6

AMPLIFIERS

This chapter covers electronic amplifiers. An electronic amplifier is a circuit using active devices such as transistors and ICs that allows an input to control a power source to produce some useful output. Industrial systems may also use hydraulic, pneumatic, or magnetic amplifiers.

6-1 SIMPLE LINEAR TYPES

A *linear amplifier* is one whose output signal is a replica of the input signal. For example, if the input signal is a sinusoid, then the output signal from the amplifier will also be a sinusoid. The purpose of a linear amplifier is to increase the level of the signal. In an industrial control system, the output from a temperature sensor may vary only several millivolts over its entire operating range. Small signals such as this must be amplified to be useful. Amplifiers are often identified according to the power level that they produce. Amplifiers that produce signals at significant levels of voltage or current are usually called *power amplifiers*. Amplifiers that work at small voltage and current levels are usually called *voltage amplifiers* or *small-signal amplifiers*.

The amount of gain that an amplifier produces is often measured in decibels (dB). The dB power gain of an amplifier is evaluated by

$$dB = 10 \times \log \frac{P_{out}}{P_{in}}$$

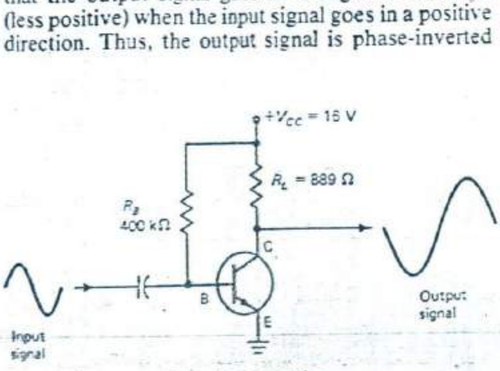
For example, if an amplifier develops a 100-W output signal when driven by a 1-W input signal, then the power ratio is 100. The common logarithm of 100 is 2; therefore the power gain of the amplifier can be stated as 20 dB. Small-signal amplifiers are usually evaluated in terms of their voltage gain. Since power varies as the square of the voltage, the dB gain is evaluated as follows:

$$dB = 20 \times \log \frac{V_{out}}{V_{in}}$$

Suppose an amplifier develops a 5-V output signal when driven by a 100-mV input signal. The voltage ratio is 50, the common logarithm of 50 is 1.7, and the gain of the amplifier is 34 dB. Since the dB system is based on power ratios, it must be adapted to work with voltage ratios by doubling the logarithm, which is equivalent to squaring the voltage ratio. However, there is an implicit assumption: that the input impedance of the amplifier is equal to the output impedance of the amplifier. This rule is commonly violated, and the dB voltage gain of a circuit is often evaluated with the preceding equation even though the impedances are not equal.

Figure 6-1 shows a simple common emitter amplifier circuit. The name *common emitter* is used since the input signal is supplied to the base circuit, the output signal is taken from the collector circuit, and the emitter terminal is grounded and is therefore common to both the input and output circuits. This circuit is useful for amplifying small signals since it can be expected to show a voltage gain of about 100 times (40 dB, ignoring the impedances). It can be viewed as a voltage amplifier even though it is based on a bipolar transistor, which is a current amplifier. This idea was discussed in Chapter 2.

The input signal is capacitively coupled in Fig. 6-1 to the base of the transistor. As the input signal goes in a positive direction it will aid the supply voltage (V_{cc}) and increase the base current. Since the transistor is controlled by base current, the collector current will also increase but many times more since there is quite a bit of current gain from the base to the collector. The increase in collector current will cause a greater drop across the collector load resistor (R_L), and therefore less voltage will drop across the transistor from collector to ground. It can be seen that the output signal goes in a negative direction (less positive) when the input signal goes in a positive direction. Thus, the output signal is phase-inverted



6-1 Common emitter voltage amplifier.

180° when compared to the input signal. As the input signal goes in a negative direction, the base current decreases, the collector current decreases, and the drop across the transistor increases, making the collector terminal go in a positive direction.

Figure 6-2 shows a graphic presentation of the amplifier performance. It consists of a collector family of characteristic curves upon which a load line has been added to show how the collector load and supply voltage will interact with the transistor to develop an output signal. Notice that one end of the load line terminates on the horizontal axis at a value of 16 V. This is the cutoff voltage and is equal to the supply voltage in Fig. 6-1. Also notice that the other end of the load line terminates on the vertical axis at a value of 18 mA. This is called the *saturation current* and represents the maximum flow with a 16-V supply and a load resistor of 889 Ω. Use Ohm's law to verify this current. An amplifier can operate anywhere along its load line between saturation and cutoff. Small-signal amplifiers usually operate near the center for best linearity and maximum signal output swing. The amplifier of Fig. 6-1 will operate where the 40-μA base curve intersects the load line due to the 16-V supply and the 400.000-Ω base-limiting resistor R_B. Use Ohm's law to verify this current.

The operating point is also called the quiescent point. It is marked with a Q in Fig. 6-2. The *quiescent value* is the steady-state or resting condition of the amplifier. Quiescent values can be measured with no input signal applied. The quiescent collector-to-emitter voltage can be found by projecting down from

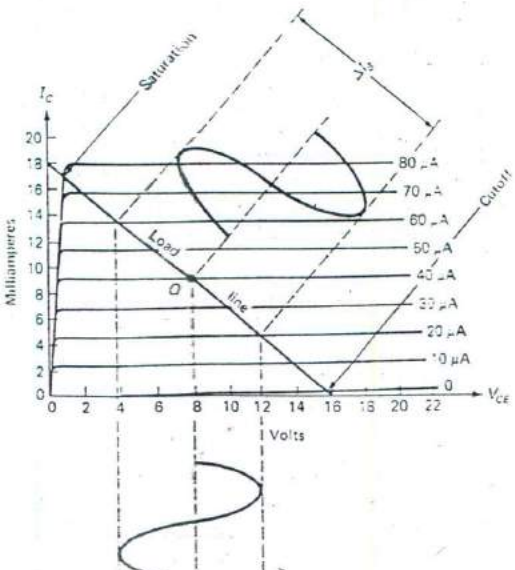


Fig. 6-2 Amplifier load line.

the Q point in Fig. 6-2 and is seen to be a little less than 8 V. The quiescent collector current is found by projecting to the left and is a little over 9 mA. What happens when an input signal is applied? An ac input signal will drive the amplifier above and below its Q point. Figure 6-2 shows the amplifier being driven with an input signal of 40 μA peak-to-peak (from 60 to 20 μA). By projecting this swing down, the output signal is shown to be about 8 V peak-to-peak. The voltage gain can be calculated if the input resistance of the base-emitter circuit is known. This resistance is approximately 2000 Ω in a circuit of this type. We can now use Ohm's law to calculate the input signal voltage required to develop a signal current of 40 μA peak-to-peak:

$$V = 40 \times 10^{-6} \times 2000 = 0.08 \text{ V}$$

The voltage gain of the amplifier is 8 divided by 0.08, or 100 times. It should now be clear how bipolar transistors can be viewed as voltage amplifiers even though they are inherently current-amplifying devices.

Figure 6-2 is also useful to explain clipping in a linear amplifier. If the input signal is too large, the output signal will be clipped. The limits at which the clipping will occur are saturation and cutoff. The amplifier will clip if the output tries to swing more than about 15 V peak-to-peak. The output signal will no longer be a good reproduction of the input. A severely clipped sine wave, for example, looks more like a square wave. This is not desirable in linear amplifiers. It is avoided by operating the amplifiers near the center of their load lines and by not allowing the input signals to become too large.

Figure 6-1 can also be evaluated with a few simple calculations. We know how to use Ohm's law to calculate the base current at 40 μA. If h_{FE} is known, the collector current is found by:

$$I_C = h_{FE} \times I_B = 230 \times 40 \times 10^{-6} = 9.2 \text{ mA}$$

You can verify h_{FE} and the quiescent collector current from Fig. 6-2. Chapter 2 covers how to calculate h_{FE} from the curves if you have forgotten. The h_{FE} varies quite a bit from transistor to transistor. Suppose the circuit is constructed with another transistor which has an h_{FE} of only 50. In this case, the collector current will be only 2 mA. The drop across the load resistor will be 2 mA × 889 Ω, or 1.78 V. This means that the drop across the transistor will be 16 − 1.78, or 14.22 V. The transistor will be operating near cutoff, which is not a good arrangement. The output signal can now be no larger than 3.56 V peak-to-peak before clipping starts to appear on the positive peaks. It should be clear that the circuit of Fig. 6-1 is not practical because it is too sensitive to h_{FE}. It is also temperature-sensitive, and its Q point will move toward saturation as it gets warmer.

Figure 6-3(a) shows an improved common emitter amplifier. This circuit is not nearly so sensitive to h_{FE} and temperature. It is improved by the addition of two resistors: one in the base circuit and one in the emitter circuit. Resistors R_{B1} and R_{B2} form a voltage divider to set the base voltage at some fraction of the supply voltage. The divider current is much greater than the base current (typically 20 times as much) and therefore behaves essentially as an unloaded divider to provide a stable base voltage. Placing different transistors into the circuit will not change the base voltage or current appreciably. Resistor R_E is added in the emitter circuit to provide negative feedback. If a transistor tends to conduct more current, the drop across the emitter resistor will increase, thus subtracting from the forward bias for the base-emitter junction. This condition tends to decrease the device current. The negative feedback stabilizes the operating point. The negative feedback increases the input impedance of the amplifier by an amount equal to h_{FE} times R_E. The negative feedback also affects the signal performance of the amplifier. The voltage gain is much lower and is approximately equal to R_L divided by R_E. However, the gain can be restored for ac signals by eliminating the ac negative feedback. This is done by adding an emitter bypass capacitor (shown as a phantom component in Fig. 6-3[*a*]). The capacitor is selected to have a reactance of about one-tenth the value of R_E at the lowest frequency of operation.

Figure 6-3(*b*) shows the PNP version of the improved amplifier circuit. It has the same ac characteristics as the NPN version but uses a negative collector supply. Compare the two versions and note that the dc currents are reversed.

The circuits discussed thus far have shown capacitors for coupling signals and for eliminating ac negative feedback. These techniques are not useful in dc amplifiers. Direct coupling must be used in dc amplifiers. One very popular direct-coupled arrangement is the Darlington amplifier shown in Fig. 6-4(*a*). The emitter of Q_1 is directly connected to the base of Q_2. This circuit provides a very high gain from the input to the output. The input impedance is also increased by this arrangement. If each transistor has an h_{FE} of 200, the overall current gain will be approximately equal to h_{FE1} times h_{FE2}, or 40,000. In practice it is difficult to achieve this much gain, however. The first transistor must be operated at a very low level to avoid driving the second transistor into saturation. Transistor gain tends to drop off at low current levels. In spite of this, the circuit still provides considerable current and voltage gain. When two transistors are connected in the Darlington arrangement, they essentially behave as a single transistor with super gain. For this reason, they are avail-

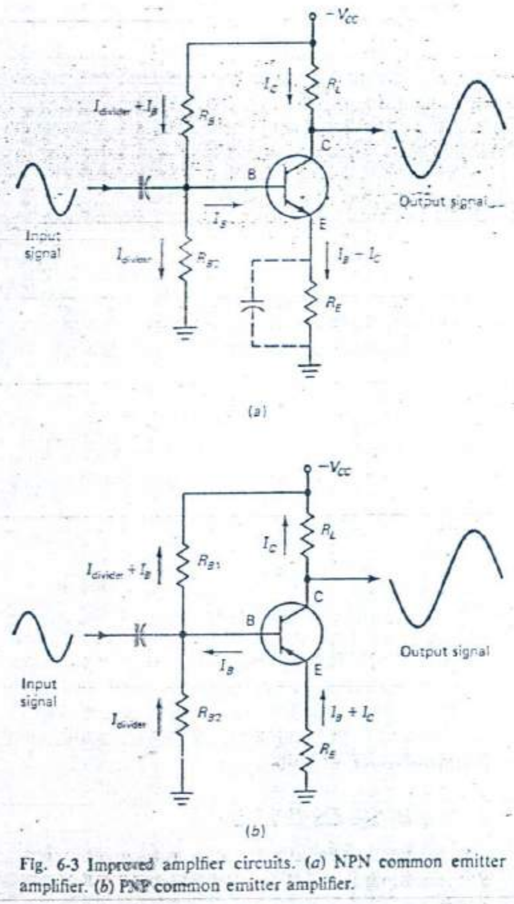


Fig. 6-3 Improved amplifier circuits. (*a*) NPN common emitter amplifier. (*b*) PNP common emitter amplifier.

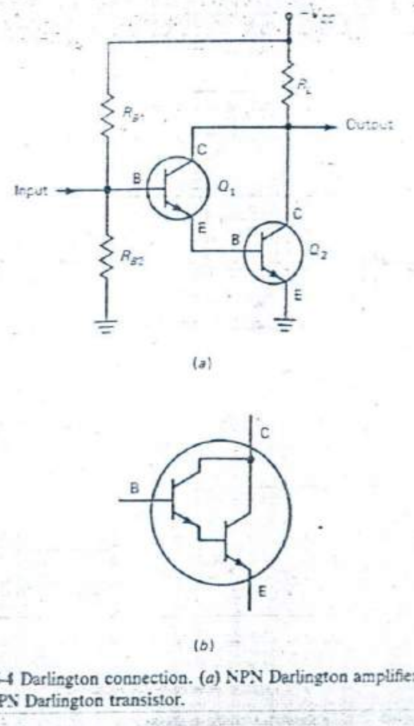


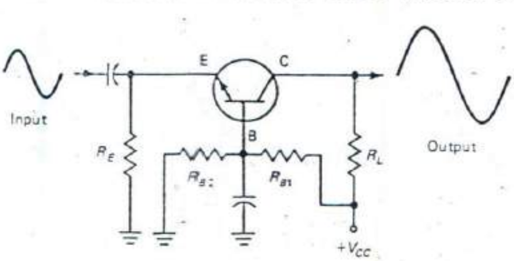
Fig. 6-4 Darlington connection. (*a*) NPN Darlington amplifier. (*b*) NPN Darlington transistor.

TABLE 6-1 AMPLIFIER CONFIGURATION SUMMARY

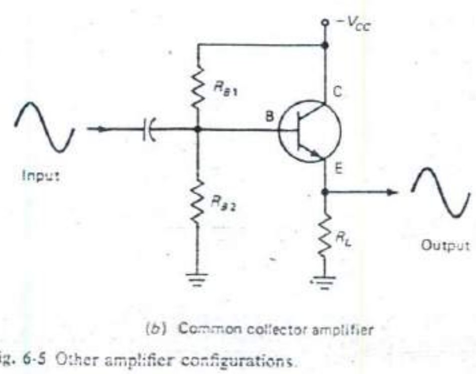
	Common Emitter	Common Base	Common Collector (Emitter Follower)
Input signal applied to	Base	Emitter	Base
Output taken from	Collector	Collector	Emitter
Current gain	High 50 to 100	Low ≈1	High 50 to 100
Voltage gain	High ≈100	High ≈100	Low ≈1
Input impedance	Medium ≈1200 Ω	Low ≈50 Ω	High ≈100 kΩ
Output impedance	Medium ≈50 kΩ	High ≈1 MΩ	Low ≈50 Ω
Application	General purpose	Low to high impedance transformer	Buffer amplifier

able packaged as a single device with one base, one emitter, and one collector lead (Fig. 6-4[*b*]). In addition, PNP Darlingtons are available.

The common emitter configuration is the most popular way to arrange an amplifier circuit. However, two other configurations exist, as shown in Fig. 6-5. The common base amplifier in Fig. 6-5(*a*) is so named because the base is common to both the input and output circuits. This configuration has a low input impedance since the input signal is applied to the emitter terminal, which is a high-current point. It is therefore limited to amplifying signal sources that have a low characteristic impedance. The output impedance is high. The output signal is not phase-inverted as it is in the common emitter amplifier. It provides no current gain but does provide voltage and power gain. The common collector amplifier configuration is shown in Fig. 6-5(*b*). It uses the emitter terminal as the output and shows no voltage gain; there is actually a slight loss in signal voltage. The output is in phase with the input, and the circuit is usually called an *emitter follower*. Even though it has no voltage gain it is still very useful, especially for eliminating loading effects on a signal source. It has a reasonably high input impedance and a low output impedance and is often used as a buffer amplifier. It does provide current gain and power gain. The circuit of Fig. 6-4(*a*) can be converted to a Darlington emitter follower by moving the load resistor to the emitter circuit of Q_2 to provide an even higher input impedance. The common emitter amplifier is the only configuration that provides both voltage and current gain. It also provides the highest power gain and is best suited to most applications. Table 6-1 provides a summary of the amplifier configurations.



(*a*) Common base amplifier



(*b*) Common collector amplifier

Fig. 6-5 Other amplifier configurations.

Other transistor types can be used to build amplifiers. For example, PNP devices can be substituted in the circuits shown by using a negative collector supply. Metallic oxide semiconductor field effect transistors (MOSFETs) can be used, although the gate-biasing techniques may be different. For those applications where a signal source has a high impedance MOSFETs are attractive. Since they are voltage-controlled transistors, they do not load the signal source as much as biopolar transistors do. They are often used in the first stage of a multistage amplifier. Bipolars will be used in the subsequent stages since they are less expensive and provide better gain. Power MOSFETs are attractive in some applications and are replacing power bipolars. Some of the industrial applications are in dc motor control, ac motor control, power supplies, induction heating, and high-frequency welding.

REVIEW QUESTIONS

1. Calculate the dB power gain for an amplifier with an input signal of 1 mW and an output signal of 5W.

2. Calculate the dB gain of an amplifier with an input signal of 0.5 V peak-to-peak and an output signal of 20 V peak-to-peak (assume the input impedance is equal to the output impedance).

3. A transistor amplifier is driven at its base terminal, and the output signal is taken from its collector terminal. What is the configuration of this amplifier?

4. What is the phase relationship between the input and the output voltages for question 3?

5. Refer to Fig. 6-1. If the drop across the collector load resistor is 12 V, what is the voltage from the collector terminal to ground?

6. Suppose the base bias resistor in Fig. 6-1 opens (infinite resistance). At which end of the load line will the amplifier operate?

7. What will be the collector voltage for question 6?

6-2 DIFFERENTIAL AMPLIFIERS

A *differential amplifier* (diff amp) responds to the difference between two input signals. Figure 6-6 shows a simple example. The circuit has two inputs and two outputs. It is possible to drive such an amplifier at one, or at both, of its inputs. It is also possible to take the output signal from one, or both, of its outputs. In a *single-ended output* the signal is being taken from one of the outputs (referenced to ground). In a *differential output* the signal is being taken from both outputs (not referenced to ground). The amplifier is very flexible in this respect since it offers four combinations of input/output conditions. Differential amplifiers are important in their own right. They are very flexible and stable and capable of rejecting some types of noise. They are even more important as an integral part of operational amplifiers, which are covered in the next section of this chapter.

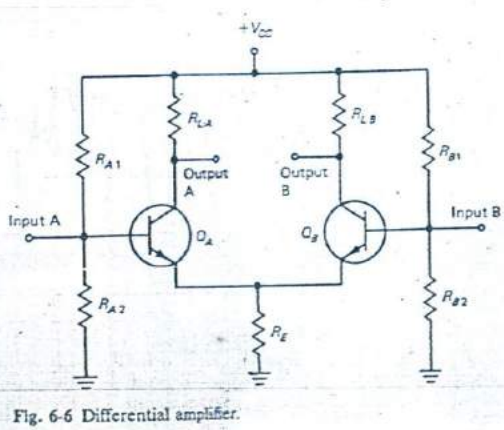


Fig. 6-6 Differential amplifier.

Direct coupling is the method used to provide dc gain. When quite a bit of gain is needed, several stages must be arranged, with the output of the first amplifier directly coupled to the input of the second amplifier and so on. When this multistage arrangement is used with the simple designs of the preceding section, drift becomes a problem. Transistors are temperature-sensitive. When many stages are direct-coupled, even a small temperature change in the first stage can drive the last stage into saturation or cutoff. Temperature compensation and negative feedback can be used to stabilize the operating point, but the circuit tends to become expensive and complicated. A differential amplifier is an attractive choice because its design inherently cancels drift due to temperature change (refer to Fig. 6-6). If the two transistors are closely matched and maintained at the same temperature, any temperature change will affect both outputs by the same amount. We will see that such changes are canceled in the differential output. Matching transistors is an expensive proposition in discrete circuitry, but fortunately it is a byproduct of monolithic circuit construction. Since all of the devices are formed at the same time, they have well-matched characteristics. They also track well in temperature since they exist together in a single monolithic structure. Thanks to integrated-circuit technology, differential and operational amplifiers have found wide application in industrial circuitry.

Suppose the amplifier of Fig. 6-6 is driven with a sinusoid at its B input only. As you would expect, the input signal will cause changes in the base current, which, in turn, will create a collector signal at output B. We can also expect that output B will be phase-inverted. Now look at Fig. 6-7. It shows that both outputs are active. Why is this so, since only input B is being driven? The reason is that both sides of the differential amplifier share emitter resistor R_E. When input B goes positive, the base, collector, and emitter currents of Q_B all increase. This causes an increase in the current through R_E which increases the voltage drop across it. This drop acts as a positive-going signal fed to the emitter of Q_A, which responds as a common base amplifier, and it produces a positive-going output at A (Fig. 6-7). Thus, both outputs are active even though only one input is driven. The same results can be obtained by driving only input A, except that output A will be out of phase and output B will be in phase with the input. Driving only one of the inputs creates a difference signal to which this type of amplifier responds.

Figure 6-7 also shows that the differential output $(A - B)$ has twice the peak-to-peak swing when compared to either single-ended output. For example, if output A has swung 2 V positive, then output B has swung 2 V negative, and the difference is $+2 - (-2) = +4$ V. What happens if both inputs are driven with the same signal? Of course, if they are driven the same, then the difference is 0 and there is no difference to amplify. Any signal applied to both inputs with exactly the same phase and voltage is

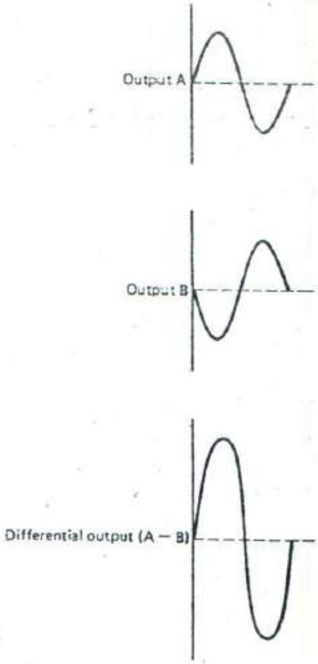


Fig. 6-7 Outputs with differential input.

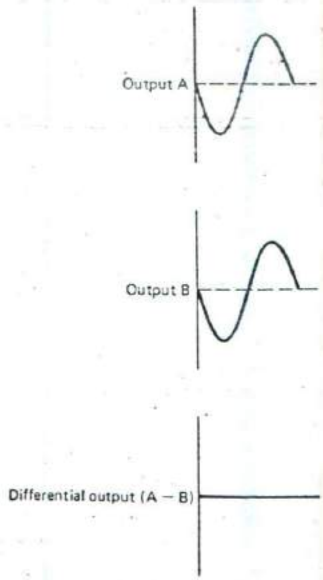


Fig. 6-8 Outputs with common mode input.

known as a *common mode signal*. Figure 6-8 shows what happens. When input A and input B are driven by a common mode signal, the outputs are in phase. If output A has swung 2 V negative, then output B has also swung 2 V negative, and the difference is $-2 - (-2) = 0$ V. So when the output is taken differentially, any common mode signal is canceled. It will be shown a little later that it is also possible to cancel common mode signals at the single-ended outputs by adding a current source to the emitter circuit of the diff amp.

The ability of differential amplifiers to cancel common mode signals is an important one. Many signals have *common mode noise*. A prime example is in the case of floating measurements, where the oscilloscope probe cannot be grounded because of the presence of ac voltages (reference to ground) at both points in a piece of equipment across which a waveform must be measured. Figure 6-9 shows how a differential amplifier can be used to make such a measurement. Notice that input A is a higher-frequency wave that is riding on a lower-frequency signal. The lower-frequency signal is undesired and is known as *noise*, or *hum*. Input B is about the same, but note that the phase of the higher-frequency signal is inverted. Now look at the output. The low-frequency noise, or hum, has been canceled, and the high-frequency signal has been amplified because it appeared as a differential signal. A differential amplifier can also eliminate high-frequency noise if it appears as a common mode signal.

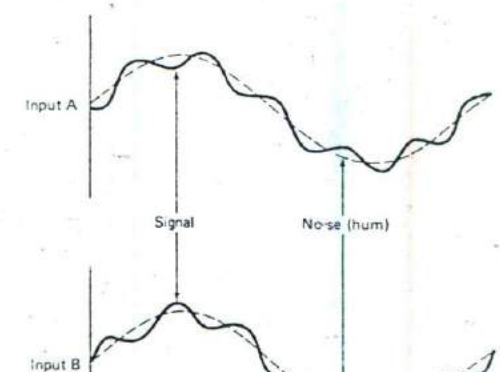


Fig. 6-9 Elimination of common mode noise.

A differential amplifier must be perfectly balanced if it is to cancel the common mode signal completely. Perfect balance can never be achieved. For this reason, a way of measuring performance is required.

The figure of merit is known as the *common mode rejection ratio* (CMRR). It is found by

$$\text{CMRR} = 20 \times \log \frac{\text{Gain (differential)}}{\text{Gain (common mode)}}$$

For example, suppose an amplifier shows a voltage gain of 300 for a differential signal and a voltage gain of 0.01 for a common mode signal. The gain ratio will be 30,000, the common logarithm is 4.5, and the CMRR is 90 dB. Note that the amplifier actually decreases the common mode signal to only 1 percent of its original value. This is the opposite of amplification and is called *attenuation*.

Figure 6-10 shows an important improvement to the differential amplifier. It is powered by a dual supply. Compare this with Fig. 6-6 and verify that the voltage divider bias has been eliminated. The negative emitter supply (V_{EE}) allows the bases of both transistors to be operated at dc ground potential. Base current is very small, and the drop across R_A and R_B is negligible. This is usually desired in a dc amplifier since the signal source is often referenced to ground. The original circuit (Fig. 6-6) is awkward to use since the inputs have a dc offset with respect to ground.

Figure 6-11 shows another improvement for the differential amplifier circuit. The emitter resistor has been replaced with a constant current source. This particular example uses a 5.7-V zener biased by the negative supply. The zener drop is applied across the base-emitter (B-E) junction of transistor Q_C and the 1000-Ω resistor. Subtracting 0.7 V for the B-E junction leaves 5 V across the resistor. Ohm's law sets the current through the resistor at 5 mA. Since the emitter and collector currents are almost equal, Q_C will conduct a total of 5 mA for both transistors in the diff amp. Assuming balance, each transistor will have 2.5 mA of emitter current. Now assume a common mode signal that is going positive. Ordinarily, both transistor currents would increase, but in this case they cannot. The constant current source supplies a total of 5 mA regardless of the change in collector-emitter resistance in the differential pair. The same thing is true if a common mode signal is negative-going and attempts to decrease the current in both transistors. The common mode signal will not appear at output A or output B. In other words, adding the constant current source provides good common mode rejection for the single-ended outputs. The previous circuits provided good common mode rejection only in the differential output. A differential input in Fig. 6-11 will unbalance the transistor currents. For example, a positive-going signal applied to one input might cause that transistor current to increase to 3 mA, in which case the other transistor current will decrease to 2 mA. The current changes will cause both differential and single-ended outputs to appear. The constant current source has a second advantage. It raises the input impedance of the differential amplifier to the megohm region. The input impedance is increased by an amount equal to h_{FE} times the impedance of the current source, which is characteristically very high.

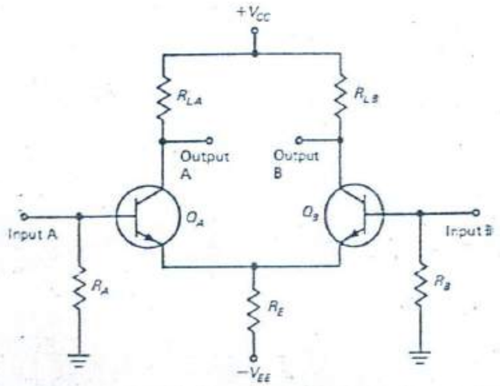


Fig. 6-10 Dual-supply differential amplifier.

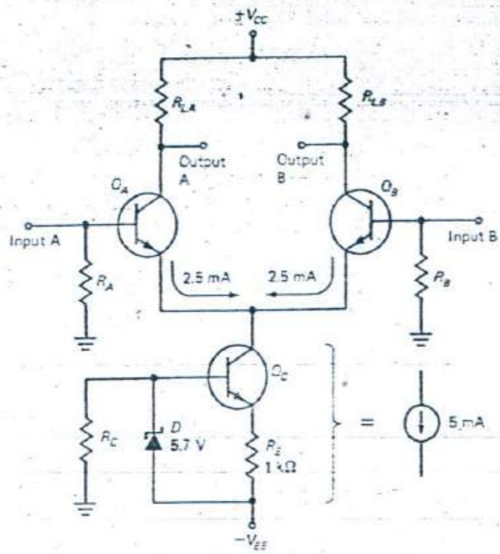


Fig. 6-11 Diff amp with constant current source.

REVIEW QUESTIONS

8. Refer to Fig. 6-6. Assume that only input A is driven with a signal. Will a signal appear at output B? If so, what is its phase relationship to the input signal?

9. The circuit of Fig. 6-6 is driven with a common mode signal. Will there by any single-ended output? Assuming balance, will there be any differential output?

10. Refer to Fig. 6-9. Suppose the low-frequency (noise) component of input signal B is out of phase

with the same component at input A. Would the amplifier be able to attenuate the noise?

11. What is the CMRR (in decibels) of an amplifier that provides a voltage gain of 100 for a differential signal and a voltage gain of 0.01 for a common mode signal?

6-3 OPERATIONAL AMPLIFIERS

The original operational amplifiers (often called op amps) were based on vacuum tubes and were used to do mathematical operations in analog computers. They were large, expensive, power-hungry, and subject to drift. The digital computer has replaced the analog computer, and solid-state devices have replaced vacuum tubes. Modern operational amplifiers, thanks to integrated circuit technology, are small, inexpensive, power-efficient, and much more stable. They have found a wide range of applications in industrial circuitry even though their first application, the analog computer, has vanished. They are direct-coupled high-gain amplifiers and are usually powered by a dual supply. Dual-supply operation allows them to conveniently amplify signals near ground potential and allows their output to swing above and below ground potential.

The major sections of a modern IC operational amplifier are shown in Fig. 6-12. There are two inputs to the first stage, which is a differential amplifier. The outputs of most op amps are single-ended. One of the differential inputs is in phase with the output and is called the *noninverting input*; it is marked plus (+). The other input is out of phase with the single-ended output, is called the *inverting input*, and is marked minus (−). An intermediate voltage amplifier follows the differential input amplifier to provide high gain. An output amplifier is the third major stage and provides a low output impedance so that the op amp can drive most loads. In addition to the terminals shown in Fig. 6-12, an op amp may have offset null terminals, frequency-compensation terminals, or gain-control terminals. The standard schematic symbol for an operational amplifier is a triangle with + and − inputs and a single output. The power supply and other connections may be omitted from some schematics for simplicity, as they are for many of the circuits shown in this book.

Figure 6-13 shows the equivalent circuit for a very popular op amp, the 741. Since it is a monolithic IC, studying the circuit in detail is not necessary. It is shown here to demonstrate how complex and expensive it would be in discrete form and to illustrate a few important concepts. Transistors Q_1 and Q_2 are the input devices and are part of a differential amplifier which acts as a voltage-to-current converter. Emitter resistors R_1 and R_2 are brought out to offset null terminals for the purpose of externally balancing the diff amp. The current signal from the differential input is sent to the second stage, consisting mainly of Q_{16} and Q_{17}, which acts as current-to-voltage converter. Capacitor C_1, a 30-pF frequency-compensation capacitor, is connected across the second stage to roll off (decrease) its frequency response 20 dB for every decade increase in frequency. A decade equals 10; thus a frequency change from 100 to 1000 Hz represents a decade increase. This gain roll-off is characteristic of most modern op amps with internal frequency compensation. It is used to ensure that the amplifier will remain stable with all feedback configurations. An unstable amplifier does not respond as planned and is useless. As we will see, op amps are almost always used with feedback. The last detail we will look at is the output circuit, which consists mainly of transistors Q_{14} and Q_{20}, which are a complementary pair. These two devices act as emitter followers to give the op amp a low-impedance output so that it can drive loads down to around 2000 Ω.

Figure 6-14 shows some popular op amp packages. The dual-in-line style is widely applied and can house one, two, or four amplifiers. The metal package is hermetically sealed and rated to operate over a slightly wider temperature range but is more expen-

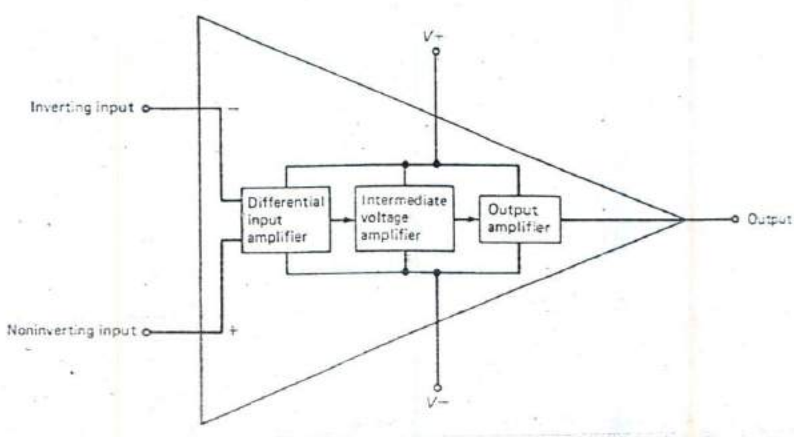


Fig. 6-12 Major sections of an operational amplifier.

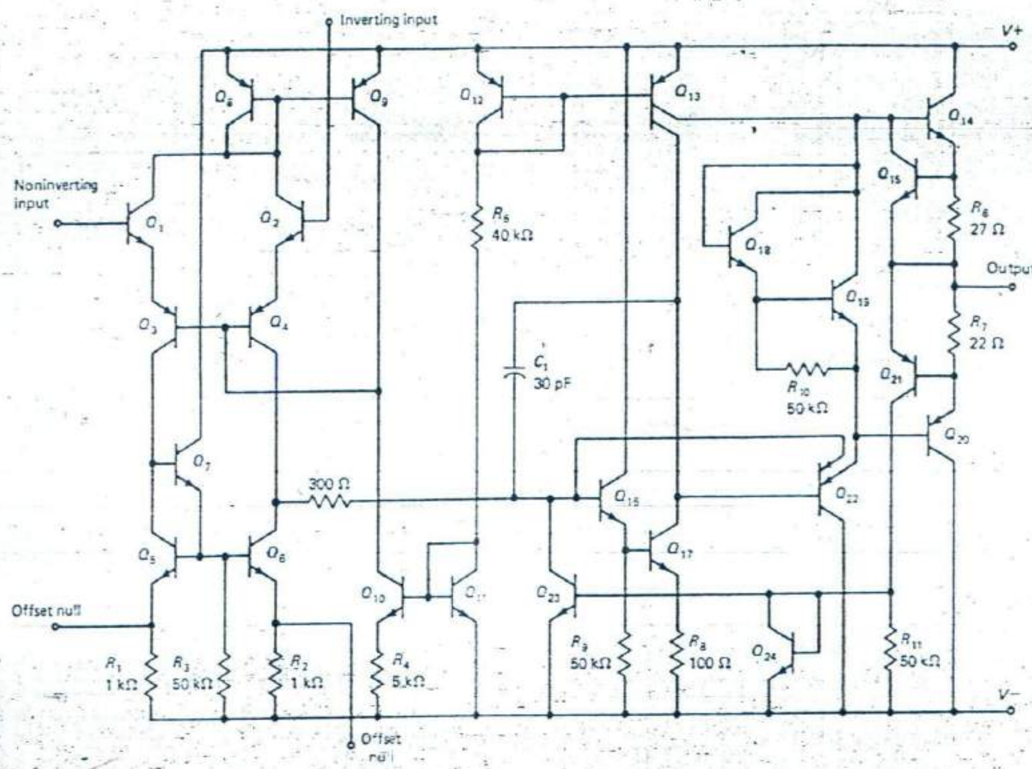


Fig. 6-13 741 operational amplifier equivalent circuit.

sive. Note that the metal case is internally connected to the $V-$ supply. The quad op amp package eliminates the offset null terminals since too many pins would be required. Many applications for operational amplifiers do not require the offset null function, and the quad package may be a good choice in those cases. The dual package uses a separate $V+$ pin for each op amp. Both pins must be energized when both amplifiers are in use.

Operational amplifiers are widely applied because they approach ideal amplifiers, especially for dc and low-frequency signals. Let's see how some specifications for common op amps compare to the ideal. The ideal amplifier has an infinite input impedance so that it can be connected to any signal source with no loading effects. An ordinary op amp, such as the 741, approaches the ideal with an input impedance of 6 MΩ. Super Beta op amps offer input impedances 10 times higher, and BI-FET operational amplifiers are available with FETs in the input amplifier and boast an input impedance of 10^{12} Ω. Another ideal amplifier characteristic is infinite gain. Common op amps provide over 100 dB of gain at low frequencies. Premium devices, called *instrumentation amplifiers*, provide 130 dB of gain at low frequencies. The ideal amplifier would infinitely reject common mode signals. Operational amplifier CMRR ranges from 90 dB for standard types to 130 dB for instrumentation amplifiers. The ideal amplifier has zero output impedance and is capable of driving any load. Common operational amplifiers, such as the 741, can supply at least 5 mA to a 2000-Ω load. Monolithic power devices are capable of up to 1 A of output current, and hybrids with even higher ratings can be found. The ideal amplifier also has infinite bandwidth, meaning it can amplify any signal frequency. In this area the typical op amp falls considerably short of the ideal, with useful gain extending only into the tens of kilohertz region. Wide-band devices are available and extend performance to the tens of megahertz. Many, many applications do not require wide-band performance, and the typical device is all that is required. Modern op amps approach the ideal closely enough to make them very attractive for many applications.

Figure 6-15(a) shows the gain versus frequency response for a 741 op amp. The gain is in excess of 100 dB for low frequencies. Around 7 Hz, the gain begins to drop at a rate of 20 dB per decade as a result of the internal frequency compensation capacitor as previously mentioned. Some op-amps without internal frequency compensation are available but are not very popular since they must be externally compensated with resistors and capacitors. This fea-

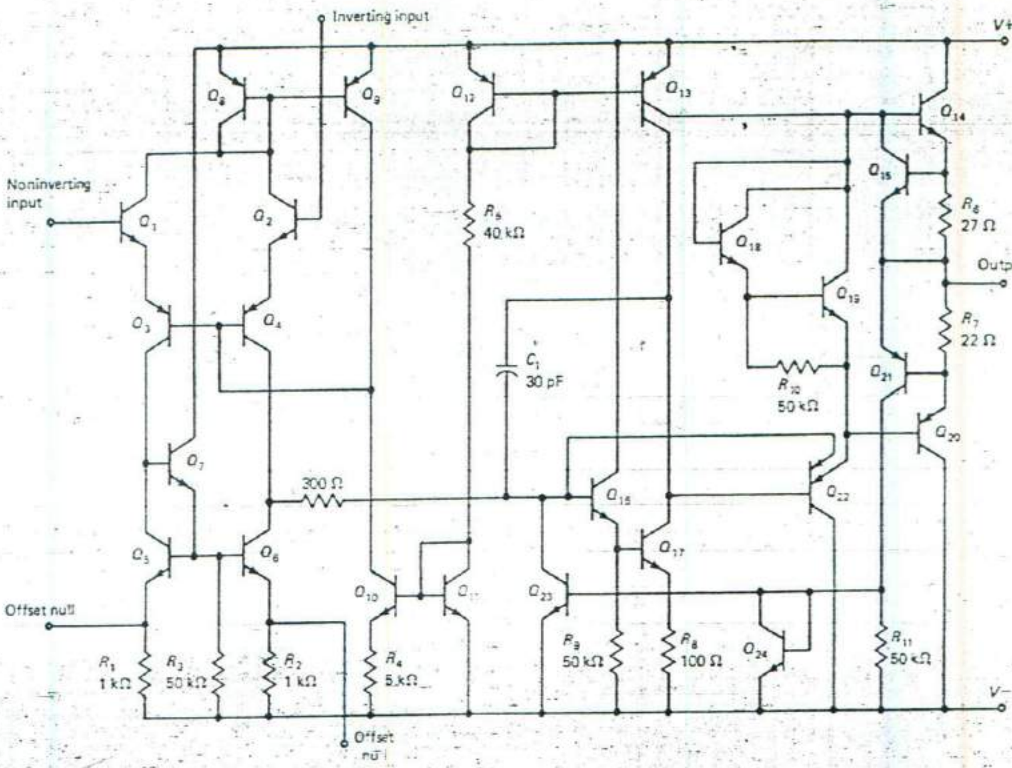


Fig. 6-13 741 operational amplifier equivalent circuit.

sive. Note that the metal case is internally connected to the $V-$ supply. The quad op amp package eliminates the offset null terminals since too many pins would be required. Many applications for operational amplifiers do not require the offset null function, and the quad package may be a good choice in those cases. The dual package uses a separate $V+$ pin for each op amp. Both pins must be energized when both amplifiers are in use.

Operational amplifiers are widely applied because they approach ideal amplifiers, especially for dc and low-frequency signals. Let's see how some specifications for common op amps compare to the ideal. The ideal amplifier has an infinite input impedance so that it can be connected to any signal source with no loading effects. An ordinary op amp, such as the 741, approaches the ideal with an input impedance of 6 MΩ. Super Beta op amps offer input impedances 10 times higher, and BI-FET operational amplifiers are available with FETs in the input amplifier and boast an input impedance of 10^{12} Ω. Another ideal amplifier characteristic is infinite gain. Common op amps provide over 100 dB of gain at low frequencies. Premium devices, called *instrumentation amplifiers*, provide 130 dB of gain at low frequencies. The ideal amplifier would infinitely reject common mode signals. Operational amplifier CMRR ranges from 90 dB for standard types to 130 dB for instrumentation amplifiers. The ideal amplifier has zero output impedance and is capable of driving any load. Common operational amplifiers, such as the 741, can supply at least 5 mA to a 2000-Ω load. Monolithic power devices are capable of up to 1 A of output current, and hybrids with even higher ratings can be found. The ideal amplifier also has infinite bandwidth, meaning it can amplify any signal frequency. In this area the typical op amp falls considerably short of the ideal, with useful gain extending only into the tens of kilohertz region. Wide-band devices are available and extend performance to the tens of megahertz. Many, many applications do not require wide-band performance, and the typical device is all that is required. Modern op amps approach the ideal closely enough to make them very attractive for many applications.

Figure 6-15(*a*) shows the gain versus frequency response for a 741 op amp. The gain is in excess of 100 dB for low frequencies. Around 7 Hz, the gain begins to drop at a rate of 20 dB per decade as a result of the internal frequency compensation capacitor as previously mentioned. Some op-amps without internal frequency compensation are available but are not very popular since they must be externally compensated with resistors and capacitors. This fea-

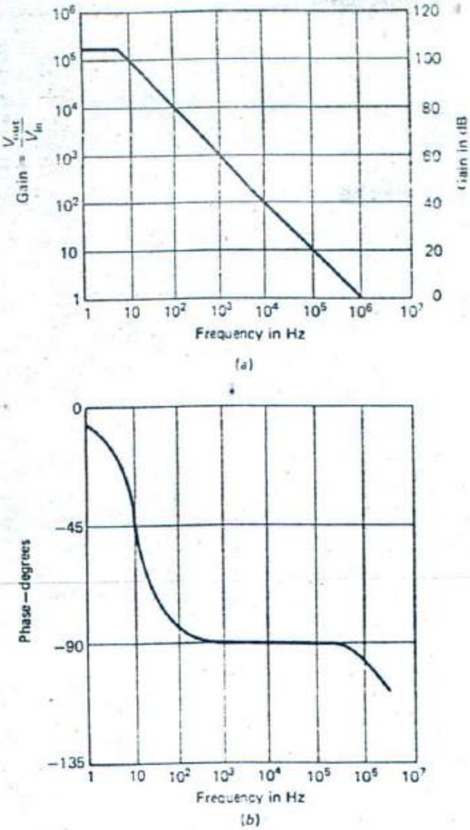


Fig. 6-15 Open loop performance curves for the 741 op amp. (*a*) Gain versus frequency. (*b*) Phase versus frequency.

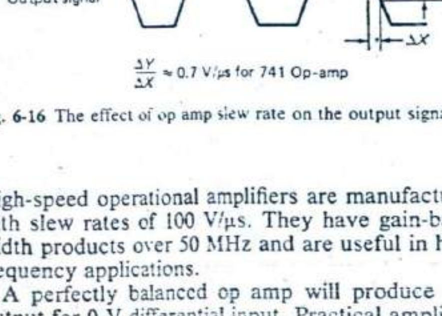


Fig. 6-16 The effect of op amp slew rate on the output signal.

High-speed operational amplifiers are manufactured with slew rates of 100 V/μs. They have gain-bandwidth products over 50 MHz and are useful in high-frequency applications.

A perfectly balanced op amp will produce 0 V output for 0-V differential input. Practical amplifiers show a slight dc voltage at the output due mainly to slight imbalances in the differential input stage. A

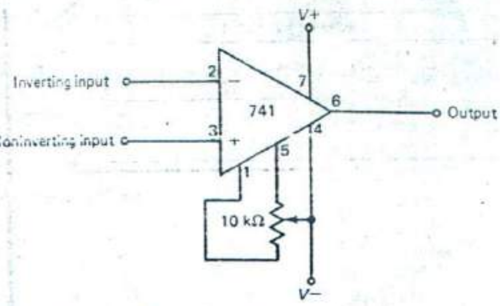


Fig. 6-17 Voltage offset null circuit.

slight differential input voltage of 1 mV or so will typically be required to zero the output. This error is called the *input offset voltage* and can be corrected in critical applications with the voltage offset null circuit shown in Fig. 6-17. A 10-kΩ potentiometer is connected across the offset null terminals 1 and 5. The wiper arm is connected to the negative supply. The potentiometer is adjusted for 0 V at pin 6 when the differential input is 0. You may wish to refer to Fig. 6-13 to verify how the offset null circuit trims the emitter resistors in the differential input amplifier to achieve balance. The offset null connection is not required in all applications, and pins 1 and 5 are left floating in those cases. The pin numbers used in Fig. 6-17 are for the mini-dip and metal packages shown in Fig. 6-14.

REVIEW QUESTIONS

12. Which op amp input is in phase with the output, and how is it usually marked on a schematic diagram?

13. If the gain of an internally compensated op amp is 80 dB at 100 Hz, what will it be at 10 kHz?

14. What is the input impedance of an ideal amplifier? Why?

15. What is the output impedance of an ideal amplifier? Why?

16. Refer to Fig. 6-15. What is the gain bandwidth product at 100 Hz? Is it constant at other frequencies?

17. Calculate the power bandwidth for an op amp with a slew rate of 100 V/μs and an output signal of 20 V peak-to-peak (10-V peak).

18. If the amplifier of question 17 has a gain bandwidth product of 65 MHz, what limits its high-frequency performance for large output signals?

6-4
OP AMP APPLICATIONS

Practical operational amplifiers approach the ideal amplifier in several important ways. The approach is close enough to allow simple and straightforward analysis techniques to be used with good results. The techniques are based on several assumptions. First, there is no input current in an ideal amplifier since it has infinite impedance. Second, the output impedance is zero. Third, the gain is infinite and reduces the differential input to zero when negative feedback is used. Refer to Fig. 6-18. Resistor R_2 provides negative feedback because it connects the output back to the inverting input. With infinite gain, it will not be possible to make the voltage at the inverting input at all different from the voltage at the noninverting input, which is at ground. Suppose, for example, that the input terminal to the left of R_1 is driven positive by a signal. The inverting input will also try to go positive, but the gain of the amplifier will immediately produce a negative-going output which feeds back through R_2 to cancel the positive change and keep the inverting input at ground potential. It is not possible to have any voltage difference across the + and − inputs since the feedback acts to cancel it. The noninverting input of Fig. 6-18 is grounded through R_3. We have assumed no input current, and there is no voltage drop across R_3. Both inputs will remain at ground potential even when the amplifier is driven with a signal. The feedback keeps the inverting input at ground potential, and that terminal is considered to be a *virtual ground*.

The voltage gain for the inverting amplifier of Fig. 6-18 is easy to derive. The virtual ground sets the current in R_1 at V_{in}/R_1 and the current in R_2 at V_{out}/R_2. Because there is no amplifier input current, these currents are equal:

$$\frac{-V_{in}}{R_1} = \frac{V_{out}}{R_2}$$

The input voltage is indicated as negative because it is inverted from the output voltage. Solving for V_{out} gives

$$V_{out} = -V_{in} \times \frac{R_2}{R_1}$$

The voltage gain for the inverting amplifier in Fig. 6-18 is equal to the feedback resistor value divided by the input resistor value. For example, if R_2 is 100 kΩ and R_1 is 1 kΩ, the voltage gain will be −100. Again, the minus sign indicates the phase inversion between the input and the output. Now refer to Fig. 6-19. The open loop gain of the typical op amp is several hundred thousand. By using negative feedback, the gain is reduced to an absolute value of

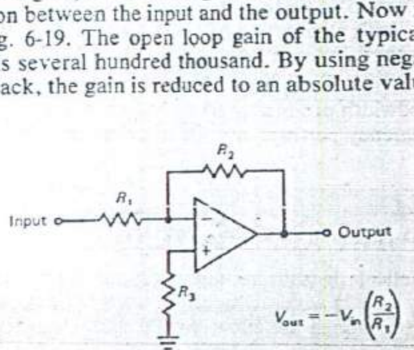


Fig. 6-18 Inverting amplifier.

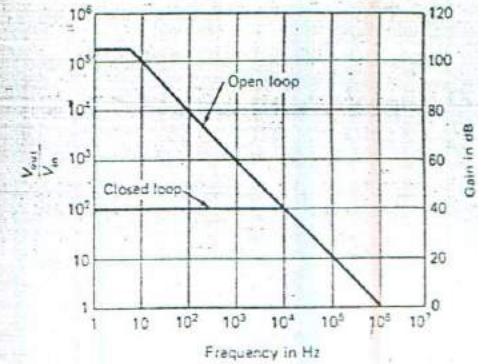


Fig. 6-19 Closed loop frequency response.

100 (the absolute value ignores the minus sign). This value is called the *closed loop gain*. One of the advantages of closing the loop and reducing the gain is increased bandwidth. Figure 6-19 shows that the closed loop gain is constant to a frequency of 10 kHz, giving a bandwidth of 10 kHz. The open loop bandwidth is only 7 Hz. This greatly improved bandwidth makes the amplifier more useful for many applications.

It is easy to predict the bandwidth when the amplifier is operated with negative feedback. Calculate the gain by using the ratio of the feedback resistor to the input resistor. Draw a line from the calculated gain value on the vertical axis until the open loop gain plot is intersected. This intersection is called the *corner frequency* and represents the closed loop bandwidth of the amplifier. The plot in Fig. 6-19 does not show it, but the gain is down 3 dB at the corner frequency. For the example given, the gain will be 37 dB (40 dB − 3 dB) at a frequency of 10 kHz. The −3-dB point is the standard limit when specifying amplifier bandwidth and is also known as the *cutoff frequency*. Note that the amplifier does not abruptly stop working beyond this frequency, but its gain drops at a rate of 20 dB per decade for signals higher in frequency.

The input impedance of the inverting amplifier in Fig. 6-18 is equal to the input resistor R_1 because of the virtual ground at the inverting input. In our example, the signal source would see a load of 1000 Ω. The output impedance of the amplifier is equal to the inherent output impedance divided by the loop gain. *Loop gain* is equal to the open loop gain divided by the closed loop gain. The open loop gain is 200,000, and the closed loop gain is 100 for our example, yielding a loop gain of 200,000/100 = 2000. The inherent output impedance is 75 Ω, and the closed loop output impedance is equal to 75/2000, or 0.04 Ω. This indicates that adding negative feedback also decreases the output impedance of the amplifier. The low output impedance is advantageous because the circuit can deliver a signal to almost any load, provided that the current capabilities of the op amp are not exceeded. Resistor R_3 in Fig. 6-18 may be replaced by a direct connection to ground with only minor impact on circuit performance. When used, it is usually selected to have a value equal to the parallel resistance of R_1 and R_2 (990 Ω in our example). The idea is to produce identical voltage drops due to the input bias current at both amplifier inputs. Even though we have assumed zero input current, real op amps do have a small input current. Balance reduces offset error. However, the input current is only 30 nA in a typical op amp, and the imbalance created by directly grounding the noninverting input can be ignored in most applications.

Thus far we have seen that it is easy to use an op amp as an inverting amplifier. The merits of negative feedback for increasing bandwidth and decreasing output impedance have been established. Now let's look at some other applications. Figure 6-20 shows a voltage follower which is a noninverting amplifier since the input signal is applied to the noninverting input. The voltage gain of the circuit is 1. This circuit is useful even though it has no voltage gain. It has a very high input impedance, which is equal to the inherent input impedance of the op amp. Since this impedance is 6 MΩ in a standard op amp, the voltage follower makes an excellent isolation amplifier. The inverting amplifier that we looked at previously exhibits a much lower input impedance due to the virtual ground created by negative feedback. A high-impedance signal source will suffer loading effects when connected to such an amplifier. Therefore, it is sometimes necessary to isolate the signal source by connecting it to a voltage follower and then connecting the voltage follower output to the next stage.

A noninverting amplifier with gain is shown in Fig. 6-21(a). The gain for this circuit is derived by starting at the inverting input. Notice that the voltage there is a function of the output voltage and the divider network formed by R_1 and R_2:

$$V_{inv} = V_{out} \times \frac{R_1}{R_1 + R_2}$$

Once again we make the assumption that both amplifier inputs are at the same potential; therefore,

$$V_{in} = V_{out} \times \frac{R_1}{R_1 + R_2}$$

Solving for V_{out} gives

$$V_{out} = V_{in} \times \frac{R_1 + R_2}{R_1}$$

Suppose R_1 in Fig. 6-21(a) is a 1-kΩ resistor and R_2 is a 22-kΩ resistor. The voltage gain will be (22 + 1)/1, or 23. The bandwidth will be 40 kHz and

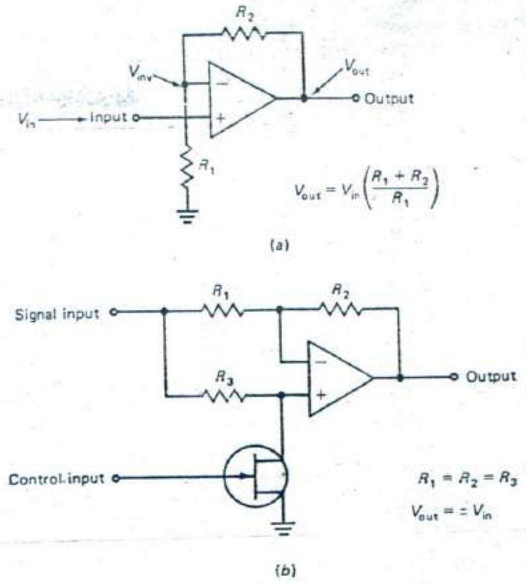


Fig. 6-21 Amplifiers. (a) Noninverting. (b) Switchable.

can be verified by using Fig. 6-19. The input impedance is equal to the input impedance of the op amp, which is high. The output impedance will be reduced by a factor equal to the loop gain and will therefore be quite low.

Figure 6-21(b) shows a switchable amplifier that will invert or not invert the input signal, depending on the control signal applied to the gate of the FET. Suppose the control signal is 0 V. The FET will be on, grounding the noninverting input of the op amp and the right end of R_3. The amplifier will function as an inverting amplifier with a gain of 1 since $R_1 = R_2$. If the control input is made negative enough to cut off the FET, the amplifier will switch to the noninverting mode since the input signal now also drives the + input of the op amp. Once again, we can assume no differential input due to the large gain of the op amp, and therefore the signal voltage at the left end of R_1 will be equal to the signal voltage at the right of R_1. There will be no signal current in R_1, and it appears as an infinite resistance as far as the signal circuit is concerned. Now, look at the gain equation in Fig. 6-21(a) to determine that the voltage gain approaches 1 as R_1 approaches infinity. Therefore, the voltage gain of the switchable amplifier is −1 or +1, depending on the control signal.

Operational amplifiers are also capable of subtraction and addition. Figure 6-22 shows an amplifier that subtracts one input signal from another and amplifies the difference. Input 1 is applied to the inverting terminal and is subtracted from input 2, which is applied to the noninverting terminal. A summing amplifier is shown in Fig. 6-23. It can be used to add two or more ac or dc input signals. All inputs are

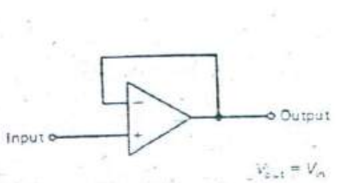


Fig. 6-20 Voltage follower.

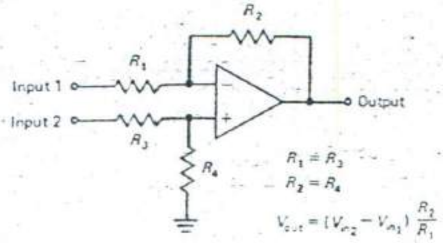


Fig. 6-22 Subtracting (differential) amplifier.

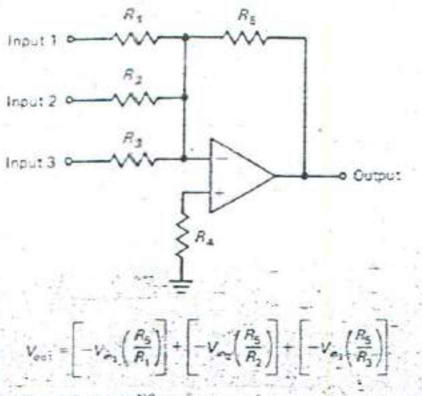


Fig. 6-23 Summing amplifier.

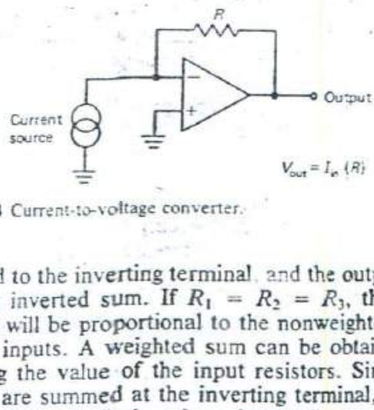


Fig. 6-24 Current-to-voltage converter.

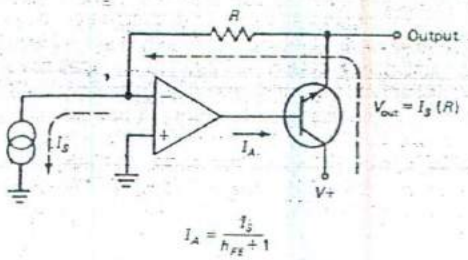


Fig. 6-25 Current-to-voltage converter with current boost.

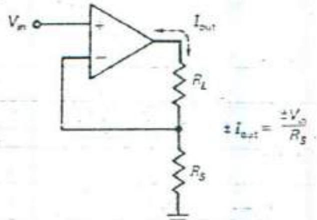


Fig. 6-26 Voltage-to-current converter.

applied to the inverting terminal, and the output will be the inverted sum. If $R_1 = R_2 = R_3$, then the output will be proportional to the nonweighted sum of the inputs. A weighted sum can be obtained by varying the value of the input resistors. Since the inputs are summed at the inverting terminal, which is a virtual ground, there is no interaction among the inputs. A signal at any input will not cause any effects at the other inputs.

An operational amplifier used as a current-to-voltage converter is shown in Fig. 6-24. The output voltage is equal to the input current times the feedback resistor. Since the inverting terminal is a virtual ground, this circuit places a short circuit across the current source. This is a valuable asset for some signal sources since they must be short-circuited for accurate measurements. Other approaches, such as using a current meter or a load resistor, may adversely affect measurement accuracy because of the loading effect of the meter or the load resistor. The internal impedance of the meter adds to the effective series resistance of the current source (assuming that it is not an ideal current source) and unloads the circuit. The unloading causes the measured current to be less than the true short-circuit current. The current-to-voltage converter circuit eliminates unloading effects for improved accuracy. Figure 6-25 shows a current-to-voltage converter circuit with current boost. The typical op amp can supply only milliamperes. This circuit extends the current capacity into the ampere range by using the current gain of an external NPN transistor. The current gain from the base terminal to the emitter terminal is $h_{FE} + 1$, and the required operational amplifier output current is reduced by this factor.

Figure 6-26 shows an op amp voltage-to-current converter. This circuit is also called a *transconductance amplifier.* Resistors R_L and R_S divide the output voltage for the inverting input. Once again, we can assume that there will be no difference between the inverting terminal voltage and the noninverting terminal voltage; therefore,

$$V_{out} = \frac{V_{in}(R_S + R_L)}{R_S}$$

The output current is predicted by Ohm's law:

$$I_{out} = \frac{V_{out}}{R_S + R_L}$$

Combining, we obtain:

$$I_{out} = \frac{V_{in}}{R_S}$$

This equation shows that the current in the load resistor (R_L) is independent of the value of the load resistor. It is a function of the input voltage and the value of the sense resistor (R_S). By making the input voltage constant, this circuit can serve as a constant current source or sink.

REVIEW QUESTIONS

19. Refer to Fig. 6-18. Calculate the voltage gain if $R_1 = 10\ k\Omega$ and $R_2 = 100\ k\Omega$. What is the input impedance of the amplifier?

20. Use Fig. 6-19 to find the bandwidth of the amplifier in question 19.

21. Refer to Fig. 6-21. Calculate the gain if $R_1 = 1\ k\Omega$ and $R_2 = 100\ k\Omega$. What is the input impedance of this circuit if the op amp is a standard type?

22. Refer to Fig. 6-22 and assume that all resistors are 10 kΩ in value. Calculate the output voltage if input 1 = −1 V and input 2 = 2 V.

23. Refer to Fig. 6-23. Assume that $R_1 = R_2 = R_3 = 10\ k\Omega$ and that $R_5 = 47\ k\Omega$. Calculate the output voltage if input 1 = +0.5 V, input 2 = +0.7 V, and input 3 = +1 V.

