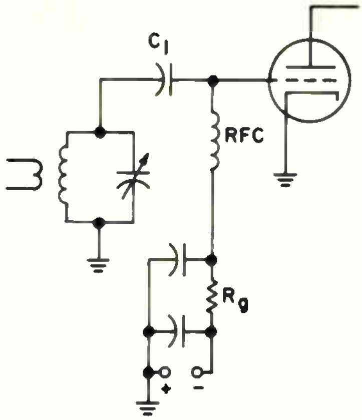


Fig. 12C. Combination external and grid-leak bias.

Combination External and Grid-Leak Bias. To protect the tube against loss of bias, a combination of external and grid-leak bias, shown in Fig. 12C, is often used. This circuit has the self-adjusting features of the circuit in Fig. 12B, and at the same time provides enough bias to protect the tube if there is a failure in the preceding stage.

Cathode Bias. The tube can also be protected against the loss of excitation by using the cathode bias combination, shown in Fig. 12D. The amount of protective bias, in either Fig. 12C or Fig. 12D, is chosen so that the plate current through the tube multiplied by the plate voltage is equal to or less than the

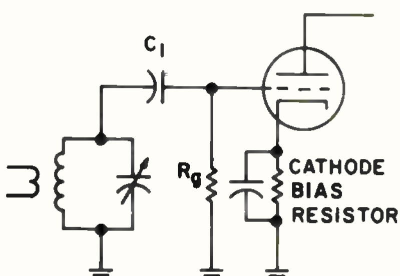


Fig. 12D. Cathode bias.

maximum safe plate dissipation of the tube. The disadvantage of cathode bias is that part of the power supplied to the plate circuit of the tube is wasted in the cathode resistor. In large high-power stages this might be a substanial amount.

Variations. The circuits in Figs. 12A through 12D show the basic class C bias methods. There are also some minor variations of these circuits.

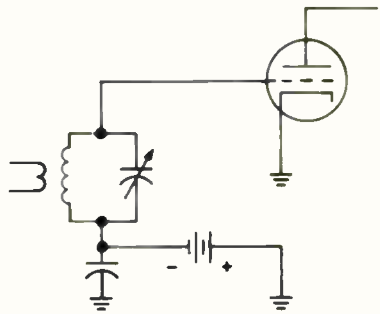


Fig. 12E. Variation of circuit shown in Fig. 12A.

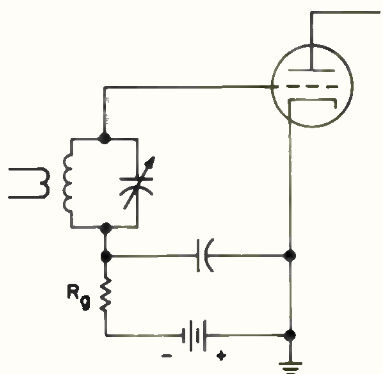


Fig. 12F. Variation of circuit shown in Fig. 12C.

For example, the circuit shown in Fig. 12A may be rearranged as in Fig. 12E, eliminating the rf choke and coupling capacitor. The circuit in Fig. 12C may be rearranged as in Fig. 12F. Perhaps we should remind you that you will find minor variations in many circuits.

Do not conclude that a circuit is necessarily different from the basic circuit you have studied just because it has been changed slightly. Study the circuit carefully and you will probably find that the method of operation is basically the same.

AMPLIFIER STABILITY

Class C amplifiers using triode tubes will go into self-oscillation easily because of feedback between the input and output circuits. As you will learn in a later lesson, a tuned-grid, tuned-plate oscillator is simply an unstable class C amplifier.

The feedback path in a triode is through the grid-to-plate capacity. Since this capacity is quite large in a triode tube, enough energy from the plate circuit can be fed back to the grid to overcome the grid-circuit losses and cause the stage to oscillate at a frequency near the resonant frequency of the tuned circuits.

Oscillation will not take place if the plate tank circuit is tuned precisely to resonance; the tank circuit must be slightly detuned to sustain oscillation. Precise adjustment, however, is very difficult. Even if you were able to make such an adjustment, the amplifier would be unstable. Slight changes in supply voltages and load would cause it to go into oscillation.

The feedback signal in a triode amplifier must be neutralized to prevent oscillation. The stage is neutralized by feeding a second signal back into the grid circuit. This second signal must be of opposite polarity and of the same amplitude as the signal fed into the grid circuit through the grid-plate capacity of the tube in order to cancel the feedback.

The most basic method of neutralization is shown in Fig. 13. In this circuit, a coil is inserted between the grid and the plate. In series with the coil is a blocking capacitor, which keeps the dc plate voltage off the grid of the tube. It has no other effect on the neutralizing circuit. The value of the coil is chosen to resonate with the grid-to-plate capacity at the frequency to which the amplifier is tuned.

The current through the coil lags the voltage 90°; the current through the capacity leads the voltage 90°. Therefore, the currents through the coil and the capacity are 180° out-of-phase and cancel each other. The disadvantage of this simple and basic method of neutralization is that it must be retuned when the operating frequency is changed. Let's look at other methods of neutralization.

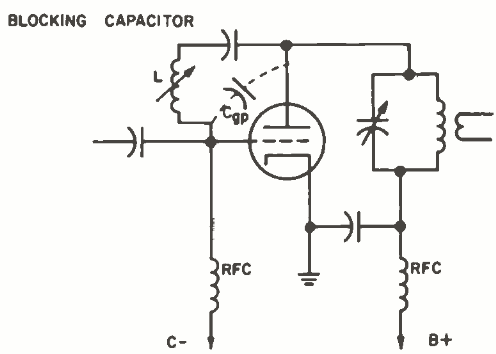


Fig. 13. Neutralization for a triode tube stage.

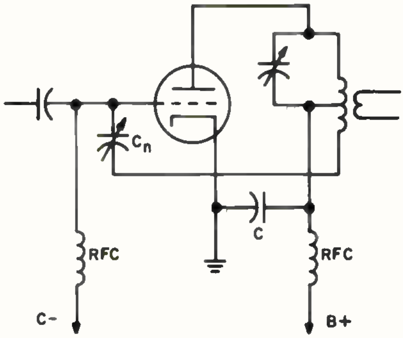


Fig. 14. Plate or "Hazeltine" neutralization.

Plate Neutralization. Fig. 14 shows a plate or "Hazeltine" neutralization. In this circuit arrangement, the coil used in the plate tank circuit is tapped, and the tap on the coil is operated at rf ground potential by grounding it through the capacitor C.

A signal voltage is developed between ground and the bottom end of the coil that is out-of-phase with the voltage at the plate end of the coil. By connecting the bottom of the tank circuit to the grid of the amplifier through the capacitor C_n, which is called the neutralizing capacitor, we can get a signal at the grid that will cancel the feedback from plate to grid through the tube capacity. Capacitor C_n is adjustable so that we can apply the exact amount of signal needed to cancel out the signal fed back through the tube.

The plate neutralization system can be considered as a balanced bridge circuit. A bridge circuit is shown in Fig. 15A. The input voltage is applied between terminals A and B and the output voltage is taken off between C and D. If the ratio of impedance Z_1 to impedance Z_2 is equal to the ratio of Z_3 to Z_4, the voltage between terminals C and D will be zero, and we say the bridge is balanced.

The plate neutralization system in Fig. 14 can be redrawn as a bridge circuit as shown in Fig. 15B. The voltage is applied to terminal A from the plate of the tube. The voltage applied to terminal B is the voltage induced in the lower half of the coil in the plate circuit of the tube. We have labeled this half of the coil L_2, and the upper half L_1. Terminal C of the bridge is connected to the grid of the tube and terminal D is grounded. L_1 is made equal to L_2 by center-tapping the coil. When C_n is adjusted to equal C_{gp}, the ratio of L_1 to L_2 will be equal to the ratio of C_{gp} to C_n. Then the bridge will be balanced, so there will be no voltage fed back to the grid circuit from the output circuit.

With this type of circuit, once the stage is neutralized it will remain neutralized over a reasonably wide frequency range, if the coil is exactly center-tapped, so that L_1 is exactly equal to L_2. If L_1 is not exactly equal to L_2, the stage can still be neutralized simply by making the ratio

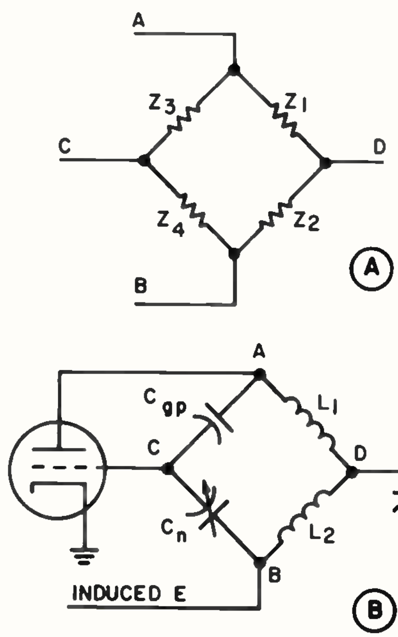


Fig. 15. Equivalent bridge arrangement of plate neutralization circuit.

of the impedance of C_{gp} to the impedance of C_n equal to the ratio of the impedance of L_1 to the impedance of L_2. If there is an appreciable difference between the values of L_1 and L_2, the frequency range over which the stage will remain neutralized becomes limited.

Another circuit for plate neutralization is shown in Fig. 16. In this circuit the center tap on the coil is not operated at rf ground potential; it is connected to B+ through an rf choke. The ground point is taken at the rotors of a split-stator tuning capacitor. A split-stator capacitor is a variable capacitor with one set of rotor plates and two sets of stator plates that are insulated from each other. The voltages at the two ends of the coil are equal and of opposite phase. The circuit in Fig. 16 can be shown as a balanced bridge by substituting the sections of the split-stator tuning capacitor for coils L_1 and

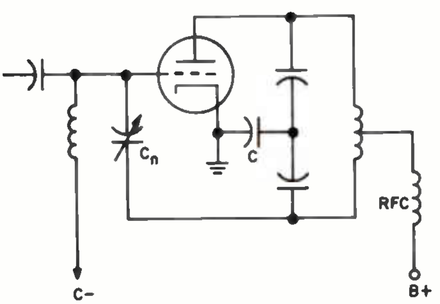


Fig. 16. Another plate neutralization circuit.

L_2 in Fig. 15B. The rf ground in both circuits is made through a capacitor (marked C in Fig. 16).

Grid Neutralization. The tapped-grid circuit arrangement shown in Fig. 17 can also be used for neutralization. This is referred to as grid neutralization or Rice neutralization. The neutralizing signal is taken from the plate and applied to one end of a tapped coil in the grid-tuned

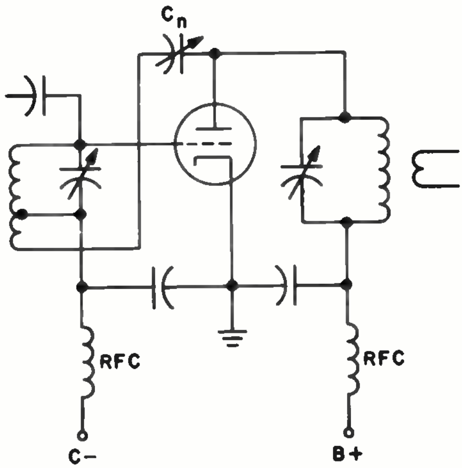


Fig. 17. Grid or "Rice" neutralization.

circuit. The rf ground connection is made to the center tap on the grid coil. The polarity of the feedback signal, introduced through the neutralizing capacitor C_n to one end of the grid coil, is in phase with the signal that is fed directly to the other end of the grid coil through the plate-grid capacity.

By properly adjusting the neutralizing capacitor C_n, voltage fed through it can be made equal to the voltage fed through the tube capacity. These two voltages will cause equal currents to flow through the

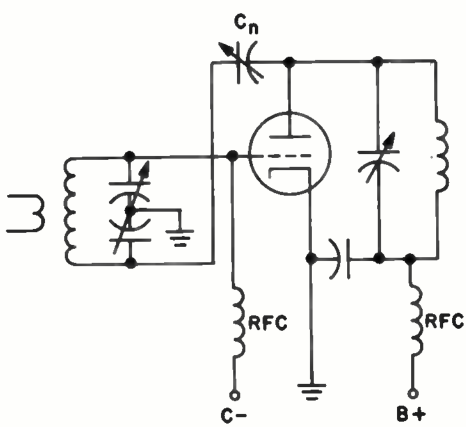


Fig. 18. Split-stator grid neutralization.

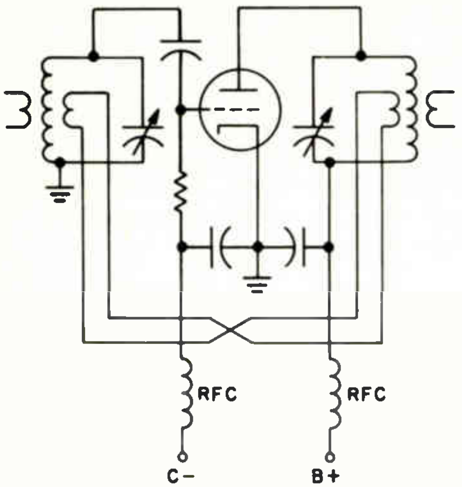


Fig. 19. Inductive neutralization.

grid coil in opposite directions. These currents will induce new voltages in the grid coil which will tend to cancel the voltage fed through C_n and C_{gp}. A split-stator version of grid neutralization is shown in Fig. 18. Its operation is essentially the same as that of Fig. 16.

Inductive Neutralization. Still another method of neutralization is shown in Fig. 19. This is referred to as inductive neutralization, because the neutralizing signal is obtained by inductive coupling between the plate and grid-tuned circuits. The signal induced in the grid circuit by the inductive link is opposite in polarity to the feedback signal, and gives feedback cancellation.

Parasitics. Neutralization of an amplifier is used to prevent oscillation at the frequency to which the grid and plate circuits are tuned, in other words, at the signal frequency. Some amplifiers go into oscillation at frequencies far removed from the desired signal frequency. Oscillations of this type are called "parasitic oscillations" or "parasitics."

The neutralization circuits we have just studied can do nothing to prevent this type of oscillation. Long leads, tube interelectrode capacities, rf chokes, and bypass capacitors are the major inductive and capacitive elements that cause parasitic oscillations.

Parasitics may exist at low or high frequencies, or at both low and high frequencies at once. They cause low operating efficiency and instability in the stage, erratic meter readings, radiation of improper carriers and sidebands, distortion, overheating of the amplifier tube, and premature breakdowns in the circuit parts. If grid-leak bias is used in the stage, parasitics will also cause changes in the grid bias.

Fig. 20A shows a typical class C amplifier stage. At the operating frequency, the grid and plate circuits are tuned by the coil and capacitor combinations L_1-C_1, and L_2-C_4. The stage is prevented from oscillating at the operating frequency by the signal fed back through neutralizing capacitor C_n.

Fig. 20B shows what the effective circuit would be if this stage were producing low-frequency parasitic oscillations. The grid circuit is now tuned by the parallel combination of the grid choke, RFC_1, and the grid bypass capacitor, C_2. Since these oscillations usually take place at frequencies below 200 kilohertz, coil L_1 has very little reactance and serves merely as a connecting lead from the grid to the junction of C_2 and RFC_1. This places the rf choke and grid bypass capacitor in parallel between grid and ground.

The tuned circuit in the plate at the low frequencies is the plate bypass C_3 and the plate rf choke. Here, too, the regular tank coil L_2 has practically no reactance at the oscillation frequency, and serves simply as a connecting lead. The neutralizing capacitor C_n is now

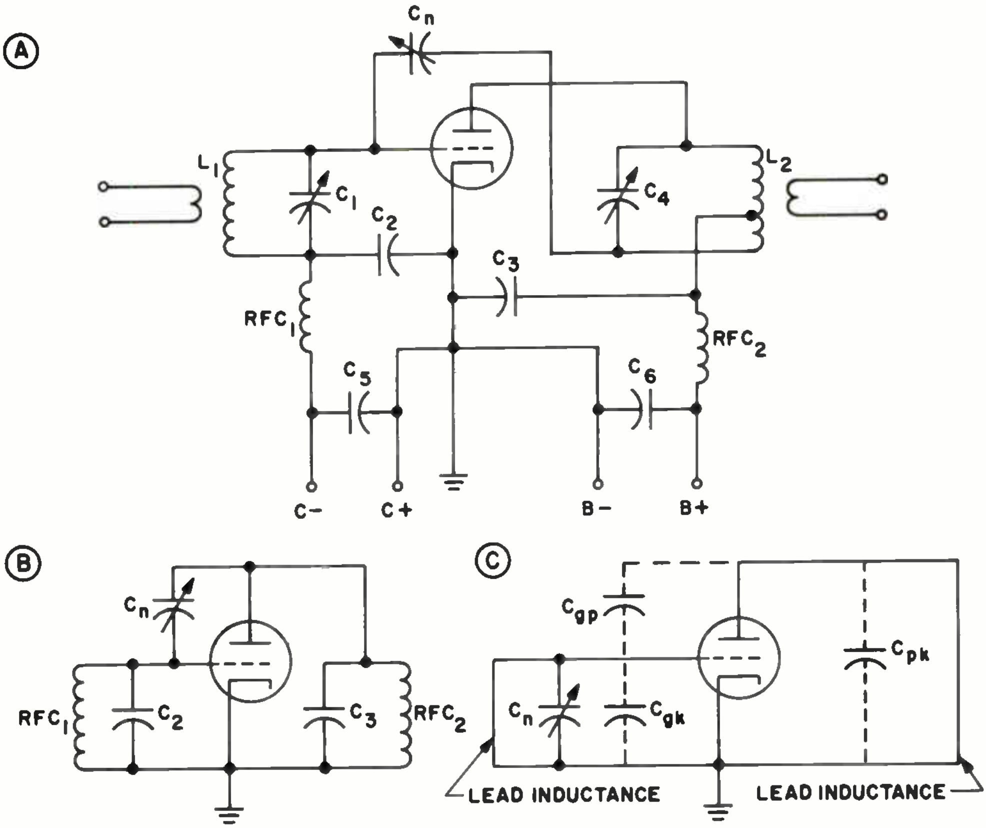


Fig. 20. A typical class C stage is shown at A; the effective circuit that produces low-frequency parasitics is shown at B; the effective circuit that produces high-frequency parasitics is shown at C.

effectively in parallel with the tube grid-plate capacity and increases rather than reduces feedback. Coils L_1 and L_2 do have a slight reactance at the parasitic frequency, so tuning capacitors C_1 and C_4 can make slight changes in the parasitic oscillation frequency.

The effective circuit for the stage, if it were producing high-frequency parasitic oscillations, is shown in Fig. 20C. In this case, the grid and plate circuits are tuned by the inductance of the leads between the tube elements and the tank circuits and the grid-to-cathode capacities. At the high frequencies, the capacities of C_1 and C_4 are so high that they act only as connecting leads in the inductive circuit. Capacitors C_2 and C_3, which are even larger in size, have practically no reactance at this frequency. The neutralizing capacitor C_n now appears between grid and ground and is effective in determining the frequency of the grid circuit. Feedback is through the capacity between grid and plate.

Preventing Parasitics. In the effective circuit of either Figs. 20B or 20C, parasitic oscillation can be prevented by making the resonant frequency of the grid circuit higher than that of the plate circuit. This may be done in Fig. 20B by making capacitor C_2 smaller than C_3 or by making R_FC_1 smaller than R_FC_2.

The most satisfactory method for suppressing very high-frequency parasitic oscillations in a class C amplifier stage is by using parasitic suppressors. The purpose of these parasitic suppressors is to increase the circuit losses at the parasitic frequency. Examples of these suppressors are shown in Fig. 21.

The suppressors are low resistances, usually around 100 ohms, in parallel with small rf chokes. At the normal operating frequency, these small coils, L_1 and L_2, have very low inductive reactances, and the signal frequency can pass through them with no loss. At the frequency of the parasitic oscillation, however, these coils have very high reactance and force the parasitic signal to flow through resistors R_1 and R_2. The loss of parasitic signal in the two resistors is great enough to prevent the tube from going into oscillation at these high frequencies.

Although a commercially manufactured transmitter should be free of parasitic oscillations, occasionally a new transmitter being tuned up for the first time will have them. Parasitics can also occur if parts are replaced by those of a different make. Whenever modifications are made in an amplifier stage, that stage should be checked for both high and low-frequency parasitics. Low-frequency parasitics will sometimes be evident as sidebands of the carrier frequency. The most common indication of high-frequency parasitics is an unusually high plate current and low output.

MULTIELEMENT TUBE STAGES

In a previous lesson you learned about the characteristics of screen grid, pentode, and beam-power tubes. Let us review briefly their characteristics with respect to their use as class C amplifiers.

If a tetrode or pentode tube is used in the stage, the screen grid of the tube acts as an electrostatic shield between the grid and plate, which reduces the grid-to-plate capacity. Therefore, tetrode and pentode tubes are less susceptible to feedback and self-oscillation, and usually do not require neutralizing.

Multielement tubes have a higher power gain than triodes. In other words, for the same amount of grid driving

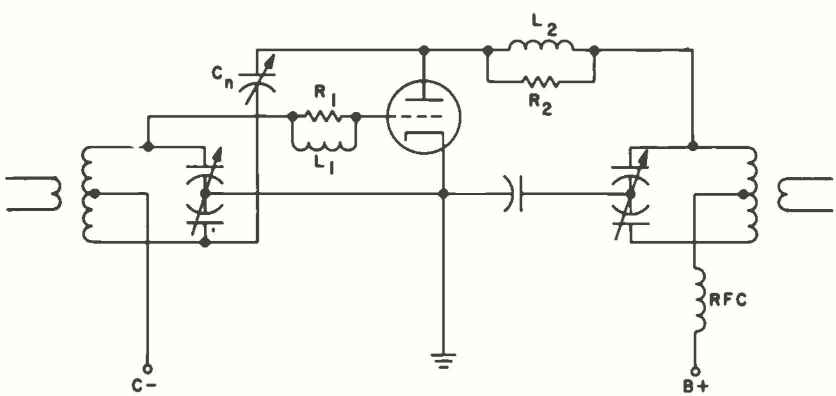


Fig. 21. Parasitic suppression methods.

power, you can get a higher power output from a stage using a multigrid tube than from one using a triode. This means that fewer stages are needed to get the desired power output. It also means that better shielding must be used between input and output circuits to prevent external feedback. Even a very small amount of feedback from the output of a stage back to the input can cause oscillation. Also because of the higher power gain, parasitic oscillations are more common in tetrode and pentode stages than in triode stages.

In screen-grid (tetrode) class C amplifiers, the minimum plate voltage must not be allowed to swing lower than the screen voltage because the screen then offers a greater attraction to the electrons than the plate. The secondary emission effect, due to electrons bouncing off the plate and being pulled to the screen instead of falling back onto the plate, determines the minimum the plate voltage can swing to. The grid excitation is adjusted so that the grid swings far enough positive so that the tube draws maximum permissible peak plate current without exceeding the dissipation rating of the plate and grid electrodes.

A pentode tube permits a greater plate voltage swing and, therefore, an even higher power gain. It does so by using a suppressor grid at cathode potential between the plate and screen grid to prevent secondary electrons from moving to the screen grid. Thus, the plate current remains independent of plate voltage to a much lower value of plate voltage. The suppressor grid forces the secondary electrons coming off the plate to return to the plate.

A beam-power tube has characteristics similar to those of a pentode. The tube elements are shaped in such a way that they control the electrons flowing between the cathode and the plate of the tube. Proper shaping of the electrodes sets up a potential barrier between screen and plate to suppress secondary emission.

You will find many multielement tubes used in transmitter equipment because of their higher power gain and simplicity of neutralization. Beam-power tetrodes are the most common.

In many circuits using multielement tubes, no neutralization is used. However, the power gain of these tubes is so great that only a small amount of feedback will set up instability and oscillations. Keeping feedback below the level that will cause instability or oscillation is a real problem. Even if a stage does not oscillate when it is first manufactured, there is no guarantee it will not be unstable when the tube in the stage is replaced. To overcome these problems, manufacturers often neutralize stages using multielement tubes.

In a class C stage using a multielement tube, the screen voltage has as much control, or more, on the plate current and power output as the actual value of the plate supply voltage. The plate supply voltage, however, must be high enough to obtain the necessary plate voltage swing across the plate-tuned circuit. Because the screen grid has so much control, the power output in some transmitters is controlled by varying the screen grid voltage.

The correct voltage must be applied to the plate of a multigrid tube at all times. If the plate voltage drops to zero or is lower than normal, the screen grid may be damaged. Under these conditions the screen current may be so high that it exceeds the safe dissipation factor.

Screen voltage and current also vary with the grid excitation, particularly if the screen voltage is obtained through a dropping resistor. An increase in grid excitation will cause the screen current to rise and the screen voltage to fall. A decrease in excitation will have the opposite effect.

When a tetrode or pentode stage is being tuned and loaded, the plate and screen voltages should be reduced. Most transmitters using tetrode or pentode power stages have provisions for reducing these voltages during tuning. A nonresonant or unloaded plate tank causes the minimum plate voltage to drop below the screen voltage. Under these conditions, the screen draws excessive current. This may destroy high-power tetrodes in a matter of a few seconds. After the tuning and loading are roughly adjusted, full voltage can be applied to the tube and the adjustments carefully peaked.

MULTITUBE STAGES

To get a higher output from a class C stage, two tubes can be connected in parallel, or in push-pull. For very high power, tubes are operated in push-pull-parallel; that is, two sets of parallel-connected tubes are operated in push-pull.

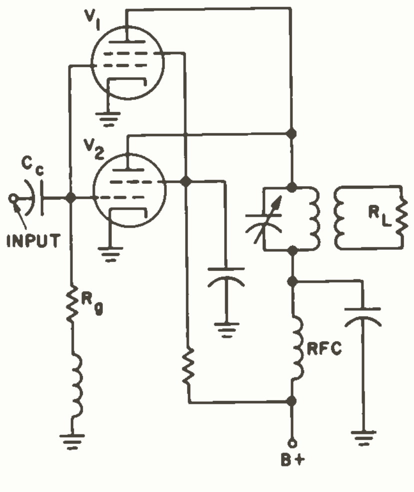


Fig. 22. Class C amplifier with two tetrodes connected in parallel.

Parallel Operation. Fig. 22 shows a stage with two tubes connected in parallel. In parallel operation, the output power is approximately twice that from a single tube, if the correct driving power is applied and the circuit components and electrode voltages have the correct values.

The grid current is doubled, because with two tubes the grid impedance is approximately halved. The driving power needed for the parallel amplifier is twice that needed for a single tube.

When grid-leak bias is used with the class C stage, the value of the grid resistor must be cut in half to get the same grid bias at twice the grid current.

The internal or plate resistance is also halved because of the parallel connection and doubling of the peak plate current. Thus, the same tuned circuit voltage is developed with twice the plate current. It is the higher amplitude plate current pulses exciting the tuned circuit that develop the added power delivered to the load in parallel operation.

Push-Pull Operation. A push-pull amplifier is shown in Fig. 23. The input excitation is applied with equal amplitude but opposite polarity to the grids of the push-pull stage. The ground point of the circuit is at the center of a split-stator variable capacitor.

As in the case of parallel tube operation, the grid and plate currents drawn are twice as great as those drawn by a single tube. To retain balanced operation of the push-pull stage where each tube performs an equal share of the work, the supply voltages are applied at the mid-

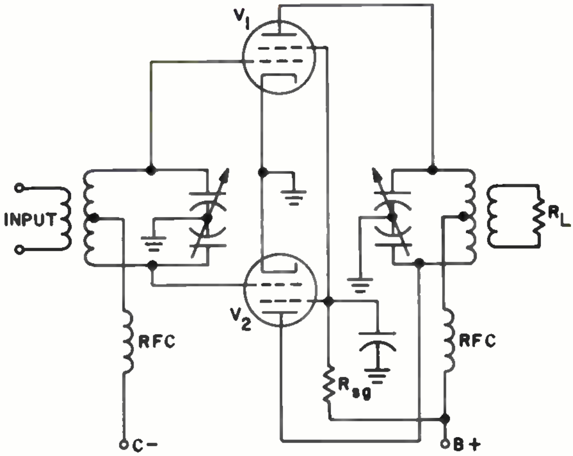


Fig. 23. Push-pull class C amplifier using tetrodes.

points of the coils so as not to imbalance the stages.

Balanced operation is necessary to prevent possible overloading of one of the tubes because of any uneven dissipation of power by the grid or plate. Imbalance can be caused by tubes that are not exactly matched or by a mismatch between the grid or plate circuit components. Thus, both the circuits and the tubes must be matched and balanced to get proper operation of the stage.

FREQUENCY MULTIPLIERS

A frequency multiplier stage is a class C amplifier that is used to generate an output signal whose frequency is some multiple of the applied signal frequency. For example, the frequency of the output of a doubler stage is twice the frequency of the input. The frequency of the output of a tripler is three times the frequency of the input. A doubler with an input frequency of 10 MHz would have an output signal of 20 MHz. A tripler with an input frequency of 10 MHz would have an output signal of 30 MHz. Multipliers can be used to generate signals of even higher multiples of their input signal frequency, but the higher the multiplication the lower the output. Thus, you can expect less output from a tripler than from a doubler using the same tube type. A multiplier generating a signal four times the frequency of the input signal would have an even lower output than a tripler using the same tube.

When a tank circuit is shock-excited into oscillation by a single current pulse, the circuit will continue to oscillate for a number of cycles. The number of cycles will depend on the losses in the circuit. Each cycle will be lower in amplitude than the preceding one because of these losses. With a high Q circuit, when the losses are low, there will be many cycles before the oscillation drops to zero.

In a frequency multiplier we take advantage of the fact that oscillation, once started, will continue for many cycles in a tank circuit. By using a tank circuit in the plate circuit of the tube that is resonant at some multiple of the input frequency, we can start the oscillation by feeding an rf signal to the grid. This will produce a plate current pulse that starts the tank circuit oscillating at its resonant frequency which may be two or three times the frequency of the input signal. This oscillation would soon die out, except on the second cycle, in the case of a doubler, or the third cycle in the case of a tripler, where the grid of the tube will be driven positive again by the rf signal. This produces another plate current pulse which adds to the energy in the tank circuit and supplies the power needed to make up the circuit losses, so the oscillation in the plate circuit continues. Now let's look at some typical frequency multiplier circuits.

Single-Tube Multipliers. The circuit of a frequency multiplier is very simple; one is shown in Fig. 24. It is even simpler than a regular class C amplifier. No neutralization is needed, even when the tube used is a triode, because self-oscillation occurs only if the input and output circuits are tuned close to the same frequency. In a doubler the output tank circuit is tuned to twice the frequency of the input circuit, in a tripler it is tuned to three times the frequency, etc.

The current waveforms in the tank circuit of a class C amplifier are compared with those in a frequency multiplier in Fig. 25. Fig. 25A shows the waveforms for a fundamental class C amplifier. The plate current pulse flows during part of each cycle of the incoming signal. The flywheel action of the tank circuit develops the fundamental sine wave output, shown by the dashed curve.

Fig. 25B shows the waveforms for a single-tube doubler circuit. The tube is

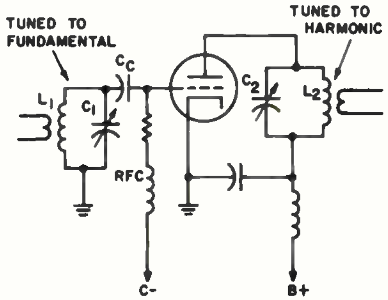


Fig. 24. Basic frequency-multiplier stage.

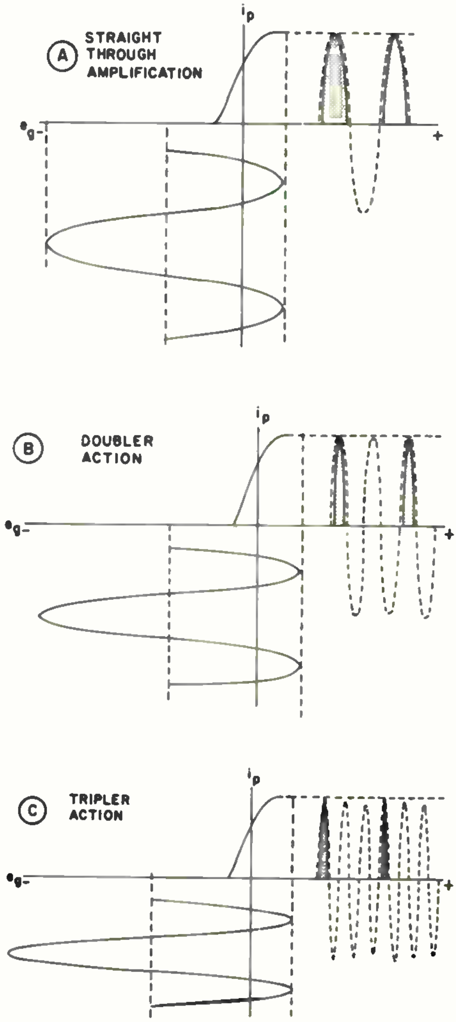
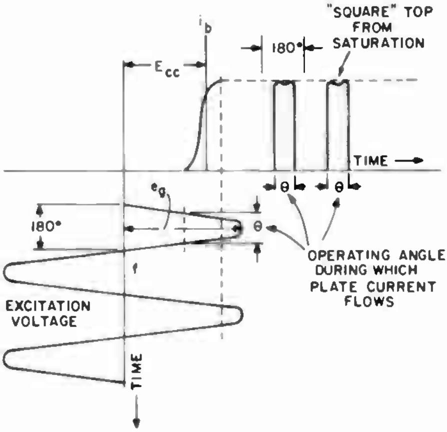


Fig. 25. Waveshapes of frequency multipliers compared to the "straight-through" amplifiers.

operated with a higher bias, so the plate current pulse flows during a smaller part of each cycle of the incoming signal, and the flywheel action of the resonant circuit carries through that cycle and another cycle before the next plate current pulse arrives. Since the plate current flows only on alternate cycles of the output, the power output and efficiency of the stage are lower than for fundamental operation. The efficiency is usually less than 50%.

The higher the harmonic signal to which the tank circuit is tuned, the lower the obtainable power output and efficiency of the class C stage. Fig. 25C shows the waveforms for a tripler. The tube is biased so that the plate-current flows a still smaller part of the cycle of the incoming signal, and the resonant circuit carries through three cycles before the next pulse arrives. Losses in the circuit cause the amplitude of each succeeding cycle to decrease. The efficiency of a tripler stage is even less than that of a doubler. The efficiency of a multiplier is kept as high as possible by using the proper values of L and C in the tank circuit and correctly shaping the plate current pulse.

The best pulse shape is a square top pulse like the ones shown in Fig. 26. This pulse shape can be obtained by operating the stage with a high bias and then driving the stage to plate-current saturation. For best doubler operation, the angle of current flow, indicated by the Greek letter θ (theta), should be somewhere between 90 and 120 degrees. With this angle of flow, the plate current pulse has a suitable and effective second harmonic content. For tripler operation, the angle of current flow is reduced to less than 90 degrees, and the third harmonic component is emphasized.



OUTPUT FREQ	OPTIMUM OPERATING ANGLE (θ)
2f	90°–120°
3f	80°–100°
4f	70°–90°
5f	60°–70°

Fig. 26. Multiplier operating characteristics.

The plate tank circuit of frequency multipliers can usually be tuned over a wide enough range to resonate at more than one harmonic of the signal at the grid. Therefore it is important that you check the output frequency to be sure you have the correct harmonic. You can do this with an absorption wavemeter. You'll learn more about this instrument later in the lesson.

Two-Tube Multipliers. A special higher powered and somewhat more efficient doubler can be obtained by using the push-push arrangement shown in Fig. 27A. In this circuit, the grids are supplied with signals in push-pull and the plates are connected in parallel.

The tubes are connected so that one will conduct on the positive alternation of the incoming signal, shown in dashed lines in Fig. 27B, and the other tube will conduct on the negative alternation.

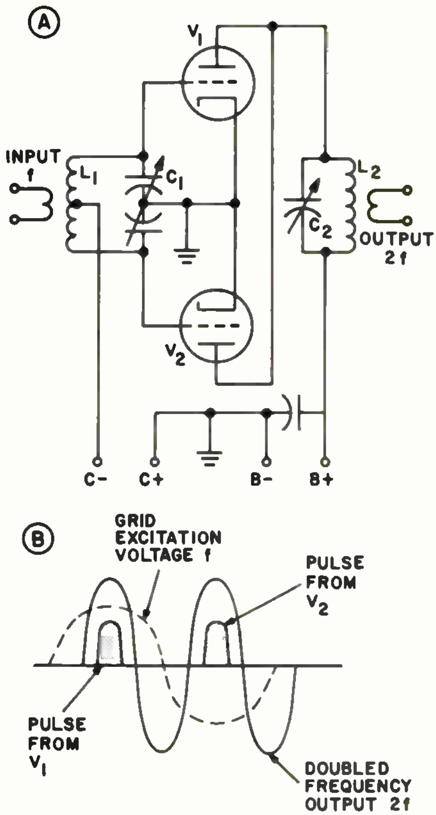


Fig. 27. Push-push doubler for even harmonics.

Thus, plate current pulses are fed to the output circuit once during each alternation of the doubled frequency. The efficiency and power output are higher than for a single-tube doubler.

The push-push frequency multiplier stage in Fig. 27A operates well on even harmonics, but not on the fundamental or odd harmonic frequencies. Frequency doublers are used more often in transmitters, especially low-frequency transmitters, than the higher harmonic generators because of the higher output and efficiency. The grids are connected in push-pull, so they must be fed with balanced signals to get the proper output signal.

Since the doubler is the most frequently used type of frequency multiplier stage, let us list some of its characteristics:

1. The plate tank circuit is tuned to twice the grid circuit frequency.
2. It does not have to be neutralized.
3. The operating angle of the plate current pulse is approximately 90°.
4. The dc bias is about 10 times the plate current cutoff value.
5. The plate current pulse has a greater harmonic content.
6. It requires a very large grid-driving signal.
7. It has a low plate efficiency compared to a fundamental class C amplifier.

As you can see, some of these characteristics vary widely from those of a class C amplifier operating as a fundamental frequency amplifier.

SELF-TEST QUESTIONS

(l) What is the disadvantage of using only grid-leak bias in a class C amplifier?

(m) What is the feedback path for oscillations near the operating frequency in a triode class C amplifier?

(n) What is the main advantage of plate neutralization over the method of connecting a coil and blocking capacitor from plate to grid?

(o) In Fig. 16, what is the phase relationship between the signal fed through C_N and the signal fed back through grid-plate capacitance?

(p) What inductive components in the grid circuit form part of the low-frequency parasitic resonant circuit?

(q) What might be the cause of unusually high plate current and low rf output from a transmitter?

(r) What two characteristics of tetrodes make them more useful as class C amplifiers than triodes?

(s) What characteristic of tetrodes make them more susceptible to parasitic oscillation than triodes?

(t) How would the values of L and C in a tank used in a parallel-connected stage compare with those used in a similar single tube circuit operating at the same frequency?

(u) How is balanced operation obtained in a push-pull stage?

(v) In general, which multiplier would have a higher output, a doubler or a tripler?

(w) Why are neutralization circuits unnecessary in a triode operated as a frequency multiplier?

Transistor Power Amplifiers

In their present state of development, transistors cannot amplify high-frequency signals to high power levels as well as vacuum tubes. However, power outputs greater than 100 watts or frequencies much above 400 MHz are seldom required in many communications applications. Chief among these is commercial mobile radio. In this application, the transistor's small size, low operating voltage, extreme ruggedness, and high overall efficiency make it ideally suited for use in mobile radio equipment.

The common emitter circuit is almost universally used for transistor rf power amplifiers because of its greater stability at radio frequencies. This circuit arrangement is often compared with the grounded cathode triode. As you might expect, it has much in common with the triode circuits you previously studied. With transistor amplifiers we are concerned with the biasing, efficiency, and stability, just as we were with the triode. In this section, we'll look at some typical circuits which illustrate the important features of transistor rf power amplifiers.

BIASING METHODS

You have learned that class C operation of a power amplifier results in the highest percentage of input power being converted to useful rf energy at the output. In the case of a transistor, where high-frequency power handling is a limitation, we are especially interested in getting the highest efficiency obtainable from the stage. It is not surprising, then, that most transistor rf amplifiers are operated class C.

Two practical methods of obtaining bias for class C operation are shown in Fig. 28. The circuits illustrated employ NPN transistors; PNP devices could just as well have been used (with all polarities reversed, of course).

In Fig. 28A, reverse bias across the emitter base junction is developed by the R_1-C_1 network in the emitter circuit. When the input signal to the stage goes sufficiently positive, base and collector currents flow over the paths indicated by the solid lines. Both these currents flow through the emitter resistor, dropping a voltage of the polarity indicated. Capacitor C_1 charges to the peak value of this voltage drop. During the time between positive-going portions of the input signal, C_1 slowly discharges through R_1

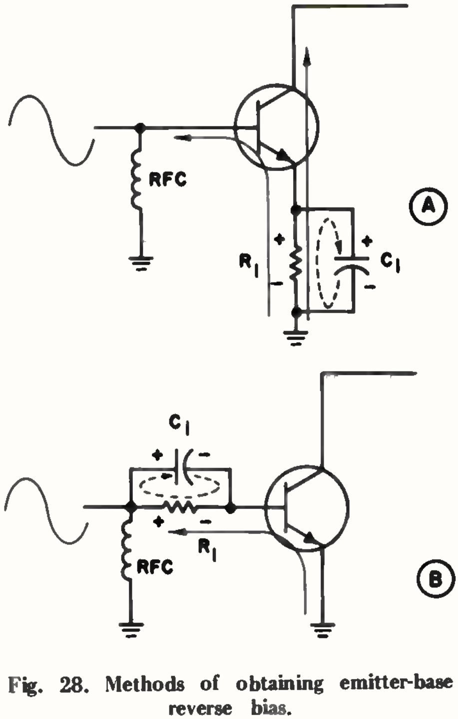


Fig. 28. Methods of obtaining emitter-base reverse bias.

as indicated by the dashed line. The values of R_1 and C_1 are such that C_1 does not discharge appreciably before the input signal again swings positive, thereby recharging C_1. The emitter is thus maintained slightly positive with respect to the base by the charge across C_1. Collector current does not flow until the input signal drives the base more positive than the emitter.

Reverse bias for the circuit of Fig. 28B is developed in the base circuit, again by a parallel combination of R_1 and C_1. The base current drawn during the positive-going portion of the input signal (indicated by the solid line) drops a voltage across R_1 as shown. C_1 charges to the peak value of this voltage drop and essentially maintains its full charge during the time between positive-going portions of the input signal. This is possible because the discharge path for C_1 (shown by the dashed line) is through R_1, the value of which is chosen to permit only a very small discharge current to flow. Notice that the polarity of the charge on C_1 is such that it subtracts from the positive-going portion of the input signal. This means that the input signal voltage must exceed the voltage across C_1 before the emitter base junction becomes forward-biased, allowing collector current to flow.

The two circuits shown in Fig. 28 depend on the presence of an input signal to develop bias. With no input signal present, zero bias is developed. Unlike vacuum tubes, however, transistors do not conduct under zero bias conditions and are therefore self-protecting.

Not only are transistors nonconducting under zero bias conditions, but also a small forward-biasing voltage must exist across the emitter base junction before collector current begins to flow. Fig. 29 shows collector current plotted against emitter-to-base voltage for a typical rf power transistor.

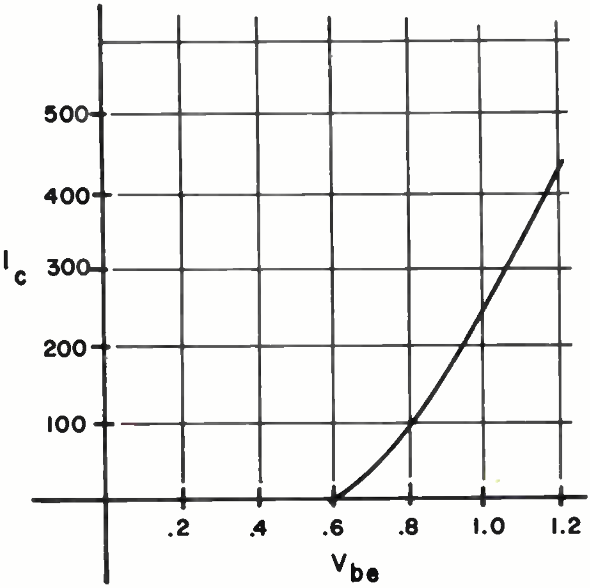


Fig. 29. Collector current plotted against emitter-to-base voltage of a typical RF power transistor.

As you can see from the graph, collector current does not begin to flow until the emitter-base voltage reaches approximately .6 volts. When we operate a power transistor with zero bias, then, we are actually biased about .6 volts below collector current cutoff.

Fig. 30 shows the graph of Fig. 29 with an input signal of 1 volt peak amplitude applied. Collector current flows only during the time the input signal is above .6 volts. The conduction angle here would be about 120° of the input cycle, well within the class C operating range. Input signal levels of such low amplitude are not unusual in power transistors because of the transistor's low input impedance. Input impedances actually range from several ohms to less than 1 ohm. With such a low input impedance, a relatively large input current is permitted to flow when the base-emitter junction of the transistor is driven positive.

Recalling that P = EI, you can readily see that the driving power to the stage is accounted for primarily by the high current flow which develops only a small voltage across the low input impedance. It follows, then, that a reverse-biased emitter-base junction is not always necessary for class C operation.

MULTITRANSISTOR AMPLIFIERS

As mentioned earlier, the power obtainable from a transistor used as an rf amplifier is rather limited as compared to vacuum tubes used in similar circuits. When more power is required than can be obtained from a single transistor, several transistors can be arranged in push-pull or parallel. In a push-pull arrangement, an input transformer is required to feed the transistors out-of-phase signals. This transformer is also required to match the relatively high output impedance of the

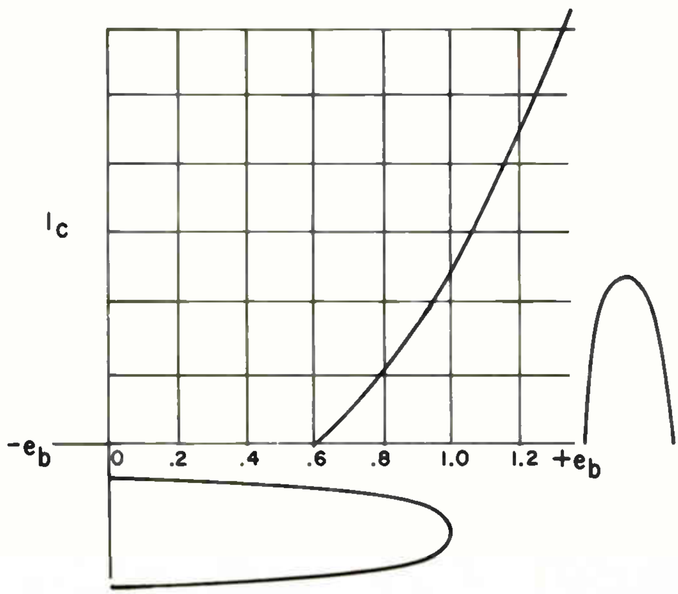


Fig. 30. Curve of Fig. B with signal applied showing collector current pulse.

driver to the very low input impedance of the push-pull stage. Such a transformer capable of operation at high frequencies is very expensive to build. For this reason, multiple transistor stages are nearly always parallel-connected.

Fig. 31 shows a two-transistor parallel-connected rf amplifier stage. C_1 and L_1 form an L network which provides the proper impedance match between the source and the input to the stage. The input signal is developed across RFC_1, amplified by the transistors, and applied to the load through the output coupling network. Notice that we can vary the operating bias of the two transistors by adjusting R_1 and R_2. These adjustments are necessary so that the two transistors will share the load equally. In practice we would insert a milliammeter in the collector or emitter circuit of each transistor and adjust R_1 and R_2 for equal currents. With the currents balanced, each transistor will be handling half of the power delivered to the output coupling network. C_4 is a coupling capacitor and may be considered a short circuit at the operating frequency. The output coupling circuit is a pi network consisting of C_5, L_2, and C_6. C_5 and C_6 are adjusted to provide the proper collector load and circuit Q for the transistors.

Another circuit employing transistors in parallel is shown in Fig. 32. Besides containing three transistors instead of two, this circuit differs from the one in Fig. 31 in two other important respects. First of all, load balancing is obtained by adjusting L_1, L_2, and L_3 in the base circuits. These adjustments vary the rf

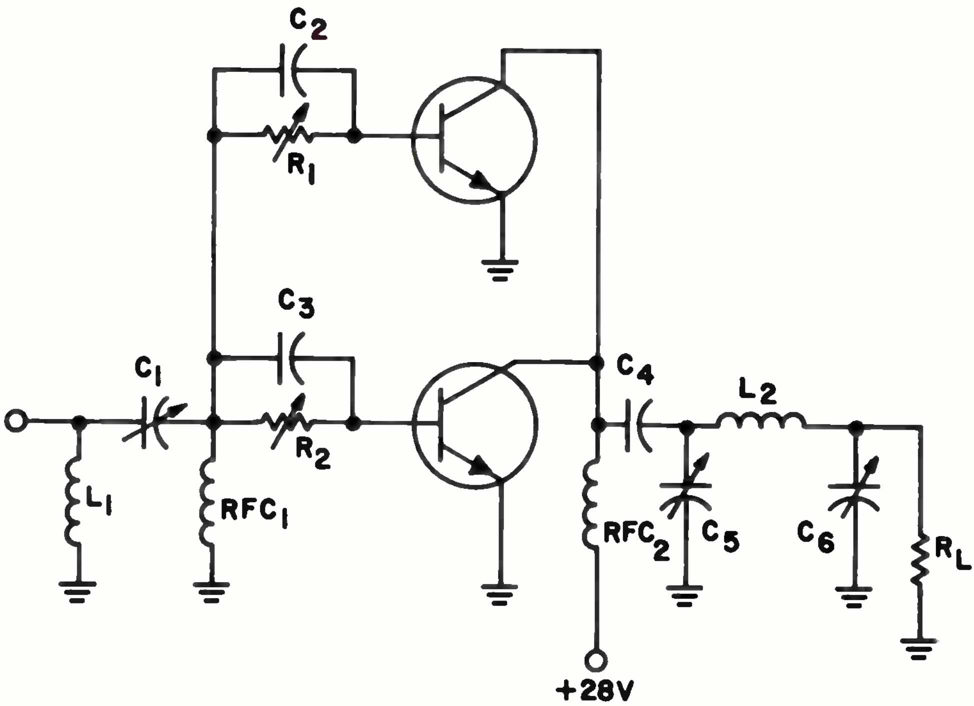


Fig. 31. Two-transistor parallel-connected RF amplifier.

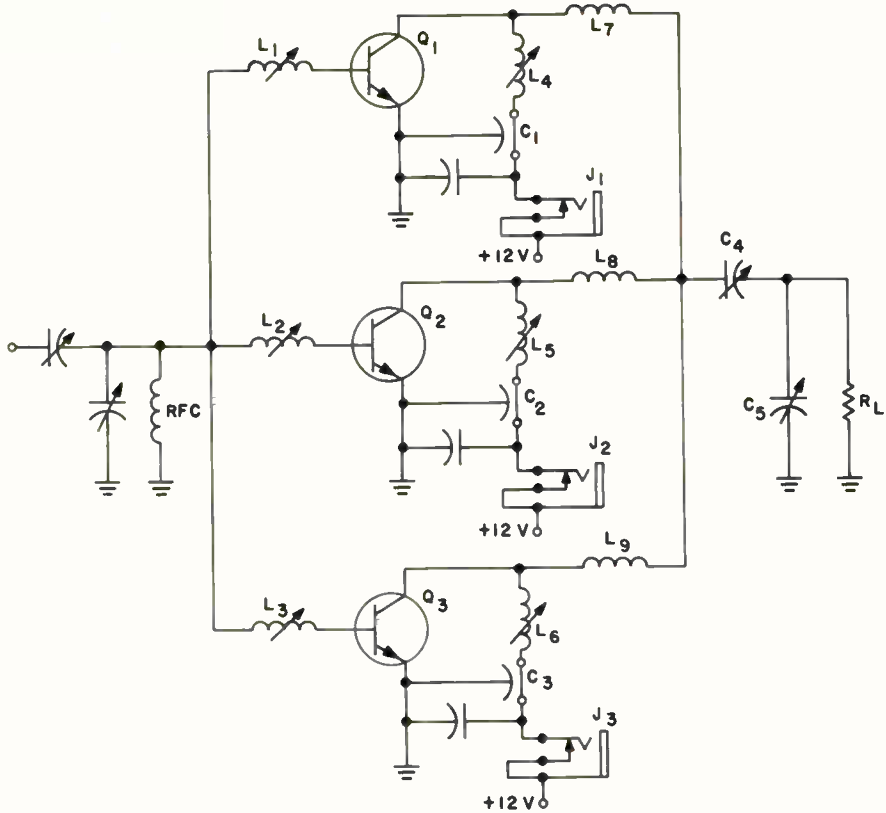


Fig. 32. Three-transistor parallel-connected RF amplifier.

drive to the transistors and equalize the load currents as previously discussed. Secondly, each of the three transistors in Fig. 32 has a separate output tank circuit connected to a common load as in Fig. 31.

The tank circuit for Q_1 is made up of L_4, L_7 and stray capacity. L_5 and L_8 are the tank inductances for Q_2; L_6 and L_9 are the tank inductances for Q_3. These tank circuits are also tuned by stray capacity. The right-hand ends of L_7, L_8 and L_9 connect to C_4 which, along with C_5, varies the coupling from the three tank circuits to the load R_L.

There are two reasons why separate tank circuits are used for the three transistors. First, each transistor is series-fed, thus eliminating the losses and other problems of an rf choke. Second, the dc collector currents are entirely separate so each transistor can operate essentially independent of the other transistors. Thus, if some trouble developed in Q_1, this stage could be disconnected and the amplifier could continue to operate at a lower power level. Directly paralleled stages would have to be completely retuned if one stage were to be removed.

Many rf power transistors have their emitters internally connected to the transistor case. This is done to eliminate the stray inductance of the emitter lead. When the case is connected to ground, as it would be in a circuit such as that shown in Fig. 32, current could not be conveniently measured in the emitter circuit. The use of separate collector loads, however, enables convenient monitoring of collector current. In Fig. 32, the jacks labeled J_1 through J_3 are provided for this purpose.

The symbol used to represent C_1 through C_3 may be unfamiliar to you. This is a special type of capacitor called a feed-thru and is often used for bypassing in high-frequency circuits. As the symbol suggests, one plate of the capacitor completes a dc path in the circuit; the other plate, usually connected to ground, surrounds the first much like the braided shield in a piece of coaxial cable.

FREQUENCY MULTIPLIERS

Fig. 33 shows two transistor frequency multiplier stages coupled together. The output of the tripler, Q_1, feeds a doubler, Q_2, to give a total frequency multiplication of six. The input signal, which we've designated as F_O, is applied to the base of Q_1. With the drive signal present, R_1 and C_1 develop a relatively high reverse bias across the emitter base junction. The high reverse bias results in a narrow conduction angle and collector current pulses with a high harmonic content. The collector tank is tuned to the third harmonic, $3F_O$, and offers a high impedance only at this frequency.

Signals at the fundamental, as well as those at other harmonics, are bypassed to ground by C_2 and C_3. The signal at the frequency $3F_O$ is inductively coupled into the base circuit of Q_2. Reverse bias for Q_2 is developed by the driving signal in a manner similar to that described for Q_1. The collector tank for Q_2 is tuned to $6F_O$ and inductively couples the signal at this frequency into the load. Undesired signals are again bypassed to ground, in this stage by C_5 and C_6.

While individual stages are seldom designed for frequency multiplications greater than three, any desired total multiplication may be obtained by connecting multipliers together. The usual arrangement in a transmitter is a straight-through class C amplifier following each one or two multiplier stages. In this

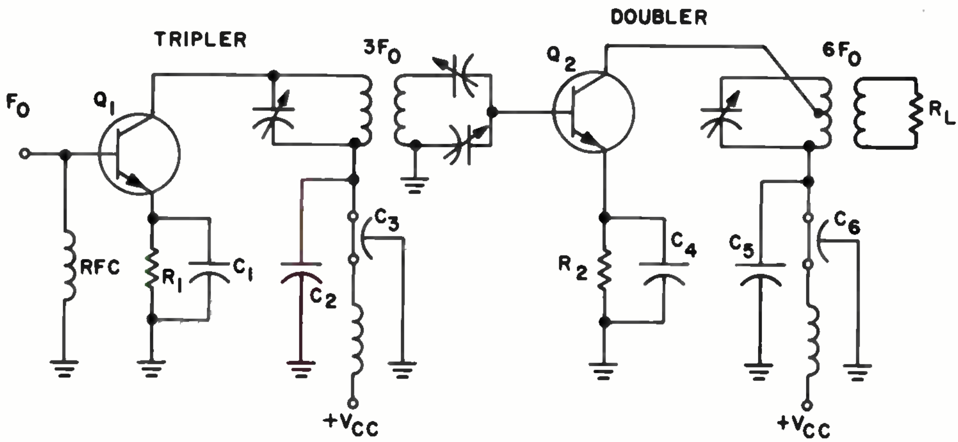


Fig. 33. Two frequency multiplier stages coupled together.

manner, the relatively low output from the multiplier is built up before being applied to the next multiplier. The straight-through amplifier also offers additional suppression to the undesired harmonics generated in the multiplier stage.

AMPLIFIER STABILITY

You learned that, in the triode power amplifier, in-phase feedback through plate-to-grid interelectrode capacitance could cause the amplifier to oscillate. To prevent these oscillations from occurring, components were added to feed back an out-of-phase voltage of equal amplitude and thus "neutralize" the interelectrode capacitance of the tube.

A similar capacitance exists between the collector and base of a transistor. However, the value of this collector-to-base capacitance in power transistors is voltage-dependent. That is, as the reverse bias across the collector base junction varies (which it normally does over the period of an operating cycle) the collector-to-base capacitance also varies. To be effective, a neutralizing circuit for a power transistor would have to continuously adjust itself to this variation. Because of this requirement, neutralization of a transistor rf amplifier is normally not practical. Instead, the need for neutralization is usually eliminated by careful circuit design using transistors with low values of interelement capacitance.

Parasitics. In the radio frequency range, the power gain of a transistor falls off rapidly as frequency increases. This characteristic of transistors works to advantage in preventing parasitic oscillations above the operating frequency. At these higher frequencies, the transistor has insufficient gain to overcome the circuit losses and sustain oscillations.

On the other hand, the power gain of a transistor is higher at the lower frequencies. To illustrate this, let's assume we have an rf power amplifier operating at 175 MHz. A typical transistor operating at this frequency might have a power gain of 5 db. This same transistor could have a gain as high as 30 db at 10 MHz. In other words, the power gain of the device is over 300 times greater at 10 MHz, the parasitic frequency, than at 175 MHz, the operating frequency. Consequently, the most common cause of instability in these power amplifiers is parasitic oscillation below the operating frequency.

The amplifier circuit shown in Fig. 34 illustrates a number of techniques used to prevent low frequency parasitics. These are discussed in the following paragraphs.

The rfc connected between base and ground (1) will at some frequency form a parallel-resonant circuit with the emitter base capacitance. To decrease the efficiency of this parasitic tank circuit, the rfc is designed to have a very low Q (high effective series resistance). Often this rfc is nothing more than a wire-wound resistor.

The emitter bypass capacitor (2) used is the smallest value which will provide effective bypassing at the operating frequency. At frequencies below the operating frequency, the reactance of this capacitor increases, resulting in degenerative feedback at these lower frequencies. This degeneration reduces the gain of the amplifier to low frequency parasitics.

The output coupling network is designed to include a portion of the network inductance in the collector dc supply line (3). With this arrangement, no

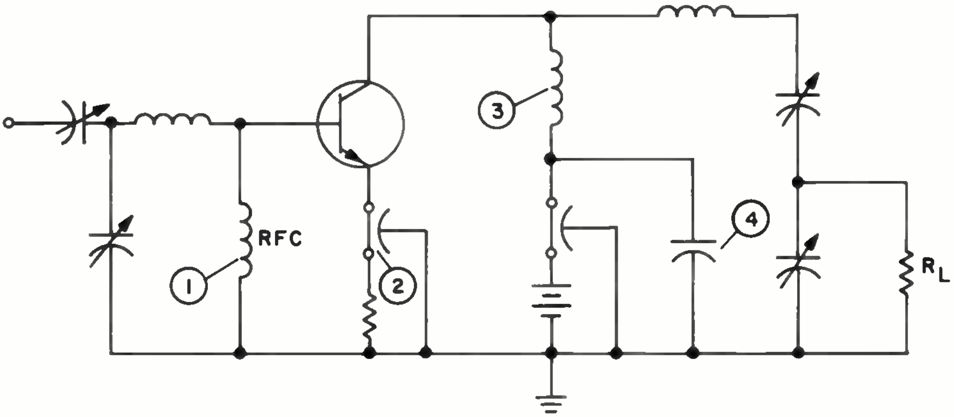


Fig. 34. Transistor power amplifier showing components used to prevent low frequency parasitics.

rfc is required in the collector circuit. Elimination of the collector rfc is desirable because this component can form a parallel-resonant circuit with the output capacitance at some relatively low frequency, thus becoming a possible source of low-frequency parasitics.

In addition to the feed-thru capacitor designed to bypass the power supply at the operating frequency, a second capacitor of larger value (4) provides a short circuit to ground at lower frequencies where parasitics usually occur. You might wonder why a larger capacitor, since it bypasses well at lower frequencies, wouldn't provide an even better bypass at the operating frequency where its X_c would be even less. The reason is that at higher radio frequencies the inductive reactance of the capacitor's leads becomes significantly large. The capacitor, instead of being a short circuit to ground, becomes an impedance to ground at these higher frequencies. The feed-thru capacitor, because of its physical construction, has very low lead inductance, and therefore provides an effective short circuit to ground at the high operating frequency. Feed-thru capacitors can only be manufactured with comparatively small values of capacitance, hence the need for the more conventional larger capacitor for the low-frequency bypass.

Fig. 35 summarizes what we have said about power supply bypassing. Shown in the figure are the equivalent bypass circuits for both the high operating frequency and the low parasitic frequency. At the operating frequency, the larger

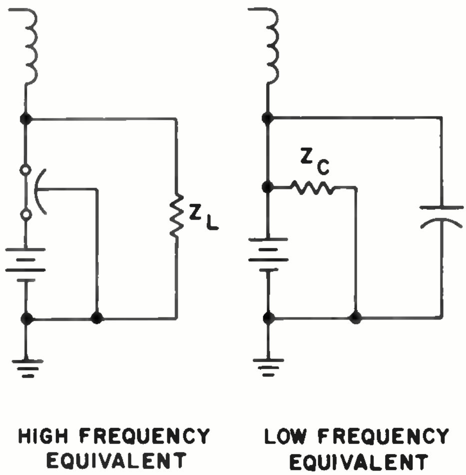


Fig. 35. Equivalent circuits for the power supply bypassing arrangement shown in Fig. 33.

capacitor appears as an inductive reactance, Z_L, and the feed-thru as an ac short circuit to ground. At a low parasitic frequency, the feed-thru appears as a high capacitive reactance, Z_C and the larger capacitor provides the ac short to ground.

SELF-TEST QUESTIONS

(x) Which transistor circuit configuration has the greatest stability at radio frequencies?

(y) What happens when loss of drive causes the bias on a transistor rf amplifier to drop to zero?

(z) Is reverse bias on the emitter base junction necessary for class C operation in a transistor rf amplifier?

(aa) What characteristic of transistor rf power amplifiers accounts for the low signal voltage developed in the input circuit?

(ab) Why is a parallel connection of transistors favored over a push-pull connection?

(ac) What are two advantages of having separate collector tank circuits for a parallel-connected transistor rf amplifier?

(ad) Why are feed-thru capacitors used for bypassing in high-frequency circuits?

(ae) What characteristic of power transistors makes neutralization impractical?

(af) What is the most common form of transistor rf amplifier instability?

(ag) Why is an additional capacitor placed in parallel with the feed-thru capacitor bypassing the power supply in Fig. 32?

Adjusting Class C Amplifiers

Adjustments to class C amplifiers in transmitter stages are performed following repairs, or routinely to compensate for normal circuit aging. The adjustment procedures for all class C stages, either frequency multipliers, intermediate amplifiers, or power output stages are basically the same. There are variations, of course; when you tune a frequency multiplier, for example, you must make certain that the plate circuit is tuned to the desired harmonic frequency.

In this section, we will go through the complete adjustment procedure for both a vacuum tube and a transistor class C amplifier stage. You should realize that adjustments such as those described in this section may be performed only by a person having the necessary authority. To obtain this authority, he must hold the proper class of FCC License or be under the direct supervision of another person who does.

THE VACUUM TUBE STAGE

Fig. 36 shows a typical class C amplifier circuit. Notice that there are current meters in the grid and plate circuits and that voltmeters are used to measure the bias, filament, and plate supply voltages. A power output stage using a screen-grid

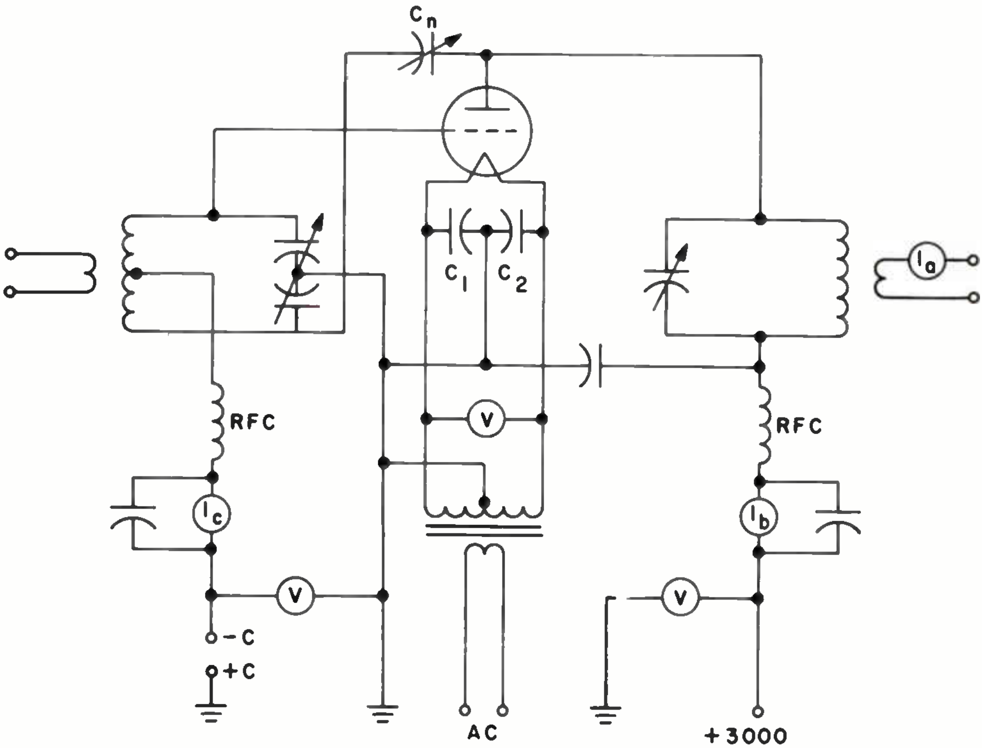


Fig. 36. Typical class C amplifier.

tube often has a current meter and voltmeter in the screen circuit. In some rf output stages, particularly in the low power exciter stages, only one voltmeter and one or perhaps two current meters are used. These meters are switched into the various stages to check the performance of the stage during operation or during the adjustment procedure.

Neutralization. The first step in the adjustment procedure is to insure that the amplifier is properly neutralized. Always check for neutralization with all interstage shields in place. If the amplifier is enclosed in a separate shield box within the transmitter cabinet, check it with the shield box closed. There will probably be stray magnetic or electrostatic coupling between output and input circuits unless all shields are in place and closed.

Neutralization can correct only for capacitive coupling directly from the grid to the plate of the tube. The simplest indication which can be used to determine correct neutralization is the grid current meter. With B+ removed from the stage, the need for neutralization will be indicated by a dip in the grid current when the plate tank is tuned through resonance. The grid current dips because the power loss in the resistance of the plate tank is greatest at resonance. Since this power is fed into the plate tank through the grid-plate capacitance, it subtracts from the grid current and causes a dip.

A second, more sensitive, method of checking for correct neutralization is by use of a wavemeter. As shown in Fig. 37, this simple device consists of a parallel L-C circuit connected to a diode and dc milliammeter. The variable capacitor, which is calibrated in units of frequency, may be adjusted to make the circuit resonant over a wide range of frequencies.

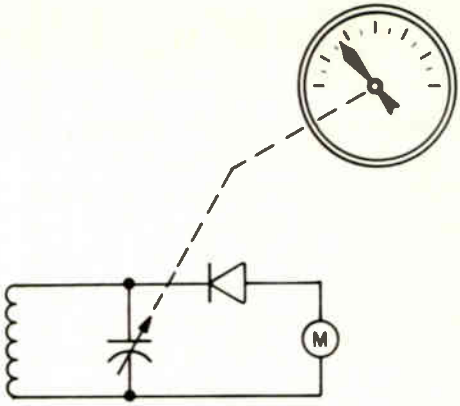


Fig. 37. Simplified schematic diagram of a wavemeter.

Any energy coupled into the tank circuit is rectified by the diode and causes the meter to deflect. In use, the wavemeter is inductively coupled to the plate tank circuit of the stage to be checked.

With plate voltage removed and grid drive applied to the stage, there should be no indication on the meter as the tuning knob is adjusted near the operating frequency. An indication on the meter indicates the presence of rf in the plate tank. This rf could only have come from the grid circuit -- coupled through interelement capacitance to the plate tank. Hence, the stage must be neutralized.

The procedure to be used in neutralizing an amplifier is as follows:

1. Remove the B+ from the stage. *Never attempt to neutralize an amplifier with the plate voltage connected.*

2. Set the neutralizing capacitor for minimum capacity.

3. Apply filament power and bias to the stage, and apply filament power, bias, and B+ to all stages ahead of the stage being neutralized.

4. Tune the grid circuit to resonance as indicated by maximum grid current.

5. Tune the plate tank circuit to resonance while watching the neutralization indicator. If it is a grid meter, it will dip; if you are using a wavemeter, it will peak.

6. Increase the capacity of the neutralizing capacitor slightly. Check grid and plate resonance; changing the neutralizing capacitor will sometimes detune both grid and plate circuits.

7. Continue to increase the neutralizing capacitance in small steps until there is no dip in the grid meter or no indication of power in the plate tank. The transmitter is then correctly neutralized.

As you come closer and closer to neutralization, make smaller and smaller changes in the neutralizing capacitor. There is only one correct setting; too much and too little capacitance are equally bad. Remember to retune both grid and plate each time you change the neutralizing capacitor. If the transmitter uses inductive link neutralization, start with maximum coupling to this link. Then reduce the coupling in small steps until you find the correct coupling.

Neutralization must be made as accurately as possible. Although steady oscillation will take place only when enough power is fed back from the output to overcome the input circuit losses, smaller amounts of feedback, which are not enough to cause steady oscillation, can still affect the operation of the stage. An amplifier operating like this is said to be "regenerative."

Several characteristics of an amplifier change when it is regenerative. One of the most pronounced is an increase in input impedance. This increase in impedance causes the Q of the grid and plate tuned circuits to increase also. The increase in Q makes the circuits selective and hard to tune. To make matters worse, changes in the plate tuning change the amount of feedback and, therefore, affect the grid tuning.

A regenerative amplifier is an unstable amplifier. A slight increase in filament current or plate voltage may cause it to go into steady oscillation. Reducing the load at the output will also cause oscillation.

In a keyed transmitter, a regenerative amplifier will cause damped oscillations every time the key is closed. As a result undesirable signals are generated. In a phone transmitter, an unstable amplifier causes still other effects.

Perfect neutralization of an amplifier is absolutely necessary. It takes time to do it right, but it does not have to be done often.

After you have completed the neutralizing procedure, apply a low plate voltage to the stage. Tune the plate tank capacitor to resonance as indicated by minimum plate current readings. You will notice that the plate current will dip sharply because the output tank circuit is not delivering power to the load. This is shown by the solid curve in Fig. 38.

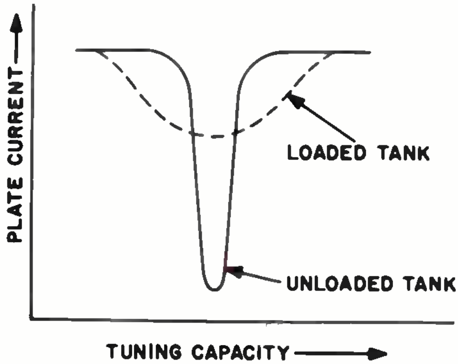


Fig. 38. How plate current dips as tank capacitor is tuned through resonance; the unbroken line shows the sharp dip that occurs if the tank is not loaded; the dashed line shows the broad dip that occurs if the tank is loaded.

The grid current meter will indicate maximum at resonance. Retune the grid tank capacitor, and increase the excitation until the grid draws the rated current.

Increase the loading in the plate tank circuit until any increase in loading causes the current through the antenna meter to drop. Increase the loading in steps. Check the plate circuit tuning for resonance each time you increase the load. Now apply the plate voltage and adjust it to the rated value.

Retune the plate tank to resonance, and then advance the loading until the tube draws the rated plate current. The minimum plate current point will not be as sharp because the tuned circuit is now loaded and more power is being fed into the load circuit. The loaded plate current tuning curve is shown by dotted lines in Fig. 38. Adjust the grid tank and excitation until the rated grid current is drawn.

Make final fine adjustments to the plate tuning and antenna loading. Be certain all meters show the recommended values for proper operation of the stage.

The output is indicated by the current readings on the rf antenna current meter. As the stage is resonated and the loading is increased, the antenna current increases, indicating power is being delivered to the antenna. The antenna meter is just as important as the plate ammeter when tuning. If the output current does not increase when the plate current increases, the plate circuit is not tuned to resonance or is overloaded. Reduce the coupling and retune the plate tank.

Parasitics. It is interesting to note that the wavemeter is also useful in detecting the presence of parasitics in the operating amplifier. As you know, these oscillations take place at a frequency far removed from the operating frequency. The most practical way to locate the oscillations is to reduce the bias of the stage so that the tube is no longer operated beyond plate-current cutoff. Then reduce the plate voltage so that the maximum plate dissipation of the tube is not exceeded. Disconnect the output and remove the drive from the stage. These changes make the circuit most favorable for the generation of parasitic oscillations.

With the wavemeter inductively coupled to the circuit suspected of oscillating, the wavemeter tuning knob is varied over its range. A meter deflection not only indicates parasitic oscillation, but the frequency may be approximately determined by the position of the tuning knob. Knowing the frequency at which parasitics are occurring often provides a clue to their origin.

THE TRANSISTOR STAGE

The transistor power amplifier, although used to some extent in low-level fixed station and broadcast transmitters, finds its widest application in low-powered mobile transmitting equipment. Since these units are operated largely by non-technical people working under less than ideal conditions, the emphasis in their design is on simplicity and reliability. Because of this emphasis, adjustment procedures for transistor class C amplifiers are usually simple and straightforward. The complete transmitter alignment of many of these units consists in its entirety of peaking the indication on a power output measuring device with as few as two transmitter adjustments.

Even in the more elaborate transmitters you'll seldom find more than one meter built into the equipment. This single meter is switched into the various points in the circuit where current or voltage is to be measured. Sometimes, all the monitoring points in the circuit are connected to a multipin jack on the transmitter chassis. When transmitter adjustments are to be made, an external meter equipped with a selector switch is plugged into this jack for monitoring.

Fig. 39 shows a parallel-connected output stage such as might be found in one of the higher powered transistor transmitters. We have shown separate meters at the various monitoring points for clarity. Before applying power to the amplifier, L_1 and L_3 should be adjusted for minimum drive to the transistors (adjusted for maximum inductance). With this accomplished, apply power and adjust the collector circuit of Q_1 to resonance. This is done by adjusting L_2 for a dip on M_1.

In like manner, adjust the collector ircuit of Q_2 to resonance using L_4 and M_2. Next adjust the coupling to the load using C_1 to obtain the rated load current as measured on M_3. Finally, adjust the base drive to the transistors using L_1 and L_3 to obtain equal collector currents at the rated value. This completes the preliminary adjustment of the stage. Since there is some interaction between the various adjustments, recheck the setting of L_2, L_4, and C_1. At all times maintain the collector currents at or below the rated value by adjusting L_1 and L_3.

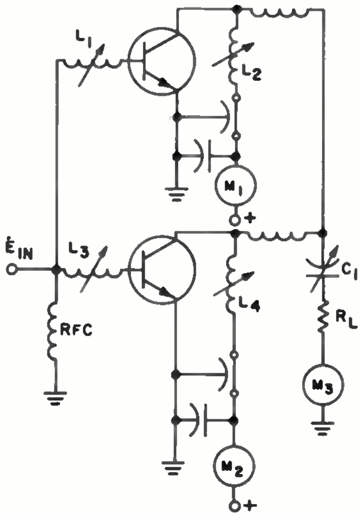


Fig. 39. Parallel-connected transistor stage.

In conclusion, the adjustments we've discussed in this section should not be considered as a procedure to be memorized and followed in any specific case. They are presented here to illustrate the basic approach to power amplifier adjustment. Before attempting any adjustment to a power amplifier, carefully consult and follow the manufacturer's literature. In making the adjustments, both the procedure and the sequence in which the steps are performed are of the greatest importance. An expensive tube or transistor may be destroyed as a result of any adjustments performed without complete knowledge of the correct procedure.

SELF-TEST QUESTIONS

(ah) How may a vacuum tube stage be checked for proper neutralization using the grid current meter as an indicator?

(ai) Why must B+ first be removed from a stage before checking the plate tank for the presence of rf?

(aj) What are some indications of a regenerative amplifier?

(ak) What tuning defect is indicated if the antenna current does not increase with an increase in plate current?

(al) Why is the drive to the transistor amplifier in Fig. 39 adjusted for minimum before collector power is applied.

Answers to Self-Test Questions

(a) Because output current flows in the amplifier only during the relatively brief conducting period.

(b) Both the bias level and the amplitude of the driving signal affect the operating angle.

(c) No. Sometimes, high power gain in an amplifier, obtained with a matched load, must be sacrificed for greater power output.

(d) The value of L and C would both have to be changed. L would be increased and C decreased.

(e) The loaded Q would decrease. The unloaded Q would be unaffected. The efficiency of the tank would increase.

(f) If the loaded Q decreases, the reflected resistance in the tank must have increased. The increased reistance in the tank causes tank impedance to decrease.

(g) By offering a high impedance only at the resonant frequency.

(h) By decreasing the spacing between the coils.

(i) Since the power supply, the tank, and the stage are in series, the amplifier is series-fed.

(j) The output connection to the tank could be tapped down on the tank coil, as is shown in Fig. 9C. This causes the internal resistance of the stage to appear to the tank as a higher value. This higher resistance seen by the tank reflects a higher resistance into the tank which increases the loaded Q.

(k) An antenna usually loads a power amplifier more heavily than the input circuit to another class C amplifier.

(l) If the drive to the stage is lost, no bias will be developed, allowing the tubes to conduct heavily. This heavy conduction could damage the tube.

(m) Through the plate-to-grid capacitance.

(n) Plate neutralization is effective over a range of frequencies, whereas the blocking capacitor and coil arrangement is effective only at one frequency.

(o) 180°. The signal fed back through a neutralizing circuit will always be 180° out-of-phase with the signal fed back through the grid to plate capacitance.

(p) The radio frequency choke.

(q) Parasitic oscillations in the amplifier.

(r) Their higher power gain and reduced grid-to-plate capacitance.

(s) Their higher power gain. Even a very small amount of feedback may cause the tetrode to oscillate.

(t) The value of L would be lower and the value of C higher. This would provide the lower value of tank impedance necessary for the parallel connected tubes.

(u) By connecting the supply voltages to the midpoints of the tank coils.

(v) The doubler. In general, the greater the frequency multiplication in a stage, the lower the power output.

(w) Since the output frequency is different from the input frequency, any signal fed back would not add to the input signal.

(x) The common emitter.

(y) The transistor stops conducting.

(z) Not always. A zero-biased transistor is already several tenths of a volt below collector current cutoff.

(aa) The low input impedance of the amplifier.

(ab) Because of the expense of the transformer required to drive a push-pull stage.

(ac) RF choke losses are eliminated and the transistors are electrically independent.

(ad) Because they have a very low lead inductance.

(ae) The collector base capacitance varies over the period of an operating cycle. Thus a varying amount of signal is fed back to the input.

(af) Low-frequency parasitics. The power gain of a transistor increases rapidly as frequency decreases. For this reason, a transistor amplifier is most susceptible to low-frequency parasitic oscillations.

(ag) Because of their physical construction, feed-thru capacitors cannot be manufactured with high values of capacitance. Proper bypass at low parasitic frequencies requires a large capacitance, so another type must be connected in parallel with the feed-thru.

(ah) With plate voltage removed, a dip in grid current when the plate tank is tuned through resonance indicates the need for neutralization.

(ai) With the B+ on the stage, rf in the tank circuit is a normal indication.

(aj) Changes in plate tuning affect grid tuning. Also, oscillations occur with reduced loading or slight changes in operating voltages.

(ak) The plate tank is not tuned to resonance or is too heavily loaded.

(al) To prevent possible excessive collector current flow due to the low impedance offered by the detuned output circuit.

NOTES

Lesson Questions

Be sure to number your Answer Sheet C205.

Place your Student Number on every Answer Sheet.

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

1. What is the primary purpose of a buffer stage?

2. What is the efficiency of a transmitter output stage if the plate supply voltage is 2000 volts, the plate current is 330 ma, and the measured power output of the stage is 500 watts?

3. What is the probable cause of trouble if the plate current of a class C stage using grid-leak bias alone rises to a very high value and the grid and output currents decrease?

4. Why is neutralization necessary in a triode amplifier?

5. Why are transistor rf amplifiers more susceptible to low-frequency parasitics than to those at higher frequencies?

6. Why does a screen-grid tube usually not require neutralization?

7. (A) If the master oscillator in a transmitter using a single doubler stage operates at 7.6 MHz, what is the transmitter output frequency?
   (B) If a transmitter output frequency of 26.4 MHz is obtained by using a single tripler stage, what is the frequency of the master oscillator?

8. What is the purpose of a Faraday screen?

9. What precaution should be taken when tuning the output of a frequency multiplier stage?

10. In adjusting a class C transmitter stage, what direction (upscale or downscale) would you expect the pointers on the following dc current meters to deflect to indicate resonance: (A) the plate current meter; (B) the grid current meter?

THE VALUE OF KNOWLEDGE

Knowledge comes in mighty handy in the practical affairs of everyday life. For instance, it increases the value of your daily work and thereby increases your earning power. It brings you the respect of others. It enables you to understand the complex events of modern life, so you can get along better with other people. Thus by bringing skill and power and understanding, knowledge gives you one essential requirement for true happiness.

But what knowledge should you look for? The first choice naturally goes to knowledge in the field of your greatest interest–electronics. Become just a little better informed than those you will work with, and your success will be assured.

It pays to know–but it pays even more to know how to use what you know. You must be able to make your knowledge of value to others, to the rest of the world, in order to get cash for knowledge.

The NRI Course gives you knowledge, and in addition shows you how to use what you learn. Master thoroughly each part of your Course, and you'll soon be getting cash for your knowledge.

nri

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-305

ACHIEVEMENT THROUGH ELECTRONICS

OSCILLATORS FOR COMMUNICATIONS EQUIPMENT

C206

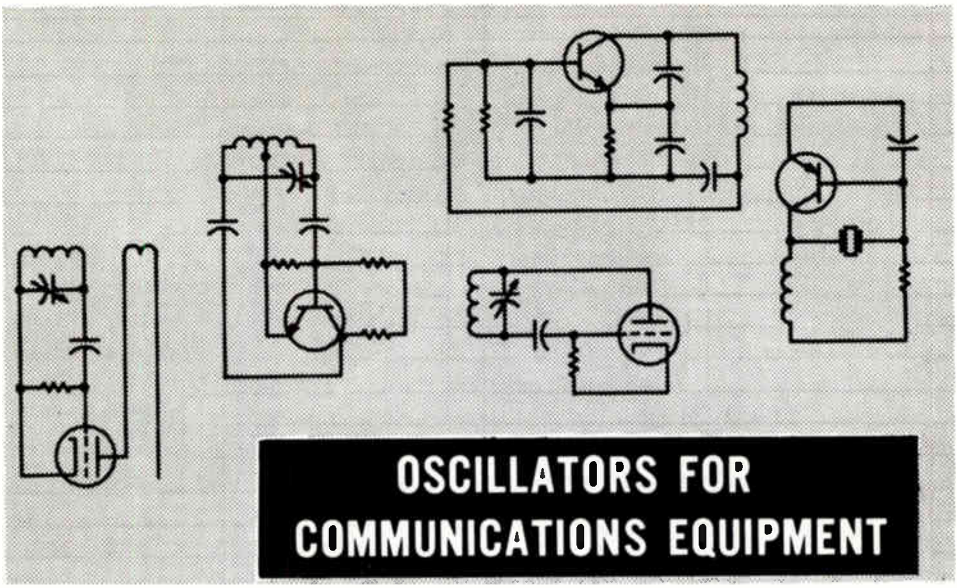
NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

OSCILLATORS FOR COMMUNICATIONS EQUIPMENT

C206

STUDY SCHEDULE

KL2M172 1972 EDITION Litho in U.S.A.



OSCILLATORS FOR COMMUNICATIONS EQUIPMENT

One of the most important circuits in electronics is the oscillator. If it were not for the oscillator, radio and television and many industrial electronics applications would not be possible. An oscillator is an amplifier which generates an ac signal. The frequency of the signal is determined by the value of the components in the oscillator circuit.

In this lesson you will begin with a study of the basic oscillator circuit. You will learn the characteristics of an oscillator and how the basic oscillator works. Then, you will study applications of the oscillator, learning the details of operation of various oscillator circuits. From there you will go into methods of controlling oscillator frequency. The lesson will conclude with a brief description of nonsinusoidal oscillators.