24. Are the inputs in question 23 weighted or nonweighted? How could the circuit be changed to achieve the other condition?

6-5 NONLINEAR APPLICATIONS

In a *nonlinear circuit* the output signal is not a replica of the input signal. For example, refer to Fig. 6-27. The input signal is a sine wave, and the output signal is a rectangular wave. This circuit is called a *comparator* and can be used to convert other waveforms to rectangular. It is also used to compare a signal to a reference voltage and change output states when the reference threshold is crossed. Since the op amp is running open loop, the gain is very high, and only a few millivolts difference between the input voltage and the reference voltage (V_{REF} in Fig. 6-27) will drive the output to positive or negative saturation. Operational amplifiers can be used as comparators, but their slew rate may limit performance for many applications. Integrated circuit comparators that have optimized characteristics such as fast switching, wide supply range, and high output current are available.

Figure 6-27 is an inverting comparator because the input is applied to the inverting input of the op amp or comparator IC. When the input signal goes more positive than V_{REF}, the output goes to negative saturation. When the input is less than V_{REF}, the output goes to positive saturation. The output waveform shows that the saturation voltages are a little less than the supply voltages. Some comparator ICs, such

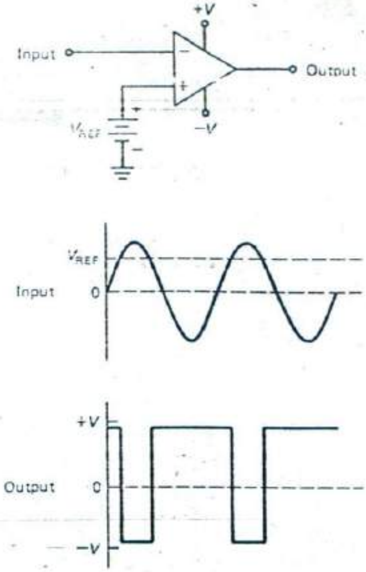


Fig. 6-27 Inverting comparator.

as the type 311, will require an external pull-up resistor on the output. This is convenient when an analog signal must be converted for another circuit where the voltage levels are different. For example, the comparator IC can be connected to a 15-V dual supply, and its output can be pulled up to a separate +5-V supply. The output will swing between 0 and +5 V. This swing is compatible with many digital circuits.

The reference voltage in Fig. 6-27 can be changed for different results. If the noninverting input of the IC is grounded, the reference voltage will be zero. This will change the duty cycle of the output waveform to 50 percent with the input waveform that is shown because the output will switch at the zero crossings. If the reference voltage is made greater than the peak value of the input, the output will remain at the positive saturation point. By making the reference voltage adjustable, the circuit could be used to produce a rectangular output with varying duty cycle. The circuit can also be reconfigured for noninverting operation, as shown in Fig. 6-28. In this example, the reference voltage is negative. When the input signal goes more negative than the reference, the output is driven to negative saturation. When the input is more positive than $-V_{REF}$ the output is at positive saturation.

Figure 6-29 shows a *window comparator*. This circuit is used to determine whether a voltage or signal is within or without a given range called the *window*. It uses two op amps or comparators and two reference voltages. The waveforms show that the output signal is 0 as long as the input is between +4.5 and +5.5 V. If the signal is outside this window, the

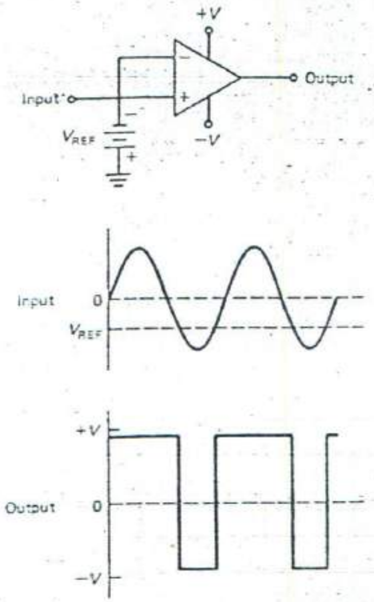


Fig. 6-28 Noninverting comparator.

output is at positive saturation. Suppose the signal is in the window. The top IC will be at negative saturation, since its noninverting input is negative (less than +5.5 V). The bottom IC is also at negative saturation because its inverting input is positive (more than +4.5 V). Both diodes are reverse-biased (off), and the combined output is zero. Now, suppose the input goes more positive than 5.5 V. The top IC goes to positive saturation, D_1 is forward-biased (on), D_2 is still off, and the combined output goes to positive saturation less the drop across D_1. If the input goes below +4.5 V, the bottom IC goes to positive saturation as its inverting input will be negative with respect to its noninverting input. This will forward-bias D_2, and the combined output will be at positive saturation less the drop across D_2. Window comparators can be used to monitor a critical voltage such as a power supply to determine whether or not it is in tolerance.

The waveforms in Fig. 6-30 show the operation of a Schmitt trigger. At first glance this circuit appears to be accomplishing the same result as the inverting comparator circuit since the input waveform is sinusoidal and the output is rectangular. However, there is a difference because positive feedback is used. Note that R_1 connects the IC output back to the noninverting input. This is positive feedback and produces two switching points: the upper threshold point (UTP) and the lower threshold point (LTP). These points are set by the supply voltages and resistors R_1 and R_2. For example, if the IC is energized by a 15-V dual suppy the output will saturate at near +14 V or −14 V. If we assume that R_1 is 22 kΩ and that R_2 is 1 kΩ, we have the information needed to calculate the two threshold points:

$$\begin{aligned} \text{UTP} &= \frac{R_2}{R_1 + R_2} \times (+V_{\text{SAT}}) \\ &= \frac{1\ \text{k}\Omega}{23\ \text{k}\Omega} \times (+14\ \text{V}) \\ &= +0.61\ \text{V} \end{aligned}$$

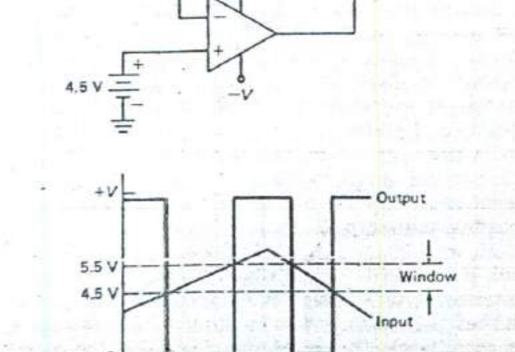


Fig. 6-29 Window comparator.

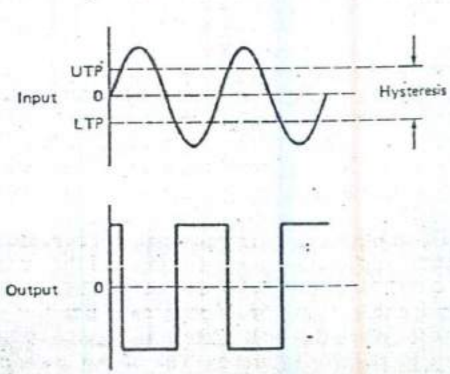


Fig. 6-30 Schmitt trigger.

The LTP is determined by using $-V_{SAT}$ in the preceding equation and calculates to -0.61 V in this example. The waveforms of Fig. 6-30 show how the threshold points work. At the beginning of the input cycle, the output is at positive saturation. Therefore, the voltage divider sets the noninverting input at the UTP. When the input signal exceeds this point, the IC output goes to negative saturation, and the voltage divider now sets the noninverting input at the LTP. The output does not switch again until the input signal goes more negative than the LTP.

The difference between the two threshold points is called *hysteresis*. For our example, the hysteresis will be equal to $+0.61 - (-0.61) = 1.22$ V. Hysteresis is desirable when signals are noisy. Refer to Fig. 6-31, which shows the noisy signal performance for a comparator circuit and for a Schmitt trigger. The output waveform in Fig. 6-31(*a*) has extra transitions. They are caused by noise when the average value of the input signal is near the reference voltage. The noise adds and subtracts from the input signal, and the instantaneous value goes above and below the reference value several times. The output frequency is higher than the input frequency. The Schmitt trigger waveforms in Fig. 6-31(*b*) show no frequency distortion. As long as the hysteresis is greater than the peak-to-peak noise amplitude, extra transitions are eliminated. The hysteresis must be less than the peak-to-peak signal, however, or the output will not switch at all.

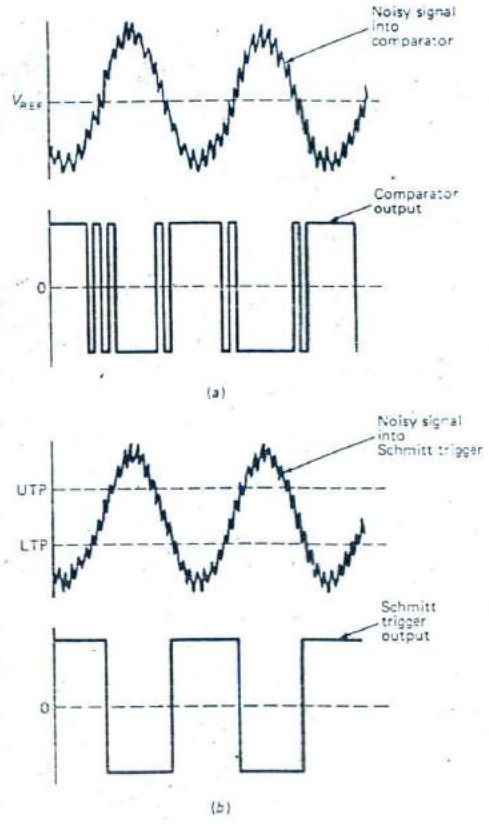


Fig. 6-31 Noisy signal performance. (*a*) Comparator. (*b*) Schmitt trigger.

Figure 6-32 shows an operational amplifier differentiator. A *differentiator* is a circuit that responds to the rate of change of the input signal. It is essentially a high-pass filter, meaning that it produces more output for signals at higher frequencies. Its instantaneous output is proportional to the instantaneous rate of change at its input (or to the slope of the input waveform):

$$V_{out} = -RC \times \frac{\Delta V_{in}}{\Delta T}$$

Assuming an input signal that is changing at a rate of 100 V/s and the circuit values shown in Fig. 6-32:

$$V_{out} = -1 \times 10^4 \times 1 \times 10^{-6} \times \frac{100 \text{ V}}{1}$$
$$= -1 \text{ V}$$

The illustration shows the input and output waveforms. The simple differentiator circuit suffers from high-frequency noise since the gain goes up as frequency does. A practical differentiator will often use a resistor in series with the input to limit the high-frequency gain.

The opposite of differentiation is *integration*. Figure 6-33 shows an op amp integrator. It is essentially a low-pass filter and produces more output for signals at lower frequencies. When an input waveform steps from 0 to -1 V, the output ramps positive at a rate of 100 V/s. The integrator output is proportional to

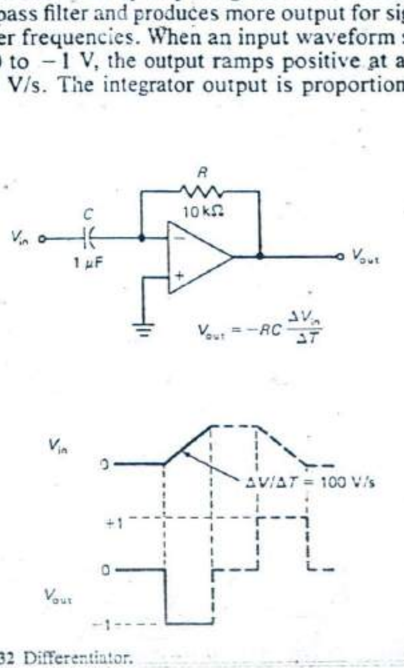


Fig. 6-32 Differentiator.

the product of the amplitude and duration of the input signal (the area under the curve). The output voltage is given by

$$V_{out} = -V_{in} \times \frac{1}{RC} \times T$$

By dropping the T term in the preceding equation, the output is expressed as a rate of change in volts per second. With the input and circuit values shown in Fig. 6-33, the rate of change will be

$$V_{out} = -(-1\ V) \times \frac{1}{1 \times 10^4 \times 1 \times 10^{-6}}$$
$$= 100\ V/s$$

The simple integrator circuit suffers from drift. Any dc offset voltage at the input will cause the integrator to drift and eventually saturate. Practical circuits usually use a high value of resistor in parallel with the feedback capacitor to reduce this drift.

Comparing the waveforms of Fig. 6-32 and Fig. 6-33 shows the opposite natures of differentiation and integration. If the differentiated function is fed to an integrator, its output will be the same as the original function. Figure 6-34 shows circuit performance with other waveforms. The differentiator output shown in Fig. 6-34(*a*) is an inverted cosine wave with a sine wave input. A cosine wave is 90° out of phase with a sine wave. The total phase shift shown is 270° since the output is differentiated and inverted (90 + 180). Note that the output is zero when the input is peaking, because the instantaneous rate of change of a sine wave at peak is zero. The maximum rate of change occurs during the zero crossing at the output peaks at this time. The integrator waveforms of Fig. 6-34(*b*) show that an inverted cosine at the input produces a sine wave at the output. Once again it is demonstrated that when a signal is differentiated and then integrated, the original signal will appear at the output of the integrator.

Integrators and comparators can be used together to achieve some other valuable functions. Figure 6-35 shows a voltage-to-frequency converter. The positive input voltage is applied to the integrator,

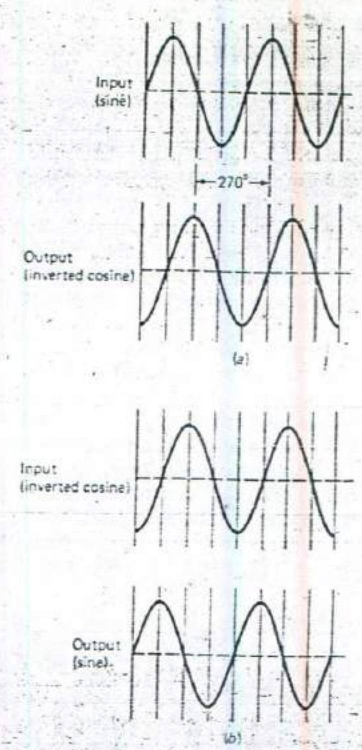


Fig. 6-34 Output waveforms. (*a*) Differentiator. (*b*) Integrator.

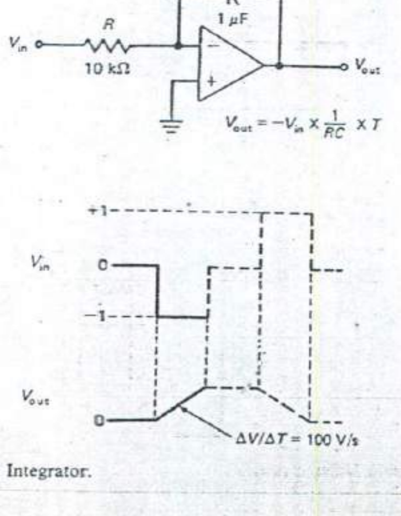


Fig. 6-33 Integrator.

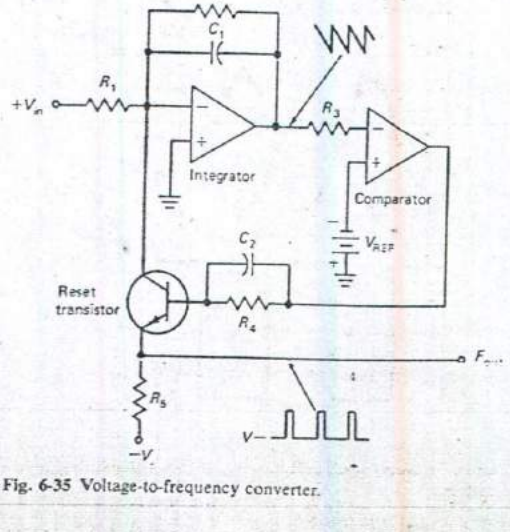


Fig. 6-35 Voltage-to-frequency converter.

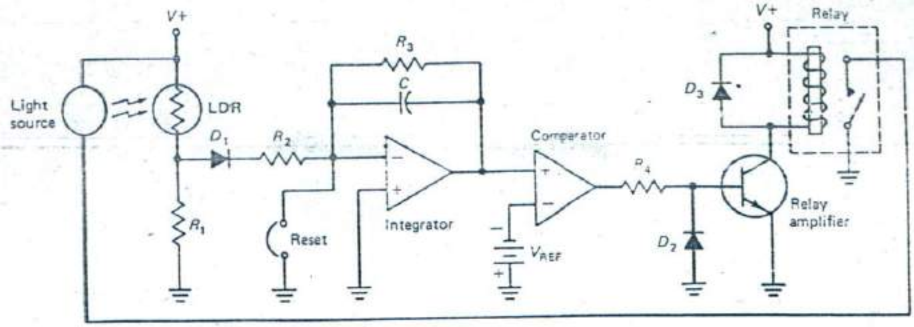


Fig. 6-36 Light integrator.

which begins ramping negative at a rate proportional to V_{in}. When the integrator ramp exceeds the value of V_{REF}, the comparator output switches to positive saturation and turns on the reset transistor. With the transistor saturated (low resistance), the integrator output is rapidly driven positive because $V-$ is applied to the input. The comparator then saturates negative, and the reset transistor turns off so the next integration cycle can begin. The output is a series of positive-going pulses, and the pulse frequency is proportional to V_{in}. For example, if the value of V_{in} is doubled the integrator will ramp negative at twice the rate, and the reset pulses will appear in half the time. The output frequency will double. Capacitor C_2 is a speed-up capacitor. It decreases the switching time of the reset transistor. The voltage-to-frequency converter can also be called an *analog-to-digital converter*. The input voltage is analog, and the output is a digital pulse train.

Figure 6-36 shows a light integrator. The light-dependent resistor (LDR) shows a decrease in resistance when it is illuminated. It acts as part of a voltage divider along with R_1. The divided positive voltage is applied to the integrator, which ramps negative at a rate that is proportional to the light intensity. When the ramp exceeds $-V_{REF}$, the comparator output saturates negative. This turns off the relay amplifier, and the contacts open and turn off the light source. Pushing the reset button will discharge the integrator and begin another cycle. The value of integration in this application is that the circuit ensures the proper light dosage regardless of supply variations, light source variations, or light transmission path changes. The integrator/comparator combination will not turn the light source off until the proper sum has been acquired. This type of a circuit can be used in various photochemical processes and is far superior to a simple timer, which is subject to errors when light intensity fluctuates. Diode D_1 prevents the divider from discharging the integrator when the light source fluctuates or is momentarily interrupted. Diode D_2 clamps the base voltage of the transistor when the comparator output is negative. Diode D_3 protects the transistor from relay coil transients.

Ordinary diode circuits cannot be used to rectify signals in the millivolt range because several tenths of a volt are required to turn diodes on. A precision rectifier circuit that overcomes this limitation is shown in Fig. 6-37. The op amp on the left, along with its diodes and resistors, forms a precision half-wave rectifier. When the input signal goes positive, the output goes negative, and D_1 turns on to conduct the feedback current. Diode D_2 is reverse-biased, and the left end of R_3 remains at zero as long as the input is positive. When the input goes negative, the output of the left op amp goes positive. Now D_1 is off, and D_2 is on and supplies the feedback current through R_2. Resistor R_1 = resistor R_2 and the signal

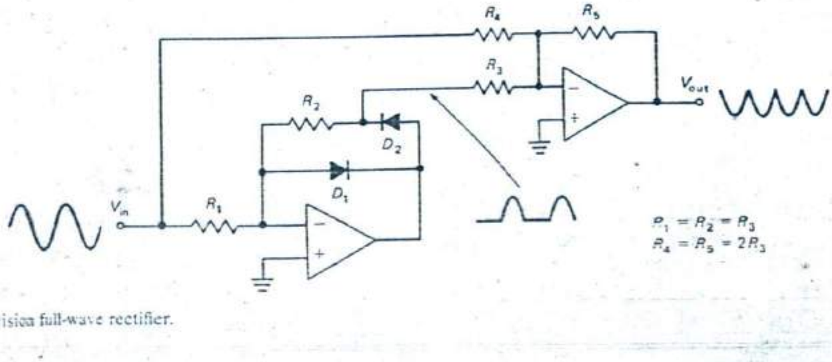


Fig. 6-37 Precision full-wave rectifier.

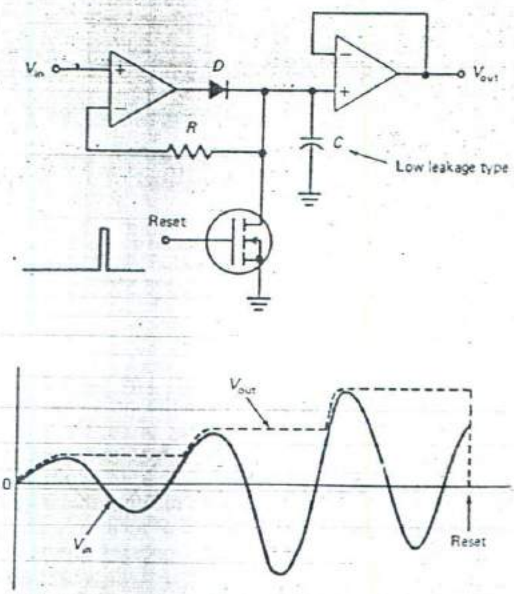


Fig. 6-38 Positive peak detector.

at the cathode of D_2 is an inverted replica of the negative portion of the input signal. If a precision half-wave rectifier is all that is required, the other op amp and its associated parts can be eliminated. Full-wave operation is realized by combining the half-wave signal with the original signal in the second op amp, which serves as a weighted adder. Its feedback resistor is twice the value of R_3, and the half-wave signal receives twice the gain that V_{in} receives.

Figure 6-38 shows a positive peak detector circuit. It "remembers" the greatest positive input value until it is reset. The op amp at the left serves as a noninverting amplifier. The diode in series with its output allows the capacitor to charge positive with respect to ground. If the input signal goes less positive, the diode turns off and prevents the capacitor from being discharged. The op amp at the right acts as a voltage follower. Its high input impedance allows the capacitor to retain its charge for long periods of time. The circuit is reset at the end of the sampling interval by applying a positive pulse to the enhancement mode MOSFET.

REVIEW QUESTIONS

25. Refer to Fig. 6-27. What will happen to the duty cycle of the output if V_{REF} is reversed?

26. What will have to be added to Fig. 6-27 if the IC is a type 311 comparator?

27. Refer to Fig. 6-29. Will the circuit work as a window comparator if the reference supplies are reversed?

28. Refer to Fig. 6-30. Assume that $R_1 = 4.7$ kΩ, $R_2 = 1$ kΩ, and V_{SAT} is equal to plus and minus 11 V. Calculate UTP and LTP.

29. What output waveform would you expect from the circuit in Fig. 6-32 with a triangular input wave?

30. Refer to Fig. 6-32. If the input is going negative at a rate of 50 V/s, what is the output voltage?

31. Refer to Fig. 6-33. If the input wave is square, what will the output be?

6-6
SPECIAL FUNCTIONS

It is often desirable to have a differential amplifier that has high gain, high input impedance, and a high CMRR. Using an operational amplifier produces conflicting design requirements in these cases since the gain is decreased by making the input resistors high in value. Conversely, making the input resistors low (for good gain) lowers the input impedance. You may wish to refer again to Fig. 6-22 to verify this point. It is also difficult to change gain because resistor ratios must be closely matched (within 0.1 percent). These op amp limitations have led to the development of a special circuit or device called the *instrumentation amplifier* (IA).

Figure 6-39(*a*) shows how an IA can be "built up" by using three op amps. The input section uses two devices, and the signal is fed to the noninverting input in each case. This meets the requirement of

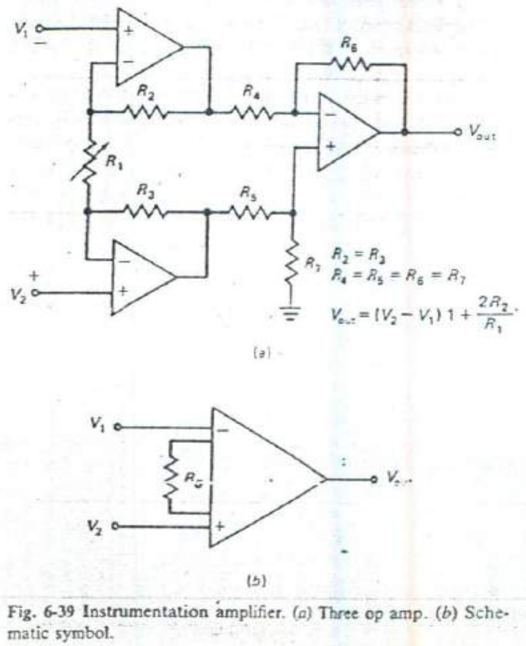


Fig. 6-39 Instrumentation amplifier. (*a*) Three op amp. (*b*) Schematic symbol.

high input impedance. The output section uses a single amplifier to combine the two signals differentially. The bottom input (V_2) is noninverting because it drives the noninverting input of the output amplifier. The top input (V_1) is inverting because it drives the inverting input of the output amplifier. Note that the gain is set by a single resistor in the sense that R_1 can be changed without the need to change R_2 or R_3. In fact, R_1 can be made adjustable, a great convenience for some applications. Changing the gain has no effect on the input impedance of the IA.

Instrumentation amplifiers are also available in a single package. They can be monolithic or hybrid integrated circuits. Figure 6-39(*b*) shows the schematic symbol commonly used for a single-package device. Resistor R_G sets the gain of the IA along with an internal precision feedback network. The National LH0038 is one example of a hybrid IA and is available in a 16-pin dual-inline package. It is shown in Fig. 6-40 in a bridge amplifier application. It features a low input offset voltage of 25 μV and a low offset drift of no more than 0.25 μV/°C. It has a CMRR of 120 dB and can amplify low-level signals in the presence of high common mode noise. The internal gain setting resistors are precision thin-film types and exhibit excellent thermal tracking. The voltage gain is set by using pins 5 through 10 on the IC package, and a range of 100 to 2000 is available. With pin 7 jumpered to pin 10 as shown in Fig. 6-40, the voltage gain is 1000.

The bridge circuit is a very popular arrangement for accurate measurements using certain types of transducers such as strain gauges. Chapter 9 will treat transducers in detail, but a brief discussion here is appropriate to demonstrate the need for precision amplification. A *strain gauge* is a wire device that is used to measure strain (elongation or compression) of some member of a physical system. If a strain stretches the wire in the gauge, it causes its resistance to increase. The bridge circuit of Fig. 6-40 is an excellent arrangement for a strain gauge since the bridge balance will be affected if the gauge is one element of the circuit, such as R_4. Also, a second identical gauge can be used in the R_3 leg of the bridge. If the two gauges track thermally, then temperature effects are canceled since bridge balance will not be affected when both elements change by the same amount. The thermal (compensation) gauge will be mounted in such a way that it will not react to the axis of strain. The output of the bridge will therefore be a function of strain only.

Ideally, the amplifier that follows a precision bridge circuit should not degrade the accuracy of the signals. It should also reject common mode hum and noise picked up on the bridge elements and the interconnecting cable. The IA of Fig. 6-40 has the high performance and thermal stability required. The guard output (pin 11) is maintained at the common mode voltage and greatly reduces noise pickup since it also maintains the shield at the common mode voltage. The output sense and ground sense terminals are utilized in those applications in which errors in load voltage can occur. For example, the output (pin 1) may be used to drive a buffer amplifier for more current capacity. The load voltage at the output of the buffer amplifier will show some error due to drops in the buffer amplifier. The error is eliminated by connecting the ground sense and output sense directly to the load.

Operational amplifiers are normally energized from a dual supply. In some cases this arrangement is not convenient. Figure 6-41 shows that it is possible to "float" an op amp across a single supply. Note that the $-V$ terminal is grounded and that a voltage divider "floats" the noninverting input terminal at half the supply voltage. The dc output voltage of the op amp will therefore be at one-half the supply voltage. This arrangement is convenient for ac amplifiers, and a coupling capacitor can be used in series with R_1 to block any dc component in the input signal. The blocking capacitor will be required if the signal source is at dc ground because the output will be driven to positive saturation without it. It may also be necessary to use a bypass capacitor across R_3 to eliminate any ac noise at the noninverting input.

A special type of amplifier called a *Norton op amp* offers biasing advantages for single-supply operation. It is a current-differencing amplifier and uses a current mirror instead of the typical differential pair to achieve the noninverting input function. Figure 6-42 shows a partial schematic for a Norton amplifier. The common emitter input amplifier uses a current source

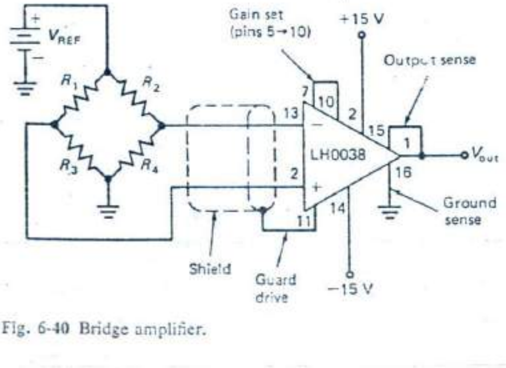


Fig. 6-40 Bridge amplifier.

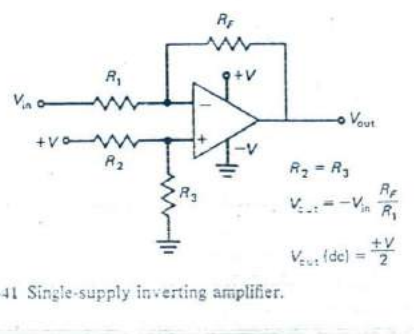


Fig. 6-41 Single-supply inverting amplifier.

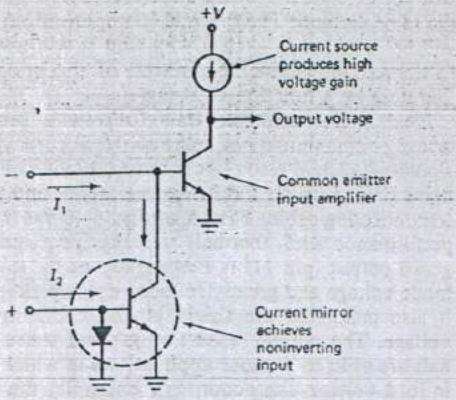


Fig. 6-42 Partial schematic of Norton amplifier.

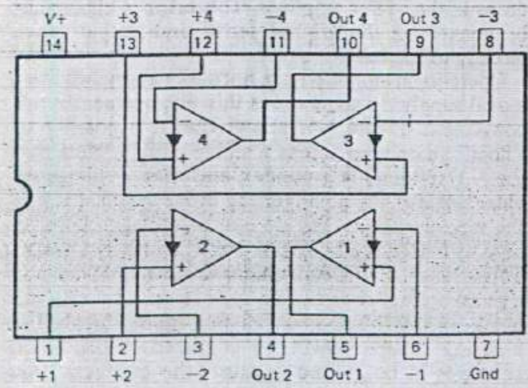


Fig. 6-43 Pin configuration of LM3900.

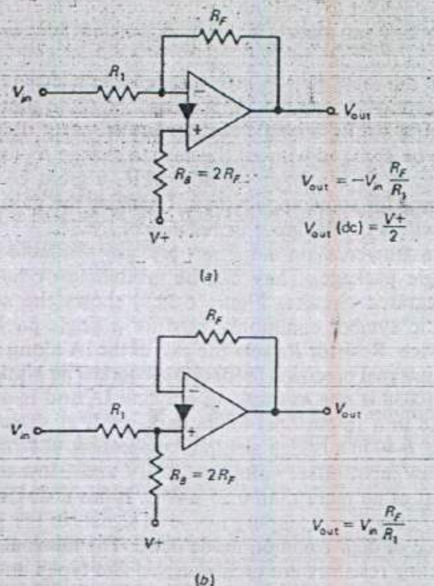


Fig. 6-44 Norton amplifier circuits. (*a*) Inverting. (*b*) Noninverting.

as a load for high-voltage gain. A current source has a very high internal impedance; any change in collector current will therefore produce a large swing in collector voltage. The noninverting input function is achieved with a *current mirror circuit*. Any current flowing into the mirror (I_2) will have the opposite effect on the output voltage when compared to any current flowing into the inverting input (I_1). For example, if I_2 increases, the mirror transistor will turn on harder and remove some of I_1 from the base circuit. This has the same effect as decreasing I_1.

The pinout for an LM3900 Norton op amp is illustrated in Fig. 6-43. It contains four separate amplifiers. It is designed for single-supply operation, pin 7 is grounded, and the positive supply is applied to pin 14. The supply range is 4 to 36 V dc. The arrow on each amplifier symbol denotes that it is a Norton type. Figure 6-44 shows two single-supply circuits utilizing the Norton op amp. The noninverting input is biased with a single resistor, R_B. When this resistor is twice the value of the feedback resistor, the dc output voltage is equal to half of the supply voltage.

Filters are circuits used to remove unwanted frequency components from a signal. They are useful for improving signal-to-noise ratio and for rejecting undesired signals. Practical filters can be built by using only passive devices such as resistors, capacitors, and inductors. Inductors are expensive components and are often physically large in low-frequency designs. It is possible to eliminate the need for inductors by using active devices such as op amps. Filters designed this way are known as *active filters*. They can save space and weight and usually cost less than passive designs. They also eliminate filter loss and are easier to apply since they have a high input impedance and a low output impedance.

Figure 6-45 shows the circuit and frequency response plot for an active low-pass filter. It is a Butterworth design and provides an attenuation of 40 dB per decade for signals beyond the cutoff frequency f_C. The plot shows that the gain is −3 dB at the cutoff frequency. Two or more filter sections can be cascaded for improved attenuation of out-of-band signals. Cascading will affect the cutoff frequency, however. If two filters with the same cutoff frequency are cascaded, the overall gain will become −6 dB at the original cutoff frequency, and the gain roll-off will be 80 dB per decade above the cutoff frequency. The cutoff frequency is always specified at the −3-dB point and will now occur at a slightly lower frequency. The circuit is easy to design. Suppose a cutoff frequency of 1 kHz is required. The design procedure can begin with a selection of resistor or capacitor values depending on which is more convenient. If a 0.005-μF capacitor is desirable for

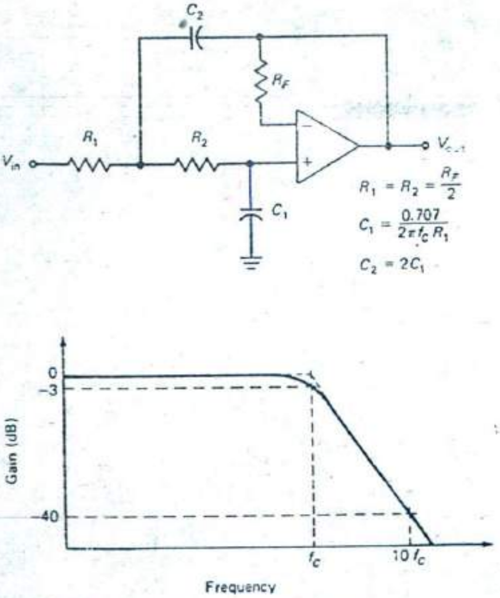


Fig. 6-45 Unity gain low-pass Butterworth filter.

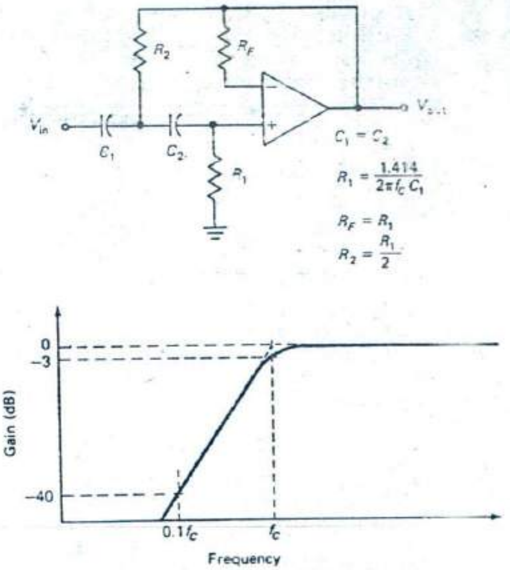


Fig. 6-46 Unity gain high-pass Butterworth filter.

C_1, the equation from Fig. 6-45 can be rearranged to solve for R_1:

$$R_1 = \frac{0.707}{6.28 \times 1000 \times 0.005 \times 10^{-6}} = 22.5\ \text{k}\Omega$$

The feedback resistor will have to be twice that value, or 45 kΩ, and C_2 will have to be twice C_1, or 0.01 μF.

The circuit for a unity gain high-pass Butterworth filter is shown in Fig. 6-46. It produces an attenuation of 40 dB per decade for signals lower than the cutoff frequency. It is useful for the reduction of low-frequency noise on a signal. Additional sections may be cascaded to achieve better attenuation for out-of-band signals. Cascading will shift the cutoff frequency upward a little. The design process is straightforward and usually begins with a selection of a capacitor value. Suppose the cutoff frequency is 1 kHz, and 0.005-μF capacitors are available:

$$R_1 = \frac{1.414}{6.28 \times 1000 \times 0.005 \times 10^{-6}} = 45\ \text{k}\Omega$$

The feedback resistor will be the same value, and R_2 will be half this value, or 22.5 kΩ.

Figure 6-47 illustrates the design of a multiple-feedback bandpass filter. This type of filter rejects frequencies above and below the pass band. It is used to select one frequency, called the *resonant frequency* (f_R), out of a range of frequencies. The figure of merit of a bandpass filter is its Q, which is

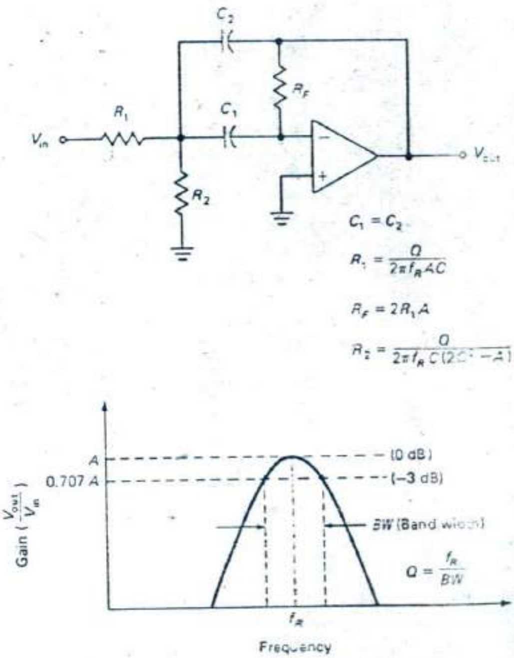


Fig. 6-47 Multiple feedback bandpass filter.

equal to f_R divided by the bandwidth. The *bandwidth* is the difference between the upper and lower cutoff points. As always, the response is −3 dB at the cutoff frequencies. The higher the Q, the more narrow the bandwidth of the filter. This particular circuit is useful for Q values up to about 20. Higher Qs are obtainable by cascading filter sections. The design process is based on the resonant frequency, the Q value, and the desired gain at resonance (A).

EXAMPLE

A filter that resonates at 1 kHz with a bandwidth of 100 Hz is required. The necessary Q is 1000/100, or 10. If the desired voltage gain at 1 kHz is 10, and 0.01-μF capacitors are selected, find the values of R_1 and R_2.

SOLUTION

The value of R_1 can be calculated first:

$$R_1 = \frac{10}{6.28 \times 1000 \times 10 \times 0.01 \times 10^{-6}} = 15.9\ \text{k}\Omega$$

The feedback resistor is next:

$$R_F = 2 \times 15.9 \times 10^3 \times 10 = 318\ \text{k}\Omega$$

Finally, R_2:

$$R_2 = \frac{10}{6.28 \times 1000 \times .01 \times 10^{-6} \times (200 - 10} = 838\ \Omega$$

Figure 6-48 depicts the circuit for a *bandstop filter*, which is used to remove one interfering frequency. It can be used to eliminate 60-Hz hum from a signal. It provides an attenuation of approximately 40 dB at the resonant frequency and a gain of 1 for signals above and below the stopband. It is designed for a given value of resonant frequency and Q. The design process begins with the selection of a convenient value of capacitor. Suppose the selected capacitor value is 0.01 μF, the resonant frequency is 1 kHz, and the stopband is 100 Hz wide. The Q is 1000/100 = 10, and the feedback resistor is found by the following equation:

$$R_F = \frac{Q}{\pi f_R C_1}$$

$$R_F = \frac{10}{3.14 \times 1000 \times 0.01 \times 10^{-6}} = 318\ \text{k}\Omega$$

R_1 is found next:

$$R_1 = \frac{R_F}{4Q^2}$$

$$R_1 = \frac{318 \times 10^3}{4 \times 10^2} = 796\ \Omega$$

The value of R_2 will also be 796 Ω, and R_3 is

$$R_3 = 2\,Q^2 R_1$$

$$R_3 = 2 \times 10^2 \times 796 = 159\ \text{k}\Omega$$

The next special function we will look at is the four-quadrant multiplier. Its transfer characteristic is

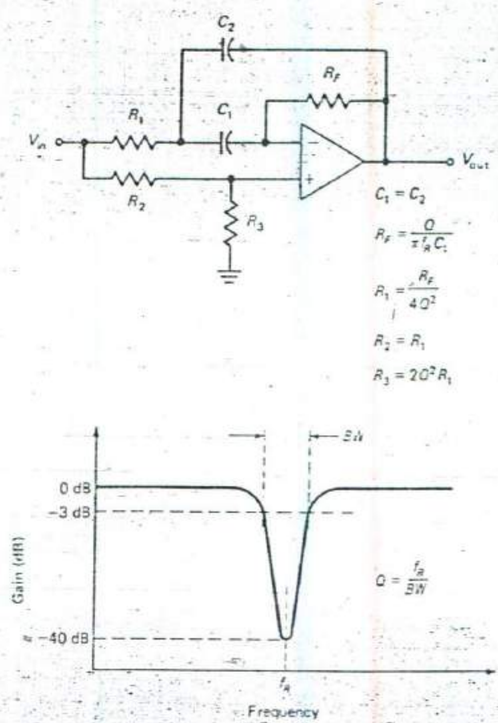


Fig. 6-48 Unity gain bandstop filter.

shown in Fig. 6-49(a), and its schematic symbol is shown in Fig. 6-49(b). Its output is the linear product of its two input voltages. Quadrant I operation results when both the Y and the X inputs are positive. Quadrant II results when X is negative and Y is positive. When both inputs are negative, the circuit operates in quadrant III. Quadrant IV means that X is positive and Y is negative. The output voltage is algebraically correct. For example, it is positive when both inputs are positive or both inputs are negative. It is negative if one of the inputs is negative. To avoid saturation with large inputs, the output is divided by 10. Therefore, with X = 10 V and Y = 10 V, the output will be equal to 10 × 10 = 100 divided by 10, or 10 V.

In addition to multiplying two signals, the four-quadrant multiplier may also be used to square a single voltage. This operation is illustrated in Fig. 6-50(a). Note that the X and Y inputs are tied together. The output signal is equal to one-tenth of the input signal squared. Note that the output of a squaring circuit can never be negative. If the input signal is a sine wave, the output signal is also a sine wave but at twice the frequency. The output sine wave for the frequency doubler also has a dc component. If this feature is undesirable, it can be removed with a coupling capacitor. Figure 6-50(b) shows a divide circuit. The output of the multiplier is summed with V_Z at the inverting terminal of an op amp. Voltage V_X is applied to the X input of the multiplier, and the

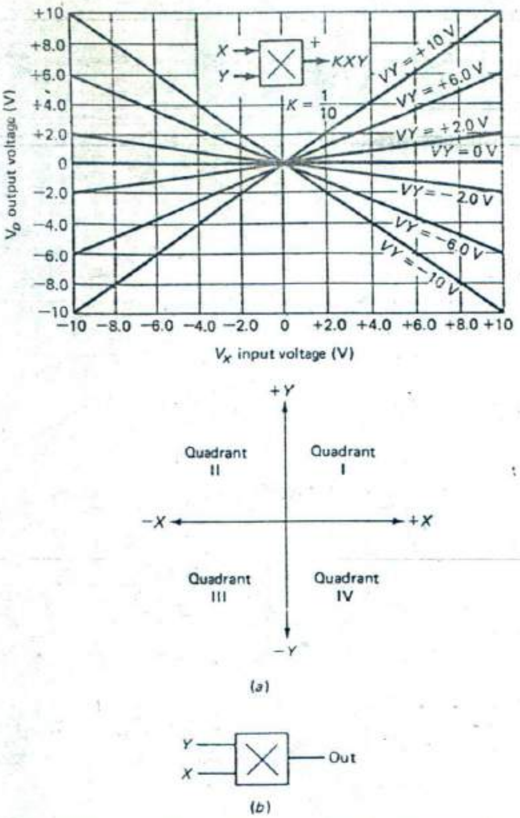


Fig. 6-49 Four quadrant multiplier. (a) Transfer characteristics. (b) Schematic symbol.

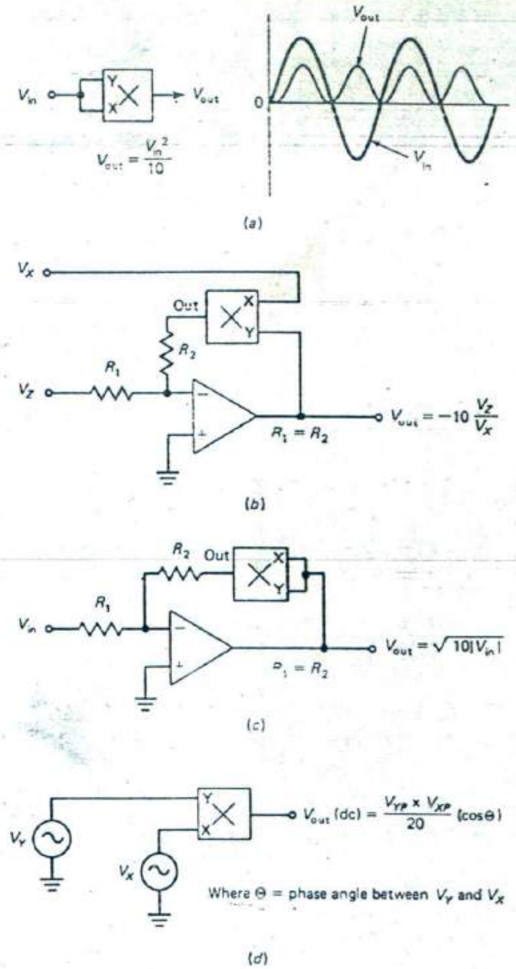


Fig. 6-50 Four quadrant multiplier applications. (a) Squaring circuit (frequency doubler). (b) Divide circuit. (c) Square root circuit. (d) Phase angle circuit.

op amp output is applied to the Y input. The overall output is inverted and is equal to ten times of V_Z divided by V_X. Figure 6-50(c) shows a square root circuit. The circuit works for negative input voltages only, and the output is equal to the square root of 10 times the absolute value of V_{in}. The last multiplier application is shown in Fig. 6-50(d). It produces a dc output voltage proportional to the peak value of two input waveforms and the cosine of the phase angle between them. For example, if one waveform is a sine and the other is a cosine, the phase angle between them will be 90°. The cosine of 90° is 0, and the dc output will be zero. If both input waveforms have a peak value of 10 V and happen to be in phase, the dc output voltage will be equal to 5 V because the cosine of 0° is 1.

Some industrial applications are hazardous and require an isolation amplifier to protect equipment from high-voltage surges and ground loops. Isolation amplifiers are also required to measure low-level signals in the presence of high common mode voltages. Malfunctions may cause power line voltages to be imposed on low-voltage signal lines. The isolation amplifier must be able to withstand such a mishap and protect expensive equipment on the other end. There are two commonly used isolator designs. In Fig. 6-51 the block diagram for an Analog Devices 284J, which is of the modulation type, is shown. High-frequency transformers couple energy into the input section. Here, a current-limited converter rectifies and filters the high frequency to provide direct current for the input circuits. Since transformer coupling is used, the input is guarded (isolated) from the primary supply, the output, and ground. The signal to be amplified is applied to a phase modulator, which is also fed by the high-frequency signal. The resulting phase-modulated signal is transformer-coupled to a phase demodulator in the output section. The demodulator recovers the original signal, and a low-

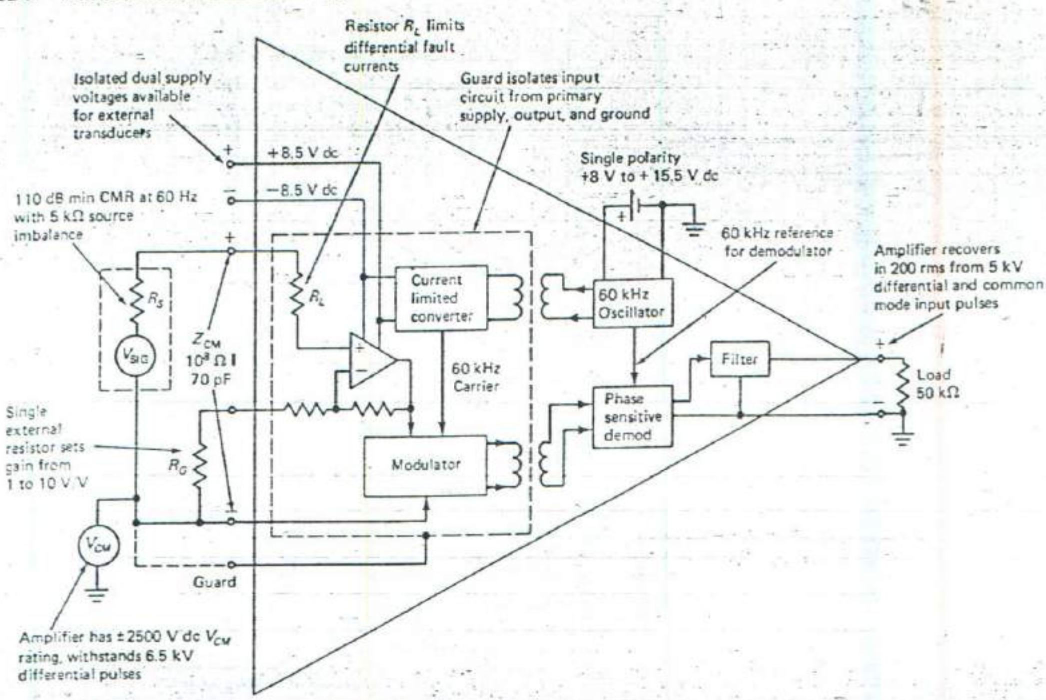


Fig. 6-51 Block diagram of 284J isolation amplifier (*Analog Devices*).

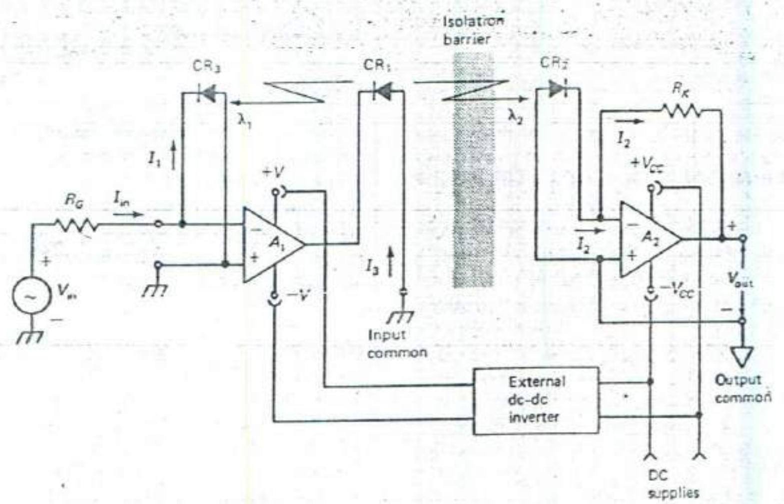


Fig. 6-52 Block diagram of an optically coupled isolation amplifier.

pass filter removes the high-frequency component. The other type of isolation amplifier is optically coupled and is shown in Fig. 6-52. It uses photodiodes and photodetectors to communicate across the isolation barrier. The optically coupled isolator is less expensive and smaller and has a wider bandwidth than the modulation type. However, it requires an external dc-to-dc converter for operation. The optically coupled types also suffer from gain inaccuracy and are more susceptible to temperature effects.

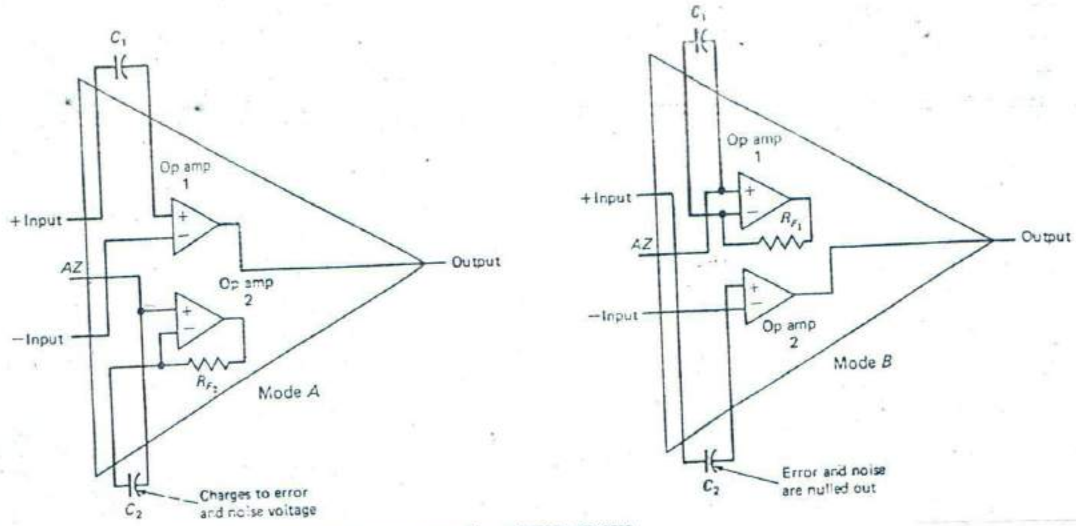


Fig. 6-53 The two half cycles of operation for a commutating autozero op amp.

Their signal-to-noise ratio is also poorer than that of the modulation types.

Temperature drift is the major enemy of high-gain dc-coupled amplifiers. An early solution was to use a *chopper amplifier.* These circuits chopped the dc signal to produce an ac signal. The ac signal was then amplified in a high-gain ac amplifier, where drift was not a problem. Finally, the ac signal was converted back to dc and filtered. Chopper stabilization has now progressed to a related but different approach called *commutating auto-zero* (CAZ) op amps. One example is the Intersil ICL7600 CAZ op amp. It features a low-input offset voltage of 2 μV. Its long-term drift is only 0.2 μV/year. It also features an input offset temperature drift of only 0.005 V/°C. It is based on two internal op amps that are connected so that when one amplifier is processing a signal the other is maintained in the auto-zero mode. Since the device auto-zeros its internal offset errors, it is extremely stable with temperature and time. Figure 6-53 shows its operating principles. The voltage at the *auto-zero* (AZ) input is the voltage to which each op amp will be zeroed. In mode A, op amp number 2 operates in unity gain and charges external capacitor C_2 to a voltage equal to the dc offset of op amp number 2 and the instantaneous low-frequency noise voltage. Later, the internal switches connect the device in the configuration of mode B. Operational amplifier number 2 now has capacitor C_2 connected in series with its noninverting input. Since it was previously charged to its offset and noise voltage, the offset and noise are nulled out.

The last special function to be covered here is the *sample and hold amplifier* shown in Fig. 6-54. It is used in those cases in which circuits or measuring devices cannot accept varying signals. It works by pulsing the sample input, which turns on the enhancement mode FET. With the FET on, the input voltage can charge or discharge the capacitor. The acquisition time can be made short by using a capacitor of reasonable value and a low impedance op amp and FET switch. After the acquisition time, the FET is turned off. The capacitor now holds the value of V_{in}. The hold time can be made long by using a high-input-impedance op amp at the output. The droop voltage is usually on the order of several millivolts per second and is caused by capacitor leakage, FET leakage, and the input current of the second amplifier. The sample and hold circuit is available on a single chip.

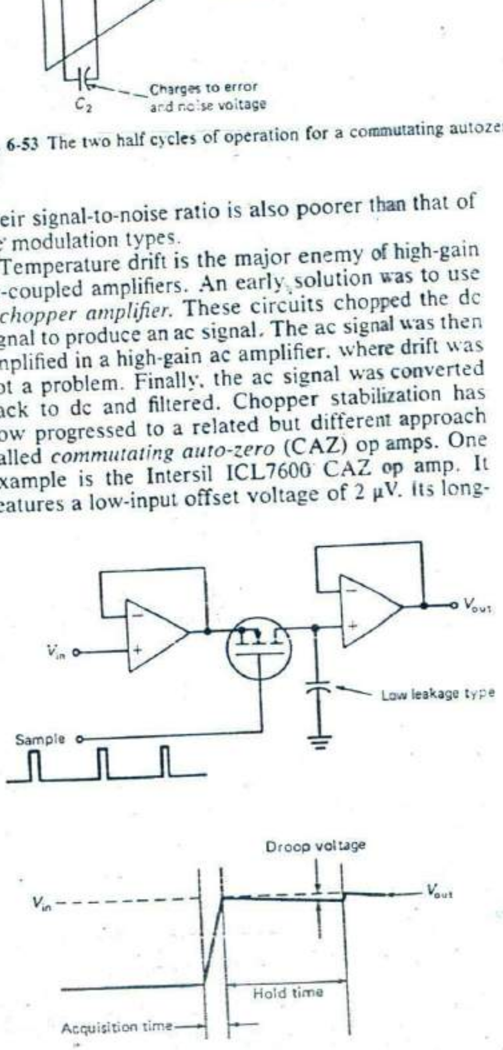


Fig. 6-54 Sample and hold amplifier.

REVIEW QUESTIONS

32. Refer to Fig. 6-39(*a*). What is the differential gain of the amplifier if $R_2 = 100\ k\Omega$ and $R_1 = 1\ k\Omega$?

33. What sets the input impedance of the IA in question 32?

34. Refer to Fig. 6-40. Assume no offset error and a balanced bridge. What is the output voltage?

35. Refer to Fig. 6-40. Will the output voltage change if R_3 and R_4 change by the same amount?

36. Refer to Fig. 6-40. Resistors R_3 and R_4 are identical strain gauges, one of which does not react to elongation. What is the function of the nonreacting gauge?

37. Refer to Fig. 6-41. What will the output voltage be if R_3 opens?

38. Why are the biasing resistors in Fig. 6-44 made twice the value of the feedback resistors?

39. Use Fig. 6-45 and design a filter that cuts off at 100 Hz by using a value of 0.1 μF for C_1.

40. What is the gain of the filter in question 39 at 20 Hz? At 100 Hz? At 1 kHz?

6-7 OSCILLATORS

An *oscillator* is a circuit that creates an ac signal or changes direct to alternating current. Oscillators are often based on amplifiers. An amplifier can be made to oscillate by adding positive (in-phase) feedback. If the gain of the amplifier is greater than the loss in the feedback network, the circuit will oscillate. Figure 6-55 shows a Wien bridge oscillator. The *Wien bridge* is made up of resistors R_1 and R_2 and capacitors C_1 and C_2. The bridge produces a zero phase shift at the resonant frequency. Frequencies higher than resonance produce a lagging response through the bridge, and frequencies lower than resonance produce a leading response. Therefore, only the resonant frequency arrives at the noninverting input of the op amp with the correct phase to sustain oscillations.

A Wien bridge oscillator is capable of very-low-distortion sinusoidal output because the bridge acts as a filter and allows in-phase feedback to occur for only a single frequency. This is an important feature in low-distortion sinusoidal oscillators because a pure sine wave has energy content at only one frequency. However, there will be considerable distortion in the circuit if the gain of the amplifier is more than that required to sustain oscillations. The circuit of Fig. 6-55 keeps the gain low by the voltage divider action of R_3, C_3, and the FET. The FET is biased by rectifying and filtering the output signal with the diodes and C_4. If the gain goes too high, the output signal becomes larger, more of it is rectified, and the

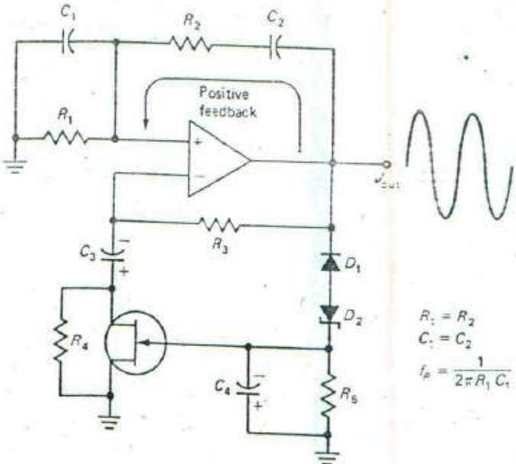


Fig. 6-55 Wien bridge oscillator.

negative gate voltage at the FET increases. This increases the drain-to-source resistance of the FET, and the negative feedback voltage at the inverting input is now larger. This decreases op amp gain, and the output is stabilized. Ideally, the automatic gain portion of the oscillator should keep the op amp gain just large enough to sustain oscillations.

The initial study of oscillators is sometimes confusing since it is difficult to understand how the circuit starts oscillating in the first place. When an oscillator is first energized, there is no strong signal at any particular frequency, but there is wide-band noise. This noise comes from the amplifying components and any parts connected to the input of the amplifier. The amplifier provides gain and wide-band noise is present at the output. The feedback network is frequency-sensitive and ensures that only one frequency arrives at the amplifier input with the correct phase (positive) relationship. This signal receives the most gain and soon appears at the output. The oscillator of Fig. 6-55 starts very quickly because the op amp gain is maximum when the circuit is first turned on.