There are many different types of oscillator circuits. For convenience in studying them, we will divide them into two types: L-C oscillators and R-C oscillators. L-C oscillators are oscillators in which inductance and capacitance are used in the frequency-determining network. R-C oscillators are oscillators in which resistance and capacitance are used in the frequency-determining network. Both types of oscillators work on the same general principle, that of feeding some of the signal from the output circuit back into the input circuit. This feedback signal enables the oscillator to go on generating its own signal. The amount of signal that must be fed back into the input depends upon a number of things, but in general it must be enough to overcome the losses in the input circuit of the oscillator.

Perhaps one of the most important considerations in an oscillator circuit is how the energy is fed from the output circuit back into the input circuit. Although it is important to feed enough signal back into the input circuit, it is even more important for the signal fed from the output back into the input to be of the correct phase. If the phase of the feedback signal is not correct, instead of aiding the input signal, it will oppose it, and the oscillator will not oscillate.

Sometimes an oscillator is considered

as a converter circuit. In other words, it converts dc into ac. The dc is supplied by the power supply to the tube or transistor used in the oscillator circuit, which changes this dc energy to ac energy.

Oscillators are the only practical means of generating high-frequency radio waves. In the early days of radio, before practical oscillators were developed, rf signals were generated by means of high-frequency generators called alternators. However, there is a limit to how high a frequency a rotating machine such as an alternator can develop, and hence most radio transmission was carried out on very low frequencies.

The most important part of the oscillator is the resonant circuit, so before we begin let's review it. In a previous lesson, you learned that the resonant tank in a tuned amplifier biased class C could be made to store energy and to deliver that energy during periods when the tube or transistor was cut off. Oscillators function similarly to class C rf amplifiers in that the tank circuit must continue to supply an output once the input signal is removed.

One characteristic of a resonant circuit that we have not discussed is its ability to produce a damped wave when it is shock-excited. We will now see what we mean by a damped wave and see how it is produced by a resonant circuit.

DAMPED WAVES

Consider the circuit shown in Fig. 1. A coil and capacitor are connected in parallel and connected to a battery through a switch. For our discussion we must assume that the switch can be opened and closed instantaneously. Now let's see what happens when we close the switch for an instant.

At the instant the switch is closed, electrons flow from the negative terminal of the battery into side A of the capacitor. At the same instant, electrons will flow out of side B of the capacitor to the positive terminal of the battery. If the resistance in the circuit, which includes the battery resistance, is very low, the capacitor can charge up almost instantly to a voltage equal to the battery voltage. Thus, terminal A of the capacitor will be negative and terminal B will be positive.

At the same time, when the switch is closed, there will be some tendency for current to flow through the coil from terminal C to terminal D. However, you will remember that one of the characteristics of a coil or inductance is that it opposes any rapid change in the current flowing through it. The instant before the switch is closed, the current flowing through the coil is zero. The coil would like to keep it that way. When the switch is closed, the inductance of the coil tries to prevent a current from building up in the coil. Actually, there will be some small current flowing through the coil from terminal C to terminal D, but if the switch is closed only for an instant, the current will not be able to build up appreciably. Therefore, at the instant the switch is opened again we have the capacitor charged, as shown in Fig. 2A,

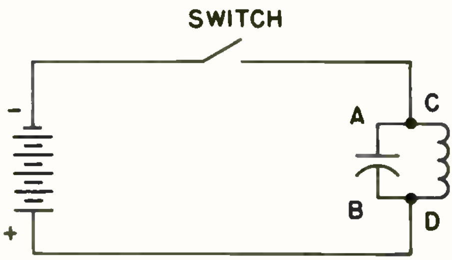


Fig. 1. A simple method of producing a damped wave.

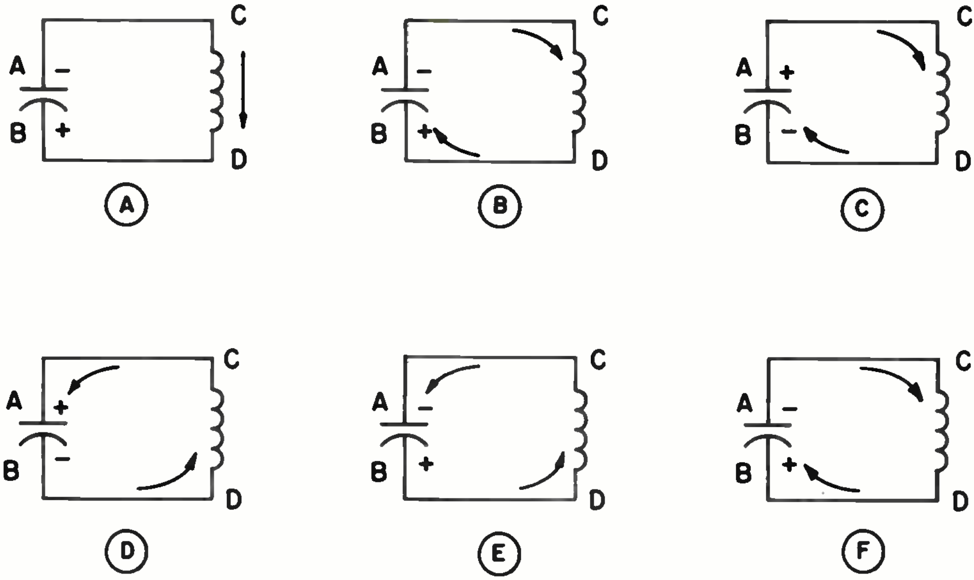


Fig. 2. How oscillation takes place in a resonant circuit.

and a small current flowing through the coil as indicated.

When the switch is opened and we have the situation shown in Fig. 2A, we have a capacitor that is charged, and immediately starts to discharge. As a result, a current flow will be set up in the circuit as shown in Fig. 2B. Now remember that a coil opposes a change in the current flowing through it. Therefore the capacitor cannot discharge instantly through the coil, but rather must build up a current in the coil which will build up a magnetic field about the coil. Eventually, the capacitor will build up a current flow in the coil and enough electrons will leave plate A to get to plate B to discharge the capacitor.

The discharge of the capacitor removes the voltage that caused current to flow through the coil. The magnetic field around the coil now collapses. The collapsing field generates an emf in the coil, which tends to keep the current flowing in the same direction as before. This continued current causes electrons to flow onto plate B of the capacitor, giving the capacitor a charge opposite to what it had at the start. This condition is shown in Fig. 2C.

After the field around the coil has collapsed, there is no emf to hold the charge on the capacitor. The capacitor now begins to discharge back through the coil as shown in Fig. 2D. The flow of current caused by the discharge of the capacitor builds up a magnetic field around the coil until the capacitor is fully discharged; the magnetic field collapses and keeps the current flowing. This current flow charges the capacitor with the same polarity it had at the instant the switch was opened. This is shown in Fig. 2E.

Again, the current will eventually drop to zero, and then the capacitor will once again begin to discharge through the coil in the opposite direction, this time with electrons flowing from plate A to plate B as shown in Fig. 2F.

Notice that in Fig. 2F we have the same situation as we had in Fig. 2B. In other words, we have gone through a complete cycle of events. The capacitor was charged with one polarity. This produced a current flow through the coil, which eventually charged the capacitor with the opposite polarity. The capacitor then began to discharge through the coil in the opposite direction, which built up a charge on it having the same polarity as the original charge placed on the capacitor. Once again this charge on the capacitor began the cycle of events all over again by attempting to discharge through the coil.

You might think that this oscillation, or backward and forward flow of current through the coil to charge and discharge the capacitor would continue indefinitely. Indeed, if we had a perfect coil and a perfect capacitor, once the oscillation was started, it would continue indefinitely. However, there is no such thing as a perfect coil or a perfect capacitor. There will be some losses in both parts, so instead of having an oscillation which continues indefinitely, we will have what is called a damped wave. The damped wave of voltage across the capacitor is shown in Fig. 3.

The important thing to notice in this damped wave is that the amplitude of each cycle is just a little bit less than the amplitude of the preceding cycle. In other words, the wave is slowly dying out because of losses in the resonant circuit. The lower the losses in the circuit, the greater the number of cycles that will occur before the wave disappears. On the other hand, the higher the losses in the circuit, the fewer the number of cycles.

If we could find some way of closing the switch in Fig. 1 for just an instant when plate A of the capacitor reaches its maximum negative charge, we could supply a small amount of energy to the resonant circuit to make up for losses in the circuit. If we continue to supply this small amount of energy once each cycle, then the resonant circuit will continue to oscillate indefinitely, and we could use it as a source of ac power. This is what an oscillator does, it supplies a pulse of energy at the correct time to make up for losses in the resonant circuit. We'll see how this is done later, but let's learn more about resonant circuits first.

FACTORS AFFECTING RESONANT CIRCUITS

There are several additional important things we should know about resonant circuits. For example, we should know the frequency at which oscillation takes place in a resonant circuit. We should also know what factors affect the loss of energy from cycle to cycle, in other words, how rapidly the wave train will die out.

Another term that we frequently encounter when dealing with resonant circuits is "period". We will now learn something about these factors.

Frequency. The frequency at which a resonant circuit oscillates will depend upon the inductance and capacitance in the circuit. We already know that resonance occurs when the inductive reactance of the coil is exactly equal to and

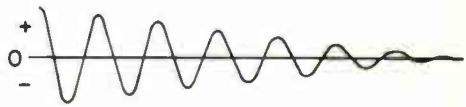


Fig. 3. Voltage across the capacitor.

canceled by the capacitive reactance of the capacitor. In other words, at resonance:

$$X_L = X_C$$

We know that the inductive reactance of a coil, X_L, is given by the formula:

$$X_L = 6.28 \times f \times L$$

and the capacitive reactance of a capacitor is given by the formula:

$$X_C = \frac{1}{6.28 \times f \times C}$$

Now, since resonance occurs when $X_L = X_C$, let's substitute the values for X_L and X_C and we will get:

$$X_L = X_C$$

$$6.28 \times f \times L = \frac{1}{6.28 \times f \times C}$$

and this can be manipulated to give us:

$$f^2 = \frac{1}{6.28^2 \times L \times C}$$

and now if we take the square root of both sides of the equation we get:

$$f = \frac{1}{6.28 \times \sqrt{L \times C}}$$

For convenience in expressing formulas of this type, the times sign is usually omitted, and in place of 6.28, the term 2π is often used, so you will usually see the formula for the frequency at which a resonant circuit will oscillate expressed as:

$$f = \frac{1}{2\pi\sqrt{LC}}$$

You should remember this formula because it is very important; but even more important, remember what the formula tells you. The formula says that the frequency of a resonant circuit varies inversely as the square root of the L-C product. Now remember as we mentioned before, when one factor varies directly with another, making one bigger makes the other bigger, and when two things vary inversely then we have the opposite situation; making one bigger makes the other smaller. Here we have a situation where the frequency varies inversely as the square root of the L-C product. This means that increasing the size of either L or C will reduce the frequency at which the resonant circuit oscillates, and reducing the size of either L or C will increase the frequency at which the resonant circuit oscillates. We can express this simply by saying: *Larger L or C, lower frequency; smaller L or C, higher frequency.*

In using this formula, the frequency of oscillation will be given in cycles per second and the value of L and C used must be in henrys and farads respectively.

Period. The period of a resonant circuit is the time it takes the resonant circuit to go through one complete oscillation. Thus, if we have a circuit that is resonant at a frequency of 1000 cycles per second, its period would be 1/1000 of a second, and if we have a resonant circuit that is resonant at a frequency of 1,000,000 cycles per second, the period would be 1/1,000,000 of a second.

The period of a resonant circuit is given by the formula:

$$P = \frac{1}{f}$$

where P represents the period of the resonant circuit in seconds and f the frequency in cycles per second.

Since in electronics we are usually dealing with comparatively high frequencies, it follows that the period of most resonant circuits will be only a very small fraction of a second. As a matter of fact, the period of many resonant circuits will be only a small fraction of a millionth of a second. Therefore, to simplify things, the microsecond is frequently used in electronics work as a unit of time. A microsecond is 1/1,000,000 (one millionth) of a second. Thus if a resonant circuit has a period of 5/1,000,000 (five millionths) of a second we can say that its period is 5 microseconds, or if another resonant circuit has a period of 1/10,000,000 (one ten millionth) of a second, we can say this period is one-tenth of a microsecond.

In order to show the cycle-time relationship, the frequency of a circuit is measured in units called HERTZ. One Hertz being equivalent to one complete cycle in one second, 1,000 cycles in one second would then be 1,000 Hertz (Hz) or one kilohertz (kHz). 1,000,000 cycles in one second would be one megahertz (MHz). These terms are replacing the older terms of kilocycles (kc) and megacycles (mc) still used in many publications.

The Q Factor. The number of cycles that will occur when a resonant circuit is shock-excited depends almost directly upon the Q of the coil. The higher the Q, the more cycles will occur.

The Q of a coil tells us essentially how good a coil we have. A coil that has a high Q has a high inductive reactance compared to the resistance of the coil. A coil with a low Q has high resistance compared with the inductive reactance.

The Q of a coil is expressed by the formula:

$$Q = \frac{X_L}{R}$$

and we can express X_L as equal to:

$$6.28 \times f \times L$$

and substituting this in the formula for the Q of a coil we get:

$$Q = \frac{6.28 \times f \times L}{R}$$

If we examine this formula, we see that the Q varies directly as the frequency and inductance and inversely as the resistance. Therefore, you might think that increasing the frequency of the resonant circuit by using a smaller capacitor in conjunction with the coil will result in a higher Q. This will often happen, but the increase in Q is not as great as might be expected, because the resistance of the coil is the ac resistance rather than the dc resistance. The ac resistance of a coil actually represents ac losses in the coil and this varies directly as frequency varies. Therefore, increasing the frequency of the resonant circuit increases the inductive reactance of the coil, but at the same time it increases the losses so that the Q normally does not increase as fast as we might expect.

In a resonant circuit with a high Q coil there will be a large number of cycles in a damped wave train set up by shock-exciting the resonant circuit. In other words, the amplitude of one cycle will be very little less than the amplitude of the preceding cycle. However, if the Q of the coil is low, then the losses in the coil will be quite high so that the amplitude of each cycle will be substantially less than the amplitude of the preceding cycle. This means that the oscillation will be damped out quite rapidly and the number of cycles that occur when the circuit is shock-excited will be somewhat limited.

In most oscillator circuits a comparatively high Q coil is used. The reason for this is that if the coil has a high Q, then only a small amount of energy must be supplied by the tube or transistor in the oscillator circuit in order to sustain oscillation. On the other hand, if the coil has a low Q, the losses in the resonant circuit will be quite high, with the result that the tube or transistor used in the oscillator circuit must supply a comparatively large amount of energy in order to keep the oscillation going.

SELF-TEST QUESTIONS

(a) What type of feedback is used in oscillator circuits?

(b) If the inductance of an L-C circuit is increased, what happens to the frequency?

(c) If the Q of the resonant circuit is increased, what happens to the damped wave train?

(d) If the resonant frequency of an L-C circuit is 2000 kHz, what is the period of one cycle?

The Basic Oscillator

The function of the switch in Fig. 1 is to supply energy of the proper phase and at the proper time to sustain oscillations in the resonant circuit. At radio frequencies it would be impossible for a mechanical switch to do this. Therefore, we must use an electronic switch such as a vacuum tube or transistor.

THE ELECTRONIC SWITCH

In order to see how the vacuum tube can be used as an electronic switch, let's go back to the basic circuit we had in Fig. 1. We have repeated this circuit as Fig. 4A. It is exactly the same as Fig. 1 except we have simply indicated where the battery voltage is to be connected instead of actually showing the battery in the circuit. In practice, we could use either a battery or the dc output of a suitable power supply. If we can momentarily close the switch, we will charge the capacitor C_1 and produce oscillation in the parallel resonant circuit consisting of C_1 and L_1. However, this oscillation will die out after a number of cycles because of the losses in a resonant circuit, unless we can find some way of supplying additional energy to the resonant circuit to make up these losses. If we could close the switch at the right instant each cycle, we could recharge capacitor C_1 once each cycle and keep the oscillation going. However, if the resonant frequency of the circuit is several hundred Hertz or higher, it would be impossible to close the switch manually at the correct instant to keep the oscillation going. As a matter of fact, it would be difficult to do this mechanically except at a very low frequency.

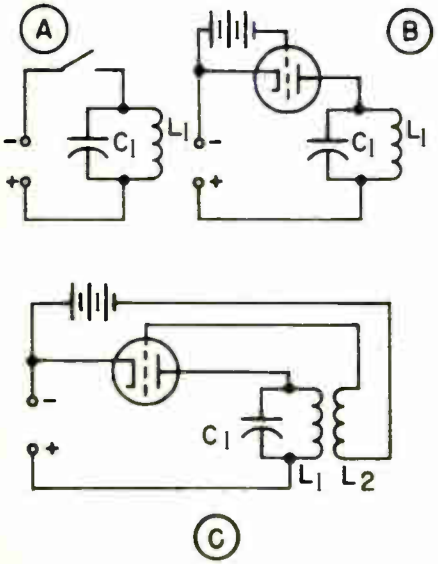


Fig. 4. Using a tube as an electronic switch to supply the losses in a resonant circuit.

In Fig. 4B, we have replaced the switch with a vacuum tube. The cathode of the vacuum tube is connected to the negative side of the power supply or battery, and the plate of the tube is connected to the resonant circuit. Between the cathode and the grid of the tube, we have connected a battery that will place a negative voltage on the grid of the tube. The battery voltage in the grid circuit is high enough to bias the tube beyond cutoff. Thus, with the circuit exactly as shown in Fig. 4B, the bias on the tube is so high that there will be no current flowing through the tube and hence no way to charge capacitor C_1 and start the resonant circuit oscillating. We have in effect the same situation as we have in Fig. 4A with the switch open.

Now let's look at the circuit shown in

Fig. 4C. Here we still have the tube connected in the circuit in exactly the same way except that we have added a coil, L_2, between the negative terminal of the grid battery and the grid of the tube. This coil is placed near L_1 so that it will be inductively coupled to L_1. Thus, if there is any change in the magnetic field about L_1, the changing flux will induce a voltage in L_2.

Now let's consider what will happen if we momentarily short the plate and cathode of the tube together. If we do this, capacitor C_1 will be charged. As soon as we remove the short, C_1 will start to discharge through L_1 and in doing so will build up a magnetic field about L_1. The changing lines of flux will cut L_2 and induce a voltage in it. This voltage in L_2 will be in series with the battery voltage applied between the cathode and grid of the tube. If the end of L_2 that connects to the grid of the tube is negative, and the other end positive, then the voltage induced in L_2 will add to the grid bias, biasing the grid still further negative so that no current can flow from the cathode to the plate of the tube. However, if the voltage induced in L_2 has a polarity such that the end of L_2 that is connected to the grid is positive, and the other end is negative, then this voltage will oppose the battery bias voltage and reduce it so that the total grid bias will be reduced below the point where the plate current is cut off, and current can flow through the tube. Therefore, by connecting L_2 with the proper polarity, we can arrange the circuit so that when the plate side of capacitor C_1 reaches its negative peak, the tube will conduct, and a burst of electrons will flow through the tube, charging C_1 still further. Thus, any loss in the charge across C_1 due to losses in the resonant circuit will be made up for by the burst of electrons flowing through the tube.

In Fig. 5, we have shown a number of sine-wave cycles such as the oscillation that might occur in the L_1-C_1 resonant circuit. The shaded pulses represent the bursts of current flowing through the tube that will reinforce the oscillation and keep it going. Notice in Fig. 5 that the burst of current flowing through the tube occurs at the correct instant to aid the oscillation. Also, notice that the current burst occurs for only a small fraction of a cycle. The current does not flow through the tube during the entire cycle.

For several reasons the oscillator circuit shown in Fig. 4C is not a practical circuit. For one thing, the battery used to provide the negative bias on the grid is somewhat cumbersome. If we were using a power supply to furnish the voltage to operate this oscillator from a power line, we would not want to be bothered with a separate battery to supply the grid bias. Furthermore, with this type of arrangement, it would be possible to pick up such a high voltage pulse in L_2 that the tube would pass an extremely high current when it was driven in a positive direction. As the grid bias battery aged and the voltage from this battery dropped, an even higher current pulse would flow through the tube. As a matter of fact, the pulse might be so high that the tube could be damaged.

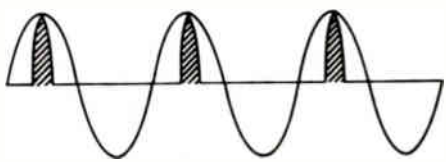


Fig. 5. The oscillator pulse is timed to occur at the peak of the oscillation in the tank circuit to reinforce the oscillation.

Both of these objections can be overcome by modifying the circuit as shown in Fig. 6. Let's look at Fig. 6A first. In Fig. 6A, you will see that we have replaced the battery in the grid circuit by a resistor, R_1, with capacitor C_2 connected across it. In other respects the circuit is identical to the circuit shown in Fig. 4C.

Let's see exactly how this circuit works. When voltage is first supplied to this circuit, there will be no grid bias on the tube. The tube starts to conduct and charges capacitor C_1. Electrons will flow into the side of this capacitor that connects to the plate of the tube and out of the other side. At the same instant, current will start to flow through L_1 and there will be a rapid change in the lines of flux about this coil. The changing magnetic lines will cut L_2, and induce a rather high voltage in it. Coil L_2 is connected so that the grid of the tube will be driven in a positive direction, which will result in a still further increase in current flowing from the cathode to the plate of the tube, which will charge C_1 still further.

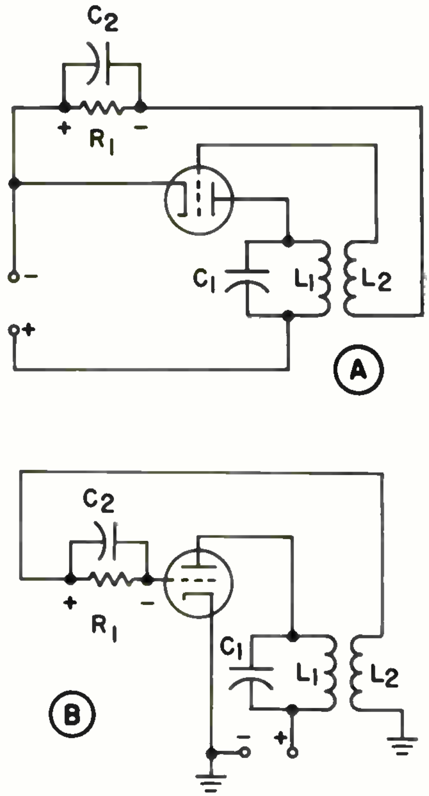


Fig. 6. A tuned-plate oscillator.

Since the grid of the tube will be driven in a positive direction, it too will attract electrons, and electrons will flow from the cathode of the tube to the grid, through L_2, and then through R_1 back to the cathode of the tube. In flowing through R_1, they will set up a voltage drop across it and charge capacitor C_2 with the polarity indicated on the diagram.

Eventually the rate at which the flux lines are cutting L_2 will decrease, so the voltage induced in coil L_2 will drop. The voltage across R_1 will cut off the flow of plate current in the tube. Capacitor C_2 starts to discharge through R_1 and keeps the grid of the tube at a high enough negative potential to keep it cut off. When this happens, we have opened the switch as in Fig. 1 and an oscillation starts in the tank circuit. The capacitor and coil begin exchanging energy back and forth. At the correct instant, once in each cycle, the grid of the tube will be driven positive by the voltage induced in L_2 by the changing flux from L_1 so that the tube will pass a burst of electrons to recharge C_1 and make up any energy lost in the tank circuit.

The oscillator we have been discussing is called a tuned-plate oscillator. In actual practice, the circuit is modified and you will usually see it like Fig. 6B. Notice that the position of the grid resistor and grid capacitor, R_1 and C_2, have been changed with reference to L_2. In other words, tracing from the grid of the tube, we come to the grid resistor and grid capacitor first and then through L_2 to ground. However, regardless of how the resistor and capacitor are connected in series with L_2, the action of the circuit is the same.

This type of oscillator has several disadvantages that can be eliminated by different circuitry. However, since it is a basic circuit and enables us to see exactly how the tube is acting as a switch, it is a good circuit to start our study of oscillators with.

SELF-REGULATION

The oscillator circuits shown in Fig. 6 are self-regulating. This means that they tend to control the flow of current through the tube themselves. For example, suppose the amplitude of the pulse picked up by L_2 should increase for any reason; if this happens, the pulse will drive the grid even more positive than normal. With a higher positive voltage on the grid, a greater number of electrons will be attracted to it. An increase in the number of electrons reaching the grid will mean that more electrons must flow through R_1. The voltage developed across R_1 depends upon two things: the size of the resistor and the number of electrons flowing through it. Therefore, if the number of electrons flowing through R_1 increases, the voltage developed across it will increase.

Notice the polarity of the voltage across R_1. The grid end of this resistor is negative, so this bias voltage tends to reduce the flow of current through the tube. Therefore, the increase in negative voltage across R_1 will subtract from the increase of positive voltage across L_2 so that the net drive voltage applied to the grid remains almost the same. Thus, even though something might cause the voltage developed in L_2 to increase, the tube will compensate for this change by developing an increased bias so that the burst of plate current flowing through the tube will remain essentially constant.

OSCILLATOR-AMPLIFIER

Up to this point, we have been considering the tube as a switch that closed at the proper instant to replenish the losses in the resonant circuit. We can also consider the tube as an amplifier that is amplifying part of its own output. For example, L_1 and L_2 in Fig. 6B are inductively coupled together. Part of the output produced across L_1 is coupled to L_2, where it is fed back into the input circuit. This signal fed into the input circuit is then amplified by the tube and fed to the resonant circuit L_1-C_1 in the output. The cycle then continues, with part of the output being coupled to L_2 and once again being fed back to the input. Thus, the oscillator can indeed be considered as an amplifier that feeds part of its own output signal back to the input, where it is amplified once again.

Of course, the signal fed back to the input must be of the proper phase to sustain oscillation. The signal must drive the grid in a positive direction when the plate current flowing through the tube should increase. Feedback of this type is called regenerative feedback. In some amplifiers a small amount of regenerative feedback is used to improve the gain of the amplifier. However, in an oscillator, enough regenerative feedback is used to start the stage oscillating, and to keep it oscillating.

OSCILLATOR FREQUENCY

You already know that the resonant frequency, f_0, of a circuit containing L and C is:

$$f_0 = \frac{1}{2\pi\sqrt{LC}}$$

The resonant frequency is also often expressed in terms of resonant angular frequency, ω_0:

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

where $\omega_0 = 2\pi \times f_0$. This expression comes from the fact that there are 2π radians in 360°. A radian is an angular measurement equal to approximately 57°. Since there are 2π radians in 360°, a vector rotating at f_0 Hertz travels through $2\pi \times f_0$ radians per second.

You might at first expect an L-C oscillator to operate at exactly ω_0, the resonant frequency of the L-C circuit. However, there is always some resistance in the circuit that affects the oscillator frequency. Furthermore, the plate resistance of the tube affects the oscillator frequency so that the actual frequency of the oscillator, ω, is:

$$\omega = \omega_0\left(1 + \frac{R}{R_p}\right)$$

where ω_0 is the angular resonant frequency of the L-C circuit, R represents the resistance in the resonant circuit, and R_p is the plate resistance of the tube.

In most oscillator stages the value of R will be small, because the Q of the oscillator coil will be high. At the same time, the plate resistance of the tube will be reasonably high so the term R/R_p will be small and ω will be almost equal to ω_0. However, the fact that R and R_p do enter into the frequency means that if either of these values change, the oscillator frequency will change. Thus, oscillator stability depends not only on keeping the values of L and C in the resonant circuit constant, but also the value of R and R_p must be kept constant.

OSCILLATOR STABILITY

One of the most important considerations in oscillator circuits is the stability of the oscillator – in other words, how stable the oscillator frequency is. The output frequency of a radio transmitter is controlled by the oscillator, and if the oscillator frequency does not remain constant, the transmitter output frequency will not be constant.

We have already pointed out that the oscillator frequency depends not only upon the inductance and capacitance in the resonant circuit, but also on other factors such as the resistance of the oscillator coil and the plate resistance of the oscillator tube. Now, let us consider each of these factors to see exactly what effect each has on the oscillator frequency.

Tank Inductance and Capacitance. The inductance in the oscillator tank circuit is made up of the inductance of the oscillator coil, plus any stray inductance in the circuit. The capacitance in the oscillator tank circuit is made up of the capacity connected across the oscillator coil plus any tube capacity that may be in parallel with the coil and capacitor, and the distributed wiring capacity in the circuit. The inductance in the circuit consists of the oscillator coil, the inductance of the leads connecting the coil to the tube and other parts in the circuit, and any inductance that other parts in the circuit may have. The capacity in the circuit consists of the capacity of the variable capacitor across the oscillator coil, the input capacitance of the tube, the stray wiring capacity in the circuit, plus any stray capacity the coil may have. This total inductance plus this total capacity are the major factors that determine the oscillator frequency.

When an oscillator is first turned on, the values of the inductance and capacitance in the tank circuit will usually change as the tube and other parts in the circuit heat. Therefore, the oscillator stability is usually measured in terms of the oscillator's ability to maintain a constant frequency after enough time has been allowed for the tube and parts to reach normal operating temperature. It is common practice in some transmitters to leave the oscillator on at all times to avoid any frequency drift during the warmup period. In some transmitters, the oscillator coil and capacitor are placed in an oven that is kept at a constant temperature by a thermostatically controlled heater to minimize changes in inductance and capacity due to temperature changes. In some oscillators, special temperature compensating capacitors are connected across the oscillator tank circuit to minimize frequency drift due to temperature changes. These capacitors usually have a negative temperature coefficient. This means that their capacity decreases as the temperature increases. By using a capacitor of this type with the correct temperature coefficient, it is possible to compensate for any increase in inductance or capacitance in other parts in the circuit as the temperature increases.

Changing a tube in the oscillator circuit can result in a change in oscillator frequency. The input capacity of the tube used in the oscillator circuit makes up part of the oscillator tank circuit. The input capacity of different tubes of the same type may vary appreciably, so putting a new tube in this or any other oscillator circuit may change the tank circuit capacitance, and hence the frequency. Therefore, if you replace the oscillator tube in a transmitter you should check the output frequency.

Tank Losses. Earlier we pointed out that the angular resonant frequency, ω_o, of the oscillator tank circuit is given by:

$$\omega_o = \frac{1}{\sqrt{LC}}$$

and at the same time, the actual frequency at which the oscillator oscillates is given by:

$$\omega = \omega_o\left(1 + \frac{R}{R_p}\right)$$

where R is the resistance in the tank circuit and R_p the plate resistance of the tube.

The term R represents the ac resistance of the tank circuit, and as such represents all the losses in the tank circuit. Thus, this term includes such factors as coil resistance and losses from the oscillator circuit due to loading of the circuit. Therefore, any change in the oscillator coil resistance will result in a change in oscillator frequency. Similarly, a change in the loading on the oscillator will result in a change in oscillator frequency. Thus, for maximum stability, the oscillator should be lightly loaded and the load on the oscillator must remain constant.

The Plate Resistance. Since the plate resistance of the tube enters into the frequency of the oscillator, any change in plate resistance will produce a change in the oscillator frequency. The plate resistance of the tube will change if either the plate or grid voltage is changed in the case of a triode, and if the grid or screen voltage (and to some extent the plate voltage) is changed in the case of a pentode. Thus, it is important that the voltages supplying the oscillator be kept constant. These voltages must also be free of hum, which actually is a changing voltage superimposed on the dc supply voltage, because the hum voltage could produce a constantly changing plate resistance which will result in a frequency-modulated signal being generated by the oscillator.

Changes in loading on the oscillator may affect the bias developed on the grid of the tube. When this happens, the grid voltage will change, causing the plate resistance and hence the frequency to shift.

Looking at the expression for the oscillator angular frequency, we see that the frequency is equal to the angular frequency of the tank circuit times one plus a fraction. Thus, the oscillator frequency will be higher than the resonant frequency of the tank circuit. Also, if the term R/R_p is small, which it usually is, the oscillator frequency will differ from the tank frequency by only a small percentage. However, at high frequencies, this small percentage or fraction can represent a great enough frequency change to cause concern. For example, if the resonant frequency of a tank circuit is 10 MHz and the value of R/R_p is only .01, .01 × 10,000 kHz represents 100 kHz so the oscillator frequency would be 10,101 kHz, or 10.1 MHz. If the value of R_p changed because of changing voltages on the tube, the oscillator might drift 50 kHz or more above or below this frequency.

TRANSISTOR OSCILLATORS

Of the three basic transistor configurations, the common emitter is the most frequently used in rf oscillator circuits. There are several reasons for this. The power, current and voltage gains of the common-emitter configuration are all greater than one, and the highest possible power gain can be had. Also in the common-emitter circuit, moderate input and output impedances make less power necessary for feedback. In the common-base configuration, low input and high output impedances inherent in the circuit cause a mismatch in the feedback circuit, producing greater losses and requiring more feedback. The current gain in the common-base circuit is less than one, even though voltage and power gains are greater than unity. A somewhat similar condition exists in a common collector circuit where high input and moderate output impedances exist. Voltage gain is less than unity, but current and power gains are greater than one.

Transistor oscillators may be designed to operate class A, B or C depending on the desired efficiency. Since rf oscillators are also amplifiers, bias supply and temperature stabilization are similar to rf amplifiers discussed in a previous lesson.

A combined voltage divider and feedback type biasing arrangement is often used because it helps produce oscillation and at the same time establishes a stable dc bias point. Emitter biasing with a bypass capacitor is also used, the operation being similar to grid leak biasing. Usually the amplitude is regulated by driving the transistor into saturation or by using special diode limiting circuits. Either shunt or series type collector feed may be used, the shunt type being preferred for greater output efficiency.

Frequency stability of the transistor oscillator is equivalent to, and sometimes greater than the electron tube oscillator. The use of lower voltages, currents and power, permits construction of better tank circuits. In particular, the low power used with transistors aids in stability due to the decrease in heat. One major disadvantage of transistors is their critical operating point. A slight bias change can cause a large shift in frequency.

The collector-to-emitter capacitance of the transistor also affects frequency stability. This internal capacitance will vary with changes in collector or emitter voltages and with temperature. In high frequency oscillators it is sometimes necessary to place a swamping capacitor across the collector to emitter leads. The total capacitance of the two in parallel results in a circuit which is less sensitive to voltage changes. The added capacitor may be a part of the tuned circuit.

Partial compensation of voltage changes may be obtained by use of a common supply. Since an increase in collector voltage tends to increase oscillator frequency and an increase in emitter voltage decreases oscillator frequency, the use of a common bias source for both the collector and emitter helps stabilize the frequency. By using a common bias source, a change in one is somewhat counteracted by the change in the other.

The three basic transistor configurations used for oscillators are shown in Fig. 7. Bias and feed arrangements are omitted for simplification. Although any of the three basic transistor configurations can be used, generally only two, the common-emitter and common-base, are used in actual practice. The common-emitter configuration offers the advantages of easily matched input and output impedances and its close parallel to the electron tube.

The major advantage of the common-base circuit is that at high frequencies collector-emitter capacitance helps feedback an in-phase voltage independently of tickler coil L_1, and oscillation is more easily obtained. In the common-emitter circuit, this capacitance feeds back an out-of-phase voltage which requires additional feedback from the tickler coil to

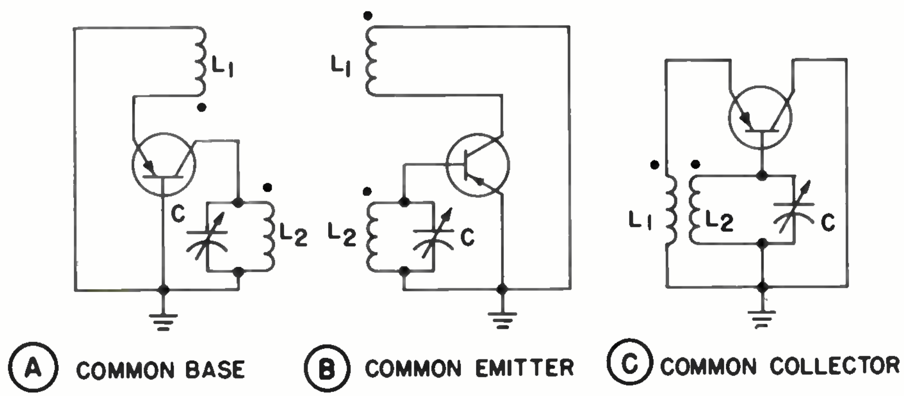


Fig. 7. Tickler coil oscillators.

overcome it. In both the common-base and common-emitter circuits, oscillation is easily sustained. This is the result of feedback provided by voltage induced through the mutual inductance of L_1 and L_2. In the common-collector circuit, the voltage gain is always less than unity, therefore feedback tends to be insufficient for stable oscillations at the lower frequencies. At the higher frequencies it is assisted by the collector-emitter capacitance. Sometimes, an external capacitor is added between the collector and emitter to give additional feedback.

Operation of the L-C circuit is similar to that of the electron tube circuit. As the oscillator is switched on, current flows through the transistor as determined by the biasing circuit. Initial current produces a feedback voltage between the collector and the emitter which is in-phase with the input circuit. As emitter current increases, collector current increases and additional feedback between L_1 and L_2 causes the emitter current to increase until saturation is reached. When saturation is reached, emitter current is no longer changing (increasing), and the induced feedback voltage is therefore reduced. At this time the collapsing field around the tank and tickler coils induce a reverse voltage into the emitter circuit which causes a decrease in the emitter current, thus causing a decreasing collector current. The decreasing current then induces a greater reverse voltage in the feedback loop driving the emitter current toward cutoff.

Although the emitter is cut off, a small reverse (leakage) current flows. This current has no effect on the operation of the circuit but it does represent a loss which lowers the efficiency. In this respect the transistor differs from the electron tube, which has zero current at cutoff.

The discharge of the tank capacitor through L_2 will cause the voltage applied to the emitter to rise from a reverse-bias to forward bias condition. Emitter and collector current start to increase and the cycle repeats itself.

The transistor oscillator circuit that most closely resembles the tuned-plate vacuum tube oscillator is the tuned-collector oscillator. This circuit is shown in Fig. 8.

Notice that in this circuit the resonant circuit consisting of C_1-L_1 is in the collector circuit of the transistor. L_2 is inductively coupled to L_1 so energy is fed from L_1 to L_2. The signal developed in L_2 is fed back to the base of the transistor.

In the operation of this oscillator, resistor R_1 and capacitor C_2 develop a bias voltage sufficient to cut off the transistor. The signal needed to overcome this cutoff bias is induced in L_2 and applied between the base and the emitter. Since this is a PNP transistor, the signal in L_2 must make the base negative and the emitter positive at the instant that a pulse of current is needed from the collector in order to sustain oscillation in the resonant circuit consisting of L_1 and C_1.

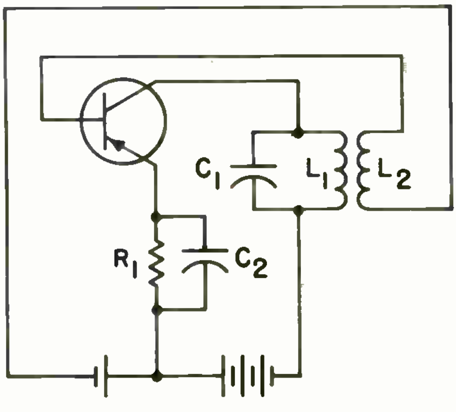


Fig. 8. A tuned-collector oscillator.

SELF-TEST QUESTIONS

(e) Describe the phase relationship between the plate current wave shape and the voltage wave shape developed across L_2 in Fig. 4C.

(f) What makes the oscillator in Fig. 6B self-regulating in regard to the amplitude of the grid signal?

(g) List three methods of reducing frequency drift due to changes in temperature in a resonant circuit.

(h) In high frequency transistor oscillators what is sometimes used to compensate for collector-to-emitter capacitance?

(i) In the oscillator circuit in Fig. 8, where is the bias developed and what component develops the signal that overcomes this bias?

Practical Oscillator Circuits

The oscillator circuits we have discussed up to this point weren't very practical. They were used to illustrate some of the basic characteristics of oscillators. Let us now look at some practical circuits actually found in communication equipment. These oscillators are grouped according to the type of resonant circuit used, inductance-capacitance (L-C) or resistance-capacitance (R-C).

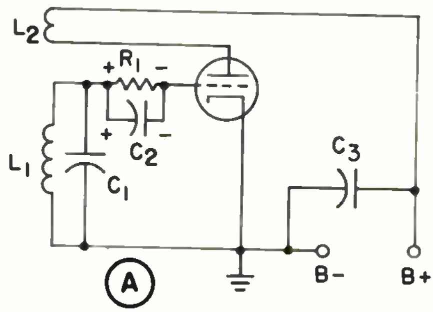
L-C OSCILLATORS

The L-C oscillators can be placed into one of two classifications: those using inductive feedback and those using capacitive feedback. The inductive feedback oscillator uses inductive coupling to return a portion of the output back to the input. The capacitive feedback circuit uses capacitive coupling to accomplish this. Although there is some difference in the circuitry involved, both types are L-C oscillators, and the net result is essentially the same.

OSCILLATORS USING INDUCTIVE FEEDBACK

One of the most important and most widely used oscillators in electronics work is the Hartley oscillator. This oscillator uses inductive feedback. The resonant circuit is placed in the grid circuit of the tube or the base circuit of the transistor instead of in the output circuit as in the case of the tuned-plate and tuned-collector oscillators. However, before we look at the Hartley oscillator let's look at another oscillator which will help you understand how the Hartley oscillator works. Let's first look at the tuned-grid oscillator.

Tuned-Grid Oscillator. Two versions of the tuned-grid oscillator are shown in Fig. 9. The circuits are basically the same; the only electrical difference is in the connection of the grid resistor R_1. In the circuit shown in Fig. 9A, R_1 is connected directly across the grid capacitor C_2, whereas in the circuit shown in Fig. 9B, R_1 is connected between the grid and the cathode of the tube. The action of R_1 is the same in both cases; it provides a path for the electrons striking the grid of the tube to get back to ground or the cathode



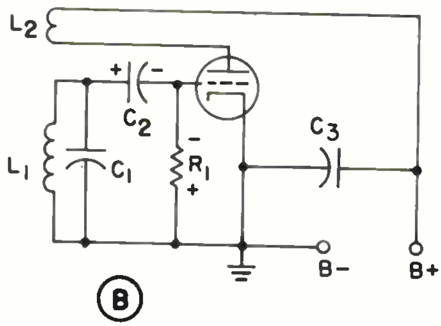


Fig. 9. Two versions of the tuned-grid oscillator.

of the tube. In the circuit of Fig. 9A, when C_2 discharges through R_1 to develop negative bias for the tube, there is no discharge through the tank circuit. In Fig. 9B, when C_2 discharges through R_1, the discharge current also flows through the tank circuit.

Actually, the biggest difference between this oscillator and the tuned-plate oscillator that you already studied is that the resonant circuit is in the grid circuit instead of the plate circuit. With this circuit, when the power is turned on, changes in plate current will set up a changing magnetic field about L_2. L_2 is inductively coupled to L_1 so the changing magnetic field about L_2 will induce a voltage in L_1. The induced voltage charges capacitor C_1, starting the oscillatory discharge in the tank circuit consisting of L_1-C_1. The voltage across C_1 becomes the grid voltage because the value of C_2 is large enough so that its reactance is so small at the frequency of oscillation, that the grid is, in effect, connected directly to C_1.

Now since the increasing plate current causes the end of C_1 that is connected to the grid through C_2 to swing in a positive direction, the grid of the tube is driven in a positive direction. Driving the grid positive produces two effects; it increases the plate current, causing C_1, and hence the grid, to be driven still further in a positive direction, and it causes grid current to flow, which charges C_2 with the polarity shown on the diagram.

Now if the plate current of the tube could keep on increasing indefinitely, the grid end of C_1 would be driven more and more positive. However, there is a limit to how high the plate current can become, because a balance will be reached between the positive voltage across C_1 and the negative voltage across C_2. When this happens, the plate current flowing through L_2 will no longer change. We will no longer have voltage induced in L_1, and C_1 will begin to discharge through L_1, setting up an oscillation in the L-C circuit. As soon as this happens, the positive voltage on the grid end of C_1 begins to disappear, and the plate current will be cut off by the negative voltage across C_2. The L-C circuit is now free to oscillate as though the tube were removed from the circuit. C_2 meanwhile starts to discharge through R_1, setting up the voltage drop across it as shown on the diagram. During the next half cycle when the voltage on the grid end of C_1 again becomes positive, it will drive the grid in a positive direction enough to let some plate current flow through the tube; this will result in a change in the field about L_2 which will induce a voltage in L_1 which drives the capacitor and the grid voltage still further in a positive direction.

The important point to remember about this oscillator is that the energy needed to sustain the oscillation in the tank circuit, consisting of L_1 and C_1, is inductively coupled to L_1 from L_2. This energy comes from the plate of the tube in the form of bursts of plate current which produce a changing magnetic field about L_2. These bursts of current are the result of the grid of the tube being driven positive by the voltage across C_1 swinging positive once each cycle.

The Hartley Oscillator. Two Hartley oscillators are shown in Fig. 10. The circuit shown in Fig. 10A uses a vacuum tube, whereas the one shown in Fig. 10B uses a transistor. Although the operation of the two circuits is so similar that if you understand one, you will understand the other, we will go through both circuits in considerable detail.

Notice the difference between the

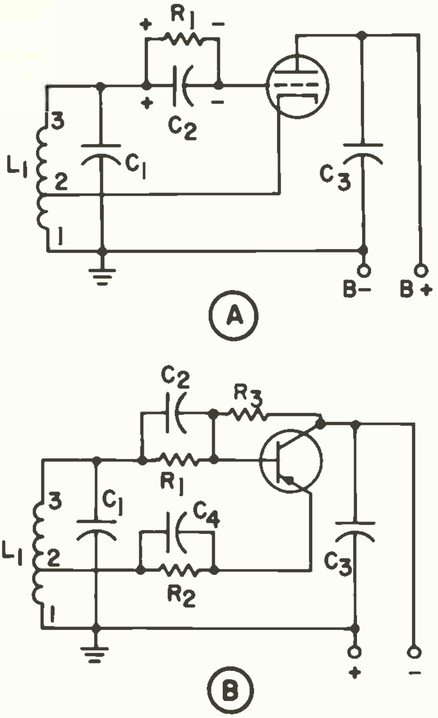


Fig. 10. Typical Hartley oscillators; a vacuum tube Hartley oscillator is shown at A, and a transistorized one at B.

Hartley oscillators and the tuned-grid oscillators. The tuned-grid oscillator has two coils, whereas the Hartley oscillator uses only a single tapped coil. In the circuit shown in Fig. 10A, the cathode of the tube is connected to the tap, and in the circuit shown in Fig. 10B, the emitter of the transistor is connected to the tap.

In the circuit shown in Fig. 10A, when the plate current starts to flow through the tube, current must flow through the lower half of the coil between terminals 1 and 2. Since the entire coil is wound on a single form, all the various turns of the coil are inductively coupled together. Therefore the increasing plate current flowing between terminals 1 and 2 will produce a changing magnetic field which will induce a voltage in the portion of the coil between terminals 2 and 3. This voltage will charge capacitor C_1 with the polarity such that the end of the capacitor that connects to the junction of C_2 and R_1 is positive. Again since the value of C_2 is chosen so that its reactance is practically zero at the oscillation frequency, the grid of the tube is in effect connected directly to C_1. This means that the increase in plate current will drive the grid of the tube in a positive direction, causing a still stronger burst of current through the coil. This in turn causes still higher induced voltage between terminals 2 and 3 which again charges capacitor C_1 still further. At the same time when the grid is driven positive, it will attract electrons, and C_2 will be charged with the polarity shown.

As in the tuned-grid oscillator, the point is eventually reached where there is a balance between the positive voltage applied to the grid by C_1 and the negative voltage applied to the grid across C_2 and R_1 so that there is no further increase in plate current. This means that the magnetic field produced by the current flowing between terminals 1 and 2 becomes constant and no further voltage will be induced in the coil. C_1 starts to discharge through the coil, and the oscillating cycle is started. Furthermore, the positive voltage on the end of C_1 that connects to the grid of the tube through C_2 disappears, and the tube stops conducting.

Again, the tube will be biased beyond cutoff by the discharge of C_2 through R_1. These electrons charge C_2 during the portion of the cycle when the grid is conducting. When grid current stops flowing, C_2 will discharge through R_1, setting up a voltage drop across it such that the grid end is negative. This voltage across R_1 maintains the bias on the grid of the oscillator tube.

In some cases you will see slight variations of the Hartley oscillator circuit. In some instances, R_1 may be connected between the grid and cathode or from the grid of the tube directly to ground. In another variation the cathode connects directly to ground, R_1 connects between the grid of the tube and ground, and then the plate of the tube connects directly back to terminal 1 of the oscillator coil. The B+ voltage is then applied to terminal 2 of the coil. This is simply a modification of the Hartley oscillator circuit; it works in exactly the same way as the Hartley oscillator shown in Fig. 10A.

In the circuit shown in Fig. 10B we have a PNP transistor. When holes begin to travel from the emitter to the collector, electrons will flow from the emitter through R_2 to terminal 2 on the coil. From terminal 2 they will flow through the coil to terminal 1 and back to the positive terminal of the battery. The electrons, in flowing through the coil from terminal 2 to terminal 1, will build up a field about this part of the coil. This field will be a changing field as the current builds up, and this will induce a voltage in the portion of the coil between terminals 2 and 3. This induced voltage will charge C_1 with the polarity such that the end connecting to terminal 3 of the coil is negative and the other end is positive. This negative voltage on one end of C_1 will be applied to the base of the transistor through capacitor C_2 because C_2 has a low reactance at the frequency of oscillation. The negative voltage on the base of the transistor will increase the forward bias across the emitter-base junction, causing an increase in the number of holes flowing from the emitter to the collector. This causes a still further increase in the electron movement from terminal 2 to terminal 1 of the coil, causing the base of the transistor to be driven still further in a negative direction.

In this circuit when the number of holes flowing from the emitter to the collector increases, terminal 3 of the coil will be driven in a negative direction, and when the holes flowing from the emitter to the collector decrease, terminal 3 will be driven in a positive direction. Remember that in a PNP transistor, driving the base in a negative direction causes the holes moving through the transistor to increase, whereas driving it in a positive direction causes the number of holes flowing from the emitter to the collector to decrease. The burst of hole movement through the transistor causes the electron movement from terminal 2 to terminal 1 of coil L_1 to flow through the coil in burst, and this burst of energy makes up for any losses in the resonant circuit consisting of L_1 and C_1.

It is interesting to note the similarity between the circuits shown in Fig. 10A and Fig. 10B. Although we have a vacuum tube used in one circuit and a transistor in the other, there is a great deal of similarity between the two circuits and the way they work. In each case we have energy lost in the resonant circuit being replaced by bursts of energy; from the tube in one case and from the transistor in the other case. Also notice that the energy is fed across only part of the coil in each case, but the voltage induced in the entire coil is enough to set up a current flow that will replace the capacitor charge that is lost because of resistance or other losses in the resonant circuit.

OSCILLATORS USING CAPACITIVE FEEDBACK

There are a number of different oscillator circuits in which capacitive feedback rather than inductive feedback (as in the preceding examples) is used to sustain oscillation. Let's look at some of them now.

Colpitts Oscillator. Perhaps the most important of the oscillators using capacitive feedback is the Colpitts oscillator shown in Fig. 11. The one in Fig. 11A uses a vacuum tube while the one in Fig. 11B uses a transistor.

The operation of the two oscillators is quite similar. When the equipment is first turned on, current flows through L_2, which is the small rf choke used to complete the cathode circuit in Fig. 11A and the emitter circuit in Fig. 11B. Current flowing through the coil produces a voltage drop across the coil, and this charges capacitor C_2. The charge on capacitor C_2 will start an oscillation in the tank circuit, which consists of coil L_1 and two capacitors, C_1 and C_2. Remember that when we have two capacitors connected in series they will act like one capacitor insofar as the coil is concerned, and the circuit will start to oscillate. The voltage developed across C_1 is the feedback voltage. It is applied between the grid and the cathode in the circuit shown in A and between the emitter and the base of the circuit shown in B. When this voltage swings in a direction that makes the end of C_1 connected to C_3 positive and the other end negative, it will increase the current flowing through the tube or transistor, causing an increase in current flow through L_2, which charges C_2 still further. When the polarity of the voltage across C_1 reverses, the voltage will oppose the current flow and in Fig. 11A simply add to the bias between the grid and the cathode, reducing the plate current to zero; or in Fig. 11B, put a reverse bias across the emitter-base junction, reducing the current flowing through the transistor to practically zero.

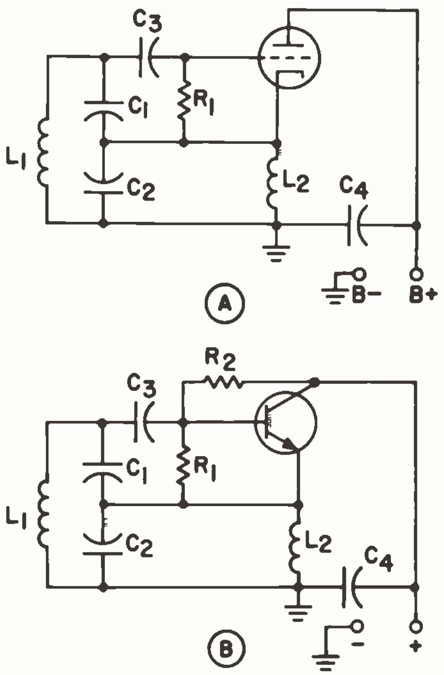


Fig. 11. Two Colpitts oscillators.

The amount of feedback voltage applied to the input of the circuit depends upon the ratio of C_1 to C_2. If C_1 is large compared to C_2, the reactance of C_1 will be low and the reactance of C_2 will be high. Most of the voltage developed across the capacitors will be developed across the higher reactance, in this case C_2. This means that the feedback voltage applied to the input will be low. However, if C_1 is small compared to C_2, the reactance of C_1 will be high compared to the reactance of C_2 and the feedback voltage supplied to the input of the circuit will be high.

The ratio of C_1 and C_2 can be altered to provide the required feedback to the input circuit to sustain oscillation. If the value of C_1 is increased and the value of C_2 decreased by the correct amount, the total capacity in the circuit formed as the result of two capacitors in series remains the same, and hence the resonant circuit of the oscillator does not change.

In some Colpitts oscillators an additional capacitor is connected directly across L_1. This is done to provide some means of changing the resonant frequency so we can vary the frequency at which the oscillator oscillates. It is impractical to try to vary both C_1 and C_2 at the same time, but an additional capacitor placed directly across the coil can be varied, and this will change the resonant frequency of the oscillator. At the same time, since C_1 and C_2 will still form a voltage divider, part of the total voltage developed across the two capacitors in series is fed back to the input circuit; this part can still be controlled by the proper selection of C_1 and C_2.

There are a number of variations of the Colpitts oscillator circuit. It is sometimes found in radio transmitting equipment that must be designed so that its frequency can be varied. The Colpitts oscillator can be designed with excellent frequency stability. By this we mean that once the oscillator is adjusted to operate at a certain frequency it will not drift from that frequency very much. Some oscillators, on the other hand, do not have good frequency stability and will drift appreciably.

Another variation of the Colpitts oscillator circuit is shown in Fig. 12. Here we have the capacitor C_1 connected across L_1 in addition to the voltage divider capacitors C_2 and C_3.

Notice that in this circuit the plate of the tube is fed back directly to L_1, C_1, and C_3 and that the choke coil L_2 has been moved from the cathode circuit to the plate circuit of the tube. The cathode in this oscillator circuit is connected directly to ground.

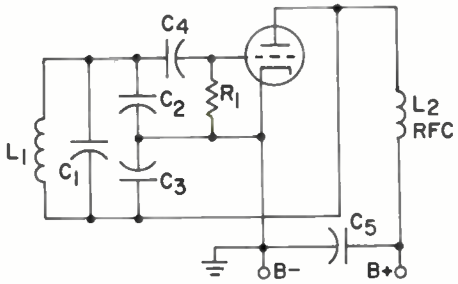
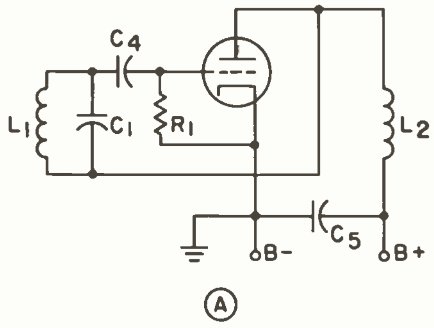


Fig. 12. A variation of the Colpitts oscillator.

In this oscillator when the plate current increases there will be a voltage developed across the rf choke, L_2, in the plate circuit of the tube, and this voltage will charge C_3. Once this capacitor is charged, oscillation starts in the circuit just as in the Colpitts oscillators shown in Fig.11.

The Ultra-Audion Oscillator. Another oscillator that uses capacitive feedback is the ultra-audion oscillator shown in Fig. 13A. When this type of oscillator was first developed, it was considered as a new type of oscillator. However, with careful analysis, we can see that it is actually a Colpitts oscillator, practically identical to the oscillator shown in Fig. 12. We have used the same designations to identify the parts in the circuits shown in Figs. 12 and 13. As you can see, the parts are all the same except for C_2 and C_3, which Fig. 13A does not seem to have. However, in Fig. 13B we have shown these two capacitors. C_2 is the grid-to-cathode capacity of the tube, and C_3 is the plate-to-cathode capacity of the tube. When we consider these two capacities, we have a capacitive voltage-divider network just like the one in Fig. 11. C_2 in Fig. 13B is between the grid and the cathode of the tube. Notice that C_2 in Fig. 12 also is in effect connected between the grid and the cathode of the



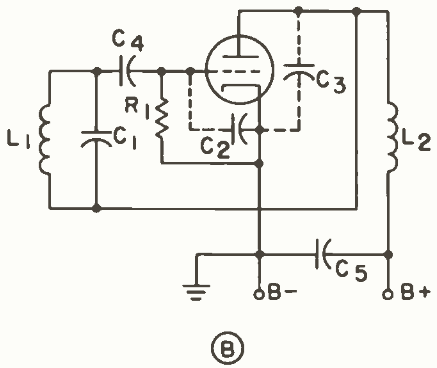


Fig. 13. The ultra-audion oscillator.

tube. C_2 and C_3 are in series in both circuits and they are connected across the tank circuit. C_3 connects directly to the lead going from the plate of the tube to one side of the resonant circuit, and C_2 connects through capacitor C_4 to the resonant circuit. Therefore, this oscillator is simply another form of Colpitts oscillator.

This type of circuit is frequently used in the vhf oscillators in the tuners of television receivers. Of course, it is usually shown in the schematic in the form shown in Fig. 13A. Manufacturers seldom draw in the distributed capacities; they expect the technician to know enough about oscillator circuits to recognize this as the ultra-audion oscillator and to know that this is simply a modified form of Colpitts oscillator.

The Electron-Coupled Oscillator. So far all of the vacuum tube oscillators we have discussed have been triode oscillators. These oscillators are widely used in receiving equipment and are sometimes found in transmitters and other rf power-generating equipment. However, they have some disadvantages, one of which is the direct coupling between the output and input circuits through the grid-to-plate capacity of the tube. Loading the output circuit of the oscillator has an effect on the input circuit and hence often results in an appreciable shift in the frequency at which the oscillator is oscillating. The net result is that triode oscillators are not stable enough for some purposes.

An oscillator that overcomes this difficulty is the electron-coupled oscillator. In this circuit a tetrode or a pentode tube is used so that the only coupling between the input and output circuits is in the electron stream flowing from the cathode to the plate of the tube. Schematic diagrams of two electron-coupled oscillators are shown in Fig. 14. The circuit shown in Fig. 14A is for an electron-coupled Hartley oscillator and the one shown in Fig. 14B is for an electron-coupled Colpitts oscillator.

The operation of these oscillators is similar to the operation of the triode oscillators, except that in the electron-coupled oscillator the screen grid of the tube acts like the plate of a triode tube. In other words, insofar as the oscillator action is concerned, we have three elements in the tube to be concerned about, the cathode, the grid and the screen grid. The screen grid acts like the plate of the oscillator tube. The oscillation is set up in this circuit by these three tube elements. However, the electron stream flows from the cathode of the tube to the plate of

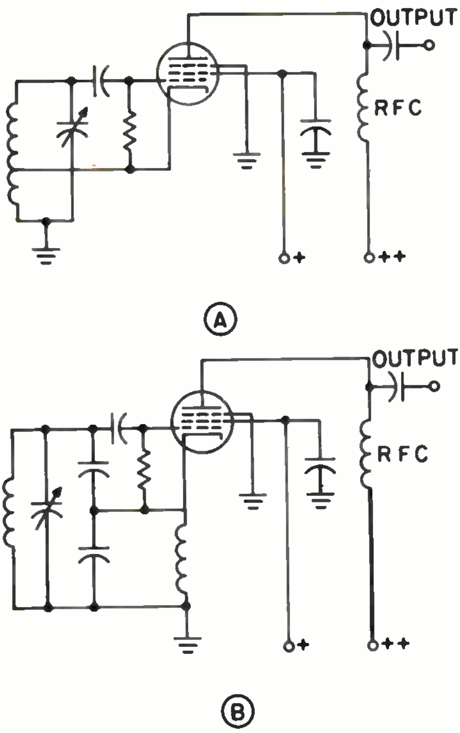


Fig. 14. Electron-coupled oscillators; the circuit at A is a Hartley oscillator.

the tube, and it is controlled by the grid. Since the grid is biased beyond cutoff during most of the rf cycle, but receives a strong positive pulse during a portion of the cycle, the plate current flowing through the tube flows in the form of pulses, which flow only when the grid of the tube is driven in a positive direction.

In the plate circuit of both oscillators we have shown an rf choke and a capacitor through which the oscillator can be coupled to the following stage. In some electron-coupled oscillators you will find a resonant circuit in the plate circuit of the tube instead of the rf chokes shown in Fig. 14. In some oscillators, this resonant circuit will be tuned to the same frequency as the resonant circuit in the grid circuit of the oscillator, but in other oscillators you'll find that it is tuned to a frequency equal to twice or three times the frequency to which the resonant circuit in the grid circuit of the tube is tuned. If we tune the resonant circuit in the plate circuit of the tube to twice the frequency of the resonant circuit in the grid circuit of the tube, we will have a frequency doubler. The resonant circuit in the plate circuit of the tube is set into oscillation by the burst of plate current flowing through the tube. However, since the resonant frequency of this circuit is twice the resonant frequency of the input circuit, the circuit in the plate circuit begins to oscillate at a frequency equal to twice the frequency being generated in the grid circuit. The oscillation in the plate circuit therefore goes through two complete cycles before a second pulse is received from the plate of the tube. This is called a frequency-doubler circuit. If the resonant circuit in the plate circuit is tuned to three times the frequency of the input circuit, the oscillation set up in the plate circuit will go through three complete cycles before it receives an additional pulse from the plate of the tube. This is called a frequency tripler. Now, you might expect this to result in a damped wave, with the amplitude of the cycles which do not receive a reinforcing pulse from the plate of the tube being considerably less than the amplitude of the particular cycles during which the pulse is received. Of course, there will be some loss in the resonant circuit and there will be some change in amplitude. However, by the use of a high Q resonant circuit in the plate circuit, the change in amplitude that occurs each cycle is very small and for all practical purposes all the cycles of the sine wave produced in the resonant circuit will have essentially the same amplitude.

R-C OSCILLATORS

Resistance-capacitance (R-C) oscillators use the charging and discharging action of a capacitor and a resistor in the feedback path to cause oscillation. The R-C oscillator is used in audio and low rf ranges. They offer an inexpensive method of obtaining a fairly stable sine wave within the range of their operation. Although many variations of the R-C oscillators exist, there are only two basic types, the phase shift and the bridge.

The phase shift uses a series of R-C phase shifting circuits between the output and the input to produce a feedback in-phase with the input. The bridge circuit usually uses an additional tube or transistor to obtain the phase shift and a bridge type circuit to control the feedback at the proper frequency.

Phase Shift. The phase shift oscillators in Fig. 15 consist of a conventional amplifier and a phase shift feedback circuit. As in L-C oscillators, the feedback must be positive. In the circuits in Fig. 15A and B, the output signal will be 180° out-of-phase with the input signal. Therefore the phase shift network must shift the phase of the feedback signal 180 degrees. This phase shift is provided by C_1, C_2, C_3 and R_1, R_2, R_3. One section, C_1-R_1, of the feedback loop is shown separately in Fig. 15C. The impedance of the circuits is capacitive and the feedback voltage, e_f, produces a current, i, through C_1-R_1 which leads e_f. The angle of the current lead is determined by the ratio of the reactance of C_1 to the resistance R_1. As the value of R_1 is reduced, the circuit will become more capacitive and more current will lead the voltage up to a maximum of 90°. If resistance is increased the current lead will decrease.

By increasing the number of R-C networks, the losses of the total feedback circuit will decrease and a less amount of phase shift will occur in each section of the feedback loop. For this reason some oscillators use five or six R-C sections. In Fig. 15 each section produces a 60 degree phase shift. The reactance of the capacitors is inversely proportional to the frequency; therefore, only one frequency

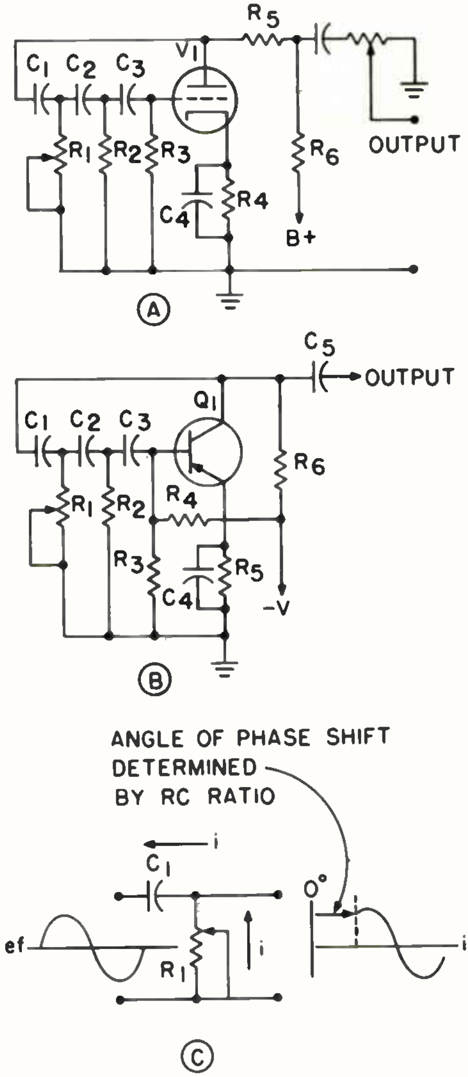


Fig. 15. Phase shift oscillators.

will pass through the feedback loop. Normally the output is fixed in frequency due to the constant value of the capacitors. A variable output may be obtained by using ganged variable capacitors or resistors since an increase in the value of either R or C will decrease the frequency.

Let us examine the operation of the circuit in Fig. 15B. R_3 and R_4 establish the base bias while R_5 and C_4 furnish thermal stabilization and furnish an rf ground to the emitter. R_6 is the load across which the output is taken and C_5 is the coupling capacitor.

Once power is supplied, operation is started by any random noise in the power source or the transistor. This noise causes a change in base current which causes a large change in collector current. This change in collector current develops an output voltage across R_6 which is 180° out-of-phase with the original change in base voltage. Part of the signal developed across R_6 is returned to the base shifted 180 degrees by the R-C network. The shift in phase through the R-C network causes the feedback to aid the output, resulting in positive feedback.

With fixed values of R and C, the 180 degree phase shift occurs at only one frequency, therefore, the output is a sine wave of fixed frequency. At all other frequencies, the reactance either increases or decreases, causing a variation in the phase relationship resulting in degenerative feedback.

The Bridge Oscillator. The Wien-bridge oscillators shown in Fig. 16 consist of low

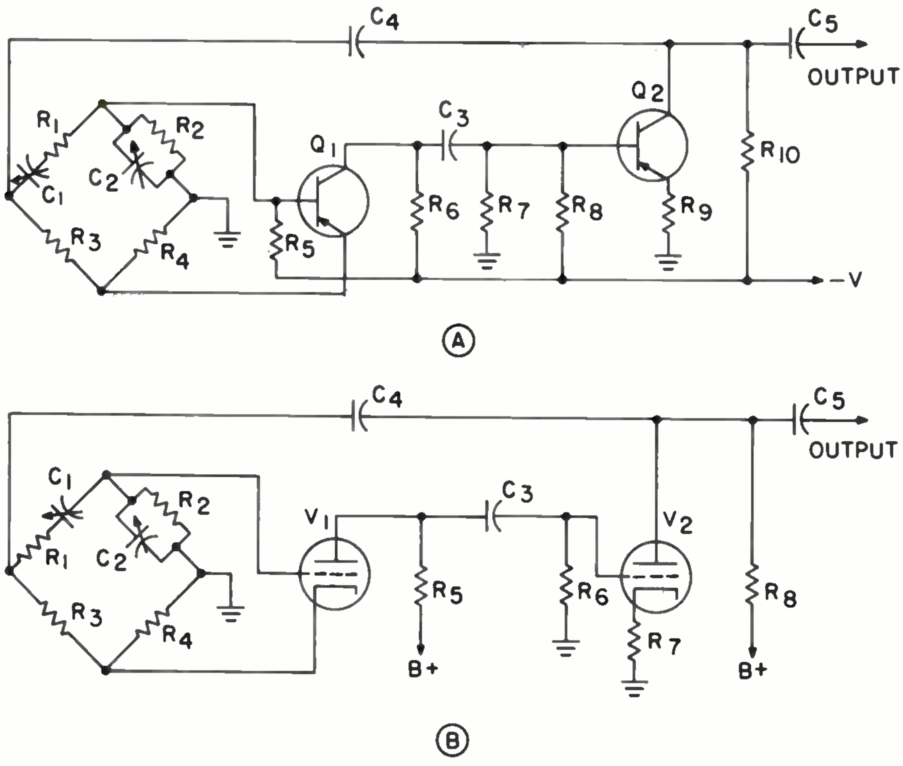


Fig. 16. Wien-bridge oscillators.

gain amplifier stages and a resistive-capacitive bridge circuit. The feedback is taken from the second amplifier and returned to the bridge circuit. Since there are two 180 degree phase shifts between the input of the first amplifier and the output of the second amplifier, the feedback will be positive. The resistive-reactive bridge circuit is designed to be balanced at the operating frequency; therefore, only feedback of the desired frequency reaches the input of the first stage. Let's examine Fig. 16A and see how this works.

The voltage divider, R_2 and R_5 supply base bias for Q_1 and R_7, and R_8 furnishes bias for Q_2. Thermal stabilization is provided by R_4 and R_9. R_3 and R_4 form a resistive leg of the bridge which is in shunt with the output of Q_2. A portion of this output coupled by C_4 and R_3 appears across R_4 as negative feedback. Since R_4 is not frequency sensitive, the negative feedback is constant regardless of the frequency of the output. At frequencies other than the operating frequency, the negative feedback furnished by R_4 will prevent oscillation. At the frequency of operation, which is controlled by the bridge reactive leg of R_1-C_1 and R_2-C_2, the positive feedback to the base of Q_1 is maximum. This in-phase feedback signal is applied to the base of Q_1 and is of sufficient amplitude to overcome the negative feedback across R_4. The total feedback is therefore positive at the operating frequency.

The amplified output of Q_1 is coupled to the base of Q_2 by C_3 and R_7. Q_2 further amplifies the signal and the voltage developed across R_{10} is coupled to the output by C_5, and a feedback signal is coupled through C_4 to Q_1.

We have not covered all of the various L-C and R-C oscillator circuits you are likely to encounter. There are many different variations of the circuits we have discussed, and some entirely different circuits. However, most of the circuits you are likely to encounter will be one of the circuits we have discussed in this section of the lesson or a variation of one of these circuits. If you come across a circuit you do not recognize immediately, first determine whether it is R-C or L-C and, in the latter case, whether capacitive or inductive feedback is used. Once you have decided on the type of feedback that is used, you should be able to figure out how the oscillator circuit works if you keep in mind that its operation is similar to one oscillator we have described in this lesson.

SELF-TEST QUESTIONS

(j) When plate current increases in the tuned grid oscillator (Fig. 9B), what happens to the voltage at the junction of R_1 and C_2?

(k) Where is the feedback voltage developed in the Colpitts oscillator (Fig. 11B)?

(l) The ultra-audion oscillator is a modification of what type oscillator?

(m) How is the 180° phase shift in feedback voltage accomplished in the phase shift oscillator?

(n) What component(s) develops the feedback voltage in the Wien-Bridge oscillator (Fig. 16)?

Crystal Oscillators

Although L-C oscillator circuits have been developed that have very good frequency stability characteristics, the master oscillators in most transmitters still use crystals. Better frequency stability is the main reason for using crystals instead of L-C resonant circuits in master oscillators. The frequency tolerance allowed by the FCC for most transmitting services is very small. The tolerance of a transmitter operating in the broadcast band, for example, is a frequency deviation of only 20 Hertz from the assigned frequency. Some services are permitted slightly more frequency tolerance, but in all cases, the tolerance is rather strict. This restriction is necessary to keep the large number of stations operating in the frequency spectrum from interfering with each other.

Several types of materials can be used for crystals. These include quartz, Rochelle salts, and tourmaline. The most often used crystal material for generating radio frequency signals is quartz. Rochelle salts work better in low-frequency applications, such as in loudspeakers and microphones. Tourmaline will work as well as quartz, but because it is a semiprecious stone, it is more expensive.

The assembly usually referred to as a crystal is composed of a small piece of crystal material mounted in a holder. The crystal material is in the form of a small slab or wafer cut from a larger crystal. The way in which the wafer is cut from the natural crystal determines many of its electrical characteristics. In this section, we will find out how the crystal works, how the crystal is used in oscillator circuits, and finally we will discuss some of the most-used crystal oscillator circuits.

THE PIEZOELECTRIC EFFECT

A quartz crystal exhibits a property called the piezoelectric (pronounced pi-ee-zo) effect when it is compressed mechanically or when a current is applied to it. To illustrate this effect, suppose we have a small crystal wafer with leads attached to its two surfaces. If we squeeze the wafer or bend it in some way, a voltage will appear between the leads. If, on the other hand, we apply a small voltage across the leads, the crystal slab will bend, expand, or contract, depending on the polarity of the applied voltage and the crystal type. These two effects are used in many types of electronic equipment.

A crystal wafer has a definite mechanical resonant frequency at which it will vibrate most readily. This resonant frequency is determined by the physical dimensions of the wafer, particularly the thickness. The thinner the wafer, the higher the resonant frequency. Thus, when an ac voltage, whose frequency is near the crystal resonant frequency, is applied, the crystal will vibrate the greatest amount and produce the greatest output. Because of the large amount of vibration at the resonant frequency, there is a limit to how thin the wafer can be before it becomes too fragile for practical use.

CRYSTAL CUTS

The natural quartz crystal, as shown in Fig. 17, is said to have 3 major axes at right angles to each other. These axes, called the X, Y, and Z axes, are shown in the illustration. The X axis is called the electrical axis; the Y axis is called the mechanical axis; and the Z axis is called the optical axis. The way that the crystal wafers are cut with respect to these axes determines many electrical characteristics, including the frequency range in which the crystal will oscillate and the amount that the resonant frequency will change with changes in the temperature (the temperature coefficient).

If the crystal wafer shown by the shaded section in Fig. 17 is cut so that its face is perpendicular to the Y axis, it is called a Y-cut crystal. Similarly, if its face is perpendicular to the X axis it is called an X-cut crystal. Crystal wafers cut perpendicular to the Z axis have no piezoelectric properties.

Crystal cuts can be made at different angles to the major axes to produce crystals having slightly different electrical characteristics. The type of cut depends on the purpose for which it is to be used.

One of the most important characteristics of a quartz crystal used in oscillator circuits is the amount the crystal frequency varies with variations in temperature. We call this the temperature coefficient of the crystal. Y-cut crystals have a range of about −25 to +100 Hertz/°C/MHz (Hertz per degree centigrade per MHz). In other words, a one-degree centigrade increase in temperature may cause the frequency to decrease as much as 25 Hertz, or increase by as much as 100 Hertz for each MHz of the frequency for which the crystal is ground. To take an extreme case, the frequency of a 4-MHz crystal might increase by 400 Hertz for each degree of temperature change. Because of this rather high temperature coefficient and also because Y-cut crystals have a tendency to oscillate at a second frequency near the design frequency, they are seldom used.

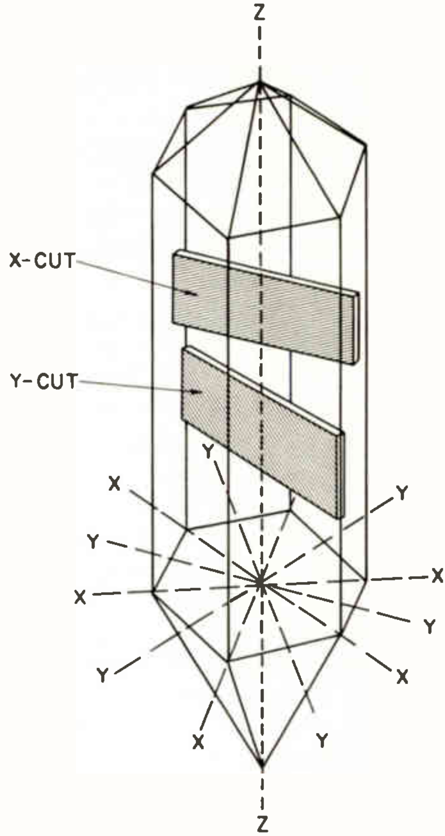


Fig. 17. Major axes quartz crystal.

X-cut crystals have a range of about −10 to −25 Hertz/°C/MHz. The minus sign means that the crystal frequency decreases with an increase in temperature. We call this a negative temperature coefficient. Even though the frequency variation is less, close temperature control is still required to keep the crystal oscillating at the correct frequency.

As an example of how temperature affects the crystal frequency, consider an X-cut crystal that has an operating frequency of 4650 kHz at 50° C. If it has a temperature coefficient of −20 Hertz per degree centigrade per MHz, let's determine what its operating frequency would be if the temperature changes ten degrees.

$$4650 \text{ kHz} = 4.65 \text{ MHz}$$

therefore the change in frequency per degree centigrade is:

$$4.65 \times 20 = 93.00 \text{ Hertz}$$

The change in frequency with a temperature change of ten degrees will be:

$$93 \times 10 = 930 \text{ Hertz}$$
$$\text{and } 930 \text{ Hertz} = .930 \text{ kHz}$$

Thus we see that with a temperature change of only 10 degrees, the frequency will change almost 1 kHz. With a 20-degree change, the frequency change would be almost 2 kHz. If the temperature increases 10°, the frequency will decrease to about 4649 kHz (4649.07 kHz), and if the temperature drops 10°, the frequency will increase to almost 4651 kHz (4650.93 kHz). Where there are many stations operating on frequencies close together, this shift in frequency could be enough to cause interference.

We can reduce the temperature effect on the resonant frequency of the crystal by cutting the crystals at angles to the major axes. Examples of such crystal cuts, called the AT-cut and the BT-cut, are shown in Fig. 18. These crystals are really Y-cut crystals with the face of the wafer at an angle of about 39° to the Z axis instead of parallel to it in the case of the AT-cut crystal, and about 45° to the Z axis in the case of the BT-cut crystal. Notice that the angle of the AT-cut crystal is opposite to that of the BT-cut crystal with respect to the Z axis. The temperature coefficient of these cuts is about ±2 parts per million at a temperature of 40°C to 50°C. Thus, the angle cut practically cancels the effects of temperature variation on the frequency of oscillation if it is operated within the temperature range.

Other angular cuts can be made to get other electrical characteristics and effects. Some examples are the CT and DT cuts used for lower frequency operation below 500 kHz. A CT-cut crystal is cut perpendicular to the BT-cut crystal, and the DT cut crystal is cut perpendicular to the AT-cut crystal. Another cut is the GT cut.

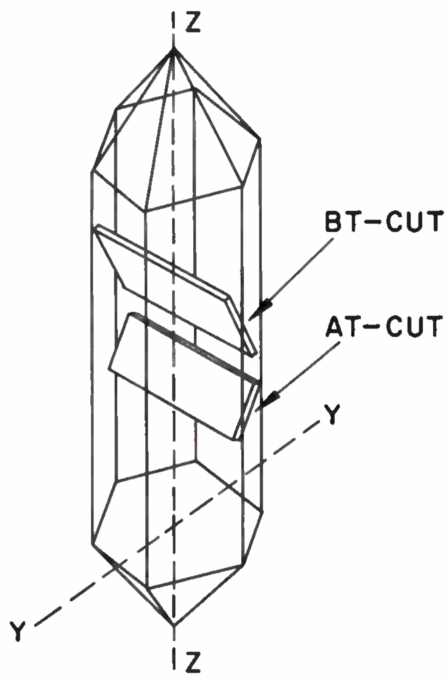


Fig. 18. AT-cut and BT-cut crystals are cut on an angle to the Z axis as shown.

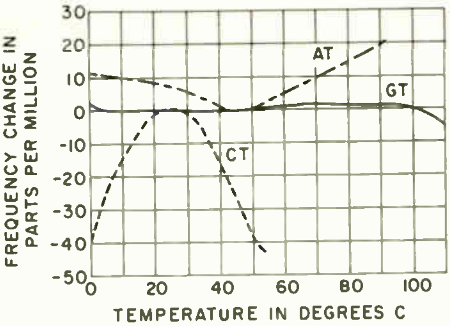


Fig. 19. Frequency changes of different crystal cuts with changes in temperature.

This crystal is cut on an angle of about 45° to either the CT- or DT-cut crystals.

Fig. 19 illustrates the variation of the resonant frequency with temperature for various crystal cuts. Notice that the frequency of the GT-cut crystal is practically constant from 0° to about 100° This cut, therefore, is the best to use in equipment that will be subjected to wide temperature variations. The other cuts have zero temperature coefficients at or near specific temperatures.

CRYSTAL HOLDERS

After the crystal has been ground, or etched by means of chemicals, until it is of the proper thickness, it is placed in a holder. The holder is then sealed so that no air, dirt, or oil can reach the crystal wafer. These can cause erratic operation of the crystal. A crystal in its hermetically sealed holder is shown in Fig. 20.

There are two ways of mounting the crystal in the holder. In one way, a very thin film of metal is formed directly on the surface of the crystal by spraying or firing. The metal can be either silver, gold, or aluminum. The crystal is supported by flexible wires fastened to this metallic film with solder having a low melting point. The leads are attached to a node or non-oscillating portion of the crystal. The node can be either in the center of the crystal face or at an edge where inhibiting the vibrations will do no harm.

Another method of mounting the crystal in a holder is to arrange it so that it is pressed between two metal plates. The metal plates make electrical contact with the crystal surfaces. Another type uses an air gap .001 to .005 inch thick between the crystal and the upper plate. This air gap produces a damping effect on the amount of crystal vibration. Often, however, the crystal holder is evacuated so that the damping will be reduced.

Regardless of the type of mounting used, the holder itself is designed to keep the crystal free of grit, dirt, and oil film. Even a speck of dirt or a greasy film can change the characteristics of the crystal. Therefore, the crystal holder should never be opened, nor should the crystal be handled. If it is ever necessary to take one apart, the crystal may be cleaned with a

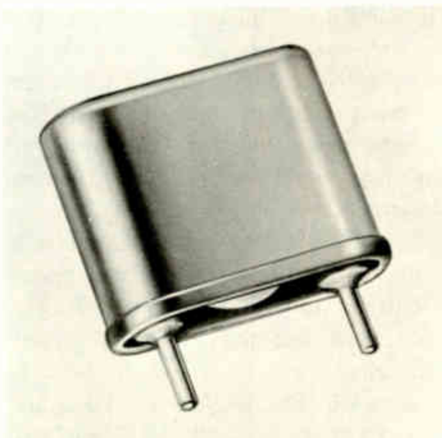


Fig. 20. A quartz crystal in its holder.

good nonflammable commercial cleaner. The crystals should be handled with clean, lint-free cloths, not with bare fingers (bare fingers leave traces of perspiration on any object they touch).

Equivalent Circuit. The symbol in Fig. 21A is used in schematic diagrams to represent the crystal in its holder. The equivalent electrical circuit of the crystal and holder assembly is shown in Fig. 21B. As you can see, we have here a series-parallel circuit composed of the series components L, R, and C_1 shunted by capacity C_2. The crystal, therefore, acts electrically as an L-C circuit; as an inductance at frequencies above the resonant frequency and as a capacitance at frequencies below resonance. The apparent inductance L is due to the mass of the crystal, resistance R is the result of internal mechanical losses, and capacity C_1 is the stiffness (piezoelectric properties) of the crystal. Capacity C_2 is the capacity between the electrode plates, with the quartz crystal acting as the dielectric.

Since the crystal acts as an electrical resonant circuit, we would expect it to be frequency selective; that is, it will oscillate more vigorously at its natural resonant frequency than at any other frequency. This is true. When an ac voltage is applied to its electrodes, the crystal generates an alternating potential of its own.

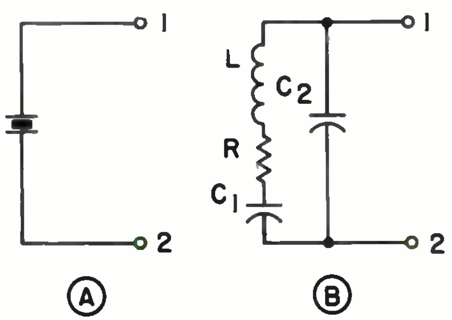


Fig. 21. The schematic symbol for a quartz crystal in its holder is shown at A. The equivalent electrical circuit of a quartz crystal is above at B.

Because of its electrical characteristics, the crystal can be used in an oscillator circuit. The electrical properties of the crystal are somewhat different from those of a typical coil and capacitor. The mechanical properties of the crystal produce a very high apparent inductance, L. Also, since the mechanical losses during vibration are small, the electrical equivalent resistance is also very small. When we have an L-C circuit containing a large inductance and a very low resistance, the Q of the circuit will be very high. This is the case with the crystal used in oscillator circuits.

Practical crystals have effective Q values which are about 100 times as great as that ordinarily obtainable with the usual inductance coil and tuning capacitor. Crystals, therefore, have extremely good frequency selectivity. The higher the Q, the better the frequency stability. Thus, if we substitute a crystal for the ordinary L-C tank circuit, we can make an oscillator that has good frequency stability.

CRYSTAL OVENS

The purpose of a crystal oven is to maintain the crystal at a constant temperature to prevent frequency drift. Some time ago, it was common practice to put the crystal, the entire master oscillator of the transmitter, and often even the buffer stage in a heat-controlled chamber. This prevented temperature variations from affecting the physical dimensions of the coil and capacitor in the plate circuit, and

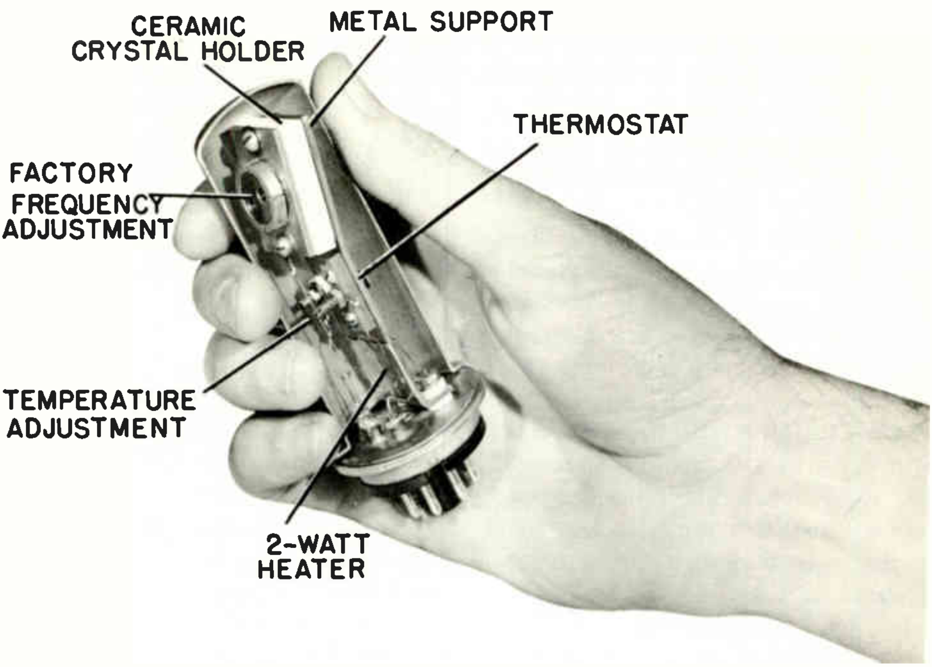


Fig. 22. A cut-away view of a crystal oven.

thus changing the oscillator frequency. This is seldom done in modern transmitters; generally only the crystal is temperature-controlled.

A typical modern crystal oven is shown in Fig. 22. The overall unit is about the size of a metal receiving tube – about 4 inches high and 1-1/4 inches wide. Smaller units are also available. The crystal oven unit fits into an octal tube socket. The unit in Fig. 22 is guaranteed to hold the transmitter frequency within 10 Hertz at any point on the broadcast band.

The crystal is contained in a ceramic holder attached to a copper support. The heater is also wrapped around this support. Thus, the metal support conducts the heat to the crystal and maintains it at a constant temperature.

The heat is controlled by a bimetallic thermostat attached to the metal support. When the temperature of the support drops a very small amount, the thermostat contacts close, and current is applied to the heater. When the support reaches a certain temperature, the thermostat contacts open. The difference between the on and off temperatures of the thermostat is very small; that is, the temperature must drop only a small amount before current is again applied to the heater coil.

The crystal frequency and thermostat temperature adjustments shown in Fig. 22 are set at the factory when the unit is assembled. It is impossible to change the adjustments because the unit is hermetically sealed. Thus, when ordering such a crystal unit, you must specify the exact crystal frequency you want and the circuit in which the crystal will be used. Crystal manufacturers will not guarantee the operating frequency of a crystal unless the crystal is adjusted in the circuit in which it will be used.

CRYSTAL OSCILLATOR CIRCUITS

To help see how the crystal oscillator works, let's look at another L-C oscillator. This oscillator is shown in Fig. 23, and is called a tuned-grid, tuned-plate oscillator. It is easy to see where this oscillator gets its name, since there are resonant circuits in both the grid circuit and the plate circuit of the tube.

The tuned-grid, tuned-plate oscillator works because of the capacity between the plate and grid of the triode tube. When the resonant circuit in the plate circuit of the tube is tuned to a frequency slightly lower than the operating frequency, it will act like an inductance. Under these conditions, the phase of the signal voltage fed from the plate of the tube back to the grid of the tube is correct to aid the ac grid voltage, and oscillation occurs.

The crystal oscillator shown in Fig. 24 is simply a modification of the tuned-grid tuned-plate oscillator shown in Fig. 23. Here a crystal has been substituted for the resonant circuit in the grid circuit of the oscillator and a milliameter is shown in the plate circuit of the stage to measure plate current. Actually almost all oscillators have some provision made for measuring plate current.

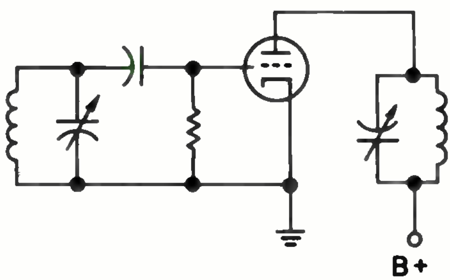


Fig. 23. A tuned-grid, tuned-plate oscillator.

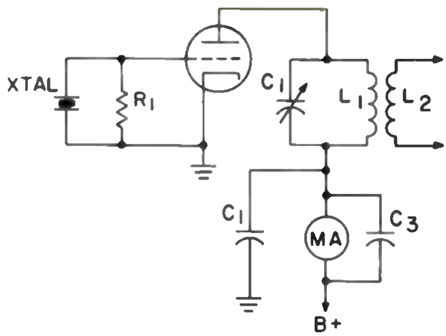


Fig. 24. A simple crystal-oscillator circuit.

Since the crystal is the equivalent of the circuit that has been removed, the crystal oscillator operates in exactly the same way as the resonant circuit in the grid of the tuned-grid, tuned-plate oscillator. The plate circuit must be tuned so that it presents an inductive load, and energy will be fed from the plate of the tube back to the grid in the correct phase to aid the ac grid voltage so oscillation will occur.

In the input circuit, the rf current flowing through the crystal itself is limited only by the resistance of R_1. This resistance is usually low enough to result in a fairly high crystal current. If the resistance is made too high, the crystal may be somewhat erratic as it starts to oscillate when power is applied to the stage. On the other hand, if the current becomes higher than about 100 milliamperes, the crystal vibrations may be so violent that the crystal may shatter, and thus destroy itself. Of course, the thinner the crystal and the higher its resonant frequency, the lower the safe current limit becomes.

Many transmitters have an rf current

meter in series with the crystal to indicate this current. To prevent possible damage to the crystal, the amplitude of oscillation in a crystal oscillator circuit must be kept at a safe level. Usually this is accomplished by keeping the plate voltage on the oscillator tube low. This low plate voltage reduces the maximum output power that can be obtained from such oscillators.

Pierce Oscillator. It is possible to place the crystal between the plate and grid circuits, as shown in Fig. 25, instead of between the grid and cathode. In this case the circuit is similar to the ultra-audion circuit. This circuit is called the *Pierce* oscillator.

Notice that the circuit contains no tank inductance or tuning capacity. The amount of feedback and the grid excitation can be controlled to some extent by adjusting the capacity of C_1. The larger this capacity, the less the feedback. The exact capacity of C_1 in most cases is not critical. Usually, when the best value is determined, crystals of slightly different frequency can be switched into the circuit without further adjustment. Again the plate voltage must be low to prevent damaging the crystal.

Crystal Oscillators Using Multi-Grid Tubes. Tetrode and pentode tubes also may be used in crystal-oscillator circuits. These tubes have less plate-to-grid capacity than do triodes, and therefore, there is less feedback current to the grid when they are used as crystal oscillators. The lower feedback means that these tubes can be operated at a higher output power than a triode without excessive rf crystal current.

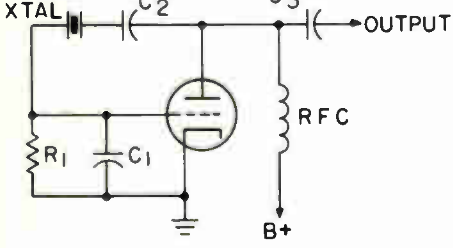


Fig. 25. The basic Pierce oscillator.

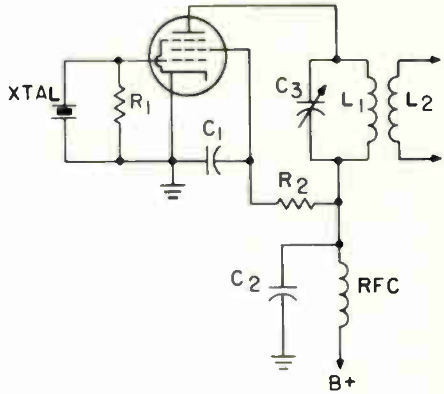


Fig. 26. Crystal oscillator using a pentode tube.

A crystal oscillator using a pentode tube is shown in the diagram of Fig. 26. Because of the limited amount of plate-grid feedback, the plate voltage can be considerably higher than for the triode oscillator. This, of course, increases the output power. The circuit may also be used for high-frequency crystals having fundamental resonant frequencies of about 10 MHz. The output in this case can be high, but the small amount of current through the crystal protects it from excessive vibration. Sometimes extra feedback is needed to produce oscillation in the circuit. This is done by connecting a small capacitance between the plate and the control grid.

Crystal Control Transistor Oscillators. A transistor tickler coil oscillator using a crystal to control feedback is shown in Fig. 27. Positive feedback from collector to base is provided through the mutual

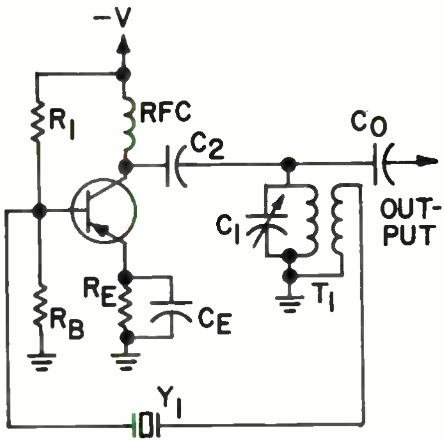


Fig. 27. Crystal tickler coil oscillator.

inductance between the windings of T_1. This provides the 180° phase shift necessary to sustain oscillation. In this circuit, the crystal acts as a series resonant circuit. At frequencies other than the resonant frequency the crystal acts as a high impedance, blocking the feedback path. At the resonant frequency the crystal offers minimum impedance to the feedback signal. The operating frequency may be varied by the use of different crystals and tuning the tank to the frequency of the crystal with C_1.

The crystal oscillator shown in Fig. 28 is a variation of the Colpitts oscillator. In this circuit the crystal acts as a parallel tuned circuit and replaces the L-C tank. Operating frequency is determined by the crystal and the capacitance in series with it. At the resonant frequency, the crystal and capacitors in parallel with it form a high impedance tank circuit. Capacitors C_1 and C_2 form a voltage divider that is center-tapped to ground. The voltage developed across C_1 is applied between the base and ground providing a 180° phase shift in the feedback path.

OVERTONE OPERATION

Most crystals will oscillate not only at their fundamental resonant frequencies, but also at odd overtones of the fundamental. A crystal with a fundamental of 4 MHz, for example, will oscillate also at frequencies near 12 MHz, 20 MHz, 28 MHz, etc. The oscillation frequency will not be an exact multiple of the fundamental oscillation frequency. Exact multiples would mean harmonic operation rather than overtone operation. Thus, if you use a crystal designed for fundamental operation as an overtone crystal, you can use the frequency markings on the crystal holder only to get an approximate idea of the actual frequency at which the oscillator is working. To determine the exact frequency, you will have to measure it with a frequency meter.

Although any crystal can be used in overtone operation, it is best to use one that has been ground specifically for this purpose. The ordinary fundamental type of crystal is somewhat unstable and hard to adjust when operated on an overtone. Most of the overtone crystals are designed

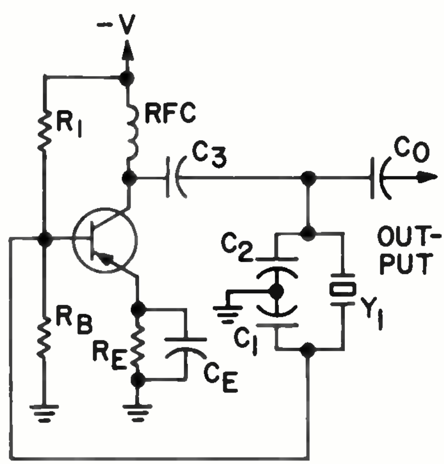


Fig. 28. Colpitts crystal oscillator.

to operate on the third overtone; operation at higher overtones is possible, but the stability and ease of adjustment becomes more critical at the higher overtones.

The chief advantage of overtone operation is that we can generate frequencies in the vhf range without the use of doubler stages. The highest fundamental oscillation frequency of a crystal is about 10 MHz. If one is ground to operate on a higher frequency then it is so thin that it breaks easily. Thus, by operating it on the third or fifth harmonic, we get a much higher frequency than we could on the fundamental. Another advantage of overtone operation is that no frequencies are generated that can cause interference with other channels. Because the number of frequency multiplier stages is reduced or completely eliminated, overtone crystals are used often in mobile, marine, and aircraft transmitters in which compactness is important.

The oscillator circuit using an overtone crystal is similar to an ordinary crystal oscillator circuit of the type you will study later. The overtone at which the crystal will operate is determined by the plate circuit resonant frequency. The circuit of an overtone oscillator is shown in Fig. 29. The oscillator uses a crystal

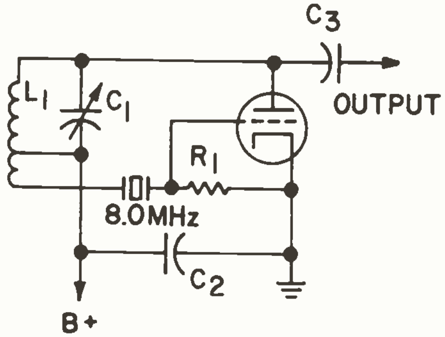


Fig. 29. An overtone oscillator circuit.

having a fundamental frequency of about 8 MHz to give a 24 MHz output when operated on the third overtone.

In Fig. 29, the circuit made up of C_1 and the upper end of L_1 is resonant at 24 MHz. The tap on L_1 is held at rf ground potential by C_2. The lower end of L_1 is inductively coupled to the upper end (usually this is a simple one-tapped coil) so energy is fed from the plate circuit back to the grid circuit to sustain oscillations. The rf signal fed back to the grid circuit will cause the crystal to oscillate on its third overtone frequency. Output from the oscillator is taken from the plate circuit through C_3.

Modern overtone crystals are capable of oscillating as high as 100 MHz on higher order harmonics, so the output of a crystal oscillator could be up in this region. However, when such high frequency signals are needed, you will often find a crystal oscillator operating at 50 MHz followed by a doubler to increase the frequency to 100 MHz, or an oscillator operating at about 33.3 MHz followed by a tripler.

An oscillator circuit that can be used to generate both even and odd harmonic frequencies of the fundamental is the tri-tet circuit shown in Fig. 30. The output signals are harmonics rather than overtones.

The tri-tet circuit is actually the crystal version of the electron-coupled circuit described earlier in this lesson. The crystal and the resonant tank L_1-C_1 are connected to the control grid, the cathode, and through C_4 to the screen grid (which acts as the oscillator plate), to form a modified tuned-plate, tuned-grid oscillator. The cathode tank circuit L_1-C_1, therefore, must be tuned to a frequency slightly lower than that of the crystal in order to be inductive.

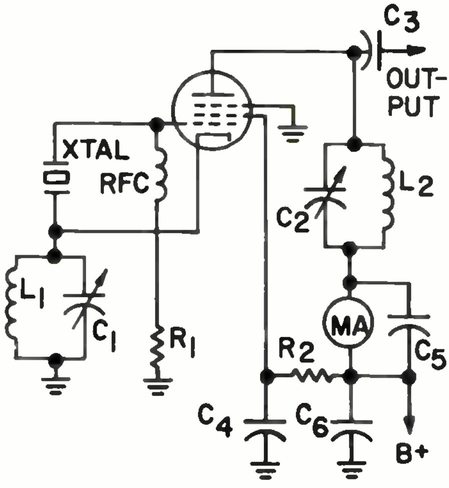


Fig. 30. The tri-tet oscillator circuit.

Since the screen grid is bypassed to ground by capacitor C_4, there is very little direct coupling between the actual tube plate and the oscillator portion of the circuit. The electron stream reaching the plate, however, arrives in the form of pulses that contain relatively large amounts of harmonic energy. If we tune the plate tank circuit L_2-C_2 to a frequency twice that of the crystal, a considerable amount of output power at the second harmonic frequency will be obtained. Even if the plate is tuned to a frequency three times that of the crystal we will get a fair amount of third harmonic output power.

Therefore, the tri-tet circuit not only behaves as a crystal-controlled oscillator, but also performs as a frequency doubler or tripler at the same time. The additional feature of the electron coupling prevents load variations from reaching the crystal and influencing the oscillator frequency. A disadvantage of this circuit is that both sides of the crystal are above rf ground potential. Another is the fact that a cathode coil is needed.

CRYSTAL-OSCILLATOR ADJUSTMENT

The tuning procedures for crystal oscillator circuits of all types are similar. The curves in Fig. 31 show how the plate current of an oscillator tube varies with changes in the tuning capacity. The solid curve is for an unloaded oscillator circuit, and the dashed curve represents the circuit loaded.

When you adjust a triode or Pierce oscillator like those shown in Figs. 24 and 26, begin first with the plate tank capacitor in the minimum capacity position. Then rotate the capacitor toward the maximum capacity position. As soon as oscillation begins, the plate current will begin to decrease. It will decrease more and more, going through points 3 and 2 in Fig. 31 as oscillation becomes stronger. If the plate capacitor is turned too far, however, oscillations will stop abruptly as at point 1 in Fig. 31. In practice, it is best not to approach point 1 too closely because minor voltage or current variations may stop the oscillator. Instead, adjust the plate tuning capacitor so that operation will be somewhere in the stable region between points 2 and 3.

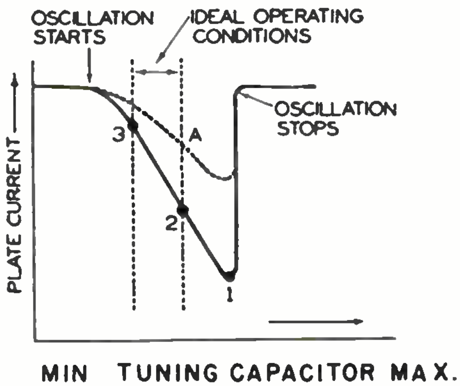


Fig. 31. Curves showing how the plate current in a crystal-oscillator circuit varies with tuning.

When a load is placed on the oscillator circuit, the plate current dip of the oscillator will not be as pronounced. It will follow the dashed curve in the diagram. As before, however, too much capacity will cause the oscillator to stop. The operating point again should be somewhere between points 2 and 3.

The transistor oscillator in Fig. 27 would be adjusted in a similar manner, C_1 being tuned to a point where the current through the tank is between points 2 and 3 in Fig. 31. C_1 is first tuned for minimum capacitance. Then the capacitance is increased until current is between the stable points of operation.

When adjusting the tri-tet circuit in Fig. 30, first set the cathode tank capacitor C_1 to a frequency higher than resonance (minimum capacity) so that oscillations will occur. In this circuit, the usual parts values are such that oscillation will be maintained over a fairly wide range of C_1 adjustment. However, the crystal current increases very rapidly as the capacity is increased. Start tuning C_1 for the minimum capacity position and progress only to the point of normal crystal current. Usually an rf milliammeter is placed in series with the crystal to measure this current. The current should be kept below 100 milliamperes.

With no load connected, the plate tank capacitor C_2 should be adjusted for a minimum value of plate current as indicated by a sharp dip in the meter reading. Now, connect the load to the circuit. This is usually the grid of the following buffer amplifier stage. With the load attached, capacitor C_2 may need readjustment to bring the plate current back to minimum. This time, however, the current value will be somewhat higher because of the loading.

Finally adjust the cathode tank capacitor C_1 for maximum harmonic power output which will be indicated by a maximum current flow in the following amplifier grid current. Also watch the crystal current to be sure that it does not rise above the safe limit.

FREQUENCY SYNTHESIZERS

In order for a transmitter to be versatile, it must be capable of operating on more than one frequency. In the crystal oscillators discussed previously we have to change the crystal each time the frequency is changed. Thus for each channel the equipment is operated on, we need a separate crystal. Transmitters that are operated on only a few channels use separate crystals for each channel. However, this arrangement is not completely satisfactory when the transmitter must be capable of operating on a large number of channels. The frequency synthesizer is one solution to this problem.

A frequency synthesizer is basically a circuit in which harmonics and subharmonics of one or more crystals are combined to provide a variety of output signals. The same principle is used that you studied earlier in the basic superheterodyne receiver. You will recall that in the mixer stage of that receiver the incoming signal was beat with the local oscillator to form two new frequencies. These new frequencies were equivalent to the sum and the difference of the incoming signal frequency and the local oscillator frequency.

Fig. 32 shows one type of frequency synthesizer. Circuit details have been omitted for simplicity. Two oscillator circuits are used with a mixer which will beat the outputs of the oscillators together and produce an output equal to

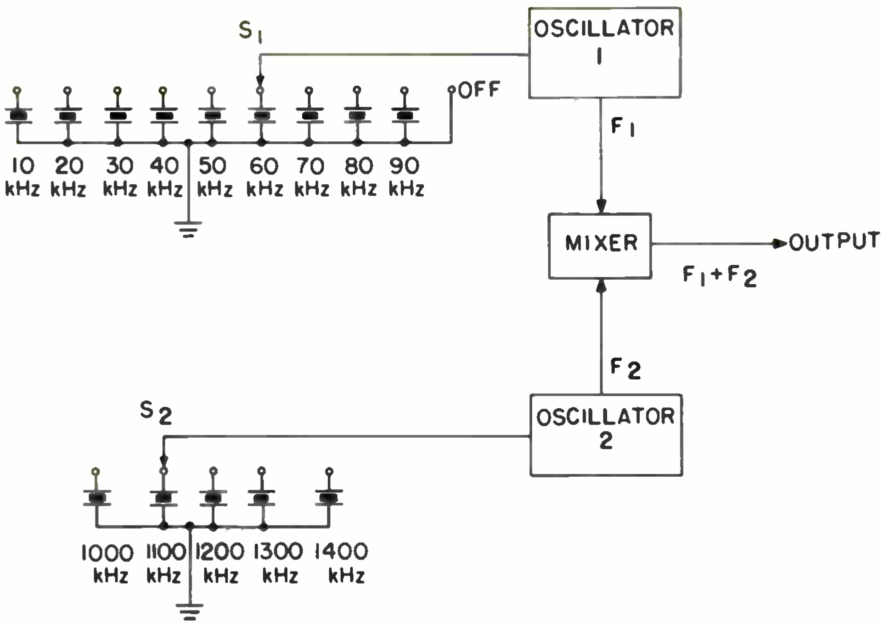


Fig. 32. Multiple crystal frequency synthesizer.

the sum of the two frequencies. Each of the oscillators is tunable to different frequencies of a selected band by switching the crystals in the oscillator circuit.

Oscillator number 1 has 9 crystals in the input. Each crystal is resonant at a different frequency, enabling the output of oscillator number 1 to be varied in 10 kHz steps from zero to 90 kHz. The tenth position of switch S_1 disables the oscillator. Oscillator number 2 has 5 crystals which vary its frequency over a 500 kHz band in 100 kHz steps.

With S_1 and S_2 in the positions shown, oscillator 1 is operating at 60 kHz and oscillator 2 at 1100 kHz. The mixer receives both outputs and beats them to form a 1160 kHz output. If S_2 was left in this position and S_1 moved, the output of the mixer could be varied from 1100 kHz (when S_1 is in the off position) to 1190 kHz in 10 kHz steps. By moving both switches, the output can be varied from 1000 kHz to 1490 kHz in 10 kHz steps. In this manner, a total of 50 different frequencies can be generated from only 14 different crystals.

This type of synthesizer has the advantage of less crystals than a conventional oscillator with similar frequency coverage.

Let's look at a block diagram of a transceiver and see another use for frequency synthesizers.

The transceiver is a compact radio station that uses some of the components for both transmitting and receiving. These units usually transmit and receive on the same frequency.

Fig. 33 is a block diagram of a transceiver. The upper channel is the receiver section while the lower channel is the transmit section.

Incoming rf is coupled from the an-

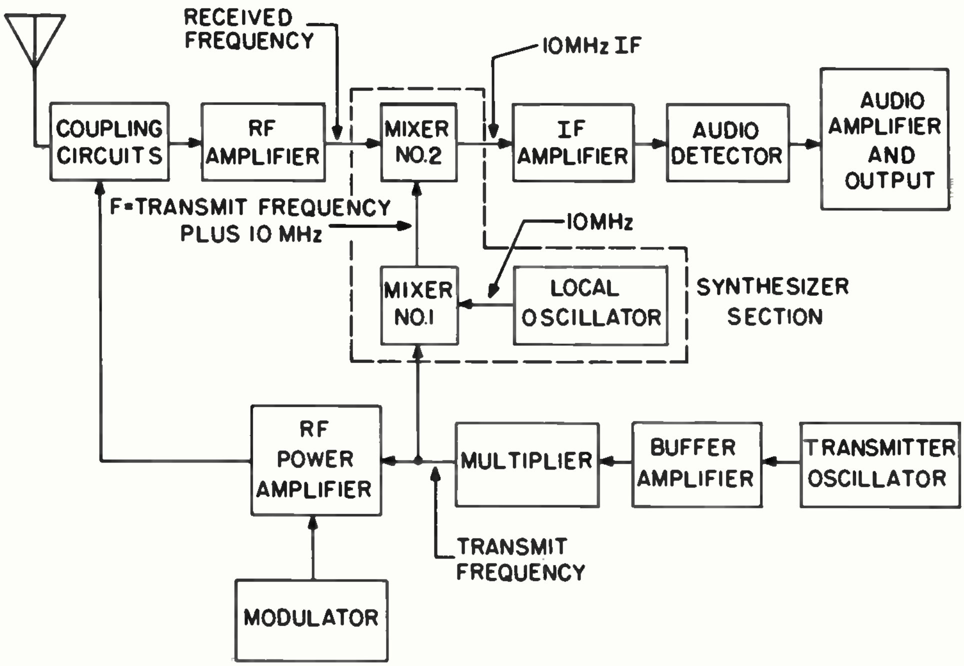


Fig. 33. Block diagram of a basic transceiver.

tenna to the rf amplifier. This amplified rf is then heterodyned in mixer number 2 with a signal from mixer number 1. We will discuss the signal from mixer number 1 shortly. The difference frequency is detected and sent through the i-f amplifiers to the audio detectors and output section.

The transmit channel gets its frequency from the transmitter oscillator. After being multiplied to the proper frequency, the rf is fed to the power amplifier. In the power amplifier the modulating signal is applied to the rf. The rf, now containing the intelligence, is coupled from the power amplifier to the antenna.

Notice that between the multiplier and the power amplifier, the transmitter frequency is fed to mixer number 1. This is part of the frequency synthesizer.

You recall that the i-f is the difference between the local oscillator and the incoming rf. Let's assume in this example that we desire a 10 MHz i-f. In that case our local oscillator frequency to mixer number 2 must be 10 MHz above the incoming rf. Since the transmitted rf and received rf must be the same frequency for stations to communicate with each other, a separate oscillator is required to generate the local oscillator frequency.

If a synthesizer were not used, the local oscillator crystal would have to be changed each time the transmitter oscillator crystal was changed in order to keep the local oscillator 10 MHz above the transmitter frequency. The synthesizer however, uses a very stable constant frequency oscillator operating at the 10 MHz i-f. The transmitted rf is fed to mixer number 1 of the synthesizer where it is mixed with the 10 MHz signal. The sum of these two frequencies is then used as the local oscillator frequency. In mixer number 2 the local oscillator frequency is mixed with the incoming rf and the difference (10 MHz) is used as the i-f.

In this manner the local oscillator frequency to the receiver is changed each time the transmitter frequency is changed. Therefore the local oscillator frequency is always 10 MHz above the received frequency. This eliminates the need to use a different crystal in the receiver local oscillator each time the transmitter frequency is changed.

There are many types of frequency synthesizer circuits in modern communication equipment. They all operate on the same principle of combining two or more frequencies to generate an output frequency. From this discussion you should be able to understand the basic fundamentals of any frequency synthesizer you encounter.

SELF–TEST QUESTIONS

(o) Why is plate voltage kept low in the Pierce oscillator?

(p) Describe how frequency is controlled in the oscillator in Fig. 27.

(q) What is the advantage of overtone operation?

(r) In the multiple crystal frequency synthesizer in Fig. 32, what is the output of the mixer if S_1 is in the 80 kHz position and S_2 is in the 1300 kHz position?

(s) In addition to performing as an oscillator, what other function is accomplished by the tri-tet oscillator?

Nonsinusoidal Oscillators

The oscillators you have studied up to this point have all been sine wave oscillators; that is, their output is a sine wave. In this section we will look at nonsinusoidal oscillators. There are many circuits that can be classed with this group of oscillators but we are primarily interested in only two; the multivibrator and the blocking oscillator. After we complete our study of the basic circuits, we will see how these circuits are used in communication equipment.

An oscillator circuit in which the output is a nonsinusoidal waveform is generally classified as a relaxation oscillator. The relaxation oscillator uses a regenerative circuit in conjunction with an R-C circuit to provide a switching action. The charge and discharge time constants of the R-C components are used to control the shape and frequency of the output waveforms.

THE MULTIVIBRATOR

The multivibrator is essentially a nonsinusoidal two-stage oscillator in which one stage conducts while the other is cut off until a point is reached at which the stages reverse their conditions. This oscillating process is normally used to produce a square wave output. A multivibrator that operates continuously with first one stage conducting and then the other is called a free running or astable multivibrator.

The multivibrator in Fig. 34 is a two stage R-C coupled, common-emitter amplifier with the output of the first stage coupled to the input of the second stage and the output of the second stage coupled to the input of the first stage. Since the signal in the collector circuit of a common-emitter amplifier is reversed in phase with respect to the input signal, a portion of the output of each stage is fed to the other stage in-phase with the signal at the base. This regenerative feedback with amplification is required for oscillation. The output of this multivibrator is a square wave whose frequency is determined by the R-C time constant in the feedback loops.

Forward bias for the base of Q_1 is obtained through the low resistance emitter-to-base junction in series with R_2 across the power supply. In a similar manner, bias for the base of Q_2 is obtained through the emitter-to-base junction and R_3. When the power supply is first energized the current that flows through each collector load resistor, R_1 and R_4, is determined by the effective resistance of Q_1 and Q_2 for a given value of base bias voltage. Due to slight differences in the transistors, more current will flow in one transistor than in the other.

For the purpose of this explanation,

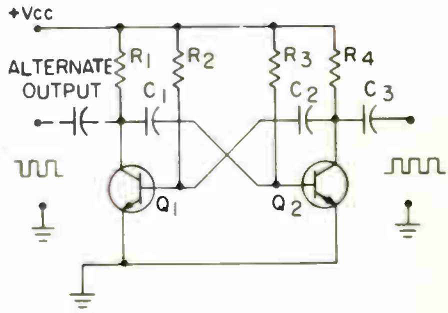


Fig. 34. Astable multivibrator.

assume that initially more collector current flows through Q_1 than Q_2. Thus as collector current in Q_1 increases, the voltage at the junction of R_1 and C_1 decreases. In other words, the collector of Q_1 becomes less positive. This negative going pulse is coupled through capacitor C_1 to the base of Q_2. As the collector current in Q_1 continues to increase, the signal coupled to the base of Q_2 continues to become more negative. As this negative signal overcomes the positive bias established by the emitter-to-base junction and R_3, collector current in Q_2 starts to decrease.

With collector current in Q_2 decreasing, the voltage at the junction of C_2 and R_4 starts to increase. This positive rise in voltage is coupled to the base of Q_1 through capacitor C_2. The positive going pulse at the base of Q_1 aids the foward bias already established by R_2 and the emitter-to-base junction of Q_1. This causes collector current in Q_1 to continue to increase. This regenerative process continues until Q_1 is driven into saturation and Q_2 is cut off.

When Q_1 is saturated, its collector current no longer increases but becomes a constant value, therefore, there is no further change in collector voltage to be coupled through C_1 to the base of Q_2. In a similar manner, Q_2 is cut off so there is no further change in its collector voltage to couple through C_2 to the base of Q_1.

With no positive pulse coupled from Q_2 to Q_1, the base of Q_1 is only a few tenths of a volt positive. C_2 quickly charges through the low resistance of the base to the emitter junction of Q_1 and R_4 to a potential approximately equal to the power supply voltage. The heavy conduction of Q_1 places its collector voltage at nearly ground potential. C_1, which was previously charged with a negative potential at the junction of C_1 and R_3, starts to discharge through R_3 at a time constant of R_3-C_1.

As C_1 discharges, the voltage at the base of Q_2 becomes less and less negative until a point is reached where reverse bias is no longer applied and Q_2 goes into conduction.

When Q_2 starts to conduct, collector current begins to flow through R_4, collector voltage at Q_2 decreases, and a negative going pulse is coupled through C_2 to the base of Q_1. As C_1 continues to discharge, collector current in Q_2 continues to increase and the pulse coupled to the base of Q_1 goes more negative. As this negative signal increases, Q_1 decreases its conduction. This results in a rise in collector voltage which is coupled to the base of Q_2, aiding the forward bias on Q_2.

This regeneration continues until Q_1 cuts off and Q_2 saturates. When Q_1 cuts off, C_1 no longer couples a positive signal to the base of Q_2. C_2 starts to discharge through R_2 at a rate equal to R_2-C_2. When C_2 is sufficiently discharged to remove the reverse bias on the base of Q_1, the transistor again starts to conduct.

This action will continue as long as the power supply voltage is present; the discharge of C_1 controlling the time that Q_2 remains cut off and C_2 controlling the time that Q_1 is cut off. In this manner, C_1 and C_2 control the width of the output pulses.

While we explained the various things happening that caused one transistor to switch from saturation to cutoff, you might have thought this action takes quite some time. Actually, the switching action is very fast. For example, with Q_1 conducting and Q_2 cut off, once the reverse bias on Q_2 disappears and Q_2 begins conduction, current rises to saturation in Q_2, and Q_1 is cut off almost instantly. The result is the output from the multivibrator is essentially a square wave. The square wave output is taken from the collector of Q_2 through C_3. A second output, reversed in phase, is available at the collector of Q_1.

The output of the multivibrator will be symmetrical, that is the two half cycles will be the same if the time constant of C_1 and R_3 is equal to the time constant of C_2 and R_2. However, if we change the time constant of either R-C circuit, the output will no longer be symmetrical. In other words, if we shorten the time constant of the C_1-R_3 network by reducing the value of C_1, it will take C_1 less time to discharge through R_3. As a result, Q_2 will be cut off for a shorter period than Q_1 and the two halves of the square wave will no longer be equal.

There are two major variations of the astable multivibrator, the monostable and the bistable. The bistable is a modification of the astable and can be used as a switching circuit. Basically it is the same circuit as the astable but provision has been made to control the change of the condition of the transistors with an input signal. Upon receipt of an input the transistors in the bistable multivibrator will change state and remain in their new condition until another input pulse is received.

The monostable multivibrator, like the bistable, is a modified astable circuit. Monostable multivibrators have only one stable state and the transistors remain in their respective states (saturated and cut off) until an input pulse is received. At that time the transistors change states and remain in their new state a length of time determined by the R-C components in the circuit. At the end of this R-C time they return to their stable condition.

THE BLOCKING OSCILLATOR

Blocking oscillators are a type of oscillator used for generating pulse waveforms of short time duration followed by a period of no output. Similar to multivibrators, blocking oscillators may be either free running or driven. Let's first look at a free running blocking oscillator, then we will examine a practical application of the circuit.

Fig. 35 shows the basic free-running blocking oscillator. Transformer T_1 provides the necessary regenerative feedback from the collector to the base of Q_1. Terminals 1 and 2 are the primary winding and connect the collector to the power supply. Terminals 3 and 4 are the secondary and furnish feedback to the base. The output is taken across the third winding at terminals 5 and 6. Notice the phase inversion between the different windings.

When the power supply is first energized, a small amount of collector current will flow through the primary of T_1. This current flow induces a voltage in the base winding (terminals 3 and 4). The induced voltage causes C_1 to charge through the low forward resistance of the base-to-

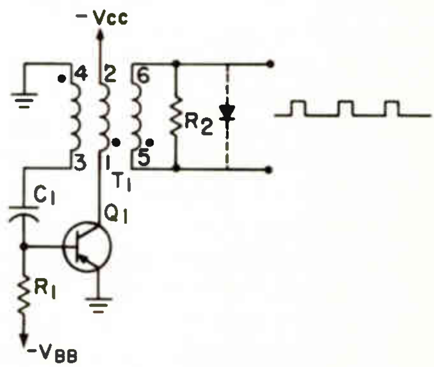


Fig. 35. Free running blocking oscillator.

emitter junction. This couples the induced voltage to the base of Q_1.

Since collector current flow is increasing, the voltage at terminal 1 will be less negative or going in a positive direction. The induced voltage in the base winding is of opposite polarity; therefore, a negative going signal is coupled to the base of Q_1. This increase in forward bias aids collector current and regeneration continues rapidly until the transistor becomes saturated.

When Q_1 reaches saturation, the current is no longer changing; therefore, there is no voltage induced in the base winding. C_1 begins to discharge through the resistor R_1, holding the base-to-emitter junction forward biased. As the charge on C_1 bleeds off, the forward bias on Q_1 decreases.

As the forward bias of Q_1 decreases and collector current is reduced, the magnetic field about the primary winding (terminals 1 and 2) collapses. The collapsing magnetic field induces a voltage in the secondary (terminals 3 and 4) which is positive at terminal 3. This induced voltage, coupled through C_1 drives Q_1 to cutoff.

Due to the reverse bias, Q_1 remains cut off until C_1 discharges through R_1 and T_1 to a point where the base of Q_1 is returned to a forward bias condition. When the forward biased condition is reached, conduction begins and the cycle repeats.

The output pulse width depends mostly on the inductance of T_1. The smaller the inductance, the more rapidly the collector current must increase to maintain a magnetizing current in T_1, the faster the collector current will reach saturation, and the shorter the pulse width. Usually, C_1 has little effect on the pulse width. However, if C_1 is small enough so that the capacitor can charge rapidly during pulse time, there will be a decrease in pulse width.

The frequency of the blocking oscillator in Fig. 35 is determined by the value of R_1-C_1. Compared to R_1, the resistance of the base winding has little effect on the discharge time of C_1.

Resistor R_2 is a damping resistor connected across the output winding of T_1 to reduce the amplitude of the reverse voltage, sometimes called the overswing, caused by the collapsing magnetic field about T_1 at the end of the output pulse. If it were not for R_2, the amplitude of the reverse voltage pulse could exceed the breakdown voltage of Q_1 and damage the transistor. In another type of damping circuit a clamping diode is placed across the collector winding or across the output winding. The diode would then shunt any reverse voltage present in these windings.

Tube type multivibrators and blocking oscillators were once widely used, but in modern equipment they are generally being replaced by transistors. The operation of the older vacuum tube units is essentially the same as the modern transistor units.

Both oscillators with a sine wave output and those with a pulse type output are widely used in communications equipment. Be sure you understand how both types work before leaving this lesson.

SELF-TEST QUESTIONS

(t) In the astable multivibrator shown in Fig. 34, what components control the length of time Q_2 is cut off?

(u) What type output is obtained from the blocking oscillator?

Answers to Self-Test Questions

(a) Regenerative or positive feedback. If positive feedback is not present, oscillations will be damped.

(b) Decrease. Frequency is equal to:

$$\frac{1}{2\pi \sqrt{LC}}$$

Therefore: if inductance increases, the frequency will decrease.

(c) There will be more cycles in the wave train. Losses in the coil of a high Q circuit are low so more cycles occur before the wave train is damped.

(d) .0000005 seconds or .5 microseconds. The formula is:

$$P = 1/F$$

$$P = \frac{1}{2{,}000{,}000}$$

$$P = .0000005 \text{ seconds.}$$

(e) They are in-phase. The voltage developed across L_2 is the feedback voltage and must aid plate current.

(f) The bias voltage developed across R_1. If the oscillator output increased, C_2 would be charged to a higher potential on positive peaks and would produce a larger bias voltage across R_1. The larger bias would reduce the amplitude of the oscillator output.

(g) 1. Oscillator remains energized at all times.
2. Oscillator coil and capacitor are placed in an oven.
3. Temperature compensating capacitors are placed across the tank.

(h) A swamping capacitor is placed across collector-to-emitter junction.

(i) R_1 and C_2 develop the negative bias. L_2 develops the signal.

(j) It goes positive. The reactance of C_2 is small at the resonant frequency of the tank; therefore, the positive voltage induced in the tank by L_2 (voltage across C_1) is coupled to the grid.

(k) C_1 develops the feedback voltage. It is applied to the base through C_3 and across R_1.

(l) The ultra-audion oscillator is a modification of the Colpitts circuit.

(m) Through the R-C phase shift network in the grid. Each pair of R-C components shift the phase a definite number of degrees. Total shift through all stages must equal 180°.

(n) R_2 and C_2 develop the feedback voltage.

(o) To prevent damage to the crystal.

(p) Frequency is controlled by the crystal, Y_1. At the resonant frequency, Y_1 offers minimum impedance to the feedback voltage. At other frequencies, impedance increases.

(q) A higher frequency can be obtained without the use of frequency multipliers.

(r) 1380 kHz.

(s) The tri-tet oscillator functions as an oscillator and a frequency multiplier.

(t) C_1 and R_3. Q_2 remains cut off until the charge on C_1 leaks off enough for Q_2 to become forward biased.

(u) A pulse output followed by a period of no output.

Lesson Questions

Be sure to number your Answer Sheet C206.

Place your Student Number on every Answer Sheet.

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

1. If the operating voltage on an oscillator tube is changed so that the plate resistance of the tube increases, what will happen to the oscillator frequency?

2. When adjusting the capacitance of the tank circuit in a crystal oscillator, what happens to the tank current as oscillations start?

3. What component of the blocking oscillator has the most effect on the output pulse width?

4. For what part of the cycle does current flow through the oscillator tube in an L-C oscillator circuit?

5. What is the advantage of a frequency synthesizer over the normal local oscillator circuit in a transceiver?

6. What will the frequency be at 65°C of a crystal oscillator using an X-cut crystal that operates at 5250 kHz at 50°C, if the temperature coefficient is -20 Hz per degree centigrade per MHz?

7. What controls the amount of feedback applied to the base of the transistor in a transistor Colpitts oscillator?

8. In a tuned L-C circuit, if capacitance is increased while inductance is held constant, what will happen to the output frequency?

9. How is positive feedback attained in the Wien-bridge oscillator?

10. How is frequency multiplication obtained in the tri-tet oscillator?

Take The Middle Course

Most of us realize the necessity for moderation in eating and drinking, but we often overlook the fact that moderation in all things is essential to happiness.

Consider, for example, the simple matter of opinions. If a man can see only his own opinions, and is unwilling to recognize that other people may also have good ideas, he is opinionated. A man with this fault is often unhappy, because he doesn't get along very well with other people. On the other hand, if a man yields his ideas to another's too readily, he is weak-kneed - and also unhappy.

If you can give and take - if you are open to reason - if you steer a middle course, you will be liked, people will be comfortable in your company, and you will be following one rule of happiness.

Let "moderation in all things" be one of the guiding principles of your life.

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-306

RADIO • TELEVISION COMMUNICATIONS

nri

LINEAR RF POWER AMPLIFIERS

21CC

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

LINEAR RF POWER AMPLIFIERS

21CC

STUDY SCHEDULE NO. 21

☐ 1. **Introduction** . . . . **Pages 1-2**

A brief discussion of why linear amplifiers are used to amplify the modulated signal at the transmitter.

☐ 2. **The Class B Linear Amplifier** . . . . **Pages 2-9**

You learn how linearity is achieved and study the requirements for the driver stage and the plate circuit, and the power considerations for a linear class B amplifier.

☐ 3. **Adjusting a Linear Amplifier** . . . . **Pages 9-12**

You learn how to interpret meter readings and study step-by-step adjustment procedures.

☐ 4. **Variations in Linear Amplifiers** . . . . **Pages 13-20**

You study push-pull, multi-grid and grounded-grid amplifiers, and the high-efficiency Doherty amplifier.

☐ 5. **Outphasing Modulation System** . . . . **Pages 21-28**

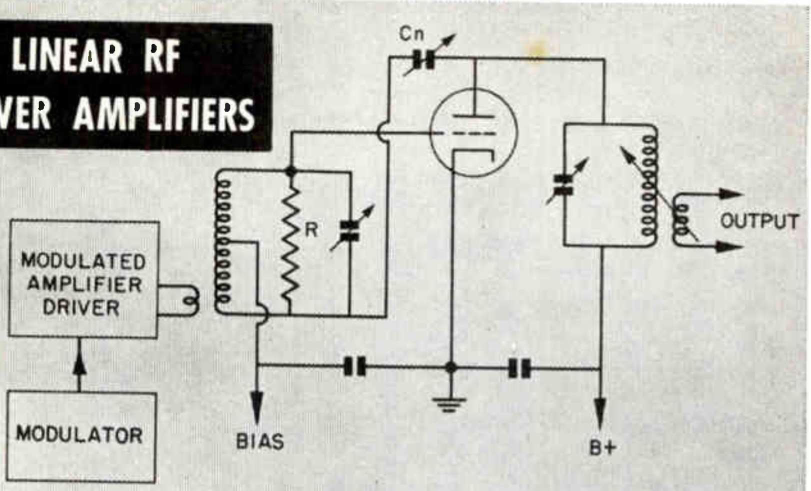
Here we take up a system of modulation, which by using phase differences, eliminates the necessity of operating the amplifiers following the modulator in a linear manner.

☐ 6. **Answer Lesson Questions.**

☐ 7. **Start Studying the Next Lesson.**

MP3M269 **1969 EDITION** Printed in U.S.A.

LINEAR RF POWER AMPLIFIERS



As you have learned, in amplitude modulation, the amplitude of the carrier signal is made to vary according to the amplitude variations of the intelligence signal. In some AM transmitters, the modulation takes place in the plate circuit of the final power amplifier stage where the power level is high, so the modulated signal is fed directly to the transmitting antenna. This is called "high-level" modulation. The modulation can also take place in an intermediate class C power amplifier stage. Then, the power of the modulated signal is increased by additional power amplifiers. This is called "low-level" modulation, because it takes place at a lower power level. The modulated signal must not be distorted by stages following the modulated stage. The signal that is fed to the transmitting antenna must be a faithful copy of the output signal from the modulator, or the demodulated signal from the receiver will also be distorted.

The stages that amplify the modulated signal cannot be operated in class C, because a class C stage is driven to saturation at the peak of each positive alternation of the input signal, which would distort the modulated signal waveform. They could be operated in class A, in which case, the output signal would be an exact duplicate of the input signal. However, we want high power output from a transmitter as well as an undistorted signal. A stage operated in class A has very low efficiency and very low power output.

It is most satisfactory to operate the power-amplifier stages following the modulated stage in class B. This gives an undistorted output signal, and more power output than for class A. The class B amplifier stage is operated on the straightest part of its transfer characteristic curve, so that there is a linear relationship between the amplified output voltage and the exciting voltage applied to the input. That is, the output voltage developed across the load is proportional to the grid voltage. An increase or decrease in the excitation voltage will produce a corresponding increase or decrease in the output voltage.

As you will learn in the following section, linearity is obtained by proper adjustment of the grid bias and the load impedance. All the stages following the modulated stage must be operated as class B linear amplifiers.

There is an exception to this, and that is in a system of modulation called "outphasing modulation." With this system of modulation, the rf power amplifiers following the modulator do not need to be linear. You will see why in the last section of the lesson.

First, you will find out how high-power linear amplifiers work, how they are adjusted, and what their advantages and disadvantages are. One of the chief disadvantages of the class B linear amplifier is that its efficiency is lower than that of a class C amplifier. However, it is possible to get higher efficiency by redesigning the stage. An example of a high-efficiency amplifier is the Doherty linear amplifier, which we will discuss.

Linearity in a transmitter can be improved by using feedback between the final linear amplifier output and the speech amplifier section. Feedback will also reduce distortion and improve frequency response and stability in the transmitter stages. We will also take up feedback circuits. First let us find out how the class B linear amplifier works.

The Class B Linear Amplifier

A schematic diagram of a class B linear amplifier stage is shown in Fig. 1. This stage has tuned resonant circuits in the input and output, and is neutralized in the same way as a class C power amplifier.

The basic difference is that the linear amplifier is biased as a class B stage. The tube, however, is not biased exactly to plate-current cut-off. The transfer characteristic curve of a vacuum tube is shown in Fig. 2. Notice that the lower end of the curve bends and is very non-linear before the plate-current cut-off point is reached. If the stage were biased exactly at cut-off, the non-linearity in this part of the curve would cause the signal to be distorted. Therefore, the bias is set a little above the actual cut-off point on the curve, so that the tube will operate only on the straight or linear part. This point is called "extended cut-off," and is shown in Fig. 2. The extended "cut-off" point is found by extending the straight part of the transfer curve until it crosses the current axis. This is the point at which plate-current cut-off would occur if the lower end of the curve were linear.

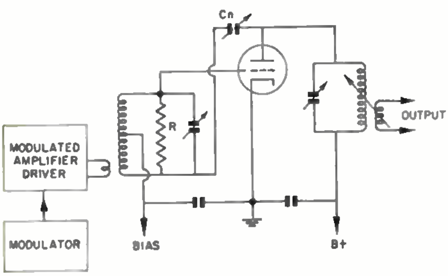


FIG. 1. Basic class B linear amplifier system.

With the grid bias set to the extended cut-off point, the plate current flows for slightly more than 180° of the input signal cycle. In class C, as you learned, the plate current flows for 120° to 150° of the cycle.

The linear power amplifier can be operated in class AB_1 or class AB_2 instead of in class B. You will remember that a class AB stage is a stage that is biased between class A and class B conditions. In class AB_1 operation, the input signal never

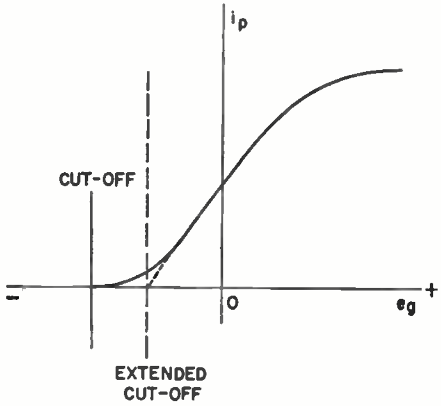


FIG. 2. A tube's e_g-i_p curve, showing the extended cut-off point.

drives the grid positive. In class AB_2 operation, the signal voltage on the grid drives the grid slightly positive on a part of the positive cycle. Class AB_2 operation delivers more power than class AB_1 but less power than class B.

The operation of the linear amplifier can be demonstrated by using the waveform illustrations in Fig. 3. With the grid bias set at the extended cut-off point, a small plate current will flow when no excitation is applied to the input. When a signal is applied to the grid, on the negative alternations the grid is driven below plate current cut-off, so no plate current flows. On the positive alternations, the input signal voltage subtracts from the grid bias, reducing the negative bias on the tube, and the plate current flows for slightly more than the entire positive half cycle. The plate current is in the form of pulses as shown in the diagram.

For distortionless output, the amplifier must operate over the straight portion of the characteristic curve. In other words, the highest peak grid voltage must not swing the plate current beyond point A. As you have learned, on 100% modulation, the

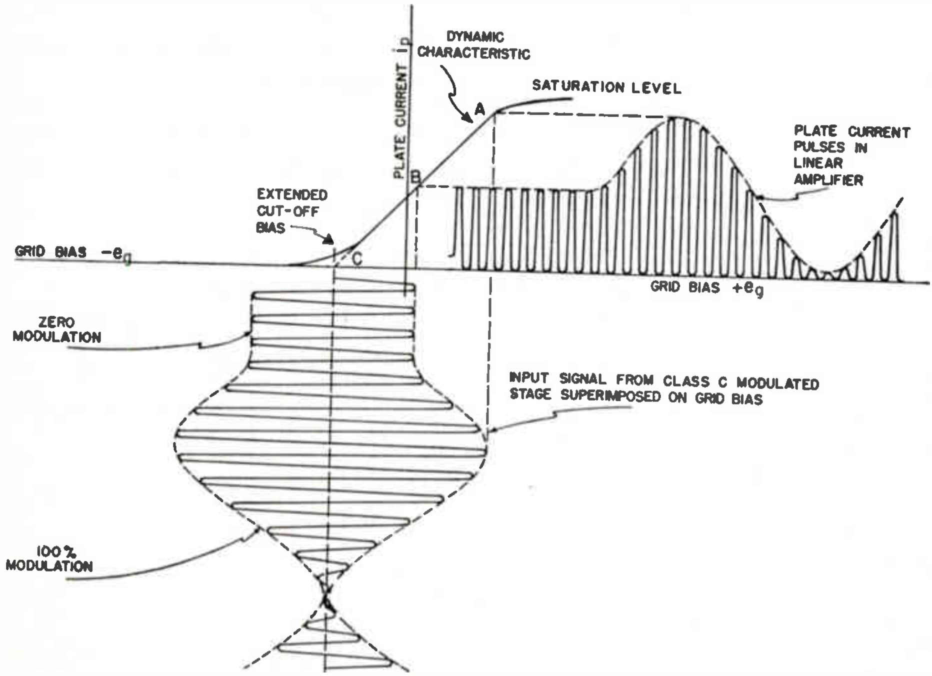


FIG. 3. Class B operation of an amplifier.

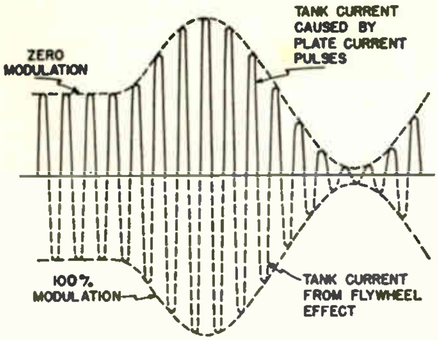


FIG. 4. The flywheel effect of the plate tank circuit in a class B amplifier supplies the missing half-cycles of plate current.

peaks are twice the unmodulated carrier value. In the diagram, the current value at point B is slightly more than half the current value at point A. Thus, if the unmodulated carrier excitation is set at point B, the positive peak at 100% modulation will swing the rf excitation to point A. During the modulation troughs when the modulation amplitude is least, the plate current will be driven near point C. For most efficient linear operation of the amplifier, the grid bias must be set properly so that point B is midway between points A and C, and the input signal must have a high enough amplitude to swing the plate current over the entire linear portion of the characteristic curve between points A and C.

The plate current pulses supply energy to the plate tank on the positive half cycles, just as in a class C amplifier. Then, on the negative half cycles when the plate current is cut off, the energy stored in the tank is fed back into the circuit to form the negative alternations of the output signal. This is shown in Fig. 4. The upper half of the output signal waveform, shown in solid lines, is the part contributed by the plate current pulses. The lower half, produced by the energy-storing (flywheel) action of the tank circuit is shown in dashed lines.

Thus, the output signal fed to the antenna or load circuit is a completely modulated signal, even though the class B rf amplifier tube feeds power to the load for only half of the input signal cycle.

The circuit shown in Fig. 1 is a single-ended stage. In your study of audio stages, you learned that for audio frequencies a class B stage must be operated in push-pull, or the sound will be highly distorted and contain many harmonics. However, this is not true at rf frequencies; a class B stage using a resonant circuit in the plate circuit can be single-ended because the resonant circuit restores the missing parts of the modulated signal wave-form. The resonant circuit also eliminates many of the undesired harmonics.

HOW LINEARITY IS ACHIEVED

The proper operation of a class B linear stage is determined by its grid bias, grid drive, plate voltage, load impedance, plate and grid currents, and the power relationships in the

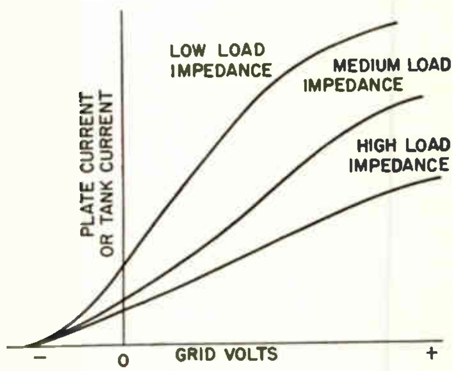


FIG. 5. Linearity curves for a class B amplifier using various values of load impedance.

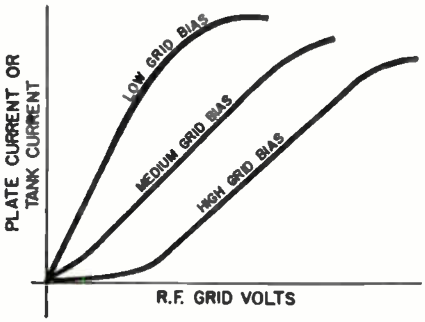


FIG. 6. The effect of various grid bias values upon the linearity of a class B amplifier.

input and output circuits. The linearity in the stage depends primarily on the grid bias and the load impedance.

The plate load impedance must be chosen to get high power output and efficiency as well as good linearity.

Load Impedance. Fig. 5 shows the effect of various load impedances on a tube operating in class B. The instantaneous grid voltage, made up of the bias voltage and the rf excitation signal, is plotted along the horizontal axis. The plate current, or proportional tank current, is plotted along the vertical axis. As you can see, the curve is not linear for low impedance but is quite linear for high impedance. The low impedance is not satisfactory, because the output would be distorted, although it is ideal for class C operation because it has high power output. The high impedance is not satisfactory either, because the plate current swing would be small and the power output very low.

Proper loading of a linear amplifier must be a compromise between linearity and power output. The medium load impedance curve in Fig. 5 is the one that is usually chosen. For a triode tube, the best value of the load impedance is equal to about twice the plate resistance of the tube.

Grid Bias. Now let us see how the grid bias affects the linearity of the stage. Fig. 6 shows the dynamic characteristic curves for various values of grid bias. With a low grid bias, the lower part of the curve is linear, but it soon folds over. If a modulated signal were applied to the input, the peak of the modulation envelope would be compressed, and the signal would be distorted.

With high grid bias, the curve is flat or compressed at the bottom, and a modulated signal would again be distorted. A medium value of bias is best for linear operation. There is a slight bend near the bottom of the medium bias curve which is kept at a minimum by operating the tube at the extended cut-off bias as discussed previously.

The grid bias for the linear amplifier must be supplied from a separate low-impedance power supply. Grid-leak bias, which is used in class C stages, cannot be used in a class B linear stage, because the excitation to the linear amplifier is not constant; the grid current flow is small for low modulation and very high for 100% modulation. Thus, the grid-leak bias itself would vary over the modulation cycle and cause distortion.

Sometimes a cathode bias resistor is used, because the average cathode current in a linear amplifier is constant and does not vary during the modulation cycle. However, a cathode bias resistor wastes too much power for use in a stage other than a low power stage, so fixed bias is more often used.

The power supply that provides the bias voltage must be a low-impedance supply, because if the impedance is low, the variation in the grid current due to the input signal will have little effect on the dc bias voltage. Also, by using a separate bias source, the bias can be easily adjusted to get the exact value required for the most linear operation of the stage.

Excitation. The excitation must be carefully set to prevent distortion during modulation peaks, because regardless of what load impedance or grid bias is used, an excessive grid excitation will cause the plate current to swing into the curved portion of the characteristic curve. At the same time, the excitation must be high enough to operate the class B stage with reasonable efficiency. Thus, the driver stage must be operated properly to get undistorted linear output.

REQUIREMENTS OF THE DRIVER STAGE

Now let's consider the requirements of the stage that drives the class B amplifier. Because grid current flows in the class B amplifier during the part of the cycle that the grid is driven positive, power is being dissipated in the grid circuit. This power must be supplied by the driver stage. The higher the grid current, the lower the impedance presented by the amplifier, and the more power the driver stage must supply.

During an rf cycle, the grid impedance of the amplifier may change from an infinite value for negative grid potential to only a few hundred ohms for a high positive potential. This varying impedance places a varying load on the driver with the result that the driver output voltage, which is fed to the amplifier to drive it, will vary. For no grid current and a light load, the driver voltage will be high; for high grid current and a heavy load, the driver voltage will be reduced. If the output impedance of the driver is too high, this effect becomes worse and the driver output voltage may become so low when the amplifier grid is drawing current, that the grid is not driven far enough positive to drive the stage.

The positive peaks of excitation will be flattened out as shown in Fig. 7. When this happens, we say the driver has poor regulation. In Fig. 7, the excitation of the grid of the class B amplifier from a poorly regulated driver is shown in heavy lines. For comparison, a perfect sinusoidal excitation voltage is shown in dashed lines. Obviously, even though the amplifier is operating over a perfectly linear dynamic characteristic curve, excitation such as that in Fig. 7 will cause serious distortion.

To minimize this grid-loading effect, the driver regulation must be made as good as possible. This is usually done by designing the driver stage so that it is capable of delivering two or three times as much power as that required to drive the grid of the amplifier. This keeps the impedance of the driver down. In addition, the input tank of the amplifier is shunted by a relatively low resistance. This resistance, shown as R in Fig. 1, absorbs considerable driver power, but reduces the wide fluctuations in grid circuit impedance. With this resistor, the input circuit impedance can change from a few hundred ohms only up to the value of the shunting resistance. Thus, the driver works into a more nearly constant load.

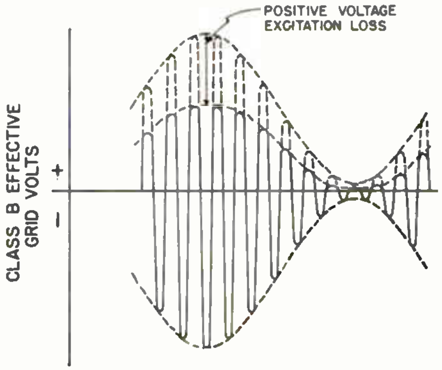


FIG. 7. How poor voltage regulation of the driver stage causes distortion of the class B amplifier excitation.

The grid-loading resistor may not be required if a special zero-bias tube is used in the class B stage. The operating bias for these tubes is very low, usually from 5 to 10 volts. The grid loads the circuit at all times, during the positive half of the input cycle so the impedance remains nearly constant during this half of the cycle. Since this is the only half of the input cycle we are interested in, it does not matter that the impedance changes during the negative half cycle.

Another way of assuring proper drive to the linear amplifier is to operate the amplifier as a class AB amplifier. If the amplifier is operated in class AB_1, the grid is not driven positive, so no grid current flows. The grid circuit impedance remains constant, and furthermore the driver needs to supply only an exciting voltage—it does not have to supply any power because no grid current flows.

Since the efficiency of a class AB_1 stage is low, a compromise arrangement is to operate the stage as a class AB_2 amplifier. Here the grid is driven positive during only part of the positive half cycle, and the power requirements are less than for class B operation. This enables us to get better driver regulation and hence a more linear output. Some modern linear amplifiers using tetrode tubes are operated as class AB amplifiers, usually class AB_2. Reasonably good efficiency can be obtained from these stages and the driving power requirements are low. Good efficiency can be obtained from class AB_2 tetrode stages because tetrode tubes have a high power sensitivity, that is, they can develop a high power output in the plate circuit with a low power input in the grid circuit.

PLATE CIRCUIT REQUIREMENTS

As in class C operation, the plate voltage in a class B or class AB linear amplifier has a decided influence on the power output. It must be high enough for maximum linear output without causing excessive plate dissipation. Naturally, the maximum rf peak voltage that can be developed across the load impedance is somewhat less than the B supply voltage.

With 100% modulation, during the negative half of the cycle of the rf exciting voltage, the tube is driven beyond cut-off, and the voltage developed across the load impedance is equal to the B supply voltage. During the positive half of the input cycle, when the tube is driven up close to the saturation point, maximum plate current flows, the tube plate resistance drops to a low value, and the plate voltage is very low. The maximum change in plate voltage is between this low value and the actual value of the B supply voltage. This maximum change occurs only during 100% modulation; for lower percentages of modulation, the plate voltage change is not as great.

The changes in the grid excitation voltage and in the output voltage are linear. In other words, a variation in the rf voltage applied to the grid will cause a corresponding variation in the plate voltage. This linear relationship depends upon the grid bias value. In a class C stage, the plate is driven to saturation in order to get a higher plate efficiency. In the saturation region, the excitation voltage and the plate voltage are no longer linear. To avoid the distortion that would result from this non-linearity, a linear amplifier is never driven into the saturation region.

If the amplifier has a truly linear characteristic, the plate current increase during the modulation crest is equal to the plate current decline during the trough. Hence, the average plate current read on a dc plate current meter is constant and should not change from no-modulation to 100% modulation. The average plate current will be equal to the peak plate current with no modulation when a linear amplifier is used to amplify a carrier and two modulation sidebands. In a later lesson you will see that in some applications of linear amplifiers the plate current does change with modulation.

POWER CONSIDERATIONS

In an earlier lesson you learned that the output power from a class C stage is proportional to the square of the plate voltage; in other words, if you double the plate voltage, the output power will increase four times. This relationship exists because the signal applied to the grid drives the tube from cut-off all the way to plate current saturation on every cycle. In a class B or a class AB linear amplifier, the output power is proportional to the square of the *exciting signal voltage*. With a constant load resistance, if the input signal voltage is doubled, the output power will increase four times because both the signal voltage and signal current developed in the output will double. Thus, the input signal has direct control over the output power.

The linear amplifier has a constant plate voltage and, when operating correctly, functions with a constant average plate current. Hence, it draws a constant amount of power from the high-voltage supply.

In your study of amplitude modulation, you learned that with 100% modulation, at the peak of the modulation the power of the carrier is four times the power without modulation. You have also learned that the plate input power of a linear stage remains constant whether or not modulation is applied. Therefore, the amplifier must have higher efficiency during modulation. (The higher the efficiency of an amplifier, the higher the power output.) When the stage is fed with an unmodulated carrier, its efficiency is 30% to 35%. For full 100% modulation its efficiency increases to 60% to 70%. The plate input power divides between the tube and the output load. With no modulation, two-thirds of the input power must be dissipated by the tube plate; at 100% modulation, only one-third must be dissipated by the tube plate. This shows that the linear amplifier tube runs cooler when it is delivering the most power output (high modulation percentage).

Since the modulation peaks of speech and music are often 10 to 20 times as great as the average signal level, the average grid excitation must be kept low. The average efficiency of a typical linear amplifier is usually not over 40%.

The power gain, which is the ratio of the output power to the driving power, of a linear amplifier is usually between 5 and 10 with triode tubes, and between 20 and 50 with tetrode and pentode tubes. The gain is low when triode tubes are used because triode tubes require a substantial driving power because of their low power sensitivity and also because the losses in the grid circuit are often quite high.

If a stage has a power sensitivity of 10, we will need a driving power equal to 1/10 of the power output. For example, if the power output of a linear amplifier with no modulation is 10 kw (10,000 watts) and the stage has a power sensitivity of 10, the driving power needed would be 10 kw ÷ 10 = 1 kw. If the efficiency of the linear amplifier is 33 1/3%, the input power to the stage would be 30 kw; 10 kw would be useful output and the other 20 kw would be dissipated as heat by the tube. If the driver is a plate-modulated class C amplifier with an efficiency of

66 2/3%, the plate input to the driver would be 1.5 kw (1500 watts). To plate modulate the driver we would need approximately 750 watts of audio power. Thus, we can 100% modulate the 10 kw output of the linear amplifier with only 750 watts of audio power.

On the other hand, if we tried to plate modulate an amplifier with an output of 10 kw we would need much more audio power. If the efficiency of the stage is 66 2/3% (a class C stage), the power input would be 15 kw and to 100% modulate the stage would require 7.5 kw of audio power. This is ten times the power needed for 100% modulation when we modulated the driver and operated the power output stage as a linear amplifier. Thus, the poor efficiency of the linear amplifier is compensated for by the lower audio power needed for 100% modulation. In high-power transmitters it is more economical to modulate one of the low power driver stages and then amplify the signal with linear amplifiers than it is to try to plate modulate a high power class C amplifier.

Adjusting a Linear Amplifier

Fig. 8 shows the schematic diagram of a typical single-ended class B linear stage. Notice the similarity between this circuit and the conventional class C amplifier stage. The plate and grid resonant circuits and neutralizing method are identical to those used for class C. The modulated rf signal is link-coupled to the grid of the amplifier from the plate circuit of the modulated stage. Thus, the modulated stage acts as the driver.

Resistor R1 connected across tuning capacitor C2 and part of coil L2 is the grid-loading resistor. It is used to prevent wide impedance variations

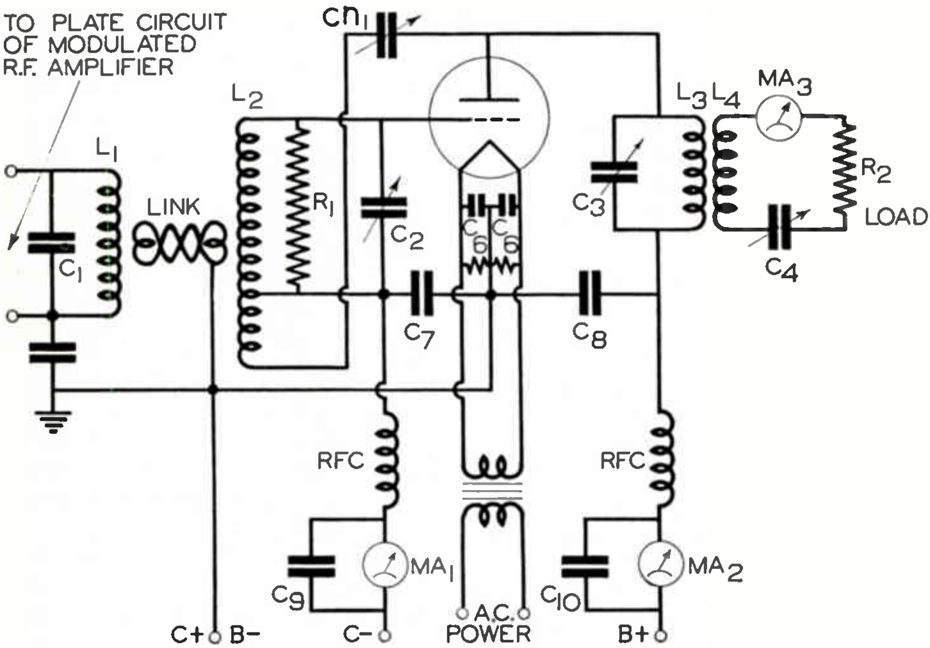


FIG. 8. A typical single-ended class B linear rf amplifier.

in the amplifier grid circuit, and thus present a more constant load to the driver stage. The bias voltage for the grid is obtained from a separate low-impedance supply.

As shown in Fig. 8 there are current meters in the circuit to indicate the operation of the amplifier. Meter MA_1 indicates the grid current, meter MA_2 indicates the plate current, and meter MA_3 indicates the load current. The meters in the grid and plate circuit are shunted by capacitors to prevent rf currents from damaging them.

INTERPRETING METER READINGS

From the readings on the grid current, plate current, and output current meters you can determine whether or not the amplifier is operating correctly. If the amplifier has been adjusted correctly and is being driven by a well regulated driver, the meter readings should be as follows:

With no modulation on the carrier, the grid current will be small and steady. The plate current will also be constant and its value should be that recommended for the rated power output. The reading on an antenna current meter should be constant. Its exact value of course will depend on the power output of the stage, the type of transmission line used to connect the amplifier to the load, and on how well the line is matched to the load. For the present, all you need be concerned about is that with no modulation the meter reading should be constant.

On 100% sine-wave modulation, the grid current reading should rise sharply to a maximum value. The plate current, however, should not change and will not if the amplifier is truly linear. Any change in the plate current reading is an indication of nonlinearity. The output load current in MA_3 should increase 22.5% with 100% sine-wave modulation.

With voice or music modulation, the grid current reading will rise and fall with the peaks of the modulation. The antenna current meter will also fluctuate, but the plate current should never vary from its normal steady reading.

Since the average power contained in voice and music is low, the antenna or load current will rise only a few percent higher than with no modulation. There will be a sharp rise in the load current value only on loud sustained passages.

The grid, plate, and output meter readings can also be used to localize defects and causes of distortion. The modulated signal fed to the linear amplifier stage must be free from distortion. Therefore, when checking any linear amplifier circuit, almost the first test to make is to see that the input signal itself is linear and undistorted. If it is distorted, check the modulated amplifier and the modulator stages. Be sure that their operating voltages and drive are correct, that the class C amplifier is not being over-modulated, and that the tuning and loading are correct for the stages.

If the input signal is not distorted, incorrect meter readings may indicate defects in the linear amplifier itself. Suppose the plate current reading increases and the pointer on the antenna current meter "kicks up" sharply when modulation is applied. A rise in antenna current is normal; the meter pointer, however, should not swing up or down sharply. (You may not notice abrupt changes in meter readings if a thermocouple meter is used.) This indicates a positive carrier shift (upward modulation). The plate and antenna current increase can be due to excess bias on the grid, which causes the tube to operate further down on the knee of the characteristic curve. This will cut off the trough of the modulated signal, causing the average plate current to rise. Parasitic oscillations, incomplete neutralization of the amplifier, and improper tuning and loading of the stage can also cause positive carrier shift.

Negative carrier shift (downward modulation) occurs when the average plate current and the antenna (output) current decrease with modulation. This is due to a defect that cuts off or distorts the peaks of the modulated signal. If an excess amount of excitation is applied to the grid, the tube will be driven to saturation, and the positive peaks will be distorted. Poor regulation in the bias or high-voltage power supplies can also cause negative carrier shift because the grid or plate voltages are not high enough to provide the peak values of the amplified waveform.

Also, if the load that the linear amplifier presents to the driver stage varies widely during the input signal cycle, the peaks of the modulated waveform at high modulation levels will be cut off. The output signal will be distorted, as indicated by a decrease in the antenna meter reading. The purpose of the shunt resistor across the coil in the grid circuit is to prevent the load on the driver stage from varying. A higher than normal load impedance for the class B linear stage or incorrect tank circuit tuning will also cause the modulation peaks to be cut off, and cause the readings on the plate and antenna current meters to decrease.

A decrease in the grid and plate current meter readings over a period of time could indicate a loss of efficiency in the driver amplifier or some defect earlier in the transmitter, causing a loss of excitation power to the linear amplifier. Also, a gradual decrease in the grid and plate current meter readings is often an indication of a weakening linear amplifier tube.

Thus, to get an undistorted output signal from a class B linear amplifier stage, the operating voltages, grid drive, and loading must be correct. Also, the stages preceding the linear amplifier must be operating properly to produce an undistorted signal of the correct amplitude to the stage. Usually the recommended operating voltages are applied to the stage and then the grid bias and the load are varied slightly on either side of the recommended values to get the greatest undistorted power output at the best efficiency. This is a part of the adjustment procedure; let us go through the complete procedure now.

ADJUSTMENT PROCEDURES

Before the class B stage itself can be adjusted, the preceding stage must be adjusted, and the class C modulated stage and the class B stage itself must be neutralized. During adjustment the grid bias and plate voltages are not set to their full final values to be sure tubes and other parts are not damaged, and to prevent the possibility of interfering with other broadcast stations. The procedure for adjusting a class B stage using triode tubes is as follows:

1. Apply grid bias of one-half its final intended value.

2. Apply a low-amplitude unmodulated rf signal to the input.

3. Increase the level of the unmodulated rf signal by tightening the coupling between the driver and the amplifier until there is grid current flow.

4. Readjust the plate tank circuit of the driver stage, and the grid tank circuit of the amplifier stage to resonance by tuning each for maximum grid current.

5. Apply plate voltage of one-half the final intended amplitude.

6. Quickly tune the plate tank to resonance as indicated by a dip to

minimum in the plate current.

7. Couple the load circuit to the amplifier by increasing the coupling between the output tank circuit and the coupling network to the antenna or the next stage; as coupling is increased, readjust the tank capacitor to resonance. Continue increasing the coupling until the plate current minimum is approximately three times the minimum without a load.

8. Apply normal plate voltage and extended cut-off grid bias to the stage.

9. Check the plate current and the output current readings. If they are both too high, reduce the excitation to the stage. If the plate current is high and the output current is low, reduce the excitation and increase the load coupling; then, increase the excitation again. If the plate current is still too high, the grid bias may be too low. In this case, to get the best possible linearity, it may be necessary to vary the bias slightly above or below the value recommended by the transmitter manufacturer.

If you get good linearity for a bias voltage near the recommended value, and the input and output power are correct with full excitation, the amplifier is properly adjusted.

If the plate current is still too high or the plate input power and the output power are below normal, the load impedance is too high. Readjust the load coupling and make a new set of linearity checks.

10. Make a final check of all meter readings with modulation applied. The grid current meter reading should change rapidly with modulation, and the plate current meter should remain steady. The load current meter reading will increase very slightly with normal modulation. However, with sustained 100% sine-wave modulation, the load current reading should rise 22.5%.

You can also use an oscilloscope to check the linearity of a linear amplifier system. To do so, you first observe the modulation envelope at the output of the driver to make certain that an undistorted signal is being applied to the linear amplifier. Then, you use the oscilloscope to check the performance of the linear amplifier itself, with 100% sinusoidal modulation applied. The major advantages of oscilloscopic checks are that they are instantaneous and do not require a tedious step-by-step measurement procedure. The oscilloscope can be used to monitor the output of the transmitter during normal operation, to provide an immediate indication of any non-linearity. You will receive detailed instructions on how to use the oscilloscope in a later lesson.

The adjustment procedure for a class B linear amplifier using pentode or tetrode tubes is the same as for triode tubes, except that some load should be connected to the stage before plate and screen voltages are applied to the stage. If you attempt to tune a pentode or tetrode rf stage without a load connected to it you may destroy the tube.

Variations in Linear Amplifiers

So far, we have been discussing single-ended linear amplifiers. Linear amplifiers may also be operated in push-pull and can use multi-grid tubes as well as triodes. There are differences in the operation of single-ended and push-pull stages, but the basic characteristics and the purposes for which they are designed are the same. We will show several examples.

PUSH-PULL LINEAR AMPLIFIERS

A class B push-pull stage using triode tubes is shown in Fig. 9. Resistors R1 and R2 load the grid and establish proper driver regulation. The two radio-frequency chokes shunting the loading resistors provide low resistance dc paths for the grid current so there will not be any unwanted grid-leak bias. Regular push-pull cross neutralization is used, and the metering arrangements are similar to those for a single-ended stage.

Advantages. The push-pull class B linear amplifier in Fig. 9 has some advantages over the single-ended type. Twice as much power can be obtained from a push-pull stage, and the Q of the tank circuit can be lower than in a single-ended stage, if the push-pull tubes are properly balanced. The Q of the tank circuit of the single-ended stage must be higher, because the single-ended class B stage depends upon flywheel effect in its plate tank circuit to establish one alternation of the rf output cycle. If the Q is too low, the flywheel effect is inadequate and the modulation envelope is distorted. This means that the second harmonic content becomes high and filters must be used to reduce second harmonic radiation. Hence, a compromise must be made in a single-ended stage between power output (which depends on the circuit loading) and the tank circuit Q.

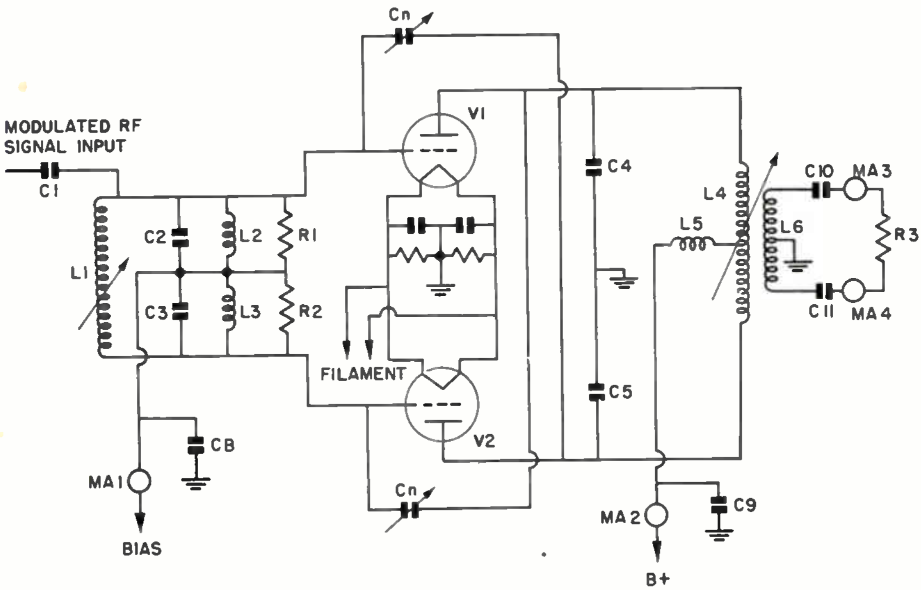


FIG. 9. Push-pull class B linear amplifier.

In a push-pull class B stage, one tube supplies energy during the positive half-cycle. Its plate current then falls to zero, and the other tube supplies energy for the negative half-cycle. Consequently, each tube contributes one-half or 180° of the complete rf cycle. Since the tubes supply energy to the tank circuit for the entire signal cycle, the tank circuit flywheel effect is not essential. The load can be coupled into the tank circuit for optimum power output and efficiency and for the best linearity. Thus, the power output is higher and the harmonic distortion is less for push-pull class B operation.

The circuit Q is still important in a push-pull stage. A low Q tank circuit will pass a considerable amount of harmonics, but by balancing the tubes properly, much of the even-harmonic energy can be eliminated. A low Q circuit will reduce the tank circuit losses and raise the efficiency. In a high power transmitter even a 5% loss in output power is a considerable amount, when it just heats the coil.

There is also less sideband clipping in a push-pull stage than in a single-ended stage. In a single-ended stage, to maintain an adequate flywheel effect, the Q of the tank circuit must be quite high and sharp. Thus, some of the higher frequency sidebands may be clipped off. With push-pull operation, the Q can be considerably less, and therefore, the tank circuit response is broad and sideband clipping is less likely to occur.

The load impedance in the push-pull stage must be chosen correctly to get the best linearity, power output, and efficiency. For a stage using triode tubes, the load impedance is about twice the plate resistance of one of the amplifier tubes. This is considerably less than the load impedance required for a push-pull class B audio amplifier in which the load impedance is four times the plate resistance of one of the tubes.

Adjustments. The adjustment procedure for the push-pull class B linear stage is similar to that described for the single-ended stage. The preliminary adjustments are made, as described previously, and then the grid and plate voltages are set to half their normal values. After the class B stage adjustments are completed, full normal operating voltages are applied. Usually small touch-up adjustments are required for best operation.

LINEAR AMPLIFIERS USING MULTI-GRID TUBES

Tetrode, pentode, and beam-power tubes are also used in linear amplifiers. Multi-grid tubes give better power gain, and less driving excitation is required for a given power output. Also, the additional electrodes give better shielding, so neutralization is not as much of a problem.

However, to prevent non-linearity, the voltage applied to the screen grid must be very well regulated. Variations in screen voltage as the modulation changed would have the same effect as variations in grid bias in a triode linear amplifier. Variations in screen voltage would cause the plate current and power output of the amplifier to change in a non-linear manner, causing a distorted modulation envelope. In low-powered amplifiers, it is common to use small voltage-regulator tubes, and in higher-powered amplifiers, the screen voltage is obtained from an extremely well-regulated power supply.

The multi-element tubes are often operated in class AB_1 or AB_2. An advantage of AB_1 operation is that no driving power is needed because the tubes will not draw grid current. Driver regulation is much less of a problem because the grid circuit in-

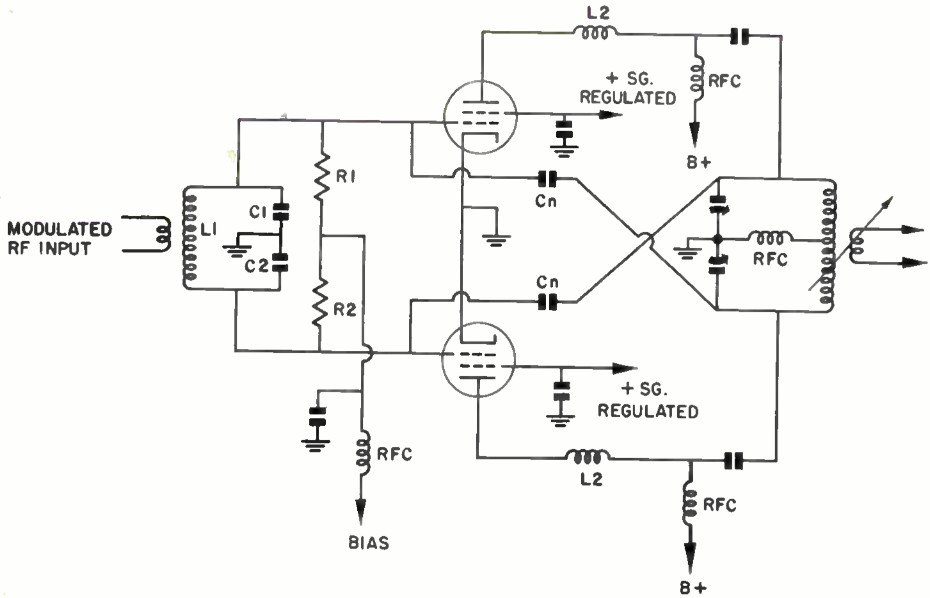


FIG. 10. A class AB_1 linear amplifier using tetrodes.

put impedance does not vary as in a class B linear amplifier, which draws a very high grid current. A class AB_2 amplifier draws some grid current, but not as much as class B.