Figure 6-56 shows another type of oscillator, known as a *multivibrator*. Resistors R_2 and R_3 provide positive feedback and upper and lower threshold points, as was discussed earlier for the Schmitt trigger. When the output is at positive saturation, current flows through R_1 to charge the capacitor. As long as the capacitor voltage is less than the UTP, the output remains at positive saturation. Eventually, the capacitor does charge to the UTP, the inverting input goes positive with respect to the noninverting input, and the output is driven to negative saturation. The capacitor now must charge in the opposite direction until the LTP is crossed. The multivibrator is not a single-frequency oscillator in the same sense as the Wien bridge circuit. It is in a category of *relaxation*

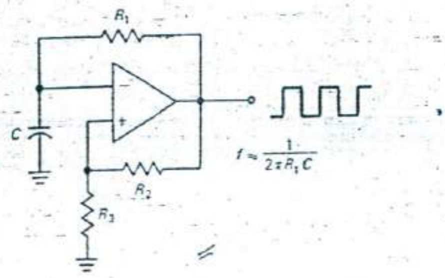


Fig. 6-56 Astable multivibrator.

oscillators, which are governed by *RC* time constants. The output frequency is a square wave at some fundamental frequency that is determined by the time constant of R_1 and C. The output also contains considerable energy at the odd harmonics of the fundamental. For example, if the fundamental frequency is 100 Hz, there will be energy at 300 Hz, 500 Hz, 700 Hz, and so on.

The 555 IC timer can also be used as an astable multivibrator. This circuit is shown in Fig. 6-57 (the 555 timer is discussed in Chapter 4). The resistors can be adjusted for an entire range of duty cycles in the output waveform. A diode will be required for duty cycles of 50 percent and less. At the moment of power on, the capacitor is discharged, holding the trigger low. This triggers the timer and establishes the capacitor charge path through R_A and R_B, assuming no diode. When the capacitor voltage reaches the threshold of ⅔ V_{CC}, the output goes low, and the discharge transistor turns on. The capacitor discharges through R_B. When the capacitor voltage drops to ⅓ V_{CC}, the trigger comparator trips and retriggers the timer. Adding the diode allows the charge path to be R_A and the diode to provide a duty cycle range of 5 to 95 percent by adjusting the ratio of the timing resistors. Resistor R_B should be no smaller than 3000 Ω to ensure reliable starting of the circuit.

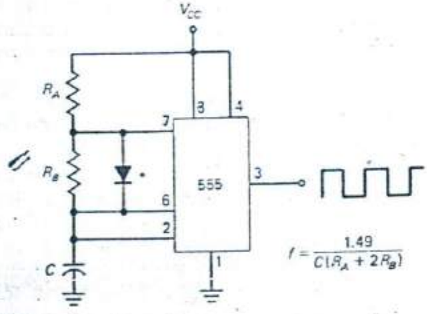


Fig. 6-57 Integrated circuit timer astable multivibrator.

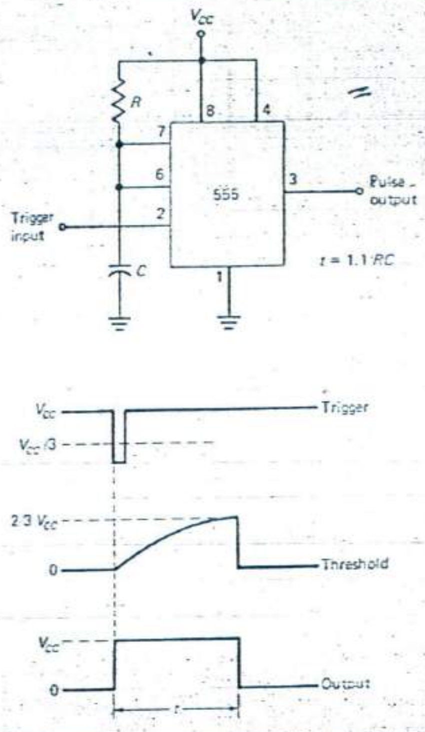


Fig. 6-58 Integrated circuit timer monostable (one-shot) multivibrator.

The IC timer may also be used in the *monostable,* or one-shot, mode to provide a timed output pulse in response to an input trigger pulse, as shown in Fig. 6-58. Prior to the trigger input pulse, the discharge transistor in the IC is on, thus short-circuiting the timing capacitor (*C*) to ground. When a trigger pulse drives pin 2 below $V_{CC}/3$, the device triggers, turning the discharge transistor off. The output goes high and remains high until the timing capacitor charges to ⅔ V_{CC} through the timing resistor (*R*). When the ⅔ V_{CC} threshold is reached, the comparator trips, in turning driving the output low and turning the discharge transistor on. The transistor quickly discharges the timing capacitor. The cycle is now complete, and the timer is ready for another input pulse. The width of the output pulse is equal to $1.1RC$. If the timing resistor is 4.7 kΩ and the timing capacitor is 0.02 μF, the pulse width will be $1.1 \times 4.7 \times 10^3 \times 0.02 \times 10^{-6} = 103.4$ μs.

Figure 6-59 shows an oscillator circuit that produces both a square and a triangular waveform output. Assume that the threshold detector output is initially at negative saturation. This will cause the integrator to begin ramping in a positive direction. It will continue to ramp until its output crosses the lower threshold point of the detector. Its output will

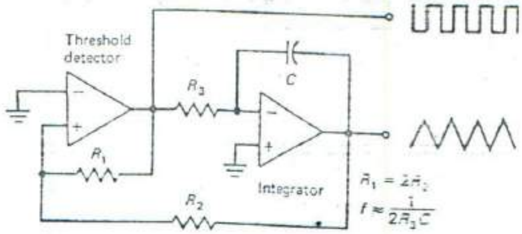


Fig. 6-59 Square and triangle wave generator.

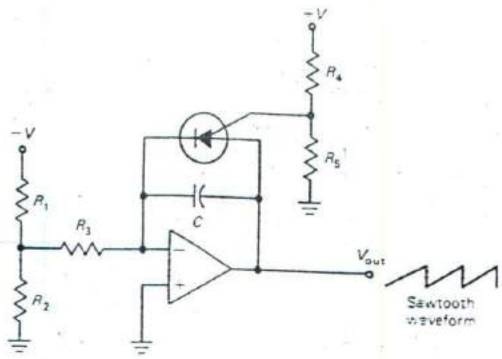


Fig. 6-60 Sawtooth oscillator.

now snap to positive saturation. The integrator will now begin ramping negative until the lower threshold point is crossed. The detector reassumes positive saturation, and another cycle begins.

A *sawtooth oscillator* is shown in Fig. 6-60. Voltage divider R_1 and R_2 apply some fraction of the negative supply to the integrator. It begins ramping positive at a rate that is determined by the magnitude of the input voltage, R_3, and the capacitor. Eventually, the anode-to-cathode voltage of the programmable unijunction transistor (PUT) will reach the breakover voltage. The PUT will turn on and quickly discharge the capacitor to about 1 V. Then the integrator will begin ramping again. The firing point of the PUT is adjusted with the positive divider made up of resistors R_4 and R_5. The output frequency is partly controlled by the gate voltage of the PUT because if it is adjusted to fire earlier, the positive ramp will be terminated sooner.

REVIEW QUESTIONS

41. What two conditions must be met for an amplifier to oscillate?

42. What is the function of the FET in Fig. 6-55?

43. The oscillator of Fig. 6-56 produces energy at a fundamental frequency and at ________ harmonic frequencies.

44. Refer to Fig. 6-57. Which resistor(s) determine(s) the length of time the output will be in the high state when the diode is used? When the diode is not used?

45. Refer to Fig. 6-60. What determines the output frequency in addition to R_3, C, R_4, and R_5?

6-8 TROUBLESHOOTING AND MAINTENANCE

Amplifier troubleshooting should begin with a thorough visual inspection. Do not forget to check cables, sockets, and connectors. Edge connectors on plug-in circuit boards are notorious for causing problems. Power down and remove the boards. Brighten the contact surfaces on the board by using an ordinary pencil eraser. Do not use highly abrasive materials. Check devices in sockets because vibration can cause them to work loose. Reseat any that look suspicious. If the pins on ICs look black, that is a clue that some corrosion has taken place. Use an IC puller or a small screwdriver and work the ICs partly out of their sockets and then firmly reseat them. This procedure produces a wiping action that will usually restore the conductivity of the contacts to an acceptable level. Of course, severe corrosion or sockets that have lost their tension will probably dictate replacement.

After you have verified that everything looks good, it is time to verify supply voltages. A circuit usually cannot work as intended if its supply voltage or voltages are out of tolerance. Too many technicians waste time troubleshooting circuits that do not have anything wrong with them. A complete power supply check may include an oscilloscope test. A dc voltmeter reading can be in the normal range when there is excess ac ripple on the supply line. This check is especially important if one of the symptoms is erratic output or ac hum in the output of the amplifier.

After the supply has been verified, it is time to identify the symptoms. Is the amplifier "dead" (no output at all)? Is the output weak? Is the amplifier unstable (acting as an oscillator)? Is the output of the amplifier distorted? Is there noise or hum on the output signal? Does the dc output drift? Is the output latched at the positive or negative rail?

Amplifier troubleshooting is usually done while working "live." If the amplifier is part of a control system with motors, hydraulic, or pneumatic actuators, you must be very careful that a circuit disturbance will not produce a dangerous mechanical output. It may be necessary to deactivate part of the system while working. However, if the system produces feedback signals, then deactivation may introduce its own set of symptoms. Refer to the manufacturer's recommendations when servicing systems of this type.

A dead amplifier is often the easiest to troubleshoot. If it is a multistage arrangement, you should view it as a signal chain. It must be verified one link at a time. Signal tracing with an oscilloscope is a useful technique. Start with the first stage and verify the presence of an input signal. Move to the output of the first stage and check again. If you find a stage where there is an input signal and no output signal, the defective stage has been isolated. It is then usually possible to take a few dc readings to isolate the defective component. For example, in a common emitter amplifier such as the one shown in Fig. 6-3, you may find no base voltage on the transistor. This situation could be caused by an open bias resistor or a short-circuited coupling capacitor. Another possibility is normal base voltage and a collector voltage equal to the supply. If the collector voltage is supposed to be about half the supply, then the transistor or the emitter resistor is open. The kinds of problems that upset dc terminal voltages in amplifier circuits (assuming normal supplies) include short-circuited or leaky coupling capacitors, short-circuited or leaky bypass capacitors, open or out of tolerance resistors, open solid-state junctions, and short-circuited solid-state junctions. Your knowledge of circuit principles and device behavior will lead you in the right direction.

If the circuit is a dc amplifier, the oscilloscope tracing technique may be less useful. It may be more desirable to use a digital voltmeter to verify the dc signal at various points in the signal chain. It is also sometimes possible to utilize circuit adjustment as a way to verify proper circuit response. For example, the circuit may have an offset null adjust potentiometer. The potentiometer setting can be changed slightly while monitoring the dc output voltage to determine whether the response is as expected. Do not use this technique until you carefully note the original setting of the potentiometer. If it is a multiturn potentiometer, carefully count the turns so you can accurately reset the circuit when you are done testing. Also make sure that the manufacturer's literature contains the adjustment procedure so you can restore the circuit to original specifications. If you find that it is possible to restore a normal dc output by making a large change in some adjustment point, you should not consider this a "fix" until you have investigated further. This condition often indicates that some component has drifted out of tolerance. It is probably going to continue drifting with time, and your fix will be temporary at best.

A weak output indicates that the gain of one of the stages is below normal or that the input signal is out of tolerance. Once again, tracing with an oscilloscope or digital voltmeter will usually isolate the stage. Low gain in a stage may be caused by a defective bypass or coupling capacitor, a bias resistor, a decoupling resistor, or a solid-state device. Even though the power supply voltages have been checked, it is possible that one stage will operate at a voltage lower than normal. Decoupling resistors are sometimes used to isolate a stage from the power bus. If a decoupling resistor increases in value, the supply voltage for that one stage will drop, and the gain will often be less than normal. When measuring gain, remember that follower-type amplifiers are expected to produce unity gain only.

An unstable output may be caused by undesired oscillations. Remember that any amplifier can become an oscillator if there is in-phase feedback and more gain than loss. Decoupling resistors are used in conjunction with bypass capacitors to prevent amplifier coupling through the power bus. An open bypass capacitor may cause ac feedback and oscillations. Another source of instability is phase error that causes feedback to become positive at some frequency. Stray capacitances in a circuit, along with resistances, form a group of *RC* lag networks that do not show on schematic diagrams. Negative feedback can become positive feedback if the total lag reaches 180°. As the frequency goes higher and higher, the individual lags can sum to produce enough phase shift to make the feedback positive. If there is enough gain at this frequency, oscillations will occur. This is why op amps are often internally compensated for a gain roll-off of 20 dB per decade. This compensation ensures that the gain will be too low for oscillations to occur at the frequency where phase shift makes the feedback become positive. If an op amp is unstable and is externally compensated, the compensation components should be checked. If the gain is too high, it may also cause oscillation for the reasons already discussed.

Distortion in an amplifier is often caused by a bias error. For example, if an amplifier is not operating near the center of its load line, then clipping of either the negative-going or the positive-going part of the signal may result. It is usually easy to trace distortion to one stage by using an oscilloscope. In some cases, it is also helpful to use a function generator to feed a known test signal through the system. Most function generators are capable of producing sine, square, and triangle waveforms. The triangle waveform is usually the best choice when looking for distortion in a linear amplifier. The sharp peaks make it easy to spot any tendency toward clipping or compression, and the straight sides make any nonlinearity apparent. Once the defective stage is isolated, dc voltage checks will often reveal which component has failed. If the dc readings are correct, a defective solid-state device should be suspected.

Noise and hum are problems that can indicate power supply difficulties, open bypass capacitors, broken ground wires and shields, or defective devices. If the noise or hum is a common mode signal that the amplifier should reject, check the op amps and IAs. Circuit imbalance is also a possible culprit when the CMRR is not what it should be. For example, the IA shown in Fig. 6-39(*a*) will not exhibit good common mode rejection if one of the four resistors (R_4 through R_7) in the output stage changes value. It may be possible to adjust the CMRR in some amplifiers. Once again, if a significant change in a potentiometer must be made to restore proper operation it may be an indication that some component is drifting. Another noise source is a defective component in the input stage of a high-gain amplifier. A resistor, transistor, diode, IC, or some other component may become a noise generator. Remember that some components such as zener diodes normally generate significant noise and must be bypassed to confine the noise. Check the appropriate bypass capacitors. A transducer (or some other signal source) connected to the amplifier input may also be defective and noisy. Check for dirty and corroded connections. Finally, the amplifier may be responding to a strong electromagnetic field from some RFI source.

Drift is often a problem in high-gain dc amplifiers. Some drift is normal. If the drift is causing problems, it must be investigated. It is usually worst when it occurs in an early stage because it is amplified by all succeeding stages. Often a pattern to the drift can provide a valuable clue. For example, the drift may always be in the same direction as the amplifier warms up. This usually indicates that some component or device is changing with temperature. You can use a source of heat or cold to help isolate a sensitive part. Some cold spray materials build up damaging static charges and attack plastics, so be careful to use an approved type. Heat sources include soldering pencils, lamps, and heat guns. Be careful not to overheat circuit boards, components, and plastic parts.

Latch-up is a problem encountered with some operational amplifiers. The output gets "stuck" at the positive rail or at the negative rail. If the inputs seem normal, check the feedback circuit components. If no problems can be found there, investigate the dual supply on power up by using a dual-trace oscilloscope. If the plus and minus voltages are not applied at the same time, latch-up may result.

REVIEW QUESTIONS

46. Refer to Fig. 6-3. Suppose R_{B2} opens up. The amplifier operating point will move toward ______.

47. Will the dc collector voltage measure normal, high, or low for the fault discussed in question 46?

48. Refer to Fig. 6-4(*a*). Will the dc output voltage measure normal, high, or low if the emitter of Q_1 opens up?

49. Refer to Fig. 6-5(*b*). Should the normal voltage gain of this amplifier be high, moderate, or near unity?

50. Refer to Fig. 6-10. What dc output voltage can you expect at *A* and *B* if the $-V_{EE}$ supply fails and goes to 0 V?

51. Refer to Fig. 6-40. What dc condition will exist at V_{out} if R_2 opens?

CHAPTER REVIEW QUESTIONS

6-1. Refer to Fig. 6-2 and assume a quiescent base current of 80 μA. Where will the amplifier operate?

6-2. What will the quiescent collector-to-emitter voltage be for question 6-1?

6-3. Refer to Fig. 6-3. Which two components establish the base voltage?

6-4. What type of coupling is required for dc amplification?

6-5. What is the function of the emitter bypass capacitor in Fig. 6-3?

6-6. What is the overall current gain in a Darlington pair in which the first transistor has an h_{FE} of 50 and the second an h_{FE} of 200?

6-7. What is the phase relationship between input and output for common base and emitter follower amplifiers?

6-8. Refer to Fig. 6-11. Assuming balance and a common mode input, what will the differential output be? What will the single-ended output be?

6-9. A particular op amp requires a 1.5-mV input differential to produce 0 output. What is this error called?

6-10. What terminals are provided on some op amp packages to eliminate the error of question 6-9?

6-11. Refer to Fig. 6-23. Input 1 is driven 0.2 V negative. What effects will be noticed at the other two inputs?

6-12. Calculate the output voltage for Fig. 6-25 if the feedback resistor is 27 Ω and the current source is conducting 0.5 A.

6-13. Refer to Fig. 6-26. If $V_{in} = 0.01$ V, $R_L = 10{,}000\ \Omega$, and $R_s = 10\ \Omega$, how much current will flow in the load resistor? How much current will flow if the load resistor is changed to 4.7 kΩ?

6-14. Refer to Fig. 6-33. If the input is +0.2 V, what are the polarity and rate of change at the output?

6-15. What is the instantaneous output voltage in question 6-14 after the input has been applied for 100 ms?

6-16. Refer to Fig. 6-35. What will happen to the output frequency if the reference voltage is made more negative?

6-17. What is the function of R_3 in Fig. 6-36?

6-18. Refer to Fig. 6-36. What will happen to the ON time of the light source if the reference voltage is made more negative?

6-19. Suppose the output waveform in Fig. 6-37 shows every other negative peak at a greater amplitude. Could this be caused by R_3 being out of tolerance?

6-20. Refer to Fig. 6-38. How would the circuit act if the diode were short-circuited?

6-21. What is the bandwidth of a bandpass filter having a Q of 20 and a resonant frequency of 100 Hz?

6-22. Two unity gain highpass filters are cascaded. Assuming that both have a cutoff frequency of 100 Hz, what is the overall response at 100 Hz? Is the cascaded cutoff frequency higher, lower, or the same?

6-23. What is the output voltage for the multiplier in Fig. 6-49 if $V_X = -5$ V and $V_Y = +6$ V? In which quadrant is the multiplier operating?

6-24. Calculate the dc output voltage for Fig. 6-50(*a*) if the input signal is a 6.32-V peak sinusoid. (Hint: Refer also to Fig. 6-50(*d*).)

6-25. Calculate the output voltage for Fig. 6-50(*c*) if the input voltage is -5 V.

6-26. What are the two types of isolation amplifiers?

6-27. Why are chopper-type amplifiers used?

6-28. What is the function of the external capacitors in the CAZ amplifier shown in Fig. 6-53?

6-29. Refer to Fig. 6-41. What dc condition will exist at V_{out} if R_3 opens?

6-30. Refer to Fig. 6-55. What ac output condition will exist if C_3 opens?

ANSWERS TO REVIEW QUESTIONS

1. 37 dB **2.** 32 dB **3.** common emitter **4.** 180° **5.** 4 V **6.** cutoff **7.** 16 V **8.** yes; 0° **9.** yes; no **10.** no **11.** 80 dB **12.** noninverting; + **13.** 40 dB **14.** infinity; to avoid loading effects **15.** zero; so it can drive any load **16.** 1 MHz; yes **17.** 1.59 MHz **18.** slew rate **19.** 10; 10 kΩ **20.** 100 kHz **21.** 101; 6 MΩ **22.** +3 V **23.** − 10.34 V **24.** nonweighted; change any of the input resistors **25.** it will decrease **26.** an output pull-up resistor **27.** no (the output is at positive saturation) **28.** +1.93 V; − 1.93 V **29.** square wave **30.** +0.5 V **31.** triangle **32.** 201 **33.** the intrinsic impedance of the op amps **34.** zero **35.** no **36.** thermal compensation **37.** positive saturation **38.** to set V_{out} at half the supply **39.** $R_1 = 11.3$ kΩ; $R_F = 22.5$ kΩ; $C_2 = 0.2$ μF **40.** 0 dB; −3 dB; −40 dB **41.** in-phase feedback and more gain than loss **42.** to control gain and minimize distortion **43.** odd **44.** R_A; R_A and R_B **45.** R_1 and R_2 **46.** saturation **47.** low **48.** high **49.** near unity **50.** equal to V_{CC} **51.** +14 V (positive saturation)

7

MAGNETIC DEVICES

This chapter deals with some important magnetic devices used in modern industry. These devices are for the most part passive and contain no moving parts. They are relatively maintenance-free and reliable. They will last indefinitely if not abused by overloading or overheating, provided they are adequately protected for the environment in which they are designed to operate. Industry relies heavily on these devices for control, transfer, and conditioning of ac power. In this chapter emphasis is on 60-Hz power, though other frequencies are mentioned where appropriate.

7-1 SIGNAL AND SPECIAL PURPOSE TRANSFORMERS

An impedance relationship exists across a transformer that is not equal to the turns ratio. The following analysis will show what is true for all power and signal transformers. Since the voltage ratio of a transformer is equal to the turns ratio,

$$\frac{V_p}{V_s} = \frac{N_p}{N_s}$$

Also, the current ratio is equal to the reciprocal of the turns ratio:

$$\frac{I_p}{I_s} = \frac{N_s}{N_p}$$

Then, by dividing one equation by the other we obtain,

$$\frac{\dfrac{V_p}{I_p}}{\dfrac{V_s}{I_s}} = \frac{N_p^2}{N_s^2}$$

Since V_p/I_p is equal to Z_p, the primary impedance, and V_s/I_s is equal to Z_s, the impedance of the secondary, by substitution we obtain

$$\frac{Z_p}{Z_s} = \left(\frac{N_p}{N_s}\right)^2$$

Therefore, the ratio of the impedances across a transformer varies as the *square* of the turns ratio.

Rearranging the terms,

$$Z_p = Z_s \times \left(\frac{N_p}{N_s}\right)^2$$

and

$$Z_s = Z_p \times \left(\frac{N_s}{N_p}\right)^2$$

If the transformer shown in Fig. 7-1 has a step-up ratio of 10 to 250 and draws 100 mA at 6 V in the primary circuit, the primary impedance is

$$Z_p = \frac{V_p}{I_p} = \frac{6}{0.1} = 60\ \Omega$$

The impedance of the secondary is

$$Z_s = 60 \times \left(\frac{250}{10}\right)^2 = 60 \times 625 = 37{,}500\ \Omega$$

For a step-down transformer with a turns ratio of 15 to 1 supplying an impedance of 2 Ω, the impedance of the primary is

$$Z_p = 2 \times \left(\frac{15}{1}\right)^2 = 2 \times 225 = 450\ \Omega$$

The value of a signal transformer is in its ability to transform a low impedance to a high impedance or vice versa. No external power is required to accomplish this feat. The isolation characteristics are also very beneficial for the elimination of ground loops that can override a small signal. Signal transformers have well-defined bandwidths. Most iron core types are limited to audio frequencies. With ultrasonic circuits of 20 kHz or more, ferrite or air cores are used to eliminate the hysteresis and eddy current losses that increase with frequency.

Maximum transfer of power from source to load is important in any circuit and especially so when moderate to high power levels are involved. The maximum transfer of power occurs when the impedance of the source is equal to the impedance of the load. The simple dc circuit of Fig. 7-2 illustrates this principle. A 10-V battery with an internal resistance R_B of 1 Ω feeds a variable load resistance R_L. When R_L is 4 Ω, the total circuit series resistance is 5 Ω ($R_B + R_L$), the current is 2 A, and there is an 8-V drop across the load. The power absorbed by the load is 8 V times 2 A or 16 W. The power dissipated

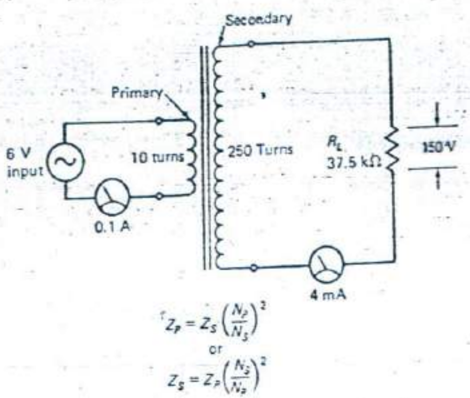


Fig. 7-1 Equivalent circuit of signal transformer impedance.

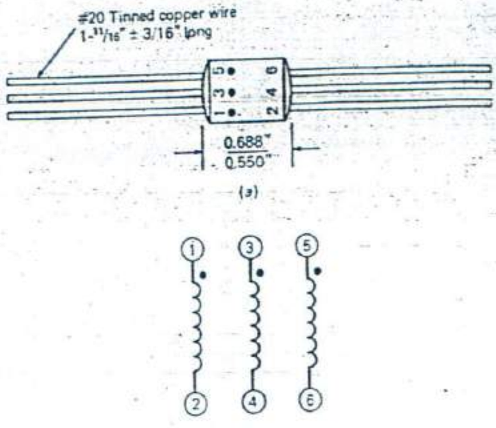
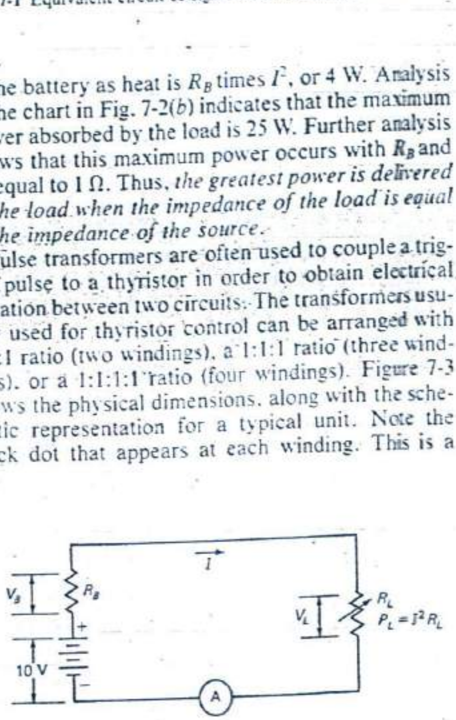


Fig. 7-3 Thyristor pulse transformer. *(a)* Physical dimensions. *(b)* Schematic symbol.

in the battery as heat is R_B times I^2, or 4 W. Analysis of the chart in Fig. 7-2(*b*) indicates that the maximum power absorbed by the load is 25 W. Further analysis shows that this maximum power occurs with R_B and R_L equal to 1 Ω. Thus, *the greatest power is delivered to the load when the impedance of the load is equal to the impedance of the source.*

Pulse transformers are often used to couple a trigger pulse to a thyristor in order to obtain electrical isolation between two circuits. The transformers usually used for thyristor control can be arranged with a 1:1 ratio (two windings), a 1:1:1 ratio (three windings), or a 1:1:1:1 ratio (four windings). Figure 7-3 shows the physical dimensions, along with the schematic representation for a typical unit. Note the black dot that appears at each winding. This is a phasing (polarity) reference dot. The dots show that the indicated terminals have the same polarity with respect to each other.

For pulse triggering a thyristor, a minimum amount of energy is required, and a minimum trigger pulse width is also necessary. A typical gate current to gate pulse width plot is shown in Fig. 7-4, along with the



V(V)	R_B	R_L	R_T	I(A)	V_B	V_L	P_L (W)
10	1	4	5	2	2	8	16
10	1	3	4	2.5	2.5	7.5	18.75
10	1	2	3	3.33	3.33	6.67	22.2
10	1	1	2	5	5	5	25
10	1	.5	1.5	6.67	6.67	3.33	22.22
10	1	.33	1.33	7.5	7.5	2.5	18.75
10	1	.25	1.25	8	8	2	16

(*b*)

Fig. 7-2 Power transfer and impedance. *(a)* Circuit. *(b)* Table of impedances.

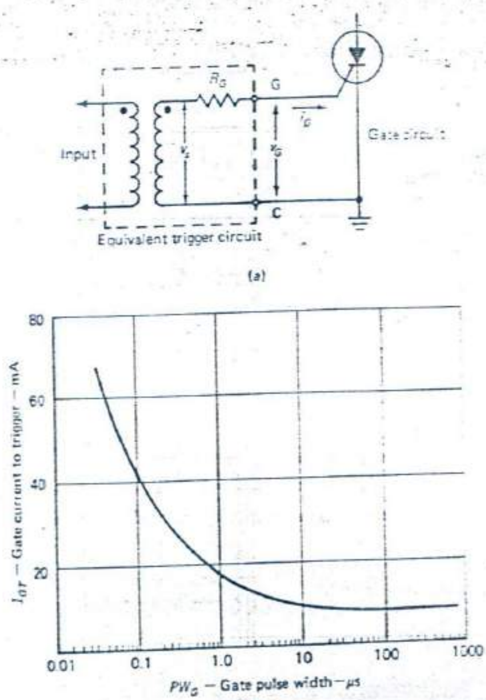


Fig. 7-4 Typical thyristor trigger circuit. *(a)* Gate circuit with pulse transformer. *(b)* Gate current to gate pulse width curve.

equivalent trigger circuit. The transformer shown in Fig. 7-3(*a*) has been specifically designed for triggering thyristors. The prime requirement of a trigger pulse transformer is efficiency. The simplest test is to use the desired trigger pulse generator to directly drive a 20-Ω resistor and then to drive the same resistor through the pulse transformer. If the pulse waveforms across the resistor are the same under both conditions, the transformer is considered to be perfect.

A similar type of pulse transformer is used to drive MOSFETs in switching and flyback converters. The windings are interleaved for lowest practical leakage inductance. When driven by these transformers, drain to source rise times of 25 nanoseconds (ns) are achieved, and useful frequencies up to 200 kHz are obtainable. Ferrite cores are commonly used at these frequencies. These devices are for pulse service and will not work as input or output transformers, which are primarily low-frequency devices.

Bucking and *boosting transformers* (sometimes called *corrector transformers*) are used to provide an economical and convenient means of boosting or bucking voltage on single- and three-phase circuits. These transformers have series-multiple 12/24- or 16/32-V secondary windings suitable for a wide variety of applications. They are connected as autotransformers in single-phase or three-phase circuits, for boosting or bucking voltages. Each *boost or buck* (B-B) transformer will have eight leads brought out

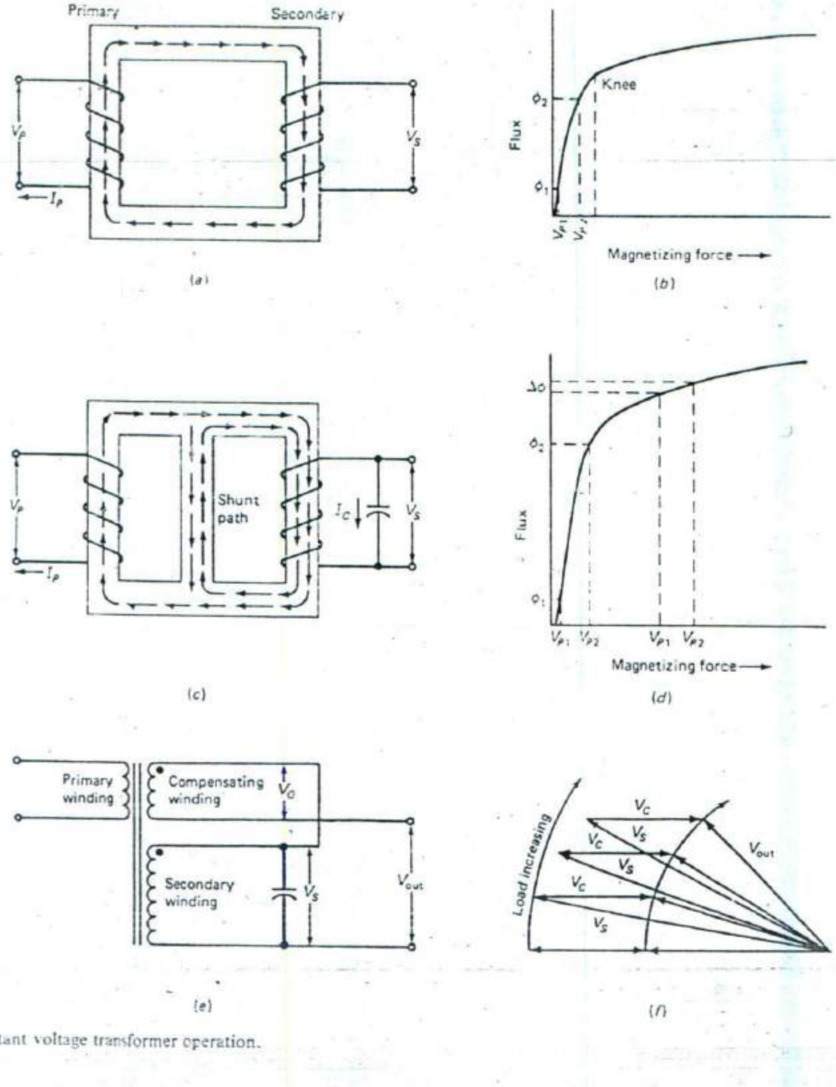


Fig. 7-· Constant voltage transformer operation.

for wiring. In accordance with American National Standards Institute (ANSI) standards, the four leads on the high-voltage winding will be marked H1, H2, H3, and H4. The low-voltage leads will be marked X1, X2, X3, and X4. A few points of interest should be noted. The *open-delta* connection requires only two B-B transformers to transform three-phase voltages. The open-delta does not provide a neutral and therefore is used only on three-phase applications in which the neutral (if one exists) is not connected to the load. The geometric neutrals of the input and output voltages on the open-delta connection are not the same. Thus, circulating currents will result if the neutral of the line is connected directly to the load: therefore, the neutral must not be connected for this type of operation.

As an example of the use of a B-B transformer, suppose we have a 60-Hz, single-phase, 12-kilovolt-ampere (kVA) load that is rated at 230 V. However, the only line available is 208 V, 60 Hz. The solution would be to select a B-B transformer rated 12/24 V with a kVA rating of 10 percent of 12 kVA or 1.2 kVA. The primary currents cancel, so only the low-voltage winding conducts full load current.

For many types of industrial applications a simple, reliable source of constant potential is a requisite for successful operation. The static-magnetic regulator, with its simplicity, ruggedness, and reliability, is commonly employed. These units have many generic names: *ferromagnetic*, *flux-former*, etc. Ferroresonant constant-voltage transformers are discussed in Chapter 5. The basic operating principle is the same: only the name is different, as they are all constant-voltage transformers.

The basic principles of constant-voltage transformer operation are illustrated in Fig. 7-5. Figure 7-5(a) reviews how the conventional transformer primary voltage (V_p) sets up magnetizing current (I_p) and a resultant flux (arrows in core) that links the secondary winding to induce a voltage (V_s). When operating below the knee of saturation (Fig. 7-5[b]), flux ϕ_1-ϕ_2, V_s is proportional to the primary voltage V_p. Above the knee, changes in V_p have little effect on V_s. With a magnetic shunt added (Fig. 7-5[c]) and a capacitor across the secondary, the capacitor current I_c generates additional flux in the secondary leg through the shunt path. The total flux in the secondary leg is now above the knee (Fig. 7-5[d]), so changes in V_p have considerably reduced effect on the secondary flux and on V_s as compared to Fig. 7-5(b). To minimize changes of V_s with V_p further, a compensating winding can be added over the primary winding and is connected in series opposition (Fig. 7-5[e]) to the secondary to subtract from V_s. Figure 7-5(f) shows a locus of points, illustrating that with increasing load current (V_p constant), the compensating winding subtracts less. This means that as the load current is increased, the compensating winding bucks out less of the secondary voltage. This effect offsets the drop in secondary voltage normally associated with increasing load.

REVIEW QUESTIONS

1. What is the primary impedance of a transformer with 300 primary turns and 20 secondary turns if the secondary load is 40 Ω?

2. The ______ characteristics of the signal transformer are useful for the elimination of ground loops.

3. Signal transformers are usually employed for frequencies below ______.

4. The maximum transfer of power is obtained when the source impedance is ______ to the load impedance.

7-2 MAGNETIC AMPLIFIERS

The *magnetic amplifier* is a device used to reproduce an applied input signal at an increased amplitude. All amplifiers do this, but the method of amplification is unique for magnetic amplifiers. The increases in amplitude that occur in magnetic amplifier circuits are produced by the variations in magnetism and inductance within the unit. The magnetic amplifier is known for its dependability, ruggedness, high efficiency, and ability to withstand high temperatures and other severe environments.

They are used as industrial regulators, relays, starters, amplifiers, servosystems, and converters, and they are employed in some computer applications. Since a good knowledge of magnetism is essential to understanding the magnetic amplifier, a brief review of basic magnetic theory is in order.

The region around a bar magnet where its influence is felt is called a *magnetic field*. This field can be viewed as a pattern of lines arranged in an orderly fashion emanating from the poles of the magnet. A similar pattern of magnetic lines will exist around a current-carrying coil. The strength of the field is called *magnetomotive force* (mmf) measured as NI, proportional to the ampere turns:

$$\text{MMF} = NI$$

where

MMF = ampere turns
N = number of turns in the coil
I = current flow in the coil, A

Another parameter to be considered is the field intensity (H), sometimes called *magnetizing force*. It is the magnetomotive force per unit length:

$$H = \frac{NI}{l}$$

where

H = ampere turns per meter
NI = ampere turns
l = length, m

Magnetic flux (Φ) is similar to the current in an electric circuit and comprises the total number of lines of force existing in the magnetic field. The unit of flux is the weber (Wb). A weber is composed of 10^8 lines. *Flux density* (*B*) is the means of measuring the amount of flux lines per unit area. In the SI (the international system of measurements), the tesla (T) is the unit of flux density. The flux density is 1 T when there is 1 Wb/m², expressed as

$$B = \frac{\Phi}{A}$$

where

B = flux density, T
Φ = flux, Wb
A = area, m²

Permeability (μ) is a comparative factor depicting the ease by which a material can conduct magnetic flux as compared to the ease of which a vacuum conducts flux. The permeability formula is

$$\mu = \frac{\Delta B}{\Delta H}$$

where

μ = permeability
ΔB = change in flux density
ΔH = change in field intensity

It should be noted that in a vacuum (or air) flux density equals field intensity at all times. Therefore, the μ (permeability) of a vacuum equals 1. Materials with a permeability of less than 1 are called *diamagnetic*, and those with a μ slightly greater than 1 are called *paramagnetic*. Materials such as iron, cobalt, and nickel with a μ much greater than 1 are referred to as *ferromagnetic*.

Reluctance is the opposition to magnetic flux offered by a magnetic material. The equation for reluctance is

$$\mathcal{R} = \frac{l}{\mu A}$$

where

$\mathcal{R}$ = reluctance
l = length, m
μ = permeability
A = area, m²

Reluctance is analogous to the resistance of an electric circuit. Up to this point, we have discussed electromagnetic circuits having fixed or slowly changing direct currents and constant or slowly changing magnetomotive forces applied. Circuits with alternating currents will now be considered.

When an alternating current flows through an air-core coil, as shown in Fig. 7-6(*a*), the flux density (*B*) increases and decreases in phase with the force (*H*) that produces it.

The *B-H* curve of an electromagnetic coil with a ferrous [illegible] is shown in Fig. 7-6(*b*). The curve forms

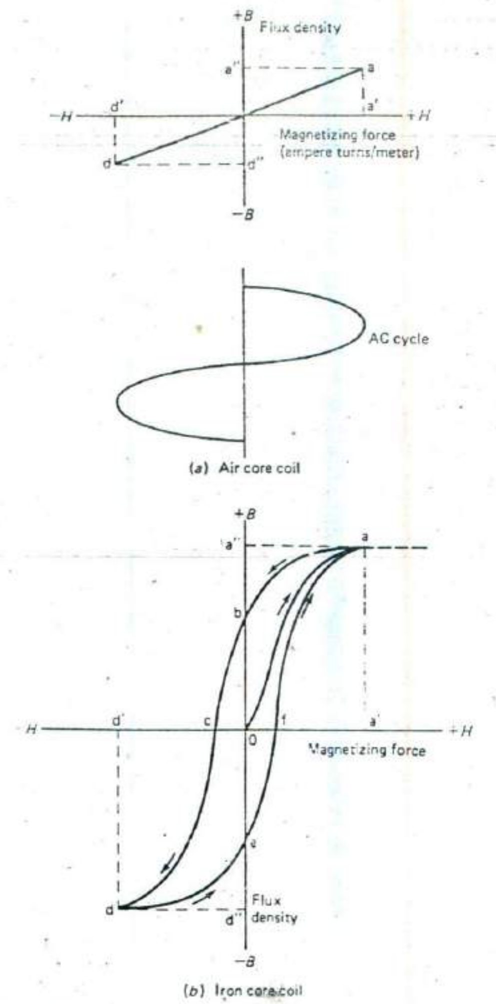


Fig. 7-6 *B-H* curves.

a loop as a result of the residual magnetism. The ability of a core to remain magnetized after the force is removed is called *retentivity*. The greater the retentivity, the greater the residual magnetism will be. It is important to note that in Fig. 7-6(*b*) if the force is increased beyond point a′, the flux density increases very little (as indicated by the dashed line) since *B* has reached saturation. The permeability (μ) is very low at this time since the change in flux density (Δ *B*) is very small. The inductance of the coil is also very low at this time. It should be noted that the permeability (μ) is a direct quantity in the equation for inductance. This, of course, also affects the coil's reactance, since

$$X_L = 2\pi fL.$$

where

f = frequency in Hz
L = inductance in henrys

The reactance of a coil is very low if its core is allowed to saturate and is very high during any time when flux density (B) changes rapidly in response to field-intensity change. Hysteresis losses depend on several factors. Among them are the core type, the temperature of the core, and the frequency of the applied voltage.

Figure 7-7 compares hysteresis loops for high, medium, and low retentivities. A core with high retentivity (Fig. 7-7[*a*]) has a wide hysteresis loop. Use of such a core results in high hysteresis loss. The use of a low-retentivity core results in a narrow loop (Fig. 7-7[*b*]) and low hysteresis loss. If a rapidly varying force is applied, the hysteresis loop will widen. The higher the frequency, the greater the amount of lag and the wider the loop. The use of ferromagnetic cores becomes impractical when the frequency goes too high.

The terms *magnetic amplifier* and *saturable reactor* are frequently used interchangeably. This practice, however, is erroneous because a magnetic amplifier is a device consisting of a combination of saturable reactors, rectifiers, resistors, and transformers. Even though the saturable reactor is the main component in all magnetic amplifiers, the term (*saturable reactor*) applies to only one part, the reactor.

Amplification occurs when a low-level signal controls a relatively large amount of power. In a saturable reactor, a control signal can determine the permeability of the core. It has been shown that a change in μ will change the inductance and thereby vary the inductive reactance (X_L) of the coil. The resistance is low compared to the inductive reactance, so the impedance of the coil is due mainly to X_L. It was also shown that when a core saturates, its μ is very low, and the impedance of the coil in this case approaches the coil's dc resistance. The ideal saturable reactor core would have the highly rectangular B-H curve, as shown in Fig. 7-8. Many materials approach this ideal curve.

When faithful reproduction of a waveform is needed, a core of nickel-iron alloy provides a linear

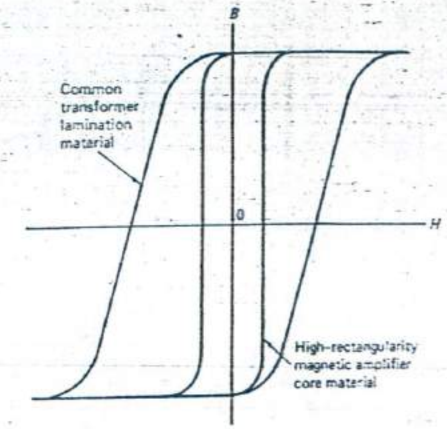


Fig. 7-8 Typical and ideal core materials.

relationship between B and H. Most saturable reactor cores use nickel-iron alloys. These are subdivided into (1) high-permeability (mumetal or Permalloy) and (2) grain-oriented alloys (orthonol, deltamax, and Permirson). High-permeability cores are used in low-level input amplifiers. The grain-oriented materials are used in cores of high-level output stages.

Two desirable features of the saturable rector core are (1) thin laminations for reduced eddy current loss and (2) gapless construction to minimize flux leakage. *Toroidal*, or circular, cores will have much less flux leakage than square or rectangular cores. Flux lines follow smooth curved paths, not sharp corners, as illustrated in Fig. 7-9.

The μ of the core in a saturable reactor is controlled by sending a dc current through a control winding. The flux produced when current flows through the dc control winding is not confined to the core. Some flux lines flow outside the core or through the insulation on the core as shown in Fig. 7-10. Consequently, the entire flux that is produced by the current in the control coil (N_C) does not pass through the load coil (N_L). The amount of leakage increases as the core approaches saturation. This leakage can be minimized by winding the control coil on the same leg of the core as the load winding (N_L), as shown in Fig. 7-11(*a*). Another method is the use of dual

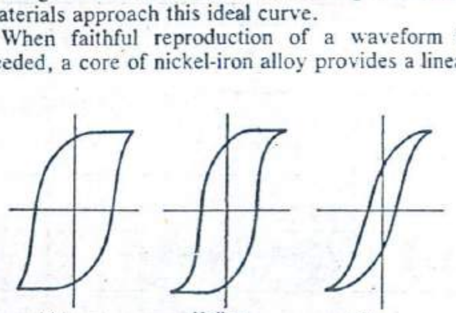


Fig. 7-7 A comparison of hysteresis loops.

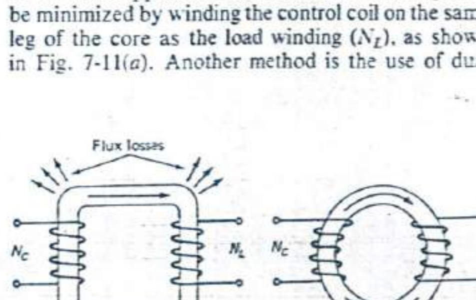


Fig. 7-9 Comparison of a rectangular and a toroidal core.

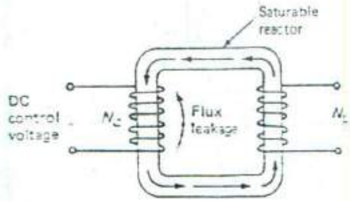


Fig. 7-10 Flux leakages in a saturable reactor core.

coils on two cores, as shown in Fig. 7-11(*b*). Both of these methods require a high amount of insulation between the coils.

Another method is the *three-legged core*, which requires less insulation. The coils are wound separately, as shown in Fig. 7-11(*c*). Reactors are classified as to whether they have a single, double, or three-legged core or according to their construction as follows:

1. Rectangular cores
   a. Stacked
   b. Spiral- or tape-wound

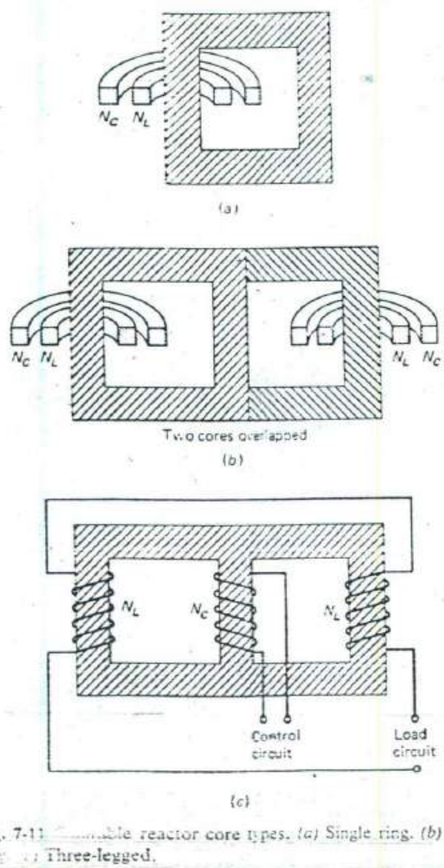


Fig. 7-11 Saturable reactor core types. (*a*) Single ring. (*b*) Twin ring. (*c*) Three-legged.

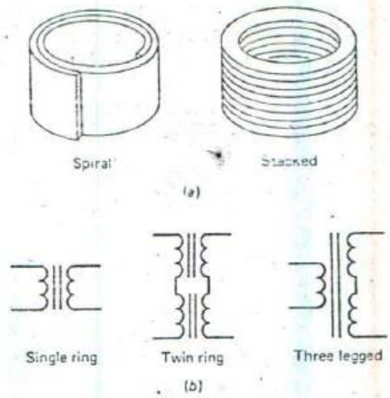


Fig. 7-12 Core types. (*a*) Construction. (*b*) Schematic representation.

2. Toroidal cores
   a. Stacked
   b. Spiral- or tape-wound

Toroidal cores are more common than rectangular cores in reactors. Figure 7-12(*a*) shows the spiral and stacked cores; Fig. 7-12(*b*) schematically illustrates the single-ring, double (twin ring), and three-legged cores. Figure 7-13(*a*) shows a saturable reactor con-

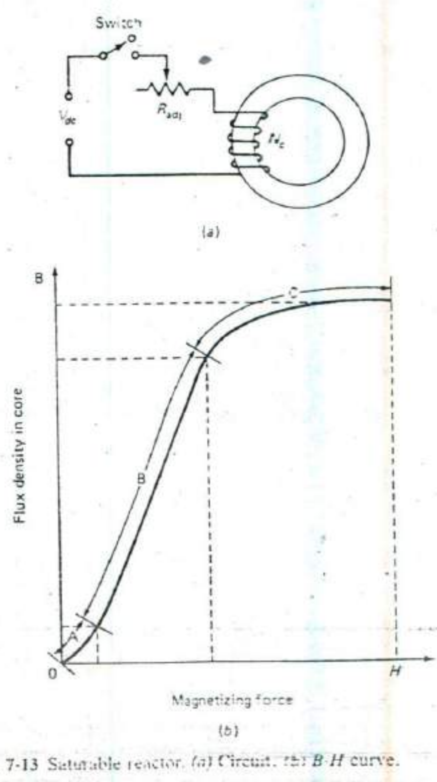


Fig. 7-13 Saturable reactor. (*a*) Circuit. (*b*) *B-H* curve.

trol circuit. It contains a ferromagnetic core with a control winding (N_C), an adjustable series resistor, a switch, and the control voltage V_{dc}.

With the switch closed, the current flowing in the control circuit depends on the amount of series resistance R_{adj}. Therefore, the magnetizing force (H) applied to the core is controlled by adjusting R_{adj}. The *B-H* curve for this core is illustrated in Fig. 7-13(*b*). As the control current is increased from zero, the flux density (B) increases slowly through region *A* of the *B-H* curve. The core permeability is relatively low in this region.

As the control current (magnetizing force H) is increased into region B, the flux density increases very rapidly. The core permeability in this area is high, because a small change in H results in a large change in B. Further increases in the control current will drive the core into saturation, area C. In area C, any increase in H results in only a very small change in B; therefore the permeability is very low. Figure 7-14 shows the addition of the load circuit to complete the saturable reactor circuit.

The load circuit consists of the load winding and a load resistor in series with an ac voltage source. The voltage developed across the load resistor is determined by the current flowing in the load circuit. This current is determined by the inductive reactance, X_L, of the load winding and the resistance of the load resistor. The inductance of a coil is directly proportional to its permeability, μ, and can be expressed as

$$L = \frac{\mu N^2 A}{l}$$

where

L = inductance, henrys (H)
μ = core permeability, B/H
N = number of turns
A = cross-sectional area, m^2
l = magnetic path length of coil, m

The inductive reactance of the coil is directly proportional to its inductance expressed by $X_L = 2\pi fL$.

It has been shown that the permeability of the core can be controlled by varying the control current, thereby changing the operating point on the *B-H* curve. If the control current is set so that the operating point is in the B area of the *B-H* curve of Fig. 7-13(*b*), the permeability of the core is very high. This results in a low current flow in the load resistor.

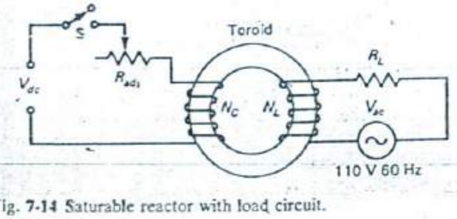


Fig. 7-14 Saturable reactor with load circuit.

and a low voltage drop occurs across the load resistor.

By increasing the control current to the point at which the core is driven into saturation, the permeability is now reduced to a low value. The inductance and inductive reactance of the load winding will also be very low. This will now permit a high load current, and most of the ac source voltage will appear across the load resistor. Accordingly, a saturable reactor uses a small control voltage variation to change or control a large ac load voltage over a wide range.

The basic circuit illustrated in Fig. 7-14 is inefficient because of the ac in the load winding. This current causes two detrimental effects:

1. Voltage is induced in the control winding from the load winding by transformer action.
2. During one half cycle, the load winding current flow will produce a flux which opposes that produced by the control winding. Power from the control winding is needed to return the core flux to its normal operating point. The problem of the opposing flux is eliminated by using a rectifier in series with the load winding. The load current becomes unidirectional, and if the rectifier (diode) is connected properly, the flux fields never oppose each other. The addition of a rectifier converts the saturable reactor into a magnetic amplifier.

Another way of reducing these effects is to use a nonpolarized three-legged core, as shown in Fig. 7-15(*a*). The control winding is wound on the center leg, one load winding on one outside leg, and the other on the opposite leg, so the flux fields cancel in the center leg, as seen in Fig. 7-15(*a*). The schematic

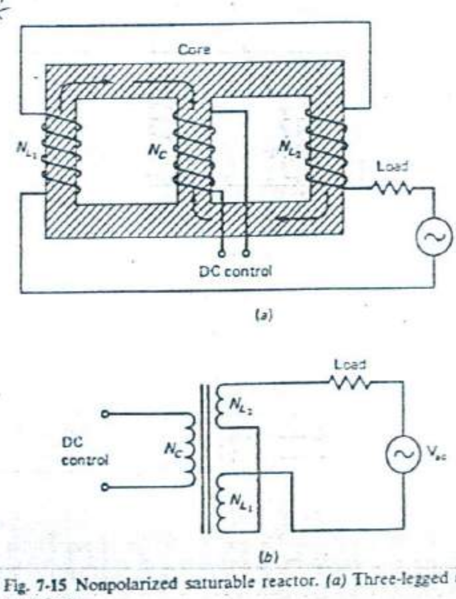


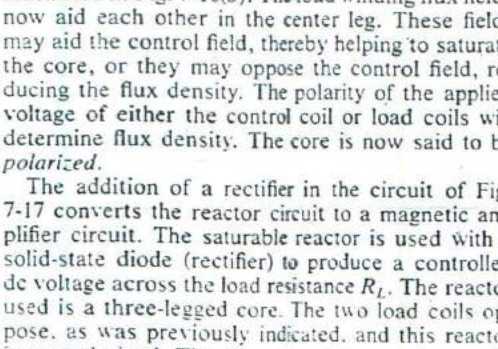
Fig. 7-15 Nonpolarized saturable reactor. (*a*) Three-legged core. (*b*) Circuit.

circuit representation of this configuration is shown in Fig. 7-15(*b*). It can be seen that the two load windings, being in opposition, induce no voltage into the control winding. When the load voltage changes (reverses), the flux fields still oppose each other in the center leg. Also it can be seen that one flux field aids the control flux, and the other opposes it. This characteristic allows a control voltage of either polarity to be used and is the reason for the term *nonpolarized three-legged core.*

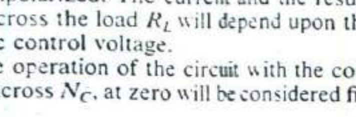
In some instances it is necessary to use a three-legged core with the load winding wound so the flux will aid in the center leg, as shown in Fig. 7-16(*a*). The schematic circuit representation of this core is illustrated in Fig. 7-16(*b*). The load winding flux fields now aid each other in the center leg. These fields may aid the control field, thereby helping to saturate the core, or they may oppose the control field, reducing the flux density. The polarity of the applied voltage of either the control coil or load coils will determine flux density. The core is now said to be *polarized.*

The addition of a rectifier in the circuit of Fig. 7-17 converts the reactor circuit to a magnetic amplifier circuit. The saturable reactor is used with a solid-state diode (rectifier) to produce a controlled dc voltage across the load resistance R_L. The reactor used is a three-legged core. The two load coils oppose, as was previously indicated, and this reactor is nonpolarized. The current and the resultant voltage across the load R_L will depend upon the level of the dc control voltage.

The operation of the circuit with the control voltage, across N_C, at zero will be considered first. When

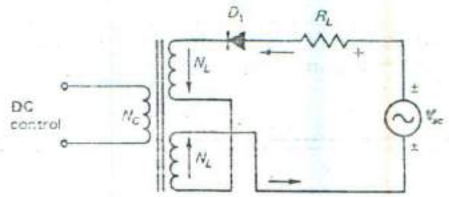


(*a*)

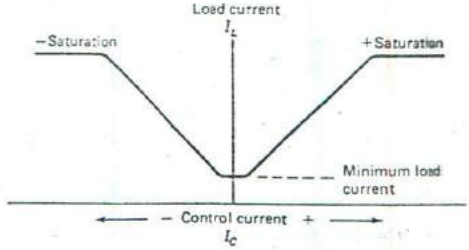


(*b*) Schematic

Fig. 7-16 Polarized saturable reactor. (*a*) Three-legged core. (*b*) Cir...



(*a*)



(*b*)

Fig. 7-17 Basic magnetic amplifier. (*a*) Nonpolarized. (*b*) Transfer curve.

the ac supply voltage is positive, current will flow through R_L and D_1, and through the load windings (N_L) back to the source. With no direct current to saturate the control core, a high permeability exists, which results in a high inductance and a high inductive reactance. The current through the load circuit will be small, and consequently the voltage across R_L will be low.

When a dc potential is applied across the control winding, current will now flow and magnetize the core. If the current is of sufficient magnitude the core becomes saturated, and μ decreases drastically. This decrease will make the inductance of the load windings approach zero, and the load current will be high. Therefore, the voltage across R_L will be high. Increasing the current beyond the saturation point produces an insignificant increase in the load current.

The dynamic characteristic curve can be plotted to illustrate the effects of varying the control current in a magnetic amplifier. This plot is shown in Fig. 7-17(*b*). Note that the direction of the control current is unimportant in a nonpolarized magnetic amplifier. The load current in Fig. 7-17(*b*) never goes to zero, but only to some minimum value. This minimum value of load current, called the *no signal* or *quiescent current,* appears when the control current is zero. The control signal is analogous to the input signal between the base and emitter of a transistor. Amplification is achieved in the transistor circuit because a small input current in the base-to-emitter circuit produces a relatively large change in the collector current. In the magnetic amplifier, small changes in the control winding current cause large changes in load current.

The load current of the magnetic amplifier is also dependent on the size of the load resistance R_L. Figure 7-18 illustrates the output curves for load current I_L for two different values of load resistances. The curve shows that the load current is greater at saturation in the magnetic amplifier that has a lower value load resister. It should be noted that operating on the steep portion of the curve (from point *a* to saturation) will produce a large change in the load current; consequently, amplification is achieved.

As with other types of amplifiers, feedback may be added to magnetic amplifiers. Positive feedback is used to increase the gain, and negative feedback is used to limit the gain or to improve the stability or linearity. Magnetic amplifiers can be classified as follows:

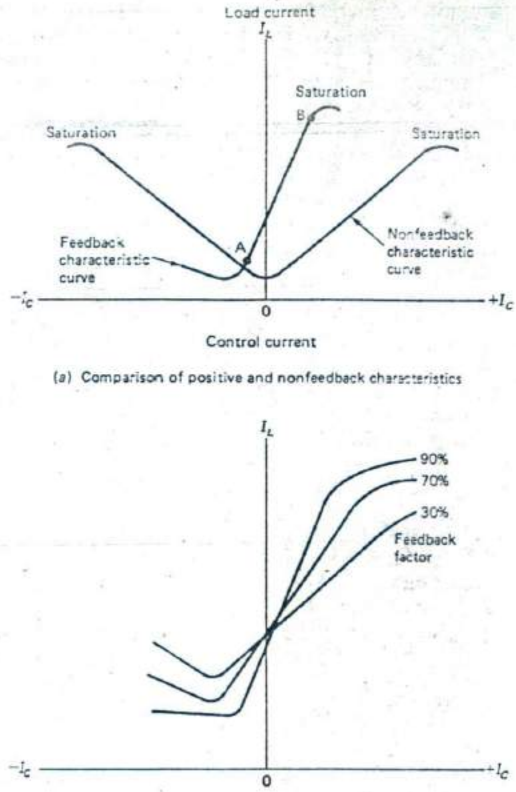
1. Without feedback
2. With external feedback
3. With internal feedback

A magnetic amplifier with external feedback may have positive or negative feedback supplied by means of an external, inductively coupled winding, with the load current flowing through it. Usually, the rectified load current flows through a feedback winding.

If this field in the feedback winding aids the control winding field, the feedback is positive. If the two fields (control and feedback) oppose each other, the feedback is negative. The effect of positive external feedback on the magnetic amplifier is shown in Fig. 7-19(*a*). A comparison of the curves shown illustrates that positive feedback produces a nonsymmetrical curve, whereas the nonfeedback curve is symmetrical. It should also be noted that the load current does not reach its minimum when the control current is zero (as in the nonfeedback case) but reaches its minimum when the control current is negative.

Figure 7-19(*b*) shows the way that the slope of the feedback curve is determined by the percentage of feedback. If the positive feedback is increased above 100 percent, the amplifier will become unstable and will probably lock up into maximum conduction even with very small control current. The feedback factor is usually kept below 85 percent to maintain good stability.

Figure 7-20 shows the schematic symbols of a reactor using a feedback winding. The terms N_C, N_F, and N_L refer to the control, feedback, and load windings, respectively. These designations will be used throughout the remainder of this section. Figure 7-21 shows a simple positive feedback magnetic amplifier with a feedback winding (N_F) added to each reactor.



(*a*) Comparison of positive and nonfeedback characteristics

(*b*) Transfer characteristics of varying amounts of positive feedback

Fig. 7-19 Magnetic amplifier feedback and nonfeedback transfer curves.

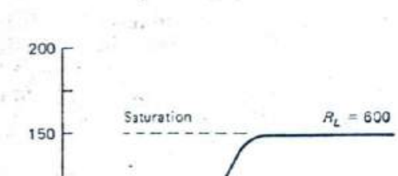


Fig. 7-18 Magnetic amplifier load curves.

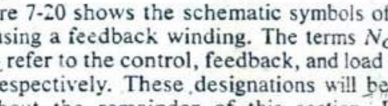


Fig. 7-20 Magnetic amplifier with external feedback.

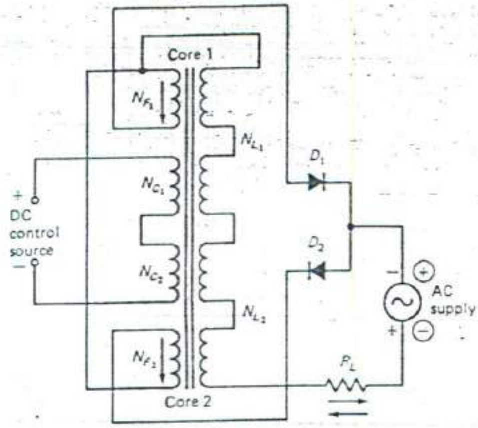
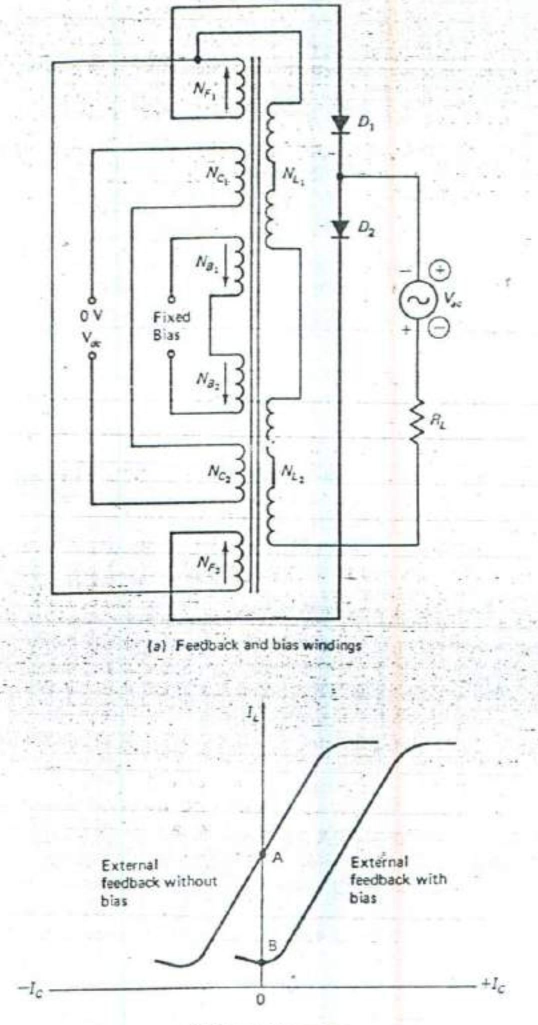


Fig. 7-21 Magnetic amplifier with feedback circuit.

To follow the circuit of Fig. 7-21, it is assumed that the ac supply is on the half cycle having the circled polarity. Current will flow from the positive end of the supply through D_2, N_{F_2}, series load windings N_{L_1} and N_{L_2} of both reactors, and load R_L back to the supply. On the other half cycle, current flows through the load R_L, N_{L_2}, N_{L_1}, the feedback winding N_{F_1}, and back to the supply through D_1. Note that the current that flows in the feedback windings will aid the flux set up in core 1 and core 2 by the dc control voltage. This causes a higher load current to flow and causes the output load voltage to increase.

If the control polarity is reversed, negative feedback occurs. The control winding flux will now oppose the flux of the feedback winding. Under this condition, an increase in the control current will increase the load winding reactance until the point where the control flux and feedback flux are equal is reached. If the control flux overrides the feedback flux, the load winding reactance will slowly decrease, causing the load current to increase. The input impedance of a magnetic amplifier is not affected by external feedback. In contrast, in other electronic amplifiers, the feedback partly determines the input impedance. Since the feedback current of a magnetic amplifier flows through an isolated winding, the input impedance is determined only by the winding resistance and any additional series resistance used to increase the response time. The time constant (T) of an inductive circuit is $T = L/R$; T is in seconds, L is in henrys, and R is in ohms.

It was mentioned previously that increased current will flow through the load winding of a magnetic amplifier using positive external feedback. This tends to make the quiescent load current too high for some applications. Quiescent current reduction may be accomplished by the addition of a bias winding, (N_B), to the core of each reactor, as shown in Fig. 7-22. The bias winding reduces the load current to an acceptable value when the control current is equal to zero. Figure 7-22(b) compares the characteristic curves with and without bias. Operating point A of Fig. 7-22(b) shows a high quiescent current with no bias. This large quiescent current is due to the flux in the core caused by the current through the feedback windings as indicated earlier.



(a) Feedback and bias windings

(b) Comparison curves

Fig. 7-22 Magnetic amplifier with external feedback and bias windings.

Operating point B shows that the quiescent current is minimum in the magnetic amplifier using feedback with a bias winding. Figure 7-22(a) shows that the flux from the bias winding opposes and cancels the flux produced by the feedback winding when the control field is at zero. This effect provides minimum flux in the reactor core and means that the permeability will be high. The inductance and inductive reactance of the load windings are maximum, and the load current, I_L, is at minimum with no control current. But once control current is applied the feedback current will increase and overcome the bias. By varying the bias current amplitude, the operating point may be adjusted to any desired value, as shown in Fig. 7-23.

Feedback in a magnetic amplifier can be accomplished by having the load winding produce the feedback directly. This is called *internal, intrinsic, electric,* or *self-saturated feedback*. This type of circuit feedback has a (rectifier) diode in series with the load so that direct current will flow through the load windings. If a three-legged core is used, the load windings are wound on the center leg of the core and are connected in series aiding. An internal feedback magnetic amplifier is one that uses a self-saturation (auto-excited) circuit because of its nature of operation. Comparison of the characteristics of internal and external feedback amplifiers yields small differences between the two. Self-saturation has the advantage of eliminating the feedback windings and also eliminates the additional resistance in series with the load. On the other hand, use of external feedback allows easy adjustment of amplifier gain.

A basic self-saturating magnetic amplifier is shown in Fig. 7-24(a). The determining factor in this self-saturating magnetic amplifier is the rectifier (diode). This configuration is a polarized magnetic amplifier; therefore, it responds differently if the control winding is reversed in polarity. The full wave output is more common and will be analyzed rather than the half-wave output of Fig. 7-24(a).

To produce a full sine wave output requires a two-reactor magnetic amplifier circuit, as shown in Fig. 7-24(b). The control input voltage for this circuit is also direct current. With no dc control voltage present and the ac supply being that of the circled polarities indicated in Fig. 7-24(b), the current will flow from the ac supply through R_L, D_1, N_{L_1}, and back to the source. On the other half cycle (uncircled polarities) current will flow through, N_{L_2}, D_2, R_L, and back to the source. The load windings N_{L_1} and N_{L_2} of Fig. 7-24(b) are wound series aiding on the center leg of their respective cores.

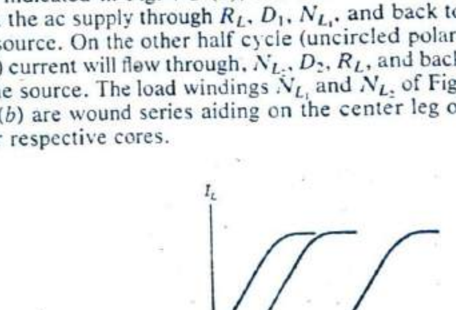


Fig. 7-23 Operating points for various values of bias.

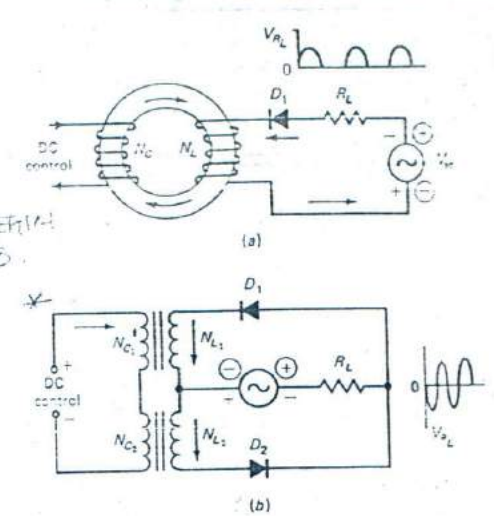


Fig. 7-24 Self-saturating magnetic amplifier. (a) Half wave. (b) Full wave.

Note that the voltage drop across the load R_L will reverse on each half cycle, producing the alternating voltage (V_{R_L}) shown in Fig. 7-24(b).

The polarity and magnitude of the input direct current to the control windings will determine the alternating current developed across the load. With the dc control polarity as shown, the flux produced in the control winding aids the load-winding flux. This will decrease the reactance of the load winding, and the load current will increase, producing a greater voltage drop across R_L. Increasing the load current increases the flux developed in the core. The load winding reactance decreases further, and the load current will increase. Therefore it can be seen that if the load flux aids the control flux, the internal feedback is positive.

When the control voltage is reversed, the control flux will oppose that produced by the load current. This causes the reactance of the load winding to be high, resulting in a low output voltage across R_L. A point where the control flux and the load flux are equal and cancel each other can be reached. The reactance will be maximum at this point, and the load voltage will be minimum. Any further increase in the control voltage will cause the flux in the core to increase, thus decreasing the reactance of the load coils. The load current will increase only a little since the increased load flux opposes the control flux and reduces the overall flux in the core. The internal feedback is now negative.

The transfer characteristic curve for Fig. 7-24(b) (with positive feedback) is shown in Fig. 7-25(a). The quiescent current (point A) is due to the high level of load flux in the core. The control flux aids the load flux, and the gain is high as a result of the positive feedback. Negative control flux opposes

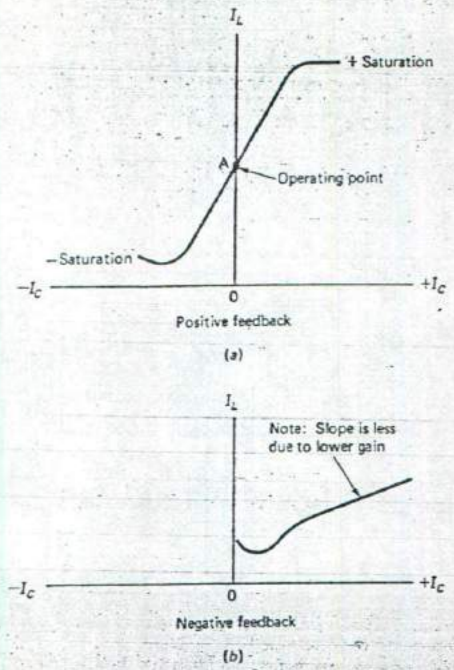


Fig. 7-25 Transfer curve of an internal feedback magnetic amplifier.

the load flux, and the resulting negative feedback results in a reduction of gain. The addition of a bias winding can be used to set the operating point, as discussed previously.

A typical application of a magnetic amplifier is illustrated in Fig. 7-26. Many servosystems (covered in Chapter 10) use a two-phase induction motor to position the load. The direction of rotation in such a motor is dependent upon the phase of the ac signal, and the speed is proportional to the signal amplitude. The circuit of Fig. 7-26 will produce the required output to drive such a servomotor. The bias windings (N_{B_1} and N_{B_2}) minimize the quiescent load-winding current. With no control signal applied, the reactance of N_{L_1} and N_{L_2} are high and equal, and the small current that flows through the field windings of the servomotor produces opposing effects. Thus the rotor will remain stationary.

The control windings of the two reactors of Fig. 7-26 are connected in series opposition, so that a given polarity of control current will increase the amount of core flux of one reactor, and the opposite polarity of control current will increase the amount of core flux in the other reactor. If the circled polarity of control voltage is applied, reactor SX_1 approaches saturation, and the reactance of N_{L_1} decreases. Since this reactor controls the current through field 1 of the servomotor, the decrease in N_{L_1} reactance will allow more current through field 1. The reactance of N_{L_2} remains high so that very little current flows through field 2. A resultant torque is generated within the motor, causing the rotor to turn the output shaft in a particular direction.

If the control winding is driven by the uncircled polarity, the reactor SX_2 will approach saturation. The reactance of N_{L_2} will decrease, allowing more current of flow through field 2 of the servomotor. Winding N_{L_1} has a high reactance, and the current through field 1 is low. A resultant torque is produced, causing the rotor to turn in a direction opposite of that previously considered. The speed of the servo is approximately proportional to the magnitude of the control (error) signal.

REVIEW QUESTIONS

5. Materials with a permeability much greater than 1 are called ______.

6. Air core coils have a higher permeability than iron core coils. (true or false)

7. Referring to Fig. 7-6(*b*), the inductance is minimum (μ is low) at what points?

8. The higher frequencies of operation will produce a ______ hysteresis loop.

9. A saturable reactor may operate in either the A or B regions of Fig. 7-13(*b*). (true or false)

10. The saturation current for a load of 800 Ω on Fig. 7-18 would be ______ than 0.1 A.

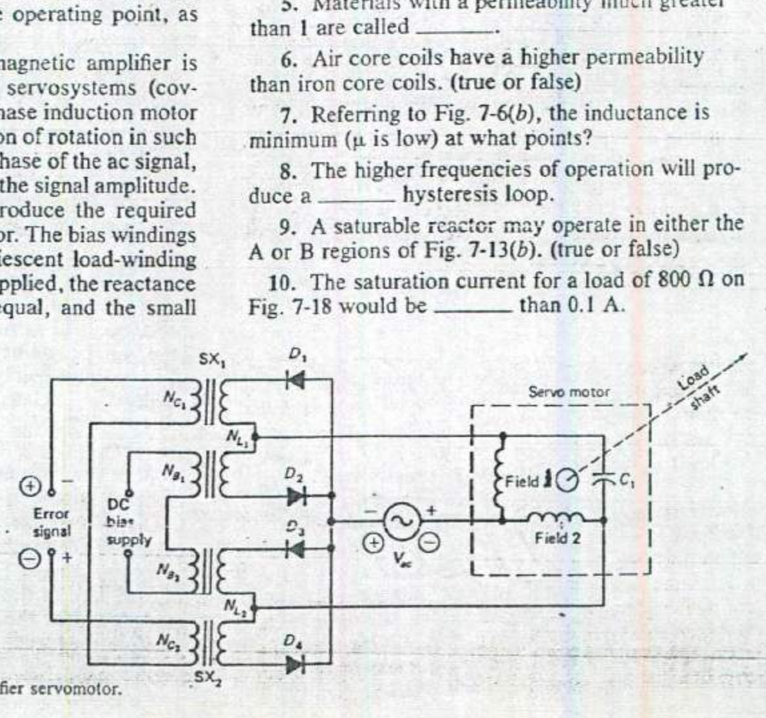


Fig. 7-26 Application of magnetic amplifier servomotor.

7-3 TROUBLESHOOTING AND MAINTENANCE

In general, magnetic devices are fairly hardy and less sensitive to abuse than solid-state devices. However, excessive current produces excessive heat, which can cause insulation failure and short-circuited turns within the inductor. As always, cooling and ventilation must be adequate and not restricted in any way. A visual inspection may show signs of overheating. A distinct "burned" odor is another clue. Some units have a built-in thermal switch to protect them. Tripping of this device can be an indication of future or current problems and should be thoroughly investigated. An inductor may fail by short-circuiting to its core (frame). This is easily checked with a high-voltage ohmmeter (megger). Most inductors have breakdown voltages greater than 500 V. Unless a very low resistance short circuit has occurred, an ohmmeter with a 10- to 30-V source is not adequate for most insulation tests for breakdowns.

The dc resistance of a high-current inductor is on the order of milliohms and would require a Wheatstone bridge or milliohmmeter to check its value adequately. If an inductance bridge is available, measurement of the inductance will also verify the integrity of the coil. In a switching power supply or in a line filter, a test frequency close to the device's operating frequency will yield a more meaningful value when using an ac bridge or impedance-measuring equipment. Do not interrupt a choke-type circuit when current is flowing. The induced voltage can be many times the applied voltage and may cause arcing and a breakdown of the insulation. Also, it creates a shock hazard, and the arc may cause eye damage. The arc caused by interrupting a 5-A current flow in a several-henry inductor is in excess of 1000 °F.

An abnormally high output voltage from a dc supply using a choke input filter can occur if the load is below the minimum for the critical inductance value used. Excessive load current can saturate the core and lower the inductance, causing abnormally high ripple.

Signal transformers should be tested for short circuits to the core or frame as well as between the windings. Since most units are low-power, the use of a signal or function generator is best suited for dynamic testing along with an oscilloscope or wide-band ac voltmeter. The pulse- or trigger-type transformer can be tested, as mentioned in Sec. 7-1. The insulation resistance of these devices is suspect in many cases and is best checked with the aid of a high-voltage ohmmeter. If the thyristor it triggered had a fault (anode-to-gate short circuit) the trigger transformer should be checked for signs of overheating (discoloration or a burnt odor). A complete test may save a future return call for a marginal system.

The saturable reactor and the magnetic amplifier require no preventive maintenance other than cleaning and inspecting any associated cooling systems. The windings should all be well insulated from each other and the core. This is easily verified with a megger. The diode (rectifier), if used, can be checked in the usual manner. The MOSFETs may find applications in magnetic circuits, especially in their constant-current mode (gate tied to source). An oscilloscope is an invaluable tool to troubleshoot these circuits. Cases of saturation or distortion can only be detected this way. Caution: The alternating current used may or may not be referenced to ground or neutral. The use of an isolation transformer may or may not be a safe way to prevent ground loops through the test equipment. Floating measurements are covered in Chapter 1. Changes in the bias source can shift the quiescent point and cause excessive no-signal load current and possible overheating. The loss of positive feedback (if external) will give a reduction in gain or sensitivity, whereas loss of negative feedback can give such an increase in gain as to lock up (latch) the system, which will become immune to all external signals. A comparison of each winding for double-reactor systems is very beneficial, and these values should be recorded for future reference.

The devices covered in this chapter are ac devices, and in most cases (except extreme short circuits, etc.) the dc resistance is not a good measure of the integrity of the device, especially where turns ratios are involved. Short circuits to the core (frame) are common, especially if overheating had occurred some time in the past. A high-voltage ohmmeter is a very useful tool for these types of checks.

REVIEW QUESTIONS

11. Suspected electric leakage of an inductor is best tested by using a ______.

12. A Wheatstone bridge can be used to check the dc resistance of an inductor. (true or false)

13. Interrupting the current through an inductor may create a dangerous ______.

14. A load waveform that appears as half-wave direct current in Fig. 7-24(*b*) is due to D_1 being open or short-circuited?

CHAPTER REVIEW PROBLEMS

7-1. The polarity of transformer windings may be indicated with phasing ______.

7-2. The constant-voltage transformer uses a magnetic ______, along with a compensating winding to stabilize the voltage.

7-3. One method used to reduce leakage flux is the use of ______ ring cores.

7-4. A basic magnetic amplifier is a saturable reactor with a ______ added.

7-5. The minimum value of load current is called the ______ or quiescent current.

7-6. Positive feedback will ______ the gain of a magnetic amplifier.

7-7. Positive feedback is kept below 85 percent to protect a magnetic amplifier from ______.

7-8. A ______ winding is used to minimize the quiescent current of the magnetic amplifier.

7-9. Negative feedback increases the input impedance of the magnetic amplifier. (true or false).

7-10. The two-reactor ______ circuit will provide a full sine wave to the load.

7-11. The servomotor magnetic amplifier typically has how many reactors?

7-12. An increase in the ripple from a filter may be due to excessive current causing it to ______.

7-13. Trigger transformers are best tested by using ______.

7-14. The ______ are usually suspect in a magnetic amplifier failure since they are the least reliable part of the system.

7-15. Loss of negative feedback can make the reactor ______.

7-16. Excessive quiescent current may indicate a loss of the ______.

ANSWERS TO REVIEW QUESTIONS

1. 9000 Ω **2.** isolation **3.** 20 kHz **4.** equal **5.** ferromagnetic **6.** false; opposite is true **7.** a or d **8.** wider **9.** true **10.** greater **11.** megger **12.** true **13.** arc **14.** open

8

OPEN-LOOP MOTOR CONTROL

This chapter treats open-loop motor control and power conversion, including inverters, choppers, and cycloconverters. It also covers phase control of dc motors and static conversion as applied in ac motor circuits. The improved reliability and increased ratings of solid-state devices have made their presence commonplace in industry. They also interface well with computer-type controls; that interface is the major thrust of automation technology.

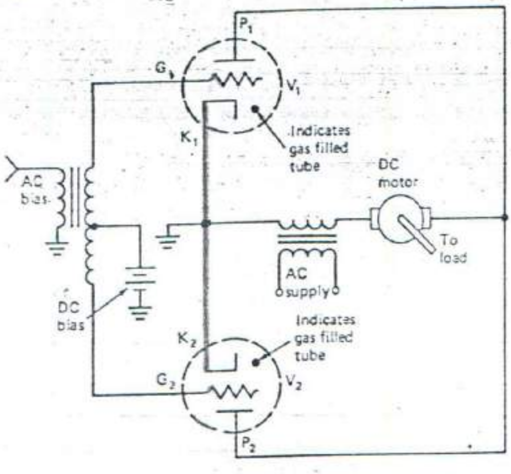
8-1 DC MOTOR PHASE CONTROL

There are two basic categories of motor control: open-loop and closed-loop. A *closed-loop system* senses the motor output and uses this information to correct the drive to the motor to eliminate error. Closed-loop systems are covered in Chapter 10. In some cases, where the load on the motor is reasonably constant, open loop motor control is adequate. The half-wave drive circuit discussed in Chapter 4 is an example of an open loop control system.

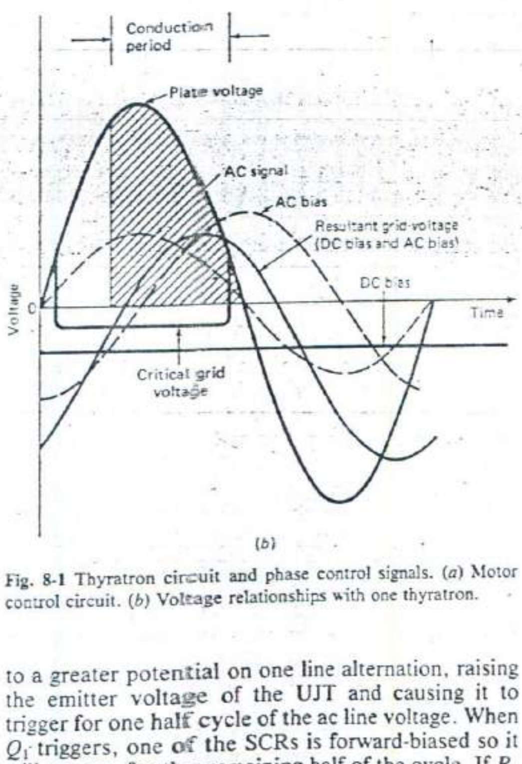
For many years, dc motors were controlled by *thyratrons*, vacuum tubes containing an inert gas, a heated cathode, an anode (plate), and a control grid located between the cathode and the anode. These tubes can control loads from a few milliamperes to several amperes. A simplified thyratron motor control circuit is shown in Fig. 8-1(*a*). The thyratron's plate-to-cathode resistance is very low when the tube is in the conducting state. The grid of the thyratron controls the point at which the tube fires or the gas in the tube ionizes. Once ionization takes place, the grid loses control and cannot stop plate current flow. The tube can only be extinguished by lowering the plate voltage below the ionization level (usually 15 V). The thyratron is either full on or full off, just like a switch. The device is operationally comparable to the solid-state silicon controlled rectifier (SCR). Motor performance is controlled by shifting the firing point of the control device, whether it is a thyratron or an SCR. This is known as *phase control* and is illustrated in Fig. 8-1(*b*). When the dc bias and ac bias add together to equal the critical grid voltage, the thyratron fires. The other thyratron fires on the next half cycle. The conduction period can be lengthened for increased motor output or shortened for decreased motor output.

A phase-sensitive control circuit capable of supplying reversible half-wave power for a permanent magnet or shunt dc motor is shown in Fig. 8-2. The circuit is called a *balanced-bridge reversing drive circuit*. It consists of two half-wave circuits back to back (SCR_1, D_1 and SCR_2, D_2) triggered by the unijunction transistor Q_1, on either the positive or negative half cycle of the applied line voltage. Which half-wave circuit fires will depend upon the direction of the imbalance of the reference bridge from the value of R_1 (the sensing element), which can be a thermistor, photodiode, potentiometer, or the output from a control-type amplifier. In some cases, an opto-isolation amplifier may be employed to avoid the direct connection between the sensor and the bridge circuit.

Resistor R_5 is set so the dc bias on the emitter of the UJT, Q_1, is just below the peak point which would trigger Q_1. Zener D_4 sets the dc voltage across R_5 and C_1 (the RC time constant for Q_1). With R_1 and R_2 equal, the bridge circuit (R_1, R_3, and R_2, R_4) will be balanced, and the UJT (Q_1) will not trigger; therefore no output will appear across the load (armature). If R_1 increases in resistance it will unbalance the bridge circuit. Capacitor C_1 will be charged



(a)



(b)

Fig. 8-1 Thyratron circuit and phase control signals. (*a*) Motor control circuit. (*b*) Voltage relationships with one thyratron.

to a greater potential on one line alternation, raising the emitter voltage of the UJT and causing it to trigger for one half cycle of the ac line voltage. When Q_1 triggers, one of the SCRs is forward-biased so it will turn on for the remaining half of the cycle. If R_1 were to decrease in resistance a similar action would occur, but with the other SCR triggering on the opposite half of the cycle. This reverses the polarity across the load, so now the motor will reverse its direction from the previous imbalance condition. Resistor R_2 is used to match the quiescent resistance of sensor R_1 at its null (balanced) position. No feedback is employed, so this is an open loop control. Any changes in the load (torque, speed, etc.) will not automatically be corrected.

If a series motor is used, a circuit similar to that of Fig. 8-3 may be employed. This circuit uses a triac in lieu of SCRs. In this circuit, the triac is triggered on either the positive or negative half cycle. The bridge rectifier (D_3—D_6) provides a dc voltage for the armature circuit. Thus, the armature current is always in the same direction. However, the field current is reversed, depending on the triac's triggering polarity. This circuit employs controls for gain, balance, and *dead band* (that range of input voltage to which a control does not respond). The circuit also provides for an analog input control signal, from a sensor, transducer, or process controller.

Figure 8-4(*a*) illustrates the internal block diagram for a silicon integrated circuit designed for phase control from ac mains with resistive or inductive loads. The IC has a voltage regulator and a voltage monitor circuit to reset timing functions and inhibit triac firing pulses when a power-up occurs. A ramp generator is provided to control motor acceleration and can be controlled by the speed program input, pin 5. Charging currents for the ramp generator are set by an external resistor for a slow ramp and internally for a fast ramp. A frequency-to-voltage converter is included to enable a rate generator (tachometer) to be used for motor speed sensing.

The control amplifier in Fig. 8-4(*a*) has differential inputs that compare the ramp voltage against the actual speed voltage (if used in a closed loop system). Synchronization of the triac pulse is achieved by delaying the pulse with reference to the zero voltage points of the line voltage. Inductive loads such as motors produce a phase lag of the load current. Under high-speed or heavy-load conditions, the triac must be fired after the load current from the previous cycle has ceased. The current synchronization input (pin 1) performs this task by ensuring that there is a voltage across the triac before a trigger pulse is permitted (when the triac is conducting only a small voltage drop appears across it). The gating pulse width is dependent upon an external capacitor connected to pin 14, which also delays the pulse from the zero voltage point.

Figure 8-4(*b*) shows a permanent magnet motor whose armature is driven from an ac supply by means of a bridge rectifier. When driving highly inductive loads of this type with a phase control circuit, the triac cannot be gated until a quarter cycle after the line zero crossing because of the electromotive force (emf) generated by the armature. Resistor R_{11} senses the motor current, and the developed voltage signal is applied to pin 3 of the IC. The motor can be rated up to 220 V, and currents up to 30 A are possible. Resistor VR_1 is used to adjust the speed of the motor, which could range up to 10,000 rpm. The circuit provides unidirectional motor operation which is adequate for many industrial applications.

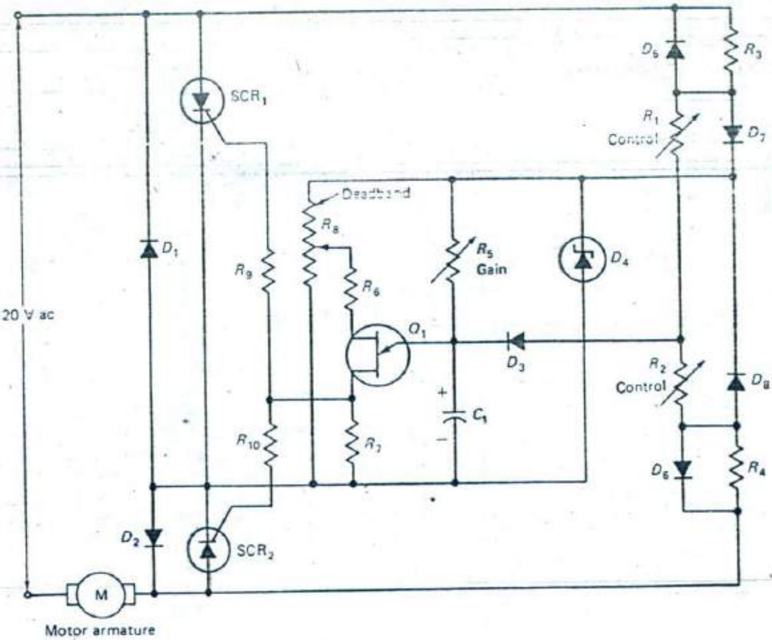


Fig. 8-2 Balanced bridge (reversing) drive for PM or shunt motors.

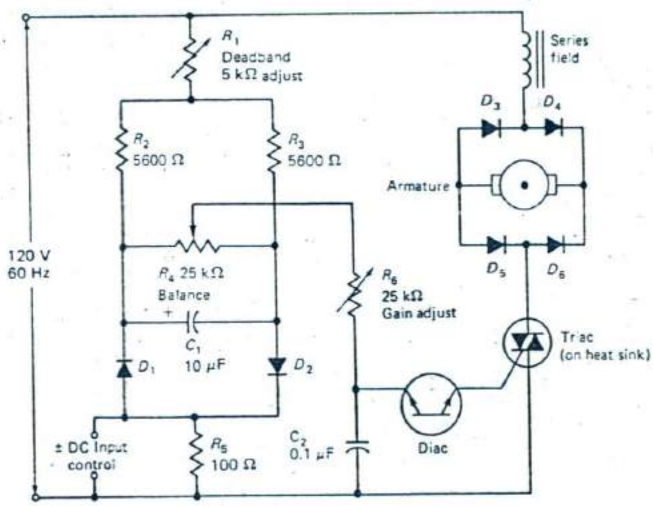


Fig. 8-3 Phase control circuit for series dc motor.

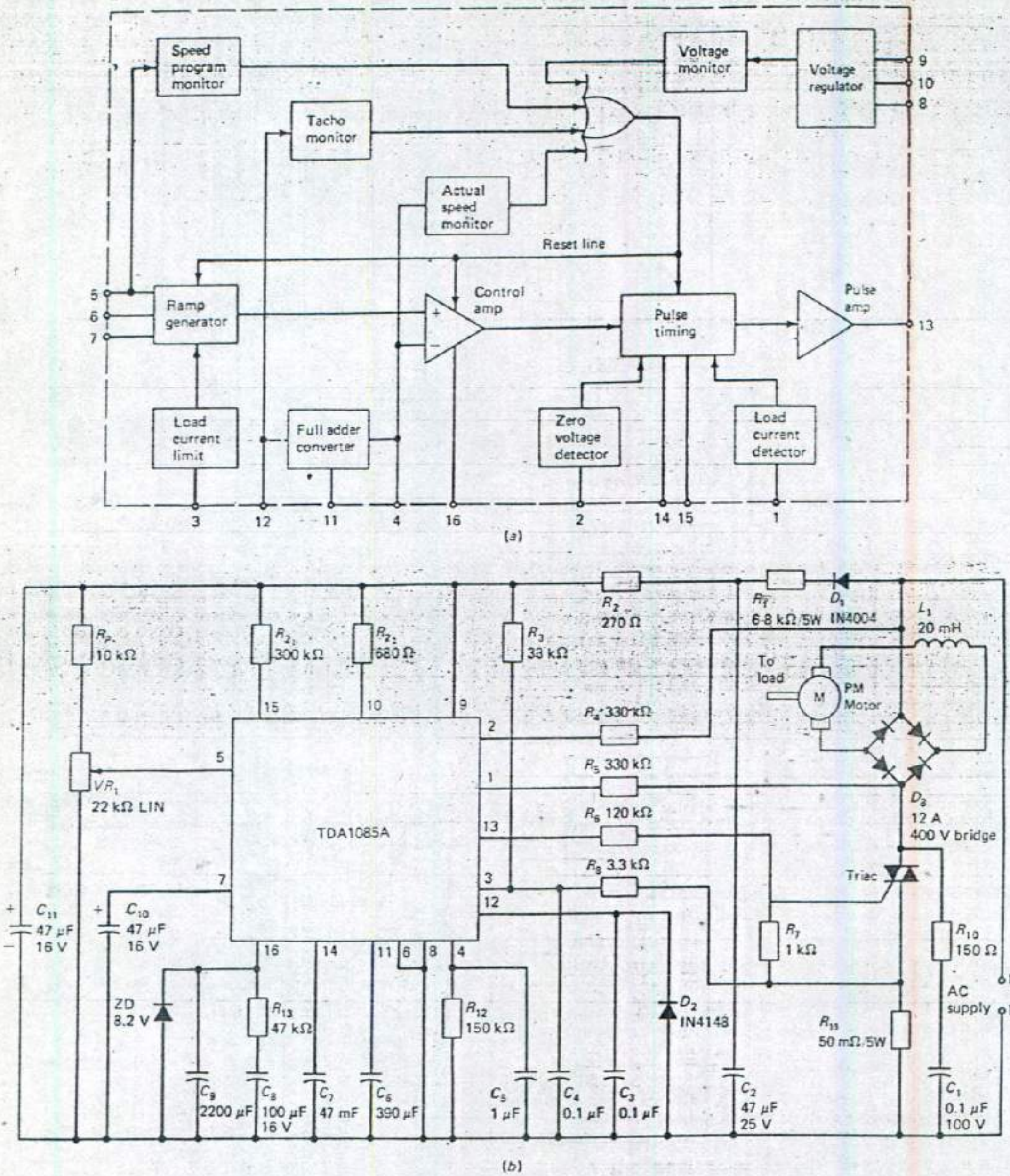


Fig. 8-4 Integrated circuit for phase control.

REVIEW QUESTIONS

1. The thyratron is a linear control device. (true or false)

2. The thyratron tube is analogous to the solid-state ______.

3. Phase control of a load is accomplished by shifting the ______ point of the thyratron or SCR.

4. In Fig. 8-2, ______ is used to balance the sensor (R_1) resistance at null.

5. In Fig. 8-2, which solid-state device produces the gating pulses for the SCRs?

6. The range of input at which a control circuit produces no output is known as its ______.

8-2
DC-DC CHOPPER CONTROL

The dc counterpart to ac phase control is the dc-dc chopper (converter). Choppers control the average load voltage by switching a fixed dc source. The switching may be accomplished by bipolar transistors, thyristors, or MOSFETs. The basic circuits for these devices are shown in Fig. 8-5(*a*). Figure 8-5(*b*) shows a typical chopper waveform. The average load voltage can be controlled by several methods. There are three techniques commonly used to vary the ratio of the switch ON time to the switch OFF time (duty cycle) of the "chopped" output waveform. The duty cycle can be adjusted by:

1. Varying the ON time (pulse width modulation).
2. Varying the OFF time (pulse rate modulation).
3. Varying both times, which is a combination of pulse width and pulse rate modulation.

Figure 8-6 shows chopper action for pulse width modulation, pulse rate modulation, and a combination of both. Note that the duty cycle controls the average voltage. Choppers can be used in variable-speed dc drives to supply the armature voltage for speed control of separately excited dc motors. They can also be used to provide a variable-supply voltage for series dc motor-speed control. The chopper offers the advantage of higher efficiency than that of traditional electromechanical devices. The improved efficiency is due to the elimination of the wasted energy when using starting and control resistances for these motors.

The thyristor makes an ideal switch for chopper applications; however, a method of turning off the thyristor must be incorporated. The turn-off methods (commutation) are discussed in section 8-3. Figure 8-7 shows a diagram of chopper control of a vehicle motor. Contacts S_2, S_3, S_4, and S_5 are field-reversing relay contacts. With S_2 and S_5 closed, the vehicular direction is forward. With S_3 and S_4 closed, the direction is reversed. This type of chopper arrangement has a typical duty cycle from 20 to 80 percent. At the stopped condition, all four relay contacts (S_2,S_3,S_4,S_5) are open. With S_2 and S_5 closed and the chopper at low speed (rate), the motor voltage averages about 20 percent of the battery voltage. This voltage may be increased to 80 percent of the battery voltage to provide greater motor output by increasing the duty cycle of the chopper. When the 80 percent point is reached, contact S_1 is closed to apply full

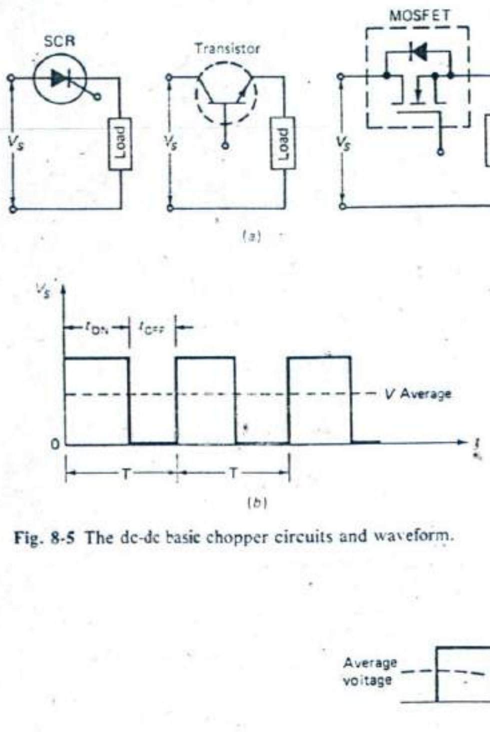


Fig. 8-5 The dc-dc basic chopper circuits and waveform.

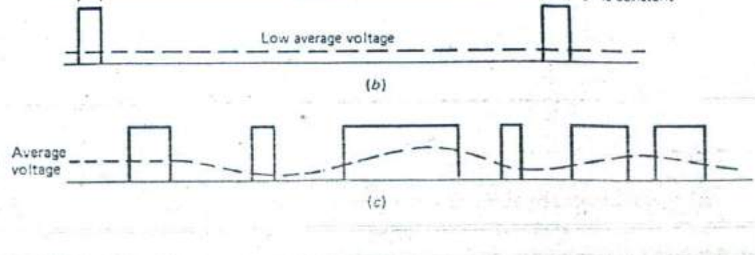
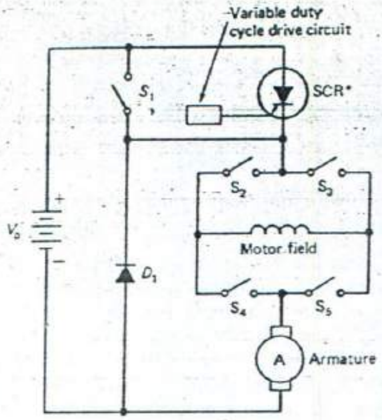


Fig. 8-6 Chopper pulse width modulation waveforms. (*a*) Pulse width modulation. (*b*) Pulse rate modulation. (*c*) Combination rate and width modulation.



*Note: SCR commutation circuit not shown

Fig. 8-7 Basic vehicle control circuit.

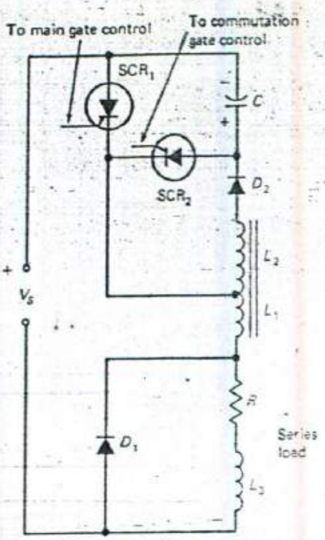


Fig. 8-8 Basic Jones chopper circuit.

battery voltage to the motor, and maximum torque will be obtained. Diode D is the freewheeling diode. Its purpose is to protect the motor from high voltage transients that can be produced when the thyristor is turned off. The chopper may use a variable-frequency constant-pulse width (rate modulation) system or a pulse width modulation control if so desired.

Choppers are an energy-efficient method for controlling a series dc motor that normally operates from a dc power source supplied by a third rail (as with electric trains) or an overhead conductor (trolley) or a battery bank, as with electric fork lifts. One basic problem encountered in chopper control is the maximum armature current that can be commutated by the thyristor. As the motor size (horsepower rating) increases, so will the locked rotor armature requirements. Another problem encountered when using thyristors in a chopper circuit is to achieve commutation, without which the control would be lost. Commutation is more or less automatic in ac power systems since the thyristors will turn off at the line zero crossings. Commutation in dc systems requires extra circuitry.

One circuit that achieves dc commutation is shown in Fig. 8-8. The Jones circuit controls the mean load voltage by varying the ratio of the ON time to the OFF time (*pulse width modulation*). To understand the operation of the Jones circuit better, six working circuits illustrating the various phases of operation are shown in Fig. 8-9. These phases represent time intervals from t_0 through t_5. Single-pole, single-throw (SPST) switches replace SCRs to indicate their state so as to simplify the analysis. The cycle starts in Fig. 8-9(a) by triggering SCR_1 at t_0. The lower plate of capacitor C will start to charge positively. By time t_1, the lower plate of the capacitor has resonantly charged, via L_2, to its peak positive voltage as shown in Fig. 8-9(b). This peak positive voltage is equal but opposite to the peak negative voltage found on the capacitor just before SCR is triggered. Turning on SCR_1 has the effect of reversing the voltage across the commutating capacitor. This voltage is held on the capacitor by the charging diode D_2. Energy is now available to commutate SCR_1, which is conducting load (motor) current through L_1 as shown in Fig. 8-9(b). Inductors L_1 and L_2 are closely coupled and form an autotransformer. When SCR_1 is on, the load current flows through L_1 and induces a positive-going voltage at the top of L_2 that charges the commutating capacitor.

At time t_2 in Fig. 8-9(c), SCR_2 is triggered on, and the capacitor is now effectively connected across SCR_1. This reverse-biases the thyristor, and SCR_1 is commutated (turned off). The load current is now being supplied by V_B and the capacitor. Figure 8-9(d) shows that the capacitor is now charged to an opposite polarity as the load current continues to flow. As the capacitor voltage reaches the supply voltage, the charging current decreases and eventually goes below the holding current for SCR_2. Thyristor SCR_2 turns off at time t_4. Now, both thyristors are off, and the motor current is decreasing. Diode D_1 is forward-biased by the emf generated by the collapsing motor field. This allows the energy stored in the motor to be dissipated. At t_5, all currents have ceased, and the circuit is ready for a gate pulse to SCR_1 to begin another cycle.

The Jones circuit provides an efficient control of motor speed by varying the duty cycle. To increase motor speed, the time between SCR_1's gating pulse and SCR_2's gating pulse will be increased. This allows motor current to flow for a greater period of time and raises the average motor voltage. The discussion has been limited to basic circuit operation.

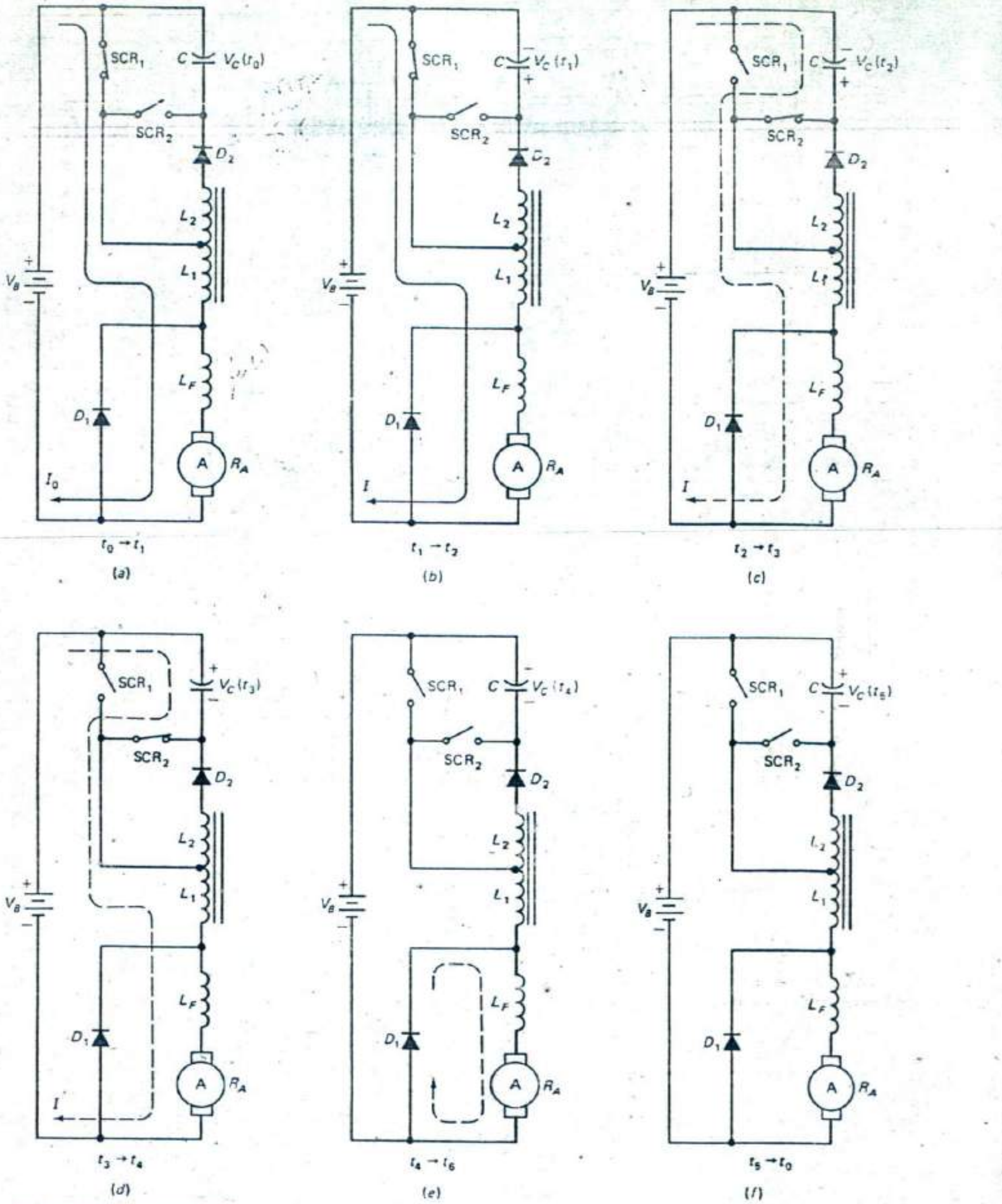


Fig. 8-9 Jones chopper working sequences.

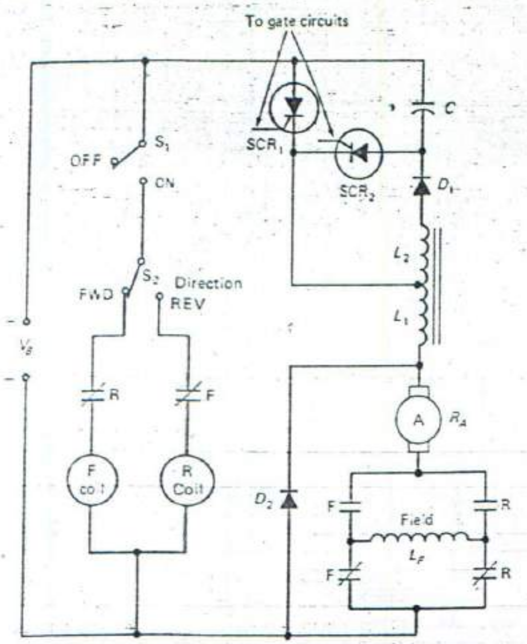


Fig. 8-10 Jones circuit modified for forward and reverse.

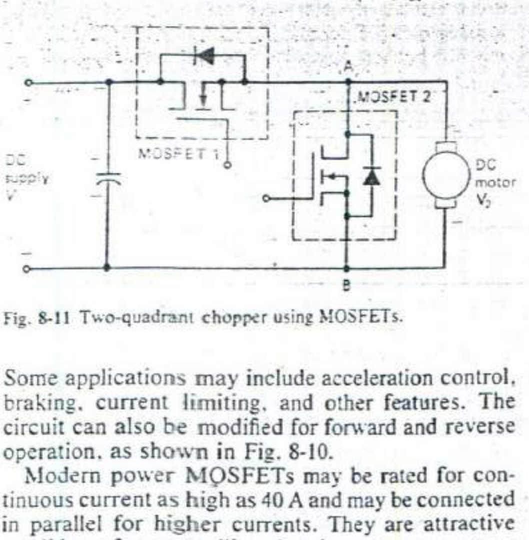


Fig. 8-11 Two-quadrant chopper using MOSFETs.

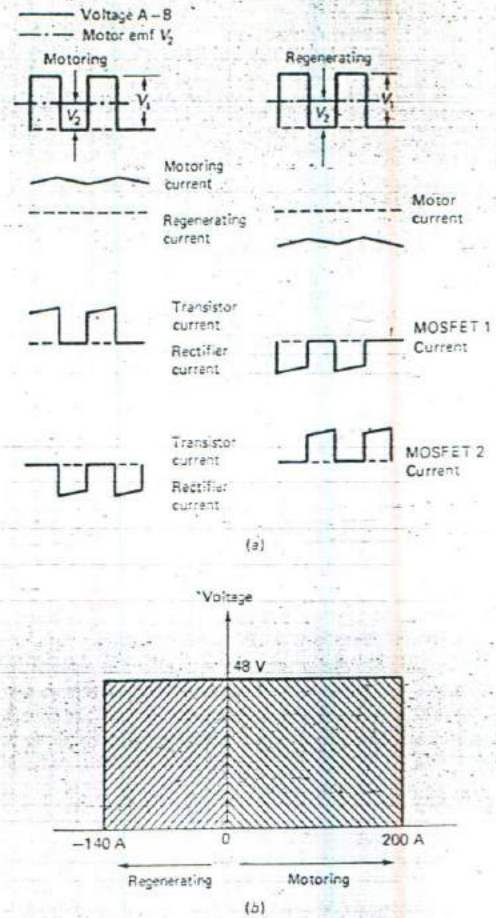


Fig. 8-12 Chopper waveforms and quadrant of operation.

Some applications may include acceleration control, braking, current limiting, and other features. The circuit can also be modified for forward and reverse operation, as shown in Fig. 8-10.

Modern power MOSFETs may be rated for continuous current as high as 40 A and may be connected in parallel for higher currents. They are attractive candidates for controlling electric motors at currents up to several hundreds of amperes. Figure 8-11 shows the basic circuit for a dc-to-dc chopper that provides continuous speed control in the motoring mode of operation (with the motor receiving power from the dc source). It also allows the motor to return regenerative energy to the dc source. This is known as a *two-quadrant chopper circuit*. Waveforms that describe the operation are shown in Fig. 8-12(*a*), and Fig. 8-12(*b*) defines the two operating quadrants of the circuit. When in the motoring mode, MOSFET 1 is switched on and off, at an appropriate repetition rate, providing control of the average voltage applied to the motor. At this time MOSFET 2 is off, and its diode acts as a conventional freewheeling diode to conduct the freewheeling motor current when MOSFET 1 is off. When the motor acts as a generator and returns energy to the dc source, MOSFET 2 is chopped on and off to control the current fed back from the motor to the source. In this mode, MOSFET 1 is off, but its internal diode is used to carry the motor current to the dc source during the intervals when MOSFET 2 is off. In order for a motor to regenerate, it is necessary for it to have either a shunt or separately excited field. A series-connected field is not feasible because of the reversing of connections required and is not practical.

Figure 8-13 shows a partial schematic for a 48-V, 200-A two-quadrant chopper based on power MOSFETs. This unit employs ten MOSFETs connected in parallel for the motoring switch and five MOSFETs connected in parallel for the regenerating switch. Continued improvements in power MOS-

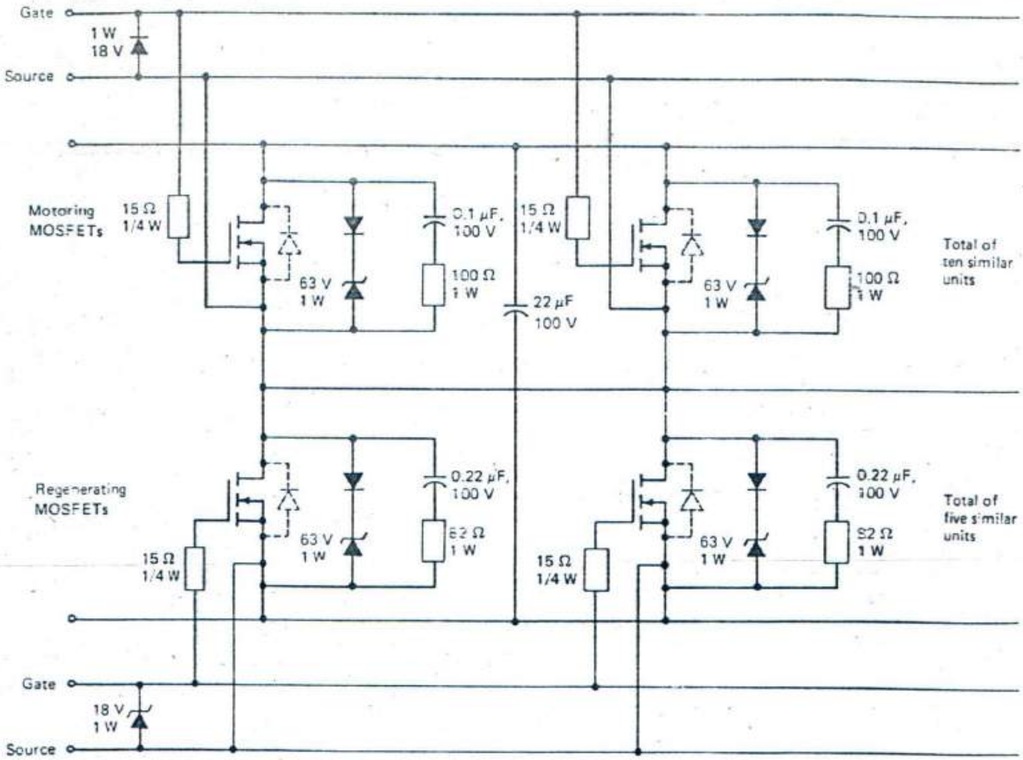


Fig. 8-13 Power circuit schematic.

FETs are expected to make motor control circuits of this type economically and technically superior to chopper systems employing transistors or thyristors.

REVIEW QUESTIONS

7. Three common solid-state choppers are the transistor, the MOSFET, and the ______.

8. Chopper circuits control the average load voltage by one of three ______ methods.

9. The freewheeling diode in the Jones circuit of Fig. 8-8 is ______.

10. In Fig. 8-7 S_1 is used for dynamic braking of the motor. (true or false)

11. A basic problem when using a thyristor in a chopper is being able to ______ the device.

12. In Fig. 8-8 L_1 and L_2 form an ______.

8-3 COMMUTATION CIRCUITS

The gate has no more control over a thyristor once the device is triggered on for any current exceeding the latching current. External means must be applied to stop the flow of current through the device. The two basic methods used for turning thyristors off are:

1. Current commutation
2. Forced commutation

Current commutation may be achieved by opening or closing a switch. Figure 8-14(*a*) shows a series switch; Fig. 8-14(*b*) shows a shunt switch. Commutation is achieved by opening the series switch or by closing the shunt switch. Each case of switch operation would produce a high value of the rate of change of voltage across the SCR. Mechanical switches are seldom used for current commutation. There are some static switching circuits, but this mode of commutation is generally not employed in industrial circuits.

Forced commutation employs a momentary reverse bias to turn off a thyristor. The reverse bias must be applied for a period longer than the device's turn-off time. Also, the rate of rise of the reapplied voltage must not exceed the critical value. With inductive loads, the stored energy of the collapsing field must be diverted from the thyristor by a freewheeling diode or some other method. The six

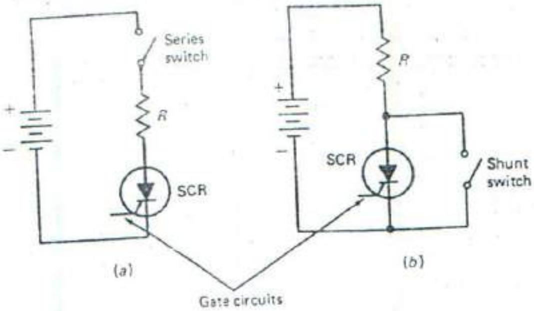


Fig. 8-14 Current commutation.

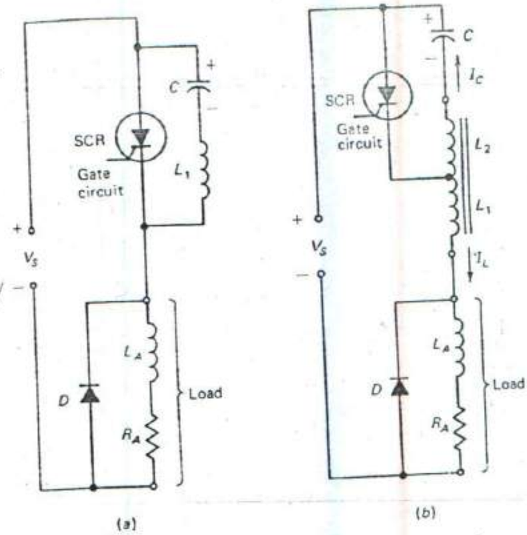


Fig. 8-16 Parallel capacitor-inductor/Morgan commutation circuits.

classes of forced commutation are given various names by different manufacturers in their literature. Once the circuits are compared, the exact method of commutation may be determined. Understanding of the commutation method used is the important point, not its name or class.

The simplest form of forced commutation is self-commutation and employs a series capacitor for resonating the load, as shown in Fig. 8-15. When the SCR is turned on, the capacitor is charged to the source voltage through the thyristor. The current will decay below the holding current as the capacitor voltage approaches the supply voltage V_s. With an underdamped resonating inductive load, the voltage on the capacitor will reverse and exceed the applied voltage V_s to assist in the thyristor turn-off. The load forms part of the tuned circuit. This method of commutation is sometimes employed in inverter circuits. In circuits with unidirectional load current, there must be a method to discharge the capacitor. Figure 8-15(*b*) shows the use of a parallel resistor; Fig. 8-15(*c*) shows the preferred method, using a second SCR in parallel with the capacitor.

Another method of forced commutation is the *LC* self-commutated or parallel capacitor-inductor circuit shown in Fig. 8-16(*a*). The previously discussed series-capacitor (resonant load) circuit has limited control range, and load variations affect its operation. Placing the underdamped *LC* circuit in parallel with the SCR, as shown in Fig. 8-16(*a*), eliminates this problem. When the power is turned on, the capacitor will charge up to source voltage V_s through inductor L and the load. When the SCR is turned on, current will flow through the load and the capacitor, and inductor L_1 will begin an oscillatory action. After one-half cycle, the stored energy will reverse-bias the SCR and reduce its current to a value less than the holding current, and the SCR will then turn off. The capacitor will start to recharge to

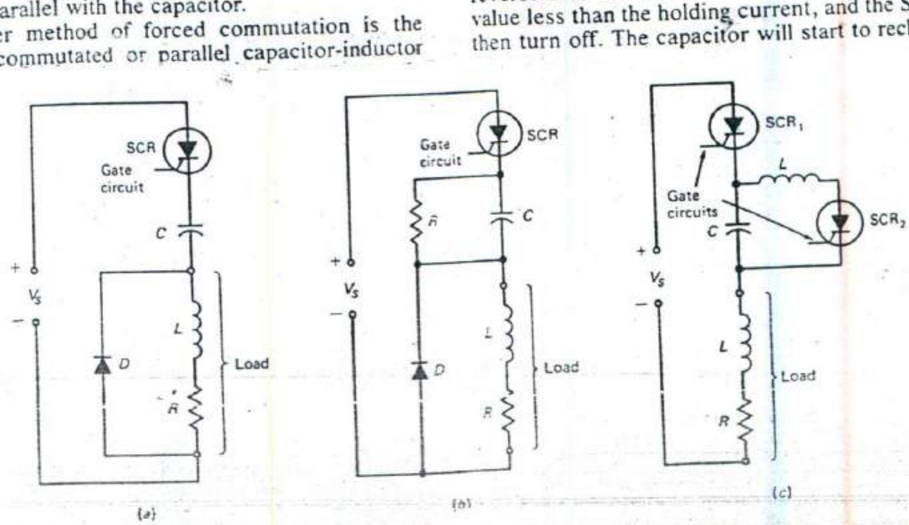


Fig. 8-15 Series capacitor commutation. (*a*) Basic circuit. (*b* and *c*) Improved circuits.

its original polarity through the inductor and the load. The circuit has several restrictions. The mode of operation can only be pulse rate modulation. The SCR can only be fired after the capacitor has fully recharged, or the circuit will fail to commutate. The ON time t_{ON} is limited to $\pi\sqrt{LC}$, and t_{OFF} is load-dependent. These restrictions of t_{ON} and t_{OFF} limit the range of load voltage control.

An improved version of this circuit is shown in Fig. 8-16(*b*). The circuit, known as the *Morgan circuit*, utilizes the properties of a saturable reactor. When the reactor is unsaturated, its inductance will be high. When the reactor is saturated, its inductance will be low. This point is covered in Chapter 7. Before the SCR is turned on, the capacitor is charged to *V* (supply) as shown. With no current flow through the tapped saturable reactor, the core is unsaturated, and the reactor is in the high-inductance state. When the SCR is triggered, the capacitor current (I_C) and the load current (I_L) flow in opposing directions through the reactor. The reactor tends to keep the rate of rise equal for both currents, and this rate of rise is determined mainly by the load inductance. As soon as full load current is flowing, capacitor current (I_C) will decrease, and load current will remain constant. The voltage induced across L_2 will initially be very small and reversed in polarity. This voltage will build in value as the charging current decreases. The period of oscillation is long, since the reactor core is unsaturated (high *L*). As the capacitor-charging current I_C goes to zero, the reactor current approaches I_L, and the reactor core saturates. Now the L_2 inductance will be very small. The period of oscillation is thus short, and the stored energy in L_2 reverse-biases the SCR and turns it off. The remaining charge on the capacitor is then dissipated in the load and is ready to recharge to the supply voltage V_S. As I_L approaches zero, the reactor core unsaturates, and the cycle can begin again when the SCR is turned on.