The class AB_1 linear amplifier may have two tubes in push-pull, or four or eight tubes in a push-pull parallel combination. Fig. 10 shows a linear amplifier with two tubes in push-pull operated in class AB_1. Resistors R1 and R2 in the grid circuit load the input, increase the circuit bandwidth, and broaden the response of the input resonant circuit so that the input does not have to be tuned. Thus, capacitors C1 and C2 can be fixed rather than variable capacitors.

Neutralization of the multi-grid tubes is not exacting, and two small fixed capacitors, such as those labeled C_n in Fig. 10, are often used.

The output circuit has parasitic chokes and a split-stator tuning capacitor, making it similar to most push-pull class C amplifier output circuits. The class AB_1 operation depends upon the loading, biasing, and excitation.

The screen voltage is obtained from a voltage-regulated supply.

GROUNDED-GRID AMPLIFIER

A linear rf amplifier may also use a grounded-grid circuit, as shown in Fig. 11. The grounded-grid circuit gives somewhat better linearity than a grounded-cathode circuit.

The exciting voltage is applied between the cathode and ground. The input of the grounded-grid amplifier presents a low impedance load for the driver. Thus, it is not necessary to

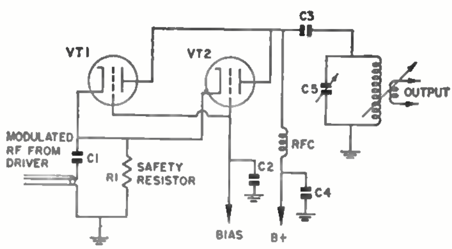


FIG. 11. A grounded-grid linear amplifier.

use a tuned circuit in the input. The driver can be connected directly between the cathode and ground, if a capacitor is used for proper dc isolation between the driver and the linear amplifier. The cathode resistor is inserted to prevent a high voltage from appearing between the cathode and ground in case an open develops between the driver and the cathode circuit of the amplifier.

A grounded-grid amplifier is very stable; much more stable than a conventional grounded-cathode amplifier. The control grid is at rf ground potential and acts as both a control element and a screen between the cathode and plate. Therefore, the stage usually does not have to be neutralized. This stage also operates better at frequencies above 50 mc than the grounded-cathode amplifier does. If it is used at these high frequencies, the inductance of the leads in the input and output circuit and the grid-to-plate capacity can provide enough feedback to cause instability and perhaps oscillation. In this case, the stage must be neutralized.

When filament-type tubes are used, the grounded-grid amplifier requires a special filament transformer that has a very low capacity between the primary and secondary windings. The capacity between the windings will shunt the signal to ground.

The stage requires more power from the driver than does a grounded-cathode stage. This additional power, however, does not represent a loss, because the extra power actually appears in the plate circuit of the stage. Thus, the output power comes partly from the driver and partly from the amplifier itself.

DOHERTY AMPLIFIER

We have learned that the average efficiency of a conventional linear amplifier is seldom better than 40%. With 100% modulation, we can get an efficiency of about 70% on modulation peaks, because then the amplifier is operated near saturation. If we could operate the amplifier near saturation at all times, its efficiency would be high, but then, on modulation peaks, the output would be distorted, because the peaks would be flattened out—the amplifier would not be able to supply the additional power demanded of it on peaks.

It is possible to get an average efficiency of from 60% to 65% by using a two-tube circuit like that shown in Fig. 12. This amplifier is called the Doherty amplifier. The circuit is arranged so that the unmodulated signal from the driver drives

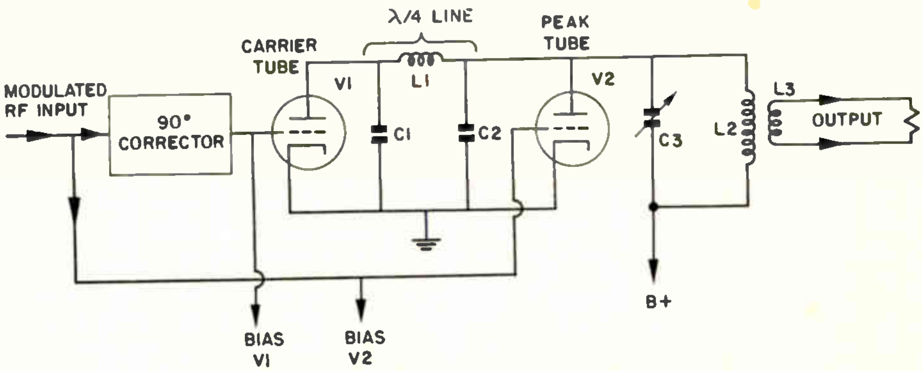


FIG. 12. A basic Doherty system.

one tube, called the carrier tube, to saturation while the other tube, called the peak tube, passes little or no plate current. Thus we have one tube operating at peak efficiency and the other tube wasting little or no power.

The two tubes are of the same types, and therefore will have almost identical characteristics. The carrier tube is biased at the extended cut-off value as a true class B linear amplifier. The peak tube has about twice as much bias applied to it so that it operates more as a class C amplifier than a class B amplifier. When the unmodulated signal from the driver is fed to this tube it barely drives it into the conducting region.

The load on both tubes is the parallel resonant circuit made up of C_2 and L_2 and the load connected to it through L_3. The output of the peak tube, V2, is connected directly to this load. The parallel resonant circuit is designed so that the load impedance will be about half what would normally be used for one of the tubes operated in the normal way as a Class B amplifier. Thus, if the load should be Z_L, the load impedance that V2 will see will be $Z_L/2$.

The load the carrier tube, V1, will see will be quite different. V1 is connected to the load through a network consisting of C_1, L_1, and C_2. This network is called an artificial transmission line because it acts like a transmission line. By selecting the values of C_1, L_1, and C_2, we can make the network act like a transmission line one quarter of a wavelength long. (The symbol λ that you see on the diagram is the Greek letter Lambda, and is used as an abbreviation for wavelength.) You will study transmission lines in detail later; for the present you need know only a few things about them.

Transmission lines have a characteristic which is known as the "surge impedance" or "characteristic impedance." It depends on the size of wire used, the spacing between the wires, and the material between the wires. The characteristic impedance is represented by the symbol Z_0. When a load Z_1 is connected across one end of a quarter-wave line, the impedance that is seen at the other end, which we will call Z_s is given by the formula:

$$Z_s = \frac{Z_0^2}{Z_1}$$

If we make the characteristic impedance of this line equal to the load that the tube should work into for normal class B operation (Z_L), and the actual load equal to half the normal class B load, then the impedance looking into the line from V1 becomes

$$\begin{aligned} Z_s &= \frac{Z_L^2}{\frac{1}{2}Z_L} \\ &= \frac{2Z_L^2}{Z_L} \\ &= 2Z_L \end{aligned}$$

Thus, the load impedance into which V1 is working is equal to twice the normal load for class B operation.

Another characteristic of a transmission line is that it delays a signal traveling through it. With a quarter-wave line between V1 and the load there will be a one-quarter cycle delay. We refer to this as a 90° delay or a 90° phase shift. One other characteristic of a quarter-wave line is that it inverts any change in the load connected across the output. If the impedance of the load is increased, at the input the line acts as if the impedance had been reduced, and if the impedance of the load at the output is reduced, at the input it acts as if it had been increased. Now let's go ahead and see exactly how this amplifier works.

When an unmodulated signal is fed from the driver to the amplifier, the signal does not drive the grid of V2 hard enough to cause this tube to conduct any appreciable amount of plate current. The signal from the driver is fed to the grid of V1 through a phase-shifting network that advances the signal phase so that the signal applied to the grid of V1 leads the signal fed to the grid of V2 by 90°. This signal drives the grid of V1 hard enough to drive this tube to plate-current saturation. Thus the tube operates with good efficiency. However, because the load impedance is twice the normal class B load impedance, the power output will be only about half of what would be obtained with the correct load impedance. The output from V1 is fed through the artificial line, which delays it 90°, to the load. Since the signal has been advanced 90° and then retarded 90°, the two phase shifts cancel, so the signal reaching the load from V1 will be in phase with any signal reaching it from V2.

When the driving signal is modulated and its amplitude starts to increase, V2 begins to conduct. This results in an increase in the signal voltage across the load because the load current increases. The quarter-wave transmission line, which is also connected to the load, sees this higher voltage. Since the extra current is not coming from this line, the higher voltage across the load has the same effect on the line as an increase in load impedance. The quarter-wave transmission line inverts this change so that at its input the impedance decreases. This means that the impedance of the load that V1 is working into goes down, and V1 supplies more power to the load.

With 100% modulation, the amplitude of the driving signal will drive V2 to saturation. Since its load impedance is only ½ the normal value of Z_L, it will supply twice the power it would if it were operated at saturation with the normal load impedance. This power supplied to the load causes the voltage across it to increase so that the value of the load impedance connected at the end of the quarter-wave transmission line appears to have doubled. This means that it is now equal to Z_L. The impedance at the other end of the line, to which V1 is connected, becomes:

$$Z_s = \frac{Z_L^2}{Z_L} = Z_L$$

Therefore V1 is now working into the load impedance it should work into for normal class B operation, and the power output from it will be twice the value with no modulation. Thus both V1 and V2 are supplying twice the power to the load that V1 supplied to it with no modulation, so the power output has increased four times, as it should with 100% modulation.

You might wonder how we can say that the efficiency of this amplifier is good when with no modulation one tube is supplying only half the normal class B power to the load and the other tube none at all. You must remember that efficiency is not a measure of the amount of power output. If the ratio of the power output to the power input is high, the stage is efficient regardless of how much power it actually delivers to the load. The efficiency simply tells you how much of the power that the stages take from the power supply is converted into a useful output signal. If the efficiency is 75% then 75% of the power input is converted to useful output and only 25% wasted. If an amplifier with 75% efficiency has a power input of 100 watts, the output will be 75 watts; the wasted power 25 watts. On the other hand, a 1000-watt amplifier with an efficiency of 30% would put out 300 watts of useful power, which is more than the preceding example, but in doing so would waste 700 watts.

Although the efficiency of the Doherty circuit is high, it has some important disadvantages. At high frequencies, the capacity in the phase-correcting circuit and the quarter-wave line is so small that even the stray capacity in the wiring and the tube capacities can affect the operation of the stage. Also, as the operating frequency increases, it becomes more and more difficult to maintain the proper phase relationships in the circuits.

Distortion will occur when a signal with a low modulation percentage is amplified by the Doherty system, because the peak tube then operates near cut-off where the characteristic of the tube is the most non-linear. Operating the carrier tube near saturation can also produce distortion. Other disadvantages of the Doherty circuit are that it presents a varying load to the preceding amplifier, and the circuit is very difficult to adjust. For these reasons, the Doherty amplifier is no longer being manufactured. However, you may find them still in use in some broadcast stations.

The class B linear circuits that we have discussed in this section are typical of those found in both low-power and high-power transmitters. The power output, of course, depends on the tube type, the component rating, and the operating voltages and currents used in the stage.

FEEDBACK SYSTEMS

You have already studied both regenerative and degenerative feedback. You will remember that when a signal is fed from the output of one stage back to the input of the stage, or to a preceding stage, we have feedback. If the polarity of the signal that is fed back is such that it aids the original signal, we have regenera-

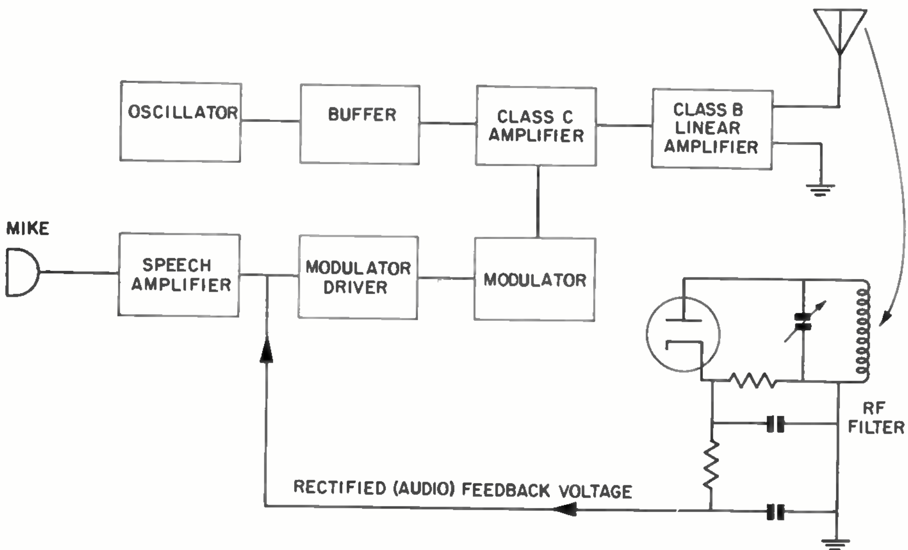


FIG. 13. Over-all rf to af feedback that reduces distortion, hum, and noise in the modulated class C and class B amplifier stages as well as in the audio-frequency amplifiers on the modulator.

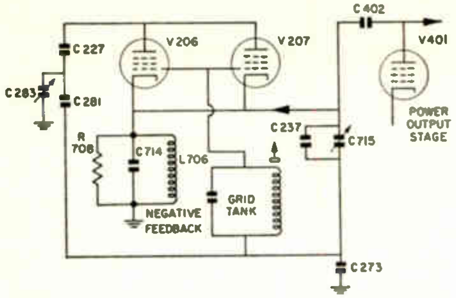


FIG. 14. The Collins rf feedback arrangement.

tive feedback, and if it opposes the original signal we have degenerative feedback, which we usually call inverse feedback. Regenerative feedback is used in oscillators. Inverse feedback is used to help reduce noise, hum, or distortion produced in amplifiers.

Inverse feedback systems are used in transmitters to reduce noise and distortion. A feedback system used in a transmitter is shown in Fig. 13. This feedback will reduce noise and distortion produced in the modulator driver, the modulator, the modulated class C stage and in the linear amplifier.

In this system there is a small resonant circuit that is tuned to the output frequency of the transmitter and loosely coupled to some portion of the output system such as the final tank circuit, or the transmission line, or picks up radiation directly from the antenna. It picks up the modulated signal, demodulates it, and filters out the rf. The demodulated signal, which is made up of the original audio signal plus any noise or distortion that has been added is then fed back to the speech amplifier or modulator of the audio system. It is reinserted 180° out of phase with the original signal.

In this system, noise or distortion originating in the modulator and its associated driver and amplifier stages, and distortion components contributed by the modulation process can be reduced to very low values.

Fig. 14 shows a system in which rf or carrier feedback is used. This is the Collins KWS-1 transmitter.

RF energy is fed back through capacitor C402 to the cathode circuit of the driver. Notice that the driver consists of two tubes, V206 and V207, connected in parallel. The feedback voltage developed between the cathodes of the driver and ground is 180° out-of-phase with the driver input signal. Hence, the feedback link includes the driver and the linear amplifier output stage. The feedback not only corrects distortion, but also improves the driver regulation, insuring a more linear operation.

The amount of feedback depends on the relative reactance of capacitors C402 and C714. Coil L706 is a radio frequency choke that provides a dc return for the cathodes. It is loaded by resistor R708 to prevent oscillation.

Outphasing Modulation System

A system of modulation in which linear amplifiers are not required has recently become more widely used. It is called the "outphasing" system or "phase-to-amplitude" modulation system.

As you have learned, in low-level modulation systems, the rf power amplifier stages must be operated linearly or the output will be distorted. However this is not the most efficient form of operation. In high-level modulation systems, the carrier is amplified before it is modulated, which eliminates the need for linear amplifiers, but much more audio power must be supplied to the class C stage to modulate it. Therefore, in low-level systems it costs more to amplify the carrier; in high-level systems it costs more to develop the high audio power required. The outphasing system combines some of the advantages of each. Let us see how.

BASIC SYSTEMS

Fig. 15A shows two ac generators connected in series across a load. If the two generators are exactly in phase, the two voltages they produce, E_1 and E_2, will add. For example, if each one is 10 volts, the total output will be 20 volts. This is shown by the vector diagram in Fig. 15B. The Greek letter phi (ϕ) is used to mean phase angle. Here it is shown as 0°.

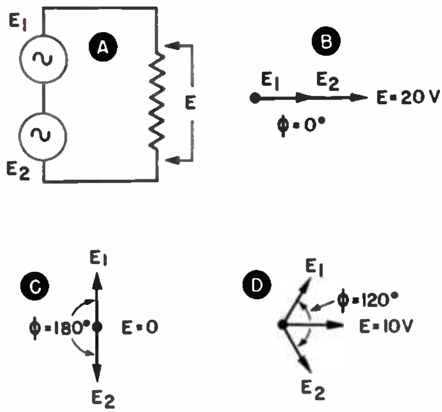


FIG. 15. Two generators connected in series across a load as in A, will produce varying voltages across the load, depending upon their phase relationship. B shows the output if they are in phase; C shows them 180° out of phase; and D shows them 120° out of phase.

If E_1 and E_2 are exactly opposite in phase, the voltages will cancel as shown in Fig. 15C. The phase angle is 180°, and the output voltage E is 0.

If the phase angle is anywhere between 0 and 180°, the output voltage will be somewhere between 0 and 20 volts. For example, if the phase angle is 120° as shown in Fig. 15D, the output voltage will be the vector sum of E_1 and E_2, or 10 volts. If the phase angle is varied, the output will also

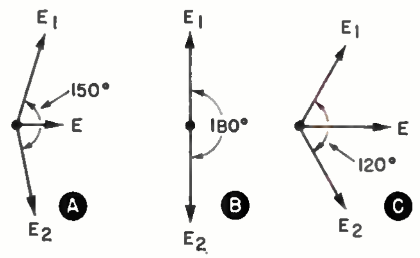


FIG. 16. Vector diagram showing how the combined output of two amplifiers depends upon the phase relationship between them.

vary in the same manner.

The generators in Fig. 15A can be replaced by vacuum tube amplifiers operating on the same frequency but out of phase by some value between 90° and 180°. If the amplifiers are then modulated with an audio signal that increases and decreases this phase difference between them in proportion to the amplitude of the audio signal, and if their outputs are then combined, the signal obtained across the output of the two amplifiers will be proportional to the modulating voltage. We will have 100% modulation if the modulating signal varies the phase of the carrier by the amount of the original phase shift away from 180°.

For example, if the amplifiers are operated with a phase difference of 150° without modulation (30° away from 180°), and 100% modulation is applied, the output will vary as shown in Fig. 16. The length of E indicates the amplitude of the output. When no modulation is applied, the carriers are 150° out of phase, and the output is equal to their vector sum, as shown in Fig. 16A.

On one half of the audio cycle, the phase difference will increase 30°, or up to 180°. The two carriers will cancel, and the output will be zero, as shown in Fig. 16B. On the other half of the audio cycle, the phase difference is decreased by 30°. This approximately doubles the output as shown in Fig. 16C.

Fig. 17 shows a block diagram of such a system. The output of an rf source is split between two amplifier branches. A phase shift of less than 180° is introduced into one branch so that the carrier will not be completely cancelled. The first tube in each string is a phase modulator. It is designed so that when an audio voltage is applied, the phase of its output will vary in step with the amplitude variations of the audio voltage. You will

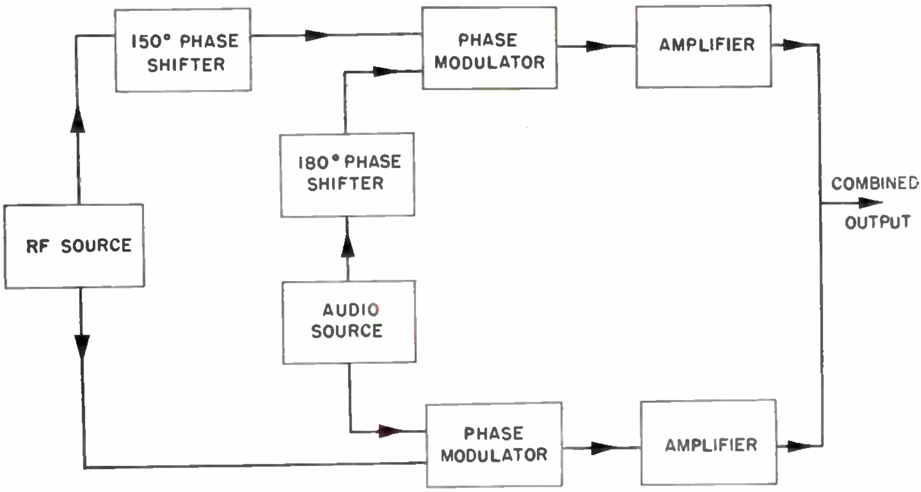


FIG. 17. Block diagram of basic outphasing modulation systems.

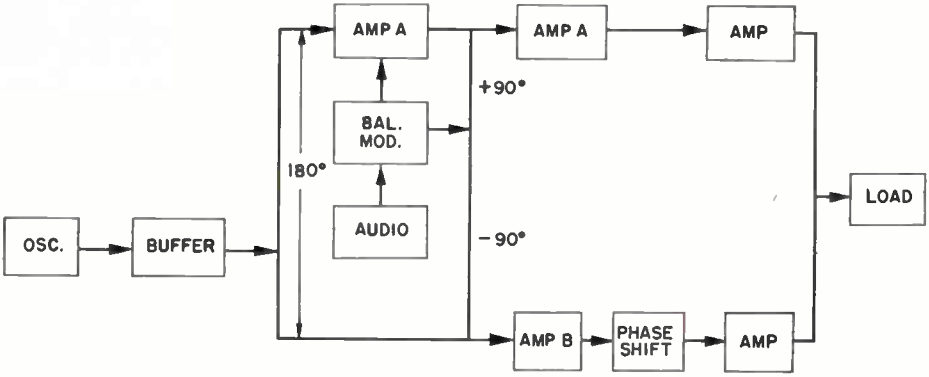


FIG. 18. Another basic outphasing modulation system.

learn how this is done in a later lesson on phase modulation. The audio voltages fed to the two phase modulators must be 180° out of phase with each other, or their effects would cancel.

The outputs of the two phase modulators are amplified, and the amplified outputs are combined.

The modulation envelope will not be distorted, regardless of whether the amplifiers are linear or not, because it is the phase not the amplitude of the signal that is varying in accordance with the audio, and the linearity of an amplifier does not affect the phase of the output.

Fig. 18 shows another arrangement for outphasing modulation. Here we have an rf source consisting of an oscillator and buffer stage. The output is fed to two amplifier strings, 180° out of phase. The audio signal is fed to a balanced modulator, which produces a double-sideband signal that is substantially free of carrier. The output of the modulator is divided between the two amplifier branches. One part is shifted +90° in phase and combined with the output of amplifier A, and the other part is shifted −90° in phase and combined with the output of amplifier B.

A small phase shift is introduced in the lower branch so that the carriers in the two branches will not cancel each other. After amplification, the two carriers are combined in the output. Since they are almost 180° out of phase with each other, the resultant carrier will be almost 90° out of phase with the carrier in each branch. Since the modulation sidebands were shifted 90° in phase at the output of the balanced modulator, they will be practically in phase in the output of the amplifier. Thus the original modulation appears in the

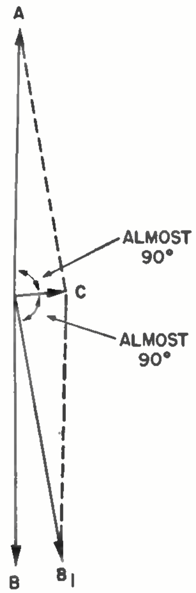
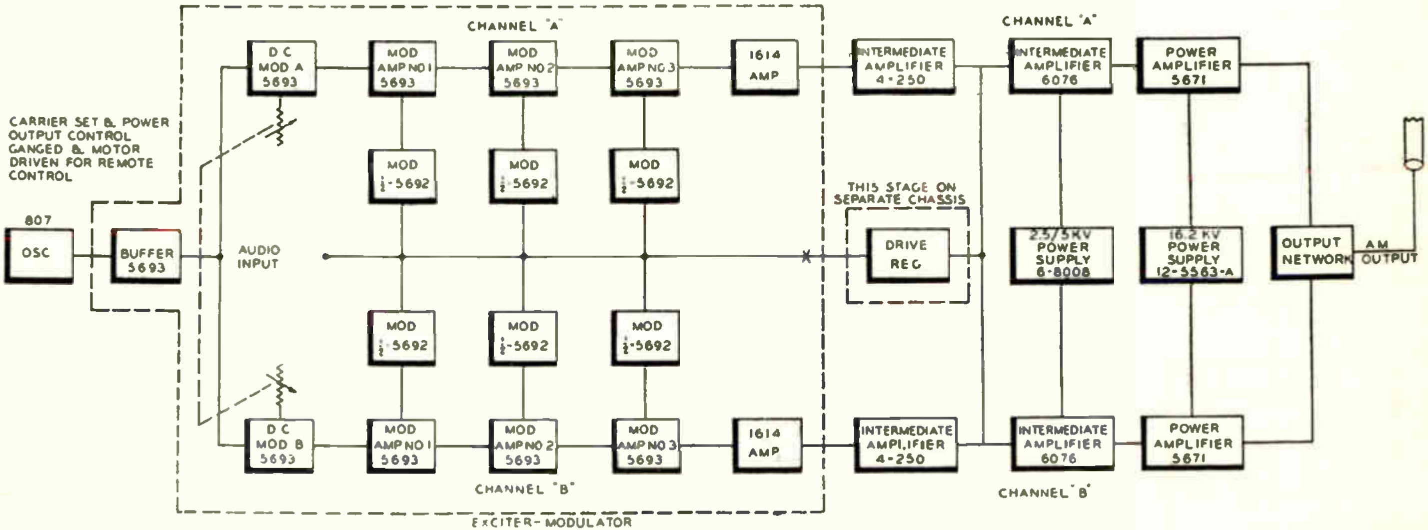
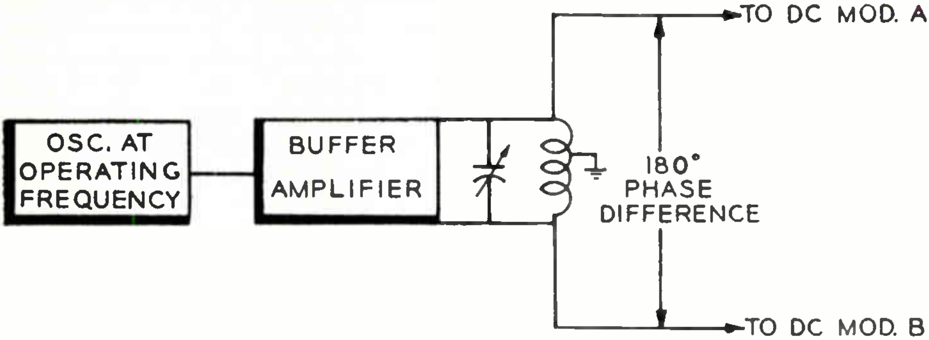


FIG. 19. Relationship of carriers in Fig. 18.



Courtesy RCA

FIG. 20. Block diagram of RCA Ampliphase transmitter.



Courtesy RCA

FIG. 21. Exciter-modulator circuit of Fig. 20.

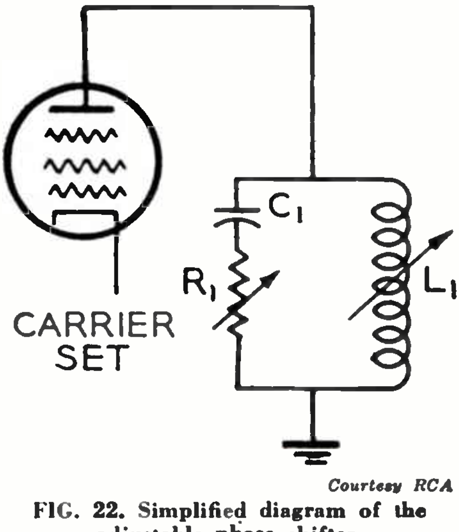
output. Fig. 19 shows the vector diagram for the carriers. A is the original carrier fed to amplifier A. B is the original carrier fed to amplifier B. B_1 is the carrier in the lower branch after being shifted slightly in phase. C is the resultant carrier in the output. As you can see, it is almost 90° different in phase from the carriers in the two amplifier branches.

THE RCA AMPLIPHASE TRANSMITTER

A modern commercial transmitter that uses this "phase-to-amplitude" system of modulation is the RCA 50-kw "Ampliphase" transmitter. A block diagram of this transmitter is shown in Fig. 20. Its operation is essentially similar to that of the circuits we have studied..

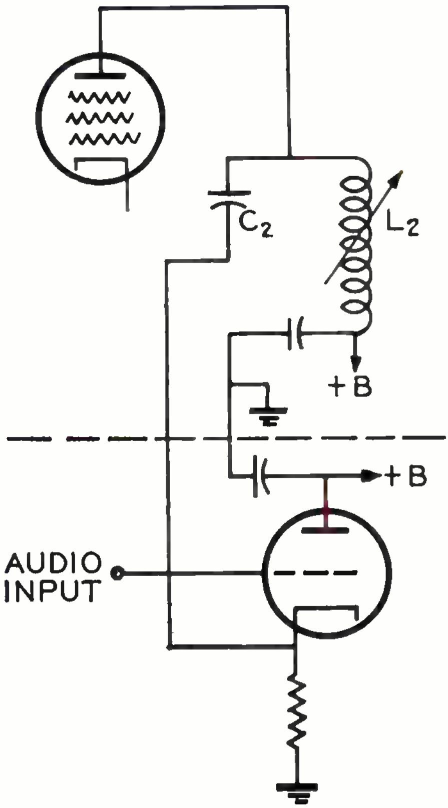
The output of a single crystal oscillator is fed to a buffer amplifier with a push-pull output tank, as shown in Fig. 21. Thus, the carrier wave is split between two rf amplifier channels, and the signal supplied to one chain is 180° out of phase with that supplied to the other. Since these two signals are 180° out of phase, no output would be obtained if they were impressed on a common load. However, if the phase difference is made less than 180°, some output will be obtained, the amount depending upon the phase angle.

So that the phase difference will be less than 180°, the first stage in each amplifier is an adjustable phase-shift amplifier. These are dc modulator A and dc modulator B in Fig. 20. A simplified diagram of the circuit is shown in Fig. 22. The values of L_1, C_1, and R_1 are chosen so that when R_1 is set to one end of its range, there will be a phase shift of +25°, and when it is set to the other end of its range, there will be a phase shift of −25°. Setting the phase-shift amplifier in one chain for a +22.5° phase shift, and the phase-shift amplifier in the other chain for a −22.5° phase



Courtesy RCA

FIG. 22. Simplified diagram of the adjustable phase shifter.



Courtesy RCA

FIG. 23. Simplified diagram of the modulated amplifier and modulator.

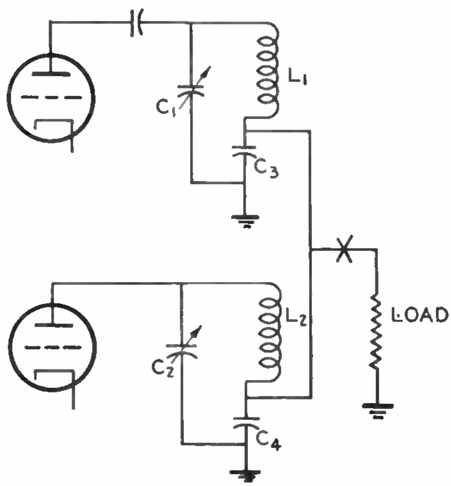
shift will give a total phase shift of 45°, and the two carriers will be 135° apart in phase instead of 180°.

The next three stages in each channel are called modulated amplifiers. They are quite similar to the adjustable phase-shift stages, except that instead of a variable resistor in the plate tank circuit there is a triode tube, as shown in Fig. 23. The triode tube acts as a variable resistance when an audio signal is applied to its grid. The audio signal applied to each modulator tube produces a phase-modulated signal in the tank circuit of its corresponding modulated amplifier.

Following the modulated amplifiers there is a conventional amplifier stage providing isolation and drive to the first intermediate power amplifiers. The signals are further amplified and then combined in the output to give an amplitude-modulated output. Fig. 24 shows a simplified diagram of the power amplifier output circuit.

The relationship between the carriers in the two channels is shown in vector form in Fig. 25. The two vectors A_1 and B_1 show the relationship of the carriers in the two branches without modulation. C_1 represents the output with no modulation. Vectors A and B show the carriers during the modulation troughs with 100% modulation, and A_2 and B_2 show the carriers on the peaks of 100% modulation. On the troughs the output is zero, and on the peaks, the output is doubled, as shown by C_2.

A very important consideration in designing transmitters is the amount of time a transmitter must be off the air if anything goes wrong. To minimize this time, the RCA Ampliphase transmitter is designed with two complete oscillators and two complete exciter-modulator sections. Either of these can be switched in with only a



Courtesy RCA

FIG. 24. Simplified diagram of the power amplifier output circuit.

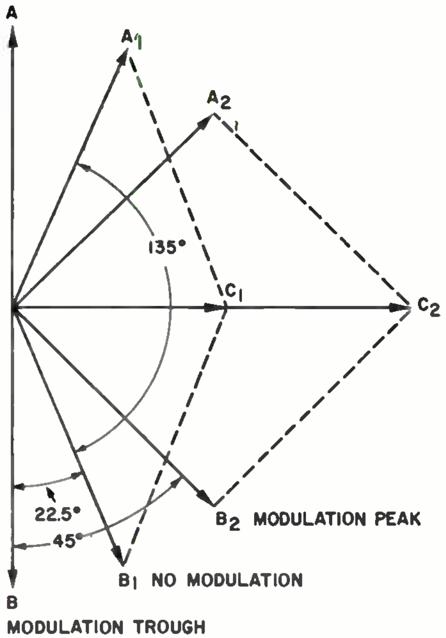


FIG. 25. Relationship of the carriers in the two amplifier channels of Fig. 20 for no modulation and 100% modulation.

momentary loss of carrier, and without cutting off the dc supply to the transmitter. Fig. 26 shows the switching arrangement.

Provision is also made for remote control of the transmitter. Because of the extra oscillator and exciter-modulator sections that can be switched in at a moment's notice, one of the big disadvantages of remote control operation is eliminated, that is, of having to lose time on the air while a maintenance man is sent to the transmitter.

Switching from local to remote control is accomplished with a single transfer switch. A safety feature is that this switch can be operated only at the transmitter. This eliminates the possibility that someone working on the equipment will be endangered by someone operating it from the remote point.

Meter readings for the total plate current, output plate voltage, driver plate voltage, carrier level, and both output cathode currents are repeated at the remote control point.

Broadly tuned band-pass coupling circuits are used to insure stability. A special transformer-type of neutralization circuit is used to make the final amplifiers completely broadband.

LOOKING AHEAD

Class B linear amplifiers are important because they are used in many AM radio broadcast stations. They are used because it is more economical to modulate a low-power class C stage and then amplify the modulated signal than it is to modulate a high power class C stage.

Remember that class B linear amplifiers are operated at extended cutoff bias. If the amplifier is properly adjusted, the plate current will not change when the amplifier is modulated. If there is a meter in the transmission connecting the transmitter to the antenna, the current reading on this meter should increase 22.5% with 100% modulation.

Distortion will be produced in a linear amplifier that is overdriven because the plate current is not able to follow the grid voltage variation if the tube is driven beyond saturation. Distortion may also be due to improper bias on the linear amplifier, or poor driver regulation.

Most linear amplifiers that you en-

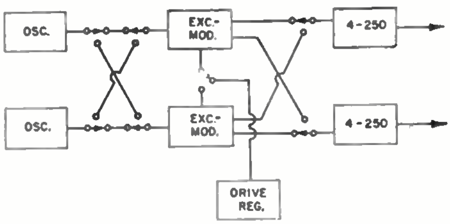


FIG. 26. Arrangement for switching in extra oscillator or exciter-modulator in Fig. 20.

counter in AM broadcast work will be class B amplifiers. However, class AB linear amplifiers are also used, particularly in single sideband applications. You will study these amplifiers later, and also class B amplifiers for TV. When you study these amplifiers, you will find that the plate current of these linear amplifiers does not remain constant when they are modulated.

One method of getting around the necessity of using linear amplifiers is to use a combination of phase and amplitude modulation, as in the Ampliphase system. You will learn more about phase modulation in a later lesson.

Lesson Questions

Be sure to number your Answer Sheet 21CC.

Place your Student Number on *every* Answer Sheet.

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

1. Why can't class C amplifiers be used to increase the power of the modulated signal from an AM transmitter?

2. What TWO operating adjustments on a properly tuned linear stage have most effect on the linearity?

3. What makes it possible to use a single tube in a class B rf linear amplifier, when two tubes in push-pull are needed in a class B audio amplifier?

4. How does the plate current meter reading in a class B linear amplifier in an AM broadcast station react when modulation is applied?

5. In which of the following operating conditions is less power dissipated in the class B linear tube: (a) *with modulation* (b) *without modulation?*

6. If the final class B linear amplifier stage in a transmitter has an output of 50 kw and an efficiency of 40%, how much power must be dissipated by the tubes?

7. List three causes of positive carrier shift in a linear amplifier stage.

8. Why are the grid and plate voltages in a linear amplifier set to one-half their normal values during adjustment?

9. What is the output voltage when two generators generating equal voltages 180° out of phase are connected in series?

10. When the output of an Ampliphase system is at its maximum, will the phase difference between the two amplifier signals be (a) *maximum* or (b) *minimum?*

THOROUGHNESS

Whatever you do, do well if you would stay on the straight road to success. The habits of carelessness and slipshod work are all too easy to acquire; beware of them as you would the plague. Men who are thorough in their work cannot remain undiscovered for long, because the demand for such men is greater than the supply.

Thoroughness is just as important in study as it is in work; what you get out of a lesson depends upon how completely you master the material presented in it. Some books, as fiction, are read hurriedly and only once, then cast away; the enduring works of literature are carefully read and reread many times but always essentially for the pleasure they give; textbooks, however, must be read quickly to get the basic ideas, then carefully many times until every important principle has been mastered.

Thoroughness in study habits leads to thoroughness in work habits, and eventually to a thorough success.

J. M. Smith

nri

RADIO • TELEVISION • COMMUNICATIONS

CURRENT, VOLTAGE AND RESISTANCE MEASUREMENTS

22CC

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

CURRENT, VOLTAGE, AND RESISTANCE MEASUREMENTS

22CC

STUDY SCHEDULE NO. 22

MP2M469 **1969 EDITION** Printed in U.S.A.

CURRENT, VOLTAGE, & RESISTANCE MEASUREMENTS

METERS play an important part in all phases of electronics. They are used to find out what is going on in an electronic circuit, for making operating adjustments, in checking performance, and in troubleshooting. In communications work you will deal with meters that are wired right into transmitter circuits to indicate operating conditions, and also with portable test instruments. Sometimes, you may want only a general indication of current, voltage, or resistance; sometimes you may have to take very accurate measurements, but whether the measurement is general or very exact, its usefulness depends on how well you understand your instrument. If you misuse a meter, you will get inaccurate readings, and may damage the meter.

In this lesson, we will discuss meters used to measure current, voltage, resistance, and power. You will learn that the same types of meter are used in all measuring circuits. It is the arrangement of the meter circuit and the way in which the meter is connected that determines the type of measurement that will be made.

We will discuss only meters that do not use vacuum tubes as part of their circuits here. In a later lesson we will take up vacuum-tube voltmeters and other instruments using vacuum tubes.

Most meters used in electronics rely on the principles of electromagnetism. That is, when current flows through a coil of wire, a magnetic field is produced around the coil that is proportional to the amount of current flowing. This principle is used in the three most common types of meters: the d'Arsonval, the magnetic vane, and the dynamometer. We will discuss these three types. No matter which type it is, the moving element is made as light as possible. Because the moving element is light it will move quickly and have a tendency to oscillate somewhat back and forth through the correct reading. There must be some means of bringing the pointer to rest quickly without oscillation after the meter has been energized, and to keep it from swinging back and forth after it is brought back to zero. This is called "damping." It can be accomplished either electrically or mechanically. You will see examples of both systems.

ACCURACY

Although it is possible to make highly accurate meters (within one-quarter of 1% or better) by hand-calibrating them with a standard meter, they are large and very expensive. For general communications work, meters having an accuracy of 2% are satisfactory. These meters are mass-produced, with printed scales, and are adjusted internally at the time they are manufactured to the required accuracy.

The accuracy of a meter movement is generally expressed as a percentage of the full-scale reading. For example, if a meter with 50 scale divisions has 2% accuracy, it is accurate within one scale division at the *full-scale* reading. This does not indicate the per cent of accuracy on the rest of the scale, because it may be off as much as one scale division at any other part of the scale also. For example, when

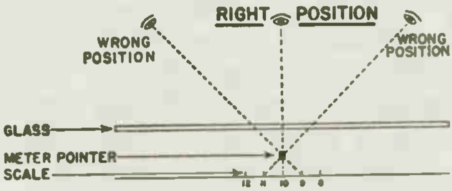
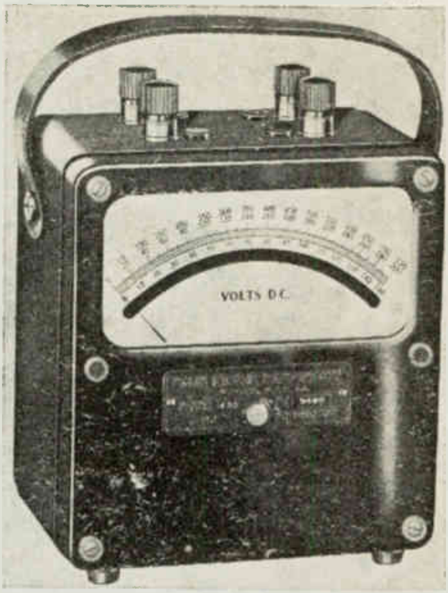


FIG. 1. When reading a meter, always look at it directly to get an accurate reading.



Courtesy Weston Electrical Instrument Co.

FIG. 2. A portable dc voltmeter. The dark arc below the scale is a mirror to make it easier to read the meter accurately.

the meter pointer is at the tenth division (1/5 of full scale), the reading may still be in error by one division; this would be an error of 10%. Most instruments used in communications work have a 2% accuracy at full scale, but not as much error as 10% at the low end of the scale.

Reading the Meter. In order to get an accurate reading from a meter, you must look at it squarely, not from an angle, because in order to swing freely, the pointer must be a little bit above the scale. As shown in Fig. 1, if you look directly down on the meter you will get one reading, for example, 10. However, if you look at it from one side, you might think the reading was 9, and from the other, you might think the reading was 11.

To help avoid this, many meters have a mirror under the pointer. The dark area under the scale in the meter shown in Fig. 2 is such a mirror. When you read the meter, you move your eye until the reflected image of the pointer disappears, and you know you are looking directly at it.

Shielding. It is sometimes desirable to shield the meter elements, because external magnetic fields produced by nearby current-carrying conductors or by the earth itself can react with them and affect the readings. There is no known insulator for magnetic lines of force, so the undesirable stray fields must be bypassed around the meter elements by a shield made of iron, which is a good conductor of magnetic lines of force. Such an arrangement is shown in Fig. 3.

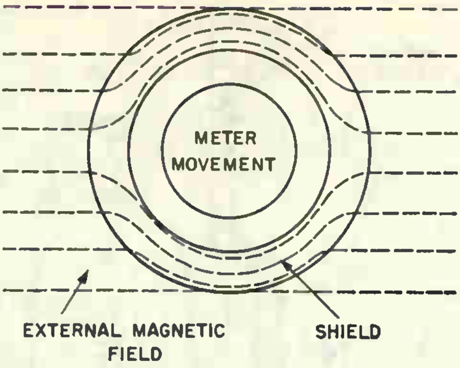


FIG. 3. A meter can be shielded from an external magnetic field by encasing it in iron, which bypasses the field around the meter movement.

Now, let's take a look at the basic types of meters, and then see how they are used to measure current, voltage, resistance, and power.

Basic Meter Movements

The three types of meters we will discuss in this section are the d'Arsonval, the magnetic vane, and the dynamometer. All three types can be used for both ac and dc measurements; however, the d'Arsonval is by far the most common for dc. When the d'Arsonval is used to measure ac, it must be used with a rectifier to change the ac to dc. Since the d'Arsonval is the most common type, let's discuss it first.

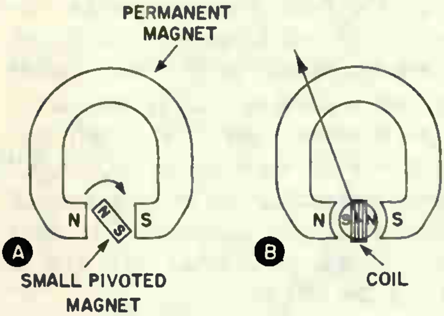


FIG. 4. A small pivoted magnet placed between the poles of a large magnet as at A, will rotate until the unlike poles are opposite each other. A pivoted coil through which current is passed will act in the same way, as shown at B.

THE D'ARSONVAL METER

In an earlier lesson, we had a quick look at the d'Arsonval meter. Let's review its action now. As you will remember, it works on the principle that like magnetic poles repel each other and unlike magnetic poles attract. If a small pivoted magnet is placed between the poles of a permanent magnet, as shown in Fig. 4A, it will move in the direction of the curved arrow.

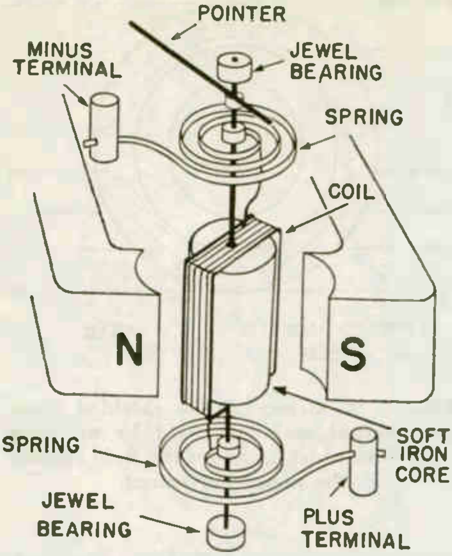


FIG. 5. How a d'Arsonval type of meter is made.

Now, suppose instead of the small magnet we put a coil between the poles of the larger magnet. As you know, when current flows through a coil it becomes magnetized. Therefore, if we send current through this coil, it will act just like the small permanent magnet in Fig. 4A. This is the way the d'Arsonval meter is made, as shown in Fig. 4B. The coil is wound on a soft iron core, and placed on a pivot, and a pointer is attached to the coil. When no current is applied, the coil and the pointer are in the position shown. When current is applied to the coil, it becomes magnetized and starts to rotate, moving the pointer.

The coil would continue to rotate until its south pole was opposite the north pole of the permanent magnet, except for the fact that springs are attached to the ends of the coil, as shown in Fig. 5, which oppose the coil movement. Therefore, the coil turns only to the point where the magnetic force caused by the current is exactly equal to the retarding force of the springs, and remains there as long as the current causing the magnetic force is applied. When current stops, the magnetic field of the coil disappears, and the springs move the coil back to its original position.

Since the magnetic force causing rotation of the coil is proportional to the amount of current flowing through the coil, one particular value of current will make the coil rotate to one particular place. A greater current will rotate the coil further, and a smaller one will rotate it less. A scale that is marked to show the amount of current that will cause any particular amount of movement is placed under the pointer.

The permanent magnet is made from a special steel or metal alloy, chosen for strong magnetic qualities and long magnetic life. The stronger the field of the permanent magnet, the more the coil will rotate for a particular current; in other words, the more sensitive the meter will be. The magnet is especially treated and aged until the field strength remains constant. The pole pieces are of soft iron, carefully shaped to give the desired magnetic distribution. If the meter scale is to be linear (that is, adjusted so that equal increases in current will produce equal increases in meter coil movement), the magnetic field must be uniform throughout the gap in which the coil turns.

So that it will turn easily, the coil is wound on a very light-weight metal form, and the coil and the form are suspended between almost frictionless pivots with jewel bearings. The number of turns used in the coil depends on the range and sensitivity desired for the meter.

The coil starts to rotate from the same position each time. When the coil rotates, one spring is wound while the other is unwound. The springs thus oppose the coil movement in either direction away from the starting position.

Naturally, these springs will not always remain perfectly balanced. Most meters have a zero adjustment to compensate for this. It is a small screw that usually protrudes through the case of the meter just above or below the meter coil. Turning this screw moves the upper spring enough to balance the springs and bring the meter pointer back to the zero position.

The springs are also used to make electrical connections to the coil. Of course, this means that they must be insulated from each other and from the meter frame.

Damping. In the d'Arsonval meter, the damping is done electrically by winding the coil on an aluminum frame. As the coil responds to the flow of current and starts to rotate, a voltage is induced in the aluminum frame as it cuts the lines of force of the permanent magnet. The induced voltage causes a current to flow in the frame, which in turn produces a magnetic field opposite to that of the permanent magnet. The opposing field produces a braking action which brings the pointer quickly to rest. When the coil comes to rest, no voltage is induced in the frame. Therefore, there is no field produced by the frame to interfere with the fields of the coil and of the permanent magnet. The same action takes place when the meter is de-energized and the pointer is returned to zero.

Another common method of creating damping is to place a resistor between the meter terminals. In this system, there is a voltage induced in the coil as it moves through the fixed field, which causes a current flow through the resistor and coil that sets up an opposing field similar to that produced by the current induced in the coil form. In both cases, damping action ceases as soon as the coil stops moving. The resistor value that will permit the most rapid coil movement without noticeable waving and still give full-scale meter reading is called the critical damping value. This value varies widely. Some meters require 10,000 ohms, others 100 ohms. Too small a resistor causes over-damping and a slow movement, whereas too large a resistor does not damp enough.

The induced current method of damping does not affect the meter range at all. The resistor method may or may not affect the current range of the meter, depending on the value of the resistance needed. You will learn more about this later on in this lesson.

THE MAGNETIC-VANE METER

Instead of having a fixed magnet and a moving coil like the d'Arsonval meter, the magnetic-vane meter has a fixed coil and a movable iron vane. It is often called an iron-vane meter. One of the best of the magnetic-vane meters is the book-type, shown in Fig. 6. In this meter, two iron vanes are used, surrounded by a coil of wire. When current flows through the coil, the vanes will be similarly magnetized so they will repel each other. It makes no difference whether the energizing current is dc or ac, the vanes will still repel each other. One vane is fixed, and the other is pivoted and attached to a pointer. The movement produced is shown on a scale under the pointer. Hair springs are used to control the motion and return the pointer to zero when no current is being applied to the coil.

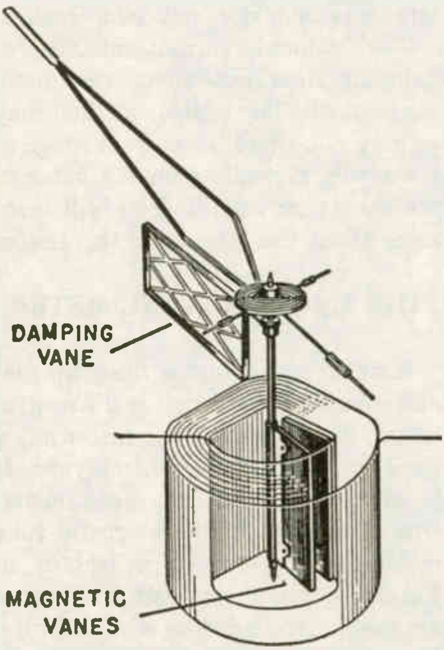
The moving vane meter is not as widely used as the d'Arsonval meter for a number of reasons. The meter cannot be made as sensitive as the d'Arsonval meter and therefore cannot be used to measure very weak currents. Also the scale is not linear; the lower quarter of the scale is usually quite compressed.

Furthermore, when used on dc, the polarity of the current through the coil may have some effect on the reading, so it is best to take a reading, reverse the polarity, then take another reading, and average the two.

In addition, the meter cannot be used on high ac frequencies because of losses in the vanes. Both eddy current and hysteresis losses become appreciable as the frequency increases. In fact, these meters are usually calibrated for use at some specific frequency and if measurements are taken at another frequency, the percentage of accuracy of the measurements will be somewhat less than the rated accuracy of the meter.

Damping. A mechanical method of damping is used in this meter. The aluminum vane shown directly under the pointer is used to slow down the movement of the pointer. This vane fits quite snugly into the space inside of the coil. Both ends of the opening are closed, so the vane, in moving through the air in the enclosed space, is held back by the air pressure developed. This effectively damps any tendency of the vane and pointer to oscillate.

Several other types of magnetic vane movements have been developed, but the book type is the most sensitive and the most accurate. All have similar characteristics that restrict their use to dc and low-frequency ac measurements.

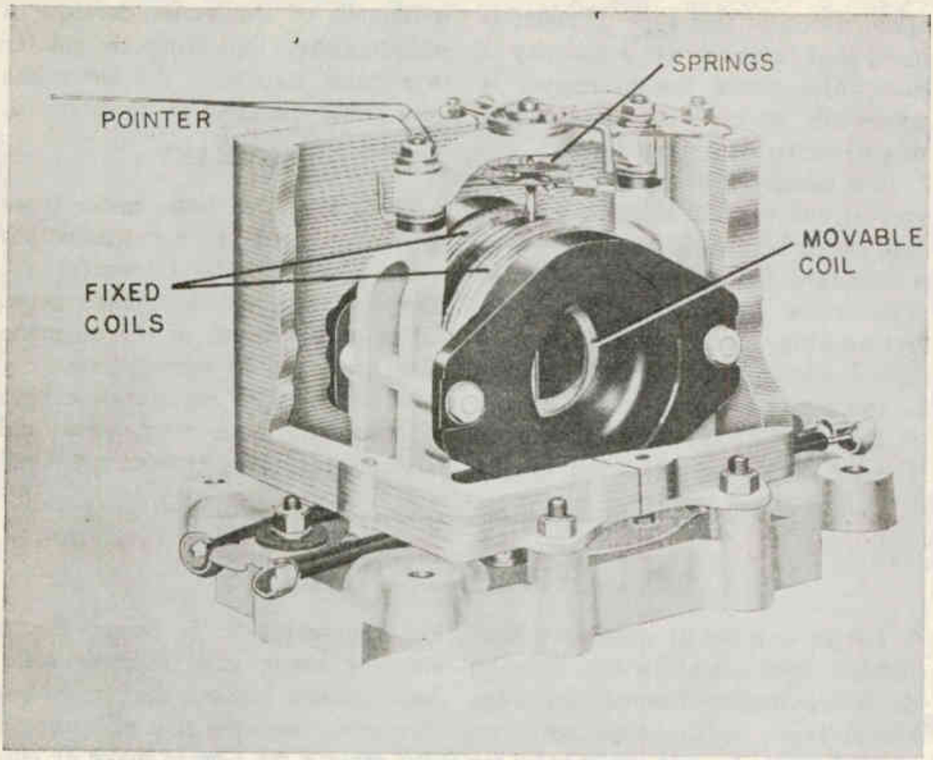


Courtesy Weston Electrical Instrument Co.

FIG. 6. A magnetic vane meter.

THE DYNAMOMETER

The dynamometer, or electrodynamometer as it is sometimes called, operates because of the reaction between the magnetic fields of a fixed coil and a movable coil when the same current flows through both of them.



Courtesy Weston Electrical Instrument Co.

FIG. 7. A dynamometer.

Actually, there are generally two fixed coils and one movable coil. Fig. 7 shows this type of meter.

The fixed and movable coils are in series, so the same current flows through them. When no current is flowing, the axis of the movable coil is at right angles to the axis of the fixed coils. This position is maintained by control springs.

When current flows through the coils, it sets up magnetic fields around them. Because of the physical position of the coils, the magnetic field of the movable coil is at right angles to the magnetic field of the stationary coils. As you know, like magnetic poles repel and unlike magnetic poles attract each other, so the two fields try to align themselves. This causes a turning force, or torque, which carries the pointer clockwise across the scale until the restraint of the springs equalizes the torque. The pointer then comes to rest. The deflection of the pointer is proportional to the square of the current. Therefore, the scale used with the dynamometer has nonlinear scale divisions.

It doesn't make any difference whether the current applied is ac or dc, because if it is ac it will change direction in the movable coil and the stationary coils at the same instant so that the two magnetic fields will still oppose each other.

Since there is no iron core to produce an economical flux, the power consumption of this type of meter is high; that is, the power sensitivity is poor. Also, since the movement is necessarily heavy, it is a slow-acting meter compared to other types.

It is usually calibrated with direct current and is often called a transfer instrument because it can be used as a standard for calibrating other ac instruments.

Damping. A mechanical method of damping is used in Fig. 7. It consists of two vanes attached to the bottom of the shaft of the moving element. The air turbulence produced by the movement of the vanes develops a retarding effect that brings the pointer to a quick stop after the meter has been either energized or de-energized.

SUMMARY

There are three basic meter types commonly used in communications work. These are the d'Arsonval, the magnetic vane, and the dynamometer. All operate because of the magnetic effect produced by current flow.

All three types require some kind of damping. Both mechanical and electrical damping systems are used

DC Measurements

The three types of meters we have just discussed can all be used to make dc measurements. However, the d'Arsonval meter is the most sensitive, and is itself a dc meter, so it is by far the most widely used for dc measurements. In fact, in communications work, you will probably use only d'Arsonval meters in making dc measurements.

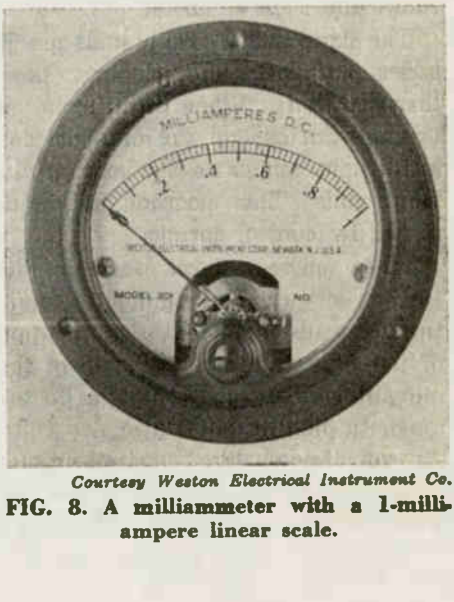
Practically all dc ammeters and voltmeters using d'Arsonval movements have linear scales. That is, the spacing between scale divisions is exactly the same over the whole scale. Let's see why.

The air gap in which the moving coil rotates is designed to give a uniform magnetic field in all of the space through which the coil moves. The torque or turning force exerted by the coil against the springs will be directly proportional to the current flowing through the coil. Since the springs allow the coil to turn by an amount proportional to the current, and the scale will be linear. Fig. 8 shows a meter with a linear scale. Each division between the longer lines represents one-tenth of a milliampere.

Now let's see how to make dc current measurements.

MEASURING DIRECT CURRENT

Current meters measure the flow of



Courtesy Weston Electrical Instrument Co.

FIG. 8. A milliammeter with a 1-milliampere linear scale.

electricity in a circuit. To make the measurement, the meter must be connected in series with the source and the load as shown in Fig. 9. Current meters that are used as operating indicators are wired into the circuit permanently. When a current meter is used as a temporary test instrument, the circuit must be broken so that the

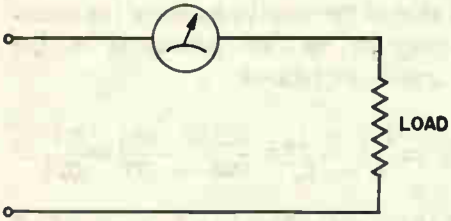


FIG. 9. A current meter is connected in series with the source and load as shown here.

meter can be connected in series with the load.

When making temporary measurements, you must be sure to use a meter with a high enough range. If you are in doubt, use a very high range, and switch to a meter with a lower range if you find that the current is low enough to permit you to do so.

The amount of current flowing in transmitter and receiver circuits varies considerably, from a few milliamperes in some low-level stages to several hundred amperes in the filament circuits of the final amplifier in a high-power transmitter.

Meters are made in a wide variety of full-scale ranges, but it would not be practical to keep on hand meters for every conceivable range. However, it is possible to extend the range of a milliammeter to measure higher currents. Let's see how this is done.

EXTENDING METER RANGES

When currents up to 5 milliamperes are to be measured, usually a meter having a basic range that covers the range to be measured is used. For example, if the currents to be measured are under 1 ma, a 1-ma meter is used. If currents up to 3 ma are to be measured, a 3-ma meter is used. When currents over 5 ma are to be measured, resistors called "current shunts" are connected in parallel with the meter movement to extend the range.

Suppose we have a 1-milliampere meter and want to measure currents up to 10 milliamperes. We can do so by putting a resistor across the meter terminals. We choose the value of the resistor so that nine-tenths of the current (9 milliamperes) coming into the resistor-meter combination will flow through the resistor and one-tenth (one milliampere) through the meter. In other words, we use the resistor to bypass nine-tenths of the current. Fig. 10 shows how this is done. Since

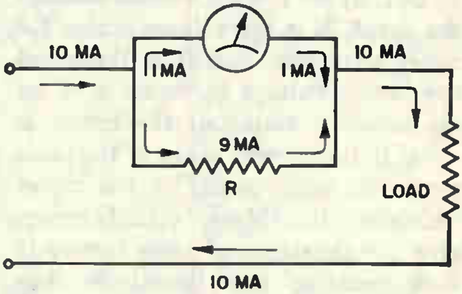


FIG. 10. How a shunt is used to increase the range of a milliammeter.

nine-tenths of the total current flows through the shunt, the shunt must have a resistance that is only one-ninth of the resistance of the meter.

Since the meter is a 1-milliampere

meter, the scale will be calibrated like the meter scale shown in Fig. 8. However, by connecting the shunt across it we have converted the meter into a 10-milliampere meter. Therefore, to determine the current flowing in the circuit you must multiply the meter reading by 10. Thus a reading of .6 milliampere on the meter indicates a current of 6 milliamperes in the circuit.

Since the current bypass resistor R makes a parallel path around the meter, it is called a shunt. The ohmic value of R is calculated so that it will pass a current that is the difference between the total current being measured and the amount of current the meter needs for full-scale deflection.

It's easy to find the shunt resistance to change the current that can be measured by a meter. For example, if we have a 1-milliampere meter, and we want to measure currents up to 25 milliamperes, we must use a shunt that will pass 24 milliamperes. To find its value, we use Ohm's Law, $R = E/I$, where I is the current through the shunt, E is the voltage across the meter terminals, and R is the shunt resistance. Voltage E, which is called the millivolt rating of the meter, is equal to the current range of the basic meter, I_m, multiplied by the meter resistance R_m. Meter manufacturers give the resistance of their meters in their catalogs and sometimes they also mark it on the back of the meter. In a few cases they also give the millivolt rating of the meter. Current I through the shunt equals the total current, I_t, minus the basic meter current, I_m. Therefore, by substituting $R_m \times I_m$ for E, and $I_t - I_m$ for I, our Ohm's Law equation can be written:

$$R = \frac{R_m \times I_m}{I_t - I_m}$$

If we figure all the current values in the same unit (amperes, milliamperes, or microamperes), the answer will come out in ohms.

Now, suppose the resistance of our 1-ma meter is 100 ohms; to find the value of the shunt necessary for measuring 25 ma, we substitute in the formula as follows:

$$R = \frac{R_m \times I_m}{I_t - I_m} = \frac{100 \times .001}{.025 - .001} = \frac{0.1}{.024} =$$

4.166 ohms, which can be rounded off to 4.2 ohms with an error of less than 1%.

We could also have calculated the resistance from the fact that the shunt must pass 24 times as much current as the meter, and therefore must have a resistance that is 1/24 of the resistance of the meter. Therefore:

$$R = \frac{100}{24} = 4.166 \text{ ohms.}$$

To find the actual total current flowing, multiply the reading on the shunted meter by the ratio of the current range of the meter with shunt to the current range without shunt. In our example, a one-milliampere meter was made into a 25-milliampere meter so this ratio is 25/1. The meter readings must be multiplied by 25 to find the actual current flow. It is important to remember at this time that the meter itself is not passing 25 milliamperes; only 1 milliampere goes through the meter and 24 milliamperes go through the shunt.

You will remember that when we

spoke about damping we said that d'Arsonval meters are sometimes damped by connecting a resistance across the meter terminals. Also we mentioned that too low a resistance would result in over-damping, which causes the meter pointer to move very slowly. When we connect a shunt across the meter terminals we connect it in parallel with the damping resistor. Since the shunt resistance is usually less than the critical damping value of the meter we end up with an over-damped meter. This situation can be corrected to some extent by connecting a small resistance in series with the meter movement.

Let's consider the example we already have discussed where we converted a 1-milliampere meter with a resistance of 100 ohms to a 25-milliampere meter. We did this by connecting a 4.166-ohm shunt across the meter terminals. This shunt becomes the damping resistor. If the value of damping resistor required for critical damping is several hundred ohms, you can see the meter will be very badly over-damped.

Suppose we connect a 20-ohm resistor in series with the meter lead. Now the total resistance of the meter is 120 ohms. To shunt this combination so the meter will read full scale in a circuit when the current flow is 25 milliamperes, we need a shunt that has a resistance 1/24 of 120 ohms.

$$R = \frac{120}{24} = 5 \text{ ohms.}$$

Now with a five-ohm shunt connected across the meter and the 20-ohm resistor we added, the meter range will be 25 milliamperes as before. However, now the damping resistor is made up of the 20-ohm resistance we connected in series with the meter, plus the 5-ohm shunt. Thus, the total damping resistance is 25 ohms, which is over five times the value it was with the 4.166-ohm shunt. The meter will still be over-damped, but not nearly as much as before.

We mentioned before that meters designed to measure currents above 5 milliamperes are usually 5-milliampere meters with a shunt. Thus, a meter with a scale from 0-100 milliamperes consists of a 5-milliampere meter as the basic meter movement, with a shunt built inside the meter case. When the meter indicates a current of 100 milliamperes, 5 milliamperes will be flowing through the meter and 95 through the shunt. Similarly in the case of a 5-ampere meter, 5 milliamperes will flow through the meter and 4.995 amperes through the shunt.

You might wonder why meters are made this way. There are two reasons, it is more practical to build one basic meter movement and extend its range by shunts than to build a large number of basic meter movements. Another reason is that if high currents were used in the basic meter, the springs which conduct the current to the coil would be quite bulky. Also we would have to use a rather large wire size to wind the coil. This would make the moving coil assembly bulky and insensitive.

In small panel instruments having a range of about 20 amperes or less, the shunt is contained within the instrument. In portable instruments of high accuracy and in panel instruments having a rating of over 20 amperes, an external shunt is generally used with the meter.

Most meters designed for use with external shunts have a sensitivity of 50 millivolts. You will remember that we said the meter sensitivity in volts is equal to $I_m \times R_m$, where I_m is the full-scale meter current and R_m is the meter resistance. Thus, the meter sensitivity simply tells us the voltage

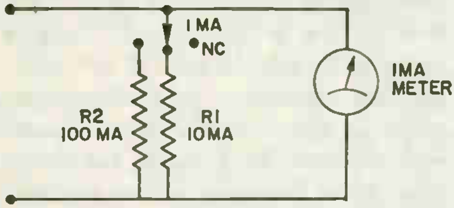


FIG. 11. A multi-range switch for connecting shunt resistances across a meter.

across the meter terminals at full-scale deflection. If we have a number of meters that are all 5-milliampere meters and all have a 50-millivolt sensitivity, they must all have the same internal resistance. Thus, a shunt designed to work with one of these meters could be used with any of them. You will find shunts made for use with meters of this type. They are usually labeled 50 millivolts and are also labeled with the current range to which they extend the meter. For example, a shunt marked 50 millivolts-20 amps, is designed for use with any 5-milliampere meters that have a sensitivity of 50 millivolts. When it is connected across the meter terminals, the meter range will be extended to 20 amperes.

Many meters have several ranges, each with a separate shunt resistance. The shunts are connected into the circuit by means of a multi-range switch, as shown in Fig. 11.

Ring Shunts. Another arrangement of shunt resistors, called the "ring shunt" is shown in Fig. 12A. In this circuit, we have a meter with a 40-ohm, 5-milliampere movement, and a ring shunt arranged to extend the scale to 25 ma, 50 ma, and 250 ma.

The range switch is shown in the position for the 25-ma range. To find what the total resistance would be, we use our formula:

$$R = \frac{R_m \times I_m}{I_t - I_m}$$

The meter resistance is 40 ohms, the meter current is 5 ma (.005 ampere), and the total current is 25 ma (.025

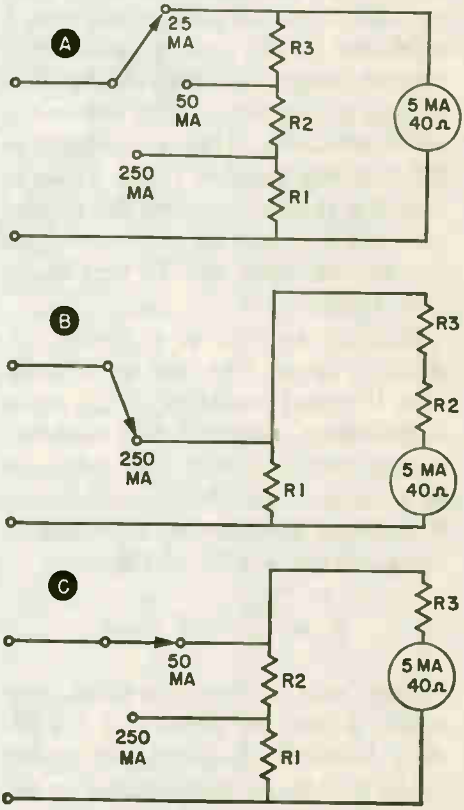


FIG. 12. How a ring shunt works.

ampere), so we have:

$$R = \frac{R_m \times I_m}{I_t - I_m} =$$

$$\frac{40 \times .005}{.025 - .005} = \frac{.2}{.02} = 10 \text{ ohms.}$$

We could also have determined the value of R using all currents in milliamperes. Thus it would work out:

$$R = \frac{40 \times 5}{25 - 5} = \frac{200}{20} = 10 \text{ ohms.}$$

Now we know that the total resistance of R1 + R2 + R3 is 10 ohms and the meter resistance is 40 ohms, so the total resistance in the circuit is 50 ohms, and we want to find the values of the individual resistances. Let's find the value of R1, first.

With the switch in the 250-ma position, the circuit could be redrawn as shown in Fig. 12B. Now the resistance of the meter is equal to 40 ohms *plus* the resistance of R2 and R3, or in other words, it is equal to 50 ohms *minus* the resistance of R1. The shunt is resistance R1 and its value can be calculated using the same formula as before. Now for R_m we substitute (50 — R1) which is equal to the total resistance in the meter circuit, 40 + R2 + R3. I_m, the current through the meter, is 5 milliamperes (.005 amps) as before, and I_t is 250 milliamperes (.250 amps). So in our formula, we have:

$$R1 = \frac{(50 - R1) \times .005}{.250 - .005} =$$

$$\frac{.25 - .005R1}{.245}$$

$$.245R1 = .25 - .005R1$$

$$.25R1 = .25$$

$$R1 = 1 \text{ ohm.}$$

The circuit with the switch in the 50-ma position can be redrawn as shown in Fig. 12C. Now the meter resistance R_m is equal to 40 ohms plus R3, which is equal to 50 ohms minus the resistance of R1 and R2, so we can find the combined resistance, R, of R1 and R2 as follows:

$$R = \frac{(50 - R) \times .005}{.05 - .005} =$$

$$\frac{.25 - .005R}{.045}$$

$$.045R = .25 - .005R$$

$$.05R = .25$$

$$R = 5 \text{ ohms.}$$

Now we know that R1 + R2 equals 5 ohms; since R1 equals 1 ohm, R2 must equal 4 ohms; and since the total of all three resistors equals 10 ohms, R3 must equal 5 ohms.

The ring shunt has two advantages over the circuit shown in Fig. 11. For one thing, the values of the resistors on the high ranges do not need to be as low. If we used the same basic 5-ma, 40-ohm meter in an arrangement like Fig. 11, we would have to have a resistor of only a fraction of an ohm on the 250-ma range. The other advantage is that the total resistance across the meter itself is the same on all ranges, and can therefore be used to provide damping.

MEASURING DC VOLTAGES

The meter in a voltmeter is actually a milliammeter or microammeter. The most commonly used meter in making dc voltage measurements is the d'Arsonval meter, which as you know, is a current-operated meter. Voltage is measured by sending current through a known resistance.

For example, if a 10,000-ohm resistor is connected across a source, and we connect a milliammeter in series with the resistance and it indicates that a current of 1 ma flows through it, you can calculate the voltage from Ohm's Law, $E = IR$. The voltage across the resistor must be $E = .001 \times 10{,}000 = 10$ volts. If we reduce the voltage, and the current drops to .5 milliamperes, (.0005 amps), we know the voltage must be

$$E = .0005 \times 10{,}000 = 5 \text{ volts}$$

Using this principle, a resistor, called a "multiplier resistor" is connected in series with a milliammeter and the combination is connected across the voltage source to be measured as shown in Fig. 13. The scale, of course, is calibrated to show the voltage rather than the current. We can do this because the current will depend directly on the voltage.

The value of resistance needed to be able to measure a certain voltage, depends upon the current range of the meter and the resistance of the meter itself, and upon the range of voltage to be measured. For example, suppose we have a 1-ma meter, and we want to use it to measure voltages from 0-1 volt. We find the value of resistance needed to limit the current flow in the meter circuit to 1 milliampere when the voltage across it is 1 volt by using Ohm's Law, $R = E/I$. The voltage range we want is 1 volt, and the current is 1 ma or .001 ampere, so we have $R = 1/.001 = 1000$ ohms. This is the total resistance in the circuit; it includes the resistance of the meter plus the resistance of the multiplier. In the example, if the meter has a resistance of 55 ohms, then the multiplier should have a resistance of 945 ohms.

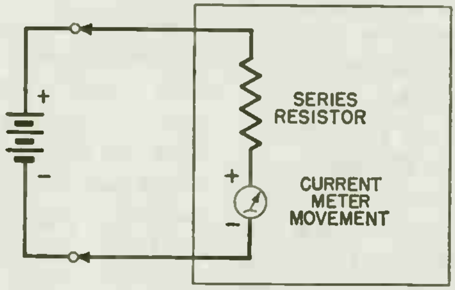


FIG. 13. How voltage can be measured by connecting a current meter and a series resistor across the source.

Let's take another example, suppose we want to use a 50-microampere meter that has a resistance of 2000 ohms to measure voltage up to 10 volts. Using Ohm's Law to find the resistance we get:

$$R = \frac{10}{.00005} = 200{,}000 \text{ ohms.}$$

So the total resistance needed is 200,000 ohms. We can ignore the meter resistance in this example because it is so small, and simply connect a 200,000-ohm resistor in series with the meter. When 10 volts is applied to the series combination, the meter will read full scale. When 5 volts is applied, it will read half scale.

Voltmeter Loading. As we have said, the meter is always placed across the line to measure voltage, rather than in series with the line as when current is being measured. As you can see, a certain amount of current must flow through the meter and its series resistor. We say that the meter is loading the circuit. Because of this, in a low-current circuit, we must use a meter with high sensitivity, or it will not indicate circuit conditions accurately. For example, a 1-ma meter will draw 1 ma of current. If the normal circuit current is 1 ampere, the additional 1 ma, which is .001 amp, drawn by the meter will be an insignificant amount. However, if the normal circuit current is only half a milliampere, then the additional 1 milliampere that the meter draws represents an increase in the total circuit current of 200%. This increase in total current will upset a high-impedance circuit and the voltage indicated on the meter will be substantially less than the voltage that is normally present in the circuit.

An indication of how much a meter will load the circuit is given by the sensitivity of the meter. A meter with a high sensitivity requires less current to operate it than one with low sensitivity and hence loads the circuit less. For example, a 50-microampere meter is more sensitive than a 1-milliampere meter. It requires only 50 microamperes to give a full-scale deflection. A 1-milliampere meter, on the other hand, requires a current of 1 milliampere, which is 20 times 50 microamperes to give a full-scale reading. Thus, a voltmeter built with a 50-microampere meter and suitable multiplier resistors will be more sensitive and load the circuit only 1/20 as much as a similar voltmeter made with a 1-milliampere meter.

Instead of giving the sensitivity of meters in terms of the current needed for a full-scale meter deflection, manufacturer's rate them in ohms per volt. Let's see what this rating means.

A 1-milliampere meter requires a current of 1 milliampere to give a full-scale deflection. To convert this meter to a voltmeter with a full-scale range of 1 volt, we connect a multiplier resistor in series with the meter. The value of this resistor is

$$R = \frac{1}{.001} = 1000 \text{ ohms.}$$

If we wanted to make a 2-volt meter, we would need a resistor

$$R = \frac{2}{.001} = 2000 \text{ ohms.}$$

Notice that this is twice 1000 ohms. If we wanted a 10-volt meter we would need 10,000 ohms, which is ten times 1000 ohms, and if we wanted a 100-volt meter, we would need a 100,000-ohm multiplier, which is 100 times 1000 ohms. Thus we say the sensitivity of the meter is 1000 ohms per volt. From this figure we can immediately tell what the total resistance of the voltmeter is.

For example, if we have a 50-volt meter with a sensitivity of 1000 ohms per volt, we know that the total resistance of the meter plus its multiplier is 50×1000 ohms, which is 50,000 ohms. Notice that this figure is based on the full-scale range of the meter and not the voltage being measured. The resistance of the meter is 50,000 ohms whether the voltage being measured is 50 volts, 40 volts, 25 volts, or any other value.

You can determine the ohms-per-volt sensitivity of any meter if you know the current required for a full-scale deflection, or, if you know the ohms-per-volt sensitivity of a meter you can determine what current it draws at full scale.

For example, if we have a 50-microampere meter used in a voltmeter, its sensitivity is

$$R = \frac{1}{.00005} = 20,000 \text{ ohms.}$$

Thus its sensitivity is 20,000 ohms per volt. If the meter has a full-scale voltage of 150 volts, its total resistance will be $20,000 \times 150 = 3,000,000$ ohms.

If you have a meter rated at 10,000 ohms per volt and want to know what current is required for a full-scale deflection, you use Ohm's Law:

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

$$I = \frac{1}{10.000} = .0001 \text{ amps}$$
$$= 100 \text{ microamps.}$$

We can convert a one-milliampere full-scale d'Arsonval meter to a two-scale voltmeter by connecting the proper value multiplier resistors in series with the instrument, as shown in Fig. 14. Remember that to get an indication of 1 volt on a 1-milliampere meter, the series resistance must be 1000 ohms. If we wish to increase the full-scale reading to say 150 volts, we merely multiply 1000 by 150, which gives 150,000 ohms as the value of the series resistor between terminals A and B.

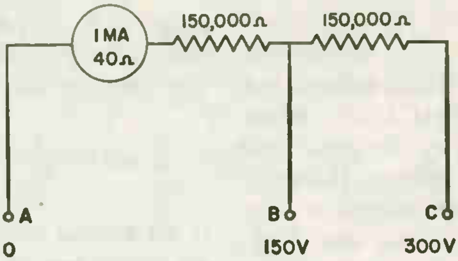


FIG. 14. How a 1-ma meter can be converted to a two-scale voltmeter.

To increase the full-scale reading to 300 volts, we add another 150,000-ohm resistor, giving a total of 300,000 ohms. In both cases, 1 milliampere of current flows through the meter, and the instrument has a sensitivity of 1000 ohms per volt on either scale.

In Fig. 14, the resistance of the 1-milliampere movement, 40 ohms, is negligible when compared with the value of the multiplier resistances. Thus, the meter resistance can be ignored. However, we would have to consider the meter resistance if we calculated the multiplier to make the meter read 1 volt full-scale deflection. The total meter circuit resistance would be 1000 ohms: 40 ohms in the meter, and 960 ohms in the multiplier.

Neglecting the meter resistance and using a 1000-ohm resistor would cause a 4% error in meter readings.

However, on the 150-volt range, neglecting the meter resistance and using a 150,000-ohm resistor would cause an error of only .02%. The percentage of error caused by neglecting the meter resistance will be approximately equal to the meter resistance divided by the multiplier resistance times 100. Although the percent of error decreases as the voltage range increases, the meter will always read low.

Meters with more sensitive movements will place less load on low-current circuits because of the much higher resistances used with them. For example, a .1-milliampere meter has a sensitivity of 10,000 ohms per volt. To extend its range to 150 volts, the resistor would be 10,000 × 150, or 1.5 megohms.

When measuring voltage, as when measuring current, if you are not sure what range to use, always use a high one, then switch to a lower one if the voltage to be measured is covered by the lower one.

Also, it would be very foolish to try to measure the output voltage of a power supply in a large transmitter or any other high-voltage source by holding the leads of a meter across it —you might even be electrocuted! If there is no permanent meter built in, and you must measure the voltage, first shut off the power. Then, discharge the filter capacitors, connect the meter across the output, turn on the power, and without touching the meter, read the voltage. Then, turn off the power, discharge the filter capacitors again, and disconnect the meter before turning the power back on.

PROTECTING THE METER

In communication circuits, special care must be taken to keep rf fields from affecting the meters. A strong field will induce rf currents in the meter wiring, which will affect the accuracy of the meter. High rf voltages may also cause the insulation to break down. This can be avoided in a dc meter in three ways: (1) connecting it so that it is at ground potential with respect to rf; (2) shunting it by an rf bypass capacitor; (3) putting an rf choke in series with it. Fig. 15A shows an example of the first method.

The meter is at the point of lowest rf potential in the cathode circuit. Although capacitor C1 does shunt rf currents to ground, it is not placed in the circuit specifically to protect the meter. Its primary function is to prevent degeneration in the circuit.

Fig. 15B shows the second method. This is used when a meter must be

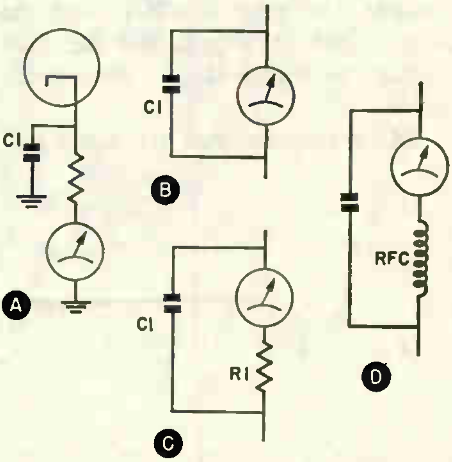


FIG. 15. How to protect a dc meter from strong rf fields.

placed in a lead that carries rf or af currents. A bypass capacitor having a low impedance compared to the meter impedance is used. The rf or af current follows the low-impedance path through the capacitor rather than the high-impedance path through the meter. This protection is sometimes increased by adding a resistor, as shown in Fig. 15C. The resistor, besides increasing the meter impedance, acts to damp out and prevent any resonance effects that might result from the parallel combination of the inductance of the meter coil and the capacitance of the bypass capacitor. Fig. 15D shows the third method. An rf choke is placed in series with the meter and within the circuit shunted by the bypass capacitor.

The bypass capacitor for rf circuits may be anywhere from .001 mfd to .01 mfd. For af circuits, the bypass capacitor should be .01 mfd to 1 mfd, depending upon the circuit. The size of the series resistor depends upon the current flowing in the circuit, but it is usually approximately equal to the meter resistance. The higher the resistor, the more it will protect the meter, but it should not be high enough to reduce the dc current too much.

DC voltmeters that are wired permanently into transmitter circuits are not put near rf fields. However, if you are making measurements with a portable instrument, you should be careful not to take your measurements where the meter can be affected by rf signals.

SUMMARY

Practically all meters used in communications work for measuring dc voltages and current have a basic d'Arsonval movement.

The range of a current meter can be extended by adding shunt resistors, the value depending upon the sensitivity of the meter and upon the current to be measured.

In a voltmeter, resistors called multipliers are added in series with the basic meter so the current through the resistor flows through the meter, and the combination is connected across the source to be measured. The meter scale is calibrated to show the voltage for the amount of current causing the pointer deflection.

For current measurements, the meter is always in series with the line; for voltage measurements, the meter is always across the line. When measuring either voltage or current always be sure to use a high enough range, or you may ruin the meter.

AC Measurements

The same basic meters can be used to measure alternating currents and voltages as are used to measure direct current and voltages. As you have learned, the magnetic vane meter and the dynamometer work on either ac or dc. Although they can be used with dc the d'Arsonval meter is so much superior for dc measurements that the magnetic vane type and the dynamometer are seldom used for dc measurements. The dynamometer is the most accurate of the meters for measuring alternating current and voltages, so it is often used as a standard for calibrating other instruments. The d'Arsonval meter can be used with copper oxide rectifiers to measure ac. This arrangement is primarily used in voltage measurements rather than current measurements. The d'Arsonval meter is used to measure alternating current by combining it with an arrangement called a thermocouple, which you will study in a minute.

MEASURING ALTERNATING CURRENT

In measuring alternating current, just as in measuring direct current, the meter is connected in series with the load. Again you must be careful to use a high enough range so you will not overload the meter.

AC meters are usually calibrated at some specific ac frequency. If the frequency at which you are taking measurements is too far removed from this, the readings will be somewhat inaccurate.

The range of an ac meter can be extended in the same way as that of a dc meter by adding a shunt resistance across the meter. The size resistance needed to extend a meter range to a given value is figured in the same way as the dc meters, using the same shunt formula.

The magnetic-vane meter and the dynamometer work directly on ac, but when the d'Arsonval meter is used, there must be some means of converting the ac to dc. One arrangement used particularly at rf frequencies is the thermocouple. Let's see how it works.

Thermocouples. The thermocouple works on the principle that when two dissimilar metals are joined and the junction is heated, there will be a dc voltage produced. The amplitude of this voltage depends upon the amount of temperature change at the junction. A sketch of a thermocouple junction is shown in Fig. 16A. Here two wires made of dissimilar metals are welded together to form a junction. The voltage produced at the junction can be measured between the other ends of the wires with a sensitive dc voltmeter.

Since current flowing through a resistance produces heat, we can get an indication of the amount of current flowing in a circuit by using a thermocouple junction along with a suitable meter as shown in Fig. 16B. The current to be measured is sent through a resistance wire or heater, producing heat. The junction of two dissimilar metal wires is brought near or actually welded to this heater. The other ends of the two wires are connected to a sensitive dc meter.

When current flows through the resistance, heat will be produced. The amount of heat produced will be proportional to the power dissipated in the resistance. Since the power dissipated in the resistance will be equal to I^2R, the heat will be proportional to the square of the current because the value of R remains constant. This heats the thermocouple junction, producing a dc voltage, which causes a dc current to flow through the thermocouple and through the meter.

Since the heat at the junction is proportional to the square of the current, and the generated dc is proportional to the heat, the meter will have what is called a square-law scale. Fig. 17 shows an example of a thermocouple meter with a square-law scale.

Square Law Meter Scales. Sometimes you'll have to use a standard dc milliammeter with a thermocouple or a square law meter. In this case the meter will be divided into equally

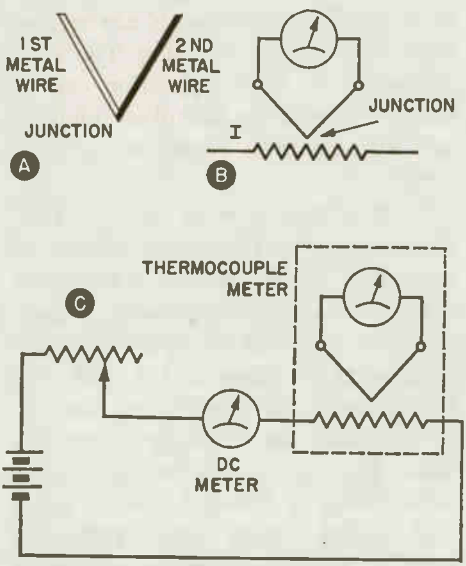
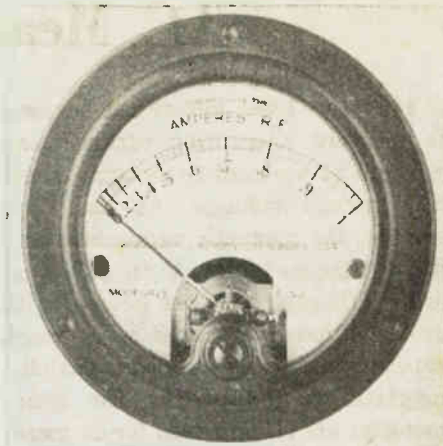


FIG. 16. How a thermocouple meter works.



Courtesy Weston Electrical Instrument Co.

FIG. 17. A thermocouple meter with a square-law scale.

squared divisions, and you can measure the current required for a full-scale deflection and then calculate the current for deflections less than full scale.

For example, assume that you have a square law meter with the scale divided into 100 equally spaced "divisions." If the full-scale reading (meter deflection of 100 divisions) is 10 ma, what is the current which corresponds to half-scale deflection (50 divisions)? To determine the unknown current, use the formula:

$$\frac{D_a}{D_b} = \frac{I_a^2}{I_b^2}$$

where I_a is the unknown current, D_a is the deflection corresponding to it, I_b is the known current and D_b is the deflection corresponding to the known current. Using values, we have:

$$\frac{50}{100} = \frac{I^2}{10^2} \text{ or } \frac{50}{100} = \frac{I_a^2}{100}$$

When transposed, we get:

$$I_a^2 = \frac{50}{100} \times 100 = 50$$

Then,

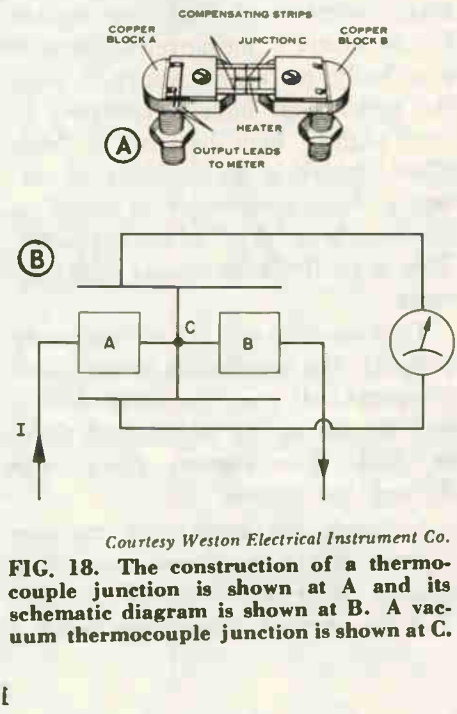
$$I = \sqrt{50}$$

which is approximately 7.1 ma.

You will remember that when we defined the ac ampere we said that an ac ampere is that ac current that will produce the same heating effect as one ampere of dc. Thus, since the relationship between ac and dc currents is based on the heating effect, we can calibrate a thermocouple-type meter with dc. A convenient circuit is shown in Fig. 16C. Here the dc current flow can be measured on a dc current meter and the reading on the meter connected across the thermocouple recorded. By adjusting the potentiometer the current can be varied so the entire thermocouple meter scale can be calibrated. Of course, thermocouple meters you buy come already calibrated, but this is how they are calibrated. Once they have been calibrated on dc they are quite accurate on ac even up into high radio frequencies. In fact, their accuracy is generally within 5% from dc up to 100 megacycles. In calibrating a thermocouple with dc, two sets of readings are usually taken, one with the current flowing in one direction and the other with it flowing in the opposite direction, and the readings are averaged.

The temperature difference between the hot junction and the free ends of the thermocouple element must not be influenced by surrounding temperature changes. To eliminate this possibility, the construction shown in Fig. 18A is used.

Fig. 18B shows the schematic diagram of this type of construction. There is a thin-walled tubular heater terminated in rather heavy copper blocks, A and B, which are so large they will not be heated by the heater. Current flowing through the heater will develop a temperature difference between the center of the heater and the blocks. The thermocouple junction C is on the center of the heater. The other ends of the wires are connected to two strips called "compensating strips." These strips are insulated from the blocks electrically by thin layers of mica, but connected to them thermally so the strips will be at the same temperature as the blocks. The heat capacity of the strips is such that the temperature difference between the ends of the thermocouple and the junction will always be the same as the temperature difference between the center of the heater and blocks A and B. Thus, if the temperature of the



Courtesy Weston Electrical Instrument Co.

FIG. 18. The construction of a thermocouple junction is shown at A and its schematic diagram is shown at B. A vacuum thermocouple junction is shown at C.

air surrounding the thermocouple changes, the temperature difference between the blocks and the center of the heater and between the junction and the ends of the thermocouple does not, so there will be no change in the potential developed by the thermocouple due to this change. Conductors are fastened to the ends of the compensating strips to conduct the potential developed by the thermocouple to the meter. If very small currents are to be measured, the thermocouple is enclosed in an evacuated glass envelope. This is designed to protect the elements from temperature changes which could be produced by warm air circulating around the thermocouple junction or the ends of the thermocouple.

Thermocouple meters are available in current ranges from less than 1 milliampere up to about 300 amperes. They are well suited for measuring radio frequency currents because of their accuracy at high frequencies. Up to about 2 megacycles, there is practically no error. Above 2 megs the meter does have a tendency to read slightly high because of "skin effect," which is the tendency of current at high frequencies to travel on the surface or "skin" of the conductor. This is particularly true at high currents.

To minimize skin effect and eddy currents, the conductors in the thermocouple unit are often made of thin-wall copper tubing plated with silver or gold. The current flows only through the plating.

Although the error may increase slightly as the frequency rises, it is usually below about 5% at 100 mc. The accuracy at frequencies below 2 mc may be as much as .5% at the temperature at which the meter was calibrated. Because of this accuracy at high frequencies, the thermocouple meter is often used as a standard for calibrating instruments at frequencies above those that can be measured on the dynamometer.

The meter used with a thermocouple must be very sensitive, because the output of the thermocouple unit may be only 15 millivolts.

Thermocouples in general must be handled with great care because they are delicate. You must avoid overloading a thermocouple because the heater is likely to burn out if subjected to more than a 40% overload.

MEASURING AC VOLTAGES

All three of the basic meter types can also be used to measure ac voltage. Just as for measuring dc voltage, a resistor is connected in series with the meter and the combination is connected across the source to be measured. The types of meters we have described in this lesson are seldom used to measure rf voltages. Such measurements will be discussed in a later lesson.

The magnetic vane meter is not suitable for most voltage measurements in communications circuits because it is difficult and costly to make a magnetic vane meter with a sensitivity better than about 5 ma. A voltmeter built around a milliammeter that required 5 ma for a full-scale deflection would have a sensitivity of only 200 ohms per volt. This would load many circuits to such an extent that the voltage reading on the meter would be much lower than the voltage normally present in the circuit.

A d'Arsonval meter can be used for ac voltage measurements by changing the current needed to operate the meter to dc. One common method of doing this is to use a copper oxide rectifier with the meter. We will study these now.

Using Rectifiers. The basic d'Arsonval meter can be used to measure ac voltages by connecting it to a single copper oxide rectifier as shown in Fig. 19A, or to a bridge circuit consisting of four copper-oxide rectifiers as shown in Fig. 19B. The rectifier consists of a number of copper discs. One side of each disc is covered with a film of copper oxide. Next to each disc is a lead washer, and the unit is held together under pressure in a clamp-like arrangement. Current will flow readily from the copper to the copper oxide, but will not flow readily in the opposite direction because the copper will readily give up its electrons but the copper oxide will not. The copper disc acts as the cathode of a rectifier, and the oxide acts as the anode. The lead disc is used as a means of contact with the copper oxide.

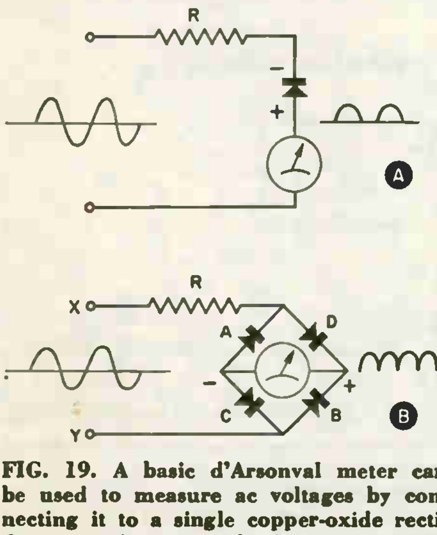


FIG. 19. A basic d'Arsonval meter can be used to measure ac voltages by connecting it to a single copper-oxide rectifier as at A, or to a bridge circuit at B.

The backward or reverse resistance of the unit to the flow of current in the opposite direction may be from 50 to 1500 times that of its forward resistance or rectifying direction.

Fig. 19A shows a single rectifier connected in series with the meter. This is a half-wave rectifier. Its output is a pulsating, direct current as shown by the waveform. The current flow is blocked during alternate half cycles.

In Fig. 19B, four copper-oxide rectifier units are connected in a full-wave bridge circuit. When point X is positive, current will flow from Y through rectifier C, through the meter from left to right, through rectifier D and resistor R to point X. When point Y is positive, current will flow from point X through resistor R and rectifier A, through the meter again from left to right, and through rectifier B to point Y.

The pointer of the meter cannot follow the pulsating direct current that appears at the input of the meter in either the full-wave or half-wave rectifier circuits. In the full-wave rectifier, the current that appears at the output of the rectifier is the average or .637 of the peak value. However, the scale is generally calibrated to read the rms or effective value of the voltage. In the half-wave rectifier the current at the output of the rectifier is only half of .637 or .318. Again the scale is calibrated to read the effective value of the voltage even though it is taken from the half cycle. The full-wave rectifier circuit is more commonly used than the half-wave circuit.

In the full-wave bridge circuit, both ac and dc are flowing simultaneously. The meter movement, however, is in that portion of the circuit where practically all the flow is in one direction. It is direct current resulting from having each alternate half of the sine wave flow through the meter element in the same direction.

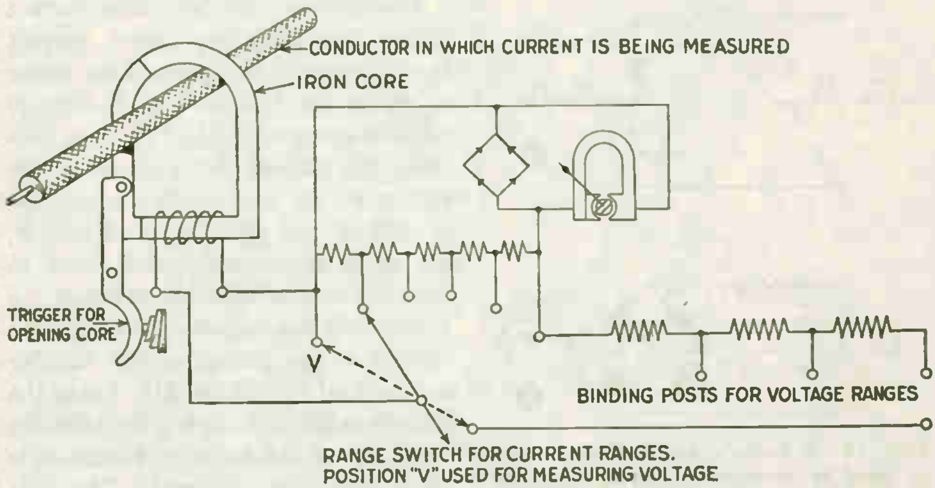
One of the advantages of the copper-oxide rectifier and d'Arsonval meter combinations for ac voltage measurements is that a very sensitive meter which takes little current from the source can be used. Another advantage is that the scale is linear. One disadvantage is known as frequency error. Because there is capacity between the oxide-coated sides and the non-oxide-coated sides of the discs in the rectifier unit, some current is bypassed, resulting in an error if the instrument is used at a frequency other than the one at which it was calibrated. Most instruments will read about 5% low for each 1000 cycles above the frequency for which the instrument is calibrated. The scale, for all practical purposes, is linear.

CLAMP-ON METER

We have mentioned several times that when current is being measured, the meter is connected in series with the line, and when voltage is being measured, the meter is connected across the line. However, there is a combination volt-ammeter that actually clamps around the line, so that the circuit need not be opened to take current measurements.

The diagram of a clamp-on meter is shown in Fig. 20. It consists of an iron-core that can be opened and clamped around the conductor carrying the current to be measured, with a coil wound around it connected to a d'Arsonval meter and bridge circuit.

The iron core and the coil form a transformer. The core is a one-turn



Courtesy Weston Electrical Instrument Co.

FIG. 20. A clamp-on meter.

primary, and the coil is the secondary.

The alternating current in the conductor produces a varying flux in the core. This in turn causes current to flow in the secondary, which is fed to the bridge circuit. The direct current output of the bridge actuates the meter movement. This type of meter is, of course, usable only with alternating current.

Both the current and the voltage range are usually wide, the current range may be as high as 1000 amperes, and the voltage range as high as 750 volts.

This instrument is a great aid in troubleshooting, especially to determine current flow taking place in a circuit in intermittent service, or to check running currents in motor circuits. It is also valuable in checking currents in three-phase circuits, which you will study later, to determine any unbalance that may exist between the phases, and in checking total input current against rated input current of a power supply. Care should be taken to see that the core laminations at the opening and hinged points are clean and sealed properly, otherwise possible obstruction at these points may cause erratic readings.

SUMMARY

The same basic meters are used for ac measurements as for dc measurements. When the d'Arsonval meter is used, the current fed to it must first be rectified. In current meters this is often done by means of a thermocouple. In voltage meters it is often done by means of copper-oxide rectifiers either singly or in a bridge circuit.

The meter is connected in series with the line to measure current and across the line to measure voltage.

When making measurements, it is always important to be sure to use a high enough range.

Resistance Measurements

Now let's see how the same basic current-operated meters can be used to measure resistance.

We know from Ohm's Law that $R = E/I$. Therefore, if we know the source voltage, and the current through a resistor, we can calculate the resistance. An ohmmeter does just this. It has its own source of dc voltage. When it is connected across a resistance, current flows through the resistance. The amount of current will be inversely proportional to the value of the resistance (the more resistance, the less current), and will determine the deflection of the pointer. The scale is calibrated to indicate directly the amount of resistance that will cause this amount of current flow. There are two general types of ohmmeter, the series type and the shunt type.

SERIES OHMMETER

The simplest type of direct-reading ohmmeter is shown in Fig. 21A. This ohmmeter is made up of a battery, a low-range milliammeter, and a combination of a fixed and a variable resistor, R1 and R2, in series.

The value of R1 plus R2 must be just enough so that with that particular battery voltage when the test leads of the instrument are touched together, making a complete circuit, the pointer of the current-operated meter movement will make a full-scale deflection. For example, if it is a 1-ma meter movement there must be 1 milliampere of current when the test leads are touched to each other. R2 is made adjustable because the battery voltage decreases with use and hence the resistance needed for a full-scale deflection will vary. Two resistors, one fixed and one variable are used to keep a minimum resistance in the circuit at all times. If you had only the variable resistance in the circuit and adjusted it so you had zero resistance in the circuit. such a high current would flow when you touched the test leads from the instrument together you could burn out the meter.

When an unknown resistance is to be measured, it is connected between the test probes as shown in Fig. 21B. Its resistance is added in series with R1 and R2, and the current will decrease accordingly, and there will be less deflection of the pointer. Since less deflection of the pointer means more resistance, the scale is printed with the zero at the right, which is the opposite of the current scale. A typical series-type ohmmeter scale is shown in Fig. 22.

To use this type of ohmmeter, the test leads are first shorted, and R2 is adjusted for a full-scale reading (zero on the Ohms scale). When the leads are connected across the unknown resistance, the current through the meter will be reduced because of the additional resistance. If the resistance being measured is equal to the resistance already in the meter circuit (the sum of the meter resistance and the resistance of R1 and R2), it will double the total resistance in the meter circuit, so the current will be cut in half, and the meter will read half scale. If the resistance being measured is less than the resistance in the meter circuit, the meter will read more than half scale. If the unknown resistance is greater than the resistance of the meter circuit, the meter will read less than half scale.

For example, the scale in Fig. 22 is for an ohmmeter with a 1-ma meter movement and a 3-volt battery. This means that the sum of R1 + R2 and the meter resistance will be 3000 ohms, because we need 3000 ohms in the circuit to limit the current to 1 ma with a 3-volt battery. We know this from Ohm's Law:

$$R = \frac{E}{I} = \frac{3}{.001} = 3000 \text{ ohms.}$$

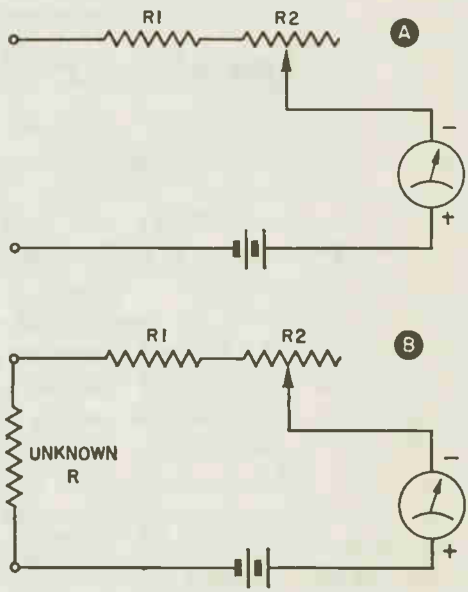


FIG. 21. A simple series-type ohmmeter.

This 3000 ohms might be made up of 100 ohms in the meter itself, a 1000-ohm potentiometer, R2, and a 1900-ohm resistor, R1. When the test leads are shorted so there is no additional resistance in the circuit, the meter will read full-scale (zero on the ohms scale). If a 3000-ohm resistor is connected between the leads, the total resistance will have doubled, the current will be cut in half, and the pointer will be at the center of the scale. As you can see, this is 3000 in Fig. 22. If the resistance being measured is 6000 ohms, the total resistance will have tripled, the current will be cut to one-third its full-scale value, and the pointer will be one-third of the way over. Thus, because of the relationship between the resistance and the current, zero resistance is at the right-hand end of the scale; 3000 ohms is represented by half the scale; the next 3000 ohms is represented by only one-sixth of the scale; the next 3000 ohms by only one-twelfth of the scale, etc.

FIG. 22. A scale for a series-type ohmmeter having a 3000-ohm center-scale value.

Higher resistance values can be measured with a series-type ohmmeter by using a higher battery voltage or a more sensitive meter. For example, if we used a 30-volt battery instead of a 3-volt battery with the 1-ma meter, then the resistance needed to limit the current to 1 ma when the test leads are shorted together would be

$$R = \frac{30}{.001} = 30{,}000 \text{ ohms.}$$

To reduce the current to .5 ma, we would need twice this resistance or 60,000 ohms. Thus, center scale on the meter would be 60,000 ohms minus 30,000 ohms (which is in the circuit at full scale) which is 30,000 ohms. The center of the scale on the ohmmeter would therefore be 30,000 ohms, ten times what it was with the 3-volt battery.

With a more sensitive meter we get the same results. Let's go back to the 3-volt battery and consider a 50-microamp meter. The resistance needed to limit the current will be:

$$R = \frac{3}{.00005} = 60{,}000 \text{ ohms.}$$

To reduce the current to half scale we would need an additional 60,000 ohms. Thus this meter would have a center scale resistance of 60,000 ohms.

Lower resistance values can be measured with a series-type meter by use of a shunt. Taking our original example of a 3-volt battery and a 1-ma meter, we can connect an additional resistor R3 in the circuit as shown in Fig. 23. Since R3 has a resistance of 333 ohms, which is 1/9 the resistance of $R1 + R2 + R_m$, nine times the current will flow through R3 that flows through the meter. Thus, with the test probes shorted together, the meter will read full scale, because 1 ma will flow through it and at the same time 9 ma will flow through R3. If we connect a resistance between the terminals that reduces the total current flow from 10 ma to 5 ma, we will get a half-scale reading on the meter, .5 ma of the current will

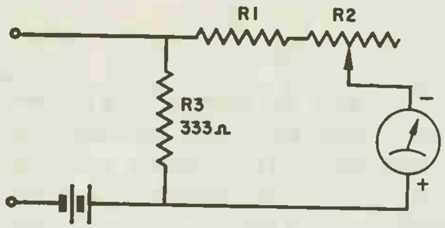


FIG. 23. Using a shunt with a series-type ohmmeter to measure lower resistances.

flow through the meter to give the half-scale reading, and the balance of the current, 4.5 ma, will flow through R3.

The resistance of R1 + R2 + R_m is 3000 ohms. This 3000 ohms is in parallel with 333 ohms. The resistance of the parallel combination is 300 ohms. With 300 ohms across the 3-volt battery, we get a total current of 10 ma which, as we said, gave us a full-scale meter reading. To cut the current in half, we need to double the resistance or add another 300 ohms between the test probes. Thus, with the shunt R3 added, center scale on the ohmmeter becomes 300 ohms. By adding a resistor that would permit still more current flow around the meter we could reduce the center scale resistance still further.

Some ohmmeters are arranged so that different values of shunt resistors can be switched into the circuit. The ohmmeter thus becomes a multi-range ohmmeter.

SHUNT OHMMETER

A different type of ohmmeter circuit is used when very low resistances are to be measured. This is the "shunt-type" shown in Fig. 24A. In this circuit, the milliammeter, the calibrating resistor R1, and the battery are connected in series and form a closed circuit even when the test leads are apart. The resistance of R1 must be just enough so that there will be enough current to cause a full-scale deflection of the meter pointer with the test leads apart (for a 1-ma meter, a 1-ma current). In actual practice R1 is made up of two resistors, a fixed resistor and a variable one to avoid the possibility of burning out the meter by setting the potentiometer so there is no resistance in series with the meter.

When the unknown resistance is connected between the test leads, it will be in parallel with the meter, as shown in Fig. 24B. This means that part of the current will flow through it, and part through the meter. If the

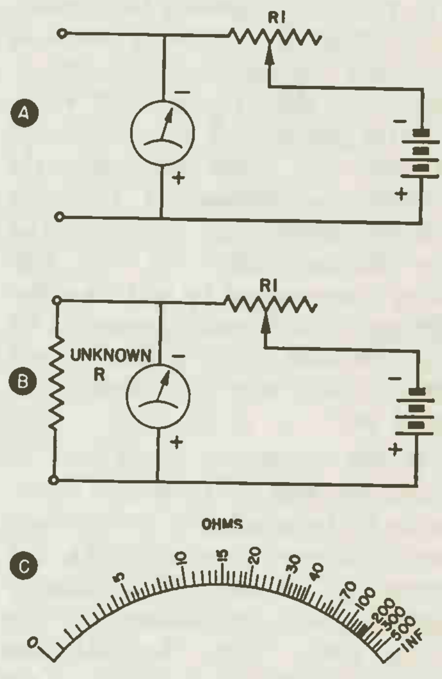


FIG. 24. A shunt-type ohmmeter circuit is shown at A and B, and the scale is shown at C.

resistance is high, most of the current will still flow through the meter, and the deflection will still be near full scale. If the resistance is low, most of the current will flow through it, not much will flow through the meter, and the deflection will be slight. If the resistance is exactly equal to the resistance of the meter itself, half the current will flow through it and half through the meter, and we will have a center-scale reading. As you can see, for low resistance values, the pointer would be at the left, and for high resistance values, it would be at the right. This means that zero would be at the left just as on a standard current scale.

A typical scale for a shunt-type ohmmeter is shown in Fig. 24C. The center-scale reading is 15, which means the resistance of the meter itself is 15 ohms. As you can see the scale is expanded at the low end just as the one for the series meter is.

The value of the resistance being measured would not appreciably affect the total resistance in the ohmmeter circuit, because the value of R1 would be so much larger than the value of the meter resistance, that for all practical purposes, it would determine the circuit current. For example, with a 1-ma meter and a 3-volt battery, the resistance of R1 plus the resistance of the meter would be 3000 ohms for full-scale deflection. Since the meter resistance is 15 ohms, R1 would be 2985 ohms. The resistance of the meter and the resistance being measured would vary from 0 to 15 ohms, depending upon the value of the resistance being measured. (Remember that the resistance of two resistors in parallel is always less than that of the smaller one). As you can see, this would not have any noticeable effect on the current drawn from the battery.

A word of caution about the use of this meter: Always turn it off when you are not using it, or you will drain the battery. You should get in the habit of switching it off after every measurement.

ACCURACY

The accuracy of any ohmmeter is limited by the stability of the battery terminal voltage. If the battery voltage is high, the meter will read low. If the battery voltage is low, the meter will read high. Practically all ohmmeters are designed to use batteries which are multiples of 1.5 volts; 1.5, 3.0, 4.5 volts, etc. Check the ohmmeter battery voltage occasionally to make sure your ohmmeter measurements will be reasonably accurate.

You can check the accuracy of the ohmmeter by measuring the value of a known resistance. A resistor with a tolerance of 1% is satisfactory. Choose a resistance value that will cause the ohmmeter pointer to indicate somewhere near the center of the ohmmeter scale.

The accuracy of most ohmmeters is only about 10% to 20%. That is, when the meter reads 100 ohms, the value of the resistance being measured may be anywhere between 80 and 120 ohms. However, this is usually as accurate as it needs to be for all practical purposes. If you must obtain a more accurate measurement, there are other instruments that can be used. You will study some of these instruments later.

THE MEGGER

Another resistance-measuring instrument you should become familiar with is the megohmmeter, or "megger," as it is called. As its name implies, it is used to measure very high resistance such as leakage in the insulation of cables, motor windings, transformers, and so forth.

Instead of a battery, it has a generator operated by a hand crank that generates about 500 volts. The meter is similar to a standard d'Arsonval meter, except it has two coils between the poles of the permanent magnet, wound on the same core, as shown in Fig. 25. Coil L1, which is called the *current coil*, is positioned so that when current flows through it, it will tend to move the pointer to the right (towards zero). Coil L2, which is called the *potential coil*, is positioned so that when current flows through it, it will move the pointer to the left (toward infinity).

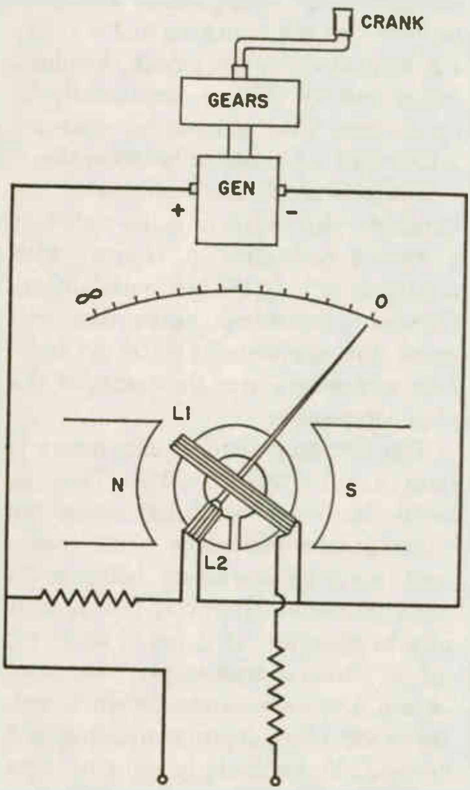


FIG. 25. How the "megger" is made.

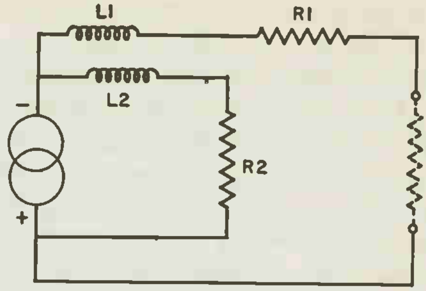


FIG. 26. Simplified schematic of the megger.

Fig. 26 shows a simplified schematic of the megger. The pointer on the megger is not restrained by springs so it is free to move to any position when the instrument is not in use. When you begin turning the hand crank the generator generates a voltage and as you can see, when nothing is connected across the terminals, the circuit will be open, and no current will flow through L1. However, current will flow through L2 and this current causes the pointer to swing to infinity. When a resistance is connected across the terminals, current will flow through L1, and this will tend to move the meter pointer towards zero. The current flowing through L1 will also tend to load the generator, which reduces the voltage and current through L2. This reduces the torque produced by L2 which has a further tendency to let the pointer move towards zero. How far it goes in this direction depends on how low the resistance is across the terminal. The lower the resistance, the more the pointer will move.

If there is any electrical path between the terminals, the meter will indicate it—up to hundreds of megohms.

For example, suppose you have a power transformer that has been subjected to undue moisture. Before placing it in operation you want to know if the moisture has caused leakage between the primary and the secondary. Connect the megger as shown in Fig. 27, and turn the crank. If there is any resistance between the primary and secondary it will be indicated on the meter scale. The design of the instrument is such that the speed of the crank has little effect on the reading as long as it is turned at a reasonable speed.

MULTIMETERS

Since the same basic meters can be used to measure current, voltage, and resistance, combination instruments, called multimeters can be designed that will measure all three. The same meter is used; it is connected in different ways by means of switches to measure current, voltage, and resistance. Multimeters are used for general maintenance and repair of electronic equipment. Fig. 28 shows the

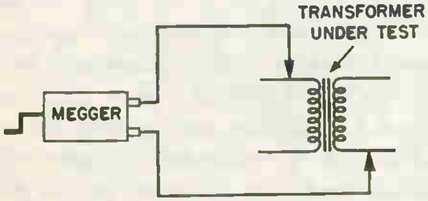


FIG. 27. How to test for leakage between transformer windings with a megger.

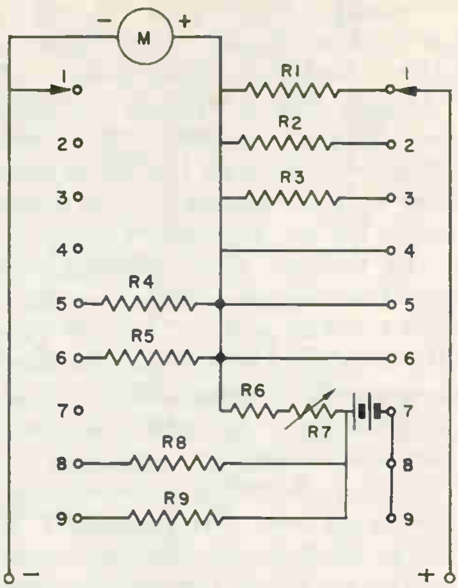


FIG. 28. Schematic diagram of a multimeter.

schematic diagram of such an instrument.

In the position shown, which we will call position 1, R1 is a voltage multiplier. It is in series with the meter. In position 2, R2 will be in series with the meter, and in position 3, R3 will be in series with the meter. If the meter is a 1-ma meter, by selecting R1 so its resistance plus the resistance of the meter is 1000 ohms, we could measure voltages up to 1 volt. By making the total resistance of the meter plus R2 equal to 10,000 ohms we could measure voltages up to ten volts in position 2, and by making R3 equal to 100,000 ohms we can measure voltages up to 100 volts in position 3.

In position 4 the meter is used as a 1-ma current meter. It is connected directly to the output terminals. In position 5, R4 is a shunt and will be connected directly across the meter terminals. In position 6, R5 will be a shunt across the meter terminals. If the resistance of the meter is 45 ohms, by making R4 equal to 5 ohms and R5 equal to .454 ohms we can measure currents up to 1 ma in position 4, up to 10 ma in position 5, and up to 100 ma in position 6.

In the last three positions, the meter is used as an ohmmeter. With a 3-volt battery and R6 + R7 plus the meter equal to 3000 ohms, we have a meter with a center scale resistance of 3000 ohms. In position 8, R8 is connected across the meter and R6 and R7. In position 9, R9 is connected across the meter and R6 and R7. If R8 is a 333-ohm resistor, the center scale resistance on this range becomes 300 ohms. By making R9 equal 30.3 ohms, the center scale resistance becomes 30 ohms.

Thus, with this arrangement we have three voltage, three current, and three resistance ranges. Additional ranges could be added by using a switch with more positions. Some of these positions could be for ac voltage measurements if we added a copper-oxide rectifier.

SUMMARY

The two main types of ohmmeters are the series type and the shunt type. Both use d'Arsonval meters. The series type has the zero at the right end of the scale; the shunt type has the zero at the left end of the scale. The shunt type is used to measure low resistance. Ohmmeters contain their own source of voltage, so the circuit in which measurements are being taken should be turned off.

An instrument called a "megger" is used to measure very high resistances. Instead of a battery, it has a 500-volt generator, and will measure resistances up to hundreds of megohms.

Current, voltage, and resistance meters are often combined in one instrument called a "multimeter."

Power Measurements

Power is the amount of electrical energy consumed in a circuit. It is measured in watts, kilowatts (thousands of watts), microwatts (millionths of a watt), or milliwatts (thousandths of a watt). In a circuit having only pure resistance, the power in watts is equal to the current in amperes multiplied by the voltage in volts, or $P = E \times I$. This is true in a dc circuit or in an ac circuit in which there is only resistance. In a circuit having inductance or capacity the phase angle between the voltage and the current must also be taken into account when figuring the power consumption. The phase angle is taken into account by multiplying the voltage and current by a figure known as the power factor, so the formula for power in an ac circuit is $P = E \times I \times PF$. The value of the power factor will be somewhere between zero and one. It is zero if the voltage and current are 90° out of phase, and one when the voltage and current are in phase. The power factor in an ac circuit is equal to the ratio of the resistance in the circuit to the impedance in the circuit. It is also equal to the cosine of the phase angle between the voltage and current.

Actually, you use this same formula for dc power, but since the power factor is always 1 in a dc circuit, it is ignored. Let's see how power is measured in dc circuits, in 60-cycle power-line circuits, in af circuits, and in rf circuits.

DC POWER MEASUREMENTS

It is easy to find the power in a dc circuit by connecting an ammeter in series with the circuit, and a voltmeter across it, then multiplying the indicated current by the indicated voltage. For example, if the voltage is 100 volts and the current is 10 amperes, the power consumed would be 1000 watts, or 1 kilowatt. (If this continued for 1 hour, we would say 1 kilowatt-hour of power had been consumed).

In making power measurements in this way, the position of the two meters in the circuit can make a difference. If they are connected as in Fig. 29A, the current drawn by the voltmeter flows through the ammeter. In a circuit where the current is low, this could add appreciably to the current indication, particularly if the voltmeter sensitivity is low. If the meters are connected as shown in Fig. 29B, the voltage indicated on the voltmeter will be slightly higher than the voltage across the load. If the current is very high, the drop across the ammeter may be enough to upset the calculation. At high currents the circuit at A should be used, at low currents, the one at B.

DC power can also be measured with a wattmeter, but it is usually simpler to measure it with a voltmeter and an ammeter.

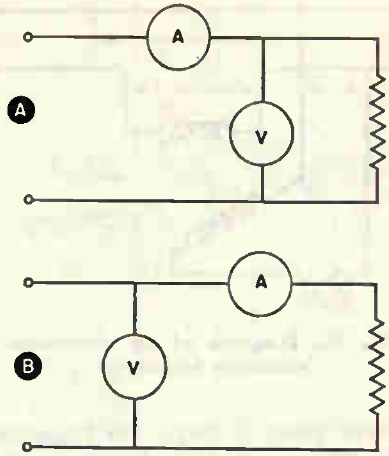


FIG. 29. Power can be measured by connecting an ammeter in series with the circuit, and a voltmeter across the circuit, then multiplying the readings. The arrangement at A is best for high currents, and the arrangement at B is best for low currents.

60-CYCLE POWER MEASUREMENTS

A wattmeter is generally used for measuring power in 60-cycle power line circuits. We cannot get the true power consumed by connecting a voltmeter and an ammeter into the circuit then multiplying the effective current by the effective voltage because we have not taken phase difference into consideration. The figure we would get by multiplying the effective current by the effective voltage is known as the apparent power and is expressed in volt-amperes.

A wattmeter takes the power factor into account. A dynamometer is the type of meter generally used as a wattmeter. Fig. 30 shows a schematic diagram of a wattmeter. The line

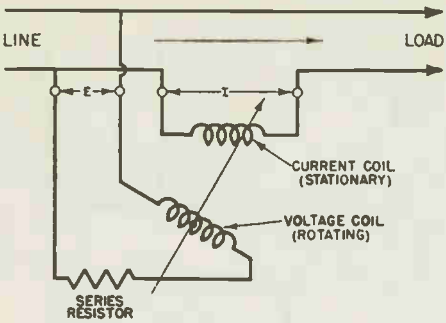


FIG. 30. Diagram of an electrodynamometer wattmeter.

current flows through the stationary coil. The movable coil is connected in series with a suitable resistance and connected across the line. The magnetic field of the stationary coil is proportional to the instantaneous current, and the magnetic field of the movable coil is proportional to the instantaneous voltage. The deflection of the pointer is proportional to the product of the magnetic fields. Since it is the instantaneous values of current and voltage that determine the deflection, it is an indication of true power.

Wattmeters are delicate instruments, and should be handled with care. They have maximum voltage and current ratings in addition to a maximum power rating. For example, a 500-watt meter should not be used to measure power of more than 500 watts. Also, if the meter is rated at a maximum voltage of 600 volts and a maximum current of 1.25 amps, you could overload the current coils by using it on a circuit when the voltage is 50 volts and the current 3 amps, even though the power in this circuit is only 150 watts. Similarly, you could overload the voltage coil by using the meter in a 1000-volt circuit when the current is .1 amp, even though the power is only 100 watts.

AF POWER MEASUREMENTS

Meters that are designed to measure the power output of audio-frequency devices such as audio amplifiers, radio receivers, etc. are designed to measure the voltage across a resistor. Since the voltage across a resistor and the current through it are in phase, the power factor will be 1, so the power will be directly proportional to the square of the voltage.

There are three general types of af power meters: the power output meter, the power level meter, and the VU meter.

Power Output Meter. The power output meter is a rectifier voltmeter with a resistor connected across it. The output from the device to be tested is applied across this resistor instead of to the usual load. The power dissipated in the resistor will be equal to $E^2 \div R$. Thus, if we know the resistance, we can determine the power by measuring the voltage across the resistance, squaring it, and dividing it by the resistance. Since the resistor is always the same, the meter can be calibrated directly in watts.

Power Level Meter. The power level meter is also a rectifier voltmeter, but it does not have a resistor built in across the meter. It must be connected across the load of the device under test. A high resistance voltmeter is used to place as little additional load on the circuit as possible.

This type of meter must always be connected across the same value of resistance for which it was designed, usually in communications work, 500 ohms. The power level meter is designed to indicate the ratio of output power to a certain reference level, and is therefore calibrated in decibels (db) rather than in watts.

A common reference level in communications work is 6 milliwatts of power into a 500-ohm load. This is called 0 db. Using the power equation $P = E^2 \div R$, we find that with 500 ohms and 6 milliwatts of power, the voltage is 1.73 volts. Therefore, by measuring the output voltage the meter can tell whether the power has gone above or below this reference level. If it is above, the meter indicates + db; if it is below, the meter indicates — db.

The advantage of this system is that it provides us with a convenient method of comparing different devices on a more or less common base.

For example, a manufacturer of a receiver may state in his specifications that the output of the receiver will be 0 db with a 100%-modulated input signal of 1 microvolt. This immediately provides us with a means of checking the receiver performance or comparing its rated sensitivity with that of another receiver.

It is important that you use this type of meter across the correct load impedance. Remember the meter is basically a voltmeter. It will indicate correctly in watts only when used across the load for which it was designed. For example, 1.73 volts across a 500-ohm load is 6 milliwatts, which we have set as zero db. Suppose we used this meter across a 10-ohm load. Now if the voltage is 1.73 volts, the meter, since it is a voltmeter, would indicate 0 db or 6 milliwatts. However, the actual power is $1.73^2 \div 10 = .3$ watt! Thus, to get an accurate reading with this instrument we must use it across the rated load of 500 ohms.

VU Meter. The VU meter is also used to indicate power ratios rather than watts of power. The unit used is called a volume unit or VU instead of a decibel, and the reference level is 1 milliwatt into 600 ohms. This is called 0 VU. With 1 milliwatt of power and 600 ohms of resistance, the voltage is .775 volts. If the voltage goes above this, the meter will indicate +VU, and if it goes below this, the meter will indicate —VU. Like the power level meter, it must be used across the correct load. These instruments are built using a special copper-oxide rectifier and an extremely sensitive meter. They are usually built right into the transmitter.

Fig. 31 shows a VU meter scale. It is only accurate in transmitters designed so that 100% modulation is obtained when the meter indicates 0 VU. The figures below the VU scale indicate the percent of modulation.

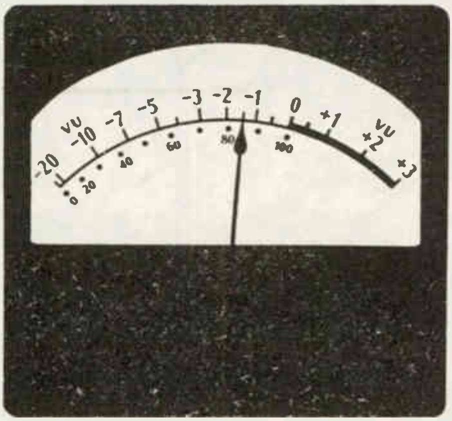


FIG. 31. A typical VU meter.

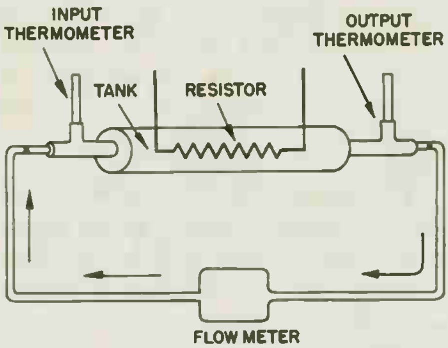


FIG. 32. How rf power can be measured without producing radiation.

RF POWER MEASUREMENTS

The radio frequency power output of a transmitter can be measured without actually producing radiation as shown in Fig. 32. Here a dummy load resistor is connected to the output circuit of the transmitter. This resistor is in a tank through which water is circulated to absorb the heat developed in the resistor. The water also goes through a flow meter that shows how much water flows through the unit in a given length of time. There are also thermometers that indicate the temperature of the water at the input and at the output of the tank.

Charts are provided with this type of measuring equipment so that you can calculate the power from the two temperatures and the water flow. This is called the calorimeter method of measuring power.

LOOKING AHEAD

In this lesson you have studied the three basic types of meters and learned how they can be used to measure current, voltage, resistance, and power. In later lessons you will learn about other types of instruments that use vacuum tubes in their circuits.

It is very important for you to understand how to use test equipment, because it is of little value unless you do.

Lesson Questions

Be sure to number your Answer Sheet 22CC.

Place your Student Number on *every* Answer Sheet.

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

1. What type of meter is generally used for dc measurements?
2. Give two reasons why the moving vane meter is not as widely used as the d'Arsonval meter.
3. Determine the resistance of the shunt needed to convert a 1-ma meter with a resistance of 100 ohms to a 5-ma meter.
4. If two ammeters connected in series with the same load, each indicate a current of 8 amps, what is the current flowing in the circuit?
5. What size multiplier resistor would you need to convert a 1-ma meter to a voltmeter with a full-scale voltage of 50 volts?
6. Find the sensitivity in ohms per volt of a voltmeter which uses a 10-microampere meter.
7. Why is a voltmeter with a high ohms-per-volt sensitivity more useful than one with a low sensitivity?
8. If the reading on a 50-volt voltmeter connected across a circuit is half-scale, indicating a voltage of 25 volts, and the meter sensitivity is 1000 ohms per volt, what is the total resistance of the meter and its multiplier?
9. Draw a schematic diagram of a series-connected ohmmeter.
10. If in a dc circuit a voltmeter connected across the load indicates 45 volts and an ammeter in series with the load indicates 7 amps, what is the power in the load?

THE VALUE OF REVIEW

Man has acquired so much new knowledge in recent years that it has become impossible for one person to know everything available about even a limited subject. Educational authorities realize this fact, and the colleges of today consider a man well-educated if he knows the elementary ideas *and knows where to find other information when he wants it.*

The field of radio and television has outgrown the memorizing ability of the human mind. Also, it is such a comprehensive field that occasionally you cannot recall important facts previously studied. Review is obviously needed.

Time spent in review several weeks or months after a book is studied will be far more profitable than an equivalent amount of extra time spent on the book initially, for your mind has then had a chance to file and store away the information secured from the first study. Each review results in more information being transferred from the textbook to your mind, and soon, with no conscious attempt to memorize, you will find yourself able to recall an amazing number of valuable facts.

J. M. Smith

nri

RADIO • TELEVISION • COMMUNICATIONS

nri

ATTENUATORS, FILTERS, AND EQUALIZERS

29 CC

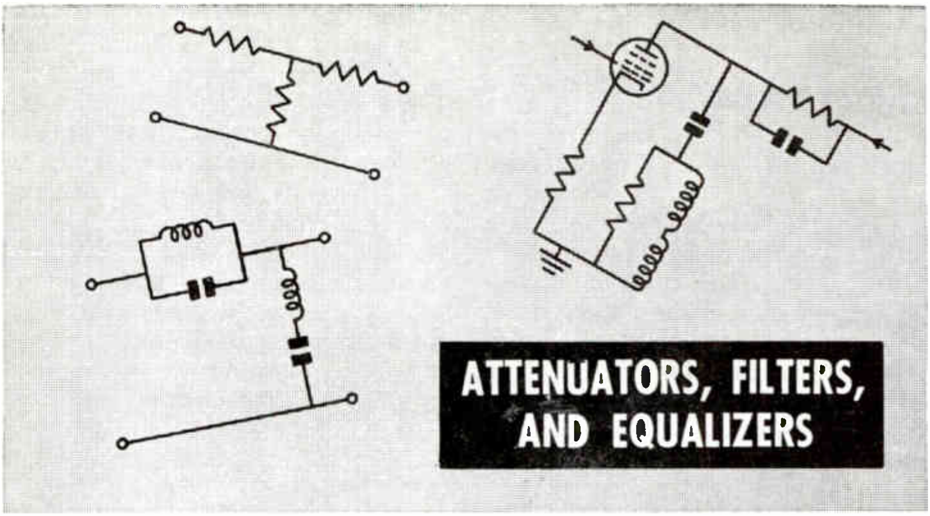
NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

ATTENUATORS, FILTERS, AND EQUALIZERS

29CC

STUDY SCHEDULE NO. 29

MPIM968 1969 EDITION Printed in U.S.A.



ATTENUATORS, FILTERS, AND EQUALIZERS

ATTENUATORS, filters, and equalizers are all basically classified as networks. A network is an arrangement of electronic components, usually resistances or reactances, designed to provide some special effect. Actually, every electronic circuit is a network of some sort. However, when an engineer or technician speaks of a network, he is usually referring to some special group of components within a circuit. Such a network might be anything from a simple arrangement of resistors used for impedance matching to a complex group of reactances providing frequency selection.

Networks of this sort will be found in every field of electronics. They provide an efficient method of controlling signal strength, noise reduction, and phase shift. They can select one particular control signal from a number of signals being transmitted simultaneously over the same control line. In fact, some type of special network will be found in all but the very simplest of circuits.

Networks can be classified into three general types: attenuators, equalizers, and filters.

Attenuators. An attenuator is a network consisting of pure resistance; its operation is independent of the frequency or the phase of the signal. Its effect on all signals is the same at all frequencies.

Filters. A filter is frequency-selective. A filter always contains reactances of some sort, and will respond to different frequencies in different ways. In addition, a filter will usually shift the phase of the signal in some way. By using the proper values of circuit components and arranging them in special ways, we can make networks that will have almost any desired effect on the signal as it passes through the circuit from the supply to the load. Filters are used in electronic equipment to pass certain frequencies and reject others. They are broadband devices in that they tend to respond to a certain band of frequencies rather than to a specific frequency.

Equalizers. Equalizers are used to correct for any undesired frequency discrimination or phase shift in other equipment or in connecting cables.

If these networks were perfect, each would do only the job it is intended to do. An attenuator would simply change the amplitude of the signal. A filter would pass a certain band of frequencies and reject all outside of this band. An equalizer would correct any undesired phase shift or frequency discrimination. Actually, leakage and stray capacitance cause each to do a little of all three. Although, in this lesson, we may speak of one of these networks as if it did its job and its job alone, you should remember that a certain proportion of undesired effects is unavoidable. The secret of successful design lies in holding these effects to a minimum. At video and rf frequencies, good shielding and the proper layout of the parts will help to prevent trouble caused by leakage. At audio frequencies, the coupling between the cores of the coils, and the coil resistance may become problems.

In earlier lessons, we learned about some of the simpler types of networks. These consisted of voltage dividers, power-supply filters, tuning circuits, and vacuum-tube coupling circuits. In this lesson, we will look at some of the more complex networks. We will learn to recognize some of the basic network configurations and their effect on a signal. We will discuss the advantages and disadvantages of the various filters and problems in designing practical filter circuits.

First, let's see what problems are involved in impedance-matching and power transfer.

Impedance-Matching and Power Transfer

In setting up and adjusting electronic equipment, impedance-matching between stages and efficient transfer of power from one stage to another or to an antenna are important problems. We can tell how effectively these problems have been solved by measuring either the signal power or the signal voltage at the input or output of the two stages and comparing the two values to see if there has been a gain or a loss.

UNITS OF MEASURE

In an earlier lesson, you learned about the unit of measurement called the decibel and how it is used to express the ratio between input and output power. You also learned that voltage ratios can be expressed in decibels, but only when the impedances across which the two voltage measurements are taken are equal.

There is another unit of measurement which you will run into from time to time called the "neper." This unit originated in Europe and differs from the decibel in that it is based upon a different system of logarithms. In computing decibels, we use the system of common logarithms, with the base 10. In computing nepers, the system used is known as the natural logarithm system in which the base is an odd number, 2.71828. One neper equals 8.686 db, and 1 db equals .1151 neper. This unit is not widely used in the United States, but you will find it in some textbooks and reference material.

POWER MEASUREMENTS

Decibel Meters. Decibel meters have been used to read the power in the line directly in db. These are ac voltmeters calibrated in db units, with the zero-power-level indication at about center scale. They are de-

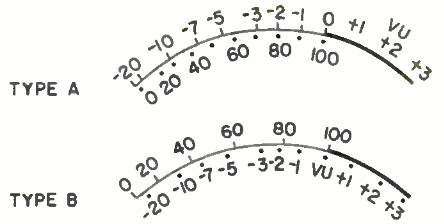


FIG. 1. Two types of VU meter scales.

signed to be connected across a specific value of impedance. Not all db meters have the same scale markings, damping, or zero levels. In the past, different services have used different values as the standard line-impedance across which the db meter is to be connected and have used different power levels as the zero reference level. Some have used .006-watt as the reference point, some .0125-watt, and others .001-watt. Some have used 500 ohms and some 600 ohms as standard line impedances.

When using a meter of this type, if the power in the line is more than the reference value, it produces a higher voltage across the line, and the meter reads to the right of the zero mark, showing a plus db reading. If the power is less than the reference value, the voltage is less, and the needle moves to less than center scale, indicating a minus db reading.

VU Meters. Since 1940, a standard type of power-level meter has been used in broadcast work and in many other audio applications. This is an ac volt-meter of the rectifier type called the volume unit or VU meter. Unlike db meters, all VU meters are built with identical characteristics. They are calibrated to read relative power levels when connected across a 600-ohm impedance line. The zero-reference level used is .001 watt (1 milliwatt). The VU meter is essentially a peak-reading meter, since with a relatively rapid pulse or steep wave, these meters rise to about 99% of the peak value very quickly, but fall off slowly.

VU meter scales are calibrated in plus or minus VU, and zero to 100% modulation. Two different types of VU meter scales are shown in Fig. 1.

The VU meter is designed to operate with a 3600-ohm external resistor as shown in Fig. 2. When the meter is connected across the line, it "sees" the 600-ohm source and load impedances in parallel. Thus, the impedance into which the meter is "looking" is 300 ohms. Since the meter has an impedance of 3900 ohms, an external 3600-ohm resistor is connected in series with the 300-ohm resistance of the parallel-connected source and load to match this impedance to that of the meter.

The resistor is not made an integral part of the meter because it may be

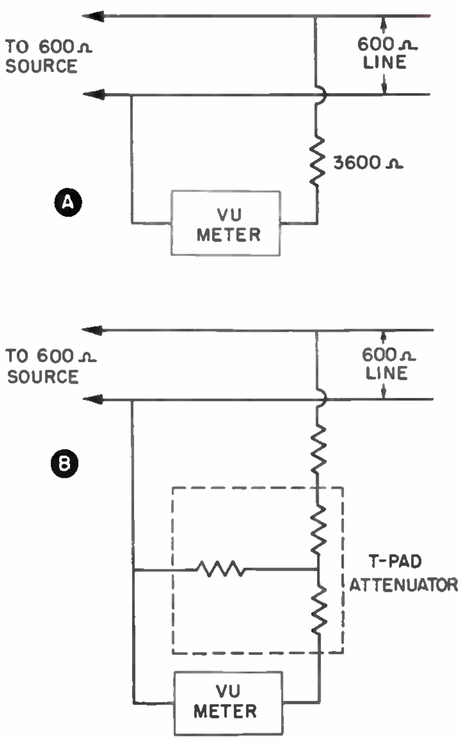


FIG. 2. A VU meter connected across a 600-ohm line is shown at A; a T-pad attenuator is added at B.

necessary to insert a T-pad attenuator in the meter circuit as shown in Fig. 2B. In broadcast work, a reading of zero VU or 100 on the percentage scale indicates 100% modulation. If the power required for 100% modulation is greater than the one milliwatt zero reference level, the meter would, of course, read farther upscale, and an attenuator called a T-pad is needed to make the meter indicate the correct percentage of modulation. Also, in some applications, the power in the line may be greater than the +3 VU at the high end of the meter scale. In these cases, a T-pad is needed to keep the meter from reading off-scale. Of course, in a case such as this, the loss introduced by the pad must be mentally added to the meter reading.

LOAD MATCHING

One of the most important considerations in the study of electronics is the transfer or transmission of power. In fact, the transmission and/or control of energy in the form of electrical power is the only purpose of any electronic system.

Let's look at some of the factors which affect this. In Fig. 3, we have shown a power source that develops an emf of 40 volts. Any source of power (generator, battery, or oscillator) has a certain amount of internal impedance, which consists of the distributed resistance, capacitance, and inductance in the source. We have assumed that our power source (let's call it a generator) has an internal impedance of 50Ω and have shown this as Z_I.

It is easy to see that the terminal voltage, E_T, is going to be equal to the emf minus the voltage across the internal impedance Z_I. The value of the voltage drop across Z_I will depend on the current flowing in the circuit as well as the value of Z_I, so the terminal voltage will vary with the current as long as the emf and Z_I remain constant. If the circuit is broken at the switch, no current will flow, there will be no voltage drop across Z_I, and E_T will be 40 volts. Now let's connect our generator to a load and see what happens.

If the selector switch is at A as shown, we will have a circuit consisting of Z_I which is equal to 50 ohms in series with Z_A, which is equal to 30 ohms. This gives us a total impedance of 80 ohms being supplied by our 40-volt emf. Through Ohm's Law ($I = E \div R$) we can determine that a current of .5 amp will flow in the circuit.

This means that we will have a voltage drop of 25 volts across Z_I, and a drop of 15 volts across Z_A. Therefore, although our power supply is actually generating 40 volts, 25 volts are dropped internally, and we have a terminal voltage, E_T, of only 15 volts.

If we look at the circuit from a power standpoint, we find that a total of

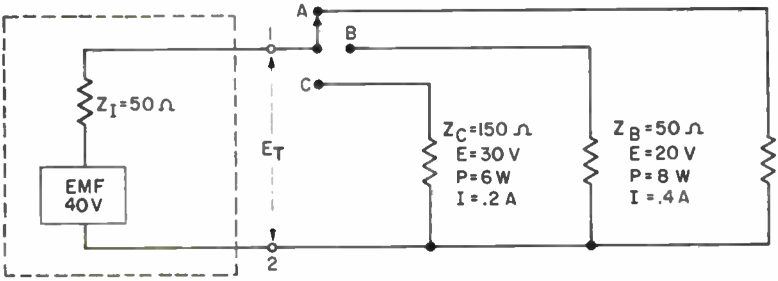


FIG. 3. Effect of load on power generation and transfer.

20 watts (40-volt emf x .5 amp) is developed by the power supply. However, 12.5 watts of this power (25 volts E_{Zl} x .5 amp) is wasted internally and only 7.5 watts is actually used by the load. This is obviously a very inefficient situation. If we have a load which consumes only 7.5 watts, it is a waste of equipment and energy to develop 20 watts to supply it.

Let's move the selector switch to Position B and see what effect a 50-ohm resistance has in our circuit. Now, our 40-volt emf will be feeding a total resistance of 100 ohms, and our circuit current will be .4 amp. We will have 20 volts dropped internally, leaving our terminal voltage at 20 volts. The total power developed will be 16 watts, and half of it, or 8 watts, will be consumed by the load. The other half will be wasted internally.

This is an improvement over the first circuit. First of all, we are developing a total power of only 16 watts, and we are getting 8 watts of this to our load. Thus, we have much better efficiency since we are wasting only 8 watts internally.

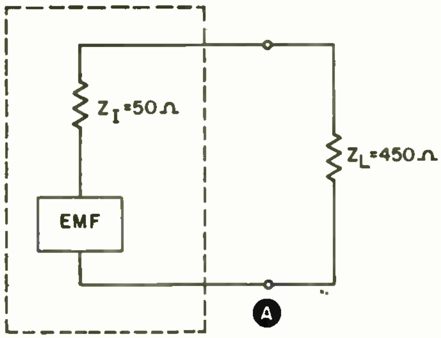
We actually have more power available in our external circuit with less total power generated. Notice that in this circuit the impedance of the load equals the impedance of the source; we say that they are "matched." Now, let's see if we can improve this still further by moving the selector switch to Position C.

In this position, the total impedance is 200 ohms, and the current will be .2 amp. The terminal voltage will be 30 volts, and the internal drop will be only 10 volts. The total power developed by the generator is only 8 watts. Of this, 6 watts is delivered to the load, while only 2 watts is wasted internally. Also, the voltage across the load has been increased. We have a much more efficient situation than we had before because we are wasting only 2 watts out of 8. However, we are delivering less power to the load than we were with circuit B when the load and source impedances were matched.

Thus, it is easy to see that a situation where the load impedance is much less than the source is always very inefficient and doesn't give much useful power. We can also see that we have the most power delivered to the load when the impedances are matched, but we do waste half of the power in the process. Further, the greatest efficiency exists when the load impedance is much larger than the source impedance although the actual power delivered to the load is less than when we have matched impedances. It is obvious that we will usually avoid having the load impedance less than the source impedance.

However, when it comes to a choice of having either maximum power with matched impedances or high efficiency with proper mismatching, we have to consider the specific application. For example, consider the circuit shown in Fig 4A, where the impedances are mismatched. If a large amount of power is required to operate the 450-ohm load, efficiency will be the most important. Let's assume that we have a large power requirement and the emf of the power source equals 10 kv. At 10 kv, 20 amperes of current will flow to the load, and a total power of 200 kw will be developed. Of this 200 kw, 20 kw will be wasted internally and 180 kw will be consumed in the load.

In the circuit of Fig. 4B we have shown the same load, but have matched the impedances. Notice that we have not shown the impedance matching system, but have just represented it by changing the value of the load resistance. In this circuit we would get more power to the load but the energy wasted would be tremendous. Our 10 kv emf would give 100 amps of current and the total power



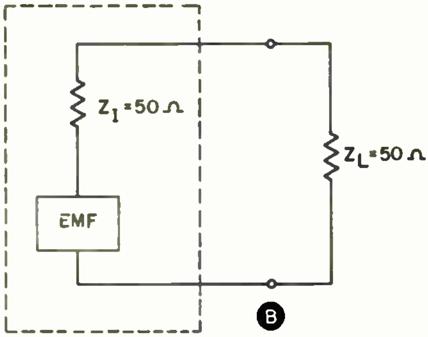


FIG. 4. Greatest efficiency with mismatched impedances, A; greatest power transfer with matched impedances, B.

developed would be 1000 kw. Of this, 500 kw would be wasted internally and 500 kw would be consumed by the load. Although we would have 320 kw more power available to do work than we had available with the mismatched impedances in Fig. 4A, we would waste 480 kw more in order to get it. The cost of producing this extra wasted energy would make it impractical to match impedances. In addition, the cost of equipment capable of handling 100 amperes of current would be much greater than that needed to handle 20 amperes. Thus, impedance matching becomes too expensive to be practical when large amounts of power must be handled.

In communications and control circuits, however, extremely small amounts of power are used, and power waste is not so expensive. Here, the important consideration is getting enough of the available power to the load for the equipment to operate properly. Often the available power is so low that unless the maximum amount is transferred to the load, the equipment will operate poorly or not at all.

For example, a speaker or speaker system must be matched to the output of the amplifier or it may not receive sufficient energy to operate properly. Similarly, a microphone must be matched to the input circuit of an amplifier or the losses will be severe.

Now that we have seen the problems involved in transferring power, let's look at the networks used to do so.

Resistive Pads and Attenuators

Resistive pads and attenuators are used to reduce or control signal levels, to match impedances, and to reduce the inductive effect of transformers coupled together through a line. Usually, they are used to reduce signal levels, to match impedances, or both.

SIMPLE ATTENUATORS

To see how an attenuator can introduce a loss without causing an impedance mismatch, it is best to examine some simple circuits. In Fig. 5A, we have a power source having a constant emf of 40 volts and an internal impedance of 50 ohms supplying power to a resistive load of 50 ohms. The load and source are matched and the power delivered to the load is 8 watts.

Suppose we want to reduce the power delivered to the load from 8 watts to 4 watts. We can do this by inserting a 43-ohm resistor in series with the load as shown in Fig. 5B. This makes a total impedance in the circuit of 143 ohms (50+50+43). The current is then .28 ampere (40÷143) and the voltage across the load is 14 volts (.28 x 50). This means the power delivered to the load is 4 watts (.28 x .28 x 50).

We can also reduce the power to 4 watts by adding a 58-ohm resistor in parallel with the load as shown in Fig. 5C. The impedance of the parallel combination will be 27 ohms, and the total impedance in the circuit will be 77 ohms (50+27). The circuit current will be .52 ampere, dividing through the parallel combination so that .28 ampere flows through the load. The voltage across the source impedance will be 26 volts (50 x .52), leaving 14 volts across the parallel combination (40—26). This means the power delivered to the load will again be 4 watts (.28 x .28 x 50).

In each case, we have delivered 4 watts to the load. However, the impedance match is destroyed in both cases. With the series resistor, in Fig. 5B, the impedance looking into the external circuit from terminals 1 and 2 is now 93 ohms (43 ohms and 50 ohms in series). With the parallel resistor, in Fig. 5C, the impedance seen looking into the circuit from terminals 1 and 2 is 27 ohms (50 ohms and 58 ohms in parallel).

Now, let us see if we can reduce the power delivered to the load from 8 watts to 4 watts and still keep the load matched to the source. Look at Fig. 5D. Here we have a 116-ohm resistor in parallel with the load, giving us a total impedance of 35 ohms for the parallel part of the circuit, and a 15-ohm resistor in series with the 35 ohms offered by the parallel combination. Now, the total impedance of the combination is 50 ohms, the total imped-

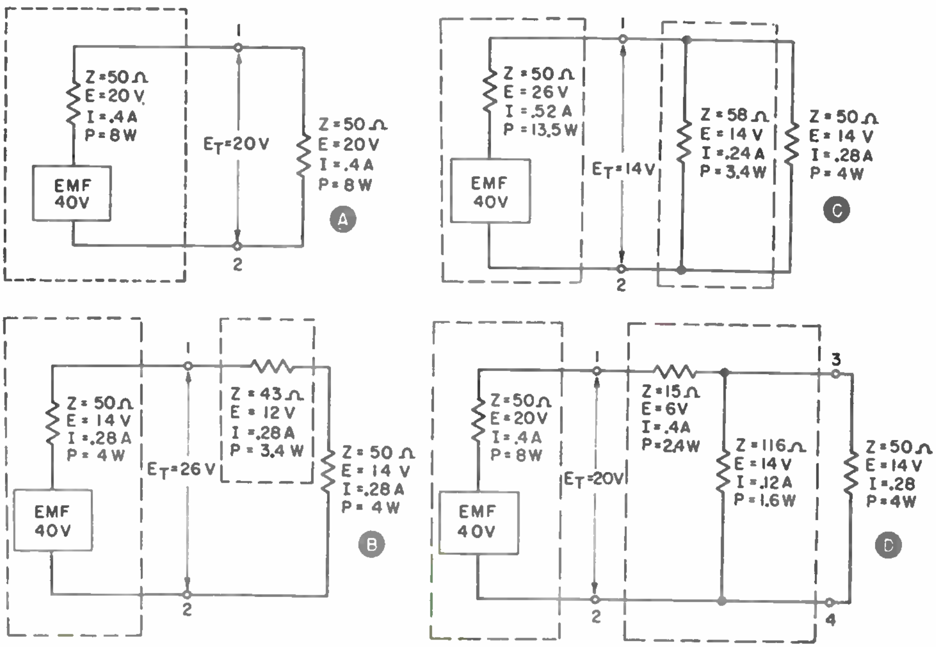


FIG. 5. Effect of dropping resistor on impedance match.

ance in the circuit is 100 ohms (50+50) as in Fig. 5A. The voltage across the load will be 14 volts, the current through it will be .28 ampere, and the power to it will be 4 watts. The total impedance as seen looking into terminals 1 and 2 is still 50 ohms. Thus, the load is still matched to the source even though the power in the load has been reduced to 4 watts.

A simple attenuator such as this that consists of one series resistance and one parallel resistance such as those we have been discussing is called an L pad.

The L Pad. Although the L pad is a definite improvement over a single resistance, it has one major disadvantage. This is the fact that it is unsymmetrical. In other words, it does not have the same impedance at both ends.

Up until now, this has not concerned us because we have been concerned only with matching a load to a source. However, in a perfect impedance match this is not enough. We do not have a perfect match unless the source sees exactly the same impedance looking towards the load as the load will see looking back toward the source.

Looking at the circuits shown in Fig. 6, you will see what we mean by this. In Fig. 6A, when the source looks into the circuit it sees an impedance of 50 ohms. In addition, if we look back into the circuit from the load we also see an impedance of 50 ohms. This is considered to be a perfect impedance match because the load sees exactly the same impedance looking toward the source as the source sees looking toward the load.

Now, let's put an L pad in the circuit as shown in Fig. 6B and see what we have. First, let's look into the circuit toward the load from terminals 1 and 2. Doing this, we see an impedance of 25 ohms in series with the parallel branch containing two 50-ohm resistors, whose combined resistance is 25 ohms. This gives us a total impedance of 50 ohms, which is equal to the source impedance, and we say that the load is matched to the source.

However, if we look back into the circuit towards the source from terminals 3 and 4, we do not see the same impedance. Now, we see an impedance of 50 ohms in parallel with an impedance of 75 ohms. This is a combined impedance of 30 ohms and does not equal the load impedance of 50 ohms. Thus, our L pad matches the load to the source, but does not match the source to the load.

Since an L pad is unsymmetrical, it cannot be used to provide a perfect impedance match between two equal impedances. However, it is often used to provide a match for two unequal impedances. Fig. 7 shows a circuit where a 50-ohm load and a 70-ohm

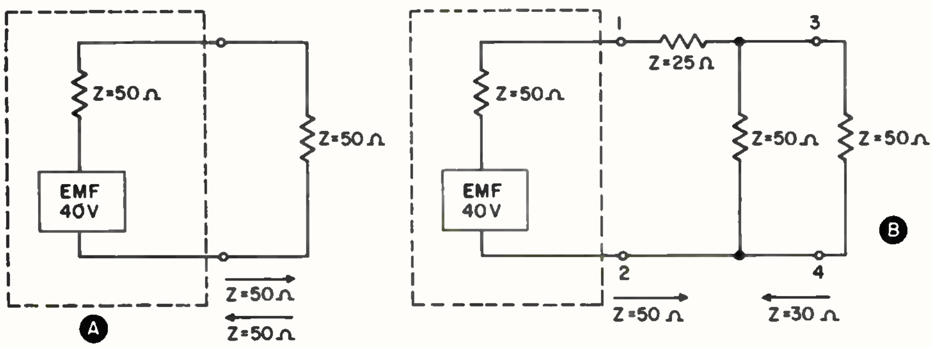


FIG. 6. Effect of unsymmetrical section on impedance match.

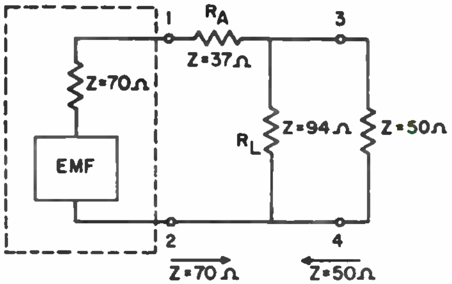


FIG. 7. Using L pad to match unequal impedances.

source have been matched by using an L pad.

Looking into the load from terminals 1 and 2, we see an impedance of approximately 70 ohms (37 ohms in series with the parallel combination of 94 ohms and 50 ohms). Looking back toward the load from the terminals 3 and 4, we see an impedance of almost exactly 50 ohms (94 ohms in parallel with the series combination of 70 and 37 ohms). Of course, the L pad always introduces a loss in the circuit and is not used to match impedance unless the loss can be tolerated or is actually desirable.

Minimum-Loss Matching Pads. By using two fairly simple formulas, we can compute the resistances for an impedance-matching L pad that will introduce the smallest possible loss. To find the value of the series or arm resistance of the pad, we use the following formula:

$$R_A = Z_I\sqrt{1 - \frac{Z_O}{Z_I}}$$

In this formula, R_A = arm resistance; Z_I = source or input impedance; and Z_O = load or output impedance.

To find the value of the parallel or leg resistor, we use the formula:

$$R_L = \frac{Z_O}{\sqrt{1 - \frac{Z_O}{Z_I}}}$$

Here, R_L = leg resistor, Z_I = source or input impedance, and Z_O = load or output impedance as before.

Let's use these two formulas to see if the L pad we used in Fig. 7 provides the smallest possible loss. In this circuit, our input impedance is 70 ohms and the output impedance is 50 ohms. Substituting these values in the formula for obtaining the arm resistance, we get:

$$R_A = Z_I\sqrt{1 - \frac{Z_O}{Z_I}}$$

$$= 70\sqrt{1 - \frac{50}{70}}$$

$$= 70\sqrt{1 - .72}$$

$$= 70\sqrt{.28}$$

$$= 70 \times .53$$

$$= 37\ \Omega$$

From this, we can see that the 37-ohm arm resistance which we used in the circuit is correct for the arm value of a minimum-loss pad.

To find the correct leg value, we substitute our 70-ohm input impedance, and our 50-ohm output impedance in the formula for finding R_L:

$$R_L = \frac{Z_O}{\sqrt{1 - \frac{Z_O}{Z_I}}}$$

$$= \frac{50}{\sqrt{1 - \frac{50}{70}}}$$

$$= \frac{50}{\sqrt{1 - .71}}$$

$$= \frac{50}{\sqrt{.29}}$$

$$= \frac{50}{.54}$$

$$= 92.5\ \Omega$$

This is equal to the leg resistor that we used in the circuit, so we know that the L pad used in Fig. 7 not only matches impedances, but it also introduces the least possible loss for a matching pad. Remember, the two formulas given here are used only to determine the values of resistance for use in a minimum-loss matching pad.

If we are going to design an impedance-matching pad, we will also be interested in knowing just how much attenuation will be introduced by our minimum-loss pad. So far in our discussion of impedance matching and attenuation, we have compared the power delivered to the load with the total power developed by our source. To do this, we have completely analyzed each circuit component as to impedance, voltage, current, and power so that you could actually see what happened to the power and where and why it happened. The amount of the losses provided by our attenuator was rated in watts. Although this is a good approach for purposes of analysis, there is a much simpler way of evaluating attenuation.

Since attenuators are so much a part of communications systems where the decibel is commonly used to express performance, it is logical that the decibel has come to be used as a measure of attenuation. Actually, the decibel is a very convenient unit for use with attenuators because we are concerned with power ratios. As we learned in an earlier lesson, the amount of change in decibels in a circuit is equal to 10 times the logarithm of the ratio of the input power and the output power. Thus, the attenuation or loss in a circuit in decibels can be found with the formula:

$$\text{db} = 10 \log \frac{P_I}{P_O}$$

This formula can be applied to any attenuator quite easily when the power in the circuit is known. However, instead of computing the necessary circuit values to obtain the power ratios in Fig. 7, we can use a more direct formula. This formula uses the input and output impedances directly, and can be used to determine the attenuation in decibels that will be introduced by any minimum-loss pad. In this formula, the loss in decibels equals:

$$\text{db} = 10 \log \left(\sqrt{\frac{Z_I}{Z_O}} + \sqrt{\frac{Z_I}{Z_O} - 1} \right)^2,$$

where Z_I = input impedance and Z_O = output impedance.

Let's apply this to the circuit in Fig. 7 to see how much attenuation our minimum-loss pad will introduce. Substituting in the equation, we have:

$$\begin{aligned}
\text{db} &= 10 \log \left(\sqrt{\frac{70}{50}} + \sqrt{\frac{70}{50} - 1} \right)^2 \\
&= 10 \log (\sqrt{1.4} + \sqrt{1.4 - 1})^2 \\
&= 10 \log (\sqrt{1.4} + \sqrt{.4})^2 \\
&= 10 \log (1.18 + .63)^2 \\
&= 10 \log (1.81)^2 \\
&= 10 \log 3.27 \\
&= 10 \times .5145 \\
&= 5.1 \text{ db loss}
\end{aligned}$$

Thus, the minimum loss possible in an impedance-matching pad for the circuit shown in Fig. 7 is 5.1 db. We have already discovered the proper values of arm and leg resistance for this pad.

SYMMETRICAL PADS

Symmetrical attenuators are similar to L pads except that they have either an extra series arm or an extra leg resistor.

The T Pad. Fig. 8 shows one of the most common types of symmetrical attenuators. This type of attenuator is usually called a T pad because the component resistors are arranged in the schematic so that they form the letter T.

As you can see, the impedances are perfectly matched because the load sees the same impedance looking towards the source from terminals 3 and 4 as the source sees looking into the load from terminals 1 and 2.

T pads can be designed to introduce any desired amount of attenuation. Let's use the circuit shown in Fig. 8 as an example and see what values we need in the T pad. First of all, the power input is 40 watts, and we want a power output of only 1 watt. Using our formula we find that:

$$db = 10 \log \frac{P_I}{P_O}$$

$$= 10 \log \frac{40}{1}$$

$$= 10 \log 40$$

$$= 10 \times 1.6021$$

$$= 16 \text{ db}$$

Now we must find the reduction factor K from the following formula:

$$K = \log^{-1} \frac{db}{20}$$

The symbol "$\log^{-1}$" is called an antilog, which means that our answer will be the number whose logarithm is $\frac{db}{20}$. Substituting our desired attenuation of 16 decibels in the formula, we find:

$$K = \log^{-1} \frac{16}{20} = \log^{-1} .8$$

Now from our log tables, we find that the number which has .8 for a logarithm is 6.31. Thus, our reduction factor K equals 6.31.

Next, we use the reduction factor K in two formulas. The first one will give us a multiplying factor for the arms, and the second will give us a multiplying factor for the leg. Multiplying these by either the source or the load impedance will give us the values for the arm and leg resistors.

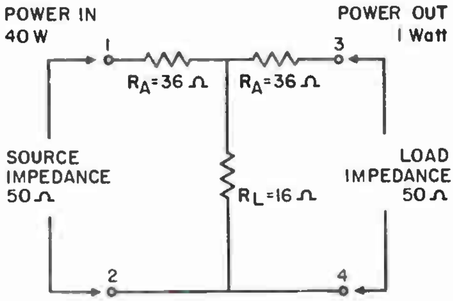


FIG. 8. T pad providing 16-db attenuation.

The formula for the arm factor is:

$$A = \frac{K - 1}{K + 1}$$

Substituting the value of K, which is 6.31 in this formula, we get:

$$A = \frac{6.31 - 1}{6.31 + 1}$$

$$= \frac{5.31}{7.31} = .72$$

Now, by multiplying our 50-ohm impedance by the arm factor A, we find the value of the arm resistors:

$$R_A = AZ = .72 \times 50 = 36\,\Omega$$

Thus, both arm resistors of the T pad must be equal to 36 ohms to give us an attenuation of 16 decibels while maintaining an impedance match between a 50-ohm source and a 50-ohm load.

The formula for the leg factor is:

$$L = \frac{2K}{K^2 - 1}$$

$$= \frac{2 \times 6.31}{6.31^2 - 1}$$

$$= \frac{12.62}{39.8 - 1}$$

$$= \frac{12.62}{38.8}$$

$$= .325$$

Our leg resistor equals:

$$R_L = LZ$$

$$= .325 \times 50$$

$$= 16\ \Omega$$

The resistance values for any T pad giving any desired value of attenuation between two equal impedances may be obtained using these formulas. Because of their importance in attenuator design, we will list them again.

Attenuation in db: $\text{db} = 10 \log \frac{P_I}{P_O}$

Reduction factor K: $K = \log^{-1} \frac{\text{db}}{20}$

Arm factor A: $A = \frac{K - 1}{K + 1}$

Leg factor L: $L = \frac{2K}{K^2 - 1}$

Arm resistors R_A: $R_A = AZ$

Leg resistor R_L: $R_L = LZ$

For convenience in designing attenuators, the multiplying factors for many different values of attenuation have been worked out in the form of tables. Although these tables do not give the values for arm and leg factors for all possible values of attenuation, they are very handy for designing most attenuators. These tables can be found in many reference books.

Attenuation db	Arm factor A	Leg factor L
0	0	∞
1	0.061	8.07
2	.115	4.27
3	.173	2.78
4	.227	2.07
5	.280	1.64
6	.333	1.33
7	.382	1.11
8	.432	0.942
9	.476	.812
10	.520	.700
20	.818	.202
30	.939	.0632
40	.980	.0200

FIG. 9. Table of multiplying factors for the design of any T pad.

A typical example giving the attenuation in decibels and the arm and leg factors is shown in Fig. 9. To use the table, we simply look down the attenuation column until we find the desired loss value. Then we follow across the table to find the arm factor and the leg factor to use with the impedance of our particular circuit in order to compute our arm and leg resistance values.

The π Pad. Another type of symmetrical attenuator is shown in Fig. 10. This is call a pi pad because the component resistances are arranged like the Greek letter π. The π pad in Fig. 10 is like the T pad shown in Fig. 8 in that it has an attenuation loss of 16 db and 50-ohm input and output impedance.

Since the pi pad and the T pad have identical behavior, we can expect a definite numerical relationship between them. We can, for example, set up a pi network like the one shown in Fig. 10 and determine the various resistance values by using the A and L factors that we derived while designing our T network. However, we must use them in a different way, as you will see.

The resistance of the arm of the pi pad is equal to the impedance Z divided by the T pad leg factor L. Similarly, each leg of the pi pad is equal to the impedance Z divided by the T pad arm factor A. In equation form these are:

$$\pi \text{ pad: } R_A = Z/L$$

$$\pi \text{ pad: } R_L = Z/A$$

It should be noted that the factors A and L are exactly the same as those used for the T pad design and are computed the same as before. However, in the pi pad the factors are used for division instead of multiplication, and L is used to find R_A, and A is used to find R_L.

Now let's compute the resistance values of the pi pad components shown in Fig. 10. Our desired attenuation is 16 db, and the source and load impedances are 50 ohms. This is the same requirement that we had in the circuit shown in Fig. 8, so we can use the same A and L factors that we previously computed. For our pi-pad arm, we use the formula:

$$R_A = Z/L$$

$$= \frac{50}{.325} = 154\ \Omega$$

Our leg resistors are found from the formula:

$$R_L = Z/A$$

$$= \frac{50}{.72}$$

$$= 70\ \Omega$$

Thus, our two legs must be 70 ohms, and our arm 154 ohms to provide the same effect with a pi pad as we had with the T pad shown in Fig. 8.

The table shown in Fig. 9 can also be used to construct pi pads. Just remember to divide the impedance by A to find R_L and by L to find R_A. In ordinary use, there is no preference between a T pad and a pi pad. These networks are interchangeable and the type of pad used is a matter of personal choice. Usually, the choice will depend upon the values of resistors that can be obtained easily.

Ganged Attenuator Pads. The input impedance of a symmetrical pad is equal to its output impedance. Thus the input of one pad can be used as the load resistance for a preceding pad. This can be kept up a great number of times and the total attenuation of the pads in tandem will be equal to the sum of the individual pad losses.

Thus, if we have three 50-ohm to 50-ohm pads (symmetrical), giving re-

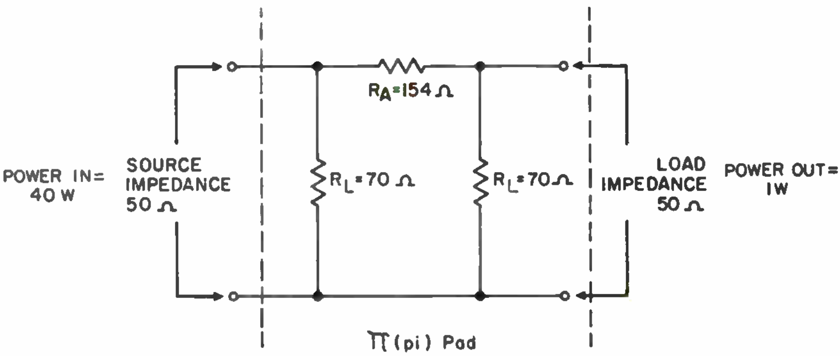


FIG. 10. Pi-pad attenuator (db = 16).

spectively 5, 8, and 4 db attenuation, we can connect these in any order and the whole combination will have an input impedance of 50 ohms. The attenuation will be 5 + 8 + 4 = 17 db when used with a 50-ohm load.

This is a very convenient feature of symmetrical pads, for it means that a great number of attenuation values can be realized with combinations of relatively few separate attenuators. Usually, if an attenuation of over 20 decibels is desired, it is better to obtain it by using two or more attenuators than by using one large-loss attenuator.

Balanced Attenuator Pads. All of the attenuator pads discussed so far have been designed for use with unbalanced lines which have one wire grounded. With balanced lines, where the wires are maintained at opposite potentials with respect to ground, these attenuators would not be suitable. The stray attenuator capacitance would upset the line balance and tend to increase noise pickup and cross talk.

The T and pi pads, however, can be balanced very easily. This is accomplished merely by splitting the series resistance arms in half, and inserting half of the resistance in the upper side of the line and the other half in the lower side of the line. Thus, the pi pad in Fig. 10 becomes the balanced pi pad shown in Fig. 11. In the same manner, the T pad of Fig. 8 is changed to the balanced version in Fig. 12. The balanced pi pad is sometimes called an O pad and the balanced T pad is sometimes called an H pad, because of their resemblance to these letters.

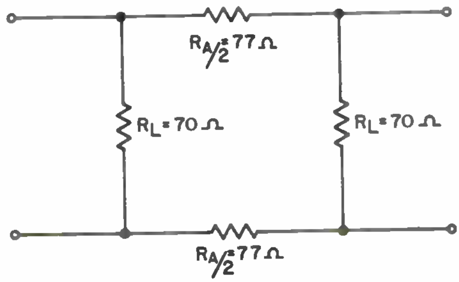


FIG. 11. Balanced pi pad—sometimes called an O pad.

UNSYMMETRICAL PADS

In addition to being used to provide attenuation between equal impedances, T, H, pi, and O pads can also be used as attenuators between unequal impedances. When used in this way, they are designed to match the unequal impedances as well as provide attenuation. We have already studied one type of impedance-matching pad, which we called a minimum-loss L pad. However, the purpose of the L pad was to provide a minimum-loss matching network, and the attenuation it offered had to be considered because it was unavoidable, although undesirable.

The purpose of the networks which we will study now is not only to match unequal impedances but also to provide more attenuation than that which would be provided by a minimum-loss pad. When we studied the L pad, we said it was unsymmetrical because it did not look the same from both ends. The T and pi pads, which are used to match unequal impedances, also look different from each end and are called unsymmetrical T or pi pads.

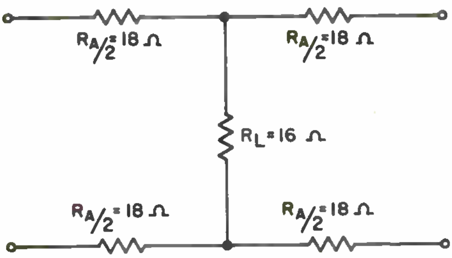


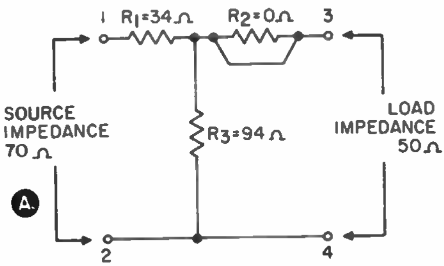
FIG. 12. Balanced T pad—sometimes called an H pad.