A parallel-capacitor commutation circuit is shown in Fig. 8-17. When SCR_1 is turned on, the capacitor is charged to the supply *V* through R_A, with the right-hand plate positive as shown. Commutation is initiated when SCR_2 is turned on, applying the capacitor across SCR_1. The discharge current of the capacitor opposes the load current in SCR_1, and the SCR commutates.

An alternative parallel-capacitor configuration, in which the capacitor-charging current flows through the load, is shown in Fig. 8-18. The capacitor is charged with the polarity indicated to the supply potential by turning on SCR_2. When the capacitor is fully charged SCR_2 will automatically commutate off. When SCR_1 is gated on, V_S supplies load current I_L. Now SCR_1 provides a discharge path for the capacitor through D_1 and L_1. The capacitor voltage will reverse as L_1 maintains current flow after the first resonant period. Now D_1 prevents the charge from being drained off the capacitor. Commutation is initiated by turning SCR_2 on. This reverse-biases SCR_1 and reduces the current below the holding value. When SCR_1 is commutated, the capacitor recharges to its original condition through SCR_2 and the load.

Another class of commutation depends on the commutation energy being supplied from an external source. There are several common configurations that may be used; two of them are shown in Fig. 8-19. In Fig. 8-19(*a*) SCR_1 is commutated off by means of an auxiliary transistor switch Q_1. The thyristor is assumed to be initially on when turn-off is desired. A signal applied to the base of Q_1 turns it on and reverse-biases the SCR. The SCR is now commutated off. The drive signal to the base of Q_1 must be of sufficient duration to ensure thyristor turn-off and of sufficient amplitude to place Q_1 in saturation. If Q_1 comes out of saturation before the thyristor turn-off is complete, commutation failure results. Therefore, Q_1 and its drive circuitry are selected to satisfy worst-case conditions for SCR turn-off under the heaviest load conditions. Figure 8-19(*b*) utilizes a pulse transformer with a square-loop *B-H* core to achieve thyristor turn-off. When the SCR is conducting, the core of the pulse transformer is saturated (low impedance) by the load current. When the time comes to turn off the SCR, the first step is to unsaturate the pulse transformer. The application

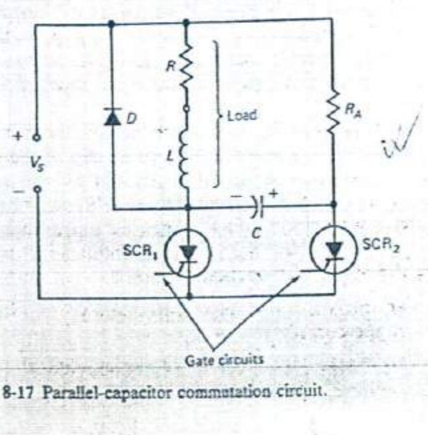


Fig. 8-17 Parallel-capacitor commutation circuit.

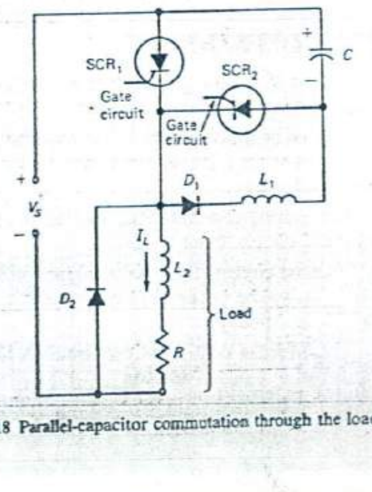


Fig. 8-18 Parallel-capacitor commutation through the load.

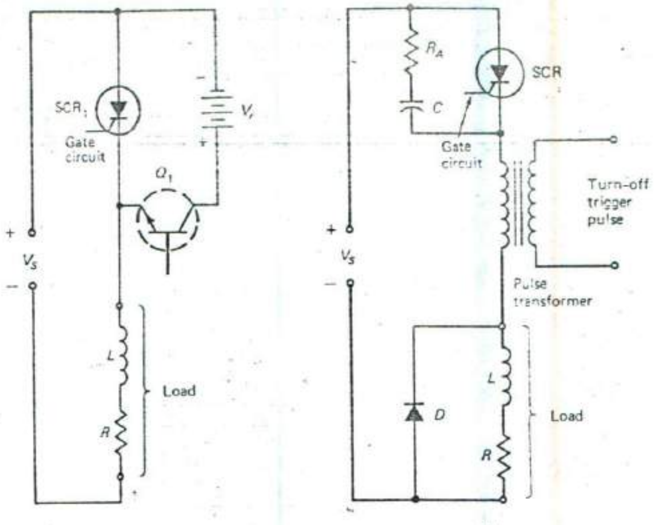


Fig. 8-19 External commutation techniques.

of a drive pulse to the primary reverses the flux in the transformer for several microseconds. A voltage pulse is developed across the pulse transformer secondary, which reverse-biases the SCR and turns it off. If an inductive load is present, a freewheeling diode is connected across the load to prevent damage to the SCR from voltage spikes developed by the inductive load discharging. External pulse commutation circuits allow both pulse width and pulse rate modulation techniques to be used. Note that the commutation is independent of the load current and the supply voltage source.

If the supply is an alternating voltage, as shown in Fig. 8-2, load current will flow only during the positive half cycle. During the negative half cycle the SCR will turn off because of the negative polarity across it. The only constraint is that the half cycle must be longer than the turn-off time of the SCR.

REVIEW QUESTIONS

13. Two methods of thyristor commutation are _______ and forced commutation.

14. Forced commutation requires that the _______ bias be applied for a longer time than the thyristor's turn-off time.

15. Resonating commutation uses an overdamped *LC* circuit. (true or false)

16. In most circuits that use a capacitor for commutation, it will be necessary to charge the capacitor to _______ polarity.

17. The Morgan commutation circuit is identified by its use of a _______ reactor.

18. In the external transistor commutation circuit, the transistor must be driven into _______ to turn off the thyristor.

8-4 STATIC FREQUENCY CONVERTERS FOR AC MOTOR CONTROL

The ac motor, especially the squirrel cage induction type, possesses many virtues in comparison to dc motors. These include significantly lower cost, weight, and inertia; higher efficiency; and fewer maintenance requirements. They are also capable of operating in dirty and explosive environments because they do not have a commutator and brushes.

In spite of the many virtues of the ac motor, the cost of converters and circuit complexity were main factors preventing the widespread application of ac drives. However, as the ratings of solid-state devices continuously improve and as their cost decreases, variable-speed ac drives are increasing in popularity. Integrated circuit technology is also assisting in this changeover as complex circuits are reduced to the chip level.

The speed of an induction motor is determined by the synchronous speed and by the slip of the rotor. The synchronous speed is dictated by the supply frequency and the number of poles. Slip can be controlled by regulating the voltage or current supplied to the motor. There are several methods for controlling the speed of induction motors:

1. Variable-voltage constant-frequency or stator voltage control
2. Variable-voltage variable-frequency control

3. Variable-current variable-frequency control
4. Regulation of the amount of slip

The term *inverter* normally refers to equipment used for transforming direct to alternating current. A cycloconverter is used for transforming a higher-frequency alternating current to a lower-frequency without any intermediate dc link.

The most commonly used ac drive system is the variable-frequency dc-link inverter. Polyphase induction motors or synchronous motors are employed because their operating characteristics are retained over the range of the inverter, which is typically from 10 to 200 Hz. Variable-frequency drives are found in machine tools and textile, paper, and steel mill equipment. In most variable-frequency drives, a constant voltage per hertz is maintained up to the rated frequency of the motor, and then the stator voltage is maintained at its rated value as the frequency is increased. Failure to maintain a constant volts/hertz ratio affects the torque output and can cause an increase in stator current and may overheat the motor.

Figure 8-20 shows a simple two-transistor inverter. The circuit uses a square wave fed through transformer T_2 to drive the bases of Q_1 and Q_2. Each transistor is alternately switched into saturation. When Q_1 is on, Q_2 is off, and the current flows through the top half of the primary of T_1. Later, Q_1 is turned off and Q_2 is turned on. The current now flows in the bottom half of the T_1's secondary. This induces a square wave in the secondary of T_1. This square wave output is filtered (L_1, C_1) to make it more sinusoidal before it is applied to the load.

A MOSFET three-phase full-wave inverter is illustrated in Fig. 8-21. This type of inverter is suitable to drive a three-phase induction motor and is capable of a variable-frequency output. This type of inverter is often referred to as a *two-on inverter* since two MOSFETs are always on. Figure 8-22 shows simplified schematics and the firing sequence required for three-phase outputs. The output waveforms are shown in Fig. 8-23. Figure 8-22(*a*) shows the state of the MOSFETs for 0° to 60°. Transistor Q_2 connects the positive end of supply V to motor terminal A while Q_3 connects the negative end of V to the B terminal. Therefore, from 0° to 60° the voltage across terminals A and B (V_{AB}) is $+V$. The voltage from terminals A to C (V_{AC}) and from C to B (V_{CB}) is $+1/2V$ since the two windings are in series across V_{AB}. The waveforms show V_{CA} and V_{BC}, the inverse of V_{AC} and V_{CB}. Hence, V_{CA} and V_{BC} are equal to $-1/2V$ for the 0° to 60° timing sequence.

Figure 8-22(*b*) shows Q_2 and Q_5 on, which makes V_{AC} equal to $+V$ (or V_{CA} equal to $-V$). Both V_{AB} and V_{BC} are $+1/2V$ for 60° to 120° as shown in Fig. 8-23 (page 192). By further analysis of Fig. 8-22(*c,d,e,f*), the three-phase composite waveforms shown in Fig. 8-23 are produced.

A similar type of inverter can be obtained by using thyristors in the three-phase configuration shown in Fig. 8-24(*a*) on page 193. The commutation and firing circuits have been omitted to simplify the explanation. The thyristors are fired in a sequence to produce positive-phase sequence output voltages V_{AB}, V_{BC}, and V_{CA} at the output terminals A, B, and C.

From Figure 8-24(*b*) it can be seen that SCR_1 con-

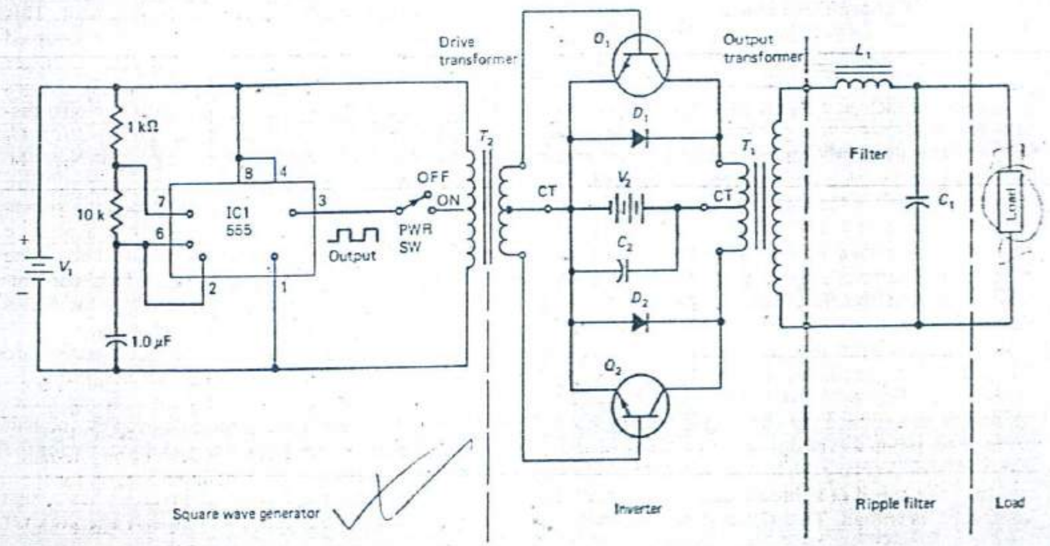


Fig. 8-20 Simple single-phase inverter.

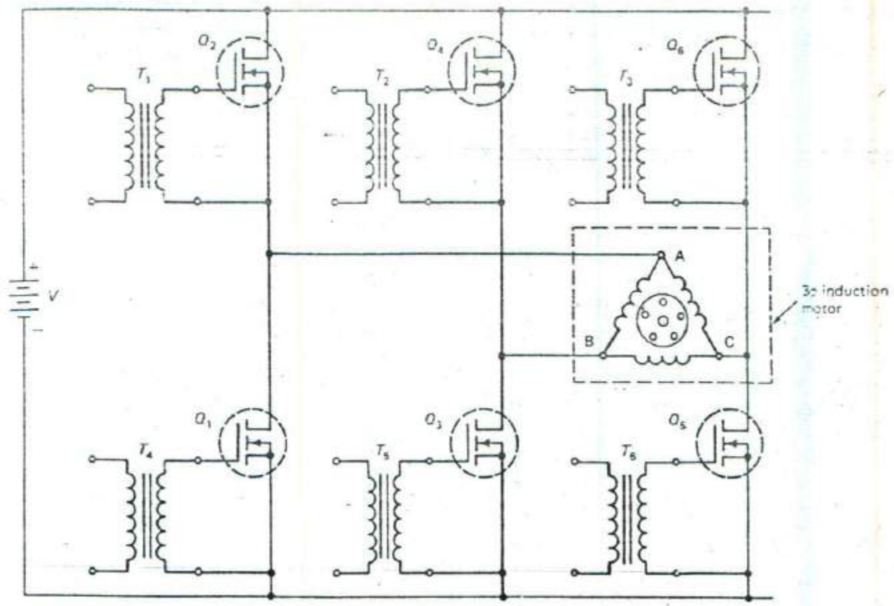


Fig. 8-21 MOSFET three-phase inverter for an induction motor.

ducts from 0 to 180°, and SCR_4 conducts from 180 to 360°. This condition relies on SCR_1 being commutated prior to SCR_4 firing; if there is any time delay in the turn-off SCR_1, a short circuit or "shoot through" can exist between the positive and negative supply rails through SCR_1 and SCR_4. The associated waveforms along with the gating pulses are shown in Fig. 8-24(*b*).

Figure 8-25(*a*) on page 194 shows a wye-connected load for the six-step three-phase inverter shown in Fig. 8-24(*a*). Figure 8-25(*b*) shows the load connections for the 0 to 60° time interval, with SCRs 1, 5, and 6 conducting. Two windings of the wye load are always in parallel, and one is in series with the parallel pair. Thus, as a voltage divider, the parallel pair drops $1/3V$ while $2/3V$ is dropped by the single winding. This assumes that the wye load is balanced. The output voltages for the load V_{AN}, V_{BN}, and V_{CN} are similar to those of Fig. 8-23. One important difference, since the load is wye-connected, is that the peak voltage (across any winding) to the neutral is $2/3V$ instead of the full V output as realized by a delta load.

The trigger circuit for the inverters is often obtained from a ring counter. A three-state ring counter is shown in Fig. 8-26 (page 195). The output frequency is determined by the applied input square wave. This input waveform is generated by a separate oscillator. Varying the frequency of the oscillator will vary the speed of a three-phase induction motor connected as a load. Two three-state ring counters may be used to trigger a six-step inverter. When the power is applied through S_2, capacitors C_1 to C_6 charge to V through T_1, T_2, or T_3 and the associated resistors. The input square wave is differentiated by C_C and R_5 to produce the spikes shown in Fig. 8-26. The spikes would trigger all three SCRs through diodes D_1 to D_3, but the positive charge on C_1 to C_3 reverse-biases the diodes and keeps them off. When the start switch (S_1) is closed, C_1 is discharged by R_6 and the other resistors in the discharge circuit. This discharge allows a positive spike to be coupled through D_1 to C_1 to fire SCR_1 and produce a trigger pulse at the secondary of T_1. This in turn discharges C_4 through SCR_1 and R_4. Capacitor C_2 is also discharged through R_2 and the 1-kΩ resistor. With C_2 discharged, D_2 is no longer reverse-biased. Now the next positive spike from the differentiator will fire SCR_2. The firing of SCR_2 through R_4 and the subsequent discharge of C_5 through R_4 produce a positive voltage pulse at the cathode of SCR_1. This pulse momentarily reverse-biases SCR_1, which commutates. The circuit operation continues as the SCRs fire in a sequence of 1–2–3–1–2, and so on.

The voltage waveforms for two-ring counters are shown in Fig. 8-27 (page 195). The (*a*) group was just discussed; the (*b*) waveforms are for a second three-state ring counter whose input is delayed by one-half period to produce the remaining input triggers for the inverter switches.

The output of these types of inverters has a high harmonic content, especially when they operate over

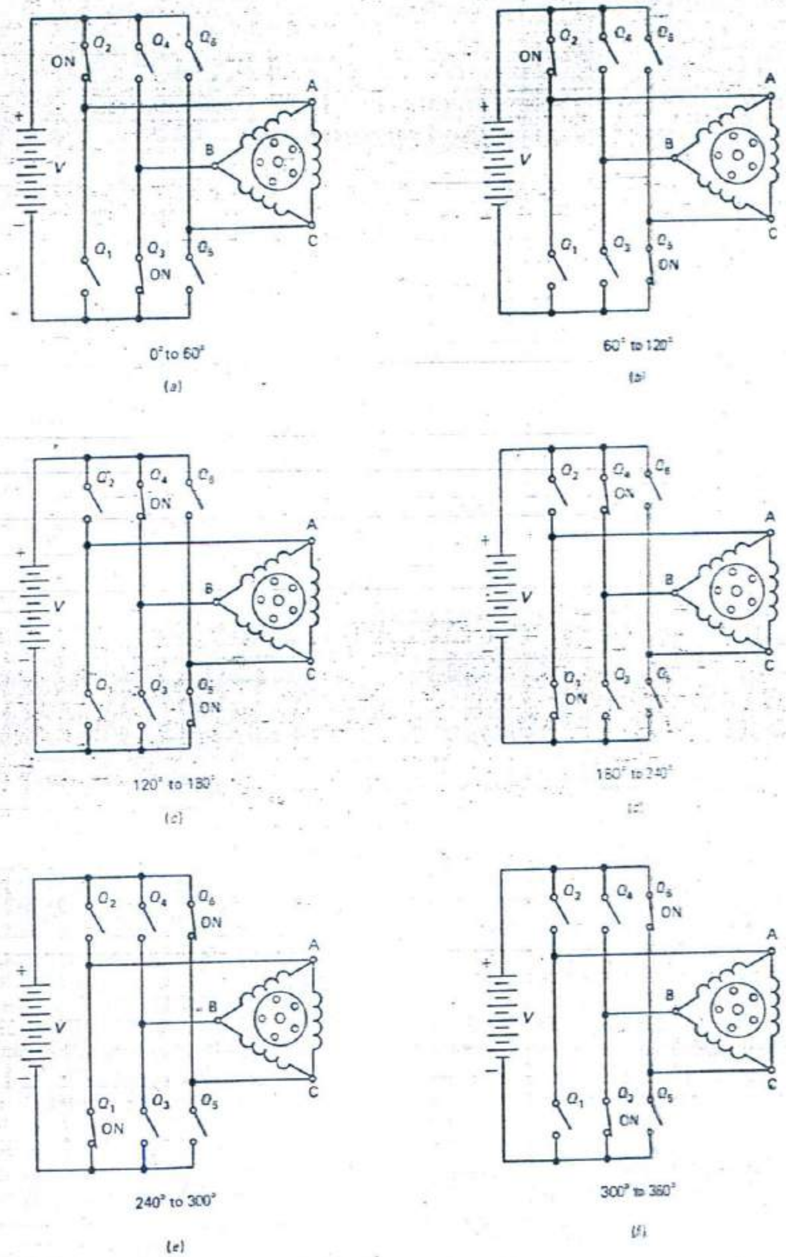


Fig. 8-22 Simplified firing sequence of Fig. 8-21.

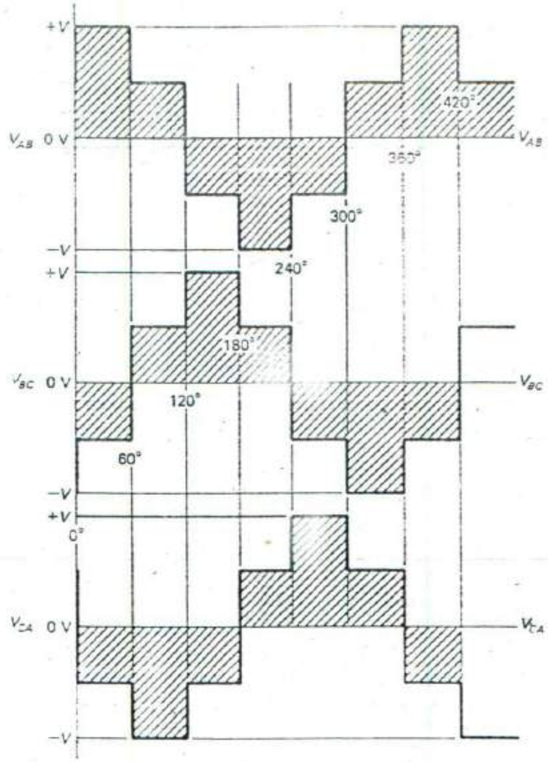


Fig. 8-23 Three-phase inverter output waveforms.

a wide frequency range such as 10 to 200 Hz. Output filtering is not practical because of the wide range of frequencies. There are several other methods that may be used to minimize the output harmonic content. One way is to use pulse width modulation techniques within the inverter. Another method is to combine a number of square-wave inverters with each one phase-shifted and fired at the desired output frequency. This method is known as *harmonic neutralization*. Figure 8-28 (page 196) shows how a three-phase inverter can be constructed by using three single-phase bridge inverters. This arrangement eliminates the third harmonics from the inverter's output.

The single-phase cycloconverter is covered in Chapter 3. To review, the cycloconverter can produce a variable-frequency output by using phase-controlled converters. They normally operate at a frequency of one-third the supply frequency or less. The major advantages for using the cycloconverter are (1) elimination of the intermediate dc link, thereby improving the overall efficiency; (2) voltage control accomplished within the converter; (3) achievement of line commutation, eliminating any forced external commutation circuitry.

The output of the single-phase, two-pulse midpoint converter (covered in Chapter 3) has a high ripple content. Increasing the number of pulses will reduce the ripple content in the load waveform. The pulse number can be increased by using converters as shown in Fig. 8-29 (page 196), arranged to form three dual converters. As shown in Fig. 8-29(*b*), this configuration uses 18 thyristors and permits only unidirectional operation. Another increase in the pulse number can be obtained by using six six-pulse midpoint converters with interphase reactors as shown in Fig. 8-30 (page 197). This unit requires 36 thyristors and also provides unidirectional operation only. Circulating currents are reduced by the interphase reactor, which presents its full reactance to the passage of circulating currents but only a quarter of its reactance to the load current.

The cycloconverters discussed up to this point are for continuous variable-frequency applications. In some applications the output frequency is a fixed percentage (less than 100) of the supply frequency. In these cases an *envelope cycloconverter*, a less complex circuit, may be employed. Figure 8-31 (page 197) shows a block diagram with logic control of the

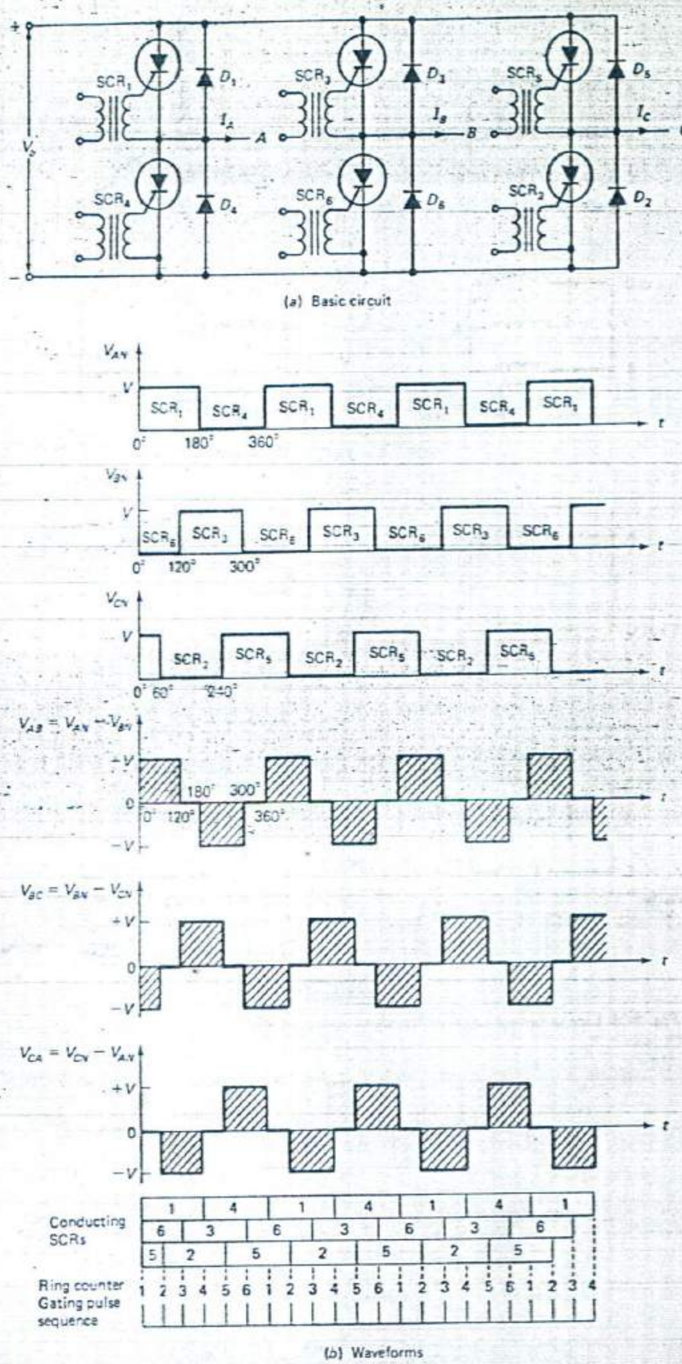


Fig. 8-24 Three-phase six-step inverter.

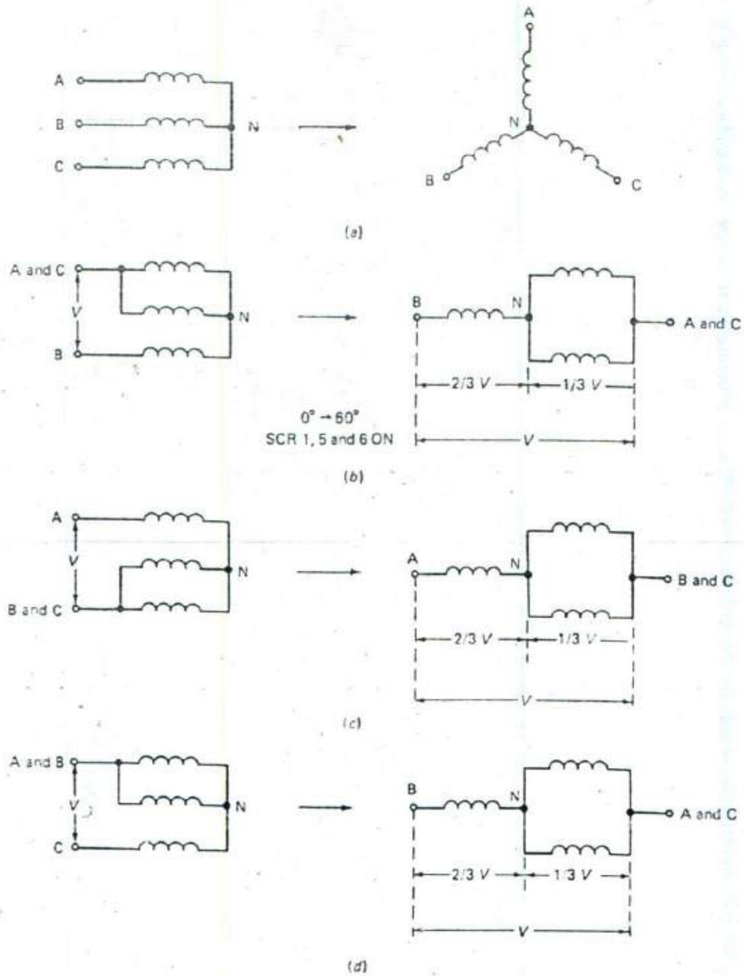


Fig. 8-25 Line to neutral voltages for wye-connected load.

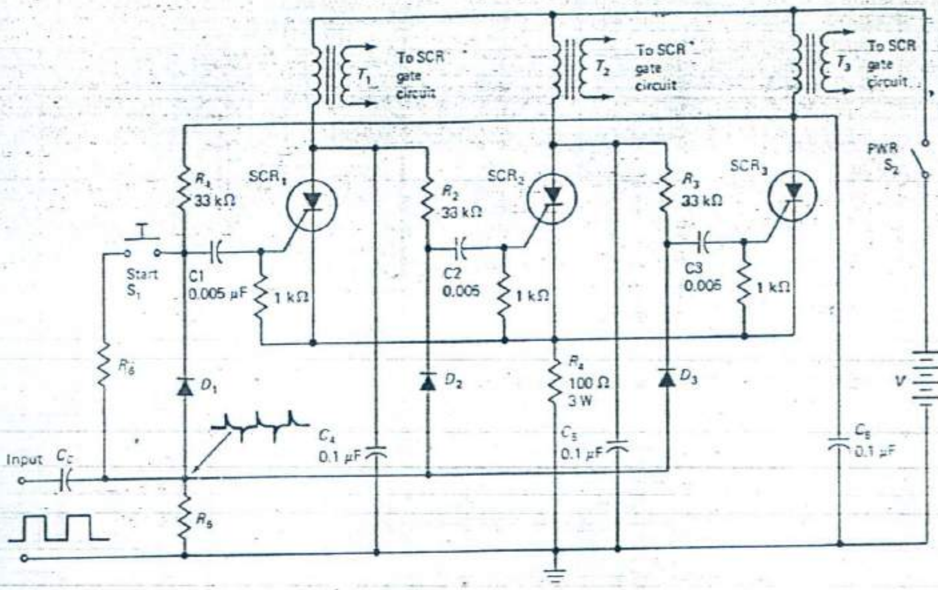


Fig. 8-26 Three-stage ring counter.

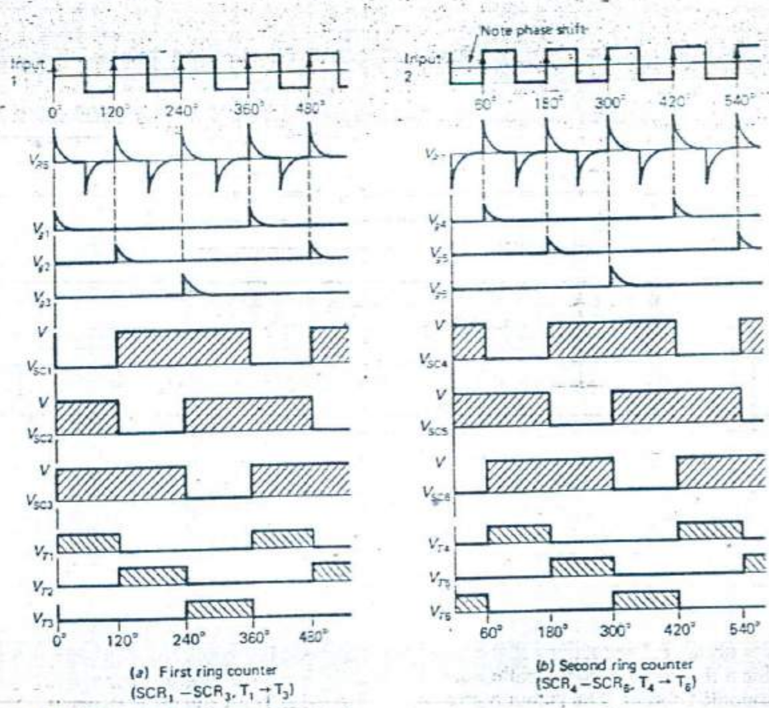


Fig. 8-27 Waveform from ring counter circuit.

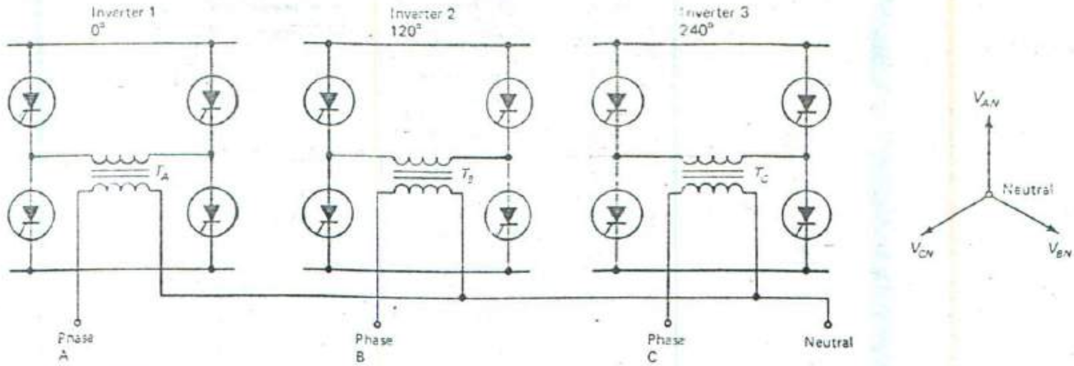


Fig. 8-28 Three single-phase inverters forming a six-step inverter.

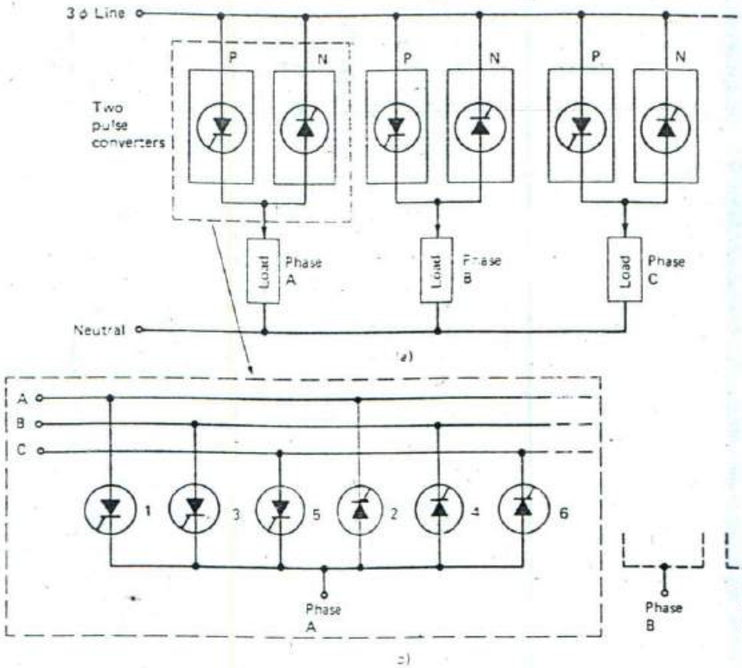


Fig. 8-29 Three two-pulse converters combined. (a) Block diagram. (b) Schematic diagram.

thyristor gates to obtain 6:1 reductions of the supply frequency. Using a three-phase configuration greatly reduces the harmonic content. The output of the six-pulse configuration is fed into a wye-double-wye transformer with four different output voltages. This arrangement produces a composite waveform that is a good approximation of a sine wave, as shown in Fig. 8-32. The waveform shows the 20-Hz envelope obtained from the 60-Hz supply. A basic disadvantage of the envelope cycloconverter is it operates only at a fixed number of frequencies when controlled by a counter with a variable number of states.

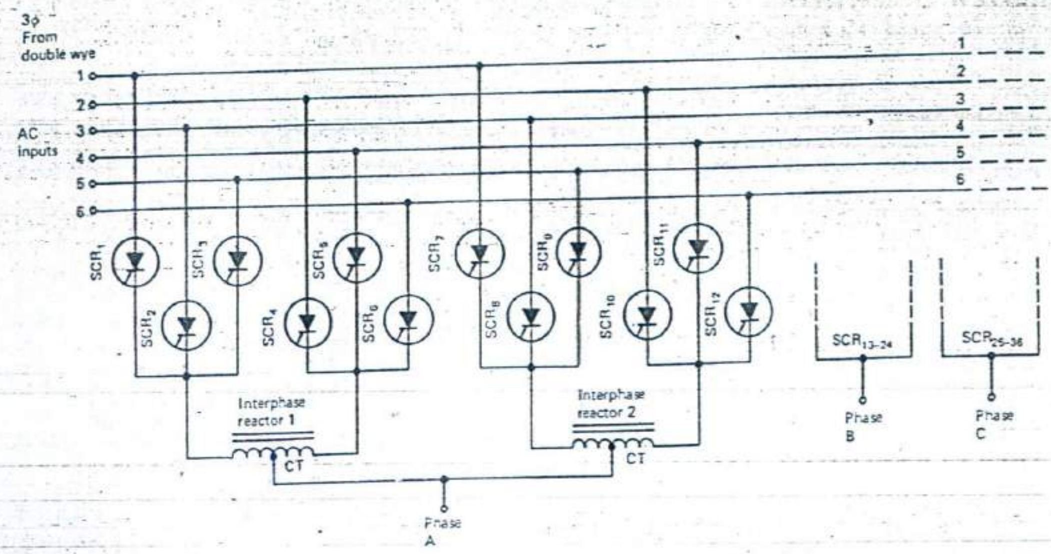


Fig. 8-30 Three-phase six-pulse cycloconverter.

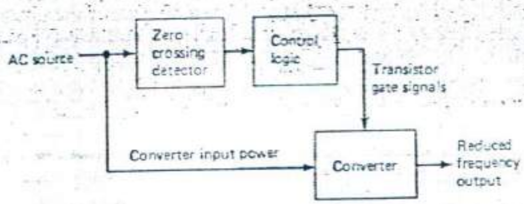


Fig. 8-31 Formation of envelope cycloconverter output.

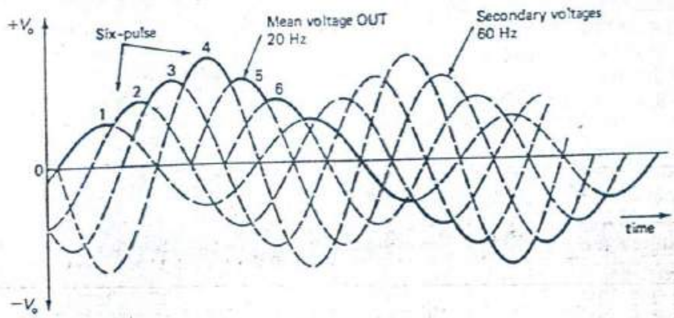


Fig. 8-32 Mean output voltages of synchronous envelope cycloconverter.

REVIEW QUESTIONS

19. The speed of induction motors is related to the line _______.

20. The _______ is used to change direct current to alternating current.

21. A cycloconverter converts a low supply frequency to a higher frequency. (true or false)

22. The cycloconverter changes frequency without any intermediate _______ link as required with an inverter circuitry.

23. Inverters typically operate from _______ Hz to 200 Hz.

24. In Fig. 8-24, shoot-through occurs if there is a delay in the _______ of the thyristor.

8-5 TROUBLESHOOTING AND MAINTENANCE

The devices used in the static converters covered in this chapter are also discussed in other chapters. One exception is the thyratron tube, whose appearance in modern industrial equipment is very unlikely. However, not all equipment is modern, and the technician must be aware that the tubes are gas-filled and exhibit a steady bluish glow when operating. In most cases, the tubes are balanced and share the load equally. If one tube is much brighter than another, a pair should be reversed to see whether the symptom follows the tube reversal. If so, the replacement of groups of tubes is recommended in this circumstance. If the problem does not follow the tube reversal, a circuit problem is indicated, so voltage and resistance checks are in order. Use the manufacturer's maintenance manual for reference. The tubes also have filaments which may be separately controlled or powered. The loss of filament power (usually less than 12 V_{ac}) will prevent tube conduction.

If possible, gather information from the operators of the equipment. This information may be useful in deciding whether a thermal problem, noise, regulation, or worn controls may exist. As with solid-state devices, heat dissipation is very important. Dirty or restricted fans and heat sinks should be checked first whenever any thermal problems exist.

Since phase-control circuits are line-synchronized, an oscilloscope triggered at the line frequency can be used for localizing any trouble. Silicon controlled rectifiers must hold off the line voltages and in most cases are in bridge-type configurations. This type of arrangement gives one a comparison device to use for checking the operation of all other devices. In the case of small motors, a power resistor or a lamp bank can be substituted for the motor to prevent overheating or damage due to latched-up conditions. If control is lost in one direction of a bidirectional control, the thyristor is usually suspect. An open device will prohibit movement in one direction, although a leaky or short-circuited device will keep the unit locked in that direction.

Most chopper circuits rely on self-commutation. In many circuits such as the Jones chopper, the capacitor is part of a resonant circuit needed for commutation, so its stability and leakage are important factors. If the capacitor is leaky, the SCR can latch up, and all control is lost. Any significant change in its value can shift the period of the resonant ringing and may not produce a voltage of sufficient amplitude to commutate the SCR. A leaky or short-circuited diode will disrupt all commutation. In the case of a freewheeling diode, a short-circuited device will stop the motor and severely overload the line and the control devices. A blown fuse or a tripped breaker can be expected in these cases. Parallel MOSFETs are employed in many applications and must be separated, at least at the drain or source lead, to check for short circuits. Remember that many devices have an integral (internal) diode that can influence the testing. In some cases, the devices can be first tested as a group. If the group test failed, the units would have to be separated to locate the one that was short-circuited. No commutation is needed for MOSFETs, so if latch-up occurs it usually indicates a short-circuited device.

Single-phase inverters must operate at their specified frequency, and any large deviation can cause a loss in efficiency. Overheating in the load can result from excess harmonic content if the inverter's filter fails to function properly. The ac current capability of any inductor or capacitor in a filter must be observed when replacements are in order. A set of manufacturer's spares is recommended in all cases. Inverters can usually be isolated into individual blocks for troubleshooting purposes. Isolation at the component level involves voltage or resistance checks. Various dc power supplies are employed for biasing and power conversion. Verification of all supply voltages must occur early in the fault analysis procedure.

Ring counters rely on the symmetry of parts (resistors, solid-state devices, capacitors, etc). Any large change in value or leakage will stop the counter or prohibit its starting. If it is stopped, one can detect the stuck stage and check the associated components. An oscilloscope can be used to trace trigger waveforms up to the gates of the thyristors or MOSFETs. If any SCR has failed, its shunt diode should also be checked. An open diode is often the cause of an SCR failure when inductive loads are involved.

In the case of large cycloconverters, the thyristors may be individually fused, and the fuses should be checked first. If a fuse is blown, the SCR should be verified before returning the unit to service. The logic used in developing the pulses can be tested to ensure that the timing and amplitude of the trigger pulses are within the manufacturer's specifications. Some units may have a test position and test points available for these purposes. In all cases, the manufacturer's maintenance manual should be consulted before any servicing is attempted.

REVIEW QUESTIONS

25. Thyratrons exhibit a ______ glow when ionized.

26. Before servicing open loop controls, a technician should obtain information from the equipment ______.

27. Latch-up in an SCR can be due to loss of commutation. (true or false)

28. Commutation in the Jones or Morgan circuit relies on energy stored in a ______.

29. Individual MOSFETs can be tested while connected in parallel in a circuit. (true or false)

30. If a capacitor opens in an inverter's filter, the load may ______.

CHAPTER REVIEW QUESTIONS

8-1. In Fig. 8-3, reversing the control voltage will reverse the motor's direction. (true or false)

8-2. In Fig. 8-4(*b*), which component senses motor current to provide overload protection?

8-3. Varying the OFF time of the waveform is called *pulse ______ modulation.*

8-4. Which solid-state switches are used in the Jones circuit?

8-5. The capacitor in the Jones circuit is usually an electrolytic type. (true or false)

8-6. Two-quadrant choppers are well suited for series-connected motors. (true or false)

8-7. When used with an alternating supply voltage, commutation is frequency-independent. (true or false)

8-8. To keep the output torque constant, an inverter must keep the ______ ratio constant.

8-9. Failure to keep the previous ratio constant may cause the motor to ______.

8-10. The two-on inverter refers to two phases being on at any time. (true or false)

8-11. The wye load voltage referenced to the neutral is ______ of the delta voltage for the two-on six-step inverter.

8-12. A method to gate inverters sequentially is to use a ______ counter.

8-13. The SCRs in a ring counter are commutated by positive pulses at their cathodes. (true or false)

8-14. Cycloconverters usually produce an output at ______ of the supply frequency or less.

8-15. Circulating currents are reduced in multiple cycloconverters by using a ______ reactor.

8-16. The ______ cycloconverter combines the output of multiple secondaries to produce the sine wave.

8-17. Thyristors in the cycloconverter may be individually ______ for overcurrent protection.

8-18. The ______ maintenance manuals should first be checked before attempting any servicing.

8-19. In most cases, the oscilloscope can be triggered from the line or ______ input of a cycloconverter.

8-20. Verification of all ______ voltages is an important early step in servicing any inverter.

ANSWERS TO REVIEW QUESTIONS

1. false 2. SCR 3. firing 4. R_2 5. Q_1; the UJT 6. deadband 7. thyristor 8. modulation 9. D_1 10. false 11. commutate 12. autotransformer 13. current 14. reverse 15. false 16. reverse/opposite 17. saturable 18. saturation 19. frequency 20. inverter 21. false 22. dc 23. 10 24. commutation 25. bluish 26. operator 27. true 28. capacitor 29. false 30. overheat

9

INPUT TRANSDUCERS

Control is based on information. Industrial automation systems must extract information from the physical process that is being controlled. Input transducers convert physical parameters into electrical signals that correspond to what is happening. In the broadest sense, a transducer *is any device that receives energy from one system and retransmits it, usually in another form, to another system. Thus, an electric motor can be viewed as a transducer. The word* sensor *is more restrictive. It refers to that part of a transducer that responds to the quantity being measured. This chapter will use the terms* transducer *and* sensor *to describe components and devices used to measure physical conditions.*

9-1 POSITION AND DISPLACEMENT

Displacement is the difference between the position of some object and a reference point. Displacement can be *linear* (straight-line) or *rotary* (angular). Potentiometric transducers can be used to measure both linear and angular displacement. Potentiometers are very common transducers in the industrial environment. Figure 9-1 shows a linear displacement potentiometer. A resistance element is shown at the top of the diagram. This element can be formed by winding resistance wire on a form or by depositing resistance material. The wiper contact moves along the resistance element in response to motion applied to the input shaft. If a voltage is applied across terminals A and B, then some portion of that voltage will appear across A and C. Most potentiometric transducers are nominally linear. If the input shaft is at its mechanical center position, half the applied voltage will appear at C. If it is at its far left position, the output voltage will be zero. If it is one-quarter from its far left position, the output voltage will be one-fourth the applied voltage. In other words, there is a linear relationship between the shaft position and the output signal.

The actual performance of a displacement potentiometer will deviate from nominal linearity. Manufacturers rate them according to worst-case deviation. A 1 percent linearity rating or better is typical for transducer service. Thus, a 100-Ω transducer may have an error of plus or minus 1 Ω at its worst-case position. Potentiometer resolution is another source of error. Suppose a 100-Ω resistance element is made up of 200 turns of wire. This means that each turn represents 0.5 Ω of resistance. As the wiper moves, the resistance across the wiper contact and either end contact changes in half-ohm steps. This is the smallest change that the transducer can resolve. It is also expressed in percentage form; the resolution would be 0.5 percent ($0.5\ \Omega/100\ \Omega \times 100$) in this example. For best accuracy, the percentages of linearity and resolution should be as small as possible.

An angular displacement potentiometer is shown in Fig. 9-2. The input shaft turns, and the wiper contact moves with it. As before, if a voltage is placed across the resistance element, the voltage at the wiper contact will be a function of shaft position or shaft displacement. The preceding discussion of linearity and resolution also applies. The angular potentiometer is usually limited to about 320° of rotation. However, gearing can be used to make a transducer with greater mechanical range. Other arrangements to provide greater mechanical range include worm drives for the wiping contact and assemblies with more than one wiper.

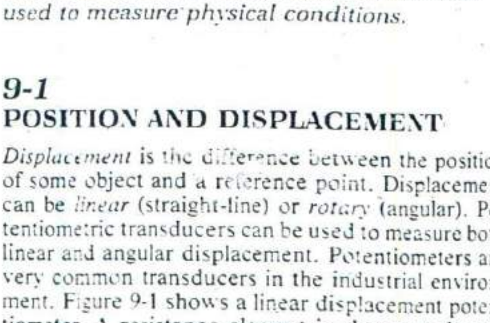


Fig. 9-1 Linear displacement potentiometer.

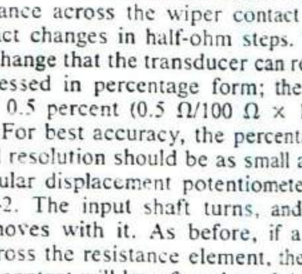


Fig. 9-2 Angular displacement potentiometer.

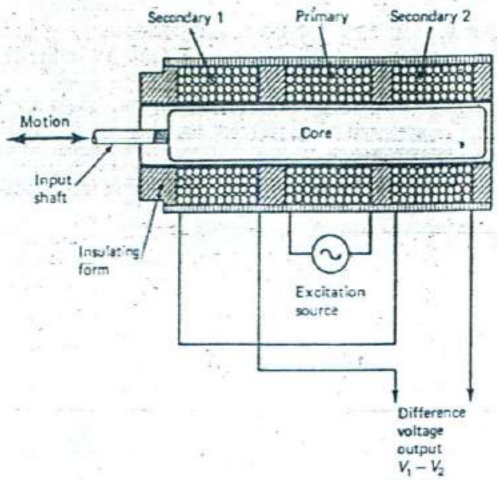


Fig. 9-3 Linear variable differential transformer (LVDT) construction.

Potentiometric transducers are relatively inexpensive and easy to apply. However, they have some limitations. For example, attempts to make the resolution very high usually result in poorer linearity. They are temperature-sensitive, a characteristic that also affects their accuracy. Potentiometers are considered to be low- to medium-accuracy transducers. The wiper contact is another limiting factor, being subject to wear and dirt and potentially producing electrical noise.

The *linear variable differential transformer* (LVDT) is more costly but outperforms the potentiometric transducer. Its construction is shown in Fig. 9-3. It consists of a primary, two secondaries, and a movable core. The primary is excited by an ac source. When the core is in its exact center location, the amplitude of the voltage induced into secondary 1 will be the same as the voltage induced into secondary 2. The secondaries are connected to phase cancel, and the output voltage will be 0 at that point. Figure 9-4 illustrates what happens to the output voltage as the core is moved to the left (Fig. 9-4[*a*]) and to the right (Fig. 9-4[*b*]). Note that the magnitude of the output voltage is a linear function of core position and that the phase is determined by the side of the null position on which the core is located.

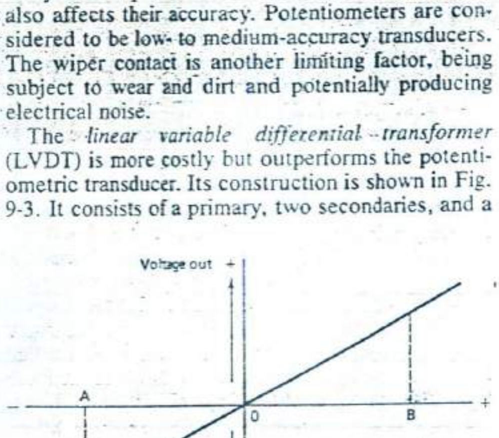


Fig. 9-4 Linear variable differential transformer (LVDT) output phase and voltage versus core position.

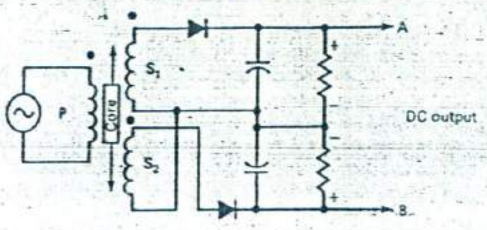


Fig. 9-5 Converting LVDT output to dc.

Figure 9-5 shows a simple circuit for converting the ac output of an LVDT to dc. With the core centered, both S_1 and S_2 produce equal amplitudes. Both half-wave rectifiers produce equal dc voltage drops across the two resistors. The polarities are opposing, so the output voltage from A to B is equal to zero. If the core is moved up, S_1 produces more voltage than S_2. The drop across the top resistor is now greater, and A becomes positive with respect to B. If the core is moved down, B becomes positive with respect to A. Although this simple arrangement works, higher performance is available with a more elaborate detector circuit. The last section of this chapter deals with transducer signal conditioning in more detail.

A *rotary variable differential transformer* is shown in Fig. 9-6. This transducer allows the measurement of angular displacement up to about 90°. This range may be extended with gearing.

Linear variable differential transformer accuracies are very good. Typical linearities are between 0.25 and 0.05 percent of full range. In general, even better linearity can be obtained by operating the LVDT over something less than its maximum range. The typical industrial LVDT has a total range of approximately plus and minus 2½ cm. Resolution is excellent, with typical specifications near 0.000001 cm.

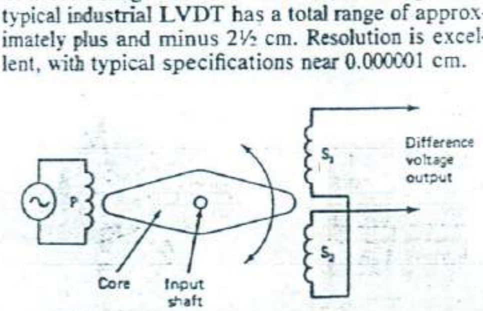


Fig. 9-6 Rotary variable differential transformer.

The excitation frequency for an LVDT varies, depending on its design and the application. For example, if the transducer must accurately track rapidly changing displacement, the higher frequencies are advantageous. Typical values range from 50 Hz to 30 kHz. The voltage applied to the primary is usually around 10 V. Reliability of the LVDT is superb, with ratings of millions of hours mean time before failure (MTBF).

Some displacement measurements involve very small movements. Strain gage transducers lend themselves to these applications. They also are noted for low cost and ease of use and, in some cases, can simply be epoxied to the physical member under measurement. Strain gages are based on the principle that the resistance of a conductor is directly proportional to its resistivity and length and inversely proportional to its cross-sectional area. If a conductor is stretched, its length will increase, thus increasing its resistance. At the same time, its cross-sectional area decreases, also increasing its resistance.

Figure 9-7(*a*) shows a bonded wire strain gage. It consists of several loops of fine wire bonded to a paper or plastic backing. Note that the sensitivity axis is parallel to the long portions of the wire runs. If the gage is slightly elongated along this axis, maximum resistance increase will result because the greatest total length of wire is involved. In Fig. 9-7(*b*) the foil-type gage is depicted. These gages are made with a printed-circuit-type process using conductive alloys rolled to a thin foil. A grid configuration is used for the strain-sensitive element to allow higher values of gage resistance while maintaining short gage lengths. Gage resistance varies from 30 to 3000 Ω with 120- and 350-Ω values being rather common. Gage lengths vary from 0.02 to 10 cm (0.008 to 4 in.). The sensitivity axis for the foil-type gage is parallel to the long runs of foil, and the alignment marks are used to accurately install the gage along the proper axis.

The typical industrial strain gage has a gage factor of 2. *Gage factor* is the resistance change ratio divided by the length change ratio:

$$\text{GF (gage factor)} = \frac{\Delta R/R}{\Delta L/L}$$

Suppose a 4-cm strain gage has a nominal resistance of 350 Ω and a gage factor of 2. How much will the resistance increase if the gage is elongated by 0.02 cm. First, find the ratio of length change by dividing 0.02 by 4. This yields 0.005. Then, multiply this ratio by the GF and the nominal resistance to find the change in resistance: $0.005 \times 2 \times 350\ \Omega = 3.5\ \Omega$. Thus, the gage will increase from 350 to 353.5 Ω when elongated by 0.02 cm.

It is obvious that the resistance change is small in strain gages. How can such a small change be converted into a useful signal? The Wheatstone bridge circuit is well known for its accuracy and sensitivity in measurement applications. Figure 9-8 shows a bridge with a strain gage serving as one of the bridge elements. If all four bridge elements equal 350 Ω, the bridge is balanced and the output voltage is zero. What happens to the output if the resistance of the gage increases to 353.5 Ω? By using the equation and input voltage shown in Fig. 9-8, you should calculate an output voltage of approximately 0.03 V. The bridge output will normally be applied to the input of an instrumentation amplifier with a gain of 100 (or more). The final output would be nearly 3 V for our example.

Another feature of the bridge circuit is that it can be accurately balanced to produce zero output for zero strain. This process is called *nulling* the bridge. In Fig. 9-8 R_2 could be replaced by a fixed resistor in series with an adjustable resistor. The adjustable resistor would allow a change of a few percentage points in that leg of the bridge. This would facilitate accurate nulling of the bridge. Finally, the bridge circuit makes temperature compensation easy. Strain

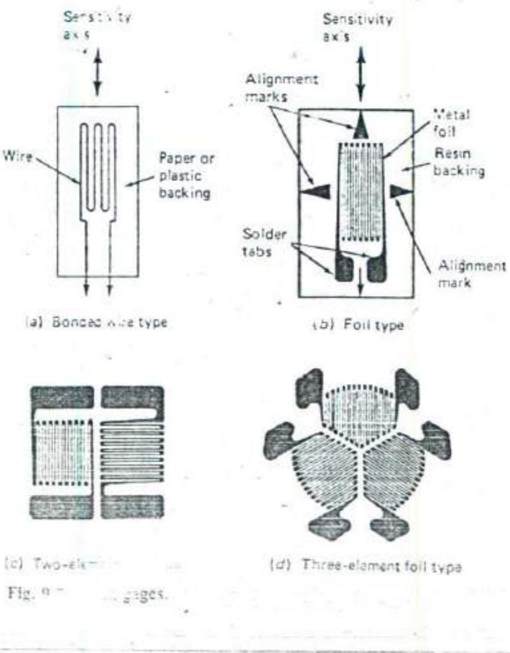
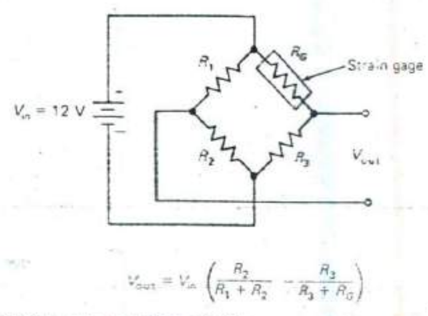


Fig. 9-[illegible] gages.



$$V_{out} = V_{in}\left(\frac{R_2}{R_1 + R_2} - \frac{R_3}{R_3 + R_G}\right)$$

Fig. 9-8 Wheatstone bridge circuit.

gages react to temperature as well as to strain. Refer again to Fig. 9-7(*c*). The two-element strain gage is often used where temperature compensation is required. With proper alignment, only one of the gages will react to strain. However, both will react to temperature and presumably by the same amount. If the two-element gage is applied to a bridge circuit such as the one shown in Fig. 9-8, one will act as a strain sensor, and the other will provide temperature compensation. For example, the gage sensitive to strain would serve as R_G, and the compensation gage would serve as R_1. The bridge equation shows that any resistance change that affects R_G and R_1 by an equal amount will produce no change in V_{out}.

Temperature compensation with a two-element gage does create some error. All gages have some sensitivity to strain perpendicular to their longitudinal axis. This is called *transverse sensitivity*. It is minimized by the gage manufacturer by placing extra material in the end loops of the conductors and by keeping the grid lines close together. The manufacturer will specify the *transverse sensitivity factor*, which is the ratio of transverse GF to longitudinal GF. The gain of the instrumentation amplifier can be adjusted to compensate for the characteristic that the temperature compensation gage also increases slightly in resistance when the assembly is elongated.

Certain semiconductor materials exhibit a characteristic known as *piezoresistance*, which is a change in resistance with strain. Semiconductor strain gages capitalize on this effect to provide very sensitive strain transducers. They have gage factors higher than those of the bonded wire or foil types. The semiconductor gage factors range from 45 to 175. This high sensitivity permits them to be used without amplifiers in some applications. Their resistance change is less linear over large ranges, however.

Figure 9-9 shows another type of displacement transducer. Capacitance is directly related to plate area and inversely related to the plate spacing. As the metal tube moves to the right, the distance between plates decreases, increasing the capacitance. The capacitive transducer can be placed into an ac bridge to provide an ac output voltage that is a function of linear displacement. Or the capacitor can be part of a tuned circuit for an oscillator. This arrangement will produce a frequency change with any change in position. Figure 9-10 shows an angular displacement capacitor. Capacitance will increase as the moving plate covers more of the fixed plate.

None of the transducers shown in this section lend themselves to the measurement of very large displacements. There are various ways to accomplish this, and Fig. 9-11 shows one example. As the material moves beneath the friction wheel, it causes it to turn. The wheel is shaft-coupled to a toothed wheel made of ferrous material. As the toothed wheel turns, it alternately provides a high and then a low reluctance path for the Hall-effect sensor. The output of the sensor is in the form of pulses. These pulses can be accumulated by counting circuits to measure extremely large displacements.

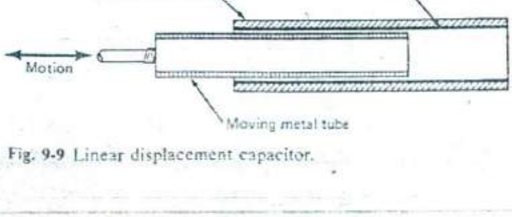


Fig. 9-9 Linear displacement capacitor.

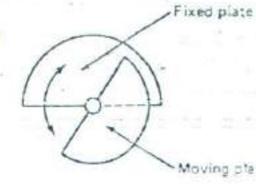


Fig. 9-10 Angular displacement capacitor.

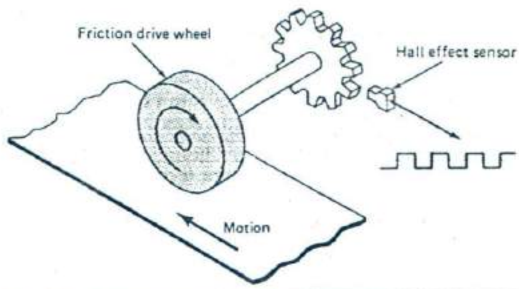


Fig. 9-11 A method of sensing large displacements.