In studying the minimum-loss L-type matching pad, we learned that there will always be a certain amount of loss in a resistive matching pad. We also learned how to compute the amount of this minimum-loss in decibels and how to determine the proper resistance values. The minimum-loss L pad could also be considered as a modified T or π pad, as shown in Fig. 13A and B. Here we have shown the minimum-loss L pad as a T pad in which one of the arms has been reduced to zero. It can also be considered to be a pi pad in which one of the leg resistors has become so large it amounts to an open circuit, as shown in Fig. 13B.

From this, we can see that by introducing an extra arm or leg resistor of appropriate size, and changing the other resistance values, we can gain more attenuation and still maintain an impedance match. As soon as we do this, we will have changed our L-pad into an unsymmetrical T or pi pad. It should be noted that, although we can increase the attenuation above the minimum-loss and maintain an impedance match by using an unsymmetrical T or pi pad, we can never decrease the loss below the amount introduced by the minimum-loss L pad.

Therefore, one of the first things to do in designing an unsymmetrical attenuator is to determine the minimum possible loss. If this amount of attenuation is satisfactory, we will use an L pad as an attenuator. If we want more attenuation than the L pad will provide, we will design an unsymmetrical T or pi pad. However, if the minimum loss is more than we want, we will have to decide whether it is advisable to use an attenuator and an amplifier to match impedances or whether it is better to use some other method of matching impedances.

L-Pad Design. Before we go into the design of an unsymmetrical T or pi pad, there is one thing we should discuss regarding the minimum-loss L pad. When we considered the L pad earlier, we studied only left-handed pads like the one shown in Fig. 14A. In this circuit, the input impedance is larger than the output impedance and the arm resistor is in series with the input impedance. However, in a circuit like the one shown in Fig. 14B, the input impedance is smaller than the output impedance, and a right-handed L pad would have to be used.



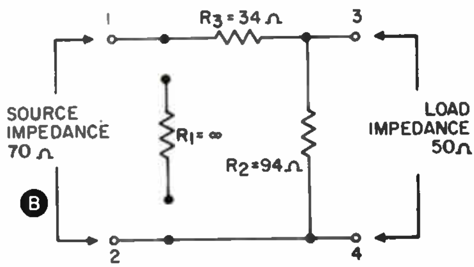


FIG. 13. The L pad can be considered as a modified T pad, as at A, or as a modified π pad, as at B.

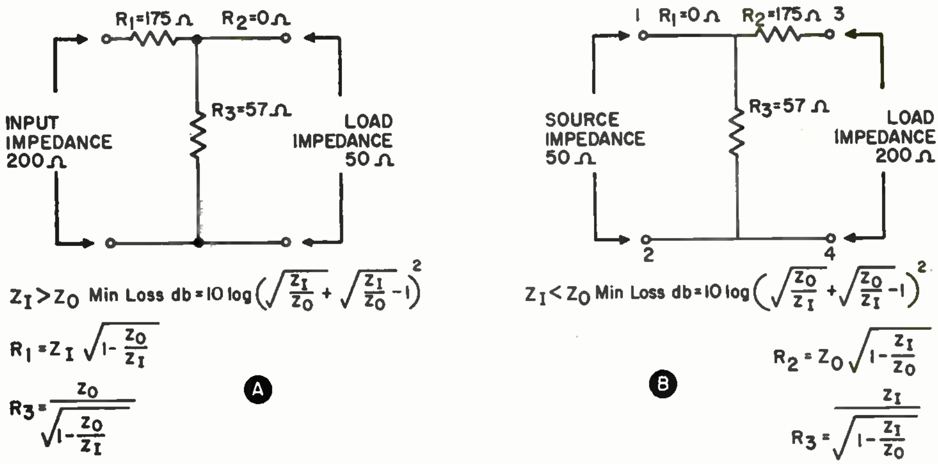


FIG. 14. Computing left-handed L pads, A; computing right-handed L pads, B.

The impedance values in the circuit shown in Fig. 14B are numerically the same, but reversed in position. The value of the arm resistor is the same as it was in a left-handed pad, but its position has changed. Because of this, our formulas for computing right-handed L pads are slightly different from those used in computing left-handed L pads. Also, the minimum loss in decibels will be figured a little differently. The formulas which we have already learned for computing left-handed L pads and their minimum loss are shown in Fig. 14A.

In a right-handed pad we compute for R2 and R3 instead of R1 and R3, and the formulas are slightly different. Now the positions of the output impedance and the input impedances in our formulas are just opposite what they were for the left-handed pads. You will also notice that the impedance relationships have been reversed in the minimum-loss formula.

Since computing the minimum loss is one of the first steps in designing any unsymmetrical attenuator network, this change in the minimum-loss formula is very important. As shown, when the input impedance, Z_I is greater than (>) the output impedance Z_O, the loss in decibels will be found from the following formula:

$$\text{Min. loss db} = 10 \log\left(\sqrt{\frac{Z_I}{Z_O}} + \sqrt{\frac{Z_I}{Z_O} - 1}\right)^2$$

However, when the input impedance Z_I is less than (<) the output impedance Z_O, the loss in decibels will be found from:

$$\text{Min. loss db} = 10 \log\left(\sqrt{\frac{Z_O}{Z_I}} + \sqrt{\frac{Z_O}{Z_I} - 1}\right)^2$$

Designing an unsymmetrical T or pi pad is simply using the proper formulas correctly. We will not attempt to show you how the formulas are derived, nor will you be expected to memorize them. The important thing is that you know how to use the formulas and where to find them when they are needed.

Unsymmetrical T or π Pads. The formulas for computing the resistance values of an unsymmetrical T pad are:

$$R_3 = \frac{2\sqrt{NZ_I Z_O}}{N - 1}$$

$$R_1 = Z_I \left(\frac{N + 1}{N - 1}\right) - R_3$$

$$R_2 = Z_O \left(\frac{N + 1}{N - 1}\right) - R_3$$

In the formulas, N is used to denote the power ratio: $P_I \div P_O$. We can find the value of N once we know our desired attenuation in decibels using the formula:

$$\text{db} = 10 \log \frac{P_I}{P_O}$$

Now let's use these formulas to compute an unsymmetrical T pad that will give us an attenuation of 15 db

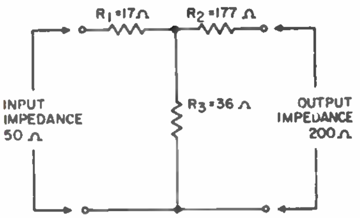


FIG. 15. Unsymmetrical T pad.

as shown in Fig. 15. The first thing we must do is determine whether or not an attenuation of 15 db with matched impedances is possible in this circuit. To do this, we compute the minimum possible loss for matched impedances. Since our input impedance is less than our output impedance, we use the formula:

$$\text{Min. loss db} = 10 \log \left(\sqrt{\frac{Z_O}{Z_I}} + \sqrt{\frac{Z_O}{Z_I} - 1}\right)^2$$

$$= 10 \log \left(\sqrt{\frac{200}{50}} + \sqrt{\frac{200}{50} - 1}\right)^2$$

$$= 10 \log (\sqrt{4} + \sqrt{4 - 1})^2$$

$$= 10 \log (2 + \sqrt{3})^2$$

$$= 10 \log (2 + 1.73)^2$$

$$= 10 \log 3.73^2$$

$$= 10 \log 13.9$$

$$= 11.43 \text{ db}$$

Since our desired attenuation of 15 db is more than this minimum of 11.43 db, we know that it is possible to construct the desired attenuator.

Now, we must find the correct value of N to use. By substituting N in place of the power ratio in our formula:

$$\text{db} = 10 \log \frac{P_I}{P_O} \quad \text{we get:}$$

$$\text{db} = 10 \log N$$

$$15 = 10 \log N$$

$$\frac{15}{10} = \log N$$

$$1.5 = \log N$$

Checking our log tables, we find that 1.5 is the logarithm of 31.7, which we can round off to 32.

Now, using our formula for the leg resistor R_3, we find:

$$R_3 = \frac{2\sqrt{NZ_I Z_O}}{N - 1}$$

$$= \frac{2\sqrt{32 \times 50 \times 200}}{32 - 1}$$

$$= \frac{2\sqrt{32 \times 10{,}000}}{31}$$

$$= \frac{2 \times 100 \sqrt{32}}{31}$$

$$= \frac{200 \times 5.65}{31}$$

$$= \frac{1130}{31}$$

$$= 36\ \Omega$$

Next, our arm resistor R_1 is:

$$R_1 = Z_I \left(\frac{N + 1}{N - 1}\right) - R_3$$

$$= 50 \left(\frac{32 + 1}{32 - 1}\right) - 36$$

$$= \frac{1650}{31} - 36$$

$$= 53 - 36$$

$$= 17\ \Omega$$

And our other arm R_2 is:

$$R_2 = Z_0\left(\frac{N+1}{N-1}\right) - R_3$$

$$= 200\left(\frac{32+1}{32-1}\right) - 36$$

$$= 200 \times \frac{33}{31} - 36$$

$$= \frac{6600}{31} - 36$$

$$= 213 - 36 = 177\ \Omega$$

If we prefer, we can use a pi pad

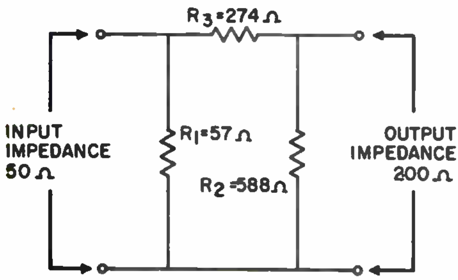


FIG. 16. Unsymmetrical pi pad.

as shown in Fig. 16. The following formulas are for computing unsymmetrical pi pads.

The arm R_3 is:

$$R_3 = \frac{N-1}{2}\sqrt{\frac{Z_I\,Z_O}{N}}$$

$$= \frac{32-1}{2}\sqrt{\frac{50 \times 200}{32}}$$

$$= \frac{31}{2}\sqrt{\frac{10{,}000}{32}}$$

$$= \frac{31}{2} \times 100\sqrt{\frac{1}{32}}$$

$$= \frac{3100}{2} \times \frac{1}{5.65}$$

$$= \frac{3100}{11.3} = 274\ \Omega$$

The leg R_1 is:

$$\frac{1}{R_1} = \frac{1}{Z_I}\left(\frac{N+1}{N-1}\right) - \frac{1}{R_3}$$

$$= \frac{1}{50}\left(\frac{32+1}{32-1}\right) - \frac{1}{274}$$

$$= \frac{1}{50} \times \frac{33}{31} - \frac{1}{274}$$

$$= \frac{33}{1550} - \frac{1}{274}$$

$$= .0213 - .0036 = 0177$$

If $\frac{1}{R_1} = .0177$, then $R_1 = \frac{1}{.0177}$

$$= 57\ \Omega$$

The leg R_2 is:

$$\frac{1}{R_2} = \frac{1}{Z_O}\left(\frac{N+1}{N-1}\right) - \frac{1}{R_3}$$

$$= \frac{1}{200} \times \frac{33}{31} - \frac{1}{274}$$

$$= \frac{33}{6200} - \frac{1}{274}$$

$$= .0053 - .0036$$

$$= .0017$$

If $\frac{1}{R_2} = .0017$, then $R_2 = \frac{1}{.0017}$

$$= 588\ \Omega$$

Thus, we can construct either a T or a pi pad that will match unequal impedances and give us 15 db attenuation. A balanced T (H) pad or a balanced pi (O) pad could be constructed by splitting the arm resistors and dividing them between the upper and lower lines.

VARIABLE ATTENUATORS

The attenuation introduced by an L pad, a T pad, or a pi pad can be made

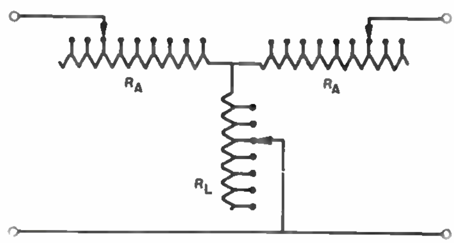


FIG. 17. The usual style of tapped T-pad variable attenuator.

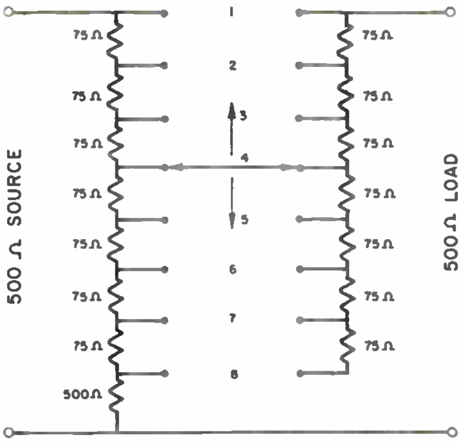


FIG. 18. An approximate T pad using only two variable resistors.

adjustable by using variable resistors for the arms and legs. The T pad of Fig. 8, for example, could be made with variable R_A and R_L values.

If this is done, however, the relations outlined in the table of Fig. 9 must be maintained or the variable pad will not have constant input and output impedances. We could do this by using taper-wound variable resistances ganged together so that all resistances varied by the proper amount when one control was operated. However, this would be quite difficult to achieve in actual practice because each resistor would have to be calibrated and wound with such precision that it would become quite costly.

A reasonable, practical design of the variable attenuator is made possible by the fact that attenuation does not need to be continuously variable. If the variations are made in equal steps of less than 2 db each, the variation will appear to be continuous for most purposes. This allows us to use a set of fixed resistors and a three-gang switch as shown in Fig. 17.

One advantage of using step attenuation in the design of the variable T pad is that such an arrangement can be calibrated very accurately. In addition, the gang switch can be made with large wiping contacts which will keep contact noise at a minimum. This is in contrast to a rheostat or a potentiometer which seldom can be reset to a given point, and usually has a high degree of contact noise.

To increase the amount of attenuation in this pad, the values of R_A must be increased and the value of R_L must be decreased. For this reason, then, it is possible to combine one of the R_A's with R_L so that the simpler construction shown in Fig. 18 can be used.

In the figure, the right-hand resistance forms the right arm of the T pad, the upper part of the left-hand resistance forms the left arm, and the lower part forms the leg. This arrangement is not a true T pad, but only an approximation since it is not possible to make the attenuator have exactly constant values of input and output resistance at all settings.

As shown in the table of Fig. 19, it varies from 336 to 861 ohms.

It will be noted that the use of this particular attenuator involves a loss of 1.9 db at the minimum setting. This is called the insertion loss of this particular device.

Decade Attenuator Pads. Suppose we have need for an attenuator adjustable in one decibel steps from 0 to

Switch position	Input Resistance	db Loss
1	336	1.9
2	432	3.5
3	522	5.0
4	605	6.5
5	680	7.9
6	748	9.3
7	808	10.8
8	861	12.4

FIG. 19. Input resistance and db loss of the approximate T pad for each switch point.

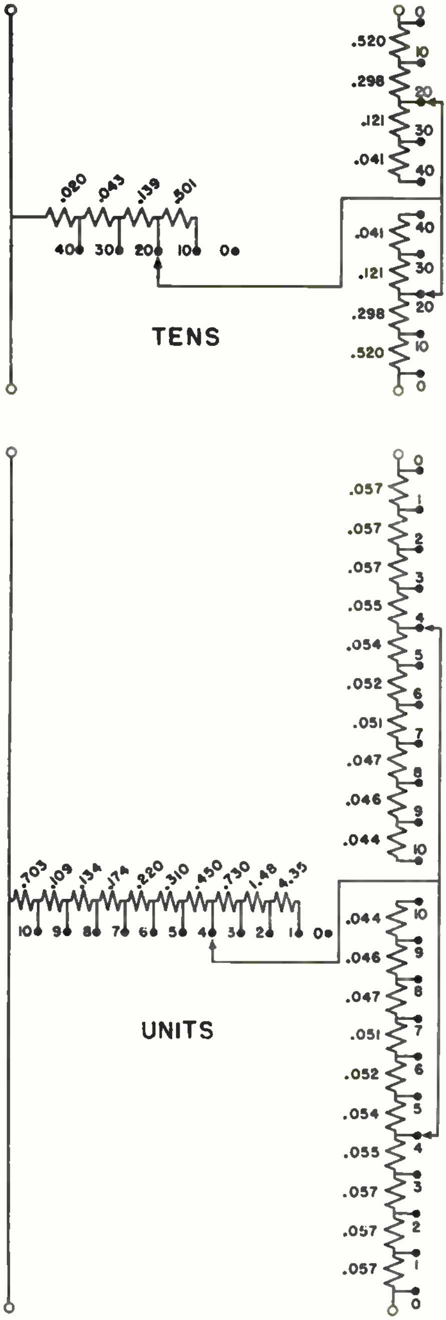


FIG. 20. A decade attenuator adjustable in 1 db steps from 0 to 50 db, designed for a source-load impedance of 1 ohm.

50 that maintains perfect impedance matching at all settings. We could, of course, design a conventional T pad having 50 steps, but this would be expensive and awkward to use.

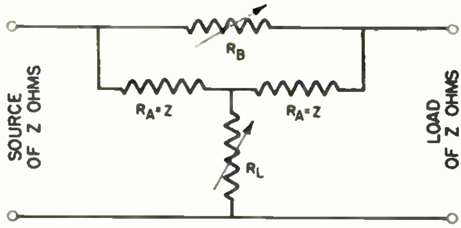


FIG. 21. The bridged-T is a true symmetrical pad with only two variable resistors.

A better method would be to build two attenuators, one having four steps of 10 decibels each, and the other having ten steps of 1 decibel each. Then, if these are constant-impedance pads, they can be connected in tandem so that the separate attenuations will add. Thus, one knob controls the tens, and the other controls the units of attenuation. Such an arrangement is called a decade attenuator.

The design for such a decade attenuator is shown in Fig. 20. This decade pad is designed to work between a source of 1 ohm and a load of 1 ohm. To find the values for any other source and load impedance, simply multiply the value shown alongside each resistor by the impedance Z.

The Bridged T. A modification of the T pad, of considerable importance because it requires only two variable resistors, is shown in Fig. 21. This is known as the bridged T. It is a symmetrical attenuator having two fixed arms, bridged from input to output by a variable arm, and a variable leg.

The resistances of the fixed arms R_A are always equal to the input-output impedance Z. The bridge and leg resistors R_B and R_L are related to the impedance Z in this way:

$$R_B = B \times Z \text{ and}$$

$$R_L = L \times Z.$$

The multiplying factors, B and L, can be determined from the table in Fig. 22. Pay particular attention to the arm and leg factors for 6-db attenuation. Both are almost exactly equal to 1. This means that 6 db of attenuation can be obtained at any impedance by connecting four resistors, all equal to the impedance, in a bridged T. Any other degree of attenuation can be obtained by changing R_B and R_L to the proper values.

You have seen how networks can be designed to give impedance-matching with minimum attenuation, or to give a desired amount of attenuation. Now let's look at some networks designed to correct uneven attenuation or phase shift at different frequencies. These are called "equalizers."

Attenuation db	Bridge-arm factor B	Leg factor L
0	0	∞
1	0.122	8.20
2	.259	3.86
3	.413	2.42
4	.585	1.71
5	.778	1.285
6	.995	1.005
7	1.239	0.808
8	1.512	.661
9	1.818	.550
10	2.162	.462
20	9.000	.111
30	30.62	.0326
40	99.00	.0101

FIG. 22. Multiplying factors for design of the bridged-T.

Equalizers

Transmission lines and certain types of equipment ordinarily introduce a certain amount of distortion in a signal because of unequal attenuation and phase shift at the various frequencies. It would be extremely difficult to design lines and equipment that would always provide a distortion-free signal. The more practical solution is to use compensating networks to introduce additional attenuation or phase shift so that all frequencies will be affected in substantially the same manner.

LINE EQUALIZERS

Long lines (such as telephone lines or cables) ordinarily attenuate high frequencies more than the low frequencies. Capacitance between the wires acts like a capacitor connected across the line ahead of the load, shorting out the high frequencies. A crude form of compensation could be obtained by connecting a capacitor in series with the line to partially block the low frequencies while passing the highs. A more practical arrangement uses an inductance-resistance combination shunted across the line to cause an equivalent loss of the low frequencies and balance the response.

Such an arrangement is the Western Electric 23A Equalizer shown in Fig. 23. As you can see, it consists of a coil and a capacitor in parallel, and a tapped series resistance. The components are mounted in an aluminum-finished metal box, and the resistor taps are brought out to terminals at one end of the box, so any needed value of resistance can be selected.

It can be easily adjusted to an unloaded line with ordinary station equipment. Suppose it is to be adjusted to a circuit consisting of the sending-station line amplifier, repeating coil, (which is an impedance-transforming device similar to a transformer), transmission line, receiving-station repeating coil, and receiving-

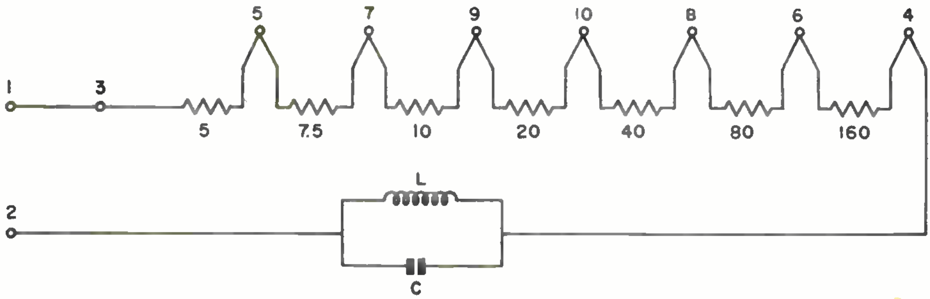


FIG. 23. The Western Electric 23A equalizer.

station line amplifier as shown in Fig. 24.

VU meters are connected across the line following the sending and receiving line amplifiers, and the receiving amplifier is terminated in a 600-ohm load. An attenuator pad is used ahead of the repeating coil at the sending end to prevent cross-talk or interference with other lines in the same cable. Another attenuator is used at the receiving end to avoid overloading the amplifier. To see what values of resistance to use in the equalizer, the output of a variable-frequency oscillator is fed into the sending amplifier, and the loss at 1000 cycles and 8000 cycles noted.

Then the equalizer is connected across the line side of the receiving repeating coil, but an adjustable resistance box is connected in place of the tapped resistor in the equalizer. That is, terminal 2 of the equalizer is connected to one side of the line. The resistance box is connected between terminal 4 of the equalizer and the other side of the line. Terminal 1 of the equalizer is left open. Various resistance values are tried (using the resistance box) until the loss is approximately the same at 1000 and 8000 cycles. Then the resistance box is removed and the portion of the fixed resistance in the equalizer with a value closest to that determined by experiment is connected in its place. This is done by connecting terminal 1 to the line from which the resistance box was removed, and moving the lead from the L-C combination from terminal 4 to the proper terminal. For example, if 12 or 13 ohms is needed, the lead would be moved from terminal 4 to terminal 7, since the resistance from terminal 7 to terminal 1 is 12.5 ohms, which is as close as we can get.

For temporary lines or intermittent service where a permanent installation is not justified, an adjustable equalizer such as the Western Electric 279A equalizer panel shown in Fig. 25 can be used to connect it to any one of several lines as the occasion requires.

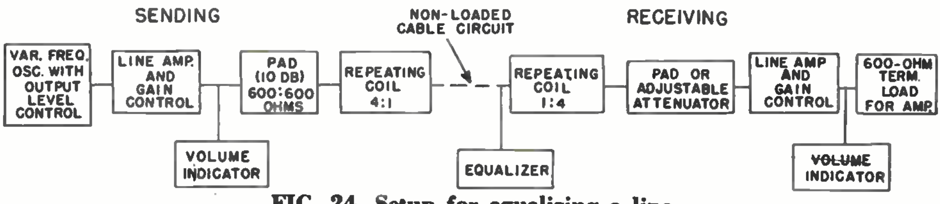


FIG. 24. Setup for equalizing a line.

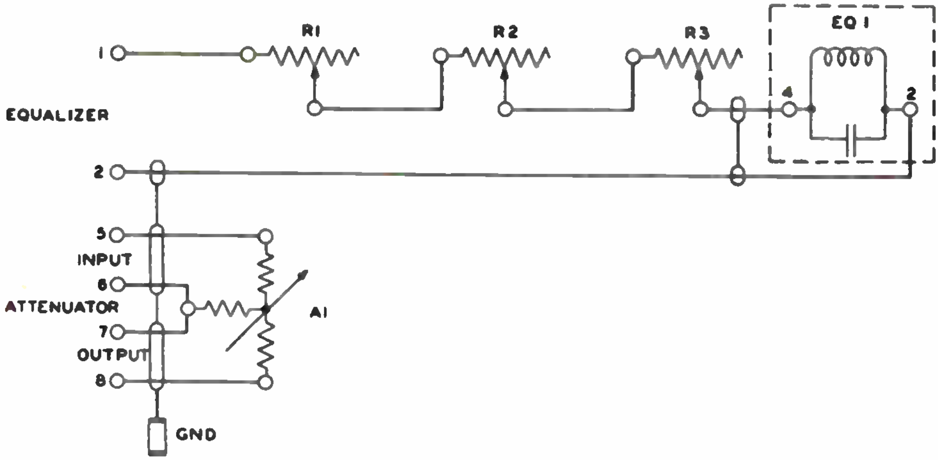


FIG. 25. The Western Electric 279A equalizer panel.

PICKUP EQUALIZERS

Another way in which equalizers are used is to compensate for the electrodynamic characteristics of phonograph and tape pickups. They are also used to compensate for the particular recording curve used in making records. These are called pickup equalizers.

Recording methods may be divided into two general classes: constant-velocity and constant-amplitude. When a record is cut, the stylus in the cutting head cuts a groove in the record. When no signal is applied, the groove forms a perfect spiral. When a signal is applied, the stylus moves in a transverse (side to side) direction, making a slight wiggle in the groove.

When a signal of constant amplitude is applied to a constant-velocity cutting head, the transverse velocity of the stylus is constant regardless of frequency, but the displacement from side to side decreases with frequency. With a constant-amplitude cutting head, the transverse displacement of the stylus is constant (for tones of constant amplitude), but the transverse velocity of the stylus varies directly with the frequency.

Just as cutting heads can be divided into two types, record pickup heads may also be divided into two classes, "velocity-operated" and "amplitude-operated." In a velocity-operated pickup, the output signal is generated by changing the flux of a permanent magnet as it passes through a coil. Different types of heads (magnetic, dynamic, variable reluctance) accomplish this in slightly different ways, but the operating principle is the same.

The voltage generated is proportional to the rate at which the flux changes, and the flux is changed by the movement of the stylus. Thus, the signal voltage generated is dependent on the speed of the stylus. This pickup is the natural opposite of a constant-velocity recorder. If a constant-level signal is recorded, a constant-level output is obtained.

The output of an amplitude-operated pickup is constant for constant groove undulation (side to side displacement) amplitude. Common amplitude-operated pickups are the crystal and ceramic types. These operate on the piezo-electric effect, generating a voltage by the physical distortion of a crystal. A constant-amplitude signal recorded with a crystal cutter will be

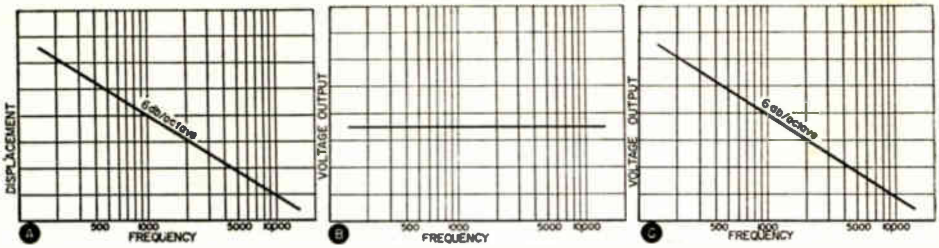


FIG. 26. A, variations of magnetic-cutter stylus displacement with frequency. B, output voltage of magnetic pickup on constant-velocity recording. C, output voltage of crystal pickup on constant velocity recording.

reproduced as a constant output voltage when played back through a crystal pickup. A constant-amplitude signal recorded with a constant-velocity cutter will be reproduced as a voltage which falls off at the rate of 6 db per octave as the frequency is increased.

Fig. 26A shows how the undulation amplitude varies with frequency in a constant-velocity record for tones of equal signal level. When a magnetic play-back head is used, the amplitude of the output signal is the same at all frequencies as shown in Fig. 26B, but with an amplitude-operated pickup, the voltage falls off as in Fig. 26C.

Fig. 27A shows the variation in stylus displacement with frequency for tones of equal amplitude when recorded with a constant-amplitude recording head. If the record is played back through a constant-amplitude pickup, the output voltage is the same at all frequencies as shown in Fig. 27B. If played back through a constant-velocity pickup, the output is as shown in Fig. 27C.

Recording Curves. In actual practice, both constant-velocity and constant-amplitude characteristics are used on parts of any record. When attempts were made to extend the high-frequency range of recordings, difficulty was experienced in keeping the high-frequency notes from dropping down to the noise level. Also, with constant-velocity recording, the stylus sometimes cut into the next groove when a loud, low note was recorded. These problems were solved by pre-emphasizing the highs and de-emphasizing the lows during recording, and using equalizers in the play-back amplifiers to make the necessary compensation.

The only trouble with this arrangement was that there was no agreement between companies as to exactly what system to use, that is, how much pre-emphasis was necessary or at what

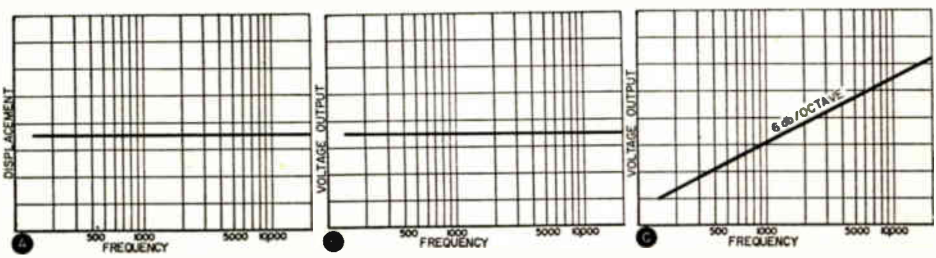


FIG. 27. A, crystal cutter stylus displacement variation with frequency for constant signal level. B, voltage output variation when record (in A) is played back through crystal pickup. C, output voltage variation when same record is played through magnetic pickup.

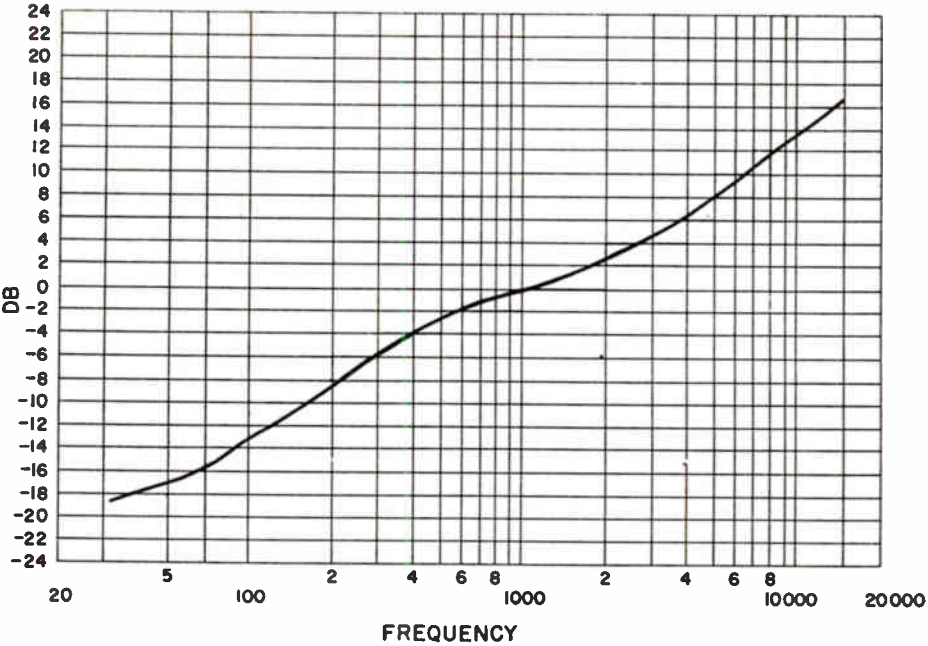


FIG. 28. The RIAA recording curve.

frequency it should start, etc. Various techniques were tried, and several recording curves were developed. As a result, it was very difficult to design playback amplifiers to reproduce all records satisfactorily with any type of pickup. Recently, the Recording Industries Association of America adopted a standard curve and recommended it to the entire recording industry. The RIAA curve, shown in Fig. 28, is identical to the NARTB (National Association of Radio and Television Broadcasters) curve and also to RCA's Orthophonic curve. The Audio Engineering Society has recommended it as a replacement for their AES curve. Nearly all the major record manufacturers now use this curve on their new records.

Equalizer Circuits. For best possible frequency response, a pickup must be terminated correctly at the input of the amplifier. The ideal pickup and termination would give a flat frequency response over the entire audio range.

Fig. 29 shows the equivalent circuit of a ceramic cartridge and its termination. The pickup can be thought of as a generator and a capacitor in series, and the load as a resistance and a capacitor in parallel. When the two are connected, an RC voltage divider results. At low frequencies, the reactance of C is so high it can be ignored, as in Fig. 29B, but its effects increase as the frequency goes up, as shown in Fig. 29C.

For any given pickup, C is fixed, as it depends on the cable length and the amplifier construction, so R is the only part that can be varied. Fig. 29D shows how varying R affects the output.

The equivalent circuit of a magnetic pickup and its termination is shown in Fig. 30A. This circuit contains inductance, capacitance, and reactance and acts as a damped resonant circuit. If R_L is too large, the output will have resonant peaks at frequencies deter-

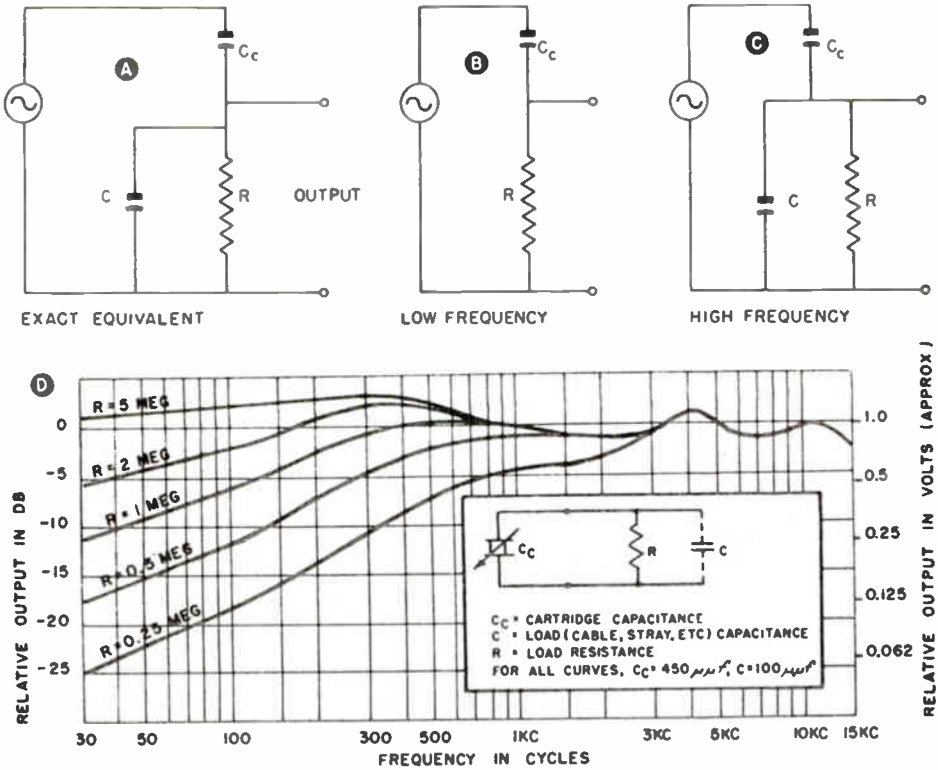


FIG. 29. A, B, C equivalent circuits of ceramic cartridge and termination; D, change in output with change in terminating resistance.

mined by the values of L and C_S. If R_L is small, C_S is effectively shorted and the inductance will cause the highs to fall off. Using a longer cable would cause C_S to increase, and there would be resonant peaks at a lower frequency. If R_L is reduced to compensate for this, the high-frequency output is reduced. The best frequency range is obtained by keeping the cable capacitance small and R_L high.

Fig. 30B shows a crystal pickup and its termination. The output of a crystal cartridge falls off rapidly at high frequencies. R_1 and C in series with R_2 correct for this. At low frequencies, capacitor C has little effect on the signal appearing at the output.

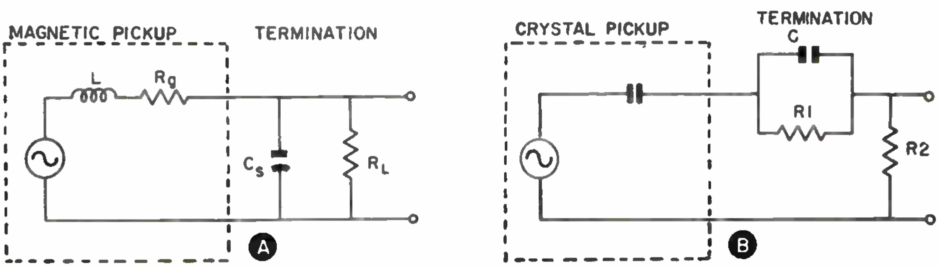


FIG. 30. Equivalent circuit and input termination for A, magnetic, and B, crystal pickup.

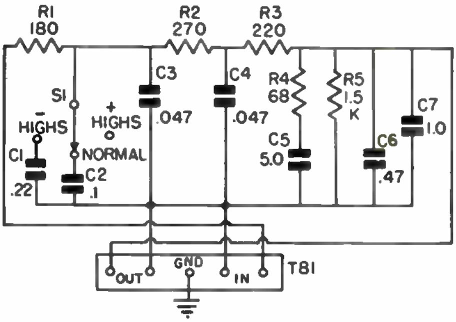


FIG. 31. The RCA MI-11888 pickup equalizer.

As the frequency goes up, the reactance of C goes down. This reduced reactance shunting resistor R_1 increases the signal at the output.

Turntable Equalizers. A typical equalizer for use with a broadcast turntable is shown in Fig. 31. This is the RCA MI-11888, designed to work with the RCA Type 70 Series turntable and the BQ-2A turntable, and is for use only with the MI-11874-4 pickup head.

The equalizer is mounted inside the turntable cabinet in the right-hand front corner, where it will be convenient for use by the operator.

Three settings of the equalizer are provided. In the diagram, the switch is shown in the "normal" position. When it is desired to accentuate the highs, the switch is moved to the "+Highs" position, which removes some of the shunt capacitance from the circuit. When the opposite effect is desired, the switch is thrown to the "—Highs" position, which adds shunt capacitance to the circuit.

The response curves for the different settings of the equalizer switch are shown in Fig. 32.

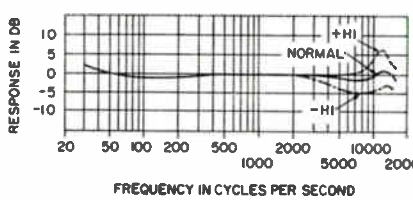


FIG. 32. Frequency response curves for the MI-11888.

TAPE SYSTEM EQUALIZERS

There are a number of factors that influence the frequency response of a record and play-back tape system. Fig. 33 shows the response curves for two standard tape speeds—15 inches per second and 7.5 inches per second. The slower-moving tape has the poorer high-frequency response. This is because, when a signal is recorded on a magnetic tape, the particles of the coating form many very small areas which act like tiny magnets having north and south poles. The slower the tape moves past the recording head or the higher the frequency of the applied signal, the shorter the magnetic bars become. This causes a canceling effect between the opposite poles, which reduces the magnetism induced in the tape.

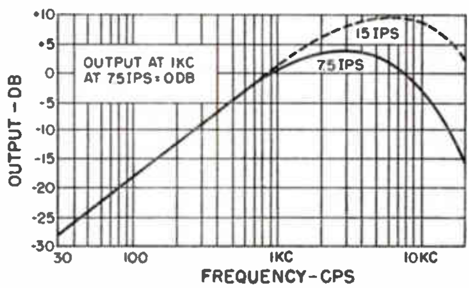


FIG. 33. Response of a record playback head.

Another cause of high-frequency loss is that the recording head is largely inductive. Thus, it has more impedance at high frequencies, decreasing the strength of the magnetic field between the pole pieces.

When the tape is played back, the voltage induced in the play-back coil varies with the rate of change of the flux lines. A high-frequency tone of the same amplitude occupies a shorter length of tape than a lower-frequency tone of the same amplitude. Thus, the

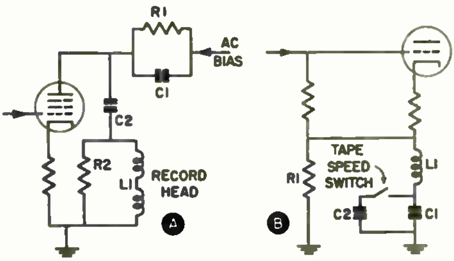


FIG. 34. Circuits for pre-emphasis.

change of flux lines is at a faster rate at higher frequencies (or a faster speed) than at a low frequency (or a slower speed).

To get a linear frequency response in a tape-recorder system, equalizing networks are necessary to boost the amplitude of the frequencies at both the high and low ends of the audio range. The usual practice is to boost the highs in the recorder and the lows (and sometimes also the highs) in the play-back amplifier. The former is called pre-emphasis and the latter, post-emphasis.

Pre-Emphasis. Two typical pre-emphasis circuits are shown in Fig. 34. In Fig. 34A, capacitor C_2 is in series with the recording head across the output of the amplifier tube. The reactance of C_2 decreases as the frequency increases, so a high-frequency signal will produce a higher current flow in the coil of the recording head than a low-frequency one will.

In Fig. 34B, a series-resonant circuit in the amplifier cathode circuit is used, At low frequencies, a degenerative voltage is developed across resistor R_1, feeding an out-of-phase signal to the grid, decreasing the gain. At higher frequencies, coil L_1 and capacitor C_1 (or C_1 and C_2 in parallel) have an increasing shunting effect on resistor R_1, decreasing the feedback, and the gain goes up. Capacitors C_1 and C_2 are connected in parallel for high tape speeds, but the switch is opened at low speed.

Post-Emphasis. Two typical post-emphasis circuits are shown in Fig. 35. In Fig. 35A, the bass response is boosted by the series network consisting of capacitor C_2 and resistors R_2 and R_3. The reactance of C_2 is so high at low frequencies that the series network of C_2 and R_2 has no shunting effect, and the lows pass unattenuated to the grid of the next stage. As the frequency goes up, the shunting effect of the two components increases, and a lower proportion of the high-frequency signal appears at the grid. Capacitor C_3 is an adjustable treble control used to correct for the very-high-frequency losses in a tape recorder system. At the higher frequencies, C_3 shunts some of the highs around R_3 and on to the grid.

In the circuit shown in Fig. 35B, the middle range of frequencies is fed back to the input circuit through re-

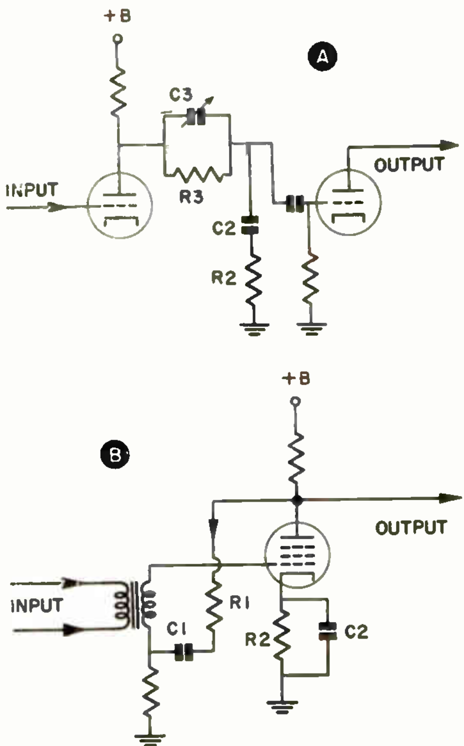


FIG. 35. Post-emphasis circuits.

sistor R_1 and capacitor C_1. There is no feedback of lows because of the rising capacitance of capacitor C_1. The feedback reduces the gain so the gain decreases with an increase in frequency.

The values of resistor R_2 and capacitor C_2 can be chosen to give a small amount of boost at very high frequencies. If C_2 is small, its low reactance at very high frequencies prevents cathode feedback and increases the high-frequency gain.

These are only some of the many uses for equalizers in broadcast service. Equalizers and filters, which we will take up in the next section of this text, are used to extend the frequency range of a sound system, to add sound effects without the use of mechanical apparatus, to remove undesirable sound effects, to improve tonal qualities resulting from poor acoustical location, and to compensate for variations in system response, acoustical conditions, and program material.

Filters

Basically, filters consist of an arrangement of reactive components that will respond in certain predetermined ways to different frequencies. Filters can be designed to pass low-frequency signals and block signals at high frequencies. Others can be designed to pass high frequencies and block low frequencies. Still others may either pass or block certain definite bands of frequencies. A fundamental rule is that a network cannot act like a filter unless it contains some reactive component, such as a coil or capacitor, that changes its reactance with frequency.

BASIC FILTER CIRCUITS

The operation of the filter depends not only upon the components used, but also on the values and circuit arrangements of the parts. Filter elements can be connected in series with the load or in parallel with the load. Let us see what effect these connections have.

High-Pass Filter. A capacitor in series with the load, or a coil in parallel with the load, as shown in Fig. 36A, forms a simple high-pass filter. At low frequencies, the current is blocked from the load by the capacitor or shorted around the load by the coil. As the frequency increases, more and more current is passed by the capacitor and reaches the load, or if a coil is used, the coil reactance goes up as

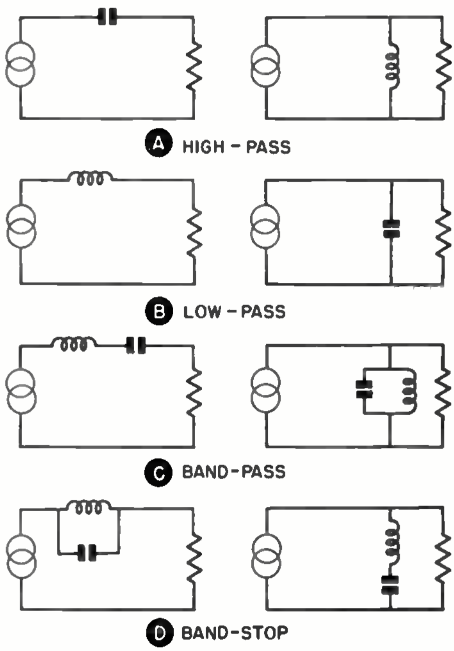


FIG. 36. Types of filters.

the frequency increases and more of the current is forced through the load.

Low-Pass Filter. If we put the capacitor in parallel with the load or the coil in series with the load as at B, the effect will be reversed, and we will have a low-pass filter.

Band-Pass Filter. Suppose we use both a coil and a capacitor as in Fig. 36C. If we put them in series with the load and each other we will have a band-pass circuit. Now, low frequencies will be blocked by the capacitor, resonant, and act like a high impedance and the current will flow through the load.

Band-Stop Filter. Band-stop filters are shown in Fig. 36D. When the coil and capacitor are connected in parallel with each other and placed in series with the load, current reaches the load through the capacitor at high frequencies and through the coil at low frequencies. At resonance, the blocking action of the circulating current in the resonant circuit prevents current from

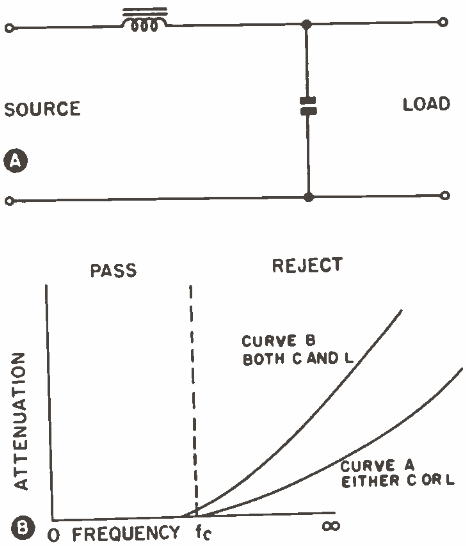


FIG. 37. Basic low-pass L filter, A; filter response curves, B.

and the higher frequencies will be blocked by the coil. However, at some intermediate frequency, the inductive and capacitive reactance will be exactly equal and opposite, the circuit will be series resonant and will pass current. This is a band-pass filter because a small band of frequencies about the resonant point will be passed. If we put the two in parallel with the load, low frequencies will be bypassed around the load by the coil, and high frequencies by the capacitor. At some frequency the two will be reaching the load. If they are connected in series with each other across the load, at resonance the current will be shorted around the load. Above and below resonance, either the coil or the capacitor will block the current and force it through the load.

These four basic circuits (low-pass, high-pass, band-pass, band-stop) illustrate the principles of filtering, but, in actual practice, we have to consider the sharpness of the cut-off action and the effect of the filter on the impedance relationship. For this rea-

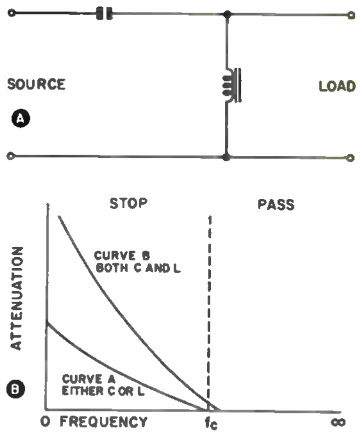


FIG. 38. Basic high-pass L filter, A; filter response curves, B.

son, most practical filter circuits are combinations of the principles we have discussed. One of the simplest of these is the L filter.

L Filter. Let's see why the circuit in Fig. 37 gives an improved response. Either the coil or capacitor alone forms a low-pass filter. With either one alone, the current in the load will decrease as the frequency goes up until, finally, at some high frequency practically no current reaches the load. By using both components, we get the benefit of two effects. If the values are chosen properly, at the cut-off frequency the coil will begin to block quite a lot of current and the capacitor will begin to pass a substantial amount of current. Thus, the coil blocks some of the current and the capacitor passes some that gets through the coil, and we have a sharper response curve as shown by curve B in Fig. 37B.

The high-pass filter shown in Fig. 38 also combines capacitance and inductance in an L-type filter arrangement. Notice that in the circuit in Fig. 38A, we use a series capacitor and a shunt coil to provide a sharper attenuation than we would get by using either component by itself. The response curves are shown in Fig. 38B.

Band-pass and band-stop filters can also be improved by arranging the components in an L network. In Fig. 39 we show the network and the curves for a band-pass filter. If the series-resonant and the parallel-resonant circuit shown in Fig. 39A are designed to resonate at the same frequency, the cut-off will be quite sharp as shown by the curves in Fig. 39B. This will result in a highly discriminating filter and a very narrow band of frequencies will be allowed to pass.

However, if we make the two circuits resonate at slightly different frequencies, we can broaden the pass-band as shown by the curves in Fig. 39C. Here, all the frequencies between the resonant frequency of the series circuit, denoted by f_1, and the resonant frequency of the parallel circuit, de-

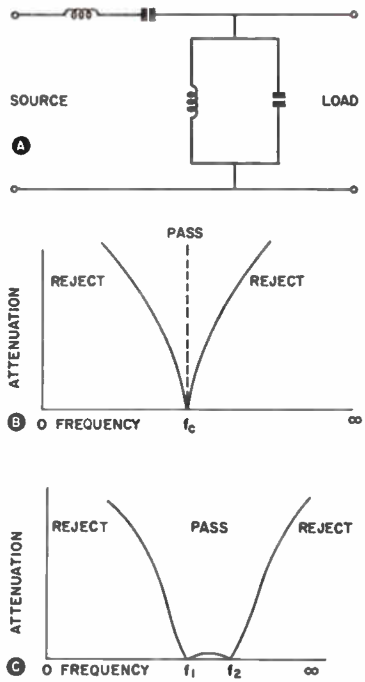


FIG. 39. L-type band-pass filter, A; sharp cut-off, narrow-band with arm and leg resonating at same frequency, B; sharp cut-off, wide-band with arm and leg resonating at different frequencies, C.

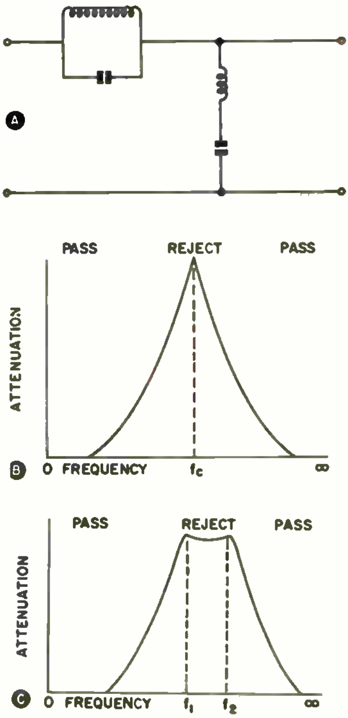


FIG. 40. L-type band-stop filter, A; sharp cut-off narrow band with arm and leg resonating at same frequency, B; sharp cut-off, wide band with arm and leg resonating at different frequencies, C.

noted by f_2, will be passed while all others will be rejected.

Similarly, a band-stop filter can be made by using the L networks shown in Fig. 40A. Notice that now the parallel resonant circuit is in series with the load and the series resonant circuit is in parallel with the load. When the parallel circuit resonates, it will tend to block current to the load. When the series circuit resonates, it will tend to short any signals away from the load. If the two circuits are designed to resonate at the same frequency, a narrow band of frequencies will be stopped as shown by Fig. 40B. If there is a difference in their resonant frequencies, the stop band will become wider as shown in Fig. 40C.

Although L pads can be designed to give sharp cut-off characteristics and provide almost any type of filtering action, they are unsymmetrical. In many filtering applications, it is very important to have the impedance matching in both directions. In such cases symmetrical filters must be used.

The low-pass L section shown in Fig. 37A can be made into a symmetrical *T section* by splitting the inductance L in two, and placing one half on each side of the shunt capacitor C. The same L section can be changed into a symmetrical *pi section* by using two capacitors of half the former capacity, and locating one at each end of the inductance.

In a similar manner, the band-pass and band-elimination L sections can also be made into symmetrical T or pi sections. A summary of the more common filter section forms, together with their attenuation characteristics, is given in Fig. 41.

MULTI-SECTION FILTERS

Now that we know that symmetrical filter sections "repeat" the load resistance for their input impedance, it is obvious that we can operate several filter sections in tandem. Thus, if we have three low-pass T sections, each designed for an impedance, let us say of 500 ohms, we can arrange them as shown in Fig. 42A. Here the second section acts as the load impedance for the first, and the input impedance of the last section is the load for the second filter section.

We may continue this arrangement for almost any number of sections. In all cases, however, the generator still works into a 500-ohm input impedance, and the impedance at the output continues to match the 500-ohm load.

Since the individual attenuations add, the tandem arrangement will attenuate the unwanted frequencies roughly three times as much as an individual section, while the wanted fre-

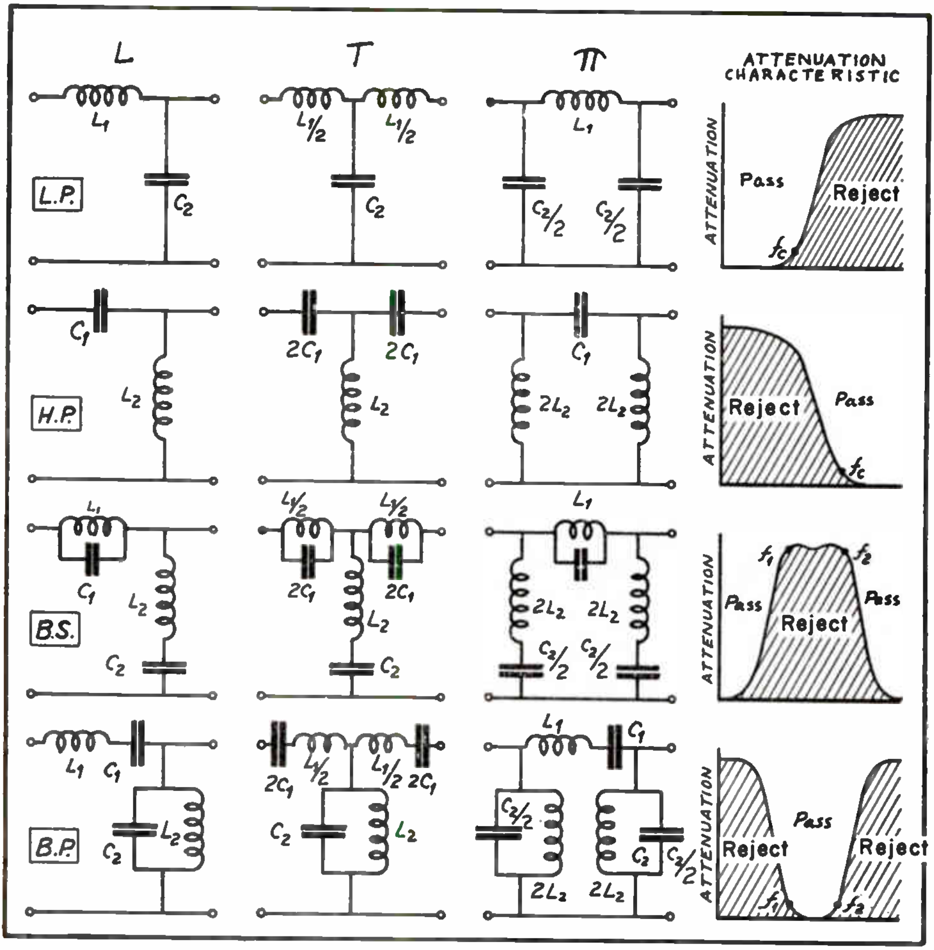


FIG. 41. The most common forms of L, T, and pi sections with their attenuation characteristics for low-pass, high-pass, band-stop, and band-pass filters.

quencies pass with little loss, and the overall response is improved. The two coils in series can be replaced by a coil of twice the inductance, as shown at B.

We may continue this arrangement for almost any number of sections. In all cases, however, the generator still works into a 500-ohm input impedance, and the impedance at the output continues to match the 500-ohm load.

In precisely the same manner, high-pass, band-pass, and band-stop multi-section filters can be made. Wherever two arm capacities appear in series, however, the effective capacity is cut in half. These two capacitors, therefore can be replaced by a single capacitor. Where two coils appear in series, they can be replaced with a single coil of twice the inductance.

In cases where it is necessary to present a very high attenuation to one band of frequencies and, at the same

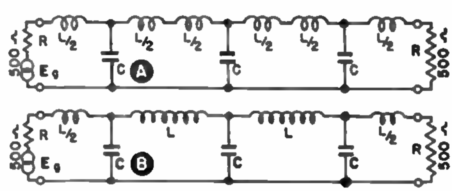


FIG. 42. Three low-pass T sections connected in tandem.

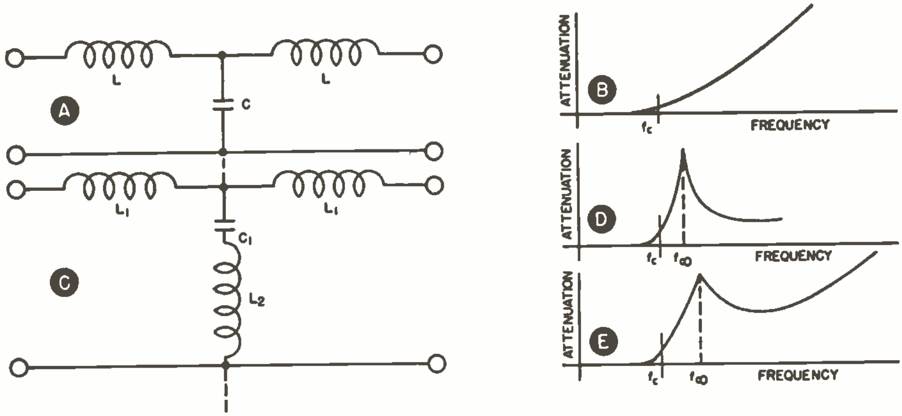


FIG. 43. A prototype section with high attenuation far from cut-off, and a derived section, with high attenuation close to cut-off, when put in tandem make a composite filter with a very sharp cut-off, and a high attenuation over the entire stop band.

time, pass an adjacent band of frequencies with little loss, we must use a filter with a sharper cut-off characteristic.

M-Derived Filters. Let us look at Fig. 43A. This is an ordinary low-pass T section, sometimes called a "constant K filter," with an attenuation characteristic like that shown in Fig. 43B. Since this is a basic filter section, it is often called a "prototype."

It is called a constant-K filter because of the fact that the product of X_L and X_C is constant at all frequencies. For example, at some frequency X_L may be 200 ohms, and X_C may be 50 ohms, and the product of the two is 10,000. At twice the frequency, X_L is 400 ohms, and X_C is 25 ohms and the product (the constant K) is still 10,000.

Let us modify the inductance and capacitance values of this prototype section so that an additional inductance L_2 can be placed in series with the capacitance C_1, as illustrated in Fig. 43C. If done properly, this does not change the section cut-off frequency, and it does not alter the response in the *pass band*.

In the *attenuation band*, however, we find that a startling change occurs. Since L_2 and C_1 form a series-resonant circuit, the section is "shorted out" at one frequency, and its response drops nearly to zero. This means that at one point in the stop band the attenuation is very high, reaching, theoretically, to infinity. In practical circuits, the attenuation does not reach infinity because of losses in the coil, but the response of such a section has a peak of great attenuation, as illustrated in Fig. 43D.

In choosing the value of L_1, C_1, and L_2 in the so-called derived filter, shown in Fig. 43C, a factor is used in special formulas to calculate their values. This factor, called the "M" factor, is determined by another formula based on how sharp the cut-off is to be. Since it is the sharpness of cut-off, or "M" factor, that determines the value of all L and C values in a derived filter, this arrangement is called an "M-derived" filter. Thus, an M-derived filter is one having infinite attenuation at some specific frequency, producing a sharper cut-off than a standard filter.

An M-derived section cannot be used alone, since after the infinite attenuation frequency, f_∞, is passed, the attenuation again drops to a low value. If, however, we connect a prototype and a derived section together in tandem, we realize a greatly improved filter performance. This means that we add the curve of Fig. 43B to that of Fig. 43D to get an over-all response like that shown in Fig. 43E.

Note that not only is the attenuation relatively high over the entire stop band, but also that the cut-off characteristic is much sharper. Notice also that the losses in the pass band near the cut-off frequency have been held to a minimum, which is impossible to realize by merely adding simple prototype sections together.

Fig. 43C does not represent the only practical type of derived filter section. The low-pass *pi-section* prototype in Fig. 44A, for instance, can be modified to appear as the derived section in Fig. 44B. The added capacitor C_2, in conjunction with coil L_1, now forms a parallel-tuned circuit, which at its resonant frequency serves to block the flow of current through the section. Here again, the result is an infinite-attenuation frequency in the stop band.

In constructing a multi-section filter, it is possible to use several derived sections. Furthermore, if different M values are used for each derived section, a separate infinite attenuation frequency for each section is obtained. In this way the over-all filter-response characteristic can have an extremely sharp cut-off, and very high attenuation at all points in the stop band. In

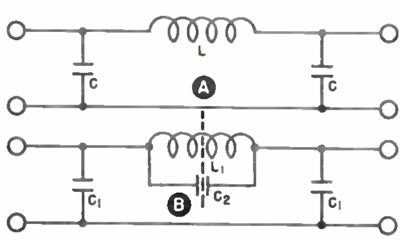


FIG. 44. A prototype low-pass pi section, and the corresponding M-derived section. Where the parallel-tuned circuit is resonant, the derived section has infinite attenuation.

general, filters that are made up of a prototype and one or more derived sections are called "composite" filters.

INPUT-OUTPUT IMPEDANCE

Earlier we stated that symmetrical filter sections like the T and π sections in Fig. 41, possess input impedances that are *nearly* equal to the respective load resistances. Let us investigate this statement further.

Any symmetrical filter section has a characteristic or "iterative" impedance that is determined entirely by the values of inductance and capacitance. *It is only when the load resistance is made equal to this inherent characteristic impedance that the input impedance of a filter assumes an identical value.* In other words, for proper filter performance, it is desirable that we use source and load impedances that match the impedance of the filter itself.

Unfortunately, the characteristic impedance of a simple prototype filter section is not constant with frequency. Look at the low-pass T section in Fig. 41 to see why this is so. In the input terminals, the left-hand inductance arm $L_1/2$, and the shunt capacity C_2, together form a series-resonant circuit. Since these two are in resonance at the section cut-off frequency, the input impedance drops to a very low value at this point. In Fig 45, the solid curve shows how the characteristic impedance of a low-pass T section decreases rapidly from a nominal value for low frequencies to a theoretical zero value at the cut-off point.

On the other hand, the low-pass *pi section* in Fig. 41 resembles a split-capacitor parallel-tuned circuit, and it behaves like a parallel resonant circuit. For low frequencies, the section has a nominal characteristic impedance. As the frequency is raised, the input impedance increases, and at the cut-off point where the elements are

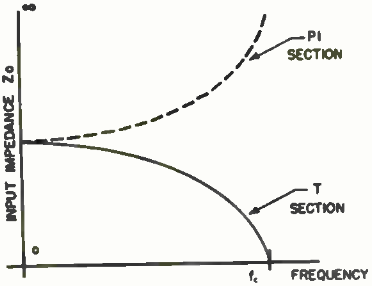


FIG. 45. The characteristic impedance of a prototype T section drops to zero at the cut-off frequency, and that of a pi section rises to a very high value.

resonant, the impedance approaches an infinite resistance. The dotted curve in Fig. 45 shows this typical impedance change for a pi section.

Although these two filter sections may have identical attenuation characteristics, their impedance variations are strikingly different. In general, it may be said that the impedance of any T section drops nearly to zero at a cut-off point, and that of any pi section rises to a very high value.

How can we match a source and load to a filter if the characteristic impedance of the filter varies so widely over the pass band?

In a great many filter applications where it is not necessary to work very close to the cut-off frequency, the change of filter impedance is not serious. In such cases, the impedance variation is ignored, and the source and load impedances are chosen to be equal to that of the filter at frequencies far removed from its cut-off frequency. This gives fair filter performance.

Under these conditions, however, the mismatch at cut-off substantially decreases the cut-off sharpness. For more accurate work, it is necessary to find a better method of matching the changing filter impedance. This *is done by a special type of T section to match a T section, or a special type of pi section to match a pi section,* called "half sections."

Terminating Half-Sections. Filters can be conveniently terminated by means of what are called half sections. Examples of half sections and how they perform are shown in Fig. 46. In A, we have shown the development of a half-section T. This filter is developed by dividing the T into two halves. Notice the value of the leg in the half sections is $2Z_2$, because the two impedances $2Z_2$ in parallel would have a value equal to Z_2, the impedance in the full T section.

Looking into terminals 1 and 2 of the left half of the T, we see the same basic arrangement as in the full T section. However, looking into the terminals 3 and 4 of the left side, the half section looks like the output of a π section filter.

Similarly, looking into terminals 1

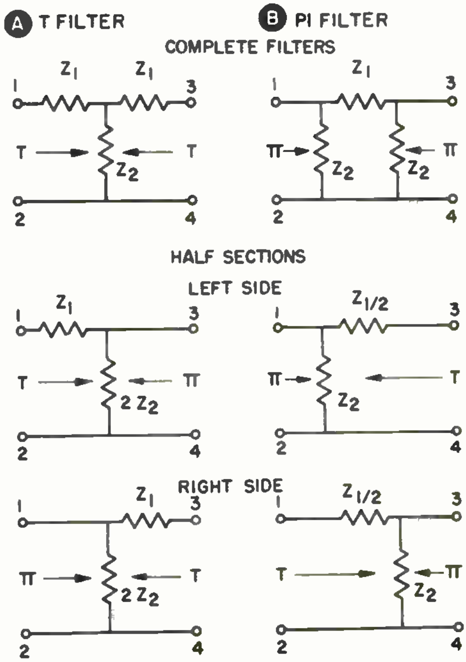


FIG. 46. How half-sections are developed from a T section (A) and a pi section (B).

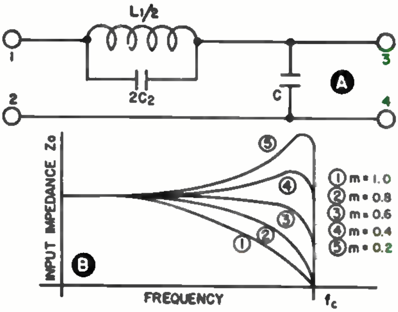


FIG. 47. A terminating half-section, and the variation of its characteristic impedance with frequency for different values of the multiplying factor M.

and 2 of the right half of the filter, we see the same basic configuration as we do at the input to a π section, and looking into terminals 3 and 4 we see it looks like a T filter.

The π filter can be broken down into two half sections as shown at B. Looking into terminals 1 and 2 of the left side, we see the circuit looks like the input to a π section, whereas looking into terminals 3 and 4 of this section we see a circuit similar to the output of a T filter. On the right side, looking into terminals 1 and 2, the circuit looks like the input to a T, and looking into 3 and 4, it looks like the output of a π filter.

Let us suppose that the M-derived pi section, shown in Fig. 44B, is split in half as indicated by the dotted line. If we consider only the right half, we have the half section shown in Fig. 47A.

Now looking into the terminals 3-4, this half section appears as a pi section. *Any full pi section, therefore, can be attached to these terminals without an impedance mismatch, since the two networks have the same impedance at all frequencies.*

Looking into the terminals 1-2, the half section resembles a T section, and the input impedance can be expected to drop to zero at the cut-off frequency. We find, nevertheless, that the manner in which this impedance drops to zero depends entirely upon the value of the multiplying factor M that we used to derive the total section in the first place.

If a value M = 1 is used, the half-section impedance varies like a prototype T section. This is shown by curve 1 in Fig. 47B. For lower values of M, the input impedance is more constant over the pass band. The input impedance variation for M = 0.8 and M = 0.6 is illustrated by curves 2 and 3, respectively. For still smaller values of M, let us say M = 0.4 and M = 0.2, the input impedance may rise to a high value before it drops abruptly to zero. See curves 4 and 5.

A value M = 0.6 gives the best performance, and results in a half-section input impedance that is very nearly constant over about 80% of the filter pass band.

We can use half sections like the one shown in Fig. 47A to match a constant-impedance generator to the variable impedance of a full pi section, or a number of pi sections in tandem. Furthermore, we can use an additional half section at the filter output to match the filter to a constant-impedance load.

A composite low-pass filter, made in this manner, is shown in Fig. 48A. The portion marked B is really the prototype pi section of Fig. 44A. The two terminating half sections, A and C, are made by splitting the derived pi section of Fig. 44B.

As two capacitors in parallel can be replaced by a single capacity, the composite filter can be simplified as shown in Fig. 48B.

Since two half sections have the same attenuation characteristics as a full-derived section, the over-all response of this filter looks like that shown in Fig. 43E. The half sections

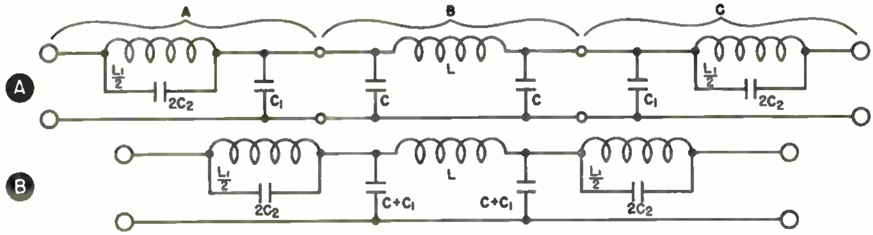


FIG. 48. A, a composite low-pass filter is made by using a prototype pi section and two constant-impedance terminating half-sections. B, the simplified filter, after combining parallel capacities.

have three uses: (a) they present a nearly constant impedance at each end of the filter; (b) they sharpen the cut-off response; (c) they supply an infinite attenuation at one point in the attenuation band.

In a similar manner, it is possible to construct half sections that accurately match a T-section filter. Thus, half sections, from the network shown in Fig. 43C, can be used to terminate a prototype T section like that shown in Fig. 43A. This gives the composite filter that is shown in Fig. 49A, which, in turn, can be simplified as shown in Fig. 49B. The half sections not only provide a good impedance match between source and filter, and between filter and load, but they also improve the general filtering action. The response curve for this composite filter is similar to Fig. 43E.

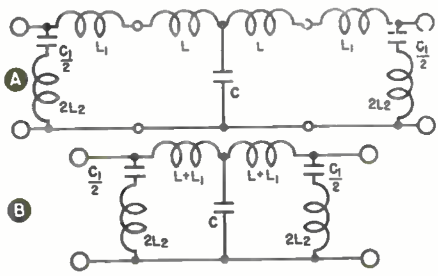


FIG. 49. The use of a T-section prototype with appropriate half sections to make a low-pass composite filter. The network is simplified by replacing two inductances in series, with a single inductance equal to their sum.

DESIGNING FILTERS

We have seen that it is the arrangement of components in a network that determines its general response characteristics. By properly arranging coils and capacitors we can make a network function as a low-pass, high-pass, band-pass, or band-stop filter. However, it is the values of the components that determine just where within the frequency range a filter will pass or reject signals. Therefore, to construct a filter we must be able to determine the values for components as well as their arrangement.

Before we can design a filter for some specific application, we must know which frequencies are to be passed, and which are to be rejected, and the amount of attenuation of unwanted frequencies necessary. In other words, we must know whether it is to be a low-pass, high-pass, band-pass, or band-stop filter. We must know the cut-off frequency of a low-pass or high-pass filter. For band-pass or band-stop filters, we need to know both the upper and lower frequency limits of the band to be passed or rejected. We should know how much attenuation of undesired frequencies is necessary, as it may be necessary to use a multi-section filter to obtain enough attenuation and a sufficiently sharp cut-off.

Let's assume, for example, that we want to construct a filter that will pass all frequencies between 0 and

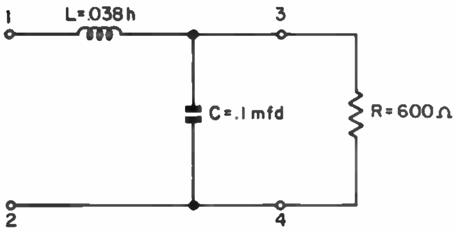


FIG. 50. Low-pass constant-K, L filter.

5000 cycles and attenuate all frequencies above 5000 cycles. Since we wish to pass low frequencies and attenuate high frequencies, we will want a low-pass filter. We know that we can obtain low-pass characteristics by using a series coil and a shunt capacitor in an L network as shown in Fig. 50. Next, we must determine the particular values of inductance and capacitance that will give us a cut-off frequency of 5000 cycles.

The L Filter. To do this, we apply our circuit values to special formulas that have been derived for filter network computation. For a low-pass filter the capacitance is found from the formula: $C = \frac{1}{\pi\ fc\ R}$; and, the inductance from the formula:

$$L = \frac{R}{\pi\ fc}.$$

In both of these formulas R is the impedance of a purely resistive load and fc is equal to the desired cut-off frequency in cycles per second. If we apply these two formulas to our circuit in Fig. 50, we find that the inductance equals:

$$L = \frac{R}{\pi\ fc} = \frac{600}{3.14 \times 5000}$$

$$= \frac{600}{15{,}700} = .038 \text{ henrys}$$

and, the capacitance equals:

$$C = \frac{1}{\pi\ fc\ R} = \frac{1}{3.14 \times 5000 \times 600}$$

$$= \frac{1}{15{,}700 \times 600} = \frac{1}{9{,}420{,}000}$$

$$= .0000001 \text{ farad or, .1 mfd.}$$

Thus, a coil of .038 henrys and a capacitor of .1 mfd arranged in an L network with a load resistance of 600 ohms gives us a low-pass filter with a cut-off frequency of 5 kc.

Converting an L Filter to a T or Pi Filter. We know from our previous discussion of filters that we can improve the cut-off curve for our filter and match impedances by converting the L network to a T or pi network. To do this, we simply use two coils of half the inductance value for the T filter as shown in Fig. 51A. Or, if we prefer the pi filter, we could use two capacitors of half the capacitance as shown in Fig. 51B. The choice of using either a T or a pi filter is simply a matter of convenience, as they give exactly the same performance. The

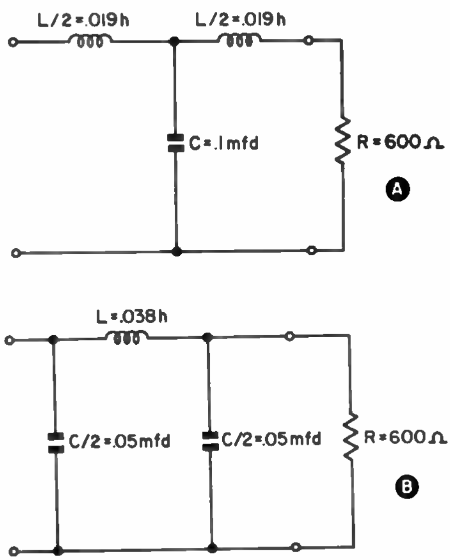


FIG. 51. Low-pass constant-K filter converted to T network, A; low-pass constant-K filter network converted to pi network, B.

choice is usually based on availability of components.

M-Derived Filter. The filter that we have just computed is a prototype, or constant-K filter. We know from the previous section on filter principles that we can get a sharper filter cut-off and obtain better impedance matching throughout the pass band by using an M-derived filter. An M-derived T section is formed by inserting an inductor in series with the leg capacitor of a prototype T section as shown in Fig. 52.

The addition of this new inductance causes the attenuation to rise sharply to a very high value as soon as the cut-off frequency is reached. This is due to the fact that the leg of the filter is designed to be series resonant at some frequency slightly above cut-off. When resonance occurs, nearly all the current is shorted around the load, giving us almost infinite attenuation.

In an ordinary constant-K filter, when the cut-off frequency is reached, the attenuation increases gradually to some maximum value as the frequency is increased above cut-off. In the M-derived filter, we actually make the attenuation curve above cut-off rise sharply by making the frequency of maximum attenuation occur nearer the cut-off frequency than it normally does. Therefore, the sharpness of the cut-off curve depends on the location of the frequency of maximum attenuation with respect to the cut-off frequency.

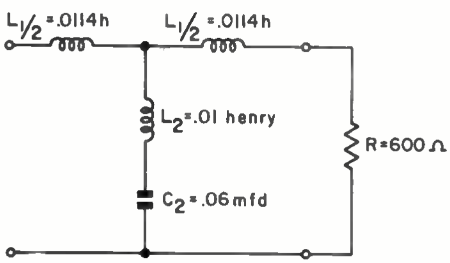


FIG. 52. Low-pass constant-K, L filter converted to M-derived T network.

The factor M from which the M-derived filter gets its name relates to the ratio of this frequency of maximum attenuation, $f\infty$, and the cut-off frequency, fc. For a low-pass filter, M is found from the following formula:

$$M = \sqrt{1 - \left(\frac{fc}{f\infty}\right)^2}$$

For a high-pass filter, M is found from the formula:

$$M = \sqrt{1 - \left(\frac{f\infty}{fc}\right)^2}$$

Converting a Constant-K Filter to an M-Derived Filter. To change an ordinary constant-K filter to an M-derived filter, we add the proper reactance component to the filter and change all the values by an amount depending on the value of M. To see how this works, let's convert the low-pass constant-K, T section, shown in Fig. 51A to the M-derived filter shown in Fig. 52.

The first thing we must do is select a value for the frequency of maximum attenuation, $f\infty$. If we make the frequency of $f\infty$ equal to 6250 cycles, we can then find the appropriate value of M from the formula:

$$M = \sqrt{1 - \left(\frac{fc}{f\infty}\right)^2} = \sqrt{1 - \left(\frac{5000}{6250}\right)^2}$$

$$= \sqrt{1 - .8^2} = \sqrt{1 - .64}$$

$$= \sqrt{.36} = .6$$

To find our new value L to use for our arm coils L_1, we simply multiply our original value of L by our value of M. This gives us:

$$L_1 = LM$$

$$= .038 \times 6$$

$$= .0228 \text{ henrys.}$$

We divide this new value L_1 by 2 to obtain the correct value for each of our arm coils. Thus our new arm inductances are .0114 henrys each.

Our capacitor C_2 is also determined by multiplying the original value of C by M. Or

$$C_2 = CM = .1 \times .6 = .06 \text{ mfd.}$$

The value of the inductance L_2 that we inserted in the leg of our network is found by applying our original value of L in the following formula:

$$L_2 = \frac{1 - M^2}{4M} \times L$$

$$= \frac{1 - .36}{2.4} \times .038$$

$$= \frac{.64}{2.4} \times .038 = \frac{.02432}{2.4}$$

or, approximately .01 henry.

If we prefer, we can use the M-derived pi section shown in Fig. 53. In this filter, the value of our capacitor C_2 used in the T section is divided by 2 to form the two legs of our M-derived pi section. We have a new capacitor C_1 in parallel with our inductance L_1 to form our filter arm. The value of L_1 is the sum of the two arm inductances $L_1/2$ used in the T section. The value of C_1 is found by substituting the original value of C that we used in our L filter in place of the value of L in the formula:

$$L_2 = \frac{1 - M^2}{4M} L \text{ or,}$$

$$C_1 = \frac{1 - M^2}{4M} C = \frac{.64}{2.4} \times .1$$

$$= \frac{.064}{2.4} = .026 \text{ mfd.}$$

These formulas, along with those used for high-pass, band-pass, and band-stop filters, are shown with their appropriate circuit diagrams in Fig. 54. Notice that in these circuits and formulas, the subscript 1 indicates an arm reactance and the subscript 2 is

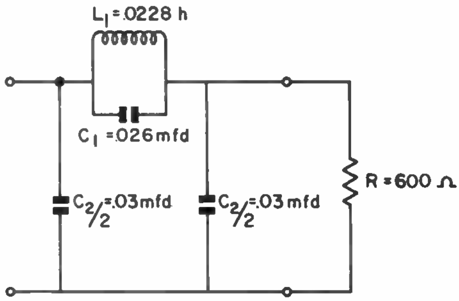


FIG. 53. Low-pass constant-K, L filter converted to M-derived pi network.

used for the leg components where necessary to avoid confusion between reactances. Although we have shown the formulas for M-derived, high- and low-pass filters, we have not shown them for band-pass and band-stop filters. The m-derived band-pass and band-stop filters become much too complex for anyone except filter design specialists.

In these formulas, the units for F, L, C, and R are in cycles-per-second, henrys, farads, and ohms, respectively.

Factors Affecting Design. The standard equations for constructing filters are based on lossless elements which are pure reactances. Thus, in actual practice, the response of a filter may be somewhat different from what was expected. How great this difference will be depends on the care that was exercised in choosing the parts and in constructing the filter.

A high-Q coil has very little loss, since it is mostly inductance and very little resistance. When a lower-Q coil is used, the losses are increased. Also, when a low-Q coil is used in a tuned circuit, the response of the circuit is broadened. It will no longer be as sensitive to frequencies at or near the resonant frequency.

There is always some capacitance between the adjacent turns of the coil winding, and this lowers the Q and increases the losses. Coils wound with wire insulated with a material having

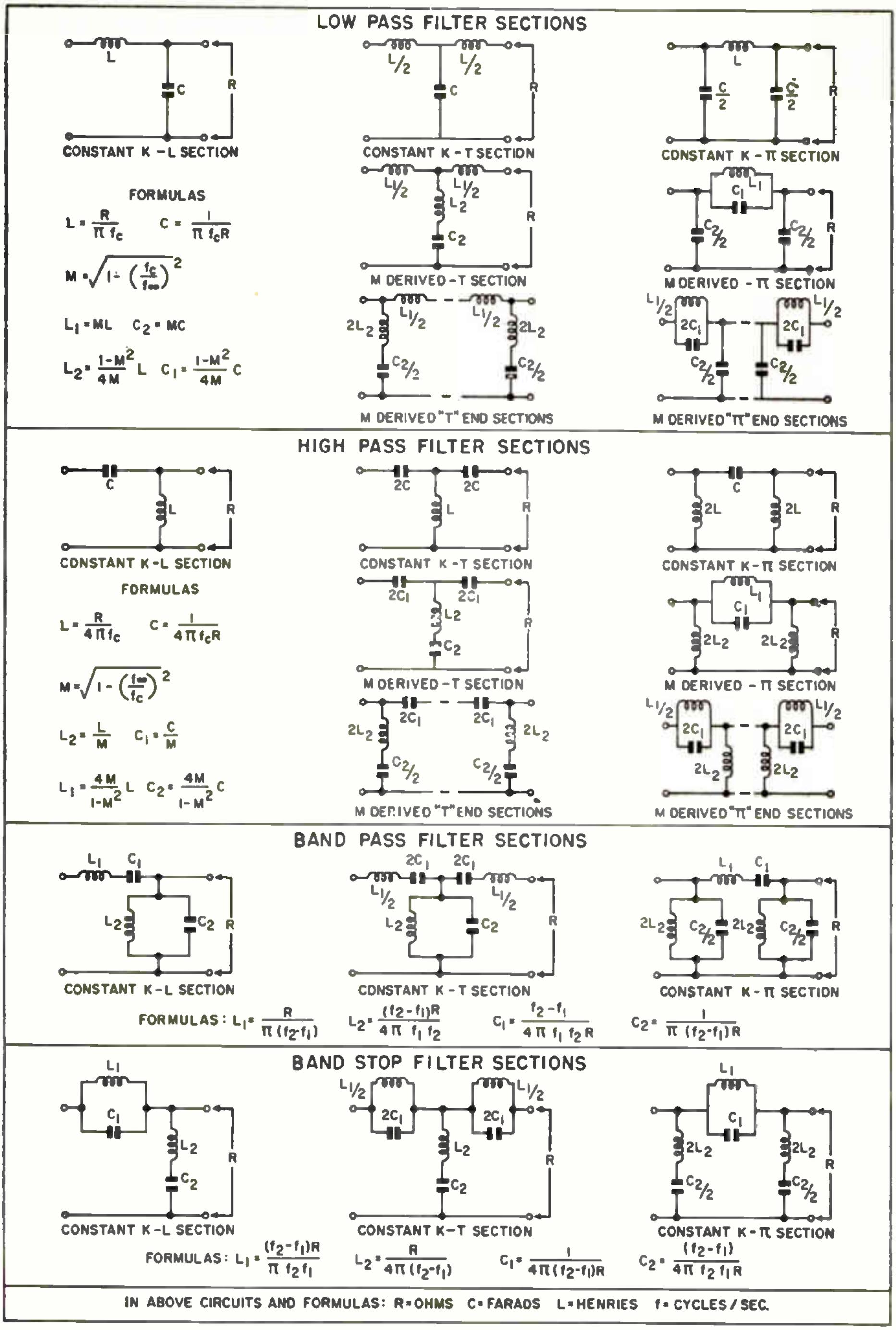


FIG. 54. Basic filter circuits and formulas.

low dielectric hysteresis loss are the best from this standpoint.