REVIEW QUESTIONS

1. Refer to Fig. 9-1. Assume that 12 V is across terminals A and B and that the slide bar is positioned at three-fourths of its right-most displacement. What voltage drop should appear from A to C?

2. Assume that the transducer in question 1 has a linearity of 1 percent. What range of output voltages would be considered normal for the data given?

3. A potentiometric transducer has a total resistance of 500 Ω, and the smallest resistance change it can produce is 1 Ω. What is its percentage of resolution?

4. Refer to Fig. 9-5. Assume that the core is 1 cm above its null position and that the drop across the top resistor is 6 V and the drop across the bottom resistor 4 V. What is the output voltage? Is A positive or negative with respect to B?

5. Assuming perfect linearity, what would happen in question 4 if the core were moved to a position 1 cm below the null position?

6. Refer to Fig. 9-8. All bridge elements are equal at 350 Ω. How much output will be produced if R_1 and R_G both increase by 1 Ω?

7. Refer to Fig. 9-7(*d*). How many sensitive axes are there?

9-2 VELOCITY AND ACCELERATION

Speed is the rate of change in displacement. It may be measured in meters per second, kilometers per hour, or centimeters per minute. *Velocity* is a measure of speed and direction. It is a vector quantity. Industrial measurements most often deal with fixed directions. A ram on a machine, for example, can only travel back and forth along a fixed path. Since the direction is established, the term *velocity* rather than speed is used to describe the rate of ram travel. One ram direction would yield a positive velocity and the opposite direction a negative velocity. Rotating machine parts also usually travel in a fixed path. Again, one direction would be measured in terms of positive velocity and the other in negative velocity. *Angular velocity* is the rate of change in angular displacement. The most common unit of angular displacement is the revolution (360°) and angular velocity is usually measured in revolutions per minute (rpm). It may also be measured in radians per second (rad/s). There are 6.28 ($2 \times \pi$) radians in 1 revolution. Linear velocity is often converted to angular velocity by an arrangement such as that shown in Fig. 9-11. Velocity information can be extracted from this arrangement by timing the sensor pulses.

High velocities are ordinarily measured by timing the period required for an object to travel from one fixed point to another. Optical detectors and timing circuits are used to make the measurement. For example, an object passes one point, where it interrupts a light beam and the timing circuit starts. Later, a second point is passed, where a second light beam is interrupted and the timing circuit stops. Very high accuracy can be achieved with such an arrangement if the distance between the two points is accurately measured and if the timing circuit is precise. Modern counting circuits can easily resolve millionths of a second. Angular velocity can also be measured by optical techniques. Figure 9-12 shows an optical tachometer that can be used to measure shaft, gear, or pulley velocity. A contrasting stripe of paint or tape is applied to the rotating part. The tachometer has a light source and lens assembly. The light reflected back into the photo detector is alternately brighter and darker because of the contrasting stripe. The detector produces one pulse for every shaft revolution. Counting and timing circuits convert the pulses to angular velocity.

Direct current tachometers are also popular for measuring angular velocity. They are usually equipped with a permanent magnet field, a rotating armature circuit, and a commutator and brush assembly. Ignoring loading effects, the output of a dc tachometer is proportional to the flux density of the field, the length of the armature circuit, and the angular velocity of the shaft. Since flux and the armature circuit are fixed for any given tachometer, the manufacturers rate them with an output constant of volts per rpm. Typical dc tachometers produce outputs from 0.01 to 0.02 V/rpm. They produce a polarity reversal when rotated backward. This is a necessary feature when the tachometer must also provide direction information to the system.

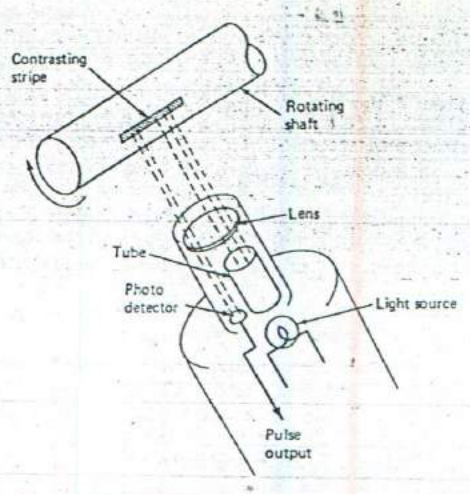


Fig. 9-12 Optical tachometer.

Direct current tachometers, because they have brushes, suffer from noise and maintenance problems. Alternating current tachometers have been developed to eliminate these problems. They employ a permanent magnet rotor and a polyphase stator. The output is ac, with both frequency and voltage proportional to the angular velocity of the input shaft. A frequency-to-voltage converter circuit is sometimes used to convert the output to a dc signal. Or the stator output can be rectified to produce a dc signal. Neither technique will produce a polarity change if the input shaft changes direction, however.

The *drag-cup tachometer* works on an induction principle. It is also called an *ac induction tachometer*. Figure 9-13 shows its construction. One set of stator coils is connected to an external excitation source. The other stator coils are positioned 90° from the excited coils and form the output circuit. The drag cup is made of copper or aluminum and is connected to the input shaft. The side view shows that there is also a laminated inner core. The drag cup has a clearance gap between this core and the stator poles. The theory of operation is shown in Fig. 9-14. With no rotation, the output is zero. An eddy current flux is set up in the drag cup, but it is at 90° to the output coils. If the shaft is turned, a second flux appears in the drag cup. It is caused by an armature reaction current flowing in the drag cup and is at

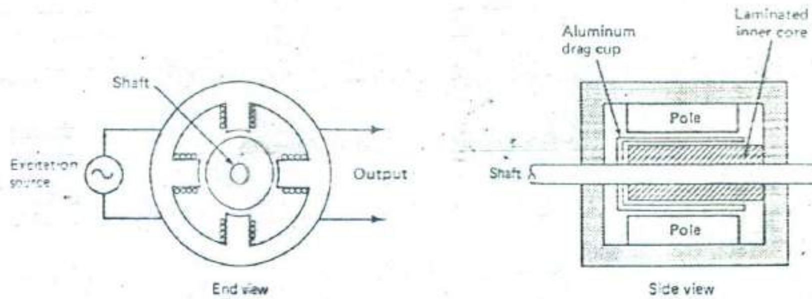


Fig. 9-13 Drag-cup tachometer.

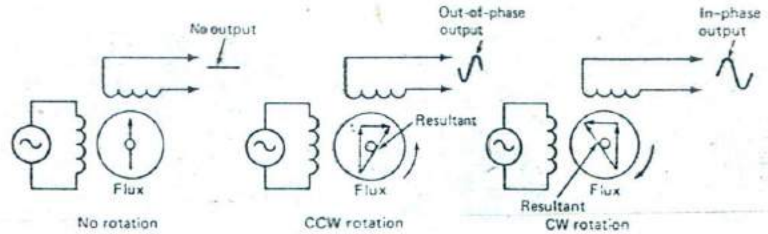


Fig. 9-14 Drag-cup tachometer operation.

right angles to the exciter coils. The resultant flux produces an output. The faster the shaft turns, the greater the reaction flux and the better the angle of the resultant for output. Therefore, the output is proportional to shaft velocity. If the shaft is reversed, the reaction flux reverses, the resultant angle shifts, and the phase of the output changes. With the appropriate phase-detector circuitry, the drag-cup tachometer can provide shaft direction information.

A linear tachometer produces an output that varies smoothly and continuously with angular velocity. For example, if the output is 10 V at 1000 rpm, it will ideally be 10.5 V at 1050 rpm and 11 V at 1100 rpm. Digital tachometers produce a fixed number of output pulses for every shaft revolution. The optical tachometer presented earlier produced one pulse per revolution. In Fig. 9-15 three pulses are produced for every revolution because there are three slots in the rotating disk. It is possible to sense direction with this type of tachometer by using two sets of staggered, overlapping slots and a dual LED/phototransistor assembly. The overlap provides a point where both beams will be on. Then one beam will be interrupted first, depending on the direction of rotation. Magnetic digital tachometers are also popular. These operate on the Hall-effect principle or the variable-reluctance principle.

Acceleration is the rate of change in velocity. An object that is accelerating is increasing its velocity with time. In *deceleration*, or negative acceleration, the object is losing velocity with time. It is measured in units of displacement per time per time. For example, gravity will cause a falling object to accelerate at 981 cm/s/s. This is often written as 981 cm/s^2 and is known as the gravitational constant g. A signal proportional to acceleration (or deceleration) can be obtained by differentiating the output of a velocity transducer. Likewise, the output of an accelerometer can be integrated to provide velocity information. Differentiators and integrators were covered in Chapter 6. Or acceleration can be measured by another indirect means. Newton's law gives us $F = ma$ (force equals mass times acceleration). If the mass is known, it is possible to measure a displacement produced by the force and derive the acceleration. Fig. 9-16 shows the basic *accelerometer,* in

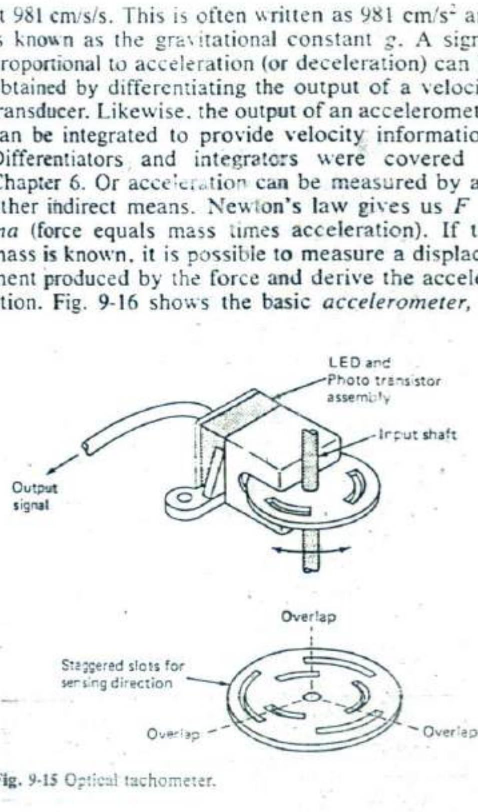


Fig. 9-15 Optical tachometer.

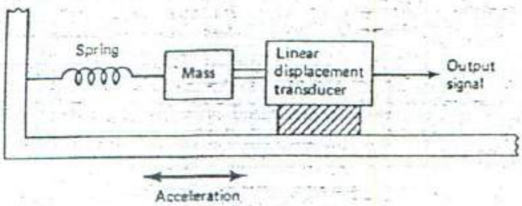


Fig. 9-16 Basic accelerometer.

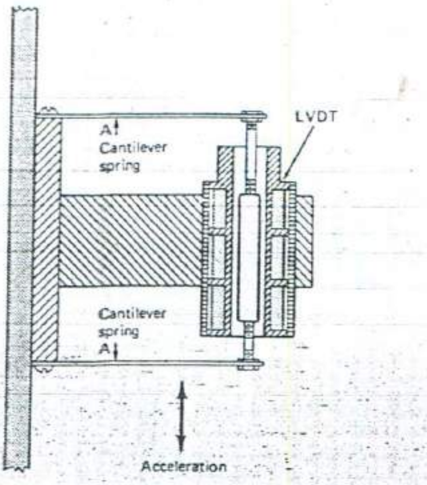


Fig. 9-17 Linear variable differential transformer (LVDT) accelerometer.

which a linear displacement transducer is mechanically coupled to a spring-loaded mass. Acceleration along the axis shown will produce a reaction force on the mass. Spring tension will allow more or less displacement of the mass, depending on the amount of force created by the acceleration. The output signal from the displacement transducer will be proportional to acceleration.

An LVDT accelerometer is shown in Fig. 9-17. The LVDT core and spring assembly form the mass. The cantilever springs resist core motion. Core displacement will be proportional to acceleration. Deceleration will cause an opposite core displacement and can therefore also be measured. If an accelerometer such as the one in Fig. 9-17 is set into motion and then suddenly stopped, it will oscillate at its natural resonant frequency. In fact, any transient mechanical input will cause oscillations in a spring-mass system. The oscillations are typically around 60 Hz for an LVDT-type accelerometer.

Figure 9-18 shows the damped sine wave produced by a spring-mass system. The natural resonant frequency is the reciprocal of one period of oscillation. The system continues to oscillate, but friction eventually brings it to a rest. This friction is called the *damping coefficient*. The greater the damping coefficient, the more quickly the system will recover from transients. Some accelerometers use a viscous material such as oil to increase the damping coefficient. However, too much damping will make the response sluggish, and the transducer will not accurately track rapid changes in acceleration. Resonant frequency and the damping coefficient are very important specifications in an accelerometer. It has been found that when acceleration is changing with time (which is normally the case) an accurate response cannot be obtained from an accelerometer unless its resonant frequency is at least several times the frequency of the acceleration change. It is also important that the mass of the accelerometer not appreciably change the mechanical response of the system to which it is connected.

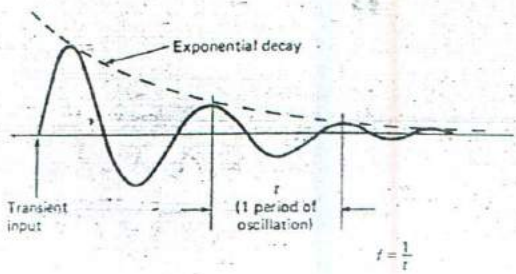


Fig. 9-18 Damped sine wave.

Figure 9-19 shows a piezoelectric accelerometer that lends itself to the very rapid accelerations involved in measuring shocks and vibrations. These

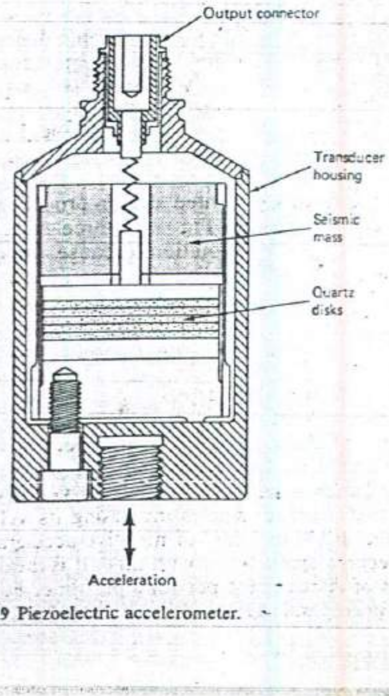


Fig. 9-19 Piezoelectric accelerometer.

transducers have natural resonant frequencies up to 40 kHz or so. The surfaces of piezoelectric materials, such as quartz, become charged when the materials are mechanically stressed. The seismic mass is supported by the transducer housing through the quartz disks. Preloading allows the accelerometer to respond to both negative and positive acceleration. The mass can be made high for large-transducer sensitivity or low for high-resonant frequency. A typical high-sensitivity unit will have a range of plus or minus 100*g* (*g* is the earth's gravitational constant given earlier) and a resonant frequency of 2 kHz. A low-sensitivity unit has a typical range of $-20{,}000g$ to $+50{,}000g$ with a resonant frequency of 40 kHz.

Angular acceleration is measured in revolutions per time per time (rev/t^2) or radians per time per time (rad/t^2). Angular acceleration is measured indirectly with an angular displacement transducer. The transducer output goes to a computer, where timing and position information can be combined to calculate acceleration and deceleration. Sophisticated machine tools and robots achieve very precise control of motion and position by using these techniques.

REVIEW QUESTIONS

8. What happens to the output of a dc tachometer when its shaft is turned in the opposite direction?

9. Why might it be said that ac tachometers are angular speed indicators rather than angular velocity indicators?

10. What is the advantage of the drag-cup tachometer over the ac tachometer?

11. When the output of a velocity transducer is differentiated, a signal proportional to ______ is the result.

12. The basic accelerometer in Fig. 9-16 measures ______ directly. The calibrated spring allows this to be interpreted as a ______ measurement, and the calibrated mass allows the output to be expressed in units of ______.

13. Refer to Fig. 9-18. The period of one oscillation cycle is 1.5 ms. What is the resonant frequency?

14. What would a larger damping coefficient achieve in Fig. 9-18?

9-3 FORCE AND FLOW

Force is measured in newtons (N); 1 N is equal to 0.225 lb. Small forces may be measured in dynes (dyn); 1 N is equal to 100,000 dyn. Force transducers are often based on displacement principles. For example, refer to Fig. 9-20, which shows a force-measuring device based on a compression spring and an

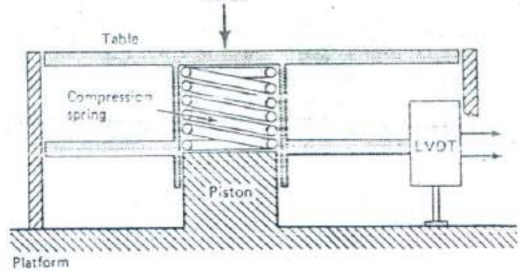
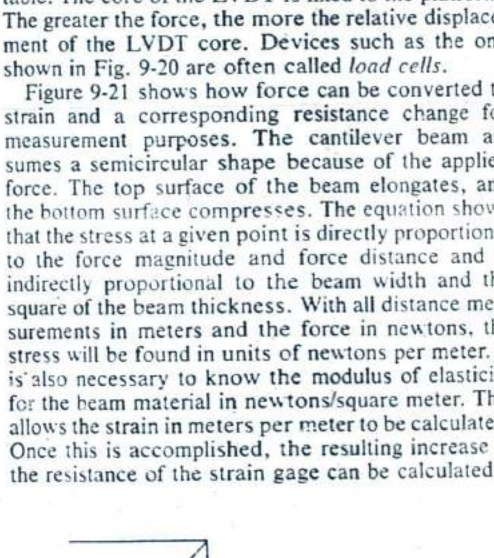


Fig. 9-20 Measuring force.

LVDT. When a force is applied to the table, the spring compresses in proportion to the amount of force. The body of the LVDT moves down with the table. The core of the LVDT is fixed to the platform. The greater the force, the more the relative displacement of the LVDT core. Devices such as the one shown in Fig. 9-20 are often called *load cells*.

Figure 9-21 shows how force can be converted to strain and a corresponding resistance change for measurement purposes. The cantilever beam assumes a semicircular shape because of the applied force. The top surface of the beam elongates, and the bottom surface compresses. The equation shows that the stress at a given point is directly proportional to the force magnitude and force distance and is indirectly proportional to the beam width and the square of the beam thickness. With all distance measurements in meters and the force in newtons, the stress will be found in units of newtons per meter. It is also necessary to know the modulus of elasticity for the beam material in newtons/square meter. This allows the strain in meters per meter to be calculated. Once this is accomplished, the resulting increase in the resistance of the strain gage can be calculated.



$$S = \frac{6FD}{WT^2}$$

S = Stress in newtons/meter

$$\text{Strain} = \frac{\text{Stress}}{\text{Modulus of elasticity}}$$

Fig. 9-21 Converting force to strain.

EXAMPLE

A beam is 2 cm wide and 0.25 cm thick. With a force of 100 N, calculate the stress on the beam at a point 10 cm from the point where the force is applied.

SOLUTION

$$S = \frac{6\,FD}{WT^2}$$

where S = stress
F = force
D = distance between strain gage and point of application of force
W = width
T = thickness

$$= \frac{6(100)(0.10)}{(0.02)(0.0025)^2}$$
$$= 4.8 \times 10^8 \text{ N/m}$$

If the beam is made of steel with a modulus of elasticity of 2×10^{11} N/m², the strain can be found by

$$\text{Strain} = \frac{S}{\text{modulus of elasticity}}$$
$$\text{Strain} = \frac{4.8 \times 10^8 \text{ N/m}}{2 \times 10^{11} \text{ N/m}^2}$$
$$= 2.4 \times 10^{-3} \text{ m/m}$$

Now, if we assume that a 120-Ω strain gage with a gage factor of 2 is mounted at the strain point, the increase in resistance due to the 100-N force may be found by

$$\Delta R = \text{GF} \times \text{strain} \times \text{resistance}$$
$$= 2 \times 2.4 \times 10^{-3} \times 120$$
$$= 0.576\ \Omega$$

A tension load cell is illustrated in Fig. 9-22. It can be used to measure the force required to pick up heavy loads in industry. The metal proving ring will elongate along the tension axis. Four strain gages are placed around the proving ring to sense the changes. The four gages are electrically arranged in the bridge configuration. Tension will slightly distort the proving ring and elongate gages R_2 and R_4, while R_1 and R_3 will be slightly compressed. The bridge is arranged to produce the maximum change in output voltage for these changes. Temperature effects are minimized by the bridge arrangement since all four gages tend to track thermally.

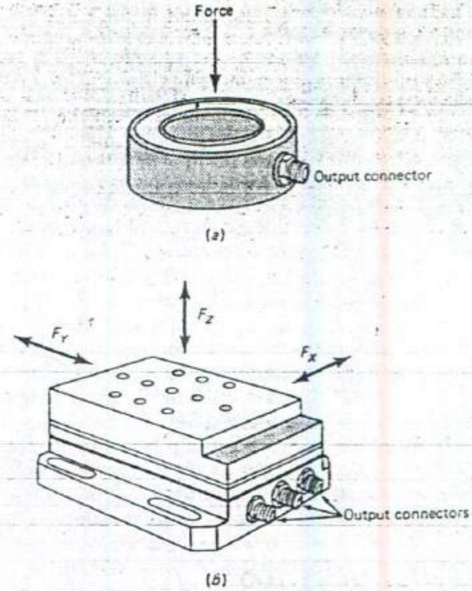


Fig. 9-23 Piezoelectric force transducers. *(a)* Load washer. *(b)* Three-component dynamometer.

Two examples of piezoelectric force transducers are shown in Fig. 9-23. The load washer type shown in Fig. 9-23(*a*) is designed to measure axial forces. It is preloaded when manufactured and can measure both tensile and compressive forces. Load washer force transducers are available with ratings from 7 kN to 1 MN. The three-component dynamometer type shown in Fig. 9-23(*b*) measures three orthogonal (right-angle) components of force. It has a natural resonant frequency of about 4 kHz and measures a

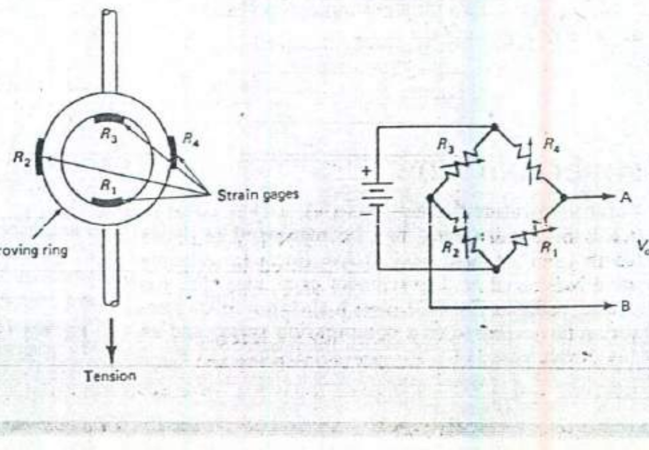


Fig. 9-22 Tension load cell.

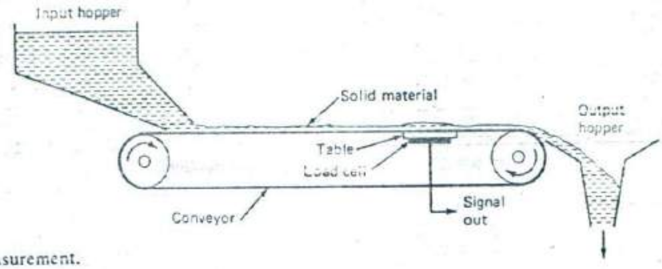


Fig. 9-24 Solid flow measurement.

maximum force of plus or minus 5 kN along any of the three axes. Typically, the work piece is mounted on the dynamometer to provide signals proportional to the forces involved in grinding and milling operations. Or the tool can be mounted on the dynamometer to measure the cutting forces produced in turning operations. Piezoelectric transducers do not have a dc frequency response and are not useful for monitoring steady-state conditions.

Force transducers can be applied in the measurement of solid flow. The units may be in kilograms/minute or in some other form; Fig. 9-24 shows one arrangement. A load cell is placed under a conveyor table to measure the force exerted by the flowing material. Another transducer, not shown, provides a signal proportional to conveyor velocity. A circuit or a computer is used to multiply the two transducer signals because the flow is directly proportional to the product of the weight and the velocity. Another technique that is used is to measure the sag of the conveyer with a displacement transducer such as an LVDT. The sag will be directly related to the weight of the material being carried. Again, when these data are combined with the velocity, flow is the result.

Fluid (liquid and gas) flow is measured in other ways. One popular way is to use some type of a flow restriction to create a differential pressure. *Bernoulli's principle* states that as the velocity of a fluid increases, its pressure decreases. Likewise, as the velocity of a fluid decreases, its pressure increases. When a fluid flows through a restriction, its velocity must increase and its pressure must drop. This is the principle behind many flowmeters. Figure 9-25 shows a venturi differential pressure flowmeter. The *venturi* is that part of the pipe where the passage necks down. This reduced area forces the fluid to increase in velocity through the venturi. Note that two pressure tubes connect bellows to two points: P_1 and P_2. Point P_1 is at a point before the venturi, and P_2 is in the venturi. When fluid is flowing, P_1 will be greater than P_2, and the pressure difference is proportional to the flow. The bellows extend in proportion to pressure. When P_1 is greater than P_2, the core in the LVDT will move to the right in Fig. 9-25. If the flow is increased, the core will move farther to the right. If the flow stops, there will be no Bernoulli effect and the core will be centered.

Some alternatives to the venturi tube are illustrated in Fig. 9-26. The orifice plate provides the necessary restriction to flow. This construction is less expensive than the venturi type. However, it sets up flow turbulence in the pipe and is less accurate. The nozzle type is a compromise between the venturi-tube and orifice-plate flowmeters. It produces less turbulence and subsequent pressure loss than the orifice type but is not quite so efficient as the venturi tube. The Dall tube has the least turbulence and insertion loss. However, it cannot be used

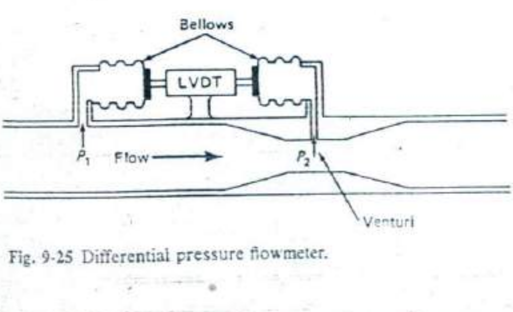


Fig. 9-25 Differential pressure flowmeter.

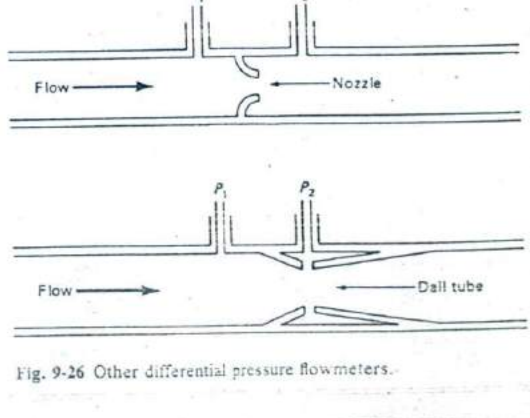


Fig. 9-26 Other differential pressure flowmeters.

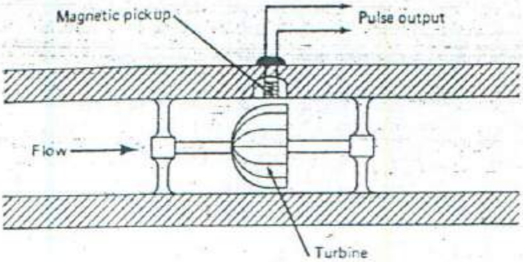


Fig. 9-27 Turbine flow measurement

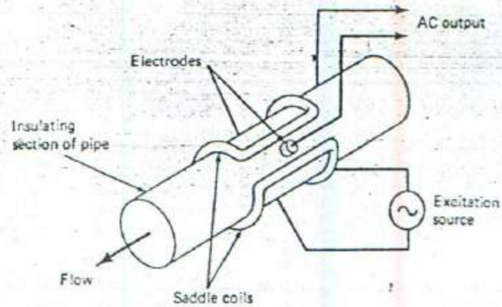


Fig. 9-29 Magnetic flowmeter.

with slurries, as the venturi tube can. A *slurry* is a suspension of solid particles in a liquid carrier such as water.

The turbine flowmeter shown in Fig. 9-27 is a very linear and accurate flow transducer. The fluid turns the turbine as it moves past the blades. The faster the flow, the greater the speed of turbine rotation. The blades of the turbine are magnetized, and pulses are induced in the magnetic pick-up coil located in the wall of the tube. These transducers are expensive and cannot be used with slurries.

Positive displacement flowmeters divide the flow into measured units and are noted for good accuracy. The *nutating disk flowmeter* is shown in Fig. 9-28. Fluid enters the left side of the housing. It then works its way into the left chamber of the inner housing, where it applies pressure to the disk. The disk wobbles (nutates) and releases measured amounts of flow into the right chamber and on to the exit pipe. As the disk nutates, it turns the top shaft and drives a pulse generator located in the top housing. The pulse rate is proportional to flow.

All of the flowmeters presented to this point offer some restriction to flow. Only one of them, the venturi type, lends itself to measuring slurry flow. Figure 9-29 shows a magnetic flowmeter that can be used with electrically conducting fluids and slurries and offers no restriction to flow. It is based on the principle of voltage induction in a conductor that is moving through a magnetic field. The faster the flow, the greater the induced voltage at the electrodes. The saddle coils provide an ac field. Direct current field flowmeters are also used but tend to cause polarization at the electrodes when used with fluids with low conductivity. Alternating current excitation also has limitations caused by dielectric losses and direct pick-up at the electrodes. In either case, the output signals are usually in the microvolt range and require quite a bit of amplification. Magnetic flowmeters are typically the most expensive to apply.

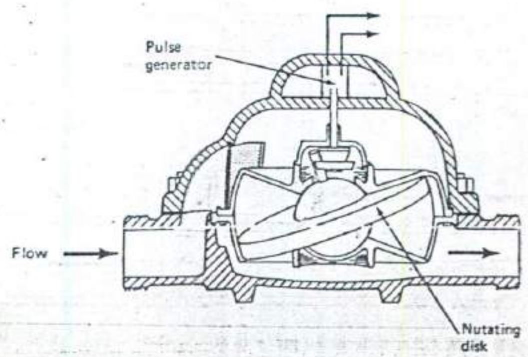


Fig. 9-28 Nutating disk flowmeter.

Other flowmeters use moving vanes or metering floats to respond to flow. They may only provide a visual indication or may be coupled to a potentiometer or an LVDT. Phototransistors have also been used to provide an electrical interface. Finally, temperature-sensing devices have also been used to measure flow. If the fluid is colder than a sensor, then the flow will remove heat from the sensor. The section on temperature transducers includes a simple circuit that detects air flow in this manner.

REVIEW QUESTIONS

15. Refer to Fig. 9-21. For what purpose might a second, identical gage be mounted next to the first gage, but with its axis of sensitivity arranged perpendicular to the first gage?

16. Refer to Fig. 9-21. Suppose the beam is 3 cm wide and 0.15 cm thick, and that a force of 5 N is applied 8 cm from the gage. What is the stress at the gage?

17. Refer to question 16. Assume that the beam is made from aluminum with a modulus of elasticity of 6×10^{10} N/m^2. What is the strain at the gage?

18. Refer to questions 16 and 17. The gage factor is 2, and the gage resistance is 120 Ω. Calculate the increase in resistance due to the 5-N force.

19. Refer to Fig. 9-22. Assume a null bridge with zero tension. What will the polarity of A with reference to B be when the proving ring is loaded?

20. What other type of transducer or information is required in Fig. 9-24 to calculate flow?

21. Refer to Fig. 9-25. The direction of flow is as shown. Which pressure is greater?

9-4 PRESSURE AND LEVEL

Pressure is defined as force per unit area. Unfortunately, there are many ways to measure pressure, and many units have evolved. The international system of units has come to the rescue and established the *pascal* (Pa) as the standard unit of pressure; 1 Pa of pressure is defined as a force of 1 N applied over an area of 1 m^2. However, standards are sometimes adopted slowly in industry, and the technician who works with pressure transducers will find several measurement units with which to cope. Industrial pressures are often measured in pounds per square inch (psi). Pascals can be obtained by multiplying psi by 6.8948×10^3.

There are three different reference conditions for pressure measurements. Gage pressure is referenced to atmospheric pressure, which is 14.70 psi (101 kPa) at sea level. Gage pressure will change with altitude. Absolute pressure is referenced to a perfect vacuum and does not change with altitude. Gage pressure can be obtained by subtracting the ambient atmospheric value from the absolute value. Likewise, absolute pressure can be found by adding the ambient pressure to the gage pressure. Differential pressure is referenced to an arbitrary value. A *g*, *a*, or *d* may be suffixed to a pressure measurement to clearly denote gage pressure, absolute pressure, or differential pressure.

The most common way to measure pressure is to use a force-summing device to convert the pressure into a displacement. Any of the displacement transducers already discussed can then provide the output signal. At very high pressures, the force-summing device may be eliminated and the pressure directly applied to a sensor based on the piezoresistive or piezoelectric effects. Figure 9-30 shows some examples of Bourdon tubes that are used to convert pressure into a proportional change in displacement. In Fig. 9-30(*a*) the cross section of a Bourdon tube is shown. Bourdon tubes are made of metal, such as steel, phosphor bronze, or brass. Figure 9-30(*b*) and 9-30(*c*) shows the circular and spiral shapes often used for these tubes. When pressure is applied, the tube tends to straighten, and a displacement results. The helical Bourdon tube shown in Fig. 9-30(*d*) produces a rotary motion when pressure is applied. Figure 9-31 shows a circular Bourdon tube coupled to an LVDT. When pressure is applied, some of the spring tension is overcome, and the core moves up.

Other force-summing devices used in pressure transducers include the bellows type and the diaphragm type. The *bellows type*, shown in Fig. 9-32(*a*), is a pressure cylinder with a thin corrugated metal wall. The bellows expand with pressure, producing a displacement to the right. Bellows can be made more sensitive than Bourdon tubes and are used at lower pressures. An auxiliary spring can be added to allow the bellows to be used at higher pressures. A pair of bellows makes a good arrangement for differential pressure measurements. (This is illustrated in Fig. 9-25.) A diaphragm force-summing device is shown in Fig. 9-32(*b*). The *diaphragm* is a flexible plate and can be made from rubber, neoprene, metal, or corrugated metal. A pair of dia-

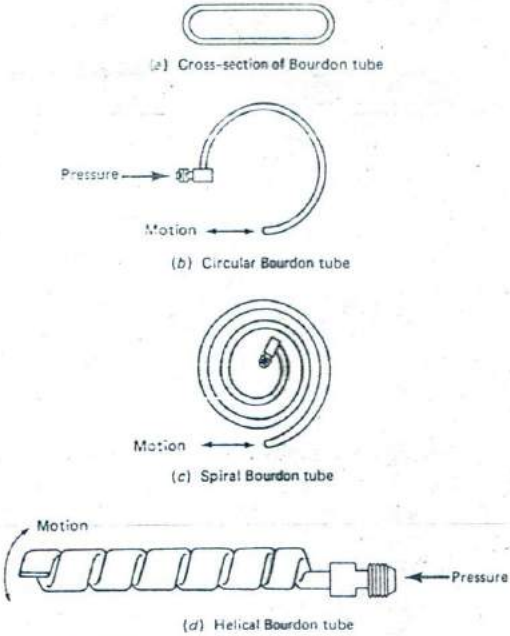


Fig. 9-30 Bourdon tubes.

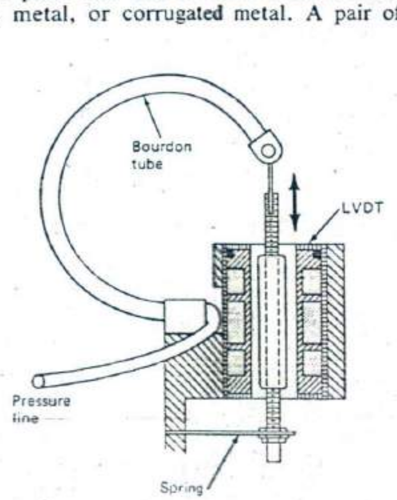


Fig. 9-31 Linear variable differential transformer (LVDT) pressure transducer.

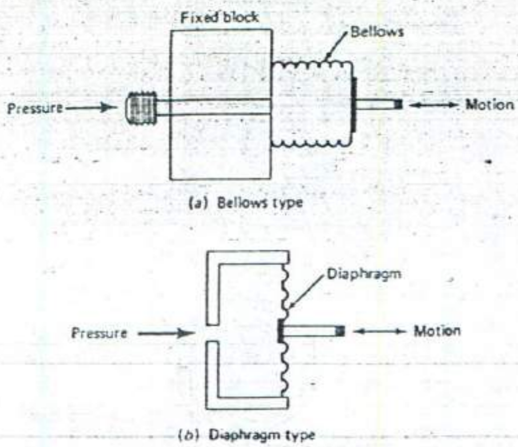


Fig. 9-32 Bellows and diaphragm pressure transducers.

phragms can be employed for differential pressure measurements.

Figure 9-33 shows a ***strain gage pressure transducer.*** The gage is bonded to a diaphragm. Pressure flexes the diaphragm, stretching the gage, and its resistance increases. The reference side of the diaphragm may be a perfect vacuum. If it is, the measured pressure will be absolute. Or a second pressure port may enter the reference side. If this is done, the output will be referenced to the second pressure and will be called *arbitrary* or a *differential pressure*. Finally, the reference side of the diaphragm can be vented to the atmosphere. Pressures measured this way are known as *gage pressures*.

Figure 9-34 shows the schematic diagram for a National Semiconductor LX04XXA monolithic pressure transducer. These units are actually piezoresistive integrated circuits. They provide an output voltage proportional to applied pressure. They also supply a separate temperature-dependent output that can be used to temperature compensate the transducer. Model LX0420A is rated to 100 pounds per square inch absolute (psia) and provides an output sensitivity of 0.2 to 0.8 mV/psi. It has a natural resonant frequency of 100 kHz. Other models are rated to 1000 psia and 3000 psia.

A *piezoelectric pressure transducer* uses a diaphragm to transmit pressure to a quartz element. An output voltage proportional to pressure is the result. Piezoelectric transducers are noted for high-frequency performance and can be used to monitor rapidly changing pressures. They are not dc sensors and cannot be used for static pressure measurements.

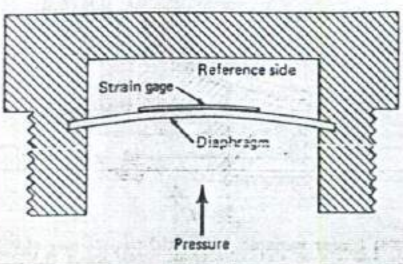


Fig. 9-33 Strain gage pressure transducer.

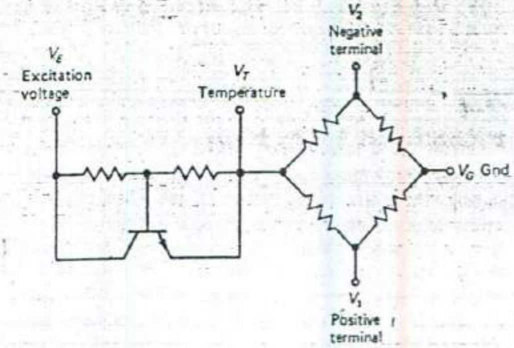
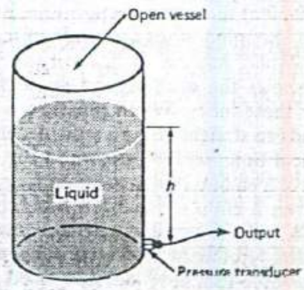


Fig. 9-34 Schematic diagram of LX04XXA monolithic pressure transducer (National Semiconductor).

Sudden pressure spikes can be damaging to some transducers. These transients are caused by pumps and valves and by resonance effects. Snubbers are sometimes used with pressure transducers to prevent transient damage. A *snubber* amounts to a restriction in a pressure port that prevents sharp increases or decreases in the pressure applied to the transducer. Snubbers are effective in preventing transducer damage, but they do limit the high-frequency performance of the measuring system.

Liquid level measurements can be accomplished indirectly with pressure transducers. Figure 9-35 shows such an application. The pressure at the transducer is proportional to the liquid density, the level above the transducer, and gravity. Note that the top of the tank is open (vented to the atmosphere). Suppose a tank contains water which has a density of



$P = dgh$
P = Pressure (N/m^2)
d = Liquid density (kg/m^3)
g = Gravitational constant (9.81 m/s^2)
h = Liquid height (m)

Fig. 9-35 Level measurement by static pressure.

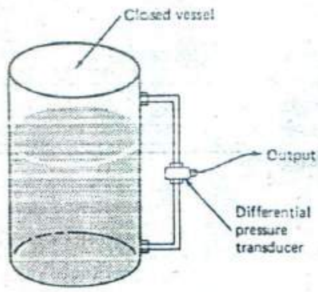


Fig. 9-36 Level measurement by differential pressure.

1000 kg/m^3 and that the level is 10 m above the pressure sensor. The pressure P at the sensor is found by

$$\begin{aligned} P &= dgh \\ &= 1000 \times 9.81 \times 10 \\ &= 98{,}100 \text{ N/m}^2 \\ P &= 98.1 \text{ kPa (14.22 psi)} \end{aligned}$$

You might wonder why the pressure is indicated in units of newtons per square meter. Recall that 1 N is defined as the force required to accelerate a 1-kg mass 1 m/s^2.

If the vessel holding the liquid is closed, a different technique is used to measure level. The height of the liquid above the pressure transducer is proportional to the difference between the pressure at the top of the tank and the static pressure at the transducer. A differential pressure transducer applied to a closed vessel is illustrated in Fig. 9-36.

Figure 9-37 shows a tank supported by load cells. The force on the load cells is proportional to the level in the tank. A summing amplifier may be used to add the signals from the individual load cells. Figure 9-38 illustrates a capacitive probe inserted into the tank. The liquid is an insulator with a dielectric constant that is different from the constant for the air or gas above the liquid. This produces a capacitance change in the probe as the level changes. The capacitive reactance of the probe also changes; the ac bridge produces an output signal that is proportional to level. In addition to load cells and capacitive probes, there are various techniques used to convert level to displacement. Floats, levers, cables, and pulley systems are employed, and the resulting displacement can be changed into an electrical signal with potentiometric transducers or LVDTs.

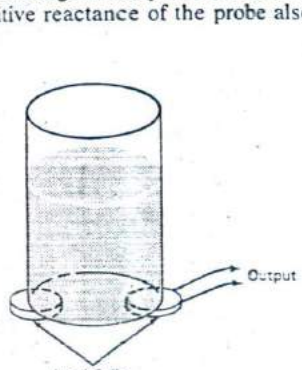


Fig. 9-37 Using load cells to measure level.

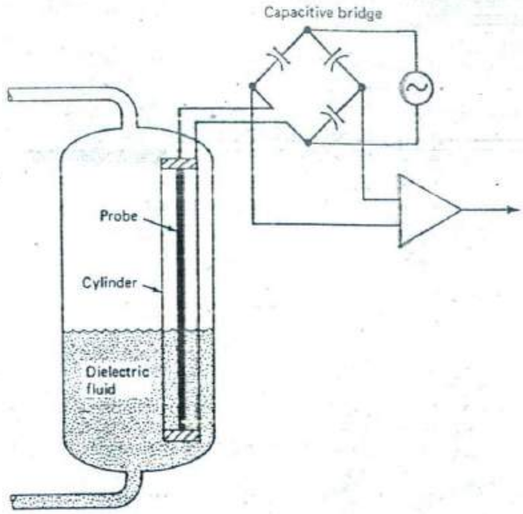


Fig. 9-38 Capacitive level measurement.

REVIEW QUESTIONS

22. A force of 1 N applied to an area of 1 m^2 is defined as a pressure of 1 ______.

23. The three ways to reference pressure are absolute, gage, and ______.

24. List three force-summing devices used in pressure-measuring transducers.

25. Refer to Fig. 9-33. The strain gage environment is a perfect vacuum. What type of pressure does the transducer measure?

26. Refer to Fig. 9-35. The tank contains water. The transducer registers a static pressure of 4 psi. How far above the transducer is the top of the water?

27. The top of a tank is sealed. What type of pressure measurement will be appropriate for evaluating the liquid level in the tank?

9-5 TEMPERATURE

There are three temperature scales widely used in industry: Fahrenheit, Celsius, and Kelvin. The *Fahrenheit* scale is the oldest and dates back to the early 1700s. It originally used the freezing point of water and the temperature of human blood as its two reference points. About 40 years later, Celsius proposed that the melting point of ice and the boiling point of water be used as reference points. His system became known as the *centigrade scale*, and in 1948 the name was officially changed to the *Celsius* scale. Lord Kelvin first proposed the concept of absolute zero in the early 1800s. His scale uses 0° to represent absolute zero. The conversions for the three scales are as follows:

$$C = 5/9(F - 32)$$
$$F = 9/5C + 32$$
$$K = C + 273.15$$

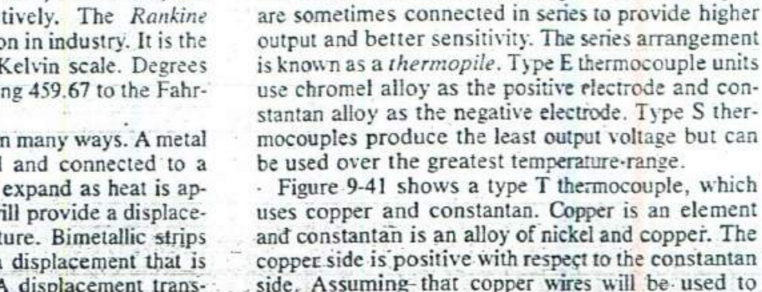
where C, F, and K are the Celsius, Fahrenheit, and Kelvin temperatures, respectively. The *Rankine scale* also finds some application in industry. It is the Fahrenheit equivalent of the Kelvin scale. Degrees Rankine may be found by adding 459.67 to the Fahrenheit value.

Temperature can be sensed in many ways. A metal tube can be filled with liquid and connected to a Bourdon tube. The liquid will expand as heat is applied, and the Bourdon tube will provide a displacement proportional to temperature. Bimetallic strips can also be used to provide a displacement that is proportional to temperature. A displacement transducer can be added to the Bourdon tube or bimetallic strip to provide an electrical output. However, it is usually easier to use a sensor that directly converts temperature into an electrical signal. Figure 9-39 shows the four common temperature sensors. The *thermocouple* produces an output voltage directly related to temperature. The *resistance temperature detector* (RTD) shows an increase in resistance with temperature. The *thermistor* has an opposite response, in that its resistance decreases with a temperature increase. Finally, the *integrated circuit sensor* produces a voltage or a current signal that increases with temperature increases. This section will deal with these four devices.

Some industrial thermocouples are shown in Fig. 9-40. They are based on the junction of two dissimilar metals. When the junction is heated, a voltage is generated; this is known as the *Seebeck effect*. The Seebeck voltage is linearly proportional for small changes in temperature. Various combinations of metals are used in thermocouples. Thermocouples are sometimes connected in series to provide higher output and better sensitivity. The series arrangement is known as a *thermopile*. Type E thermocouple units use chromel alloy as the positive electrode and constantan alloy as the negative electrode. Type S thermocouples produce the least output voltage but can be used over the greatest temperature range.

Figure 9-41 shows a type T thermocouple, which uses copper and constantan. Copper is an element and constantan is an alloy of nickel and copper. The copper side is positive with respect to the constantan side. Assuming that copper wires will be used to connect the thermocouple to the next circuit, a second copper-constantan junction is unavoidable as the illustration shows. This second junction is called the *reference junction*. It generates a Seebeck voltage that opposes the voltage generated by the sensing junction. If both junctions are at the same tempera-

	Thermocouple	RTD	Thermistor	IC Sensor



	Thermocouple	RTD	Thermistor	IC Sensor
Advantages	☐ Self-powered ☐ Simple ☐ Rugged ☐ Inexpensive ☐ Wide variety ☐ Wide temperature range	☐ Most stable ☐ Most accurate ☐ More linear than thermocouple	☐ High output ☐ Fast ☐ Two-wire ohms measurement	☐ Most linear ☐ Highest output ☐ Inexpensive
Disadvantages	☐ Non-linear ☐ Low voltage ☐ Reference required ☐ Least stable ☐ Least sensitive	☐ Expensive ☐ Power supply required ☐ Small ΔR ☐ Low absolute resistance ☐ Self-heating	☐ Non-linear ☐ Limited temperature range ☐ Fragile ☐ Power supply required ☐ Self-heating	☐ $T < 200°C$ ☐ Power supply required ☐ Slow ☐ Self-heating ☐ Limited configurations

Fig. 9-39 Four common temperature sensors (Omega Engineering).

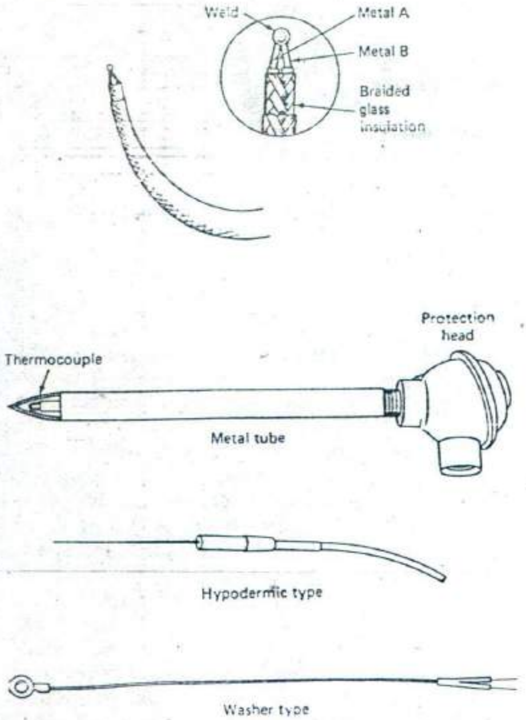


Fig. 9-40 Industrial thermocouples.

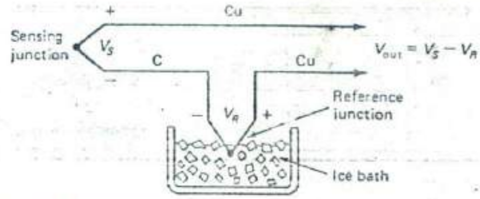


Fig. 9-42 Cold-junction compensation.

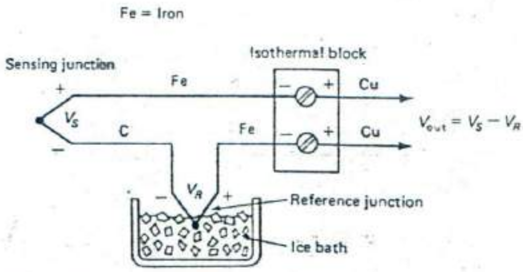


Fig. 9-43 Four-junction circuit.

ture, V_{out} will be zero. If the sensing junction is at a higher temperature, V_{out} will be proportional to the differences between the two junction temperatures. The problem is that the temperature cannot be derived directly from the output voltage alone. It is subject to an error caused by the voltage produced by the reference junction.

One solution to the problem is shown in Fig. 9-42. The reference junction is placed in an ice bath to keep it at a known temperature. This process is known as *cold junction compensation*. The reference junction is maintained at 0°C, and the reference voltage is now predictable from the calibration curve of the type T thermocouple. The reference voltage is subtracted from V_{out}, and the temperature of the sensing junction is found from the calibration curve. When copper is not one of the thermocouple metals, the four-junction circuit of Fig. 9-43 results. The type J thermocouple uses iron and constantan. When it is connected to copper wires, two iron-copper junctions result. These junctions present no additional compensation problems, however, because of the isothermal block. This block is made of a material that is a poor conductor of electricity but a good conductor of heat. Both iron-copper junctions will therefore be at the same temperature and generate the same Seebeck voltage. Note that these two voltages will cancel. Also notice that cold-junction compensation is used at the reference junction.

It is obvious that ice baths are not the most convenient way to compensate the reference junction. This technique is used in the calibration laboratory. The industrial environment demands a different approach; Fig. 9-44 shows one possibility. The isothermal block contains two reference junctions and a thermistor. The resistance of the thermistor is a function of temperature. A circuit is used to sense

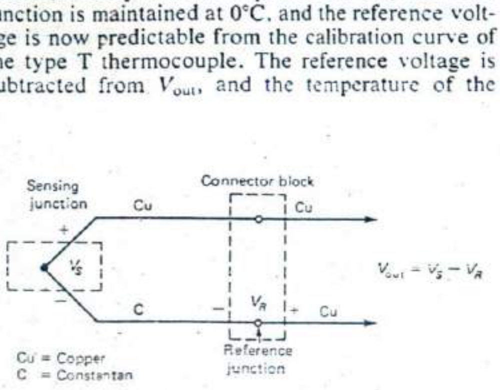


Fig. 9-41 Thermocouple structure.

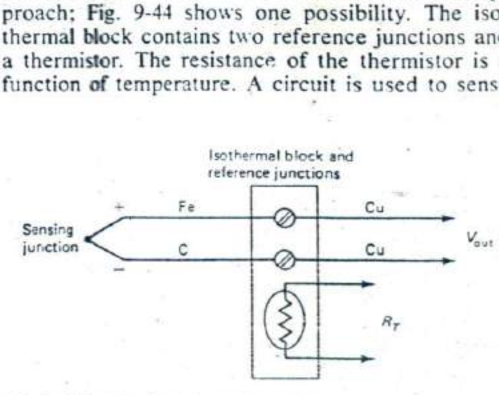


Fig. 9-44 Sensing the reference junctions.

this resistance and to compensate for the voltage introduced by the two reference junctions. This arrangement is sometimes called *electronic ice point reference*. If the sensor is interfaced to a computer, the reference temperature will be converted to a reference voltage and then subtracted from V_{out}. This process is known as *software compensation*. The question now is why bother with any of these processes since the thermistor appears to be capable of sensing absolute temperature with no compensation problems? Thermocouples are useful over a much wider temperature range than the other three sensors. They can be optimized for various atmospheres and are rugged and inexpensive. They lend themselves to monitoring a large number of locations. An isothermal block with one temperature sensor can provide compensation for several units. The term *zone block* is often used in this application. A scanner circuit using reed relays selects one junction from the zone block at a time. If software compensation is used, the individual thermocouples do not have to be of the same type. Different correction voltages for the various metals will be stored in computer memory. The outstanding advantages of thermocouples outweigh their disadvantages for many industrial applications.

Metals exhibit a *positive temperature coefficient;* their resistance increases with temperature. This effect is exploited in resistance temperature detectors (RTDs). They are usually made from platinum, which can maintain its stability at high temperatures. Figure 9-45 shows several styles of platinum RTDs. The glass-encapsulated type is bifilar-wound with platinum wire on a glass or ceramic bobbin. This type of

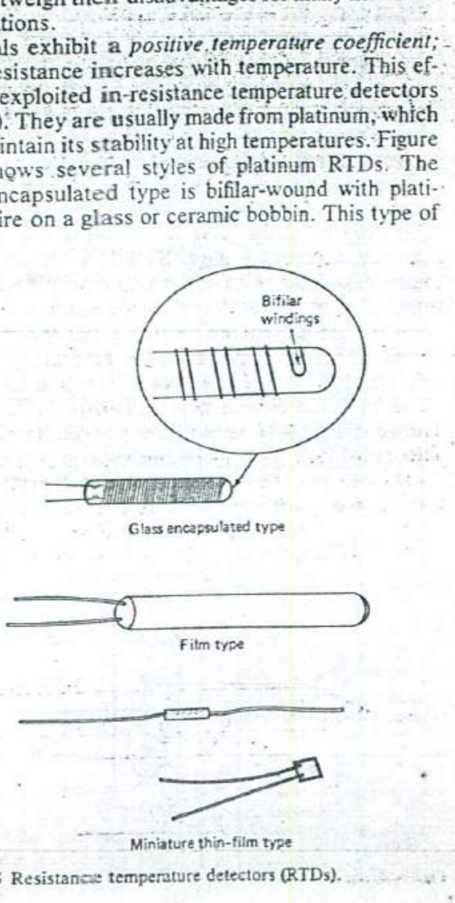


Fig. 9-45 Resistance temperature detectors (RTDs).

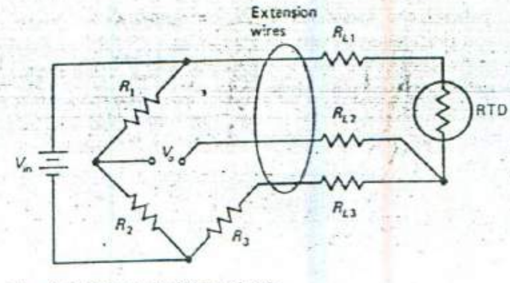


Fig. 9-46 Three-wire bridge circuit.

winding reduces magnetic pick-up, and a sensor less susceptible to electrical noise results. The assembly is then sealed with molten glass. The film types are manufactured with a platinum film on an alumina substrate. If the film is screen-deposited, it is a *thick-film* type. If it is vapor-deposited, a *thin-film* type results. The film types are less costly, yet are as accurate as the wire-wound types. They can be made very small, making their response time faster as a result of the low thermal mass.

Platinum RTDs are available from 10 Ω to several thousand ohms. The most popular value is 100 Ω at 0°C. Platinum's temperature coefficient is +0.00385. Thus, the typical RTD will increase its resistance by 0.385 Ω/°C. This small change in resistance demands the accuracy and sensitivity of a bridge circuit. The sensor is usually mounted away from the bridge so that the bridge resistors are not subjected to a temperature that would cause them to drift or to be damaged. The extension wires are a source of error because of their resistance. They also exhibit a positive temperature coefficient. The effects of the extension wires can be minimized with the three-wire bridge circuit of Fig. 9-46. The resistance of the extension wires is represented by R_{L1} and R_{L2}. If the wires are matched in length and material, their effects are canceled because each is in an opposite leg of the bridge. Lead R_{L3} is a sensor lead and carries little or no current. Therefore, its resistance is not a source of error. Unfortunately, the three-wire bridge circuit creates a nonlinear relationship between resistance change of the RTD and output voltage. For this reason, the four-lead circuit of Fig. 9-47 is pre-

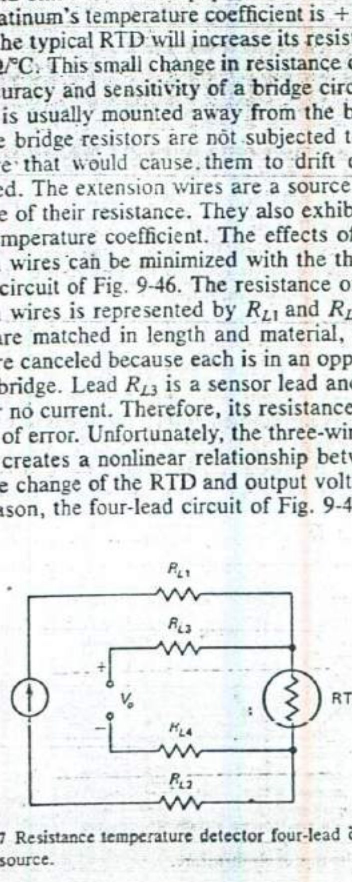
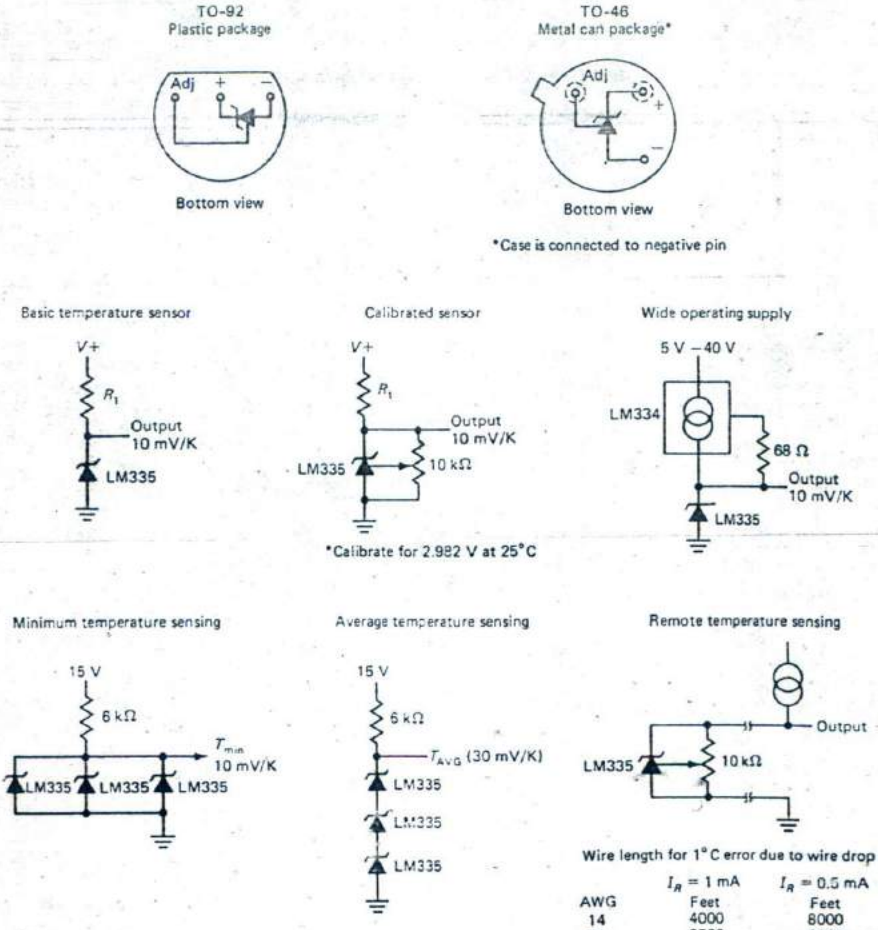


Fig. 9-47 Resistance temperature detector four-lead circuit with current source.



Wire length for 1° C error due to wire drop

AWG	I_R = 1 mA Feet	I_R = 0.5 mA Feet
14	4000	8000
16	2500	5000
18	1600	3200
20	1000	2000
22	625	1250
24	400	800

Fig. 9-48 Integrated circuit LM335 temperature sensor (National Semiconductor).

ferred for those applications requiring greatest accuracy. A constant current source supplies the RTD. The resistances of R_{L1} and R_{L2} are no longer a factor since the current will be held constant. Leads R_{L3} and R_{L4} are the sense leads. Their resistance is also not a factor, assuming that V_O is applied to a very high input impedance circuit such as an instrumentation amplifier.

A *thermistor* is a negative-coefficient sensor made from semiconducting material. Oxides of titanium, iron, and nickel are among the materials used. They are very sensitive, with temperature coefficients that range from −2 to −6 percent/°C. As such, thermistors are capable of detecting minute changes in temperature. Their resistance at 25°C ranges from 100 Ω to 100 kΩ, with 5000 Ω being a very common value. Their linearity is the poorest of those of all temperature sensors, and they are susceptible to permanent decalibration if exposed to high temperatures. The normal limit is 200°C, but they are also subject to decalibration if operated below, but near, their upper limit for extended periods of time. They can be built as small as 0.1 mm (about 0.005 in.), and their small thermal mass provides a very fast response time. They are available in a wide variety of shapes and sizes. Thermistors are much more delicate than thermocouples and RTDs.

The higher resistance of thermistors makes them less error-prone than RTDs. Lead resistance is not nearly so significant, and a simple two-wire extended

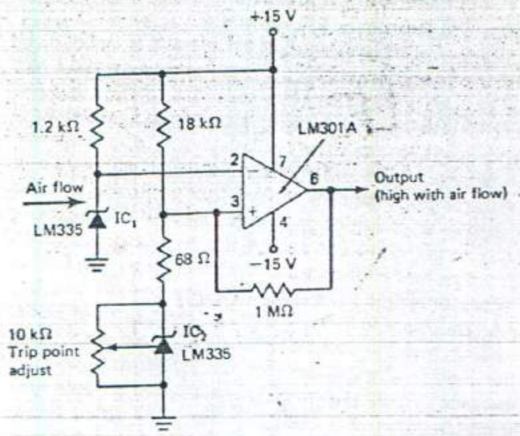


Fig. 9-49 Air flow detector.

bridge connection is usually adequate. However, they are subject to self-heating error. The bridge voltage is usually reduced to a low value to minimize this effect, or a pulsed supply may be used.

Integrated-circuit (IC) temperature sensors eliminate the linearity errors associated with thermistors. However, as semiconductor devices, they exhibit the other limitations. They are available in both voltage- and current-output configurations. Figure 9-48 shows the National Semiconductor LM335 IC temperature sensor. It provides a proportional output of 10 mV/°K. It operates as a two-terminal zener. Its dynamic impedance is less than 1 Ω, and it operates over a current range of 400 μA to 5 mA with virtually no change in performance. When calibrated at 25°C, it typically shows less than 1°C error over a 100° range. Its usable range is −10 to +100°C, and an LM135 is also available with a range of −55 to +150°C.

Figure 9-48 (p. 217) shows how to apply and calibrate the IC temperature sensor. The minimum temperature-sensing circuit is feasible because of the low dynamic impedance of the sensors. The coolest sensor will set the output voltage. The average circuit simply adds the individual output voltages. A simple potentiometer circuit provides one-point calibration. Single-point calibration works because the output of the sensor is proportional to absolute temperature with an extrapolated output of 0 V at 0 K. Errors in output voltage versus temperature are only slope errors. Thus, a slope calibration at one temperature corrects all temperatures.

Self-heating errors can be reduced by operating the IC at the minimum current suitable for the application. Sufficient current must be available to drive the sensor and the calibration pot at the maximum operating temperature. Self-heating can be exploited in some applications. Figure 9-49 shows a detector in which an air flow is directed onto an LM335 sensor. If the flow stops, the sensor temperature increases as a result of self-heating, and its output will go in a positive direction. This voltage is applied to the inverting input of the comparator. The LM301A comparator output will go negative when the threshold set by the voltage divider at its noninverting input is crossed. The second LM335 sensor provides trip point adjustment and adjusts the positive threshold for changes in ambient temperature. It also must be exposed to the air flow. Hysteresis is provided by the 1-MΩ feedback resistor. This circuit does not provide a linear measure of air flow but provides a negative-going output if the air flow stops or falls below the trip point value.

REVIEW QUESTIONS

28. Refer to Fig. 9-43. Ideally, how much error is introduced by the two iron-copper junctions? Why?

29. Refer to Fig. 9-44. The output of the thermistor goes to a computer, which uses stored information to correct for the reference temperature. This technique is known as ______ compensation.

30. Examine Fig. 9-46. Why is R_{L3} not significant?

31. Refer to Figs. 9-46 and 9-47. Which circuit is more accurate?

32. Refer to the calibrated sensor circuit shown in Fig. 9-48. What is its output at 100°C?

33. Refer to Fig. 9-49. The ambient air temperature increases. What happens to the threshold voltage at the positive input of the comparator? What is the net effect of this?

34. Refer to Fig. 9-49. What circuit feature reduces the possibility of multiple output pulses from the comparator as the air flow approaches the trip point?

9-6 MISCELLANEOUS MEASUREMENTS

Humidity is the moisture content of air. *Relative humidity* is the ratio of water vapor pressure in the atmosphere to that of the saturated water vapor pressure of the atmosphere at the same temperature. It is usually expressed in percentage form: 0 percent means there is no water vapor at all in the air, and 100 percent indicates that the air is holding all the water vapor that it can at that temperature. When the relative humidity is 100 percent, any drop in air temperature will initiate condensation of some of the water vapor. The temperature at which this occurs is known as the *dew point*. Relative humidity may also be used to express moisture content in artificial environments, as well as in gas or gaseous mixtures. Relative humidity affects electromagnetic propagation, ballistics, aerodynamics, and many industrial processes.

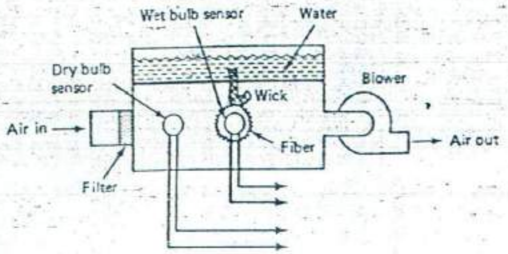


Fig. 9-50 Psychrometer.

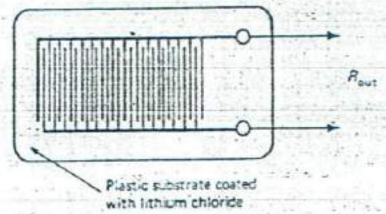


Fig. 9-51 Resistive humidity sensor.

Relative humidity may be measured with a *psychrometer*, as shown in Fig. 9-50. Air, or gas, is drawn into a chamber containing two temperature sensors. Thermistors or IC sensors are suitable for this application. One sensor is dry and measures the air temperature. The other sensor is encased in wet fiber. A wick and a reservoir maintain the wet condition. Water will evaporate from the wet fiber and cool the sensor. With low relative humidity, the water will evaporate quickly, and the wet sensor will achieve a temperature substantially lower than that of the dry sensor. High ambient air temperature also speeds the evaporation and increases the temperature difference. High relative humidity slows the evaporation, producing a lower temperature difference. Table 9-1 shows the relationship between temperature difference, air temperature, and relative humidity. More detailed data will be required for accurate measurements than are shown in this table. The data can be stored in a computer, and, with the proper sensor interface, an automatic measuring system can be realized.

The psychrometer is not a particularly convenient instrument. Hygrometers provide a more simple alternative for measuring relative humidity. Hygroscopic materials that readily absorb moisture from the air are available. For example, human hair is hygroscopic and can be used as a relative humidity sensor. A hair increases in length by about 3 percent when the relative humidity changes from 0 to 100 percent. The hair can be placed under tension, and an LVDT can be used to translate its length into an electrical signal. Resistance hygrometers also use hygroscopic materials to sense relative humidity. Figure 9-51 shows an example. The plastic substrate holds interlaced foil electrodes and is coated with lithium chloride. As the lithium chloride absorbs moisture from the air, the resistance between the electrodes drops. The performance of a resistance hygrometer is illustrated by Fig. 9-52. The resistance ranges from 80 kΩ at 100 percent to over 300 MΩ at 20 percent relative humidity.

Robotics has increased the interest in ***proximity sensors.*** Such sensors can help a robot find an object and are also useful for detecting obstructions and human personnel that have entered the work envelope. Figure 9-53 shows a simplified block diagram for an ultrasonic proximity sensor. A burst generator produces seven cycles at a 30-kHz frequency. This frequency is above the human range of hearing and is therefore considered ultrasonic. The burst rate is 2 Hz. The bursts are amplified and applied to an output transducer, where they become ultrasound. The sound waves travel at approximately 340 m/s. The reflected waves arrive at an input transducer and are changed back into an electrical signal. The time (t) between the transmit burst and the receive burst can be used to calculate distance. For example, if the time is 30 ms the total distance traveled is $340 \times 0.03 = 10.2$ m. The object that caused the reflection is half that distance away: 5.1 m (16.7 ft). Sound velocity changes with the temperature of the atmosphere. It travels at 331 m/s at 0°C and at 386 m/s at 100°C. A temperature sensor can be used to correct for this effect.

TABLE 9-1 RELATIVE HUMIDITY LOOK-UP TABLE

Dry Bulb, °C	Dry Bulb-Wet Bulb Difference, °C: 0.56	2.78	5.56	8.33	11.1	13.9
4.4	92	60	—	—	—	—
10.0	93	68	38	12	—	—
15.6	94	73	49	26	6	—
21.1	95	77	55	37	20	3
26.7	96	79	61	44	29	16
32.2	96	81	65	50	36	24
37.8	96	83	68	54	42	31

Relative Humidity, %

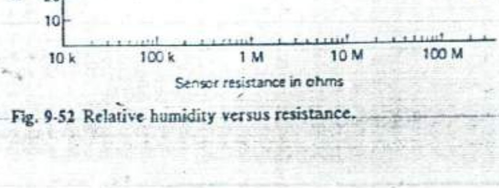


Fig. 9-52 Relative humidity versus resistance.

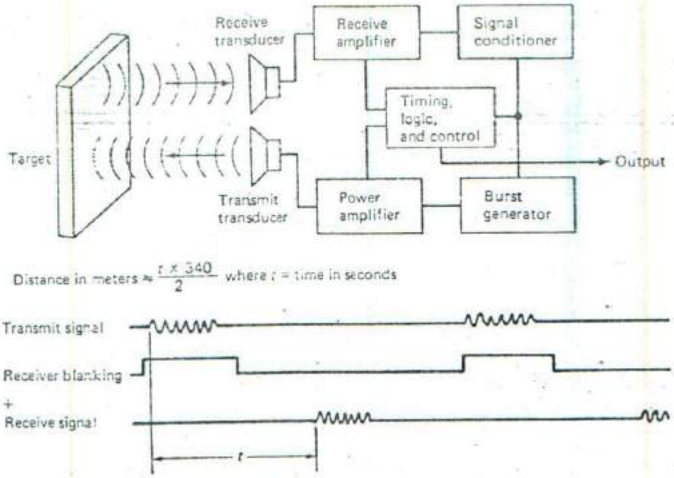


Fig. 9-53 Ultrasonic proximity sensor.

Some of the waveforms in an ultrasonic proximity sensor are shown in Fig. 9-53. A receiver blanking pulse turns the receiver off during the time the burst is applied to the output transducer. This pulse protects the receiver circuits from overload. It is possible to use a single transducer for both transmit and receive. A single piezoelectric element can be switched between the receiver input and the power amplifier output. It will be switched to the output only during the time of the transmit burst. The gain of the amplifier may be controlled by a ramp waveform not shown in the illustration. The strength of the received signal falls off as the inverse square of the total distance traveled. Therefore, for measuring distant objects, high receiver gain is needed. However, with high receiver gain, extraneous reflections and other sounds may give a false indication. The solution is to ramp the gain up as time increases. The receiver gain is set low for the time immediately following the transmit burst. As time increases, the receiver gain is also increased to compensate for the path loss.

Pulsed infrared systems are also finding increased application. They use light-emitting diode transmitters that operate below the frequency range of the visible spectrum. Phototransistors with high infrared sensitivity are used to receive the reflected signals. These devices are described in the discussion of optoelectronics. Infrared systems make excellent motion detectors and are used in industrial robotic, security, and safety installations.

The sensing of near ferrous objects can be based on magnetic principles. Figure 9-54 shows a *reluctance proximity sensor.* The assembly uses a permanent magnet and a core with a coil wound on it. The magnet produces a flux that surrounds the turns of the coil. There is no coil output since the flux is static. When a ferrous object enters the field of the magnet, flux distortion is produced and a cutting action results. The graph shows the coil voltage with an object approaching the sensor field and then leaving the sensor field.

Sensing of ferrous or nonferrous objects can be accomplished with the *eddy current killed oscillator* (ECKO) sensor shown in Fig. 9-55. An oscillator provides an ac signal for a coil which in turn generates an electromagnetic field. When a target is intercepted by this field, eddy currents are induced in the metal. These currents represent a circuit loss, and the amplitude of the oscillations decreases. In fact, if the coupling between the coil and the target is tight and if the target is made from a metal with large eddy

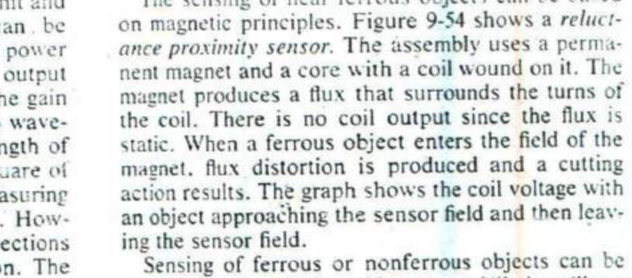


Fig. 9-54 Reluctance pick-up proximity sensor.

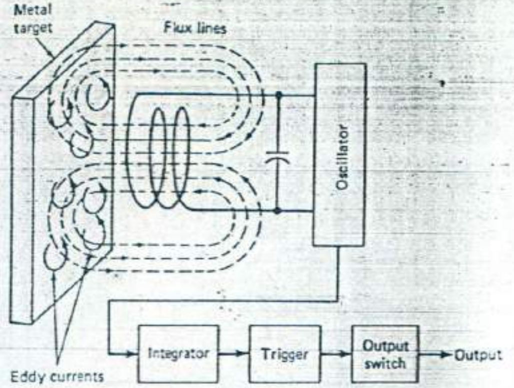


Fig. 9-55 Eddy current killed oscillator proximity sensor.

current losses, the oscillator will stop working. This effect established the name *killed oscillator*. It is not necessary to kill the oscillator to trigger the output, however. The integrator output can trip the trigger before oscillations cease altogether.

Figure 9-56 shows a tilt transducer. *Tilt* is any departure from the horizontal and is usually measured in degrees or radians. The tilt cell contains an electrolyte that conducts electricity. There are also three electrodes sealed in the tilt cell. When the cell is horizontal, the bubble is centered as shown, and electrodes A and B have the same exposure to the electrolyte. Any tilt will cause the bubble to move left or right. Suppose the seating plane tilts clockwise. This tilt causes the bubble to move to the left, decreasing the electrolyte contact on electrode A and increasing the contact area on electrode B. The resistance from the main electrode increases to A and decreases to B. The bridge circuit takes advantage of this dual effect, and an output voltage proportional to tilt results.

Atomic radiation must be sensed in some industrial environments for the protection of human personnel. Radiation may also be used in some measuring applications such as *thickness gaging*, in which a material placed between a radioactive source and a sensor absorbs an amount of energy proportional to its

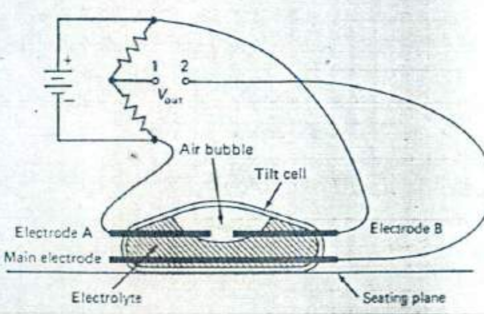


Fig. 9-56 Tilt transducer.

density and thickness. If the density is a known constant, the thickness of a material can be gaged by the output of a radioactive sensor.

Ionization transducers are used to measure atomic radiation. They are often tubes filled with inert or organic gas. Atomic particles enter the tube and collide with the gas molecules, creating free electrons and positive ions. The tube also contains electrodes. A voltage across the electrodes produces current pulses due to the free electrons and ions. Sensitivity may be improved by operating some tube types at a higher voltage. This higher voltage will accelerate any dislodged electrons to create more collisions, and an avalanche will result. Therefore, the current pulse will be larger. Geiger-Muller tubes are operated at very high potentials. They avalanche over the entire electrode area, and a single event becomes self-perpetuating. They require some means to stop the action after each event. Organic materials are used inside the tube to provide a self-quenching action after each discharge.

Solid-state radiation detectors are based on a reverse-biased junction. Normally, little or no current flows. Atomic particles entering the depletion region raise the energy level of electrons from the valence band to the conduction band. In addition to free electrons, holes are created at these sites. The holes and electrons serve as carriers to support the flow of current, and output pulses result.

Some substances absorb atomic energy and reemit the energy as photons. The flashes of light that result can be coupled by fiber-optic cable to a photomultiplier tube. The high gain of the photomultiplier tube produces a very sensitive radiation detector. These detectors are known as *scintillation counters*.

Not all thickness-gaging systems use radiation. Other techniques include variable reluctance, variable inductance, and variable capacitance. The inductance and reluctance types work with metallic materials, and the capacitance sensor is suited to gaging nonmetallic materials.

REVIEW QUESTIONS

35. What is the name of the instrument that uses wet and dry sensors to measure relative humidity?

36. What types of humidity sensors use hygroscopic materials?

37. Refer to Fig. 9-53. If $t = 20$ ms, how far away is the target?

38. Refer to Fig. 9-53. The burst rate is 2 Hz. What is the greatest distance that can be measured?

39. Reluctance proximity sensors are useful for detecting ______ objects.

40. The ECKO proximity sensor can detect ______ and ______ objects.

41. Refer to Fig. 9-56. The bridge output is zero when the seating plane is horizontal. What is the

polarity of terminal 1 with respect to terminal 2 if a counterclockwise tilt is introduced? A clockwise tilt?

9-7 SIGNAL CONDITIONING

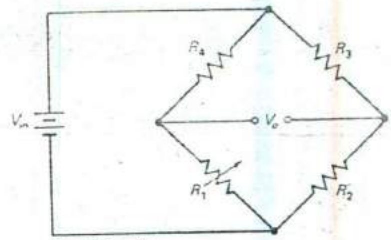
Signal conditioning provides any required gain, isolation, noise rejection, offsetting, or linearization of the output of a transducer. Gain is required in those cases in which the output is too small to be directly useful in a measurement or control system. The required gain is usually supplied by operational amplifiers or instrumentation amplifiers. These circuits are discussed in Chapter 6. The required gain may also be supplied by special signal-conditioning devices, several of which are covered in this section. Isolation is required in those applications in which ground loops must be eliminated. Noise rejection includes common mode nulling and low-pass filtering.

Offsetting is required when the level of a signal must be shifted by some predictable amount. For example, an application may require the measurement of small changes about some large initial value. Offsetting may also be required to remove an error, as in thermocouple cold-junction compensation. Another example is the conversion of one measurement scale to another, such as gage pressure to absolute pressure or degrees Celsius to degrees Kelvin. Finally, offsetting also includes the conversion of a voltage signal to a current signal for transmission purposes. This topic will be covered later in this section.

Linearization may be accomplished by digital or analog methods. The digital approach lends itself to applications in which the transducer output is applied to the input of a computer. The computer may linearize the readings by performing mathematical operations on them. Another computer approach is to convert each digitized value to a corresponding corrected value by using readings stored in computer memory. The linearized values are stored in a lookup table, such as Table 9-1. Analog linearization uses amplifiers and other circuits that have a nonlinear response that is complementary to the characteristic curve of the transducer. For example, an amplifier with a logarithmic response can be used to linearize a sensor with an exponential output.

In addition to transducer nonlinearity, circuit nonlinearities can also introduce significant errors in some cases. Let's review the basic instrumentation bridge circuit shown in Fig. 9-57. When $R_1/R_4 = R_2/R_3$ the bridge is at null and $V_O = 0$. If the ratio R_2/R_3 is fixed at K, then a null condition guarantees that $R_1 = KR_4$. If R_1 is unknown, it can be measured by nulling the bridge by adjusting R_4. If R_4 is calibrated, the unknown is found by multiplying by K. Null-type measurements are typically used in feedback systems.

Figure 9-57 shows a different set of equations because transducer measurements involve the deviation of one or more of the bridge elements. The equations show that the output voltage is not a linear function of the resistance change in R_1. Assume that all of the bridge elements are nominally equal and that R_1 is variable with a fractional deviation of X. Note that the final equation shows that the output voltage will be approximately equal to $V_{in}/4 \times X$ for very small values of X. As X gets larger, the nonlinearity becomes more evident. As a demonstration, the following values of V_O were calculated for $V_{in} = 10$ V:



$$R_1 = R(1+X)$$

$$V_o = \frac{R(1+X)}{R+R(1+X)} V_{in} - \frac{V_{in}}{2}$$

$$= \frac{2+2X-2-X}{2(2+X)} V_{in}$$

$$= \frac{V_{in}}{4} \cdot \frac{X}{1+\frac{X}{2}}$$

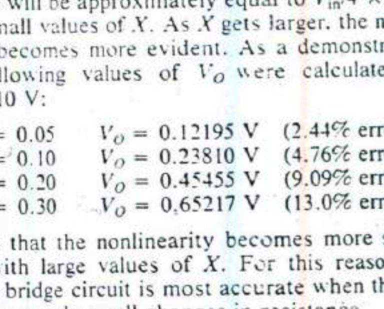
$$\approx \frac{V_{in}}{4} \cdot X \text{ (when } X \ll 1)$$

Fig. 9-57 Typical instrumentation bridge circuit.

$X = 0.05$	$V_O = 0.12195$ V	(2.44% error)
$X = 0.10$	$V_O = 0.23810$ V	(4.76% error)
$X = 0.20$	$V_O = 0.45455$ V	(9.09% error)
$X = 0.30$	$V_O = 0.65217$ V	(13.0% error)

Notice that the nonlinearity becomes more significant with large values of X. For this reason, the simple bridge circuit is most accurate when the sensor shows only small changes in resistance.

Figure 9-58 shows a linear bridge circuit. This ar-



$$V_O = \left[\frac{R(1+X)}{2R} - \frac{R(1-X)}{2R}\right] V_{in}$$

$$= \left[\frac{1+X-1+X}{2}\right] V_{in}$$

$$= X\ V_{in}$$

Fig. 9-58 Linear bridge.

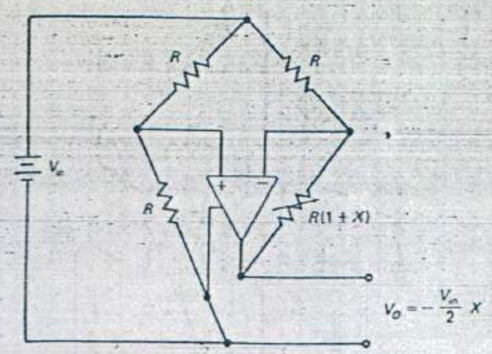


Fig. 9-59 Active bridge.

rangement might represent two identical, two-element strain gages mounted on opposite sides of a cantilever beam. As the beam stretches on one side, it compresses on the other side. This would increase the resistance of two bridge elements and decrease the resistance of the other two bridge elements. The output of this bridge is linear and is four times the magnitude of a bridge where only one element changes resistance.

Unfortunately, the circuit of Fig. 9-58 does not lend itself to all measurement applications. Another technique that improves linearity is to make the resistance ratios of the top elements to the bottom bridge elements large. This process, however, decreases the bridge sensitivity. Figure 9-59 shows an active bridge with good linearity and sensitivity. An operational amplifier enforces a null condition even for large values of X. It adds a variable voltage in series with the sensor to maintain the null. This voltage is also the output voltage and is inherently linear with changes in X and twice the magnitude of the basic circuit of Fig. 9-57.

The bridge excitation voltage (or current) is another source of error. It must be well regulated for high accuracy. Figure 9-60 shows the application of a 2B31 transducer signal conditioning module manufactured by Analog Devices. It solves many of the problems associated with accurate measurements. This module also provides gain and stable bridge excitation for measurements using resistive sensors such as RTDs or strain gage sensors. It features a high CMRR (140 dB) and active low-pass filtering for a low-noise output. Pin 29 is available for offsetting the output over a plus and minus 10 V range. If no offsetting is required, pin 29 is grounded. The gain can be set from 0 to 66 dB by varying the resistor across pins 10 and 11.

Linearization is provided by feeding back a small percentage of the amplifier output to modulate the bridge excitation voltage. The sense of the feedback is determined by whether the nonlinearity is concave

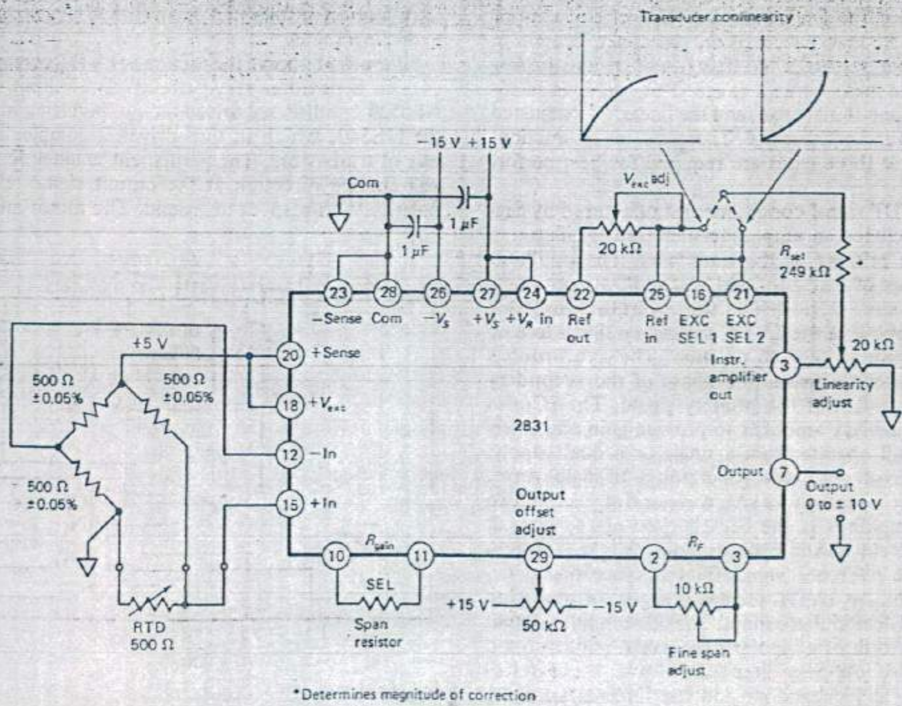


Fig. 9-60 2B31 Linearizing circuit (Analog Devices).

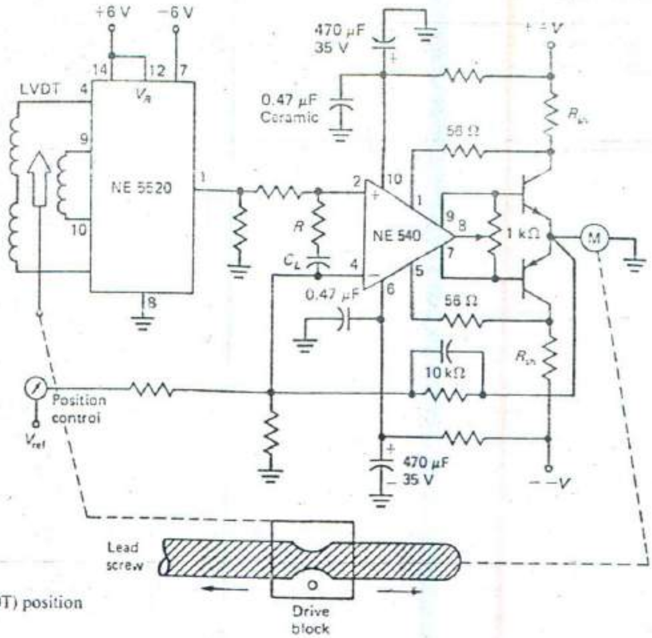


Fig. 9-61 Linear variable differential transformer (LVDT) position servo (Signetics).

upward or downward. Figure 9-60 shows the appropriate jumper connection for each case. The magnitude of the correction is set by R_{sel}, and the linearity pot is used for final trimming. The circuit can be adjusted by using a precision resistance decade in place of the sensor. The offset is adjusted at the low end of the measurement range. The fine span is adjusted at one-third range, and the linearity is adjusted at the top end of the range. There is some interaction, and two or three trials are required for best performance.

An LVDT signal conditioner manufactured by Signetics includes an amplitude-stabilized oscillator to drive the primary of the transducer. The oscillator has a range of 1 kHz to 20 kHz. The IC also contains a synchronous demodulator to convert the amplitude and the phase of the LVDT secondary signal to a dc voltage proportional to position. The synchronous demodulator compares the phase of the secondary signal with that of the primary signal. The IC also has an auxiliary amplifier to provide gain and filtering. It will operate from a single or a dual supply. The internal oscillator generates a triangle wave, which is converted to a sine wave and applied to two driver amplifiers in the IC. Their output appears at two different pins in phase opposition. One pin provides the reference signal for the synchronous demodulator. An LVDT secondary signal is applied to a second demodulator input. The demodulator output is in the form of a bipolar full-wave rectified signal. The active low-pass filter formed with the external resistors and capacitors and the internal auxiliary amplifier removes the carrier component (ripple) and provides gain. The dc output signal is equal to half the reference voltage when the LVDT core is at null. Core motion results in a dc shift that is proportional to displacement.

Figure 9-61 shows the Signetics LVDT conditioner used in a position servo circuit. The output of the NE5520 conditioner drives the noninverting input of the NE540, which in turn drives a complementary pair of transistors. The permanent magnet dc motor will change direction as the output signal reverses polarity with respect to ground. The motor mechan-

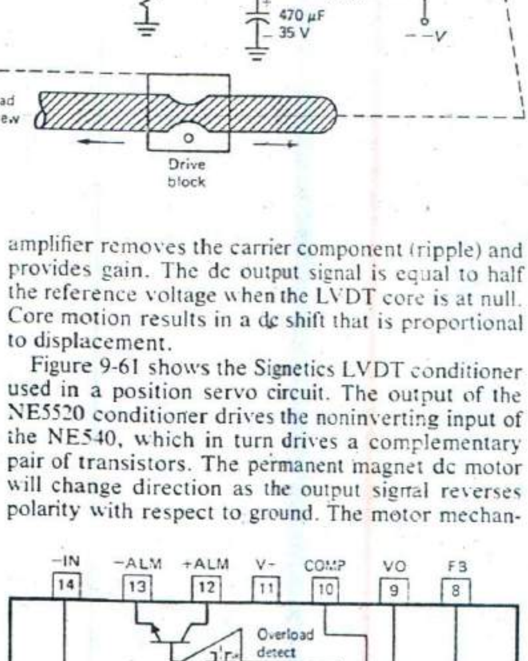


Fig. 9-62 Thermocouple amplifier (Analog Devices).

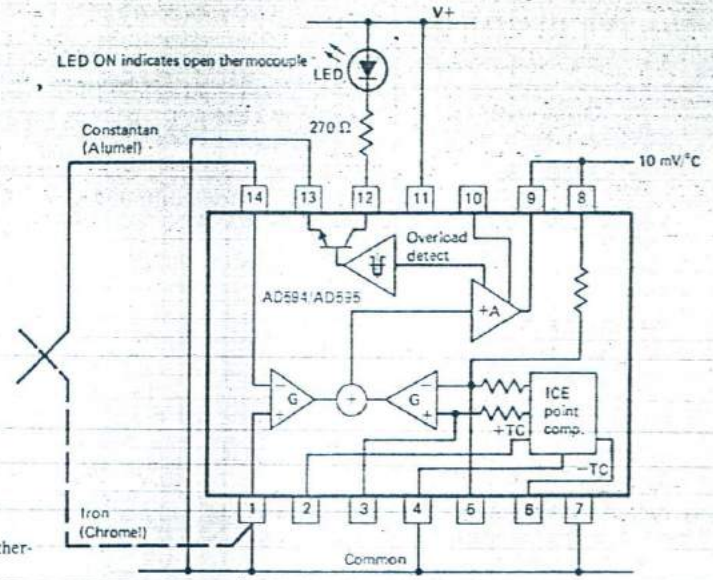


Fig. 9-63 Single supply circuit with open thermocouple indicator (Analog Devices).

ically drives a lead screw, and the drive block is mechanically coupled to the core of the LVDT position sensor. The inverting input of the NE540 is established by V_{ref} and the position control. Any discrepancy between the setting of the position control and the actual position sensed by the LVDT will result in a differential signal at the NE540. The motor will run in a direction to reduce the discrepancy. Servos are covered in detail in Chapter 10.

The Analog Devices AD594/AD595 monolithic thermocouple amplifier is shown in Fig. 9-62. The 594 is precalibrated by laser wafer trimming to match the characteristics of type J sensors, and the 595 matches type K thermocouples. The chip contains a complete instrumentation amplifier and a cold-junction compensator. It also includes a thermocouple failure alarm circuit that activates if one or both thermocouple leads open. Figure 9-63 shows the device interfaced to a thermocouple. The thermocouple leads are soldered directly to pins 1 and 14, producing copper-constantan (or copper-alumel) and copper-iron (or copper-chromel) reference junctions. These junctions are isothermal with the IC itself, and the ice point compensator offsets their Seebeek voltages. Special solders are recommended for these connections. They should be composed of noncorrosive rosin flux and an alloy of 95 percent tin with 5 percent antimony, 95 percent tin with 5 percent silver, or 90 percent tin with 10 percent lead. Ordinary electronic solder is usually 60 percent tin and 40 percent lead.

Many transducer signals in industry are converted to current signals for transmission to a distant controller or computer. The standards are 4 to 20 mA and 10 to 50 mA, with the 4 to 20 range being the most common. Current transmission offers several advantages. The signal is not affected by noise, drops in the line, stray thermocouples, contact resistance, or contact emf. Only two wires are needed, and an open circuit fault is easily detected by a 0-mA signal condition. The National Semiconductor LH0045 is a two-wire transmitter. It is a linear IC that converts the voltage signal from a sensor or a bridge to current. A single twisted pair of wires is all that is required for both signal output and supply circuits. The device contains an internal reference to power the bridge, an input amplifier, and an output current source. The output is adjustable to meet either industry current standard. It interfaces easily with thermocouples, strain gages, RTDs, and thermistors.

REVIEW QUESTIONS

42. Refer to Fig. 9-57. The bridge is at null when V_O = ______.

43. Refer to Fig. 9-57. Calculate the actual value of V_O when $X = 0.03$ for $V_{in} = 5$ V. Calculate the nominal value by using the approximation. Calculate the percentage of error.

44. Refer to Fig. 9-57. What happens to the nonlinearity as X increases?

45. Refer to Fig. 9-58. Calculate V_O for $V_{in} =$ 5 V and $X = 0.1$. For $X = 0.2$. Is V_O a linear function of X?

46. Refer to Fig. 9-59. Calculate V_O for $V_{in} =$ 5 V and $X = 0.07$. Is V_O a linear function of X?

CHAPTER REVIEW QUESTIONS

9-1. Stretching a conductor will ______ its resistance.

9-2. An industrial strain gage has a GF of 2 and a resistance of 350 Ω. If its normal length is 2 cm, what will its resistance be if it is stretched to 2.02 cm?

9-3. What output voltage would be produced by the data in question 9-2 if the gage were in a Wheatstone bridge circuit with three fixed 350-Ω resistors?

9-4. Assume that a two-element strain gage is used for temperature compensation. What would have to be done to the gain of the instrumentation amplifier to compensate for the transverse sensitivity factor?

9-5. Why might dual-staggered slots be used in an optical tachometer?

9-6. When the output of an accelerometer is integrated, a signal proportional to ______ results.

9-7. The LVDT accelerometers do not lend themselves to vibration analysis because of their rather ______ resonant frequency.

9-8. The principle that relates fluid flow rate and pressure is attributed to ______.

9-9. Identify a transducer that divides flow into measured units.

9-10. Which temperature sensor converts heat to a voltage?

9-11. Which temperature sensor has the greatest temperature range?

9-12. The signal generated by a thermocouple is due to the ______ effect.

9-13. Which temperature sensor requires compensation?

9-14. A series arrangement of thermocouples is known as a ______.

9-15. Low thermal mass is desired in a sensor requiring a ______ response time.

9-16. Why do wire-type RTDs use bifilar windings?

9-17. A platinum RTD is rated at 100 Ω at 0°C. Assuming a positive temperature coefficient of 0.00385, what is its resistance at 50°C?

9-18. Low-resistance sensors, such as RTDs, require three or four wire connections to eliminate the error due to ______ resistance.

9-19. Radiation detectors that use a gas-filled tube are based on the ______ principle.

9-20. Atomic radiation entering a semiconductor depletion region can cause the junction resistance to ______.

9-21. What type of radiation detector utilizes light?

9-22. Linearization based on look-up tables is a(n) ______ technique.

9-23. Refer to Fig. 9-63. Why must the thermocouple leads be soldered at the IC pins and not to copper extension wires?

9-24. Why do the current transmission standards use 4 or 10 mA to represent zero rather than 0 mA?

ANSWERS TO REVIEW QUESTIONS

1. 9 V **2.** 8.88 to 9.12 V **3.** 0.2 percent **4.** 2 V, positive **5.** A would be 2 V negative with respect to B **6.** none **7.** 3 **8.** the polarity reverses **9.** they cannot indicate shaft direction **10.** it provides direction information **11.** acceleration **12.** displacement; force; acceleration **13.** 667 Hz **14.** fewer cycles of oscillation (more rapid decay) **15.** for temperature compensation **16.** 3.56×10^7 N/m **17.** 5.93×10^{-4} m/m **18.** 0.142 Ω **19.** negative **20.** velocity **21.** P_T **22.** pascal **23.** differential **24.** Bourdon tube; bellows; diaphragm **25.** absolute **26.** 2.81 m **27.** differential **28.** none; the voltages cancel **29.** software **30.** it is a sense lead (little or no current) **31.** Fig. 9-47 **32.** 3.732 V **33.** it increases; IC_1 will have to become warmer to trip the comparator **34.** hysteresis **35.** psychrometer **36.** hygrometers **37.** 3.4 m **38.** 85 m (in practice it is much less) **39.** ferrous **40.** ferrous; nonferrous **41.** positive; negative **42.** zero **43.** 0.03695 V; 0.03750 V; 1.47 percent **44.** it increases **45.** 0.5 V; 1.0 V; yes **46.** −0.175 V; yes

10

SERVOMECHANISMS

This chapter discusses servomechanisms and the components that make up servosystems. Some components are electrical, some electronic, and some mechanical. Servomechanisms *are systems that position an object by comparing position feedback signals with command signals. Systems that use feedback are* closed-loop systems; *those that do not use feedback are* open-loop systems. *Feedback can be continuous or discontinuous. This chapter will only deal with continuous types of control systems.*

10-1 POTENTIOMETERS AND ENCODERS

Precision potentiometers are simple rotary devices for obtaining shaft position information. The most straightforward application is the conversion of mechanical position to a voltage. Basically, a precision potentiometer consists of a resistive element with a movable arm, or *slider,* in contact with the element. As the arm (slider) rotates, the resistance varies between the end of the resistive element and the slider, indicating shaft position. The resistive element can be made of wire, conductive film, or a cermet element.

Potentiometers used for servomechanisms are generally about ⅞ to 3⁵⁄₁₆ in. (22.2 to 84.125 mm) in diameter. The early models were mostly of the wirewound type. Current technology provides other choices such as conductive plastic, which offers a better temperature stability, longer life, and lower sensitivity to the environment. Potentiometers can be excited with alternating or direct current. Single-turn potentiometers have a rotation that is usually limited to 350°. Some models have continuous rotation with no internal stops. Potentiometers may be ganged so that a single shaft will rotate several sliders. *Multiturn potentiometers* are limited to 3, 5, 10, 15, 25, or 40 revolutions before hitting internal stops. Figure 10-1 shows the internal construction of a ten-turn potentiometer. The winding (resistance element) is in the form of a helix, and the contact assembly is such that the slider travels a helical path while making contact with the resistance element.

A linear potentiometer produces a resistance change that is linearly related to its shaft position. A position of one-half rotation will produce 50 percent of maximum resistance, and a position of three-fourths rotation will produce 75 percent of maximum resistance, and so on. *Linearity* is specified as the deviation (in percentage of the total resistance) of the actual resistance at any point from the expected resistance. This is called *normal,* or independent, linearity. A standard value of linearity is 0.1 percent with 0.01 percent types available. If a load resistance is placed between the slider and one end of a linear potentiometer, as shown in Fig. 10-2, the potentiometer will no longer be linear. The magnitude of loading error is a function of the ratio of the slider load to the total resistance of the potentiometer. This error varies inversely with the load ratio. That is, a small load ratio will produce a large error.

There are several other important characteristics of potentiometers, such as resolution, noise, and mechanical tolerances. *Resolution* in a potentiometer is the minimum change of resistance output expressed

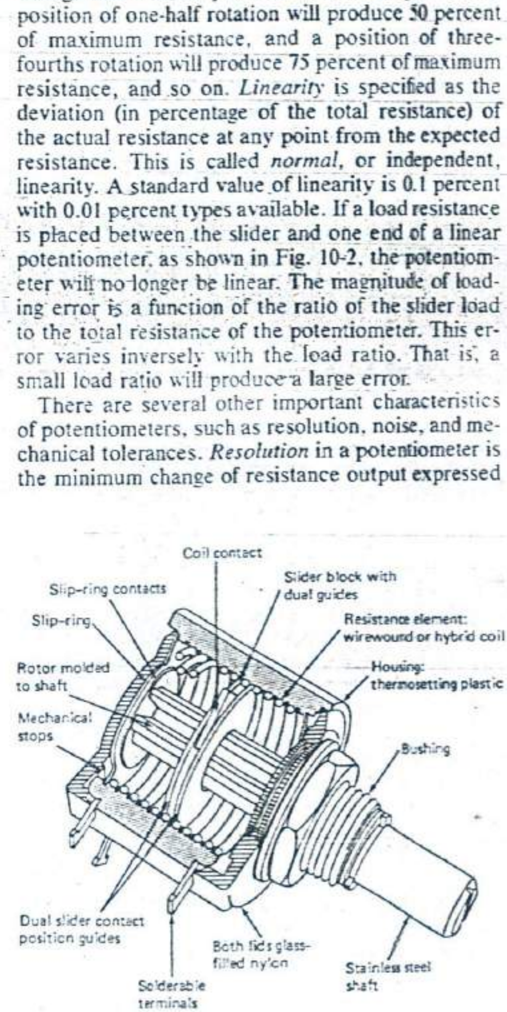


Fig. 10-1 Multiturn potentiometer.

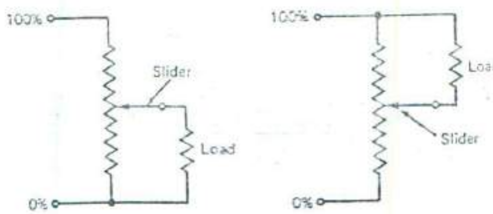


Fig. 10-2 Potentiometer loading.

as a percentage of its total resistance. It is dependent on the number of turns of wire per inch on the winding and the arc diameter of the slider.

A nonwire potentiometer is stepless and has essentially infinite resolution. The typical resolution of a wire-wound potentiometer is 0.05 percent. Many high-gain servomechanisms have a tendency to "hunt" between the turns of wire on the potentiometer, seeking a voltage that does not exist in the surface. Plastic conductive and cermet units have all but eliminated this problem. A multiturn wire-wound potentiometer may also be employed because of its better resolution.

Noise in a potentiometer appears as spurious unwanted voltages. If a wire-wound potentiometer is excited with direct current, for example, the finite resolution of the potentiometer will cause a ripple voltage to appear at the slider as the shaft is rotated. In some systems this noise can cause problems, especially if the potentiometer's slider is worn and the noise becomes excessive. There will also be an increase in the hunting of the system. This noise is all but eliminated with the nonwire potentiometer. Mechanical tolerances of potentiometers are important in servomechanism applications. They are manufactured with accurate external surfaces to allow interchangeable units and to permit low backlash (see Chapter 10) and matching coupling surfaces.

Why use wire-wound potentiometers? Two reasons are apparent. First, the wire-wound potentiometer can be made with very low values, such as 10 Ω. Values less than 1000 Ω are hard to obtain with non-wire-wound types. The second reason is that nonlinear functions are needed in many servomechanism systems, and it is all but impossible to duplicate the specifications obtainable from wire-wound units. Nonlinear potentiometers can be made by winding the element on a mandrel which has the slope of the function. Examples of mathematical functions that are available include the tangent, secant, cosecant, square root, and inverse functions. The use of taps on a linear winding is another method used to obtain nonlinear functions. Sections of the resistance between taps can be loaded with external resistors to warp the resistance to obtain a given function. From 4 to 10 taps are normally used, as shown in Fig. 10-3.

A relatively new potentiometer that solves numerous linear control applications is the *rectilinear potentiometer.* This unit has strokes (travels) from ½ to 6 in. (12.7 to 152 mm), with an independent linearity of 0.1 percent.

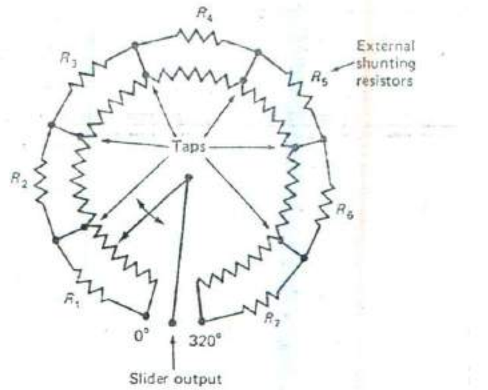


Fig. 10-3 Tapped potentiometer.

Optical encoders may also be used to provide position feedback. The primary parts of an optical encoder are shown in Fig. 10-4. Two types of rotary optical encoders available are *incremental* (outputs a fixed number of N pulses per revolution) and *absolute* (the output is a unique code for each angular position). *Incremental optical encoders* generally provide two signals which are in quadrature (90° phase difference). The phase difference provides directional information. The output waveforms may be square wave or sine wave, as illustrated in Fig. 10-5. The marker pulse is used to provide index information and is typically 180 electrical degrees. Not all optical encoders provide a marker pulse.

Figure 10-6 shows a 28-mm two-channel incremental optical encoder. The emitter end plate contains two light-emitting diode (LED) light sources (with molded lenses) to form a parallel beam for each channel. The code wheel is a metal disk which, in this case, has 500 equally spaced slits around its circumference. An aperture with a matching pattern is positioned on the stationary phase plate. The light beam

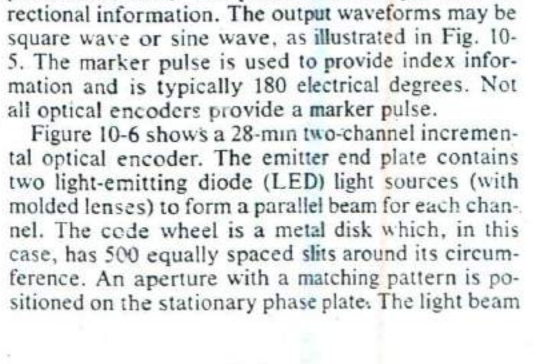


Fig. 10-4 Primary elements of an optical encoder.

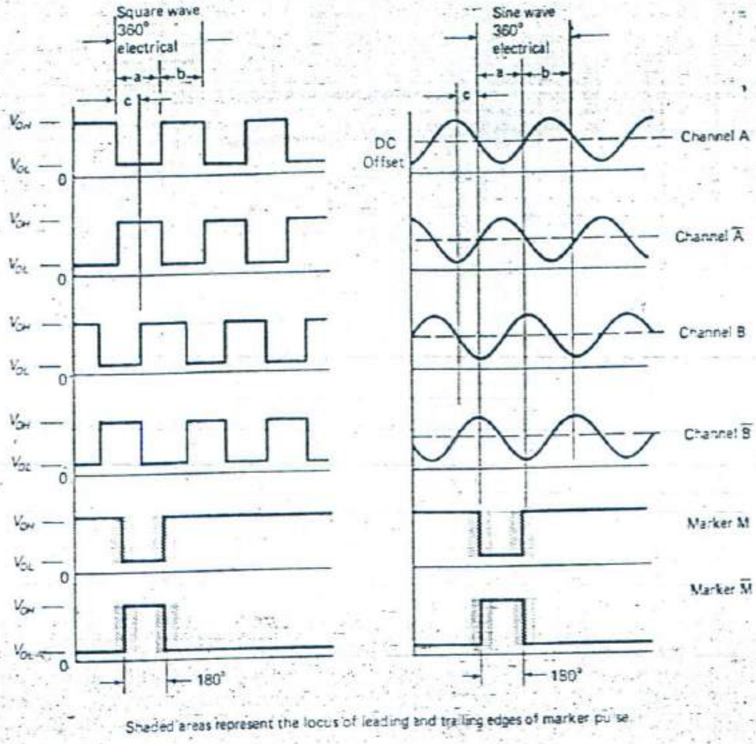


Fig. 10-5 Incremental encoder output waveforms.

is transmitted only when the slits in the code wheel and the aperture line up; therefore, during a complete shaft revolution, there will be 500 alternating light and dark periods. A molded lens beneath the phase plate aperture collects the modulated light into the photodiode detector. Each channel consists of an integrated circuit with two photodiodes and amplifiers, a comparator, and output circuitry as shown in Fig. 10-7. The two quadrature output signals are indicated at their respective channels. The direction of rotation is determined by observing which of the channels is the leading waveform.

Figure 10-8(a) shows a circuit approach to interface to a microprocessor. The logic gates and microprocessors are discussed in the next two chapters. An encoder used to provide position information for a shearing process is shown in Fig. 10-8(b). The material moves the measuring wheel, and the encoder provides pulses to the computer. When the desired count is reached, the shear solenoid is activated.

The *absolute optical encoder* makes absolute position information available in the form of binary, gray, or binary-coded decimal (BCD) formats, depending on the coding pattern of the optical encoder's disk. These codes are covered in Chapter 11. The advantage of the absolute encoder is that it can maintain position information even during a power

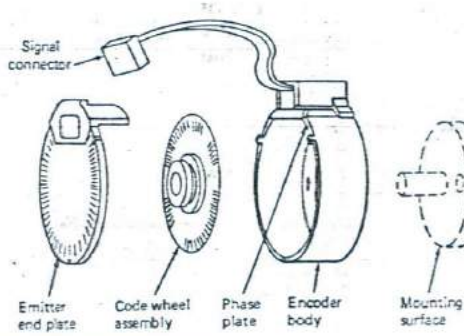


Fig. 10-6 Optical encoder kit.

failure, whereas the incremental type of encoder must be rotated to the index marker for initialization once power is restored. In the absolute encoder, there are signal output lines (one for each bit position) going to the digital processor. An optical encoder disk is illustrated in Fig. 10-9(a). The parallel digital data (four lines in this case) require conditioning and eventually are interfaced to a computer.

Many industrial processes and computer controls also require linear motion inputs. They may be from X-Y tables, plotters, quality control (QC) equipment,

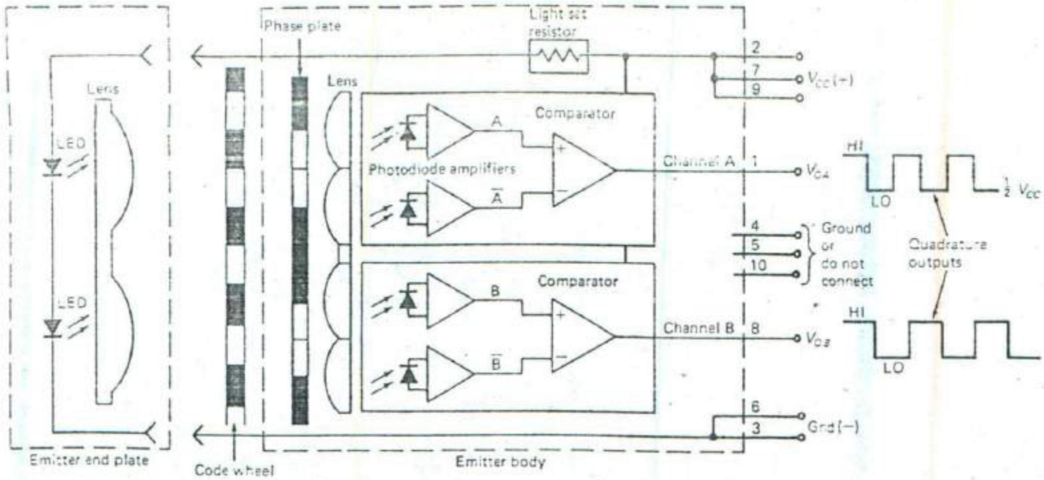


Fig. 10-7 Incremental encoder block schematic.

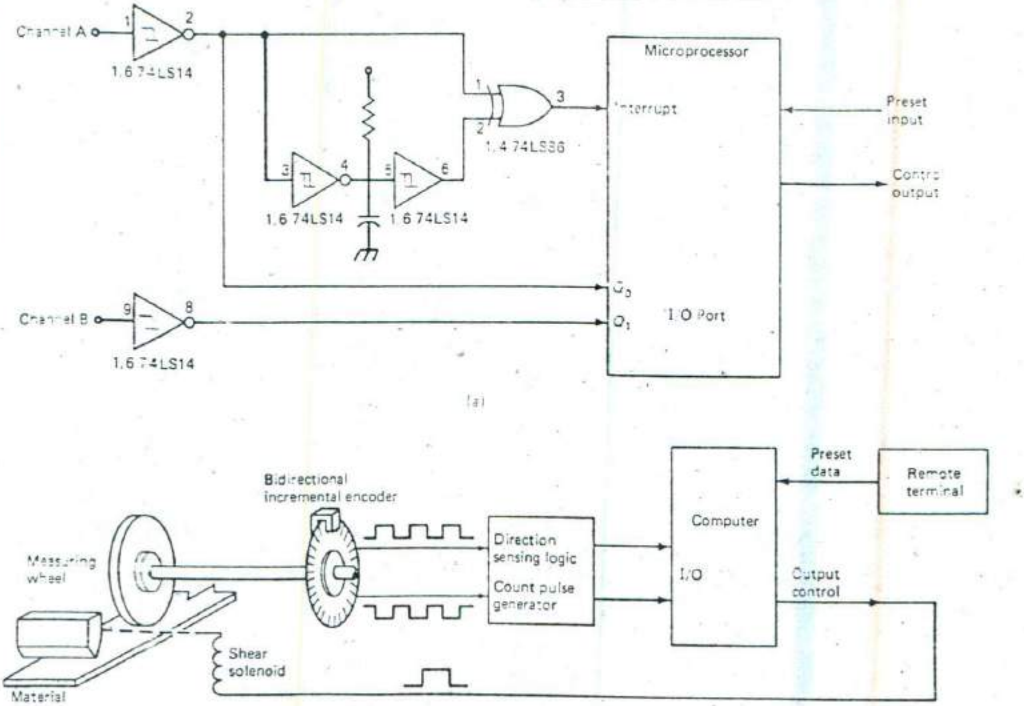
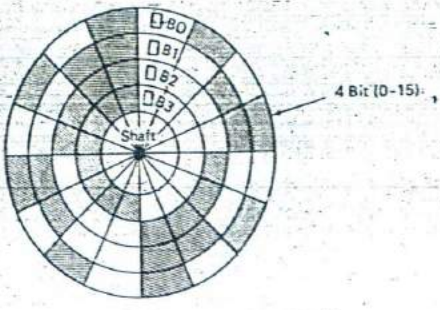


Fig. 10-8 Incremental encoder processing. (a) Electronic processing. (b) Feedback control.

machine tools, etc. These units have a similar type of optoelectronics and also produce quadrature signals with an optional reference just as the incremental encoders do. High-resolution units using special signal electronics provide resolutions as fine as 0.5 μm (2000 pulses/mm) at 300 kHz. A wide variety of linear encoder lengths are also available. Velocity information can also be derived from the incremental signals.

Position sensing varies from application to appli-



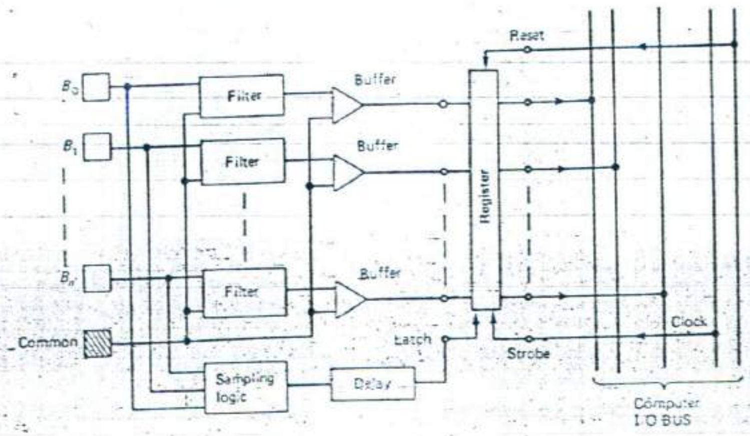


Fig. 10-9 Absolute encoder. (a) Simple encoder disk. (b) Typical encoder interface.

cation. Potentiometers are used for systems that have low resolution requirements. In some cases, a potentiometer is used as a coarse position sensor, and an optical encoder is used for the high-resolution positioning. Areas with heavy contamination (such as oil and dust) are a problem for optical encoders. The problem can be solved with special enclosures.

REVIEW QUESTIONS

1. List three common types of potentiometers.

2. The movable contact on a potentiometer is called the *wiper* or ______.

3. Suppose a 10 kΩ linear potentiometer has a total travel of 350.° What is its nominal resistance from the wiper contact to the far-end contact when it is rotated 270°?

4. The deviation of a potentiometer from its nominal straight-line resistance is rated as a percentage and is known as its ______ specification.

5. A noisy or dirty wiper on a potentiometer can cause a feedback system to ______.

6. Taps may be used to linearize a potentiometer. (true or false)

7. The two outputs from an incremental encoder are ______ degrees apart.

10-2
SYNCHROS AND RESOLVERS

The term *synchro* is a generic name for a family of inductive devices which can be connected in various ways to form shaft angle measurements. All of these devices work on essentially the same principle, which is that of a rotating transformer. A synchro looks like an ac motor and consists of a rotor and a stator. Synchros vary in diameter from 0.5 to 3.7 in. (12.7 to 94 mm).

Internally, most synchros are similar in construction. They have a rotor with one or three windings (depending on the synchro type) capable of revolving inside a fixed stator. There are two common types of rotors: the salient pole and the wound rotor. During one complete cycle, the magnetic polarity of the

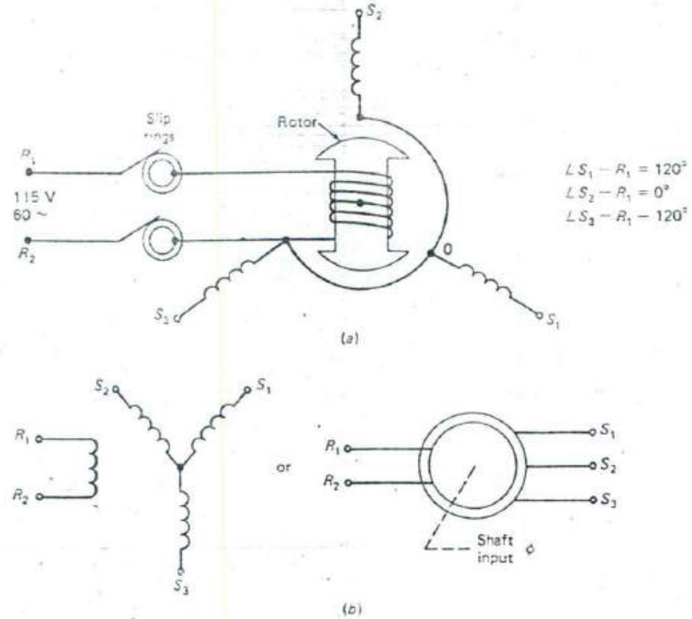


Fig. 10-10 Synchro. (a) Electric circuit. (b) Schematic symbol.

TABLE 10-1 SYNCHRO TYPES

Functional Classification	Military Abbreviations	Input	Output
Torque transmitter	TX	Rotor positioned mechanically or manually by information to be transmitted	Electric output from stator identifying rotor position supplied to torque receiver, torque differential transmitter, or torque differential receiver
Control transmitter	CX	Same as TX	Electric output same as TX but supplied only to control transformer or control differential transmitter
Torque differential transmitter	TDX	TX output applied to stator; rotor positioned according to amount; data from TX must be modified	Electric output from rotor (representing angle equal to algebraic sum or difference or rotor position angle and angular data from TX) supplied to torque receivers, another TDX, or a torque differential receiver
Control differential transmitter	CDX	Same as TDX but data usually supplied by CX	Same as TDX but supplied only to control transformer or another CDX
Torque receiver	TR	Electrical angular position data from TX or TDX supplied to stator	Rotor assumes position determined by electric input supplied
Torque differential receiver	TDR	Electrical data supplied from two TDXs, two TXs, or one TX and one TDX (one connected to rotor, one to stator)	Rotor assumes position equal to algebraic sum or difference of two angular inputs
Control transformer	CT	Electrical data from CX or CDX applied to stator; rotor positioned mechanically or manually	Electric output from rotor (proportional to sine of the difference between rotor angular position and electric input angle)

rotor changes from zero to maximum in one direction, back to zero, then to maximum in the opposite direction, and then back to zero. The wound rotor is used in most synchro control transformers. It often consists of three coils arranged so that their axes are displaced from each other by 120°. One end of each coil terminates at one of the three slip rings on the shaft, and the other ends are connected together.

The stator of a synchro is a cylindrical structure of slotted laminations with three Y connected coils wound with their axes 120° apart. The stator windings are not connected directly to the ac power source. Their excitation is supplied by the ac magnetic field of the rotor.

The synchro may be viewed as a variable coupling transformer. The rotor is energized by an ac voltage, and the coupling between the rotor and the stator windings varies as a trigonometric or linear function of the rotor position. Figure 10-10 shows a synchro schematic diagram. Synchro systems consist of two or more interconnected synchros. Units are grouped together according to their intended function. The seven common types are listed in Table 10-1.