When two coils or two circuits are physically near each other, usually some coupling exists, although none may have been intended. The magnetic field about one coil or circuit induces a current in the other. Also the metallic parts of the two circuits form a small capacitance through which energy may be transferred. The Q of a coil is lowered by the loss of energy coupled out of it, and the response of the filter may be seriously altered.

This effect can be minimized by the proper placement of the parts (physical separation, mounting coils at right angles to each other, etc.) and by careful shielding. Shielding a coil lowers its inductance and also results in an energy loss because of the resistance of the shield, which increases the effective resistance of the coil. These effects lower the Q of the coil, but the reduction will be slight if the spacing between shield and coil is equal to at least the coil diameter at the coil ends and at least half this distance at the coil sides.

At power supply or audio frequencies, eddy current losses in iron-core coils may be reduced by constructing the cores of thin metal sheets called laminations. At higher frequencies, it is not practical to try to make laminations small enough or thin enough for use in coil cores. Coils used at these frequencies are air-core coils or have cores made from powdered iron mixed with a suitable binder The binder holds the particles together so that they can be shaped to the form desired and also insulates them from one another to prevent eddy currents. A type often used in communications equipment is the toroid coil. It is a single-layer or multi-layer coil wound on a doughnut-shaped powdered-iron core. An advantage of this type is that, so long as the core is unbroken, all of the lines of force are in the core and none outside. As a result, shielding is seldom necessary to prevent coupling between a toroid coil and another coil or circuit.

MEASURING FILTER RESPONSE

When checking filter response, great care must be exercised in setting up the equipment to avoid any stray coupling or leakage. All the units in the test setup must be bonded together and a good common ground provided.

Insertion Loss. The insertion loss of a filter is simply the loss in output, usually expressed in db, caused by inserting a filter in a circuit When the filter is to be tested before installa-

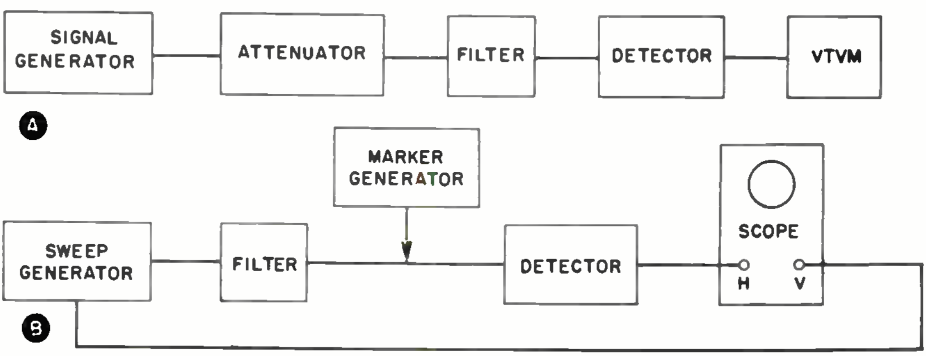


FIG. 55. Test setups for measuring (A) insertion loss, and (B) frequency response of a filter.

tion, a setup such as that in Fig. 55A can be used to measure this insertion loss. The signal source can be a signal generator, with or without modulation. For low frequencies, the output meter may be a db meter or a vtvm. For rf frequencies, a vtvm should be used. In some cases, the needed output is most conveniently obtained by feeding the filter output to a good receiver.

Using the lowest signal level that provides a usable indication with the filter in the circuit, the output is noted. Without changing the amplitude of the input signal, the filter is removed from the circuit, the circuit is reconnected, and the output is noted. The difference in the two readings is the insertion loss of the filter.

The attenuator at the input of the filter is provided to reduce the input signal to the lowest amplitude that will give a usable indication on the meter with the filter in the circuit. Unless the input signal is kept as low as possible, there may be so much leakage across the filter that the results obtained are inaccurate.

It is important for the shielding for the test setup to be as good as it can be made, particularly when dealing with rf frequencies. With rf frequencies, it is difficult to obtain accurate results at best, and without adequate shielding, it becomes impossible.

In most cases, the most satisfactory check of the insertion loss of a filter is simply to try the filter in the circuit and see if it does the job for which it was designed. Ordinarily, you'll know how much attenuation of unwanted frequencies is necessary and how much attenuation of wanted frequencies can be tolerated. If the circuit works satisfactorily after the filter is installed, you won't usually need to know or care exactly what the insertion loss measures.

Filter Response. The response curve of a filter can be checked either by a point-to-point method or by using a sweep signal generator and an oscilloscope to obtain a visual indication of the response curve. In the first method, the signal generator is connected to the input of the filter, and its frequency set to some point within the expected pass band. An output meter is connected to the output circuit and the signal generator set to give a convenient indication. Then the signal generator is set to the lowest frequency at which there is any response, and the frequency is advanced in equal steps, and the output at each frequency recorded. These readings are plotted on graph paper in db, the frequency being plotted horizontally and the output vertically.

A more convenient method of checking response is shown in Fig. 55B. Here the response curve can be observed and the effects of any adjustments noted. A marker generator set to the proper frequency and loosely coupled to the filter can be used to inject a pip at the cut-off frequency or any other point of interest within the band.

Lesson Questions No. 29

Be sure to number your Answer Sheet 29CC.

Place your Student Number on *every* Answer Sheet.

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

1. What are two reasons for using attenuators?

2. Use the table in Fig. 9 to compute the values of R_A (the arm resistors) and of R_L (the leg resistor) in a T-type resistive pad in which the output and input are both 60 ohms, and attenuation is 8 db.

3. Complete the following: A minimum loss L pad is used to (a) match unequal impedances; (b) provide attenuation and maintain an impedance match between equal impedances; (c) filter out undesirable harmonics.

4. What requirement must a network have before it can be considered a filter?

5. Name the four basic types of filter circuits.

6. Name the filter circuit you would use if you wanted to pass only the band of frequencies between 3,000 cycles and 10,000 cycles.

7. If we connect three low-pass T sections in tandem, what effect will this have on the attenuation of the unwanted frequencies as compared with a single low-pass, T-section filter?

8. What information must you have before you can design a filter circuit?

9. Compute the values of L and C in a simple, constant-K, L-section, high-pass filter having a characteristic impedance of 100 ohms and a cut-off frequency of 100 cycles.

10. Complete the following: A long line, such as a telephone line (a) attenuates the low frequencies more than the highs; (b) attenuates the high frequencies more than the lows; (c) attenuates the highs and lows about equally.

"WISHERS" AND "DOERS"

How often have you said, "I wish I had more money?" Thousands of times, possibly. But do you realize that if you are living in a town of, let us say, 5,000 inhabitants, there are exactly 4,999 others in your town who are saying exactly the same thing?

And yet, of these 5,000 "wishers," only about 100 are going to do something about it. The others are going to continue being "wishers."

Now, any man who shows enough "get-up-and-go" spirit to undertake this course proves that he is not a mere "wisher." When you enrolled, you showed that you wanted to be a "go-getter." Your job now is to keep going forward on the road you have mapped out for yourself.

Every lesson in this course, every job you work hard to get, is a step along this road. So don't let yourself wish that the lessons were easier, or that you could become successful without studying, or that jobs would come looking for you. Stay out of the class of the "wisher," and stay in the class of the "doer."

J. M. Smith

nri

ACHIEVEMENT THROUGH ELECTRONICS

nri

OPERATIONAL AMPLIFIERS

K309

NATIONAL RADIO INSTITUTE • WASHINGTON, D. C.

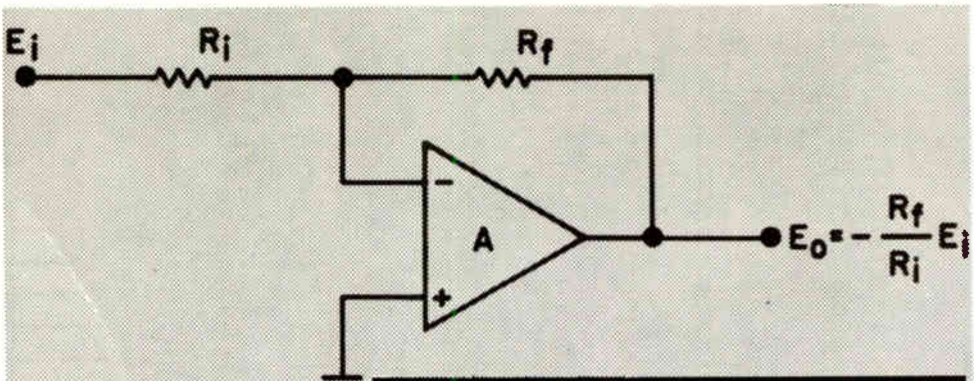
OPERATIONAL AMPLIFIERS

K309

STUDY SCHEDULE

CEC5C675 1975 EDITION Litho in U.S.A.

3-16-76



OPERATIONAL AMPLIFIERS

This lesson is your detailed study of one of the most important types of amplifiers used in electronics today – the operational amplifier. Its application is almost unlimited because of its great versatility. This type of amplifier is also the key element in all analog computers. The operational amplifier, in its numerous configurations, can help perform many mathematical operations.

Since you previously had a brief introduction to the operational amplifier, you should already have a basic idea of what it is and how it is used. However, this lesson will expand your basic knowledge of this important circuit.

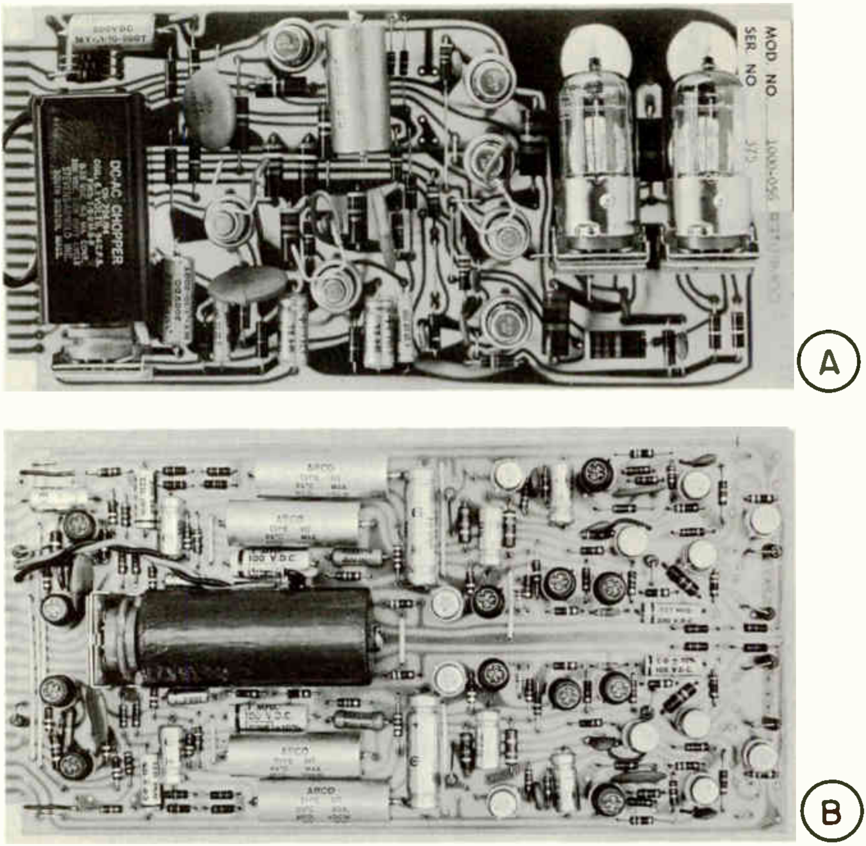
Like any other amplifier, an operational amplifier has an input and an output. Amplifiers generally increase the level of the input signal and present it at a higher level at the output; therefore the operational amplifier circuit has gain and amplifies small signals that are applied to its input.

An operational amplifier is basically a linear amplifier which accurately reproduces the shape of the input signal. However, it possesses special characteristics which set it apart from more conventional amplifiers and permit it to be a much more versatile circuit.

The term "operational" came about because of the early use of an amplifier to perform mathematical operations. Today these amplifiers still perform their original operations, but they have also become adapted to be used for many other functions.

Operational amplifiers, commonly called op amps, come in many different forms. For example, the very earliest types of op amps were made with vacuum tubes. A typical vacuum tube op amp is shown in Fig. 1A, where you can see the vacuum tubes, resistors and capacitors mounted on a printed circuit board. The seven circular metallic devices are miniature vacuum tubes. The large rectangular object is a mechanical chopper. While you may still encounter tube type op amps in some older equipment, the majority of operational amplifiers in use today are transistorized units. A transistorized operational amplifier is shown in Fig. 1B. The large cylinder is a chopper. Again, all of the components are mounted on a printed circuit board.

Today integrated circuit operational



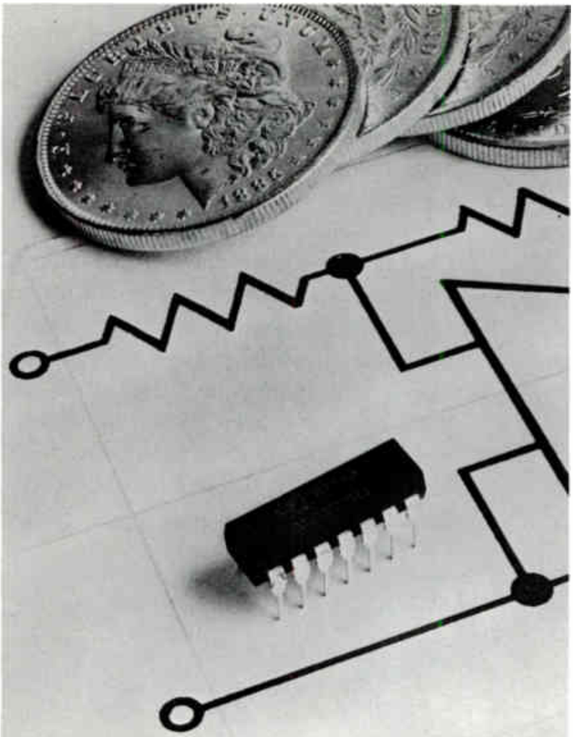
Courtesy Hybrid Systems, Inc.

Fig. 1. A tube type operational amplifier (A) and a transistorized operational amplifier (B).

amplifiers are rapidly replacing the discrete component types in many applications. An integrated circuit operational amplifier is an entire miniature transistor amplifier constructed on a single chip of silicon, in much the same way a transistor is made, and housed in a small multi-lead package like that shown in Fig. 2. Notice the extremely small size of this amplifier. Such an amplifier is as sophisticated as and generally superior to the tube and transistor operational amplifiers of Fig. 1. Later in this lesson we will take a look at some of the circuit details of transistor operational amplifiers. The emphasis in this lesson will be on the transistor type circuit; these are the circuits you will most encounter in your work, in both discrete and miniature integrated circuit form.

An amplifier has to have certain distinct characteristics before it can be called an op amp. The extent to which it meets these requirements indicates how good an operational amplifier it is. Even though the perfect operational amplifier cannot be made, its value is based on how closely it can approach the characteristics of the ideal amplifier.

One characteristic of an ideal oper-



Courtesy Texas Instruments

Fig. 2. Integrated circuit operational amplifier in a dual in-line package.

ational amplifier is that it must be a dc amplifier, that is, it must amplify slowly varying signals or dc levels. In order to amplify or pass very low frequency ac signals or dc levels, direct coupled circuitry is usually used.

The ideal op amp should also have a zero output voltage if the input voltage is zero. Many amplifiers do not have this characteristic. With a zero voltage input signal applied, many amplifiers have a fixed dc output voltage. When the input signal is applied the output varies above and below this fixed dc output level. Although this is not detrimental to circuit performance, it is more convenient for an op amp to have the output voltage zero when the input is zero. If only ac signals are to be used, then a capacitor can remove this output level. However, remember that we want our op amp to have the capability of measuring, or amplifying, dc signals.

Thirdly, the ideal op amp must have very high input impedance. When connected to a source of signal voltage, its impedance is so high that it will not draw any current from the source; its input impedance is infinite.

Ideally the op amp must also have low drift. In other words, it must be a perfect dc amplifier whose output does not vary with changes in circuitry or surroundings.

The ideal op amp will have a zero output impedance. This means that if any load resistance is connected to it, the output will not be loaded down since there is no internal circuit resistance present. With a zero output impedance, the op amp is a perfect voltage source with an output voltage that remains constant regardless of the load connected to it.

Another ideal factor is minimum power consumption. A good op amp will not draw a lot of current from its power supplies during operation.

*

Optimally, the amplifier should be a differential amplifier. Op amps can be constructed with a single-ended rather than a differential input. However, most operational amplifiers you encounter will be differential amplifiers.

In order to function at its best, the op amp must have very high voltage gain. A minimum figure is 1000 (most have gains higher than 10,000).

As noted before, the perfect op amp does not exist. Even so, actual operational drawbacks are hard to find. Just what degree of departure from the ideal that can be tolerated depends strictly upon the amp's specific application. Each requirement for an op amp will have to be investigated to determine those characteristics needed or desired. For example, a particular requirement may call for an op amp with input impedance greater than one megohm, voltage gain greater than 25,000 and output impedance less than 200 ohms. Even though these characteristics are less than ideal, they accurately satisfy the requirements of the particular application.

3-16-76

Fundamentals

An op amp is nothing more than a combination of some of the individual amplifier circuits you studied before. For example, a typical op amp may consist of a differential amplifier direct coupled to an emitter-follower which drives a complementary symmetry output circuit. All of these amplifiers and fundamentals you learned in previous lessons. Here we will collect this information and show you how an op amp is formed.

To begin, however, we will emphasize only the fundamentals of the op amp. Once you understand the overall concept, you can begin to study the detailed circuit operation.

In most of this lesson you will see the op amp designated by the symbol in Fig. 3. The letter A inside the triangle designates the gain of the amplifier.

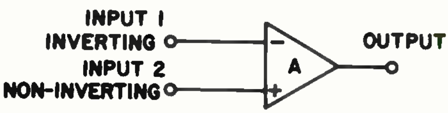


Fig. 3. Basic operational amplifier symbol.

This op amp has two inputs; therefore, it is a differential amplifier. The − and + signs on the inputs designate the inverting and non-inverting input lines. Input 1 is designated with a negative sign, indicating that it is the inverting input. If a signal is applied to input 1, the output will be 180° out-of-phase with the input. When input 1 is used, the op amp is an inverter. Input 2 is designated by a positive sign; it is the non-inverting input. Any input signal applied here will appear in the same phase in the output.

If we apply an input signal to the amplifier, an output signal will be generated. The input signal will be amplified by an amount equal to gain A of the amplifier. If the input signal is very small, it will appear greatly enlarged at the output because of the high gain.

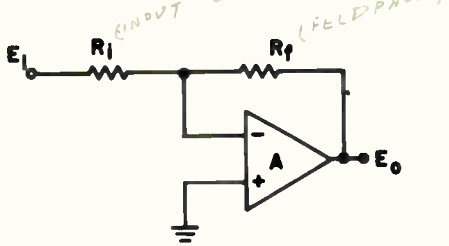
If the input signal to the op amp is too high, the amplifier output will be clipped. For example, if we apply a 1 volt input signal to an op amp with a gain of 50,000, our output should theoretically be equal to the input multiplied by the gain, or 50,000 volts.

This, of course, cannot actually happen. The output of any amplifier is limited by the power supply voltages. Most op amps are powered with both positive and negative voltages. A typical transistor op amp, for example, uses power supplies of +15 and −15 volts. For this reason the output cannot swing any greater than ±15 volts (for a sine wave signal, no more than 30 volts peak-to-peak). Should the input voltage be too high and try to cause an output voltage beyond the power supply capability of the amp, the output signal will be clipped on both positive and negative peaks. This is due to the saturation of the output transistors.

Since its gain is so high, the op amp can be used satisfactorily with only very low level signals. The gain is also unstable, varying greatly from one amplifier unit to the other of the same type. One operational amplifier may have a gain of 25,000 while an identical amplifier may have a gain of 35,000. While the circuits of the two amplifiers may be identical, various characteristics of the components could cause large gain differences. For many applications this is a disadvantage.

To overcome this problem we generally use negative feedback with op amps; some of the output voltage from the amplifier is fed back to the inverting input through a resistor. As you learned in the previous lesson, negative feedback reduces the overall gain of an amplifier and at the same time stabilizes it. It also improves frequency response

The most common op amp circuit is shown in Fig. 4. Here the non-inverting input to the amplifier is grounded. This eliminates the differential characteristics of the amplifier so we can use it as a single-ended circuit.



$$A = -\frac{R_f}{R_I} \qquad E_o = -\frac{R_f}{R_I}(E_I)$$

Fig. 4. Standard operational amplifier circuit.

Negative feedback is provided by the feedback resistor R_f connected between the output and the inverting input. An input resistor designated as R_i is also connected to the inverting input. The free end of this resistor is the input to the circuit. With this arrangement the gain of the circuit is determined strictly by the ratio of the feedback to input resistances: gain = R_f/R_i.

Negative feedback makes the gain of the circuit completely predictable. To control or set the gain of this amplifier, we simply select appropriate values of feedback and input resistors. For example, if the feedback resistor R_f is equal to 100K-ohms and resistor R_i is equal to 10K-ohms, the gain of the circuit is 100K/10K = 10. Keep this fact in mind.

Another important characteristic of this circuit is that it is an inverter. Since we are using the inverting input of the amplifier, the output signal will be 180° out-of-phase with the input signal. When expressing the gain, a negative sign is indicative of this inversion.

In Fig. 5 we show all the currents and voltages in a standard op amp circuit. The input resistor is designated R_i while the feedback resistor is designated R_f. Current flowing in the input resistor is designated as I_i; current flowing in the feedback is I_f. I_b is the current flowing into the input. It is generally the base current of a transistor. The amplifier input voltage, or that voltage appearing directly at the input of the amplifier, we label E_b. The input voltage to the entire operational amplifier circuit is designated as E_i. The output voltage is E_o.

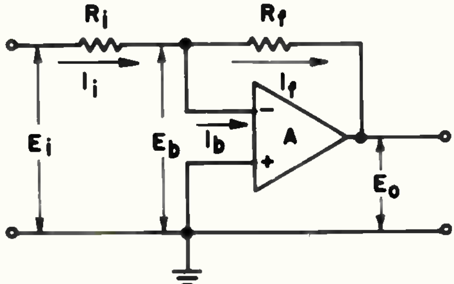


Fig. 5. Currents and voltages in an op amp.

Using these currents and voltages and some simple Ohm's Law relations, let's again see how the gain formula for this circuit is obtained. First of all, we know that the output is equal to the input voltage multiplied by the gain. The output voltage E_o, then, is equal to E_b times $-A$. To find the value of E_b, we algebraically rearrange the formula to produce $E_b = -E_o/A$. Remember that the negative sign indicates inversion.

From earlier lessons you found that the voltage across any component is equal to the difference in the two voltages at the ends, with respect to ground. For example, if we measure 10 volts at one end of a resistor and 7 volts at the other, the voltage across that resistor is 3.

With this in mind, we can write a formula for the voltage across the input and feedback resistors. The voltage across the input resistance R_i is equal to $E_i - E_b$. The voltage across the feedback resistor R_f is equal to $E_b - E_o$. Knowing the voltages across the resistance, we can write a simple Ohm's Law expression for the current through each resistance. For example, the current I_i through the input resistor is equal to the voltage across it $(E_i - E_b)$ divided by R_i. The current I_f through the feedback resistor is equal to the voltage across it, $(E_b - E_o)$ divided by the resistor R_f.

Refer to Fig. 5, where the current flowing into the amplifier is I_b. If we assume that our op amp is perfect and has an infinite input impedance, we can conclude that it will not draw current from the input source; therefore, $I_b = 0$. Because of this the input current I_i also flows in the feedback resistor R_f. This means that the input and feedback currents are equal. We can write a simple expression for their equivalence:

$$I_i = I_f$$

or

$$\frac{E_i - E_b}{R_i} = \frac{E_b - E_o}{R_f}$$

Now, let's make another assumption based on the characteristics of an ideal op amp. Earlier we said that the input voltage E_b is equal to the output voltage E_o divided by the gain $-A$. The gain is so high that the input voltage E_b is extremely small. It is so small, in fact, that for all practical purposes we can call it zero. If we assume a zero value for E_b, we can remove E_b figures from the formula. Thus simplified it reads:

$$E_i/R_i = -E_o/R_f$$

From this we can find that the output voltage is equal to the gain of the circuit, the ratio of the feedback to input resistances, multiplied by the input to the circuit:

$$E_o = -\frac{R_f}{R_i}(E_i)$$

The negative sign again indicates inversion. As you can see, we have used only algebra and Ohm's Law to determine the gain of the op amp circuit.

Keep in mind several things. We assumed that the input current to the amplifier was zero. As a result, we neglected it. We also assumed that the gain of the amplifier was infinitely high. Even though we had to use these assumptions to approach the ideal, the formula given will be quite satisfactory for an op amp with a gain of 1000 or more.

There is one very important characteristic of the op amp in its standard configuration. As mentioned earlier, the input impedance to a good op amp is very high. This, of course, is true in the circuit of Fig. 5. However, the input impedance of the circuit is not the same as the input impedance of the amplifier. The feedback and input resistor connections determine what the input impedance will be.

The input voltage to the amplifier (E_b) is near 0 volts, or ground. Since E_b gives the same effect as a true ground, we can more precisely call it a virtual ground. This is the case of the junction of the feedback and input resistors at the inverting input of the amplifier in Fig. 5. This point is called the summary junction because it is the input and output of all the currents in the circuit.

Since we can consider the right hand of resistor R_i as being connected to ground, the input impedance of the amplifier circuit is equal to the value of resistor R_i. If R_i is 100K-ohms, then the input impedance of the circuit is also 100K-ohms. This is a usually desirable characteristic of this circuit because the exact input impedance to the circuit is known. In the case of the input impedance of the amplifier, we do not always know its exact value. Of course, if it is high enough it will not affect the circuit. Therefore, we need not always be concerned with it. However, the input resistance produced by resistor R_i often is not high enough to suit the application. In that case we have to make a compromise in the gain of the circuit by adjusting the values of R_f and R_i to give an appropriate input resistance.

Now let's see how a typical op amp circuit works. Fig. 6A shows an op amp with a 60K feedback resistor and a 20K input resistor. From this circuit you know that the input impedance is equal to 20K, the value of the input resistor. We can figure the gain of the circuit by finding the ratio of the feedback to the input resistance (60K/20K = 3). The circuit will amplify an input signal by a factor of three.

Fig. 6B is a sine wave input signal of 8 volts peak-to-peak. If we apply the peak signal voltage to the input of the amplifier circuit, the output shown in Fig. 6B

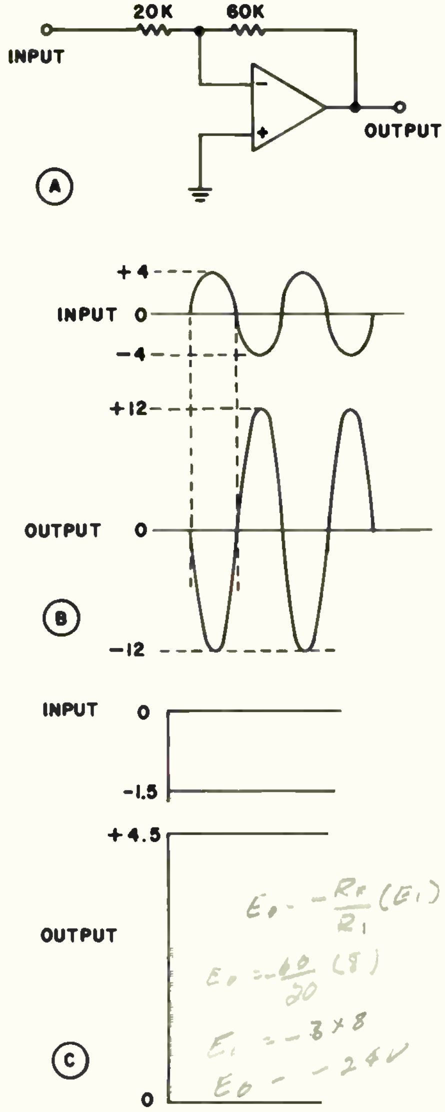


Fig. 6. A conventional op amp circuit (A), typical ac input-output waveforms (B) and dc input-output levels (C).

will be produced. Notice that the amplitude of the output is 3 times that of the input. The peak output voltage is 12 volts or 24 volts peak-to-peak. Notice also that the input and output signals are 180° out-of-phase, designating the inverting characteristic of the op amp circuit.

The op amp can also amplify dc signals. Fig. 6C shows a negative 1.5 volt input applied to the amplifier circuit. This is amplified and inverted to produce an output of +4.5 volts. If the input signal is removed from the circuit and the input line grounded, then the output of the amplifier should be 0 volts. Because of unbalanced amplifier circuits, the output may be only near 0 volts. In most op amps it is necessary to reduce the output to exactly zero; some form of compensating voltage must be applied to the circuit to bring the output voltage to zero.

A popular variation of the circuit shown in Fig. 6 is one where both the input and feedback resistors are equal. Since these resistors are equal, the gain of the circuit is equal to 1. We say that the amplifier has unity gain. An op amp connected this way is called a unity gain inverting amplifier. Since it has no gain you might think that it has no application. However, it is widely used where we want to maintain the amplitude level of a signal and at the same time obtain 180° phase inversion. The circuit acts as a good isolation amplifier with a relatively high input impedance and low output impedance.

In many ways this circuit also acts like the emitter-follower circuit you studied in a previous lesson. The gain is 1, but in this case the amplifier produces inversion which the emitter-follower does not. The low output impedance gives the circuit the ability to provide power amplification.

It is important to note that we can obtain gains less than or greater than 1 because we set the amplifier gain strictly by the ratio of the feedback to input resistors. If we make the input resistance greater than the feedback resistance, the gain will be a fraction. If we put a 2 volt peak-to-peak signal into an amplifier with a gain of .5, the output will be half this, or 1 volt peak-to-peak. The amplifier, with its gain of less than 1, is actually producing a loss. Even though this circuit appears to be ineffective, it is often used in analog computers.

SELF-TEST QUESTIONS

(a) True or false: An op amp is usually a differential amplifier.

(b) True or false: The gain of an op amp circuit is equal to the ratio of the input to feedback resistors. RF : RIN

(c) True or false: An op amp can amplify both ac and dc signals.

(d) True or false: The output signal of an op amp is inverted, or 180° out-of-phase with the input.

(e) True or false: An op amp cannot be connected to provide a gain less than one.

(f) What is the gain of an op amp whose input resistor is 12K and feedback resistor is 84K? 7

(g) An op amp circuit, like that in Fig. 4, has a gain of 4. The input signal is a dc voltage of −3.5. What is the amplitude and polarity of the output?

(h) Name five important characteristics of an op amp.

-16-

1: HIGH OPEN LOOP GAIN
2: HIGH INPUT IMPEDANCE
3: LOW OUTPUT IMPEDANCE
4: WIDE FREQ. RESPONSE
5. DIRECT COUPLING
6: LOW DRIFT
7: LOW POWER CONSUMPTION
8: DIFFERENTIAL INPUTS.

E0 = -GAIN (E1)
E0 = -4(-3.5
E0 = 14

Applications

Early use of op amps was limited primarily to analog computers. Vacuum tube op amps were large, expensive and complex. In order to obtain the accuracies necessary in an analog computer, op amps had to be quite sophisticated. But because of their high cost and complexity, they were not practical for more common electronic applications. However, as transistor op amps were developed, many of the problems associated with the vacuum tube types were eliminated.

For that reason, the op amp has become a practical circuit element. Today it is a common building block, not only in analog computers, but also in many other electronic devices and systems. Design engineers now use op amps as a matter of course in equipment they design.

The op amp is quite versatile in electronic circuits. By varying the types of components in the input and feedback circuits and their particular connections, an extremely wide variety of different operations can be performed. In fact, the op amp is practically unlimited in its application.

In your work you will discover op amps in several major connections. You have already seen the most common op amp connection, where the amplifier provides gain and phase inversion. Now let's examine some other typical op amp connections.

NON-INVERTING OP AMP CIRCUIT

There are times when the phase inversion of the common op amp is not desirable. In such applications we might want a high input impedance, a low output impedance, a particular gain, and no inversion of the signal. This operation can be accomplished by providing the gain in an op amp circuit and compensating for the phase inversion with a unity gain inverter.

Fig. 7 shows how this is done. Suppose we want a gain of 3 with no phase inversion. Amplifier 1 provides the gain of 3. Since this circuit produces phase inversion, we follow it with a unity gain inverter stage to remove the inversion produced by the first amplifier. While this circuit is used in some situations, it wastes parts and power. There is a simpler way to eliminate phase inversion.

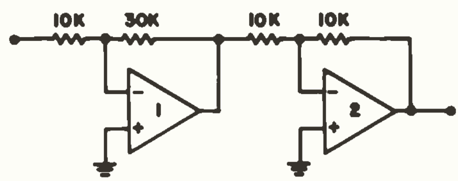
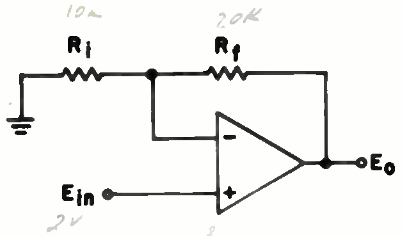


Fig. 7. Using a unity gain inverter to correct for inversion.

In Fig. 8 we show a non-inverting op amp circuit. In this circuit the non-inverting (+) input is not grounded. Instead, we apply the input voltage to this terminal. Input and feedback resistors are connected to the inverting input, but the end of input resistor R_i (normally connected to a signal) is connected to ground. This connection provides a gain that can be adjusted by setting the values of resistors R_f and R_i. The output voltage of this circuit is equal



$$E_o = (1 + \frac{R_f}{R_i}) E_{in}$$

Fig. 8. A non-inverting op amp circuit.

to the input voltage multiplied by $(1 + R_f/R_i)$. As you can see the gain is again a function of the ratio of the feedback to input resistances. But because of the particular connections, a value of 1 must be added.

Let's take a typical example and calculate the gain of this non-inverting circuit. Assume feedback resistor R_f is equal to 20K and input resistor R_i is 10K. If we have an input voltage of 2 volts, what will the output voltage be with the gain obtained?

First let's calculate the gain of the circuit: $(1 + R_f/R_i) = (1 + 20K/10K) = (1 + 2) = 3$. If the input voltage is 2 volts, then the output voltage will be 2 times 3, or 6 volts. Keep in mind that this is a non-inverting circuit. If the input is a sine wave, then the output will be in phase with the input. If the input is a dc voltage, the input and output polarities will be the same.

One of the biggest advantages of this circuit is the high input impedance. With the standard op amp configuration, the input impedance is equal to the value of the input resistor. For some applications this may be very low, causing undesired loading on the driving source. However, in this circuit the input signal is applied directly to the non-inverting input. This input is usually a very high impedance (an infinite impedance for an ideal op amp). Therefore, very little loading of the driving source will occur.

OPERATIONAL AMPLIFIER FOLLOWER

Take a close look at the gain formula for the non-inverting amplifier circuit in Fig. 8. The gain is equal to 1 plus the ratio of the feedback to the input resistances. What would happen if we were to increase the value of resistor R_i in this formula? The ratio of R_f to R_i would decrease. If we made R_i infinite, or disconnected it from the circuit altogether, the gain of this amplifier circuit would be 1. When this condition occurs, we have what is known as an op amp follower.

You already know about cathode and emitter-follower circuits. These are non-inverting buffer circuits whose output is very nearly equal to the input. The non-inverting op amp circuit is easily converted into a follower type circuit. The input resistor is removed and 100% feedback from output to input through resistor R_f is provided to the inverting input terminal. For this type of circuit the feedback resistance can be eliminated and the output connected directly to the inverting input.

The typical follower circuit is shown in Fig. 9. In this circuit the output voltage is equal to the input voltage. The input impedance is extremely high; the output impedance is very low. While the circuit

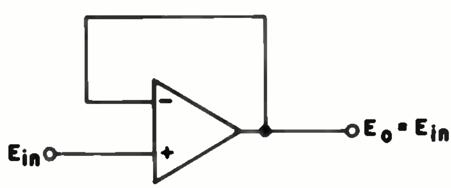


Fig. 9. An op amp follower.

25K

provides no gain or phase inversion, it is useful as an isolation amplifier and a power amplifier.

A particular voltage source may have a high output impedance and, therefore, will be loaded substantially when connected to a load. This may be an undesirable situation, but it can be eliminated with the follower. The high impedance source is connected to the very high input impedance of the op amp follower. Because of this high input impedance, it will not substantially load the voltage source. The output voltage of the op amp is equal to the input. However, its output impedance is very low and can supply substantial current to a load.

A DIFFERENTIAL AMPLIFIER CONNECTION

Since most op amps are differential amplifiers, they have both inverting and non-inverting inputs. In the circuits that we have discussed, however, we have not used this differential input capability. In the standard op amp connection, the non-inverting input is connected to ground. The differential op amp amplifies the voltage difference between the two inputs. If one of these inputs is connected to ground, the difference will simply be the amount of voltage applied to the other input. This forms what is known as a single-ended amplifier input. This means that both the input and output signals are measured with respect to some common ground reference.

There are occasions when it is desirable to use the differential capabilities of the op amp. At the same time, we want to take advantage of the fact that negative feedback around the amplifier can stabilize the circuit and make the gain predictable. The circuit shown in Fig. 10

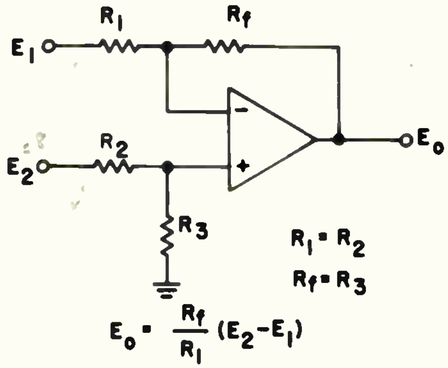


Fig. 10. Using an op amp as a differential amp.

produces this differential amplifying capability with controlled gain.

In this circuit the two input resistors are R_1 and R_2. These two resistors are generally made equal. The feedback resistor R_f is connected as usual. Resistor R_3 is connected between the non-inverting input and ground. Resistors R_3 and R_f also are made equal to each other. The gain of the circuit is expressed as R_f/R_1 or R_3/R_2. The output voltage of this circuit is equal to the gain multiplied by the difference of input voltages E_1 and E_2. This gives us the output expression

$$E_o = \frac{R_f}{R_1}(E_2 - E_1)$$

We might also call this circuit a subtraction circuit, since we subtract input voltage E_1 from E_2 in the process of calculating the output.

Let's assume that resistors R_3 and R_f are equal to 100K and R_1 and R_2 are 25K. The gain of the circuit then is 100K/25K = 4. Now let's assume input voltages of E_1 = 3 volts, and E_2 = 8 volts. The output voltage of the amplifier is the difference between the two input voltages (5) multiplied by the gain of the circuit (4), or 20 volts.

Let's take another example of output voltage calculation of a differential amplifier circuit. Assume that the values of R_f and R_3 are 500K-ohms. Input resistors R_1 and R_2 are also equal to 500K-ohms. Input voltage E_1 = +13 volts, and input voltage E_2 = −2 volts. What is the output voltage? Using the output formula

$$E_o = \frac{R_f}{R_1}(E_2 - E_1)$$

we get

$$E_o = \frac{500K}{500K}(-2 - 13)$$

$$E_o = 1\,(-15) = -15 \text{ volts}$$

This circuit performs a very accurate algebriac subtraction operation.

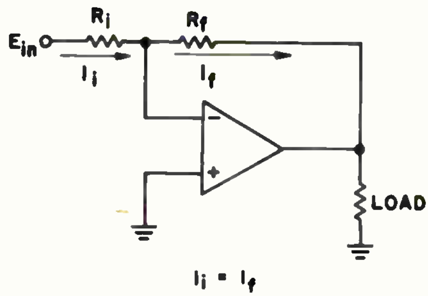


Fig. 11. A conventional op amp circuit with a load showing that the input and feedback currents are equal.

AN OP AMP CONSTANT CURRENT SOURCE

In an earlier lesson, you learned how a transistor can be connected to form a constant current source. Now you will find out how an op amp can be connected to perform this same function.

The conventional op amp circuit is shown in Fig. 11. The input is applied to resistor R_i and the gain is determined by the ratio of R_f to R_i. The output voltage is generally applied to a load connected between the amplifier output and ground as shown. Earlier we mentioned that the input and feedback currents were equal. Because of the high input impedance of the op amp, almost no current flows into the inverting input. Therefore the input and feedback currents I_i and I_f in Fig. 11 are equal.

Using this fact, we can produce the constant current source op amp circuit. Fig. 12 shows the arrangement. First we cause a fixed current to flow through input resistor R_i. Since the value of R_i is constant, we can cause the current flow through it to be constant by applying a constant voltage to it. The Zener diode in Fig. 12 is used to provide the constant voltage to resistor R_i.

This fixed current also flows in the feedback resistor. Instead of connecting the load directly between the output of the amplifier and ground, we connect the load as the feedback element. In this way the constant current, determined by the

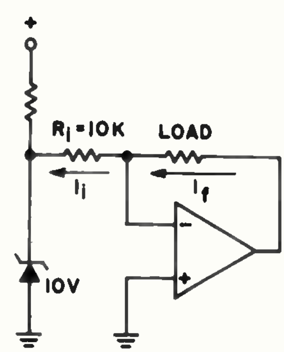


Fig. 12. A constant current source made with an op amp.

Zener diode and the input resistor, also flows through the load. The load resistance can be varied over a wide range of values but the current flowing through it will remain constant. For example, if R_i is a 10K resistor and the Zener diode provides a voltage of 10 volts, then the input current is equal to 10/10K = 1 milliampere. A constant current of 1 milliampere will flow through the load. This current will remain constant despite variations in load resistance.

VARYING THE GAIN OF AN OP AMP

In all the op amp circuits we have discussed, the gain is fixed by the values of the feedback and input resistors in the circuit. However, there are many occasions where we wish to control the gain or vary it continuously over a wide range of values. The most obvious way of varying the gain of any op amp circuit is to replace one of the fixed resistors with a potentiometer connected as a variable resistor.

A potentiometer is nothing more than a variable voltage divider, a resistive network whose output is less than its input by a factor determined by the resistance values used. The "pot" is a variable voltage divider with a continually adjustable resistance ratio.

Fig. 13A shows a typical op amp circuit with a variable resistor at both the input and the feedback positions. However, to control the gain of the amplifier, a variable resistance will usually be required in only one of these positions; a fixed resistor can be used in the other. By varying the input or the feedback resistance, the ratio of the two resistors will change and the gain will vary. Continuous adjustment of the gain is possible with this method.

With this arrangement it is more often desirable to control the gain with a potentiometer in the feedback path than in the input circuit. By varying the potentiometer in the feedback path, we obtain linear control of the gain. If we adjusted the input resistance, the gain would not be linear; it would not vary in a straight line with the value of the resistance. This is because the gain is a function of the ratio of the two resistances.

Also in this arrangement, the input resistance varies as the gain changes, altering the loading of the driving circuit.

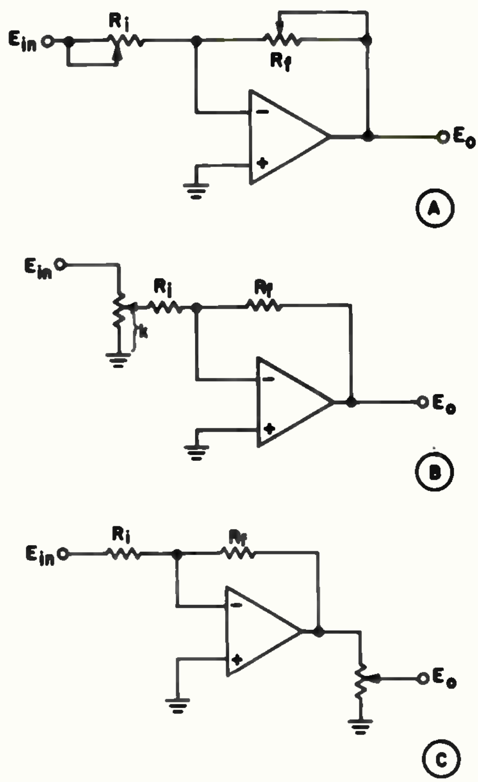


Fig. 13. Methods of varying the gain of an op amp.

The input impedance can even be set to zero, placing a dead short across the source driving the amplifier. In most instances, however, this is not a desirable situation.

An easy method of controlling the gain of an op amp circuit is to place a potentiometer before or after the circuit, as shown in Fig. 13B and C. In both circuits the op amp gain is fixed by the ratio of the feedback and input resistances; the amplitude of either the input or output is controlled with a potentiometer.

Fig. 14 shows a potentiometer connected as a variable voltage divider. The input voltage is applied between the upper end of the potentiometer and ground, causing current to flow through the resistance. The output voltage is taken between ground and the variable arm.

With the arm all the way to the top of the resistance, the output voltage will be equal to the input voltage. If we move the arm to the bottom of the resistance, the output voltage will be zero. Positioning anywhere between the two extremes causes the output voltage to be some fraction of the input voltage. For example, if the input voltage is 100 volts and we set the potentiometer arm to the center of the resistance value, the output voltage will be 50 volts.

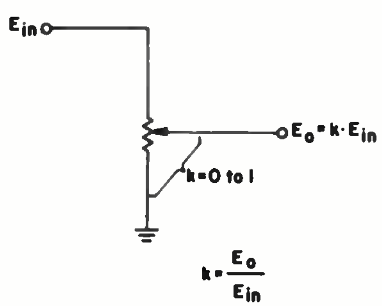


Fig. 14. The input-output relationship of a potentiometer.

In effect we are tapping off a portion of the voltage dropped across the resistance of the pot. Expressed as a fraction, the percentage of voltage tapped off between the arm of the pot and ground is known as the pot coefficient and designated k. In our example, where we assumed a 100 volt input, the coefficient is .5. Multiplying the pot coefficient by the input voltage gives us the output voltage value. With an input voltage of 100 volts and a pot coefficient of .25, the output voltage will be 25 volts. Since the pot is continuously variable, we can set any coefficient between 0 and 1, where the output and input voltages are equal.

Therefore, by connecting the pot at the input or the output of an op amp, we can vary the gain of the amp. The overall gain of the circuit shown in either Fig. 13B or C is equal to the amplifier gain multiplied by the pot coefficient. In other words, the overall circuit gain is equal to

$$\frac{R_f}{R_i}k$$

When the pot is used this way, the overall gain of the circuit can never be more than the total gain of the op amp.

One of the most useful circuits for varying the gain in an op amp is shown in Fig. 15. In this type of circuit, the input and feedback resistors are normally equal. Notice that the feedback resistor is indirectly connected to the output of the op amp through the arm of the pot. With this arrangement the gain of the amplifier circuit is equal to the reciprocal of the pot coefficient. If, for example, the pot

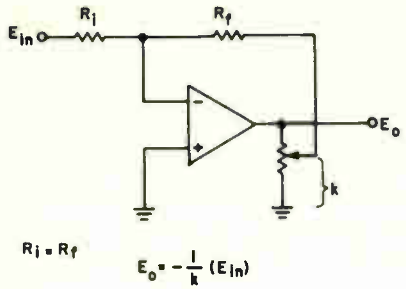


Fig. 15. A variable, high gain, op amp circuit.

coefficient k is .5, the gain of the amplifier is 1/.5 = 2. Since the coefficient of a potentiometer can be adjusted to any value between zero and 1, we can theoretically obtain any gain between 1 and infinity; the only limiting factor is the resolution of the pot setting.

For some potentiometers it is difficult to obtain very small variations in coefficients. We may want to obtain a gain of 1000 with our amplifier, in which case the coefficient would have to be .001. This coefficient is the ratio of 1 to the desired amplifier gain (k = 1/gain = 1/1000 = .001).

The output of voltage from the amplifier circuit, shown in Fig. 15, is equal to the gain of the circuit (1/k) multiplied by the input voltage. Keep in mind that the gain of the amplifier should not be adjusted so high that, with the given input voltage, it causes the output amplifier to swing to an output voltage beyond the power supply voltages used.

SELF-TEST QUESTIONS

(i) The non-inverting amplifier circuit of Fig. 8 has a feedback resistor of 120K and an input resistor of 30K. What is the gain?

(j) The differential amplifier circuit of Fig. 10 has values of $R_f = R_3 = 75K$ and $R_1 = R_2 = 15K$. The input voltages are $E_2 = -23$ and $E_1 = -16$ volts. What is the amplitude and polarity of the output?

(k) True or false: An op amp follower inverts the input signal.

(l) True or false: The load in an op amp constant current source is connected between the output and ground.

(m) A pot has an input voltage of 27 volts and an output of 9 volts. What is the pot coefficient k?

(n) The pot coefficient in the circuit of Fig. 15 is .004. What is the circuit gain?

3-18-76

Characteristics and Specifications

Most of the characteristics and specifications that we discuss in this section arise from the fact that an op amp is an imperfect device. Like any other electronic circuit, it is made from components that have specific tolerances on their values; undesirable characteristics cannot be completely eliminated in their manufacture.

Despite their imperfection, op amps are still quite useful; their usefulness depends entirely upon their application. By knowing and understanding the characteristics and specifications of op amps, you will be able to compare them and determine whether a particular amplifier is suitable for a given application.

INPUT OFFSET VOLTAGE

Almost all op amps have a differential input stage. As you have seen in previous applications, the two inputs providing inverting and non-inverting operation permit the op amp to be used in a wide variety of circuits. The desirable characteristics of a differential amplifier are taken advantage of in the op amp. As you recall, any differential amplifier amplifies the difference voltage between the two inputs. If we should connect both inputs to ground (0 volts), the difference between the two inputs should be zero; the output voltage would also be zero.

This, however, is only theoretically true. Because of small differences in the characteristics of the transistors used in the input differential amplifier, there will be a small difference of voltage that will be amplified by the high gain of the circuit and appear as a dc output voltage. Therefore, even when both inputs are at ground, there will be a measurable output voltage.

The small difference voltage, that appears in the input differential amplifier when both inputs are connected to ground, is caused by the differences in the emitter-base voltages of the two input transistors. If the emitter-base voltages, designated V_{be}, are exactly equal, there will be no difference between the two input voltages and the output voltage will be zero. When separate transistors are used to form the input differential amplifier, it is very difficult to find two transistors with identical V_{be}'s. For that reason there will be a small difference, appearing as an input signal that is amplified by the high gain of the circuit and appears as a measurable dc voltage at the output. The amount of voltage that must be applied between the input terminals to offset this output voltage is known as the input offset voltage.

In the design of a transistor op amp, the manufacturer normally tries to choose matched input transistors. Special test set-ups are used to compare the emitter-base voltages for the input transistors to match them as closely as possible. This reduces the input offset voltage. Some manufacturers produce special differential transistors with an extremely small difference voltage. Integrated circuit op amps, which you will study later, are very good in this respect. The transistors of the input amplifier are constructed on the same chip of silicon material and exhibit almost identical characteristics, reducing this input offset voltage to practically zero.

In a good op amp, this input offset voltage varies anywhere from several hundred microvolts to about 10-20 millivolts. When you consider the fact that the open-loop gain of an op amp is many thousands, this voltage can be amplified to a substantial signal at the output. For this reason we must generally take steps to correct this condition.

Fig. 16 shows a circuit that can be used to compensate for input offset voltage. This circuit is a standard inverting op amp circuit with feedback and input resistors. When we ground the input resistor as shown, the output voltage should be zero. In a practical circuit the output will not be zero, however, due to the input offset. If, for example, R_f = 100K and R_i = 100 ohms, the circuit gain would be 100K/100 = 1000. With an input offset voltage of 5 millivolts, the output would be .005 X 1000, or 5 volts. This is far from zero as you can see.

To compensate for this, we can apply a correction voltage from a potentiometer, as shown. The pot is connected to both plus and minus power supply voltages. This enables us to apply either a negative or a positive voltage to resistor R_1 which feeds the offset voltage into the op amp.

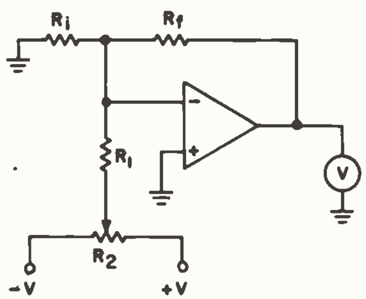


Fig. 16. Correction for input offset voltage.

Since we do not know in which direction the offset voltage will force the output, we can apply the desired correction voltage by making the input voltage polarity and the amplitude variable with the pot. With the input resistance R_i grounded and a voltmeter connected to the output, potentiometer R_2 is adjusted until the output voltage is equal to zero. The input offset voltage is then corrected.

The input offset voltage limits the magnitude of the input signal that the amplifier can amplify. For example, if the input offset voltage were 5 millivolts, we could not handle or amplify dc signals of 2 millivolts amplitude. Such small level signals, approximately the same magnitude of the offset voltage, will be confused with the offset voltage. An unpredictable output can result. However, by balancing out the offset voltage with a circuit like that in Fig. 16, we can then handle small amplitude signals. In noncritical applications where very low gain is used and where the signal voltages are many times larger than the offset, it is often possible to disregard the need for input offset voltage correction if a good op amp is used.

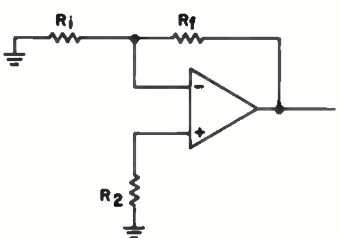
INPUT OFFSET CURRENT

The input offset current is the difference between the two input currents with both of the op amp inputs connected to ground. This difference in base currents is amplified and the result is an output voltage. The average of the two input base currents is called the input bias current. It is found by adding the input currents and dividing the sum by two.

This offset current is a function of the balance, or matching, of the input transistors in the differential input stage. Instead of being the result of differences in V_{be}, as with input offset voltage, the input offset current is a function of the degree of matching between the Betas and leakage currents of the input transistors. Because the two transistors used in the differential stage will not have equal Betas or collector-base leakage currents, different amounts of base current will flow when the inputs are connected to ground. This difference between the two currents (the input offset current) is amplified and produces an offset output voltage.

This current can be compensated for in the same way we correct for offset voltage (Fig. 16). Because of the directions of the current and possible opposite voltage offsets, it may be necessary to apply correction input to both the inverting and non-inverting inputs in order to compensate for both effects. Normally, however, the arrangement shown in Fig. 16 will enable you to null out the effect of input offset voltages and currents.

Fig. 17 shows an op amp circuit that can be used to help minimize the input offset current effects. Instead of connecting the non-inverting input directly to ground, as in previous circuits, we insert a resistor between this input and ground. By making this resistance equal to the parallel combination of the normal input and feedback resistances, the input currents can be made very nearly equal so that their results will cancel. The result will be a minimum of output voltage offset due to this input offset current. By making R_2 in Fig. 17 fully adjustable, it is possible to bring the offset to practically zero. If the input and feedback resistors are both 100K, the parallel combination will be equal to one-half this value. R_2 in the circuit would be made 50K for minimum offset current effects.



$$R_2 = \frac{R_f R_i}{R_f + R_i}$$

Fig. 17. Input offset current compensation.

INPUT IMPEDANCE

One of the most important characteristics of an op amp is the input impedance, the resistance at the input of the amplifier. In other words, each differential input represents a certain amount of resistance to ground to some driving source. It is desirable to have this input impedance as high as possible so that the circuit does not load or draw current from the source driving it.

Whenever this input resistance is connected to an oscillator, amplifier or other voltage source, current will flow through the input impedance and effectively load the driving source. If the input impedance is too low, excessive loading of the driving source may occur and result in reduced input voltage. Therefore the higher the input impedance, the lower the current drawn from the source.

Most typical op amps with no feedback connections have an input impedance of from 20K to 500K-ohms. This is the impedance of a typical bipolar transistor op amp. Higher input impedances can be obtained with Darlington input differential amplifiers. When field effect transistors are used for the input differential amplifier, extremely high values of input impedance (even thousands of megohms) are possible.

There is one special technique used with op amps to increase the input impedance. Known as bootstrapping, this technique involves taking some of the output voltage of the same polarity as the input voltage and feeding it back to the input. This compensates for the amount of current being drawn by that input from the source. In other words, we supply a current of proper polarity from the output of the amplifier to the input to furnish the current to the input normally drawn from the source. By effectively canceling out the current drawn from the source, the apparent input impedance becomes infinite.

The circuit shown in Fig. 18 uses this bootstrapping technique. Notice that resistors R_1, R_2, R_3 and R_4 have been added to form a voltage divider which taps off part of the output voltage which is fed back to offset the need for input current from the signal source.

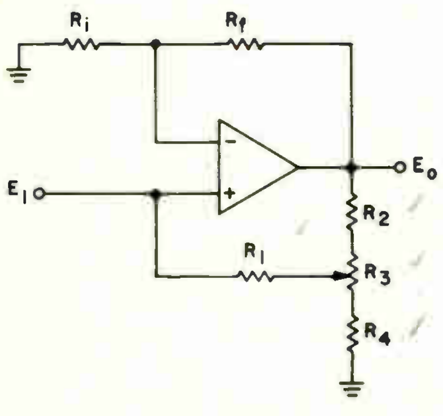


Fig. 18. A non-inverting op amp with bootstrapping to increase input impedance.

OUTPUT IMPEDANCE

The output impedance is the value of the effective internal resistance of the amplifier when it is acting as a voltage generator. The output of an op amp appears to be a source of voltage for whatever load it drives. The output impedance of the op amp appears in series with this load. Ideally we would like to have a very low output impedance so that little or none of the voltage produced is dropped across the internal impedance of the amplifier; almost all of it should appear across the load.

The output impedance of a typical transistor op amp without feedback is generally several hundred ohms. However, this value is reduced substantially when feedback is used. By using negative feedback, as we normally do with an op amp, output impedances of less than 1 ohm are easily obtainable. If the gain of the amplifier is very high, output impedances of .01 ohm are not too difficult to obtain.

POWER CONSUMPTION

Power consumption is the dc power required to operate the amplifier with the input and output voltages at zero and no load current flowing. This is the quiescent amount of current drawn by the amplifier when no input signal is applied or output is being developed.

Remembering that power is equal to the product of the voltage and the current, you can easily find the power consumed by an op amp. Most op amps use two power supplies, one to supply a negative voltage and the other a positive voltage. For example, a typical op amp may require + and – 15 volt power supplies. If the current drawn from each of these 15-volt power supplies is 2 milliamperes, power consumption would be 30 volts × .002 ampere = .06 watt, or 60 milliwatts.

INPUT VOLTAGE RANGE

The input voltage range is a measure of the maximum amount of voltage that can be applied between the two inputs of the op amp. If this value is exceeded, damage may occur to the input differential amplifier transistors. Normally this voltage range is restricted by the reverse emitter-base breakdown voltage of the input transistors.

OUTPUT VOLTAGE SWING

The output voltage swing is the peak or maximum output voltage, with reference to zero, that can occur without clipping or other distortion. This voltage swing is generally limited by the power supply voltages. For example, in an op amp with power supplies of ±15 volts, the maximum output voltage swing would be very nearly equal to 30 volts. For an ac sine wave at the output, this would be a 30-volt peak-to-peak value. Normally, the maximum output voltage swing is less than the supply voltage by an amount equal to the drop across the saturated output transistor and any series resistance.

When the input voltage is low enough so that, when it is amplified by gain of the amplifier it produces an output voltage within the output voltage swing, the amplifier is said to be operating in the linear region. That is to say, the output voltage is an enlarged copy of the input voltage. The amplifier does not distort the signal.

However, if the input voltage is made too large, the gain of the amplifier may be such that we exceed output voltage swing capabilities of the amplifier. In this case, the amplifier output will saturate. If the input signal is a sine wave, both the positive and negative peaks of the output will be clipped because the output transistors go into saturation. To avoid going beyond the output voltage swing capability, always check to be sure that the output voltage is less than the maximum value when you multiply the input voltage by the gain of the amplifier.

COMMON MODE REJECTION

The inputs to an op amp are the inputs to a differential stage. As you recall from your study of differential amplifiers, it is the voltage difference between the two inputs that is amplified by the stage and fed to the output. This means that if both input signals are equal, the output voltage should be zero.

The ability of the amplifier to produce a very small or zero output voltage for equal input voltages is known as the common mode rejection. The ratio of the common mode input voltage to the output voltage produced is known as the common mode rejection ratio. For example, if we apply 10 volts to both inputs of the amplifier and the resulting output voltage is 1 millivolt, the common mode rejection ratio is 10/.001 = 10,000.

Normally the common mode rejection ratio is specified in decibels. The common mode rejection ratio that we just calculated, expressed in decibels, would be equal to 20 log 10,000 = 20 × 4 = 80 db. The larger the common mode rejection ratio figure, the better the rejection ratio and the quality of the amplifier. Common mode rejection ratios of well over 100 db are not uncommon in op amps.

POWER SUPPLY REGULATION

The input offset voltage of an op amp is sensitive to changes in the power supply voltages. We call this op amp characteristic the power supply rejection ratio. This is basically the ratio of the change in input offset voltage to the change in the supply voltage producing it. Normally, the op amp has two power supplies, one held constant and the other varied in voltage. Therefore, any input offset voltage change is noticeable. The ratio of the input offset voltage change to the supply voltage that caused it is the power supply sensitivity.

The power supply sensitivity is a function of the balance or matching in the input differential amplifier. If the differential amplifier transistors are as closely matched as other components in the circuit, the change in input offset voltage for a power supply voltage change will be very small. The quality of the amplifier is inversely related to this variation.

OPEN-LOOP DC VOLTAGE GAIN

The voltage gain of the op amp without external feedback is called the open-loop dc voltage gain. As we mentioned before, the open-loop gain of an op amp is generally very large. Gains of 100,000 or more are not uncommon. However, most op amps will use some form of negative feedback to reduce the amplifier gain and make it more predictable.

As you saw earlier, negative feedback through a resistance between the output and the inverting input makes the gain strictly a function of the ratio of the feedback to an input resistance. It is very desirable to be able to control the gain in this way. In order for this relationship to hold true, however, the open-loop gain of the amplifier should be at least 100 times the desired closed-loop gain of the amplifier.

NOISE

All amplifiers are sensitive to and produce a certain amount of noise. Noise is a term that describes any stray signal voltage (produced within the amplifier or external to it) that is amplified and appears in the output. Noise is an unwanted signal, generally measured in terms of rms voltage, that may interfere with the amplification of a specific input. If the noise level of the amplifier is higher than the small signal to be amplified, the input signal will be completely masked; only an amplified noise output will occur. Therefore, it is important that the noise in the amplifier be minimized.

There are several types of noise which can occur in op amps. We generally break these down into noise components associated with specific frequency ranges. One of these, called the wide band noise spectrum, extends from about 1 kHz to 100 kHz. Noise produced in this frequency range is generally from thermal effects in resistors and from within semiconductor components, such as diodes and resistors. This type of noise is internally generated.

Another noise spectrum extends from the frequencies of 10 Hz to 1 kHz. Noise in this region is usually from the ac power line and frequencies harmonically related to it.

In the frequency range from 1 Hz to 10 Hz, noise is caused by structural imperfections in the components (particularly transistors) and is known as flicker noise.

In the frequencies above 100 kHz we move into the radio frequency spectrum, and amplifiers can pick up rf signals. Of course, this is undesirable. It is best to minimize the amount of noise produced both within the amplifier and external to it. This is particularly true if extremely low level signals are to be amplified properly.

DRIFT

Drift is a form of very low-frequency noise; it is the slowly varying change in the input offset voltages and currents in an op amp due to fluctuation in temperature, time and power supply voltages. We want to minimize drift in the design and application of the amplifier because variations in offset voltages and currents are amplified and interfere with the normal input signal.

No relationship exists between the input offset voltage drift and input offset current drift. As temperature, time or power supply voltages change, each of these factors will vary independently, often in opposite polarity to the other. Since drift is an undesirable characteristic of the op amp, every means must be taken to minimize or eliminate it. Let's now discuss the causes of drift and how it can be minimized.

A major cause for drift is temperature change. Electronic equipment using op amps can be subjected to a wide range of temperatures. Temperature effects can come from the environment or from heat in adjacent electronic components or equipment. Such temperature effects must be considered when designing and using an op amp.

Most operational amplifiers and components are available in two basic temperature ranges, the commercial range from 0°C to 70°C and the military temperature range of −55°C to +125°C. These are the typical temperature ranges over which the op amp must work for these specified applications. Since the temperature ranges are so wide, care must be taken to protect the amplifier.

Temperature variations cause changes in the input offset voltage and current which can produce undesirable amplifier action. As the temperature varies, the emitter-base voltage drops of the differential input transistors vary. Normally the emitter-base voltage drop decreases 2 millivolts for every °C increase in temperature. If the two differential input transistors are accurately matched, their emitter-base voltage drops will decrease together as temperature increases, maintaining a balance and keeping the input offset voltage reduced to a minimum.

Recall that the input offset voltage is due to the differences between the emitter-base voltage drops in the input differential amplifier. It is extremely difficult to control the manufacture of these transistors so that the emitter-base voltage drops track together with temperature. For that reason the input offset voltage will change when the temperature changes. Emitter-base junction voltages are so sensitive to temperature changes that a temperature difference of only .01°C between the junctions of two transistors in the input stage can produce an input offset voltage of 24 microvolts.

As you can see, it is important to hold the junction temperature of the two transistors to exactly the same value. Keeping the input offset voltage low, when individual transistors are used in the differential stage, is quite difficult. Generally the transistors can be mounted on a common heat sink. Also there are special differential transistor pairs manufactured, where both are made from the same silicon chip. Often special heating elements or ovens are used on the differential input stage to keep the temperature at a constant value to reduce the changes in the offset voltage that produce undesirable drift.

Drift due to changes in input offset voltage is generally specified in millivolts per °C of temperature change. For example, a typical drift may be .05 millivolt per °C. This can be evaluated by varying the temperature, noting the input offset voltage drift and then calculating the change for a given increment of temperature.

Fig. 19 shows a curve relating the change in input offset voltage to a change in temperature. At 25°C (room temperature), the input offset voltage is nulled to zero. Normally this is done by using the compensation circuit shown in Fig. 16. As the temperature is lowered, the offset voltage increases in the negative direction. Raising the temperature from 25°C causes an increased offset voltage in the positive direction. As you can see, the input offset compensation voltage applied by the circuit in Fig. 16 is effective only at one temperature; the output will drift due to the change in offset voltage at other temperatures.

Notice in this curve that the offset voltage does not change linearly with temperature. In other words, equal temperature change does not produce a proportional change in input offset voltage. In addition, the offset voltage in one amplifier varies independently from that in another.

The input offset voltage also tends to drift with time. That is, the input offset voltage will change over a long period of time. Normally this condition is expressed in microvolts of offset change per days or per thousand hours. This change is generally caused by the physical changes inside transistors because of their age. In most op amps, this characteristic is not particularly detrimental to the operation of the circuit. Only in the most

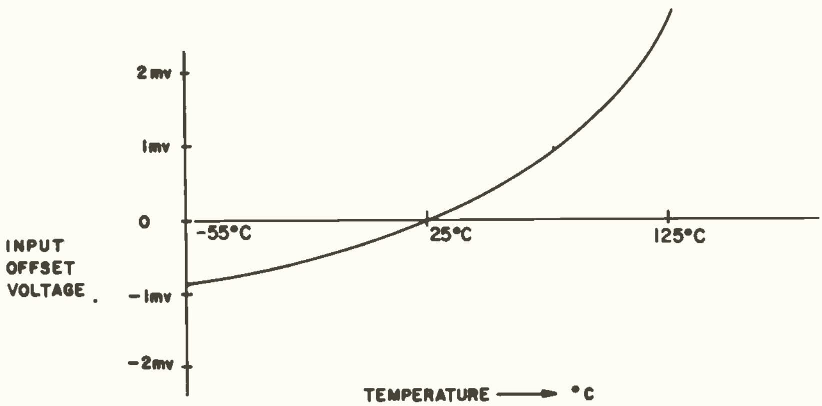


Fig. 19. Input offset voltage vs. temperature.

critical applications is the offset change with time a crucial factor.

The input offset current in an op amp also drifts with temperature and time changes. As you recall from our earlier discussion, the input offset current is a function of both the gain, or Betas, of the input transistors and the leakage currents. Both these factors increase when temperature rises. In silicon transistors leakage current generally doubles for every 10°C increase. If the input transistors are matched so that their Betas and leakage currents vary the same amount with changes in temperature or age, then the offset current will not change. However, it is extremely difficult to match the input transistors this closely; they will change differently with temperature, increasing the amount of input offset current with an increase or decrease in temperature.

Changes in Beta or leakage with temperature do not directly correspond to changes in input offset voltage. In one particular situation, they may vary together and their effects may add; in another case their effects may oppose one another. It depends upon the individual components and the circuit in which they are used.

Generally the same precautions used to minimize input offset voltage changes with temperature can be used in minimizing changes in input offset current. Mounting the transistors together on a common heat sink or using differential transistor pairs made on the same silicon chip will help eliminate the effects of temperature changes.

As with input offset voltage changes, the offset current will vary due to aging of the transistors. There is very little that can be done except to try to match the transistors for aging effects as closely as possible. These are usually long term, very slow drift effects which can be neglected for most applications.

FREQUENCY RESPONSE

As you learned in an earlier lesson on amplifiers, amplifier circuits generally respond to a specific band of frequencies. Amplifiers are classified according to the frequency range in which they amplify best. There are audio amplifiers, video amplifiers, rf amplifiers and dc amplifiers. The op amp is a special case in that it covers a wide range of frequencies. This wide frequency response capability adds to the versatility of the op amp.

We defined frequency response in terms of bandwidth. Bandwidth, as you recall, is the difference between the upper and lower three db cutoff points on the amplifier response curve. For op amps the lower frequency limit is dc. The upper 3 db point primarily depends upon the internal capacitances of the op amp circuit and components as well as external capacitance, inductance and other stray impedances. Op amps have been constructed to perform at frequencies up to 100 MHz.

Now let's take a look at a typical frequency response curve for an op amp. In Fig. 20 the upper curve shows the frequency response of the op amp operating with no feedback. The curve is designated as A_1. This is the open loop response that indicates how the gain of the amplifier varies with frequency. Notice that the gain remains constant for only a very narrow frequency range before it begins to decrease linearly as the frequency is increased. The 3 db down point for the open-loop gain occurs at an extremely low frequency, f_1, that is about 100 Hz or less. Beyond this point

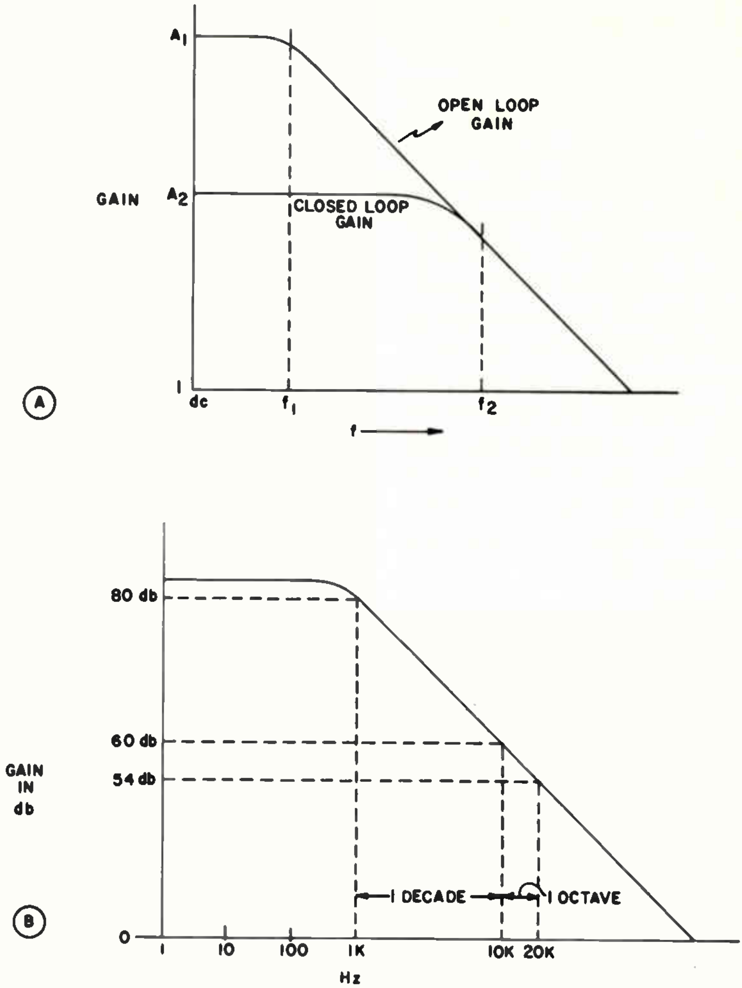


Fig. 20. Op amp frequency response (A) and an op amp frequency response curve that illustrates the roll-off characteristic (B).

the gain decreases in a straight line as the frequency increases until the gain is equal to one (unity).

The lower curve shown in Fig. 20A is the closed-loop frequency response of the op amp. If we apply negative feedback to the op amp, we immediately reduce its gain. At the same time we also increase its bandwidth. Notice that the 3 db down point (f_2) for the closed-loop gain response is far greater than that of the open loop gain response. Here we are trading off gain for frequency response by adding the negative feedback. Notice that the initial gain is lower, but that beyond the 3 db point, the closed-loop gain response merges with that of the open loop gain response and the gain drops linearly as the frequency increases.

The rate of decrease in gain with respect to frequency is generally called the roll-off. The gain roll-off characteristics of an op amp are generally set so that the gain decreases at a rate of 6 db per octave. This means that the voltage gain drops by a factor of 2 each time the frequency doubles. Expressing this another way we can say that the gain roll-off occurs at a rate of 20 db per decade; when the frequency is increased by a factor of ten, the gain is decreased by a factor of ten.

Fig. 20B illustrates this gain roll-off characteristic. The gain in this curve is expressed in db and the frequency in Hz. If the frequency should increase from 1 kHz to 10 kHz, we can say that the frequency increased by a decade or by a factor of ten. The gain of the amplifier at 1 kHz is 80 db, and it drops to 60 db at the 10 kHz point. 80 db represents a gain of 10,000 while 60 db represents a gain of 1,000. Therefore, as the frequency increased by a factor of ten, the gain dropped by a factor of ten.

Notice on this same curve that as we increase the frequency from 10 kHz to 20 kHz the gain drops a small amount. We have effectively doubled the frequency which represents one octave. The gain has dropped 6 db, representing a factor of 2. (60 db represents a gain of 1,000 while 54 db represents a gain of 500.)

There is one specification of an op amp that helps define the relationship between gain and frequency. Known as the gain-bandwidth product, this term will show you the general relationship between the gain of the amplifier and its bandwidth when negative feedback is applied.

The gain-bandwidth product of an amplifier is a constant which is determined by the amplifier design. This means that as we change the gain of the amplifier, the bandwidth must also change in order to keep the product of the two constant. For example, an amplifier with a gain-bandwidth product of 10^6 would have a bandwidth of 10^6 Hz at unity gain. Multiplying the gain of one by the bandwidth of 10^6 Hz we get 10^6 gain-bandwidth product. Now assume that we increase the gain of the amplifier to 100. As you would expect from looking at the curve of Fig. 20B, increasing the gain would have the effect of decreasing the bandwidth. With a gain-bandwidth product of 10^6 and a gain of 100, the bandwidth must be $10^6/100 = 10^4$ Hz or 10 kHz.

SLEWING RATE

The slew rate limit of an op amp is the maximum time rate of change of the output voltage for a step input voltage. This is sometimes known as the velocity limit. The slewing rate is usually specified in terms of the change in output voltage per unit of time; volts per microsecond is typical. It is closely related to the frequency response of an amplifier.

In a previous lesson you studied the square wave testing of amplifiers where a step voltage or square wave was applied to the input and the output of the amplifier then noted. The change in shape of the square wave by the amplifier gives much useful information about the response characteristics of the amplifier. A step response signal (square wave) contains a large number of high-frequency harmonic sine waves. If the amplifier passes these high-frequency harmonics faithfully, the output waveform should be the same as the input waveform. However, since some of the high-frequency components will be attenuated by the limited bandwidth of the amplifier, the output waveform will actually be different from the input waveform. The amount of difference determines the bandwidth of the amplifier. This can be calculated by measuring the output rise time of the square wave of the amplifier.

Fig. 21 shows a step voltage input applied to an op amp. Notice that the rise time for the pulse is zero. In a practical situation, the rise time of the input pulse would be finite. It is usually made so much smaller than the expected output rise time that its effects can be eliminated.

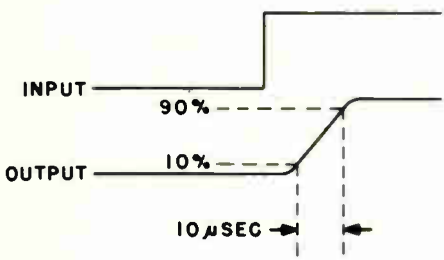


Fig. 21. The step response of an op amp.

The output of the op amp is also shown in Fig. 21. Notice that the output has a very definitely measurable output rise time. These waveforms might be displayed on an oscilloscope and the calibrated horizontal sweep and crt screen will be used to measure the rise time between the 10% and 90% points of the waveform.

Frequency Compensation. Because of the very high gain and frequency response characteristics of an op amp, the circuits are often susceptible to instability when feedback is used. Instability refers to the tendency of an amplifier to oscillate. Naturally any oscillation in an amplifier is undesirable. In op amps oscillation can be eliminated by using frequency compensation techniques.

As you know, the gain of an amplifier varies with the frequency. When feedback is used, the gain is relatively constant over a range of frequencies. Then as the frequency approaches an upper limit, the gain begins to roll off at a constant rate.

As you recall from your previous studies of amplifiers, as frequency increases the gain drops off and substantial phase shift is introduced. At the half power (3 db down) point the phase shift between the input and output signals will be 45°. This phase shift will increase further as the frequency is increased and the gain continues to roll off. It is possible that at some point during the roll-off the phase shift will equal 180°. Combining this with the 180° phase shift or inversion of the op amp and assuming a feedback connection, the output signal will effectively be in phase with the input. If the gain is equal to or greater than 1 at the point where the 180° phase shift occurs, then the amplifier will oscillate. The frequency of oscillation will be the frequency at which the phase shift is proper for in-phase feedback.

To avoid this unstable condition at high frequencies, it is often necessary to connect external resistor-capacitor networks to an op amp to shape the output response curve. This makes the gain roll off at a rate where the phase shift will not be 180° when the gain of the amplifier is equal to or greater than unity.

The phase shift characteristics of an amplifier depend upon the amplifier circuit and the various external resistor-capacitor networks used. Occasionally at high frequencies, these characteristics cause the phase shift to approach 180° when the amplifier gain is equal to or greater than one. Even without feedback an amplifier can oscillate in the open-loop condition. This is due to stray capacities and coupling between circuits of the amplifier caused by the common power supply impedance.

Such instabilities can be eliminated by using an external frequency compensation network that forces the amplifier to roll off at a 6 db per octave rate. Doing this makes the amplifier appear to have the frequency response of a simple R-C integrator network. Such a network has a maximum phase shift of 90°. Since the amplifier can never achieve the 180° phase shift, it will be stable over its useful frequency range.

SELF-TEST QUESTIONS

(o) What is the basic cause of input offset voltage?

(p) What causes input offset current?

(q) Why is a high input impedance desirable?

(r) What is meant by rise time?

(s) What determines the maximum output voltage swing of an op amp?

(t) What causes drift in an op amp?

(u) Does the circuit of Fig. 16 compensate for the effects of drift?

(v) A certain op amp has a gain of 120 db at 100 Hz. What is its gain at 1 kHz if the gain rolls off at 6 db per octave?

(w) Why are external frequency compensation networks often used with op amps?

3-18-76

3-18-76

Typical Circuit Techniques

Fig. 22 is a generalized block diagram of an op amp, showing its three stages. The first stage is called the input stage. It is always a differential amplifier. Usually a current source feeds the emitters of the differential transistors. This current source is either a single high value resistor or a transistor, forming an active current source. This stage is important because it determines the input characteristics of the op amp. The input offset voltage, input impedance and common mode rejection, for example, are primarily determined by how the input stage is designed.

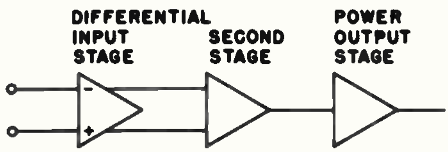


Fig. 22. An op amp block diagram.

Since op amps are dc amplifiers, the first stage is directly coupled to the second stage. The second stage provides more amplification and produces the proper level translation for the dc signals on which the ac input signals are riding. Without level translation there is a buildup in the dc voltage coupled from stage to stage. Also the output will not swing near ground so that the output is zero when the input is equal to zero volts. We could have shown the second stage in Fig. 22 as two blocks, one showing a gain stage and the other a level translator or shifter.

In many op amps there are more than two stages of gain. Usually the differential signal is eliminated after the second or third stage and only a single-ended output is fed to some form of power output stage.

The power output stage has two functions. First it tends to maximize the output voltage the amplifier can handle. (A large output voltage is a desirable op amp characteristic.) It also determines the output impedance (which must be as low as possible) and the current handling capabilities of the amplifier. In many cases the output stage must also provide some form of short circuit protection.

BASIC CIRCUITS

Let's now discuss in detail some amplifier circuits which illustrate how the blocks in Fig. 22 are filled. Fig. 23 is the circuit of a simple but complete op amp which illustrates most of the principles involved. In this circuit the first stage is made up of transistors Q_1 and Q_2, which form a differential amplifier. The input to transistor Q_1 is an inverting input, while the input to transistor Q_2 is the non-inverting input. The emitter of both stages of the transistors of the differential amplifier must be fed by some form of current source. In this case, a very simple current source is made by having a large value resistance (R_4) connected to minus supply $-V_{ee}$. Resistors R_1 and R_2 are the load resistors for the first stage transistor.

The collectors of transistors Q_1 and Q_2 are direct-coupled to the bases of transistors Q_3 and Q_4, which form the second stage. The second stage emitters are also fed by a current source, in this case

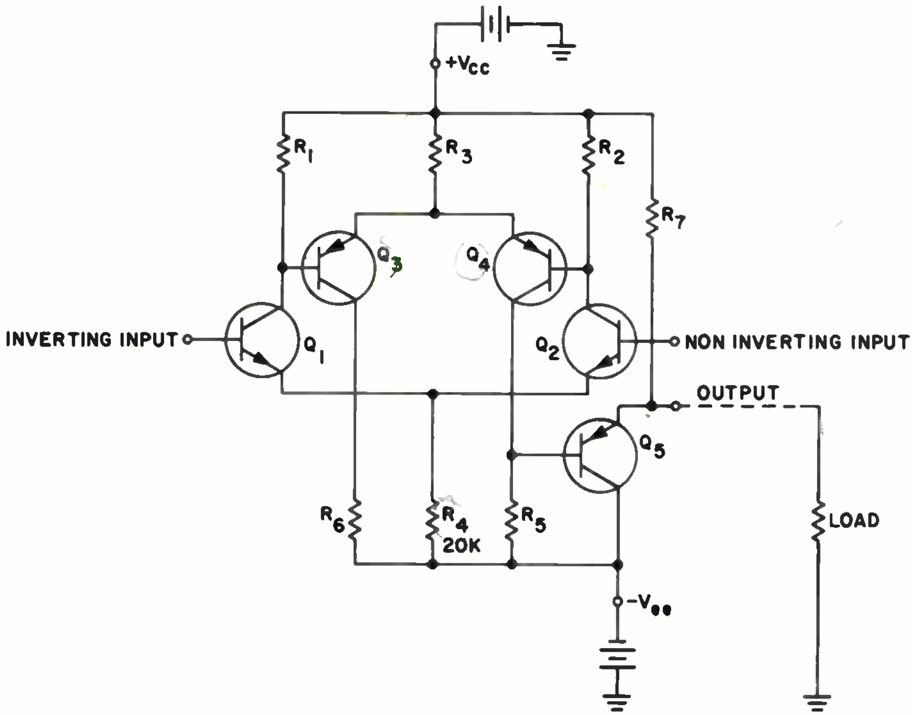


Fig. 23. A simple operational amplifier.

formed by a voltage $+V_{cc}$ and resistor R_3.

In this simple amplifier, NPN transistors are used for the first stage and PNP's are used for the second stage. The use of PNP transistors in the second stage automatically gives us the ability to level-shift the voltage. Remember that the voltage builds up only if we use transistors of the same polarity.

The collector voltages of the second stage are shifted to a level very near ground. The output of the second stage comes from the collectors of transistors Q_3 and Q_4. However, the output from the collector of transistor Q_3 is not used; it is simply returned through resistor R_6 to $-V_{ee}$. Meanwhile the output of transistor Q_4 goes to load resistor R_5 which is coupled to output transistor Q_5. As you can see from the diagram, the output stage in this op amp is made up of transistor Q_5 connected as an emitter-follower.

This amplifier, having most of the desired characteristics of an op amp, is suitable for general purposes. The transistors of the first stage must be matched very carefully in order to obtain a low input offset voltage and current.

To get an idea of the level of performance of this simple circuit, let's outline its characteristics: input impedance, 100K; input current, 600 nanoamps; equivalent input offset voltage, 1 millivolt; voltage gain, 73 decibels. This amplifier uses a simple differential amplifier for an input stage. However, in many cases a more sophisticated input stage design is required to increase the input impedance, increase the common mode rejection ratio or obtain lower values of input bias current.

The high input impedance and the low bias current are direct results of each other. To achieve a high value of common mode rejection ratio in both cases, an active current source is used to feed the emitters. The active current source is usually made up of a transistor and some resistors.

HIGH INPUT IMPEDANCE

In order to obtain a high input impedance, three methods are commonly used: (1) Darlington connected transistors in the differential input stage, (2) a field effect transistor differential amplifier input stage and (3) a standard transistor connection. We will now discuss these three approaches and the active current sources.

In Fig. 24 is an input stage using both a Darlington connected differential amplifier stage and an active current source. The current source is formed by transistor Q_5; resistors R_1, R_2 and R_3; and diode D_1. The purpose of the current source is to produce a constant current feeding the two emitters of the differential Darlington pair. This current should be stable enough not to change with temperature or bias conditions. Resistors R_1 and R_2 form a voltage divider taken from the power supply. The

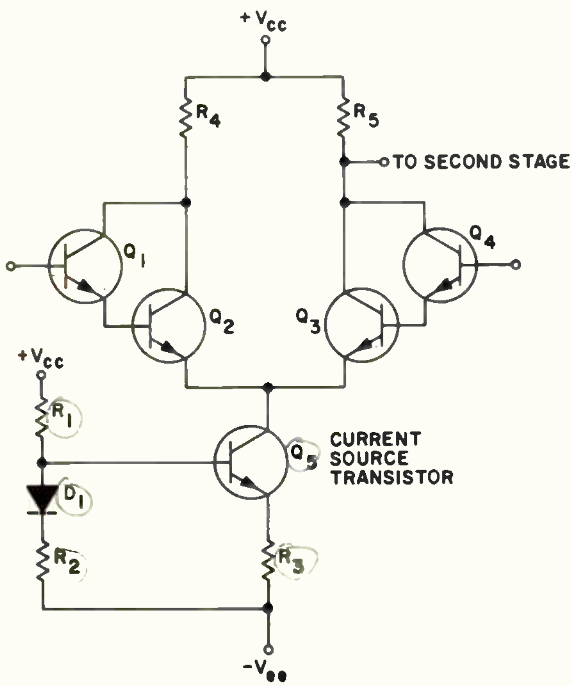


Fig. 24. A Darlington stage and an active current source.

voltage at the base of Q_5 is determined by this voltage divider. This means that the voltage drop across R_3 will also be determined by this divider, since the base potential and emitter potential will be the same. This now establishes a constant voltage across R_3 which implies a constant emitter current (emitter current = emitter voltage/R_3). The collector current is the emitter current multiplied by the alpha of the transistor. The output at the collector is a very high impedance, which simulates an ideal current source.

The use of a transistor as a current source gives us much better current stability than a high value resistor connected to the power supply. Diode D_1 is used in order to achieve a certain degree of temperature compensation. As you know, the emitter-base junction of a transistor has a negative temperature coefficient. That is, V_{be} decreases with an increase in temperature. Diode D_1 tends to cancel out the variation in temperature of the emitter-base junction by introducing an equal voltage with a polarity that cancels the effect. In this way, we obtain a more stable current with regard to temperature changes. For example, if the temperature increases, V_{be} will decrease, causing more voltage to appear across R_3 and increasing the current in Q_5. However, the drop across D_1 decreases, reducing the voltage applied to the base of Q_5 and the voltage appearing across R_3. Therefore, the current remains constant. We must be careful to match the two temperature coefficients of the emitter-base junction of the current source transistor with the one of diode D_1.

The differential amplifier stage in Fig. 24 is made up of Darlington pairs Q_1 and Q_2 on one side and Q_3 and Q_4 on the other. The effect is to achieve a much higher value of Beta. Therefore, for a certain emitter current, the amount of base current that flows into the input transistors will be much less than with a single transistor. This, of course, achieves a small value of input current and a much higher input impedance.

The use of the Darlington connection for the input stage of an op amp is limited because of the input offset voltage and the input offset temperature coefficient. The reason is that the input offset voltage is the net difference in V_{be} drops from one input to the other. Since we now have four transistors, it is more difficult to obtain a low value of input offset voltage. Also the V_{be} of the four transistors must be very closely matched for temperature coefficients, more so than if we had a standard differential pair.

In the input stage shown in Fig. 24, resistors R_4 and R_5 are the load resistors. The output from R_5 is used to drive a second single-ended stage. We could go to a differential second stage with PNP transistors as we did before, then convert to a single-ended output after the second stage.

Another method of obtaining a high input impedance is to use a standard differential amplifier stage, operate the transistors at a very low value of collector current and manufacture the transistors so that their Betas are extremely high at these low values of collector current. This is sometimes referred to as the super-Beta technique. It is very difficult to make transistors that have a very high Beta at low values of collector current. In order to get a high Beta, we must make the base of the transistors so thin that the reverse breakdown characteristics are severely restricted. Therefore, we must sacrifice the magnitude of our input voltage to low values in order to get the higher gain.

The use of field effect transistors to obtain high input impedance is shown in the complete op amp diagram of Fig. 25. This op amp offers a much higher performance than the first one we discussed. Transistor Q_5 forms the active current source that we discussed in the previous section. Diode D_1, the 3.3K resistor and the 224K resistor form the voltage divider used to establish the emitter voltage reference. As before, D_1 is used for emitter current stabilization with temperature. The collector of Q_5 is connected to the emitters of Q_2 and Q_4 and through 3K stabilizing resistors to the sources of field effect transistors Q_1 and Q_3. The two input signals are applied to the gates of transistors Q_1 and Q_3 which are N-channel field effect transistors. The sources of Q_1 and Q_3 are connected to the bases of Q_2 and Q_4, directly coupled.

As you know from previous lessons,

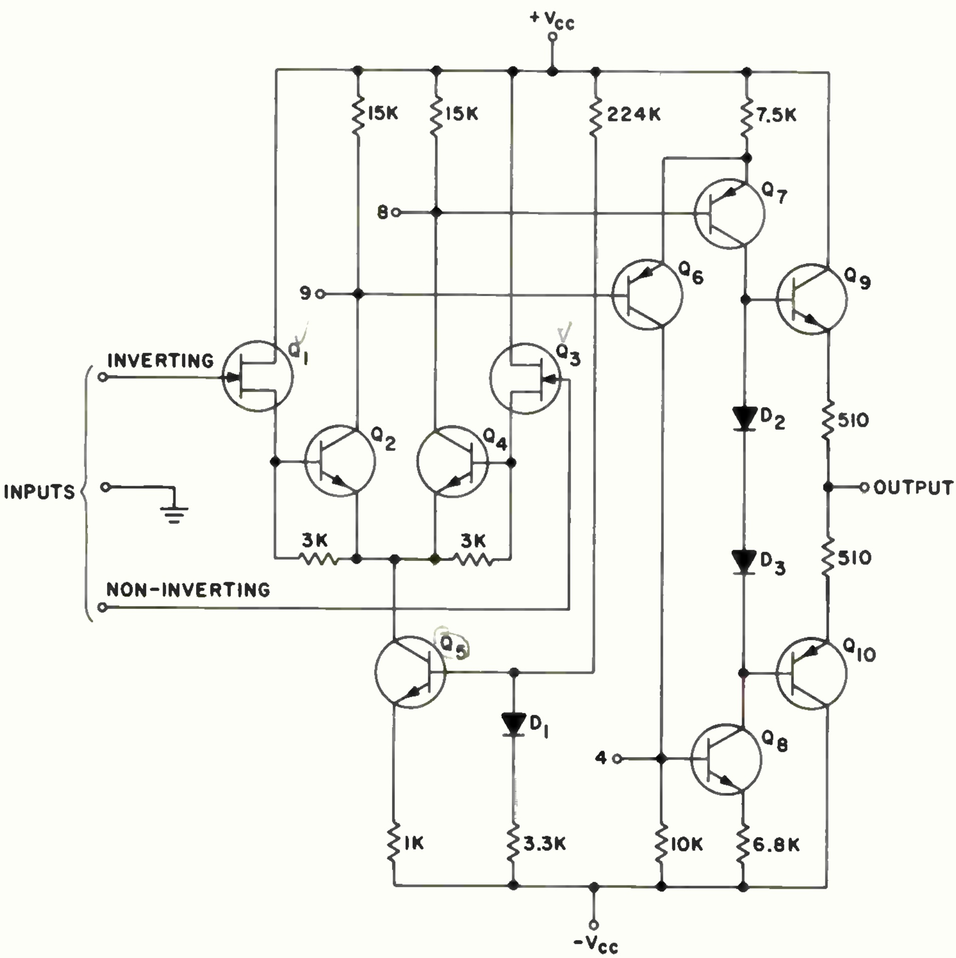


Fig. 25. Operational amplifier with FET's in input.

the field effect transistor is similar to a vacuum tube in its input impedance characteristics. You can typically obtain hundreds of megohms input impedance from a field effect transistor. Using this compound stage with field effect as well as bipolar transistors, the problems of input offset voltage drift and input offset voltage are quite difficult. The field effect transistors, in particular, do not have good tracking ability with temperature (for temperatures above about 70°C, the tracking is very poor). However, for commercial applications they give extremely low values of input bias currents and extremely high values of input impedance.

The 15K resistors in the collector circuits of Q_2 and Q_4 are the load resistors of the first differential stage. The two collectors of Q_2 and Q_4 are connected respectively to the bases of transistors Q_7 and Q_6 which form a modified differential amplifier. Transistors Q_6 and Q_7 are PNP devices. Here we use the complementary symmetry connection to obtain a dc level shift. This achieves direct coupling and level shifting at the same time. The output stage or power stage is formed by transistors Q_8, Q_9 and Q_{10}. The collector of Q_6 is connected to the base of Q_8, while the collector of Q_7 is connected directly to the base of Q_9. Q_9 and Q_{10} form a class B output stage. Q_8 is used to produce an out-of-phase signal into the base of Q_{10} so that we form a push-pull stage with Q_9 and Q_{10}. The output stage is essentially a complementary emitter-follower. Diodes D_2 and D_3 are placed between the two bases to reduce crossover distortion.

In a class B amplifier the transistors are biased to cutoff, conducting only on one-half cycle of the input driving signal. In a complementary symmetry class B amplifier, the positive half-cycle is amplified by the NPN transistor while the negative half-cycle is amplified by the PNP transistor. This is true for Q_9 and Q_{10} in Fig. 25; when one is conducting, the other is cut off.

With this arrangement, crossover distortion occurs. It is caused by the emitter-base characteristics of the transistors used. A silicon transistor requires a voltage of .6 to .7 volt across this junction before it conducts. With only .2 volt of forward bias, it is still cut off. The depletion layer in the junction effectively causes the transistor to be reverse-biased by about .6 to .7 volt. For that reason, a transistor is not operated exactly at cutoff as required for class B conditions. The result is that the input waveform must go more positive or more negative than .6 to .7 volt before the NPN or PNP transistor will conduct. During the zero crossing period of the input sine wave, neither transistor conducts, so the output signal appears as in Fig. 26.

This distortion can be eliminated by applying just enough forward bias to the transistors so that they operate right on the threshold of conduction, a condition required for proper class B operation. In Fig. 25 diodes D_2 and D_3 provide this forward bias. Each of these silicon diodes has a drop of .6 volt, for a total of 1.2 volts between the bases of Q_9 and Q_{10}.

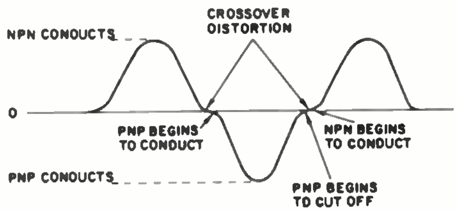


Fig. 26. Crossover distortion in a class B transistor amplifier.

This is just enough to overcome the depletion layer effects of Q_9 and Q_{10} in series. Now when Q_7 and Q_8 drive Q_9 and Q_{10} in the proper direction, they will conduct without producing the crossover distortion. The output is taken from the common emitters Q_9 and Q_{10}, through 510-ohm resistors which are there for short circuit protection.

In this amplifier we have seen several techniques. One of them is the use of field effect transistors in the input for high input impedance and low input bias current. Another is the use of an active current source. The second stage is formed of PNP transistors to obtain level shifting and some degree of gain. The last stage is a class B push-pull stage formed with a complementary pair of transistors.

To obtain an idea of the level of performance of the op amp just described, let's look at some of its characteristics: open-loop voltage gain, 100 db; input impedance, 100,000 megohms; input bias current, 20 picoamps; input offset current, 2 picoamps; input offset voltage, 10 millivolts; output voltage swing, 50 volts peak-to-peak. In this amplifier the voltage gain, input resistance and input bias current are considerably better than in the other one. However, the input offset voltage of 10 millivolts is not as low as it was in the previous amplifier.

IMPROVED OUTPUT STAGES

One of the most important op amp requirements is that it be able to produce a fairly high amount of output voltage into a low value of resistance. What this means is that a power amplifier must be used. The output stage in most op amps can supply a reasonable amount of power to a load, but for some applications, an external power booster amplifier must be used.

A typical booster output stage of an op amp is shown in Fig. 27A. This circuit is basically a current amplifier which provides wideband unity gain and peak currents up to ±200 milliamps into a 50-ohm load. Many common op amps can produce only a maximum of ±10 volts across a resistor of 3K or 4K. As you can see, this represents a very small amount of current. While using the technique that we will describe, currents of much higher values can be driven into a much lower load resistance. This configuration does not provide any voltage gain; the output voltage will be equal to the input voltage from an op amp.

The circuit in Fig. 27A is basically a set of cascaded complementary emitter-followers. It is made complementary so that the emitter-base voltage drops of Q_1, Q_3 and Q_2, Q_4 cancel each other. This makes the output voltage equal to the input voltage. The amplifier then has exactly unity gain. But because of its high input impedance and low output impedance, the circuit provides power amplification. When the output voltage from an op amp is applied to the input, the output voltage and polarity from the booster will be exactly the same. However, it is now capable of driving a larger current through the load than the op amp is.

Assume that the input to the booster circuit is at ground. This means that both transistors Q_1 and Q_2 become forward-biased. The voltage between the input and the emitter of Q_1 and Q_2 will be approximately .6 volt. This voltage is applied to the bases of Q_3 and Q_4. However, this voltage is insufficient to cause conduction of Q_3 and Q_4 so no current flows through the output load.

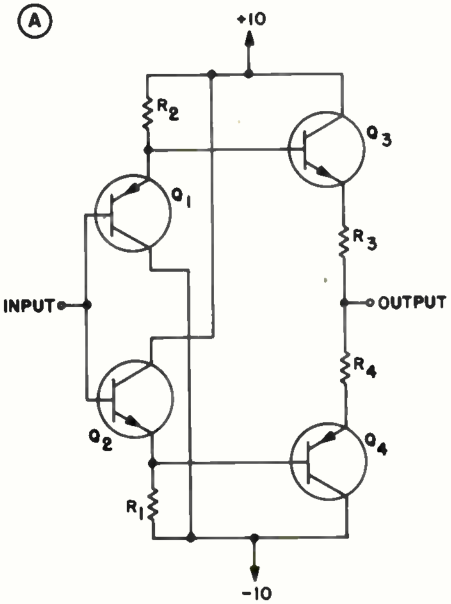


Fig. 27A. A current booster amplifier.

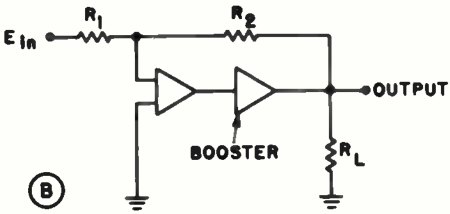


Fig. 27B. How the booster is normally connected to an op amp.

The output voltage at this time is essentially zero.

Now assume that we apply an input voltage of −5 volts. This causes transistor Q_1 to conduct more, producing a voltage of approximately −4.4 volts at the emitter of Q_1 and the base of Q_3. This voltage reverse-biases Q_3 so that it is cut off. The −5 volts at the input causes Q_2 to conduct less. The emitter voltage of Q_2 is approximately .6 volt more negative than the input, or −5.6 volts. This causes the emitter-base junction of Q_4 to conduct. Therefore, the output voltage is approximately .6 volt different, making the output −5 volts. The emitter-base voltage drop of Q_4 effectively cancels the emitter-base drop of Q_2, making the output voltage equal to the input voltage.

When +5 volts is applied to the input, Q_1 conducts less and the voltage at the base of Q_3 becomes approximately +5.6 volts. Q_3 conducts and its emitter is .6 volt less than its base, so the output voltage is +5 volts. As you can see, this is a unity gain dc power amplifier. Resistors R_3 and R_4 in the circuit are generally low in value so they do not contribute substantially to a change in the output voltage. They are used primarily for short-circuit protection. If the output accidentally short-circuits, the current flow through transistors Q_3 and Q_4 will be limited by R_3 and R_4.

For proper operation of this circuit, the emitter-base voltage characteristics of all the transistors must be closely matched so that their effects will cancel. The greatest error is introduced because the emitter-base voltage drops for NPN and PNP silicon devices are slightly different. For example, the V_{BE} of the NPN will be .6 volt; the V_{BE} of the PNP transistor will be .64 volt under the same conditions of bias. The difference would be .04 volt. This small error is generally of little consequence. Its effect can be minimized by using the booster amplifier within the feedback loop of the op amp to which it is connected. Fig. 27B shows how the booster is normally connected. The output of an op amp is fed into the booster and the load connected to the booster output. Instead of the normal feedback resistance being connected from the inverting input to the op amp output, the normal connection to the op amp output is connected to the booster. The small offset voltage produced by the imperfectly matched transistors is somewhat overcome by the use of this negative feedback.

CHOPPER STABILIZED AMPLIFIERS

A chopper stabilized amplifier is a dc amplifier in which modulation techniques are used to minimize the effects of drift. Variations in the temperature, power supply voltages and component characteristics will cause the output voltage of a dc amplifier to drift. This is an undesirable condition as it introduces an error in the output. Careful circuit design and accurate matching of components aid in minimizing drift. However, for critical applications it must be reduced further. To do this, the dc input to be amplified is passed through a chopper where it is converted into a square wave in the audio frequency range. This square wave is amplified by a standard ac amplifier. The ac output then is rectified, or converted back into dc, by the chopper; the resulting dc voltage is amplified in the standard dc amplifier.

It is the drift in the input stage of a dc amplifier that gives us the most trouble. The small drift voltage that occurs will be amplified many times by the remaining stages of amplification. Therefore, even a small input drift voltage can appear as a large offset at the output. If we can minimize the effect of drift in the input stage, then any drift in the remaining amplifier stages will have a minimum effect on the output voltage. We do this by converting the dc input voltage to ac early in the amplifier, where drift has no effect. We then convert this back to a dc signal of substantial amplitude and feed it the rest of the way through the dc amplifier.

Most op amp applications can make use of the standard unstabilized circuitry. However, for very critical applications, chopper stabilization is used to reduce drift to practically zero.

Fig. 28 shows a chopper stabilized op amp. The dc portion of the amplifier consists of the differential stage made up of Q_1 and Q_2. Transistor Q_3 and its associated components make up a temperature-compensated current source for the differential input stage. The output of the differential stage is taken from the collector of Q_2 and fed to a complementary amplifier stage, Q_4. Additional gain is produced in this stage and its output is coupled to transistor Q_5 and the emitter-follower output circuit.

The input signal applied to this circuit is actually fed to two places. First the high-frequency components of the input are passed through capacitor C_5 to the differential input stage. Any dc at the input is eliminated by capacitor C_5. The values of capacitor C_5 and resistor R_3 are chosen to act as a high-pass filter that passes only frequencies above a certain cutoff point. All dc and very low-frequency ac signals are passed through R_1 and C_1, a low-pass filter, to a chopper.

As you recall from a previous lesson, a chopper is nothing more than a switch that alternately connects and shorts out the dc signal at the input of an ac amplifier. In this circuit transistor Q_6 is used as a shunt switch. When Q_6 is open,

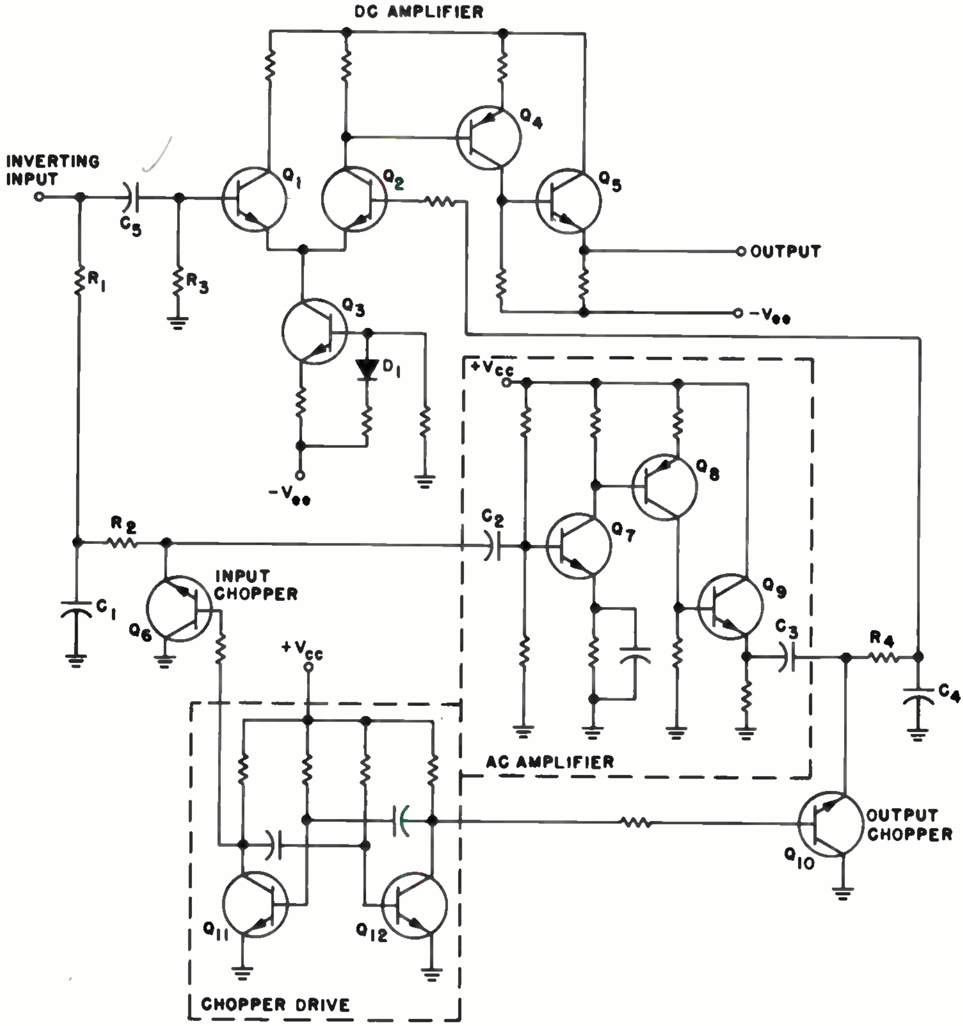


Fig. 28. A chopper-stabilized amplifier.

it has no effect on the circuit and the input voltage is applied directly to the ac amplifier through C_2. When Q_6 conducts, it shorts out the signal so that none is applied to the ac amplifier. Transistor Q_6 is turned off and on very rapidly by a free-running multivibrator circuit used as a chopper driver. This consists of transistors Q_{11} and Q_{12} and their associated components. The outputs at the collectors of Q_{11} and Q_{12} are square-wave signals that switch between ground and $+V_{CC}$. This is used to turn transistor Q_6 off and on. A rate of about 400 Hz is used.

Notice that in this circuit Q_6 is connected in a configuration inverted from that normally used. To turn a transistor on, we usually forward-bias its emitter-base junction. In this circuit we are forward-biasing the collector-base junction. Most transistors are somewhat symmetrical; the collector can be used as the emitter and vice versa. By using this connection, there will be an extremely low voltage drop between the emitter and collector when the transistor is turned on. This permits the transistor to function as an almost perfect on/off switch. When it is on, its resistance is practically zero; when it is off, almost infinite.

The 400 Hz signal appearing at the emitter of Q_6 is passed through the ac amplifier made up of Q_7, Q_8 and Q_9. Here it is amplified and fed to the output chopper, which is another inverted transistor, Q_{10}. Q_{10} effectively rectifies the ac output while R_4 and C_4 filter it into a dc voltage to be applied to the other input of the differential amplifier. Here the dc signal undergoes further amplification in transistors Q_2, Q_4 and Q_5.

Because of the tremendous improvements recently made in electronic components and circuit techniques, chopper stabilized amplifiers are actually being used less. It is now possible to obtain closely matched transistors and stable power supplies that reduce drift effects to a point where chopper stabilization is unnecessary. In fact, many of today's integrated circuit op amps have characteristics which approach those of a standard chopper stabilized amplifier. However, for extremely critical operations, chopper stabilization may be used to completely eliminate the effects of drift.

SELT-TEST QUESTIONS

(x) Name three ways a high input impedance can be obtained in an op amp.

(y) What is the purpose of a booster amplifier?

(z) What transistor characteristic causes crossover distortion in a class B amplifier?

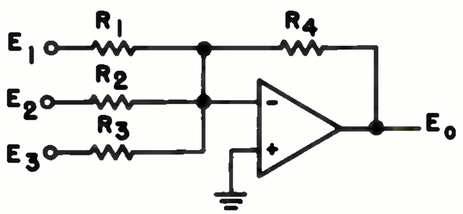
3-18-76

Common Uses

Today amps are used in a wide variety of ways. They appear in almost every phase of electronics because they are capable of performing so many useful functions. The op amp is a key element in the analog computer because it can perform so many useful mathematical calculations. It can perform multiplication and division by a constant simply by selecting the gain. It can also add or subtract. Since one of the major uses for op amp is performing mathematical operations, we will show you several of these useful applications.

MATHEMATICAL OPERATIONS

A typical op amp and summer is shown in Fig. 29. This circuit is basically that of the simple inverting op amp you studied before. The only difference is that we have added more than one input resistance. In this case we have added three, and to each supplied an input voltage. The op amp will sum the three input voltages after multiplying each by a gain that is the ratio of the feedback resistance divided by the associated input resistance.



$$E_o = -\left[\frac{R_4}{R_1}(E_1) + \frac{R_4}{R_2}(E_2) + \frac{R_4}{R_3}(E_3)\right]$$

Fig. 29. An op amp summer.

The mathematical output expression is shown in Fig. 29. Notice that the negative sign indicates that the sum is inverted. This amplifier circuit can be used to perform algebraic addition or subtraction operations. Any number of input resistances may be connected to the summing junction.

To show how this circuit works, let's assume some values for R_1, R_2, R_3 and R_4 in Fig. 29. Assume that all the resistors are 100K. Next assume the input of E_1 equals 2 volts, E_2 equals 3 volts, and E_3 equals 4 volts. Now, using all of these values, what is the output voltage? Find its magnitude and its polarity.

To solve this problem, all you have to do is fill in the appropriate values in the output expression shown in Fig. 29. We will do this after trying to understand exactly what is happening in this circuit. Keep in mind that each of the input voltages is going to be multiplied by a coefficient that is the gain of the amplifier for that input. Recall that the gain of the op amp is a function of its feedback and input resistance ratio. However, since we have made each input resistor the same size as the feedback resistor, this ratio will be 1:1. Therefore, in this circuit the output voltage is equal to:

$$E_o = -\left[\frac{100K}{100K}(2) + \frac{100K}{100K}(3) + \frac{100K}{100K}(4)\right]$$

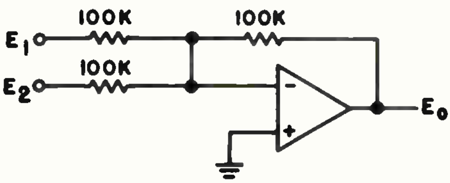
$$E_o = -\left[1(2) + 1(3) + 1(4)\right]$$

$$E_o = -(2 + 3 + 4) = -9$$

The solution to this problem simply involves basic addition. The amplifier performs the addition. Just keep in mind that the amplifier circuit still inverts and that this inversion must be considered when the absolute algebraic or polarity of the signal is important.

In a previous section you were introduced to the differential op amp circuit that is used to take the difference between two input voltages. This circuit could also be used to perform subtraction. However, we can also use a summer circuit to perform subtraction.

Fig. 30 shows a two-input summer that can be used for subtraction. Notice in this circuit that both the input and feedback resistances are equal so that the input voltages are multiplied by 1. To use the summer to subtract, we must feed voltages of opposite polarity into the two inputs. The op amp produces the difference between the two input signals as an output with a polarity opposite to the larger voltage.



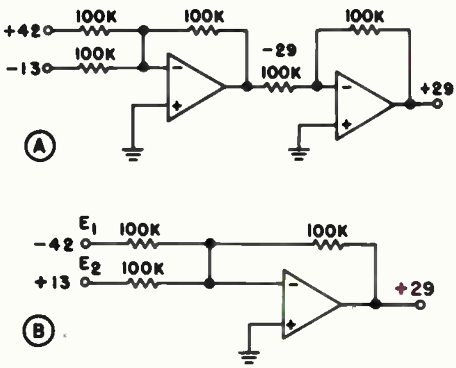
$E_0 = -(E_1 + E_2)$

-27 +18

Fig. 30. Two-input summer for subtraction.

As an example, let's subtract 18 from 27. E_1 will be -27 volts and E_2 will be 18. Since the polarities of the two voltages are different, the amplifier is going to produce an output voltage equal to the sum of the two (-9 volts). Because the larger voltage is negative, the output voltage will be inverted by the amplifier to become positive. The op amp performs algebraic addition; with it we can add numbers of any polarity. When we add numbers of opposite polarity we are producing subtraction. In other words, the process of subtraction is a form of algebraic addition. This op amp circuit inverts so that the polarity of the output voltage is significant. Therefore be sure to invert the output voltage that you observe to obtain the true algebraic value.

By properly assigning the polarities of the input voltages the inverting characteristics of the amplifier can be used to provide proper polarity output. For example, if we wish to solve $42 - 13$, one way is to make 13 a negative voltage and 42 a positive voltage. As in Fig. 31A, the circuit will produce an output voltage equal to -29. Note that the input and feedback resistances are all equal to one another so that no coefficient multiplication takes place. Although this circuit produces the proper magnitude output voltage of 29, the polarity is algebraically incorrect. The reason for this is simply the inversion of the op amp. We can correct the situation by passing the signal at the output of the summer through a simple unity gain inverter. This will produce the proper output voltage without changing the magnitude.



$-(E_1 + E_2) = -(-42 + 13) = (42 - 13) = 29$

Fig. 31. Two ways of producing correct algebraic subtraction.

However, we can produce the proper polarity output voltage with a single summer amplifier if we correctly assign the polarities of the input voltages. This is shown in Fig. 31B. The two input voltage magnitudes are the same as before, 42 and 13. However, we make the larger of the two voltages the negative so that the polarity of the inverted output will be correct. This way we let the inverting characteristics of the amplifier work for us.

CALCULATING OP AMP NETWORKS

By now you should be able to calculate the output voltage and polarity of an op amp network. Knowing the values and polarities of the input voltages and the values of the input and feedback resistances, you should be able to take a circuit like that in Fig. 32 and compute the output voltage and polarity. Notice that this circuit uses a wide variety of op amp configurations. While at first this network may seem somewhat complex, keep in mind that it uses nothing but the simple individual op amp circuits that you have already studied. By breaking the problem down and calculating the voltage at each point in the circuit, it is very easy to arrive at the correct output voltage and polarity. Let's go through this circuit and calculate the output voltage, assuming the input voltages shown.

Amplifier 1 is a simple inverting amplifier whose gain is the ratio of the feedback to input resistors. This ratio is 40K/25K = 1.6. This amplifier multiplies the input voltage by a gain of 1.6. Multiplying our input voltage of 40 by 1.6 gives us an output voltage of 64 volts. Keep in mind that the amplifier inverts so the output is negative since the input voltage is positive. This means that −64 volts appears across the potentiometer. This pot is set to a coefficient k of .2. This simply indicates that 20% of the voltage applied to the pot is tapped off and appears at the arm. To find the voltage at the arm we simply multiply −64 by .2 to get −12.8 volts. This voltage is applied to one end of a differential amplifier. Notice the differential amplifier uses all 10K resistors so that no gain is provided. Instead, the output voltage will be the difference between the two input voltages.

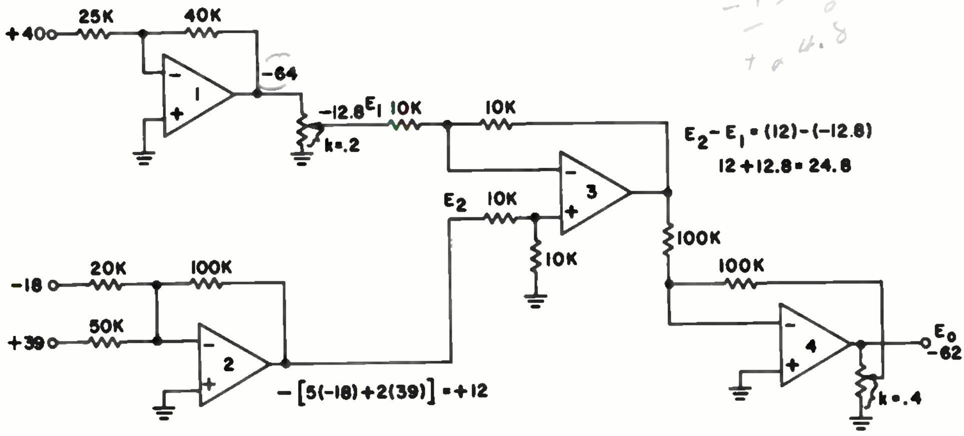


Fig. 32. An analog computing circuit using a variety of op amp circuits.

Amplifier 2 is a summer to which two input voltages are applied. The resistor values in the circuit provide gain multiplication for each input. Again the gain of each input is equal to the ratio of the feedback to the input resistor. The gain of the upper input is 100K/20K = 5. This means that the input voltage, −18, is multiplied by a gain of 5. The gain of the lower input is 100K/50K = 2, meaning that the input voltage, +39, is multiplied by 2. Making the calculation using the standard output formula, we get an output voltage of 12.

Now we have both inputs to the differential amplifier and we can calculate the output. Recall from your knowledge of this circuit that the output is equal to $(E_2 - E_1)$, or the lower input voltage minus the upper input voltage. Plugging our two input voltage values into this formula, we find that the output voltage is 24.8. Notice that since we are subtracting two input voltages of opposite polarity, the overall effect is algebraic addition. The output of amplifier 3 then is a positive 24.8 volts. This is now fed to amplifier number 4, where it is further multiplied in gain. Notice that the circuit used here is the one whose gain is determined by the coefficient of the output potentiometer. The gain of this circuit is equal to 1/k = 1/.4 = 2.5. The output voltage then is the inverted product of 24.8 multiplied by 2.5, or 62 volts.

INTEGRATOR CIRCUITS

In several previous lessons the subject of integrators was discussed. An integrator, in its simplest form, is a resistor-capacitor network whose time constant is made long with respect to the period of the input signal applied to it. This circuit, then, produces an output that is a function of the mathematical integral of the input signal. Integration is a mathematical function; specifically it is a form of calculus. However, we'll take a simpler approach (eliminating a calculus discussion) toward understanding the integrator circuit.

A simple integrator circuit is shown in Fig. 33. The input voltage is applied across a resistor and capacitor connected in series and the output voltage is taken from across the capacitor. If the input signal applied to this integrator network is a sine wave, the output signal will also be a sine wave. Because of the reactance in the circuit, the output voltage will be out-of-phase with the input voltage. Specifically, the output voltage will lag the input voltage by some angle between zero and 90°. The longer the time constant of the circuit, compared to the period of the incoming signal, the closer the output voltage approaches the 90° lag point. At this time the reactance of the capacitor is generally so low, compared to the resistance value, that the output

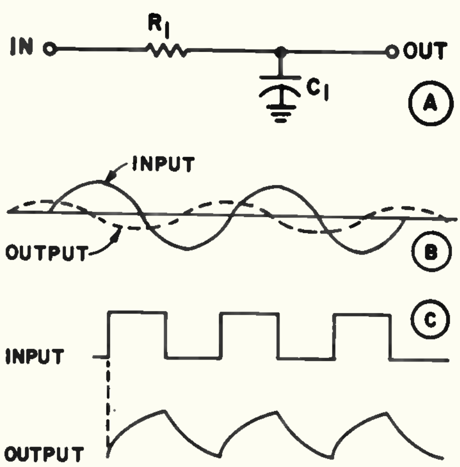


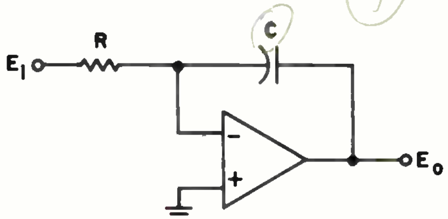
Fig. 33. A simple R-C integrator (A), sine wave input/output (B) and square wave input (C).

voltage is much smaller than the input voltage. Theoretically, if the phase shift were a total of 90°, the output voltage would be zero. However, a practical integrator is one whose phase shift approaches 90° closely, but also one that still has a measurable output voltage.

If the input voltage applied to an integrator is a square wave like that shown in Fig. 33C, the output voltage will be a triangular or sawtooth type waveform. The larger the time constant, the closer the output approaches a linear sawtooth or triangular waveform.

The integrator circuit shown in Fig. 33 is a useful network. It can be used to perform phase-shifting and wave-shaping operations. However, it is limited in that its output voltage is much smaller than its input voltage and it does not provide perfect integration. In many applications it is desirable to have exactly a 90° phase shift, or a perfectly linear sawtooth or triangular wave output. The op amp integrator that we will describe next can be used for such applications.

A typical op amp integrator is shown in Fig. 34. This is the standard op amp configuration except that we use a feedback capacitor instead of a feedback resistor. The result is a nearly perfect



$$E_O = -\frac{E_I \cdot t}{RC}$$

Fig. 34. An op amp integrator.

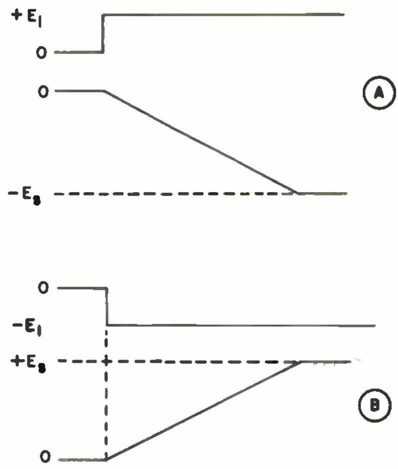


Fig. 35. Integrator input-output waveforms.

integrator circuit. Disregarding amplifier phase shift, the output is shifted exactly 90° when a sine wave is applied to the input of this circuit. The output amplitude can be adjusted to a desired level by choosing the proper values of resistance and capacitance. When a square wave or dc step voltage is applied, the output is nearly a perfect linear sawtooth or ramp waveform. The rate of change of the output voltage depends upon the input voltage and the gain of the integrator which is determined by the R-C values in the circuit.

Now let's take a look at the operation of the op amp integrator circuit of Fig. 34. We are going to apply a step voltage to the input where the voltage switches rapidly from zero to some dc voltage level and remains there. The output waveform of the integrator, under these conditions, is shown in Fig. 35. In Fig. 35A, both the input voltage (E_1) and the output voltage are zero.

When the input voltage steps rapidly to the positive voltage level E_1, the output begins to rise slowly in a negative direction. The change of voltage is linear with respect to time, producing a straight negative-going ramp. As long as the input voltage is held at the E_1 level, the output will slowly rise due to the charging of the feedback capacitance through the input resistor. Soon we will reach a point where the output voltage can no longer continue to increase due to the saturation of the output amplifier of the op amp. When this occurs the output levels off at an output voltage $-E_S$, which is equal to or slightly less than the negative power supply voltage of the op amp.

Fig. 35B shows the input and output waveforms for a negative-going step voltage input. When the input voltage steps to a negative E_1 level, the output voltage begins to rise linearly from zero. As long as voltage is applied, the output will increase in a straight line to eventually reach the saturation point that is nearly equal to the positive supply voltage of the op amp.

The value of the output voltage is very predictable if we know the values of the input resistor, the feedback capacitor, the input voltage and the time during which the input voltage is applied. All of these factors are related by the formula:

$$E_{out} = -\frac{1}{R\text{-}C}(E_1)(t)$$

This formula tells us the value of the output voltage at a specific time (t) after input voltage E_1 is applied. The 1/R-C value is the gain of the integrator stage. The gain of this circuit is equal to the reciprocal of the time constant. If we are using a 1 mfd capacitor and a 1 megohm input resistance, the integrator gain is $1/(10^6 \times 1 \times 10^{-6}) = 1$. With a gain of 1, the output voltage with a step voltage input is simply equal to $-E_1(t)$.

Let's assume that we have an input voltage of +10 volts and that we allow the circuit to integrate for a period of 5 seconds. To determine the output voltage, we multiply the input voltage by the time and invert the result. In this case, $E_{out} = -10 \times 5 = -50$ volts. We can stop the integration after 5 seconds by removing the input voltage or by shorting the capacitor.

Fig. 36 shows the output waveforms you can expect. The waveform at A shows the step voltage input that switches from zero to +10 volts. The output voltage begins to increase in negative value. If the input is allowed to be present for 5 seconds, the output voltage will go to −50 volts. If we remove the input voltage at this point, the output will remain at its −50 volt level. This is

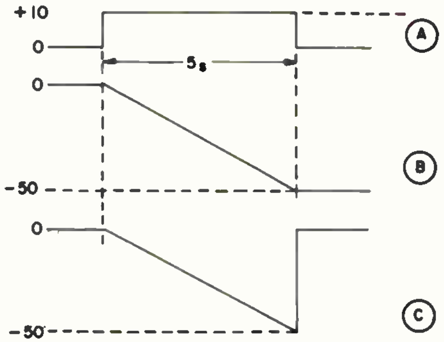


Fig. 36. Integrator step input (A), output where input is removed after 5 seconds (B) and output where feedback capacitor is shorted after 5 seconds (C).

because the capacitor retains the charge it received (−50 volts) while the input was connected. Over a long period of time, the charge on this capacitor will leak off due to the internal resistance of the capacitance, the input impedance of the op amp and other factors.

Fig. 36C shows the result of shorting the output feedback capacitor after a 5 second period. The output voltage again goes negative to reach a value of −50 at the end of 5 seconds. If we short the output capacitor with a switch of some type, the capacitor discharges quickly through the short circuit. The output voltage drops to zero despite the fact that the input is maintained at this value.

What happens to the output voltage when we change the gain of the amplifier? Let's change the values of the input resistor and feedback capacitor and note the effect this has on the output voltage when we integrate the +10 volt dc input for a period of 5 seconds. Assume that our new input resistance is 100K-ohms and our feedback capacitor is .1 mfd or 10^{-7}. The R-C time constant then is $10^5 \times 10^{-7} = 10^{-2}$. Then the gain of the integrator is $1/\text{R-C} = 1/10^{-2} = 1/.01 = 100$. We now multiply our input voltage by 100 and the time the input voltage appears at the integrator. The output voltage should theoretically rise to a value of $10 \times 100 \times 5 = 5000$ volts. However, this is an unreasonable value because of the output voltage swing limitations. Even if this is a high voltage amplifier where the output voltage swing may be ±100 volts, the output will quickly rise and saturate at its negative level shortly after the input voltage is applied.

The gain of an integrator, the input voltage value and the time the integrator is allowed to observe the input voltage determine the rate at which the output voltage changes. If the gain and the input voltage are high, the output voltage changes rapidly. If we are not careful to control the input voltage, gain and time duration of integration, it is quite easy to saturate the amplifier or to cause it to try to produce an output voltage beyond its capabilities. Fig. 37 shows the output

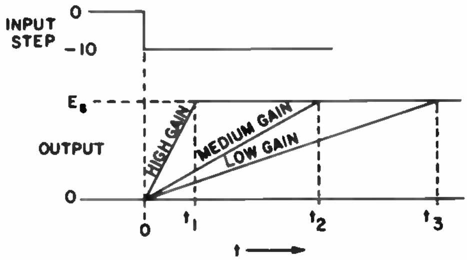


Fig. 37. How integrator gain affects the output ramp.

voltage of an integrator with a −10 volt input step. The outputs produced with high, medium and low gains are shown. The high gain is a result of a short R-C time constant. The capacitor charges quickly, so the output voltage rises rapidly and saturates at time t_1. Making the R-C time constant longer lowers the gain. Since it takes the capacitor longer to charge, saturation doesn't occur until time t_2. A very low gain, meaning a long time constant, produces a very gradual output ramp that reaches saturation at t_3.

Fig. 38 shows how the op amp can be used with sine wave signals. If we assume that the feedback capacitor is 1 mfd and the input resistor is 1 megohm, the gain of the op amp is $1/\text{R-C} = 1/(10^6 \times 10^{-6}) = 1$. This means that the output will be equal to the input. However, because the circuit is an integrator, the input and output waveforms will be out-of-phase. A perfect integrator would produce an output that lags the input by 90°. Keep in

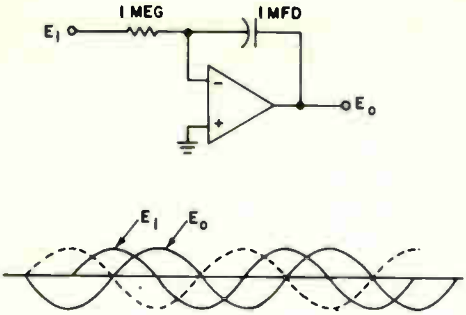


Fig. 38. Integrator operation with sine wave signals.

mind that our op amp integrator also inverts the signal; the output is actually shifted by 90° and then inverted 180°.

The input and output voltages of the integrator are also shown. The output voltage, as it would appear without inversion, is marked by the dashed line. As you can see, it lags the input signal by 90°. However, inversion adds another 180° phase shift, which causes the output voltage to appear as though it is leading the input signal by 90°. We call the output waveform a cosine waveform to distinguish it from the sine wave input. By adjusting the values of input resistance and feedback capacitance, the gain of the amplifier can be adjusted to produce any desired output amplitude within the amplifier's capability.

ELECTRONIC COMPARATORS

An electronic comparator is a circuit that compares one voltage with another and produces an output signal that indicates when the two input voltages are equal. A good electronic comparator will also produce an output signal that will permit you to know whether one of the input signals is less than or greater than the other. One of the input signals to the comparator is the reference signal. It is the standard by which we will compare another input signal, generally a changing or varying signal.

The simplest form of a comparator is the diode circuit shown in Fig. 39. A diode is biased through a resistance with a 20-volt battery which represents the reference voltage. We are going to compare this reference voltage with the input signal, applied to the cathode of the diode that is a positive-going ramp voltage. This could possibly be the output from an integrator being driven from a negative dc source.

With the input voltage at zero, the diode is forward-biased. If we assume this to be a perfect diode that has no voltage drop across it, the output voltage will also be zero. As the input voltage begins to rise, the diode still conducts; therefore, the input is effectively connected directly to the output through the diode.

The output exactly follows the shape of the input signal. The output remains the same as the input only while the input voltage is less negative than the reference voltage, keeping the diode forward-biased. As soon as the input voltage reaches a value of 20 volts, the diode no longer conducts; there is insufficient voltage across it to cause it to be forward-biased. As the input voltage rises beyond 20, the diode is still cut off so the output voltage remains at 20 volts. The

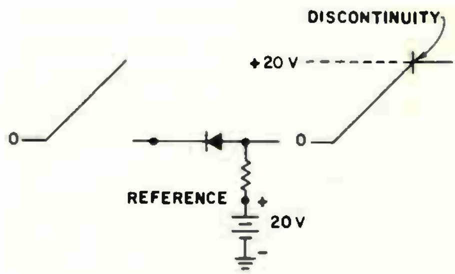


Fig. 39. A simple comparator circuit.

output, in effect, sees the 20 volts from the reference battery through the resistor.

The output signal flattens off at this point. The point where the waveform stops rising in voltage and flattens out is known as a discontinuity. As long as the input voltage is less than the reference voltage, the output follows the input. However, when the two are equal, a discontinuity in the output occurs. This discontinuity represents the point where the reference and input voltages are equal. Beyond this the output waveform flattens out, indicating that the input voltage is greater than the reference.

While the circuit of Fig. 39 is definitely a comparator, its usefulness is somewhat limited. In a practical circuit, there is a finite voltage drop across the diode. For a silicon diode this drop may be .6 or .7 volt. As a result the diode will not stop conducting until the input voltage is approximately .6 or .7 volt above the reference voltage. For this reason there is a substantial error in the comparison. This is particularly true if the reference voltage is a small voltage that is the same order of magnitude as the diode voltage drop. For very large voltages, the percentage error would be much smaller.

In addition, the output waveform is not in the most easily used form. In a practical circuit, the discontinuity point is not a sharp and clearly defined point. Because of the stray capacitance in this circuit, the point is rounded and introduces some ambiguity in the exact point of equality. This circuit is adequate for simple comparisons, but for critical comparisons, a more sophisticated circuit is required.

Consider for a moment how we may use a differential amplifier as a comparator. A differential amplifier is normally designed to handle dc signals. Therefore we may apply the reference voltage to one of the differential inputs. Since the output of the amplifier is equal to the difference between the two inputs multiplied by the gain, applying a varying signal to the other input will make the differential amplifier output indicate when the two are equal. When the two signals are equal, the output of the amplifier will, of course, be zero. The higher the gain of the amplifier, the larger the output voltage swing will be for a small difference between the two input signals. Remember that this circuit amplifies the difference voltage between the two input signals.

According to this principle an op amp operated in the differential mode with no feedback should provide excellent comparison. By using the full open-loop gain of the amplifier, only millivolts (or perhaps microvolts) of difference signal is required to cause the amplifier output to swing between its two maximum output voltage limits.

Fig. 40A shows an op amp connected as a comparator. Notice that absolutely no feedback connection is provided. The reference signal is connected to one of the inputs while the varying signal, to be compared to the reference signal, is applied to the other input.

If we use the input signal assumed earlier for the simple comparator, we can get some idea as to the sensitivity of this circuit. When the input voltage at the inverting input is 0 volts, the reference being 20 volts, the difference signal is 20 volts. With this much input signal multiplied by the full open-loop gain of the amplifier, the output will swing to its positive saturation level.

With an open-loop gain of 10,000, the output voltage would try to swing toward a value of 200,000. However, the ampli-

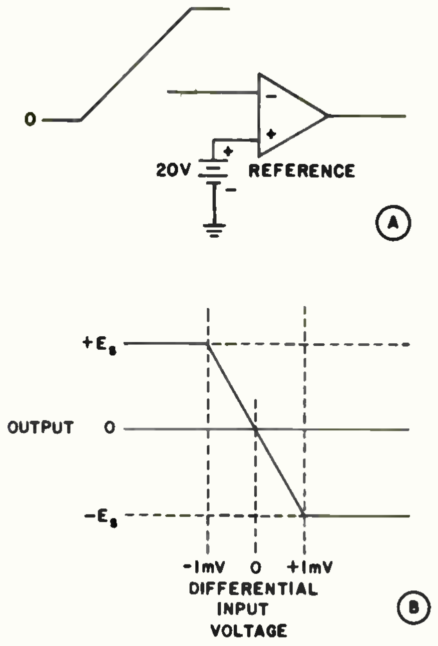


Fig. 40. An op amp comparator (A) and its output voltage (B).

fier will simply saturate at its positive output swing point. You can see this in the graph of Fig. 40B. As the input signal begins to rise from zero in a positive direction, the difference voltage between the two inputs grows smaller. However, during most of the rise, the difference voltage between the two is still large enough to keep the amplifier output in saturation.

Soon we begin to reach a point where the input voltage approaches the value of the reference. When the input voltage becomes equal to the reference voltage, the output of the amplifier will be zero. As you can see from the graph in Fig. 40B, the output voltage drops from its positive saturation level to zero when the two voltages are equal.

As the input voltage continues to swing in a positive direction, it becomes greater than the reference voltage and the difference signal between the two is amplified. As soon as the difference signal reaches a large enough value, the output of the amplifier swings to the negative saturation region. As you can see from the curve, the gain of the amplifier is so high that it takes only a 1 millivolt difference between the two input signals to cause the amplifier output to swing to one of its saturation levels.

For that reason the sensitivity of this circuit enables us to detect the equality between two signals within very close ranges. With only 1 millivolt of error, it is possible to compare very low level signals with good precision. When the output of the comparator sets at one of its saturation levels, it means that the input signal is either greater than or less than the reference voltage. The other saturation level is reached whenever the input signals are again unequal. When the two input signals are equal, the output of the comparator is zero. Simply by looking at the output signal, you can tell whether the input voltage is less than, equal to, or greater than the reference voltage.

There are some very important facts to keep in mind here. The rapid switching of the op amp output is due to its high gain or sensitivity. Remember that we are amplifying the difference voltage between the two inputs. With the high open-loop gain, the output can swing from its negative to positive output saturation level with only an extremely small difference signal. For most typical op amps, a difference voltage in the millivolt region is sufficient to cause the amplifier to swing from one extreme to the other. There is some error introduced, but it is extremely small; it is negligible for almost all practical electronic circuits.

The sensitivity of the op amp com-

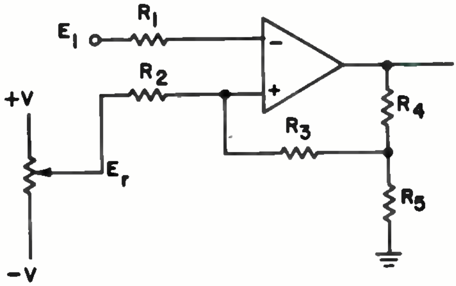


Fig. 41. Adding feedback to raise the trigger level and minimize the effects of noise.

parator can also be a disadvantage. Many times noise signals appearing on the inputs can cause the comparator to trigger falsely, producing undesirable output voltage transitions.

To overcome the noise sensitivity of the comparator, we can add feedback to the circuit. Fig. 41 shows a differential op amp comparator where the input voltage E_1 is applied to the inverting input through R_1. The reference voltage is obtained from a potentiometer that is connected to both positive and negative power supplies. The arm of the pot can be set to any voltage between the upper (positive) and lower (negative) voltage levels and still include the zero volts at the center of the pot. This provides a convenient means of adjusting the reference voltage level for the circuit. The reference voltage is applied through R_2 to the non-inverting input of the op amp.

Let's assume that the op amp has an initial sensitivity of 10 millivolts. This means that a change or difference of 10 millivolts between the two inputs will cause the amplifier to switch from one output level to the other. If noise exceeding 10 millivolts appears on the input line, the comparator output will trigger erratically and give us a false indication of the comparison of the input and reference signals.

To get a true comparison, we apply feedback to the non-inverting input. A small portion of the output voltage is tapped off by the voltage divider made up of R_4 and R_5. The voltage is fed back through R_3 to the non-inverting input along with the reference voltage. Here the feedback voltage effectively adds to the reference voltage, increasing the amount of difference voltage required to cause the amplifier to trigger. If we feed back 15 millivolts from the output, the overall threshold for the circuit becomes 10 + 15 or 25 millivolts. This means that the difference between the two input signals must be at least 25 millivolts in order to cause the output voltage to switch from one level to the other.

This introduces a little more error in our comparison. However, the increased voltage difference required for output switching minimizes the effect that noise has on the circuit. In other words, the circuit can now tolerate a higher level of noise on the input without producing erratic switching. The feedback effectively reduces the sensitivity of the comparator but improves its noise immunity. By replacing resistors R_4 and R_5 with a potentiometer, the sensitivity can be made adjustable.

Fig. 42 shows another way of using an op amp as a comparator. The non-inverting input is connected to ground as it normally is in a standard op amp

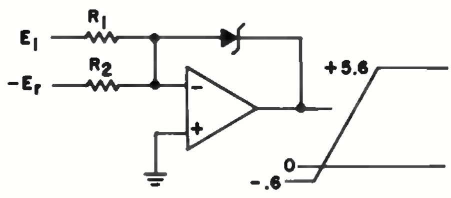


Fig. 42. Comparator made with an op amp summer.

configuration. The input and reference voltages to be compared are applied to the inverting input through summing resistors R_1 and R_2. In order for this circuit to properly perform algebraic addition (subtraction), the polarities of the two voltages to be compared must be opposites.

In the comparator circuit you studied in Fig. 40, the polarities of the input and reference voltages were the same. If the input voltage was positive, the reference voltage would also be positive. In the same way, a negative input voltage and a negative reference could also be used to obtain proper operation. Here we are effectively using the algebraic addition characteristics of the op amp to compare the two input signals. When the two input signals E_1 and $-E_R$ are equal in magnitude, the sum of the inputs at the inverting input will be zero, making the output zero. If the input voltage E_1 is greater in magnitude than the negative reference voltage, the resulting input is positive. This causes the output to swing to its negative saturation level. When E_1 is smaller, the net input voltage is negative and the output voltage swings to its positive output extreme.

In this circuit we are using a Zener diode in the feedback path. This prevents the amplifier from swinging between its two saturation levels, limiting the output voltage levels. When the output swings in the negative direction, it causes the Zener diode to be forward-biased. The output voltage is limited to .6 volt. When the output voltage swings positive, the Zener diode becomes reverse-biased in its Zener mode and holds the output voltage to a level equal to the Zener voltage. In this case we are using a 5.6 volt Zener. This practice of limiting, or clamping, the output voltage to a specific level is often used to make the output voltage of the comparator compatible with other electronic circuits that it will drive.

COMPARATOR APPLICATIONS

The op amp comparator circuits that we have discussed here have a wide variety of applications. They are useful in any situation where you want to compare the voltage level of two inputs and produce an output that detects this comparison. There are many electronic control applications where it is necessary to detect when two voltage levels are equal.

The comparator may be used to monitor the output voltage of a power supply driving a critical instrument. The reference voltage is set equal to the desired power supply output voltage while the output voltage is applied to the comparator input. As long as the output voltage is above the reference voltage, the output of the comparator is considered to be *off*. However, if the power supply voltage should drop, the comparator would detect this change in voltage and cause its output to switch to an *on* condition. The output could then be used to turn on a light that would signal a low voltage condition that may be undesirable in the system. A comparator used in this way is called a threshold detector. Comparators find a wide variety of uses in control applications such as this.

The output voltage of a comparator can also be used to make a decision regarding the *off* or *on* of a specific logic circuit, based upon the relative amplitudes of certain input and reference voltages. For example, if the input voltage is above the reference voltage, a particular electronic circuit may be permitted to operate. However, if the input voltage drops below the reference voltage, the output of the comparator either turns off the electronic circuit to prevent its operation or enables another electronic circuit to perform an alternate function.

A very common application for the voltage comparator is waveform generation. For example, the comparator can be used to change a sine wave signal into a square wave. Fig. 43 shows a sine wave signal applied to the input of a comparator. The sine wave varies symmetrically above and below zero. The dashed line in the figure represents an applied dc reference voltage. Whenever the sine wave and reference voltages are equal, the output of the comparator will switch and produce a square wave output signal. Also notice in Fig. 43 that each time the sine wave and reference voltages are equal, the output of the comparator switches to produce a clean square wave. The frequency of the square wave is the same as that of the input sine wave. However, the duty cycle of the output waveform depends upon the setting of the reference voltage. A 50% duty cycle square wave can be generated from the sine wave by making the reference input equal to zero. Whenever the sine wave voltage crosses the zero axis, the comparator will switch and produce a 50% duty cycle square wave. Remember that output clamping with a Zener diode can be used to produce desired, limited voltage levels for driving other circuitry.

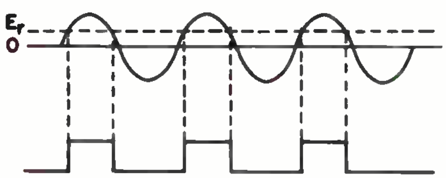


Fig. 43. Input-output waveforms of a comparator used for wave-shaping.

SELF-TEST QUESTIONS

(aa) A 3-input summer amplifier has input resistors $R_1 = 10K$, $R_2 = 5K$, and $R_3 = 2K$ to which are applied voltages of $E_1 = -10$, $E_2 = +6$ and $E_3 = +8$. The feedback resistor R_f is 20K. What are the amplitude and polarity of the output voltage?

(ab) True or false: An op amp summer can perform any algebraic addition problem.

(ac) True or false: A summer can have any number of inputs.

(ad) True or false: The output voltage rate of change in an integrator can be varied by changing the size of the feedback capacitor.

(ae) True or false: An op amp integrator does not invert.

(af) An integrator has a feedback capacitor of 1 mfd and an input resistor of 100K. What is the gain?

(ag) If the integrator in (af) is allowed to integrate a $-.3$ volt input for 7 seconds, what will the output voltage be?

(ah) True or false: In a summer type comparator the input and reference voltages must be of opposite polarity.

3-22-76

Tests and Measurements

The ultimate test of an op amp is its performance in the particular application for which it was selected. However, there are individual tests that can be made to test the individual specifications to determine its suitability for a given job. In this section we are going to discuss measurements of op amp characteristics. Adequate prediction of performance requires measurement of many of the parameters that will be discussed in this section.

MEASURING OPEN-LOOP GAIN

Open-loop gain of op amps can be a difficult parameter to measure; gain is usually very large and oscillation problems are often encountered. As you know, the dc gain of op amps is fairly large. At low frequencies it starts to decrease, or roll off, at a fixed rate. Therefore, we must measure open-loop gain at a very low frequency or at dc. Measurement at dc poses difficult problems because the offset voltage also gets amplified. If the offset voltage is too large, it could cause the output voltage to swing to either its positive or negative saturation level. Therefore, it is better to measure the open-loop gain using ac at some low frequency.

A common circuit used for the measurement of open-loop voltage gain is shown in Fig. 44. The input source is a sine wave at 10 Hz that is applied to the non-inverting input through a resistive divider formed by R_1 and R_2. This resistive divider is designed to reduce and accurately control the input voltage to the circuit. The op amp input voltage is adjusted to some low value such as 100 microvolts.

At first glance it may seem as if the amplifier is not in the open-loop mode because there is a 1 megohm resistor connected from the inverting input to the output and a capacitor of 1,000 microfarads connected from the input to ground. Notice that this test circuit has the same configuration as a non-inverting amplifier circuit: the input signal is applied to the non-inverting input while a feedback network is connected between the output and the inverting input. The only exception here is the use of a capacitor between the input and ground rather than a resistor. This capacitor is an effective impedance with ac signals so it works just as well as a resistor. The value of its reactance at 10 Hz is used with the 1 megohm feedback value to determine the gain. The gain called for by this feedback configuration at 10 Hz is equal to

$$\frac{E_o}{E_i} = \left(1 + \frac{R_f}{X_c}\right)$$

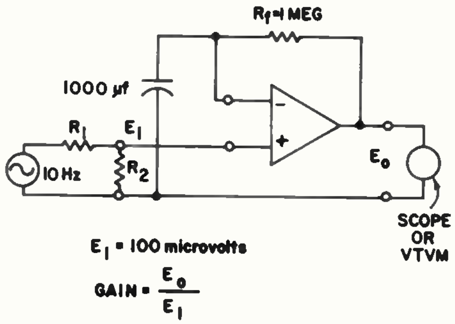


Fig. 44. Open-loop gain measurement.

The capacitive reactance of a 1,000 mfd capacitor at 10 Hz is

$$X_c = \frac{.159}{fC} = \frac{.159}{10 \times 1000 \times 10^{-6}} = \frac{.159}{10^{-2}} =$$

$$.159 \times 10^2 = 15.9 \text{ ohms}$$

Rounding off this value to 16 and using the other values given, we find that this feedback yields a gain of

$$\left(1 + \frac{R_f}{X_c}\right) = \left(1 + \frac{1{,}000{,}000}{16}\right) = 62{,}500$$

or approx. 96 db

In all probability, however, this gain is beyond the capability of the circuit. In other words, even though we have what appears to be closed-loop gain, it is high enough so that the limiting factor will be the open-loop op amp gain. If the op amp has a very large open-loop gain, the values of R_f in Fig. 44 should be further increased. Op amp, open-loop, frequency response is usually flat up to about 100 Hz. Therefore the low-frequency measurement is valid and convenient for measurements with standard instruments.

To measure the gain, we first connect a signal generator, whose output is 10 Hz, to the input and adjust it until the input voltage to the amplifier is 100 microvolts. This can only be done with knowledge of the value of the voltage divider formed by resistors R_1 and R_2. Since 100 microvolts is usually too small to observe directly, you can set the generator output to some more easily measured value and then rely upon the voltage divider ratio to produce an input voltage of 100 microvolts. Next, a high impedance voltmeter or an oscilloscope is connected to the output and the output voltage is measured. The gain is the ratio of the output to input voltage.

If an oscilloscope is used for output measurement, the peak-to-peak value is normally recorded. This must be converted to rms, or the same input voltage units, to obtain the correct value of gain. It is important that the power supply be properly decoupled, or filtered, with large capacitors to ground as close to the amplifier as possible. Should it be required, it is also important to compensate for amplifier oscillation. These precautions should be taken in almost all the measurement circuits we describe.

Gain Stability vs. Temperature. The temperature coefficient, or variation of the gain with temperature, can be measured with the same circuit as Fig. 44. All you have to do is vary the temperature of the amplifier and measure the gain at the various temperatures. A gain vs. temperature curve can then be plotted. The stability factor is a number obtained by dividing the change in gain by the change in temperature. The smaller this number, the better the stability.

Gain Stability vs. Power Supply Changes. The same procedure can be used to measure the variation of gain with changes in power supply voltage. This is often very important because the amplifier specification sheet may call for a gain with a given power supply setting different from the setting being used. To do this measuring, we must know the typical and worst case values for the various power supplies. In general, the gain of the op amp increases as the power supply voltage increases. However, we must be careful to see that the maximum rating of the op amp is not exceeded. The stability factor is a number obtained by dividing the change in gain by the change in power supply voltage setting. The lower this number is, the better the stability.

IMPEDANCE MEASUREMENTS

Open-loop Input Impedance. The open-loop input impedance is measured in a circuit like the one used for the open-loop gain test. The only exception is that a variable resistor (R_3) is placed in series with the input. Refer to Fig. 45. The input voltage is again adjusted to be 100 microvolts at about 10 Hz. With the value of resistor R_3 reduced to zero, the output voltage is measured. The second step is to increase the variable of R_3 until the output drops by 10%. In other words, if the output swings 10 volts peak-to-peak, we increase the value of R_3 until the output drops to 9 volts peak-to-peak. We are decreasing the output by decreasing the value of the input with a voltage divider, formed by the series R_3 and the input impedance to the amplifier. Since the output of the amplifier has dropped 10%, we conclude that the resistance into the input terminal is equal to 9 times the value of R_3 at that setting. All you have to do now is measure the value of R_3 and multiply it by 9. This will give you the input impedance of the amplifier.

The key to understanding this method of measuring the input impedance is to realize that the output voltage of the amplifier drops because of the introduction of variable resistor R_3. When it drops 10%, the relationship between R_3 and the input impedance of the amplifier is 9 to 1.

In order to use this method, we must have an initial idea of what the input impedance is going to be so we can choose the right magnitude for R_3. For example, this input impedance can be as large as 1,000 megohms in FET input amplifiers. This would mean that the value of the resistor must also be extremely large. Therefore, this method may be impractical for some high input impedance op amps.

Output Impedance. The output impedance measurement is similar to the measurement of the input impedance, but simpler. We'll use the circuit shown in Fig. 46 to illustrate.

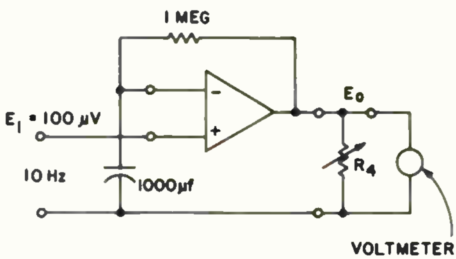


Fig. 46. Measuring output impedance.

We again adjust the input voltage to 100 microvolts but, this time, connect a variable load resistance (R_4) to the output. With the load resistance disconnected, we measure the output voltage. We then connect R_4 to the output and adjust its value until the output voltage drops 10%. For this procedure we must be sure that the output signal is not clipped or otherwise distorted. We are

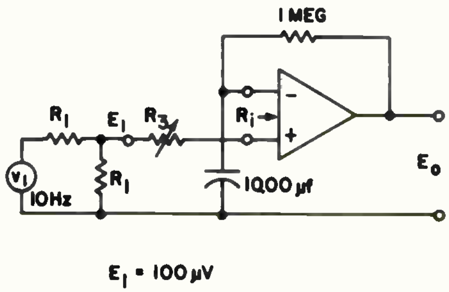


Fig. 45. Measuring input impedance.

using the output impedance of the amplifier and R_4 as a voltage divider whose output voltage drops as we decrease R_4. When the output voltage drops 10%, the output impedance of the amplifier is equal to the value of the load resistance divided by 9.

INPUT VOLTAGE MEASUREMENTS

Input Offset Voltage. The input offset voltage is very conveniently measured at dc using the circuit shown in Fig. 47. In this circuit we have an amplifier connected in the inverting configuration with an input resistance equal to 100 ohms and a feedback resistance equal to 10K. Therefore, the gain of the circuit is 10K/100 = 100.

With the input shorted to ground, the input to the circuit is equal to the input offset voltage of the amplifier. Since this input offset voltage is usually in the millivolt region, it is difficult to measure; it is best to amplify it and measure the enlarged version at the output. This is what this circuit does. The output voltage, if measured with a suitable voltmeter, will be equal to the input offset voltage multiplied by the gain of the amplifier. What we have done is to use the amplifier to increase its own input voltage by shorting the inputs. Now, if the input voltage is as small as 1 millivolt, it will appear in the output as a more conveniently measured voltage of 100 millivolts.

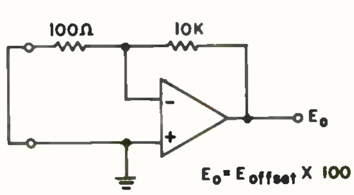


Fig. 47. Input offset voltage measurement.

Input Offset Voltage Drift vs. Temperature. The input voltage drift versus temperature can also be measured in the circuit of Fig. 47. All you have to do is vary the temperature of the amplifier and note the offset voltage change. Then compute the change in offset voltage divided by the change in temperature. The smaller this ratio, the better the matching of V_{be} of the input transistors with temperature.

Input Voltage Drift vs. Supply. The variation of input offset voltage with power supply voltage changes is commonly called the power supply rejection ratio. We measure the input offset voltage with different values of power supply voltages. Then we divide the change in offset voltage by the change in power supply voltage producing that change. The resulting ratio tells us the stability. A low ratio value is desirable. In some cases it is best to vary only one of the two supply voltages at a time.

Input Voltage Drift vs. Time. To measure the input voltage drift versus time, attach a strip chart recorder to the output and let the amplifier run. Then, checking the strip recorder output, we find the maximum output change for the prescribed time.

A strip chart recorder is a special type of voltmeter that records the output voltage in ink on a strip of graph paper. The paper is moved very slowly past the pen that marks a line or curve on it. The position of the pen is determined by the input voltage; the position of the mark on the paper shows the voltage variation with time.

Input Bias Current. The measurement of the input bias current is done best as in Fig. 48A. Rather than measuring the

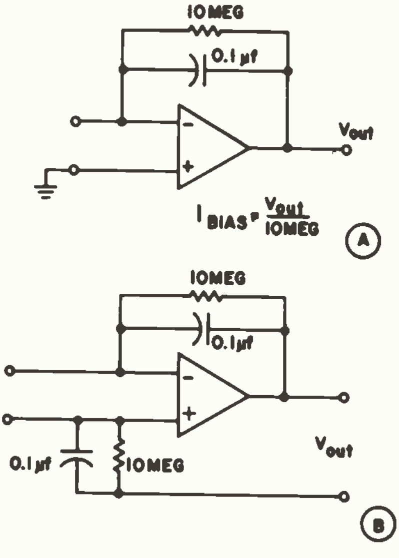


Fig. 48. Input bias current measurement (A) and input offset current measurement (B).

input bias current directly at the input with some form of current meter, we insert a 10 megohm resistor from the input to the output. The other input of the amplifier under test is returned to ground. This circuit takes advantage of the fact that the input bias current, the current that flows into or out of the input terminal, must flow through the 10 megohm resistor and appear as an output voltage. So the approach here is very similar to the one used for the input offset voltage. We convert the current into a fairly large voltage rather than a current of a few nanoamperes, which is very difficult to measure.

To use the circuit shown in Fig. 48A, measure the output voltage (E_0) with a suitable oscilloscope or voltmeter. The input-output current will be equal to the magnitude of the output voltage divided by 10 megohms. The reason for this is that the input bias current is flowing through the 10 megohm resistor. We can use Ohm's Law to determine this current. Let us assume that the output voltage of the amplifier (which we are testing under the conditions of Fig. 48A) equals 1 volt. When we divide this voltage by 1 megohm, we get .1 microampere, or 100 nanoamperes. This is the value of the input bias current flowing into the inverting input.

Using this method we can obtain the current flowing into only one of the input terminals. The current flowing into the other terminal could have a different value. Therefore in order to find the non-inverting input bias current, we must do a similar measurement. To do this we must ground the inverting input previously used and connect the 10 megohm resistor between the other input and ground. Now we obtain another value for input bias, different from the first one. The .1 mfd capacitor across the 10 megohm resistor is used to eliminate any noise or ac signal that could occur.

Input Offset Current. As you know, the average of the two input bias currents is called the input offset current. One way we could measure the input offset current is to individually measure the two input bias currents, add them and divide by two. However, the circuit shown in Fig. 48B shows a more direct way of obtaining the input offset current in one measurement.

We have a balanced input with a 10 megohm resistor and a .1 mfd capacitor connected to both inputs. This way we obtain an output voltage that is proportional to the difference between the two bias currents. To obtain the true input offset current, the value of this output current must be divided by two. In this circuit, the input offset current equals the output voltage divided by 10 megohms, as before.

In order to measure the input bias current drift and input offset current drift with temperature, we place the amplifier in a temperature-controlled environment and note the change in input offset current divided by the change in temperature. That way we get what we call the "temperature coefficient" of the input bias current and of the input offset current.

Often it is important to find the input current drift with power supply voltage changes. Depending upon the amplifier requirements, we vary the power supply output ±10, or 20%, while we monitor the input offset current. We then divide the change in input offset current, caused by the power supply variation, by the change in power supply magnitude. The resulting ratio gives an indication of the stability. A low ratio is most desirable.

FREQUENCY RESPONSE

In order to measure the frequency response, we normally connect the amplifier in the desired configuration. Then we apply a sine wave of varying frequencies to the input while we monitor the output voltage and the gain. However, in order to find the maximum frequency response, it is better to connect the amplifier in the unity gain configuration shown in Fig. 49A.

As we explained before, if feedback is used in an amplifier, the magnitude of the gain decreases and the bandwidth increases. In the extreme case of unity gain, the amplifier frequency response curve extends all the way to the point where it meets the open-loop roll-off curve. This is shown in Fig. 49B. In this figure we show the frequency response for open-loop gain, a gain of 100, and the frequency response of the unity gain amplifier which intersects the axis at the maximum frequency at which the amplifier can be used (F_{max}).

To find the maximum frequency response, we use a signal generator with a variable frequency output. The generator should be set at a fairly low output, e.g., 30 millivolts. The output of the amplifier is monitored with an oscilloscope or voltmeter and the frequency is increased until the output voltage drops 3 db. At this frequency the response will equal the F_{max} of the amplifier. Of course, if some gain is required the procedure will be the

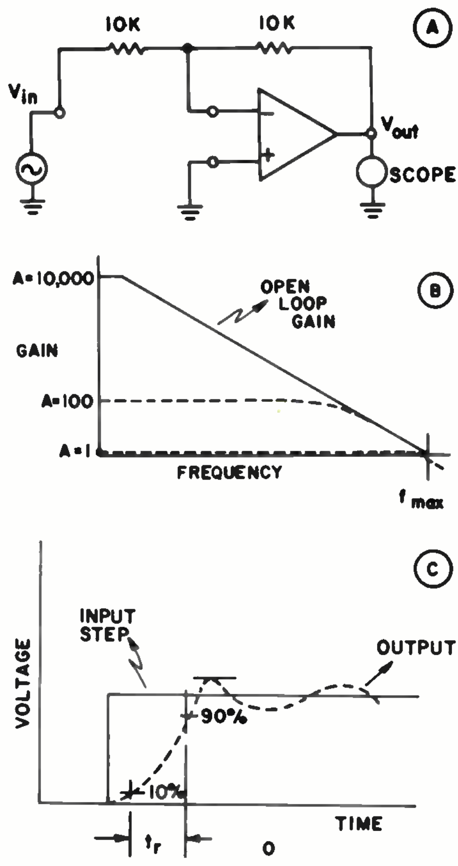


Fig. 49. Frequency response measurement.

same. The frequency response is always taken when the gain drops to 3 db below the low-frequency or dc value. As you can see, this frequency will decrease as the closed-loop gain of the amplifier increases.

Rise Time and Delay Time. The same circuit that was used for the frequency response measurement is also used to find the rise time of the amplifier. However, instead of connecting an input sine wave, we connect a square wave of about 100 millivolts peak-to-peak. We then monitor the output with an oscilloscope. The procedure is to observe the output on an expanded time scale and measure the output rise time from 10% to 90% of its rise. If the square wave input rise is fast compared to the rise time of the amplifier, the rise time of the square wave input can be neglected. For example, if we use a generator with a rise time of 10 nanoseconds and the rise time seen on the oscilloscope is 1 microsecond, the 10 nanoseconds can be neglected.

In Fig. 49C we show a picture of what can be seen on a dual channel oscilloscope. The first signal is the input step, or square wave, which we have shown as an ideal condition with zero rise time. The dashed line is what could be the amplifier output. As we mentioned before, the rise time is measured from the 10% point to the 90% point.

It is also possible that some degree of overshoot will be noticed rather than a smooth signal that simply follows the input. In many cases the overshoot limits must be specified by expressing the amount of overshoot as a percentage of the output voltage. Take the case where the input voltage is 100 millivolts and the output voltage is also 100 millivolts because the circuit is connected with the unity gain configuration. If the output in this situation overshoots to 110 millivolts and then comes back down, the amount of overshoot is 10% (10 millivolts/100 millivolts). In the case of many high-frequency amplifiers, the overshoot can go up to 40%.

The frequency response and rise time characteristics of an op amp are intimately related. In general, amplifiers with very high-frequency response have much faster rise times than amplifiers that have a low-frequency response. The rise time and frequency response are closely related; if one is known, the other can be calculated. The formula $F = .35/tr$ shows this relationship. F is the 3 db down frequency, or the amplifier bandwidth in MHz, where the t_r is the rise time expressed in microseconds. For example, if the rise time is .5 microsecond, the bandwidth, or 3 db cutoff, is $.35/.5 = .7$ MHz. By turning the formula around ($t_r = .35/f$), you can calculate the rise time for a given bandwidth.

SUMMARY

In this section we have attempted to describe the procedures for testing, measuring and evaluating some op amp parameters. In many cases such comprehensive tests are not necessary. The most important tests that guarantee op amp performance are:

1. Open-loop gain.
2. Input offset voltage and input offset voltage drift with temperature.
3. Input offset current.
4. Input impedance.

In applications where we are interested in specific characteristics, they must be measured. If we have an op amp that will change very much with time, the variations of gain, offset voltage and current with power supply voltage change become very critical. However, in most applications the parameters given are all that are needed to guarantee good op amp performance.

A lot of these test circuits can be combined with switching arrangements to form a piece of test equipment that can quickly perform all of the op amp tests given here. These automatic test sets for op amps are able to sequence from one test to the other by themselves. By using comparators they provide a light signal of whether all the tests have been passed; if any test was not passed, a red light indicates a no-go condition.

A photograph of one of these units is shown in Fig. 50. The square push buttons are used to select the desired test while the rotary switches select various input parameters. The meter is used for voltage and current measurements. This unit tests integrated circuit op amps that are plugged into the test socket below the push buttons. You are going to study integrated circuits in detail in a later lesson. You will see then how useful this device is in testing integrated circuit op amps.

Courtesy Philbrick/Nexus-Teledyne

Fig. 50. An automatic operational amplifier tester.

SELF-TEST QUESTIONS

(ai) Why is the open-loop gain measurement made at a very low frequency?

(aj) The output rise time of an op amp is 400 nanoseconds (.4 microseconds). What is the 3 db bandwidth?

Answers to Self-Test Questions

(a) True.
(b) False. The gain is the ratio of the feedback to the input resistor.
(c) True.
(d) True.
(e) False.
(f) Gain = R_f/R_1 = 84K/12K = 7.
(g) $E_o = -\text{gain} \times E_1 = -4(-3.5) = +14$ volts.
(h) High open-loop gain, high input impedance, low output impedance, wide frequency response, direct coupling, low drift, low power consumption, differential inputs.
(i) Gain = $1 + R_f/R_1 = 1 + 120K/30K = 1 + 4 = 5$.
(j) $E_o = R_f/R_1\ (E_2 - E_1) = 75K/15K\ [-23 - (-16)] = 5(-23 + 16) = 5\ (-7) = -35$ volts.
(k) False.
(l) False. The load is connected between the output and the inverting input.
(m) $k = E_o/E_1 = 9/27 = .333$.
(n) Gain = $1/k = 1/.004 = 250$.
(o) Unmatched emitter-base voltage drop in the differential input transistors is the basic cause of input offset voltage.
(p) Unmatched leakage and Beta in the input transistors cause input offset current.
(q) A high input impedance does not load the driving source.
(r) Rise time is the time it takes the amplifier output to rise from 10% to 90% of its final amplitude in response to a step input.
(s) The power supply voltages determine the output swing limit.
(t) Drift is caused by changes in component characteristics with temperature, time and power supply variations.
(u) No. This circuit can adjust the offset to produce a zero output for zero input, but drift will change the characteristics and make the output above or below zero.
(v) Six db per octave is equivalent to 20 db per decade. 100 Hz to 1 kHz is a decade change so the gain drops by 20 db or from 120 db to 100 db.
(w) Frequency compensation networks insure a 6 db per octave roll-off to prevent oscillation.
(x) High input impedance can be obtained by using FET input transistors, the Darlington connection or super-high Beta transistors at low current levels.
(y) A booster amplifier is used with an op amp to permit it to develop its normal output voltage across a low resistance load. It is a power amplifier.
(z) The internal self-bias, due to the depletion layer, causes crossover distortion.
(aa) $E_o = -\left[E_1(R_f/R_1) + E_2(R_f/R_2) + E_3(R_f/R_3)\right]$

$E_o = -\left[-10(20K/10K) + 6(20K/5K) + 8(20K/2K)\right]$

$E_o = -\left[-10(2) + 6(4) + 8(10)\right]$

$E_o = -(-20 + 24 + 80) = -84$.
(ab) True.
(ac) True.
(ad) True. Changing the feedback capacity varies the gain so the rate of output voltage change can be controlled.

(ae) False. The op amp integrator does invert.

(af) Gain = 1/R-C = 1/(100K X 1 mfd) equals
$1/(100{,}000 \times 1 \times 10^{-6}) =$
$1/(10^5 \times 10^{-6}) =$
$1/10^{-1} = 1/.1 = 10.$

(ag) $E_o = -1/R\text{-}C \cdot E_{in} \cdot t = -10 (-.3)(7) = 21$ volts.

(ah) True.

(ai) Since the gain begins to roll off at a very low frequency, to obtain a measurement of the maximum open-loop gain, the test should be performed at dc or some low frequency before this roll-off occurs.

(aj) Bandwidth $f = .35/t_r = .35/.4 =$ .875 MHz or 875 kHz.

Lesson Questions

Be sure to number your Answer Sheet K309.

Place your Student Number on every Answer Sheet.

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

1. What three transistor characteristics affect input offset voltage and current?

2. What two external conditions are most responsible for changes in op amp characteristics?

3. A two-input op amp summer has input resistors R_1 = 50K, R_2 = 40K to which are applied input voltages of E_1 = −30 volts and E_2 = +40 volts. The feedback resistor R_f is 200K. What are the output voltage amplitude and polarity?

4. Name two ways of getting high input impedance in an op amp.

5. An input voltage of −2.3 volts is applied to an op amp follower. What is the output voltage?

6. A triangular waveform is applied to a comparator whose reference input is grounded. The output signal will be:

   (a) A triangular wave
   (b) A sine wave
   (c) A square wave
   (d) A fixed dc voltage

7. The output saturation limit on an op amp integrator is ±50 volts. The integrator gain is 5, the input voltage is +2 and the circuit integrates for 4 seconds. Is the output in saturation?

8. The 10% to 90% output rise time of an op amp is .08 microseconds. What is the upper 3 db down bandwidth?

9. As the closed-loop gain of an op amp circuit increases, the bandwidth

   (a) increases
   (b) decreases

10. Name five mathematical operations that an op amp can perform.

3-25-76

$E_0 \quad R_c\,(K_1)(T)$

$E_0 \quad 5(+2)(4)$

$E_0 = 40v$

$F = \frac{1.3}{t_n}$

$1 \quad \frac{.22}{.08} = 2.375\ \text{kHz}$

COMPETITION

When a competitor opens a shop in your neighborhood, your first reactions are probably the same as those of most people – you feel that he is "cutting in" on your trade and that, by fair or foul means, he may run you out of business. However, there is another view to take of this problem.

First, forget your fears! A mind frozen by mistrust and hate is incapable of reasoning; it will lead you to the very downfall you fear. Face the facts: someone else is in the same business, so you must make your services so much better than his that you get your share of the work.

Welcome the competition as a spur – something to force you to your best efforts – something to make you become more careful, more efficient, more alert. You will find that honest competition adds enjoyment to your work.

And, another thing, force your competitor to rise to your level to survive – don't stoop to his. Do your best work and you'll find that your fears were not justified – there is plenty of business for the man who can deliver the goods!

Innovation in learning
by a McGraw-Hill
Continuing Education School

139-LT-710(506)