The conventional *synchro transmitter* (TX) uses a salient pole rotor with skewed slots. When an ac excitation voltage is applied to the rotor, the resultant current produces a magnetic field and by transformer action induces voltages in the stator coils. The effective voltage induced in any stator coil depends upon the angular position of the coil's axis with respect to the rotor axis. When the maximum coil voltage is known, the induced voltage at any angular displacement can be determined. Figure 10-11 shows the voltages induced in one stator coil as the rotor is turned to different positions.

The turns ratio between the rotor and stator is such that when single-phase 115-V excitation is applied to the rotor, the highest value of effective voltage induced in any one coil will be 52 V. Because the common connection between the stator coils is not accessible, it is only possible to measure the stator coil-to-coil effective voltage. Figure 10-12 shows how these voltages vary as the rotor is turned. Values are shown above the line when the terminal-to-terminal voltage is in phase with the R_1 to R_2 voltage and below the line when the voltage is 180° out of phase with the R_1 and R_2 voltage. Therefore, negative values indicate a phase reversal. As an example, when the shaft (Fig. 10-12) is turned 30° from the reference (zero degree) position, the S_1 to S_3 voltage will be about 45 V and in phase with the R_1 to R_2 voltage. The S_1 to S_2 voltage will be about 90 V and is 180° out of phase with the R_1 to R_2 voltage. Although the curves of Fig. 10-12 resemble time graphs of ac volt-

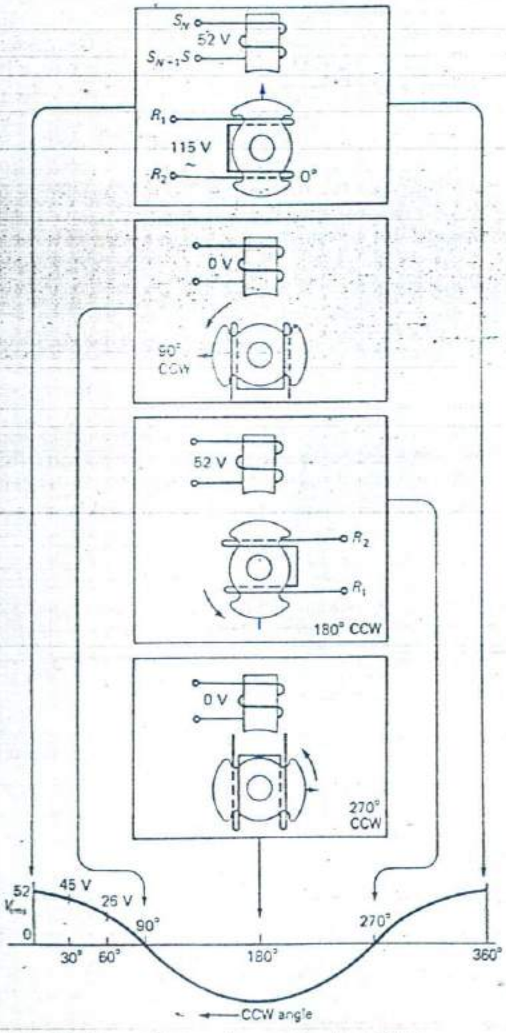


Fig. 10-11 Curve of stator voltage versus rotor position.

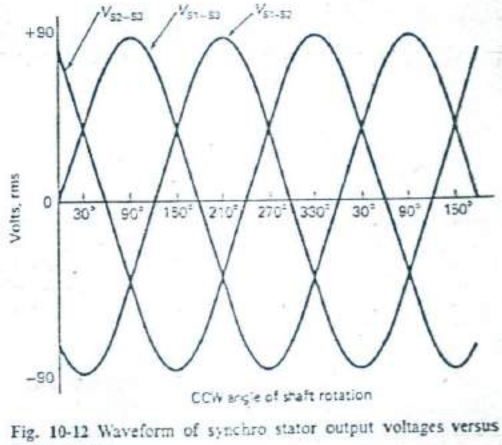


Fig. 10-12 Waveform of synchro stator output voltages versus shaft rotation.

ages, they show only the variations in the effective voltage amplitude and phase as a function of the mechanical rotor position.

It should be noted that the synchro is *not* a three-phase machine or generator. In a three-phase machine, there are three voltages *equal* in magnitude, displaced from each other by 120 electrical degrees. With the synchro, which is a single-phase device, the three stator voltages vary in magnitude, and one stator coil is in phase or 180° out of phase with another coil, as illustrated in Fig. 10-12.

In general, if the rotor of a synchro is excited by 60-Hz or 400-Hz ac (called the *reference voltage*), the voltage induced in any stator winding will be proportional to the cosine of the angle between the rotor coil axis and the stator axis, as was indicated in Fig. 10-11. The voltages induced across any pair of stator terminals (S_1 to S_2, S_1 to S_3, S_2 to S_3) will be the sum or difference, depending on the phase, of the voltages across the coils measured.

For example, if a reference voltage $V \sin(\omega t)$ excites the rotor of a synchro (R_1 to R_2), the stator terminals will have a voltage of the following form:

$$V(S_1 \text{ to } S_3) = V \sin(\omega t) \sin\theta$$
$$V(S_1 \text{ to } S_2) = V \sin(\omega t) \sin(\theta + 120)$$
$$V(S_2 \text{ to } S_3) = V \sin(\omega t) \sin(\theta + 240)$$

where θ = synchro shaft angle.

Note: The expression $V \sin(\omega t)$ predicts the instantaneous voltage of a sine wave at time (t) where V represents the maximum voltage and $\omega = 2\pi f$. The rms stator voltages are given by:

$$V_{rms}(S_1 \text{ to } S_3) = 0.707\, V \sin\theta$$
$$V_{rms}(S_1 \text{ to } S_2) = 0.707\, V \sin(\theta + 120)$$
$$V_{rms}(S_2 \text{ to } S_3) = 0.707\, V \sin(\theta + 240)$$

These voltages are known as the *synchro format voltages* and will be referred to as such from now on. Synchros are divided into two basic types, torque synchros and control synchros. *Torque synchros* are required if it is necessary to transmit angular displacement information from a shaft of one synchro to the shaft of another synchro without using any additional amplifiers or gearing. The two most common torque synchros connected in a repeater system are the *torque transmitter* (TX) and the *torque receiver* (TR). Figure 10-13 shows a TX and TR connected as a repeater system. In the repeater system, the rotor of the transmitter (TX) is excited with a reference voltage to produce the synchro format voltages on output terminals S_1, S_2, and S_3. The stator voltages induced in the receiver (TR) stator coils as a result of its rotor excitation will be equal to the voltages induced by the transmitter stator current. In this balanced condition, shown in Fig. 10-13, there is no current flow in the stator coils or in the stator interconnections. The total current drawn is that used by the excitation of the two rotors. Therefore, the transmitter will supply current only when the receiver rotor is out of alignment with the transmitter rotor. These repeater systems are accurate to ±1° and are used in systems in which a rotating device's output is required to position a remote pointer. Today, most of these units have been replaced by a synchro-to-digital converter driving an LED display of the digital readout of angular position. There also exists a digital-to-synchro transmitter to convert digital data to synchro format voltages to drive a remote electromechanical pointer. A mix of these devices is not uncommon. Digital synchro converters will be covered shortly.

Two other members of the torque synchro family are worth mentioning. The first device is the *differential synchro transmitter* (TDX). The stator is quite similar to that of the transmitter and receiver just discussed. It has three sets of coils wound around the stator frame to produce poles 120° apart. The rotor is quite different from those of conventional synchro units. Electrically, it has three sets of coils

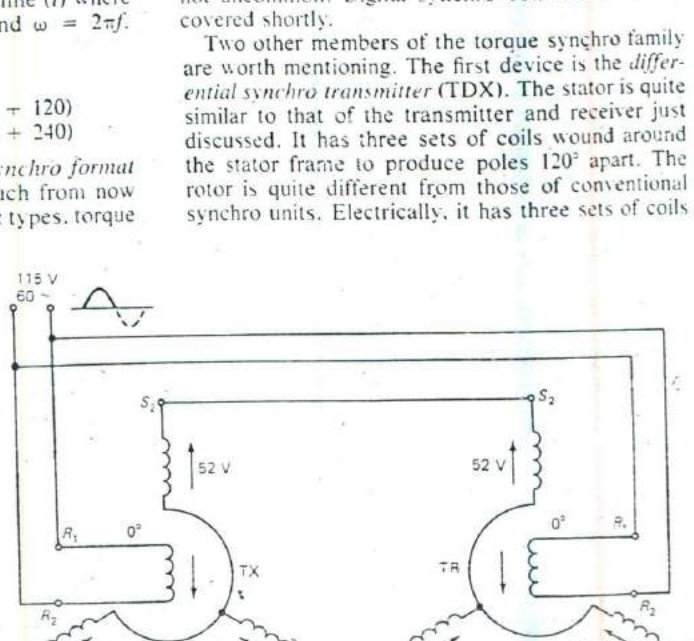


Fig. 10-13 Synchro transmitter and receiver pair with both rotors at 0°.

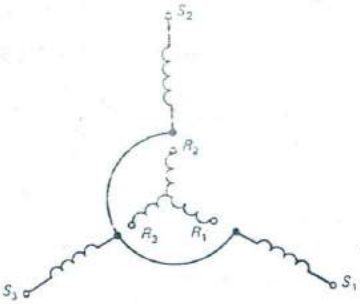


Fig. 10-14 Schematic diagram of differential synchro.

wound in slots equally spaced on the rotor and connected to produce poles 120° apart. The schematic of the differential transmitter is shown in Fig. 10-14.

The TDX usually obtains its input from a torque transmitter (TX) and produces electrical synchro format signals. The power comes from the synchro stator outputs (there is no reference applied to the TDX). The TDX may be connected to add or subtract two inputs. Figure 10-15(*a*) shows a connection for subtraction. If a mechanical input of 75° is applied to the TX and its output signals go to the TDX stator, the TDX subtracts its own mechanical input (30°) and transmits the results to the TR, which indicates the system's mechanical output by position of its rotor (45°), as indicated in Fig. 10-15(*a*). In some cases, the system is set up for addition. This is done by reversing the S_1 and S_3 leads from the TX to the TDX stator, and from the TDX rotor to the TR stators. This will result in the behavior shown in Fig. 10-15(*b*). With the same mechanical inputs of 75° and 30°, the receiver will provide an output equal to the sum of signals by turning to 105°. A wiring schematic for the subtraction system is shown in Fig. 10-16. The torque synchro system is suitable only for very light loads and is never really accurate. In addition the torque system places a drag on the associated equipment it is measuring.

When larger amounts of power and more accuracy are required, torque synchros give way to the *control synchros*. These devices are used for providing and handling control signals to a servo power amplifier when more power and accurate angular displacement of a large load are required. The control synchros are not designed to handle any mechanical load. The two most common control synchros are the *control transmitter* (CX) and the *control transformer* (CT). The CT develops an ac rotor output voltage that is proportional to the relative shaft angles between the synchro transmitter and the control transformer. The devices are normally connected as shown in Fig. 10-17. The output of the CX (transmitter) is fed to the stators of the CT (transformer). The CT is a high-impedance version of the torque receiver with its

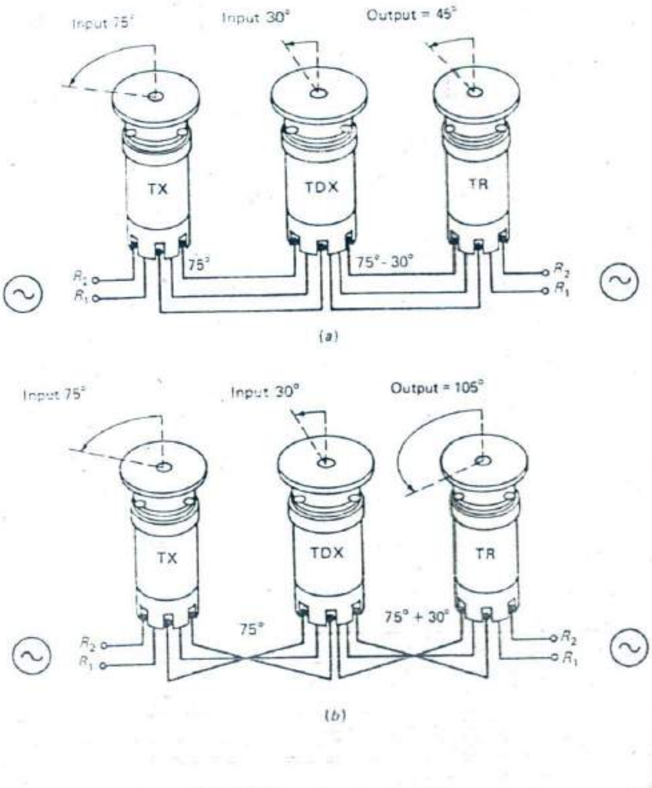


Fig. 10-15 Applications of the differential synchros. (a) Subtraction with TDX. (b) Addition with TDX.

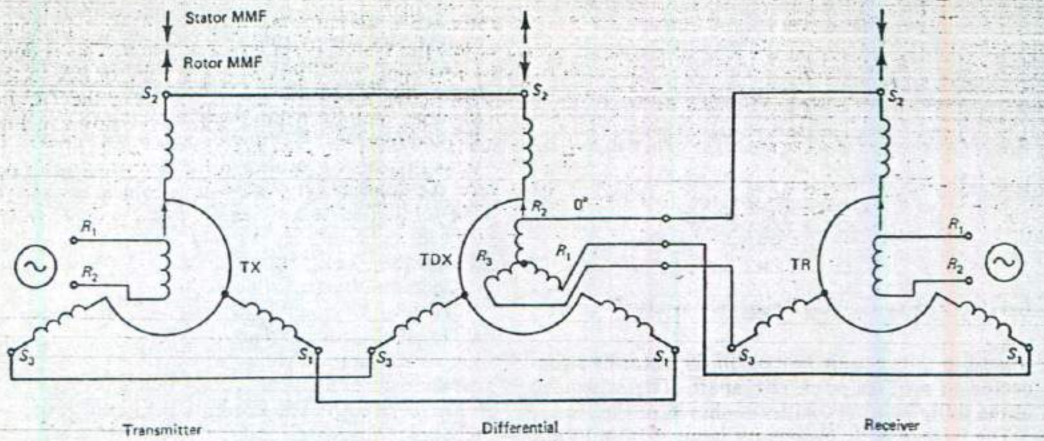


Fig. 10-16 Schematic diagram of a subtraction TDX system.

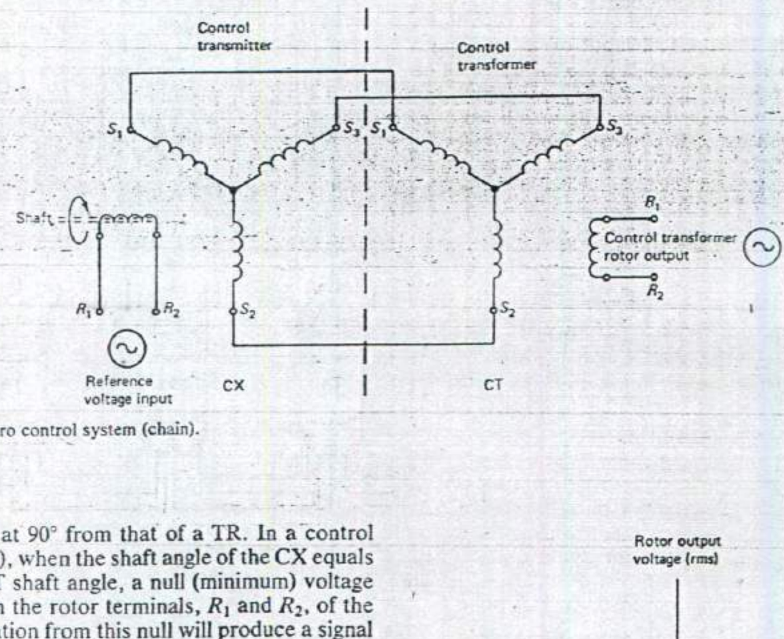


Fig. 10-17 Synchro control system (chain).

rotor aligned at 90° from that of a TR. In a control system (chain), when the shaft angle of the CX equals that of the CT shaft angle, a null (minimum) voltage will appear on the rotor terminals, R_1 and R_2, of the CT. Any variation from this null will produce a signal in the CT rotor whose phase will depend on in which direction it is moved off null. Figure 10-18 shows the output of a CT rotor as it travels near alignment (null) with the transmitter rotor. Typically, for a 115-V CT the null voltage would be about 30 mV rms.

A simple closed loop servo system using a CX and CT control system is shown in Fig. 10-19. When the shaft of the CX is turned to some angle, the S_1, S_2, and S_3 outputs provide the synchro format voltages previously mentioned. These voltages are transmitted to the CT stators S_1, S_2, and S_3. If the CT is not at the input angle θ, a voltage will be produced at the output of the CT rotor winding. This signal (error)

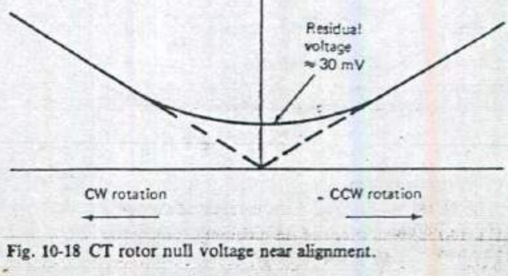


Fig. 10-18 CT rotor null voltage near alignment.

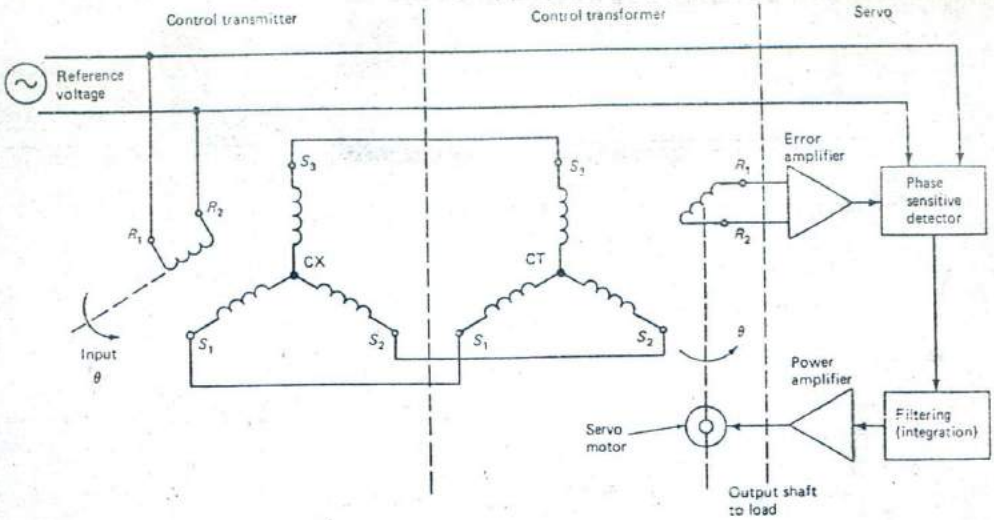


Fig. 10-19 Simple servo system using control synchros.

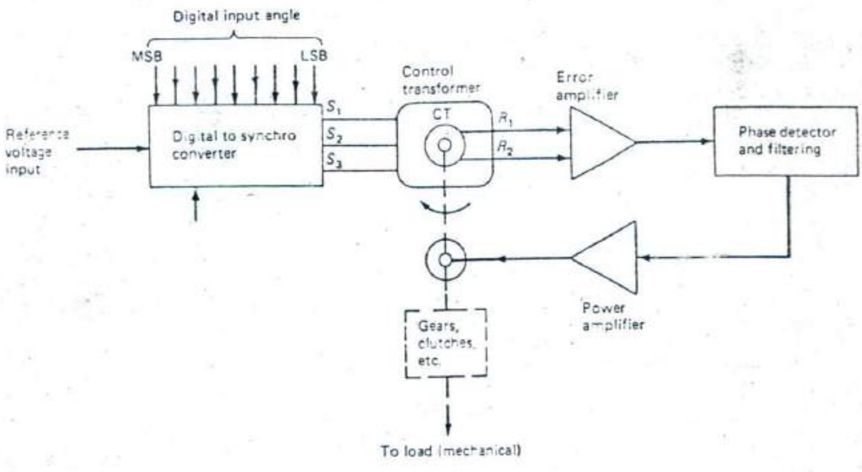


Fig. 10-20 Digital-to-synchro converter.

is amplified, phase-detected and fed to a servo amplifier to cause a servomotor to position a load and the CT shaft to a position where the CT rotor output is minimum (null). The direction in which the motor turns toward angle θ is determined by the phase of the CT rotor signal with respect to the reference voltage.

In some later applications, a digital-to-synchro converter (a solid-state CX) can be used, with the input angle in digital form as shown in Fig. 10-20. There also exists a *control differential transmitter* (CDX), which is the control equivalent of the TDX previously discussed and is used to add or subtract an additional shaft angle.

The resolver is basically a trigonometric function generator that resolves for an angle θ, the hypotenuse, or sides of a right triangle. Although it resembles a synchro device outwardly, internally the resolver is quite different. There are a wide variety of winding and ratio configurations available. The most common has two isolated primary (rotor) windings at right angles to each other. The two windings of the stator are isolated and are also placed 90° apart. A *resolver* is illustrated in Fig. 10-21. It solves the unknowns of a right triangle. If θ is the shaft input angle, one can determine B and A directly from a resolver by applying voltage C as an input to the stator, positioning the rotor to angle 6, and reading A and B as outputs

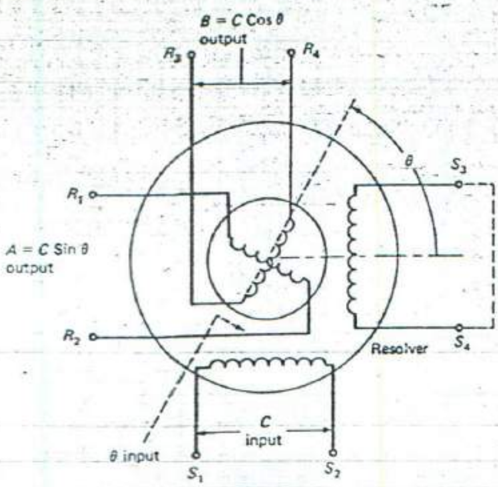


Fig. 10-21 Schematic diagram of a resolver. Either rotor or stator can be the primary.

from the rotor windings. To be more mathematically correct, if the stator is excited with $V \sin \omega t$, the resolver format voltages will be

$$V_{R1-R2} = V \sin \omega t \sin \theta$$
$$V_{R4-R3} = V \sin \omega t \cos \theta$$

For example,

$$V_{R1-R2} = 100 \times 0.500 = 50 \text{ V}$$
$$V_{R3-R4} = 100 \times 0.866 = 86.6 \text{ V}$$

where

$$\text{resolver shaft angle} = 30^\circ$$
$$V \sin \omega t = 100 \text{ V}$$
$$\sin \theta = \sin 30^\circ = 0.500$$
$$\cos \theta = \cos 30^\circ = 0.866$$

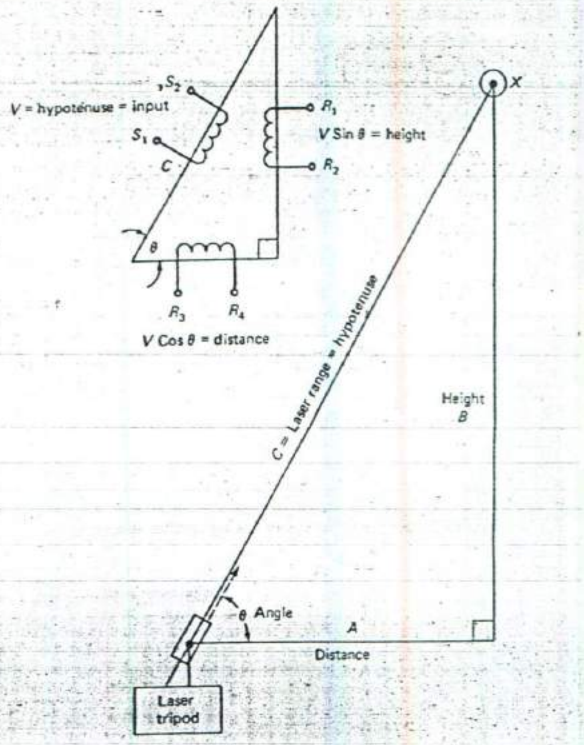


Fig. 10-22 Example of computing resolver for determining height.

The unused stator rotor is usually short-circuited. These voltages represent the rectangular or cartesian coordinates of the point.

Figure 10-22 is an example of polar-to-cartesian conversion used to determine the precise height of

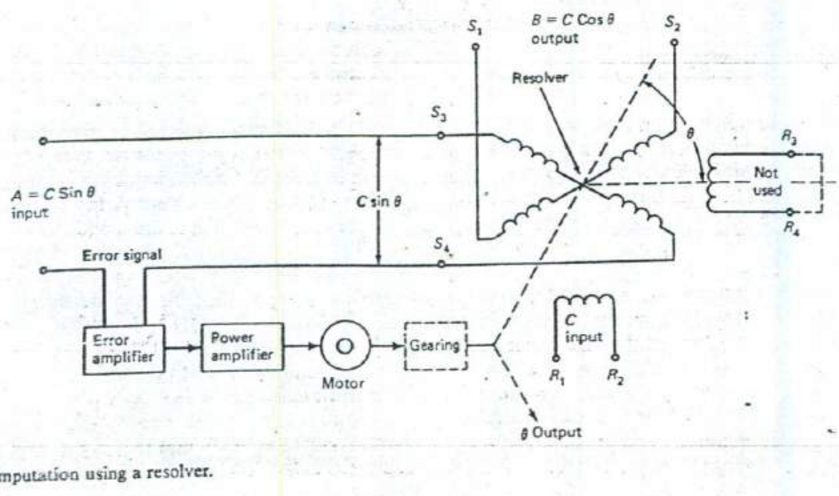


Fig. 10-23 Computation using a resolver.

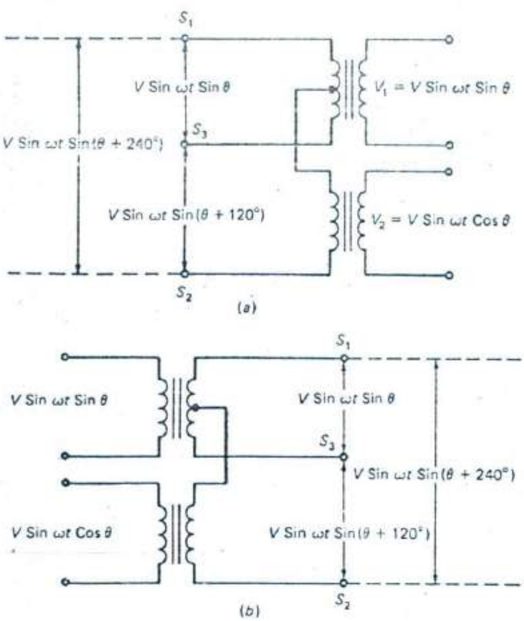


Fig. 10-24 Scott connected transformers.

an object by using an adjustable laser. The output voltage $VR_1 - VR_2$ is proportional to the height, and output $VR_3 - VR_4$ is proportional to the ground distance when a voltage V (equal to the laser range) is applied to $S_1 - S_2$, and the resolver is rotated to an angle θ. A servo system to perform this function is shown in Fig. 10-23. The servomechanism is satisfied only when the resolver angle is equal to θ, which occurs when the input to the error amplifier is at a null. One output is $B = \cos\theta$, which is the voltage proportional to the height of the Fig. 10-22, when a voltage C (equivalent to the laser range) is applied to the stator as shown.

By using all four windings of the resolver, a two-dimensional space problem can be solved. If the stators are excited with a voltage representing X and Y, respectively, when the shaft is positioned to an angle ψ, the voltages produced at the rotor terminals will be

$$V_{R1-R3} = X\cos\Psi + Y\sin\Psi$$
$$V_{R4-R2} = Y\cos\Psi - X\sin\Psi$$

The applications of these mathematical functions are commonly used in guidance and robot control systems.

It is possible to convert synchro input signals into resolver format signals and to convert resolver format signals into synchro format signals. These conversions are generally done by using Scott connected or Scott T transformers. In many cases this is or was done to upgrade a system to use synchro-to-digital (SDC) or digital-to-synchro converters (DSC), which use signals in the resolver format. Figure 10-24 shows two common Scott transformer connections. In Fig. 10-24(a) the synchro-to-resolver connection and the resolver output format voltages obtained are shown. In Fig. 10-24(b) the resolver-to-synchro Scott connected transformers, which are simply the inverse of the synchro-to-resolver case, are illustrated.

REVIEW QUESTIONS

8. The stator windings S_1, S_2, and S_3 of a synchro are connected to a three-phase power source. (true or false)

9. Referring to Fig. 10-11, the voltage at the 180° point should be ______.

10. In Fig. 10-12, the voltage of S_1, S_2 at 270° is out of phase with R_1—R_2 voltage. (true or false)

11. Referring to Fig. 10-13, if TX is moved clockwise, TR will move ______.

12. If the differential synchro system of Fig. 10-15(b) has an input of 130°, the receiver will be at ______.

13. A ______ transformer is used to input angular signals into a servo power system.

14. The control transformer error signal is composed of a voltage and ______ signals.

10-3 SERVOMOTORS AND RATE GENERATORS

The requirements for most small servomechanisms are met by ac two-phase induction motors. Their mechanical output power varies from 0.5 to 100 W. Above 10 W, most two-phase servomotors are cooled by a separate motor-driven blower included in the same housing with the servomotor. A 10-W frame will deliver about 25 W output with the added blower cooling. Direct current servomotors vary in size from 1/20 hp to many horsepower and are generally used in large power servomechanisms.

A typical connection for a two-phase motor is shown in Fig. 10-25. A voltage V_M is applied to the main (fixed or reference) winding. Voltage V_C is supplied from the controller, which is usually an amplifier. The magnitude of V_C is a function of the degree of action required of the motor. The windings are usually identical and equally rated. The voltages V_M and V_C must be in synchronism and are derived from the same ac source. They must also be in time quadrature, which may be produced by introducing a 90° phase shift in the amplifier or by connecting a suitable capacitor in series with the main (reference) phase V_M. When V_C has a voltage value leading V_M by approximately 90°, rotation in one direction is obtained; when V_C lags V_M, rotation in the other direction will occur. Since torque is a function of both V_M and V_C, changing the magnitude of V_C changes the developed torque of the motor. Some

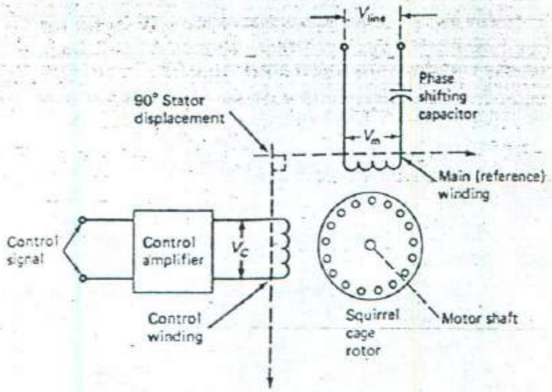


Fig. 10-25 Conventional two-phase servomotor circuit.

servomotors are designed with a center-tapped winding to be fed from a push-pull output amplifier as shown in Fig. 10-26.

The speed-torque characteristic curves for a typical ac servomotor are shown in Fig. 10-27. They are typical of servomotor characteristics. The torque is large at zero speed to aid in servo static sensitivity, to give internal damping for the servomechanism, and to prevent single phasing (the tendency of the rotor to continue to rotate when one winding is opened and the other winding remains excited) of the servomotor. Two-phase servomotors are inherently high-speed, low-torque devices and are geared down to drive the load.

Although two-phase servomotors are available in a wide variety of configurations, the most popular type has a squirrel cage rotor with a low ratio of rotor-to-fame diameter and high rotor resistance. This type gives the best overall performance and is efficient for converting input watts to shaft torque. There are other motor configurations that are used in specific applications. These include drag-cup motors and solid iron rotors.

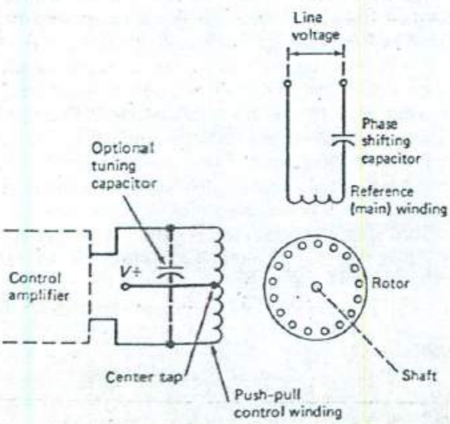


Fig. 10-26 Center-tap (push-pull) two-phase servomotor.

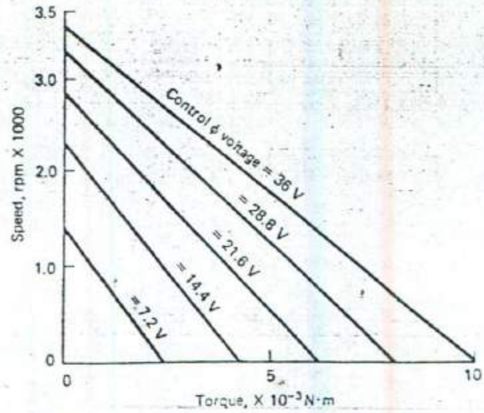


Fig. 10-27 Torque-speed characteristics of a typical servomotor.

The drag-cup motor shown in Fig. 10-28 is constructed with its rotor made of copper, aluminum, or an alloy. For a given size and weight as compared to the squirrel cage motor, it generally has lower torque. The heavy iron laminations are stationary, only the lightweight cup rotates, and the inertia is very small. The solid iron rotor core is used to carry both flux and induced rotor currents. This motor was designed for operation from the output of vacuum tube amplifiers and is rarely encountered today. The two-phase induction motor consists of two input windings (coils in slots of a laminated-iron structure) spaced 90 electrical degrees apart. Under a balanced condition, the windings are excited with equal voltages, 90° apart in time phase. The motor currents therefore generate magnetic fields in the air gap which are also in space and time quadrature. As with the induction motors, the rotating speed N_S is

$$N_S = \frac{120f}{p}$$

where f = frequency of line
p = number of poles

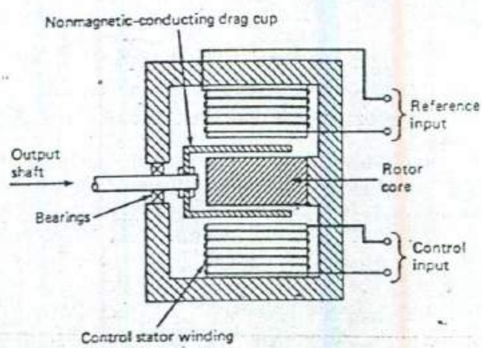
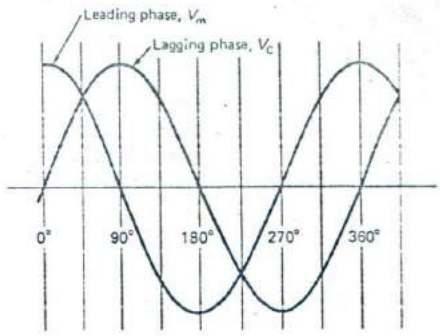


Fig. 10-28 Drag-cup two-phase servomotor.



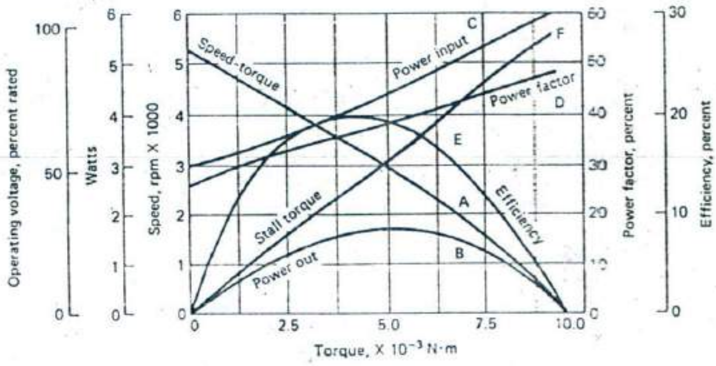


Fig. 10-29 Waveforms and characteristic curves for two-phase servomotor.

A 4-pole, 60-Hz winding causes the resultant field to rotate at 1800 rpm, and a 4-pole 400-Hz winding rotates at 12,000 rpm. Figure 10-29(*a*) shows the two-phase stator currents which develop the rotating field in a 2-pole machine. This rotating flux field induces a voltage in the rotor conductors with a magnitude proportional to the relative speed. The rotor voltages in turn cause currents; as a result a torque is developed by interaction of the current-carrying conductors and the rotating field. This drags the rotor along after the synchronous field of the stator. Since the rotor must overcome friction, it cannot reach synchronous speed. The difference between actual and synchronous speed is known as *slip*:

$$\text{Slip} = \frac{\text{synchronous} - \text{actual speed}}{\text{synchronous speed}}$$

In servomotors, the no-load speed is approximately five-sixths synchronous speed, corresponding to a slip of one-sixth, or 16.7 percent.

Figure 10-29(*b*) presents the characteristics of a typical servomotor from no load to the stalled condition. Curve *A* shows the variation of torque for rated phase voltage and varying load. As previously shown in Fig. 10-27, it is nearly linear. Curve *B* shows the power output and has a parabolic shape which peaks near one-half of the no-load speed. Curves *C* and *D* show the power input and the power factor, respectively. Curve *E* shows that the efficiency peaks near one-half of the no-load speed. Curve *F* plots the variations of stall torque as a function of control voltage with the fixed-phase voltage constant. This curve is linear and is a measure of the motor's stiffness. A servomotor should develop torque with a minimum amount of input wattage.

In ac servomotors, as mentioned earlier, there has to be a phase shift of the voltage on the main (reference) winding with respect to the control winding. There is no simple method of maintaining this phase shift for all motor speeds. Some designs use the two-capacitor method shown in Fig. 10-30, which gives good results on small servomotors.

A dc motor can be controlled by varying either the field current or the armature current. The types of dc servomotors are the series motor, the shunt motor, and the permanent magnet (PM) motor. These motors offer higher efficiency than an ac motor of the same size, but radio frequency interference (RFI) is a problem in some applications.

Most of the dc servomotors used in low-power applications are of the PM type. The ease of con-

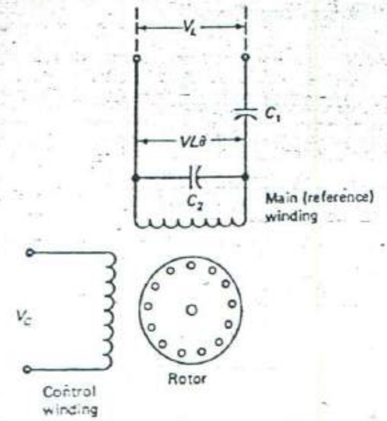


Fig. 10-30 Two-capacitor method of phase shifting main excitation.

trollable speed, along with the linear torque-speed control curve, makes the PM motor ideal for servomechanism applications. The characteristic curves are shown in Fig. 10-31. The speed-torque curve is quite similar to that of the ac servomotor presented earlier in this section. These motors are available in 6-, 12-, and 24-V models, making them applicable to solid-state circuitry. By comparison, the dc motor has some advantages over the ac motor. The dc motor inertia is greater than that of the ac motor (inertia is covered in the next section). This greater inertia is due to the wound armature and commutator, which produce a heavier rotor. The dc motor does not require any standby power; however, the ac servomotor continuously draws power for its main (reference) winding.

Figure 10-32 shows a modern dc servomotor that is only one-third as thick as a conventional motor. Such motors are popular in compact systems, such as numerical control and robotics systems, and in automatic control machines. At 24 V and 3.8 A, it develops 1.5 kilogram-centimeters (kg · cm) of torque. Larger units with up to 4.5 kW of output power are available.

The dc servomotor in some modern servomechanisms may be one of the brushless types (covered previously) which lend themselves to easy computer control. Along the same lines, the stepping motor has become a valuable type to be used as a servo-type motor. This type of application will be covered in a later section.

In the past, large dc servo systems employed constant armature current, with the control signal applied to the field winding. These systems operated with vacuum tubes, thyratrons, or amplidynes providing the field drive. The power required for field control is only a fraction of the power required for armature control. Power amplification of the order of 20,000:1 can easily be obtained.

A *rate generator* (tachometer) is an electromechanical device resembling a small motor, which produces an output voltage that is proportional to its shaft speed that can be read out or used for closed loop speed control or stabilization. The dc rate generator is usually separately excited (shunt-wound) or is a permanent magnet generator. Though any dc generator can be used with a calibrated meter to indicate speed, more precise units are required for control-system applications. Special consideration is given to certain electrical and mechanical characteristics. The electrical output signal should be a noise-free voltage that varies linearly with speed. Mechanically, the unit should run quietly and smoothly, with low drag and low inertia. The relationship between voltage and speed can be expressed mathematically as

$$V = KN$$

where N = rate generator's speed
K = constant of proportionality between the voltage and speed

Most dc rate generators have an output sensitivity of about ½ V/100 rpm. Other important factors that differ from those of a standard dc machine are func-

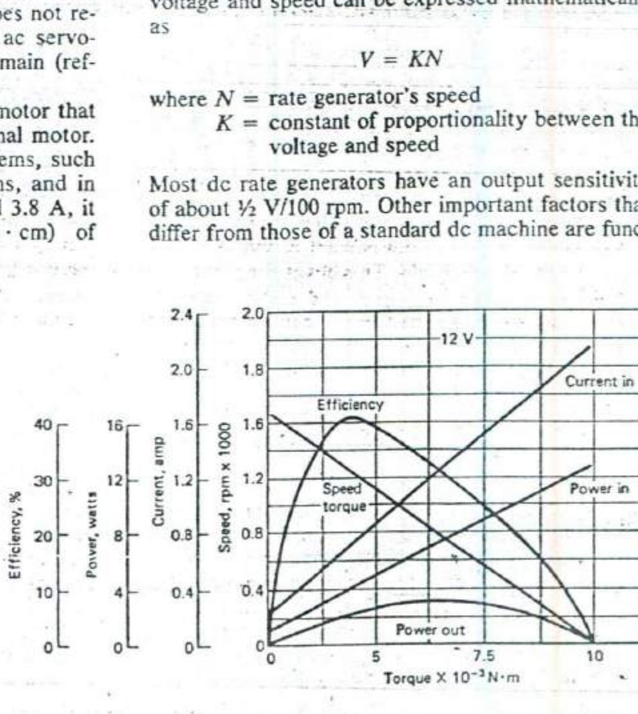


Fig. 10-31 Characteristic curves for a dc servomotor.

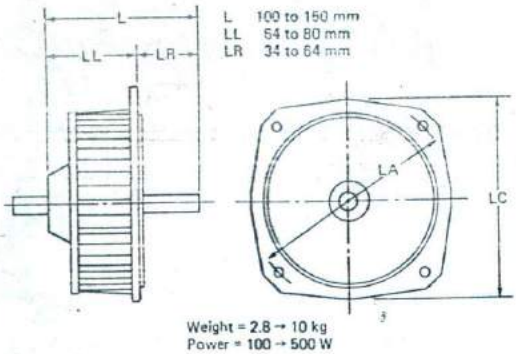


Fig. 10-32 Space saver dc servomotor.

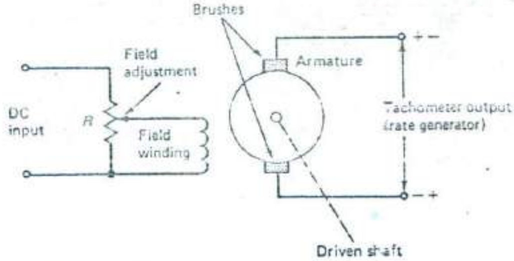


Fig. 10-33 Separately excited dc rate generator.

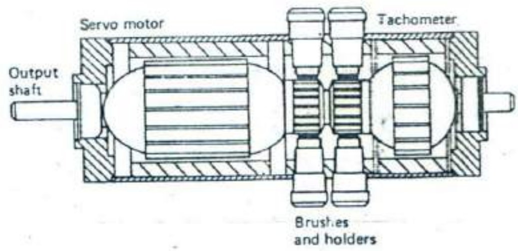


Fig. 10-34 Servomotor and rate generator on common shaft.

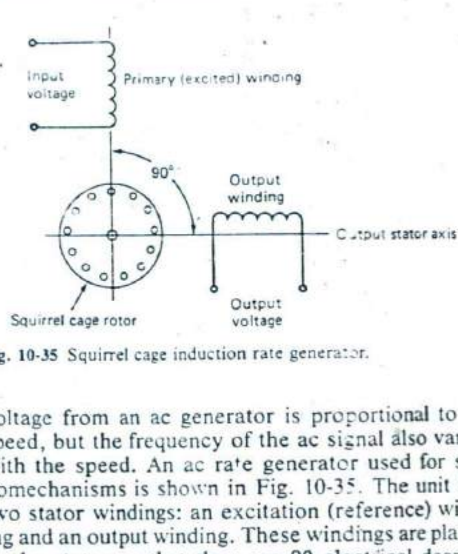


Fig. 10-35 Squirrel cage induction rate generator.

tions of the mechanical and electrical quality. Cogging is prevented by skewing the armature slots, and friction is reduced by limiting brush pressure and by using precision high-quality bearings.

Some of the advantages of the dc rate generator are the following:

1. Freedom from waveform and phase-shift problems
2. Absence of any residual zero-speed voltage (present in ac units)
3. High output gradients: up to 20 V/1000 rpm
4. Easier temperature compensation than in an ac rate generator

Some of the disadvantages of using dc rate generators are the following:

1. Brush problems: contact, vibration, arcing, and breaking contact. The position of the brushes must be exactly at neutral for reversible units.
2. Noise generation: filtering of high-frequency brush commutation noise is necessary.
3. Output ripple is undesirable and must be attenuated in some systems.
4. Brush friction and hysteresis effects require a higher driving torque.

The selection of a separately excited or a permanent magnet rate generator is based on factors such as line-voltage variations, ambient temperature variations, and the possiblity of demagnetization of the PM type. The output of the separately excited rate generator can be line-voltage-compensated, as indicated in Fig. 10-33, whereas the output of a PM type can only be attenuated.

Since a majority of rate generators (tachometers) are closely connected to the servomotor, many devices combine the motor and the tachometer on the same shaft. An example is shown in Fig. 10-34.

The ac rate generator must produce a sinusoidal signal output of *constant frequency* whose amplitude is proportional to its speed. The amplitude of the voltage from an ac generator is proportional to its speed, but the frequency of the ac signal also varies with the speed. An ac rate generator used for servomechanisms is shown in Fig. 10-35. The unit has two stator windings: an excitation (reference) winding and an output winding. These windings are placed in the stator so that they are 90 electrical degrees apart. Because these coils are at right angles to each other, no output voltage is induced when the rotor is stationary. When the rotor turns, the flux produced by the eddy currents in the rotor is along the secondary axis and produces a voltage at the reference frequency in the output winding. The magnitude of the output voltage is proportional to the rotational speed. The direction of shaft rotation is indicated by the phase of the output voltage (compared to the reference voltage). If the output voltage is in phase with the reference, the direction is said to be *positive*.

If the output is 180° out of phase, the direction is said to be *negative*. The sensitivity of ac rate generators ranges from 1/10 to 1 V/100 rpm, with output impedances of 100 to 1000 Ω. The output voltage is mathematically expressed as it is for the dc rate generator.

There are three types of errors associated with the induction-type ac rate generator: residual voltage at zero speed, nonlinearity, and voltage and phase errors at low speed. The zero speed output has a detrimental effect on a servo system's performance. The fixed and variable residual components as a function of rotor position are shown in Fig. 10-36(*a*). These residual voltages are minimized by the manufacturer by using precision machining techniques. The residual output voltage must be considered when replacements are required. Figure 10-36(*b*) shows a simple compensation network that may be employed to cancel the fixed component of the residual voltage.

If an ac rate generator is directly coupled to the servomotor shaft it adds directly to the inertia. For this reason, the drag-cup rate generator finds wide application. The drag-cup rate generator consists of the same type of stator (two-winding) as the two-phase induction rate generator just discussed. Its rotor consists of a thin nonmagnetic conducting material of aluminum or copper. Drag-cup rotors yield maximum uniformity. This is accomplished by various manufacturing techniques and by having matched temperature coefficients for all mating parts.

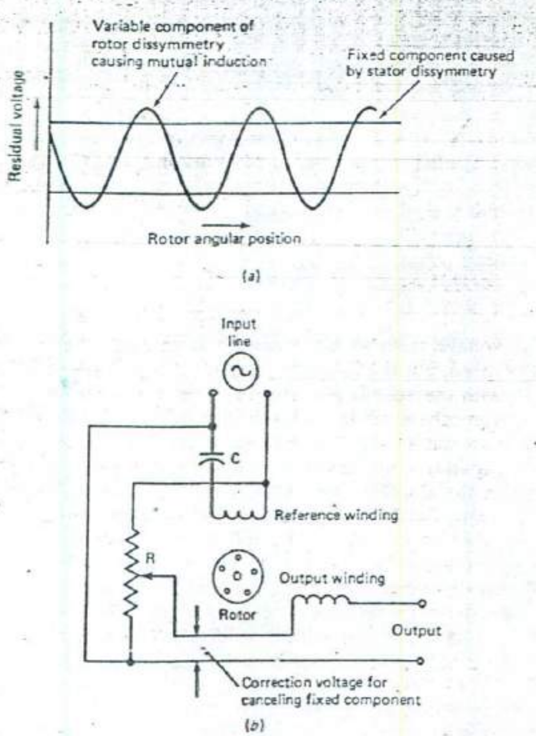


Fig. 10-36 Fixed and variable components of residual voltage and canceling circuit.

No output voltage is induced when the drag-cup is stationary. Upon rotation, the eddy currents induced in the rotor cup distort the path of the flux so that a voltage proportional to shaft speed appears at the output winding. A notch or some other dissymmetry may be added to the rotor to cancel the inherent output dissymmetry.

A wide variety of incremental encoders are available and are being integrated into servomechanisms as rate generators. These devices are discussed in Chapters 3, 13, and 14. In many cases their usage is dictated by the environment of the application.

REVIEW QUESTIONS

15. The phase relationship between the reference voltage and the control voltage in a two-phase servomotor is ______ degrees.

16. The reference voltage is usually shifted by use of a ______.

17. Single phasing of a servomotor is a method of braking. (true or false)

18. The rotor of a two-phase servomotor is either a squirrel cage or a ______.

19. A 2-pole, 400-Hz motor will rotate at ______ rpm.

20. If the motor in question 19 rotates at 20,000 rpm, the slip is ______ percent.

21. The resistor in Fig. 10-36(*b*) is used for phase-shifting purposes. (true or false)

10-4
MECHANICAL COMPONENTS

A servomechanism will usually contain mechanical components such as gears, couplings, bearings, limit stops, and clutches. These parts are manufactured to tight tolerances for ease of assembly and for conformance with performance specifications.

Couplings are used to connect the ends of two shafts together so they always rotate at the same speed with the same angular position. Most couplings fall into one of the four following categories:

1. Rigid (sleeve) coupling
2. Flexible three-piece (Oldham)
3. Flexible-bellows, spring, one piece
4. Flexible-sleeve, one piece
5. Universal joints

The *rigid (sleeve) coupling*, as the name implies, is a one-piece coupling that rigidly couples two shafts together. Each end has a set screw to secure the sleeve to each shaft. One type requires that both shafts be of exactly the same diameter; no misalignment is allowed. Another type of adapter coupling allows mating of two different diameter shafts. In specifying couplings, the letters *OD* stand for outside diameter, and the letter *B* indicates the bore (inside diameter). Dimensions are specified in inches and millimeters by most manufacturers. Mixing of sizes, that is, inches with millimeters, though it may seem tolerable, is not good practice. The set screws usually require use of an allen wrench for any adjustment. These will also be in inch or metric sizes and will require the appropriate tool. Some very-high-speed high-torque applications will use a recessed inner set screw with a second set screw on top.

A *flexible three-piece coupling* is also known as an *Oldham-type coupling*. This coupling allows for a slight angular or lateral misalignment of the shafts being coupled. It consists of two end hubs with machined surfaces to receive the center interlocking floating member. The floating (center) member, called the *torque disk*, is usually not metallic, but a plastic material such as delrin. The machined parts must be handled with care, and distortion will cause backlash in the coupling. *Backlash* is the play or lost motion that occurs between two loosely fitting parts.

The *bellows* or *spring coupling* consists of two hubs connected by a flexible metal bellows or a spring. It also allows for shaft misalignment. However, flexing of the bellows can cause metal fatigue, so these couplings are found only in low-torque applications. The spring coupling is preferred for high-speed applications and acts as a shock absorber.

A *flexible one-piece coupling* consists of a flexible element that may be polyurethane, rubber, or neoprene. This coupling can accommodate shafts that are out of alignment by as much as 1 in. (25.4 mm). They are quiet-running and absorb end play.

Universal joints can be either single or double types, as indicated in Fig. 10-37. The *single joint* will operate at angles up to 30°, and the *double joint* can couple shaft angles approaching 90°. Figure 10-37(*c*) shows a typical position for each type of universal joint.

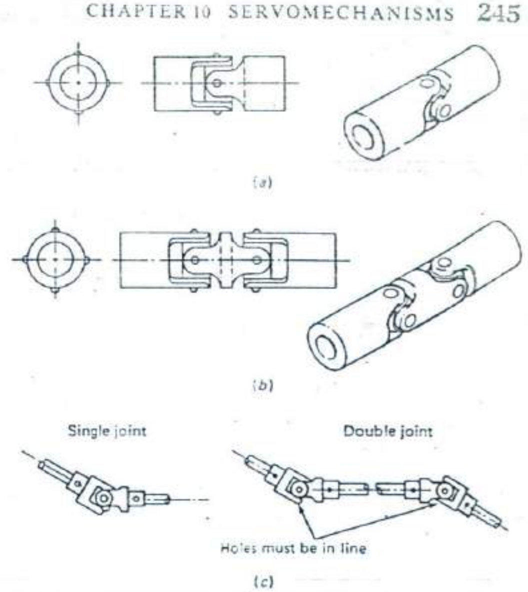


Fig. 10-37 Universal joints. (a) Single. (b) Double. (c) Typical applications.

The *clutch* can be thought of as a special type of coupling device. The most commonly used clutch in servomechanisms is the *slip clutch*. The clutch slips when the shafts reach a torque limit. Without the clutch, the servo may stall, potentially resulting in damage to the gearing or motor if the rotation energy is not dissipated in the clutch. Some models can be adjusted at the installation; others are fixed and are not field-adjustable. Occasionally, an electrically operated magnetic clutch is employed. For example, it could be used to disconnect a hand crank so it does not turn during normal system operation.

Bearings are an important part of any servomechanism. Sleeve bearings are usually oil-impregnated and are not commonly found in precision servomechanisms. For the most part, *ball bearings* are used because of their low-friction and low-wearing characteristics. Two common types of ball bearing mounts are the plain and flanged types. The *flange* type is popular because of the ease of snapping it in and out of place in its housing.

Some servomechanisms employ *limit stops* as a mechanical safety feature to prevent a shaft from rotating past a particular point. They are often used in conjunction with clutches to avoid damage. Some limit stops allow many revolutions before stopping by using a traveling nut on a screw thread. The threaded shaft is stopped when the traveling nut contacts either stop plate. In some cases, rubber or springs are used to absorb the shock when the stop is reached.

Gearing is required in a servomechanism to convert the high-speed low-torque power from the servomotor to a lower-speed higher-torque power to the controlled shaft. Many types of gears are available, including spur, helical, worm, bevel, internal pinion, and gear racks. Of all the gears mentioned, the *spur gear* is the most commonly employed. A *worm* is a gear with teeth in the form of screwthreads. Figure 10-38(*a*) shows a worm gear and a spur gear that can be mated to obtain a right-angle drive. Because of their high friction, right-angle drives are rarely found in servomechanisms. *Bevel gears* are conical in form and operate at intersecting axes, usually at right angles (Fig. 10-38[*b*]). *Internal gears* are usually limited to planetary drives and must mesh with an external gear. Figure 10-38(*c*) is an illustration of an external gear. A *helical gear* is cylindrical and has either right-hand or left-hand teeth, as shown in Fig. 10-38(*d*). They may be operated on parallel or crossed axes. Crossed helical gears are sometimes called *spiral gears*. When two gears run together, the one with

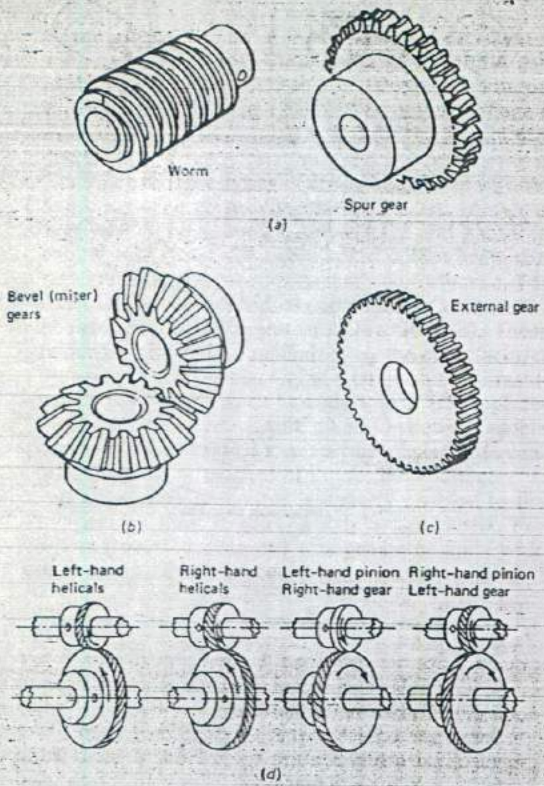


Fig. 10-38 Typical servo gears. (a) Worm gears. (b) Bevel (miter) gears. (c) External gear. (d) Helical gears.

the larger number of teeth may be called the *gear*, and the one with the smaller number of teeth may be called the *pinion*. A *rack* is a gear with teeth spaced along a straight line.

Uniform clearance of gears is critical in determining backlash between two gears. Backlash is proportional to the difference between the tooth space and the mating gear's tooth thickness. Increasing clearance will increase the difference and the backlash. Wear also increases the backlash, and so does loss of lubrication.

Differential gearing is used to mechanically add or subtract the angular position of two shafts. Differentials are usually made from bevel gears, as indicated in Fig. 10-39. Differentials are available with speed ratios from 1:1 to 3000:1.

Some dynamic characteristics of gearing are important in servomechanism applications. Assume a pinion of diameter d_1 is driving gear d_2. An external torque T_1 is applied to the pinion shaft, causing a rotation with an angular velocity ω_1. The torque on the shaft is equal to the force developed at a point on the circumference of the gear times half the gear diameter:

$$T_1 = F\frac{d_1}{2},$$

Therefore,

$$F = \frac{2T_1}{d_1}$$

Similarly, with the second gear:

$$T_2 = F\frac{d_2}{2}$$

For meshed gears, the forces are equal and substituting for F gives

$$T_2 = T_1\frac{d_2}{d_1}$$

For proper gear meshing to take place, all teeth must be the same size, and the number of teeth on each gear is proportional to the pitch diameter. The *gear ratio* is the ratio of the number of teeth on the second gear to the number of teeth on the first gear. This leads to

$$\frac{d_2}{d_1} = \frac{N_2}{N_1}$$

Combining with the torque equation,

$$\frac{T_2}{T_1} = \frac{d_2}{d_1} = \frac{N_2}{N_1}$$

These equations show that the ratio of the torque is proportional to the ratio of the diameters, which is also proportional to the gear ratio. It can be shown that

$$\frac{\alpha_2}{\alpha_1} = \frac{\omega_2}{\omega_1} = \frac{\theta_2}{\theta_1} = \frac{N_1}{N_2}.$$

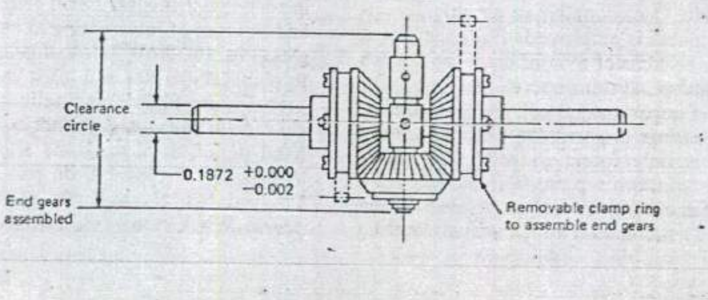


Fig. 10-39 Differential.

where α_1 and α_2 = angular acceleration of gears 1 and 2, respectively
θ_1 and θ_2 = angular position of gear 1 and gear 2
ω_1 and ω_2 = angular velocities of gear 1 and gear 2

Hence, the ratio of accelerations, positions, or angular velocities is inversely proportional to the ratio of the number of gear teeth. If the moment of inertia of the load is defined as

$$J = \frac{1}{2} Mr^2$$

where J = inertia, kilogram-meter² (kg · m²)
M = mass, kg
r = radius of gear, m

The angular acceleration of a gear is equal to torque divided by inertia:

$$\alpha_1 = \frac{T_1}{J_1}$$

and

$$\alpha_2 = \frac{T_2}{J_2}$$

Substituting, we obtain

$$\frac{\alpha_2}{\alpha_1} = \frac{T_2 J_1}{T_1 J_2} = \frac{N_2 J_1}{N_1 J_2} = \frac{N_1}{N_2}$$

which results in

$$J_1 = J_2 \left(\frac{N_1}{N_2}\right)^2$$

Therefore, using a step-down gear box reduces the apparent inertia affecting the servomotor by a factor equal to the square of the gear ratio.

For example, if a rotating load has an inertia of 250 kg · m², through a 40:1 gear ratio, what is the resulting inertia on the motor?

$$J_m = J_l \left(\frac{1}{40}\right)^2$$

$$J_m = \frac{250 \text{ kg} \cdot \text{m}^2}{40^2} = 0.156 \text{ kg} \cdot \text{m}^2$$

where J_m = inertia of motor
J_l = inertia of load

The gearing also provides an improvement in load torque as compared to motor torque. The torque and inertia advantages are the main reason for the use of gear boxes in a servomechanism. However, do not forget that these advantages are obtained at the sacrifice of speed.

REVIEW QUESTIONS

22. The B diameter on a coupling signifies the _______ diameter.

23. The bellows coupling is only used in high-torque applications. (true or false)

24. Large angular or lateral alignment corrections are obtained by using _______.

25. The mechanism that allows slip to occur when torque limits are reached is called a _______.

26. Torque multiplication is obtained at the expense of _______.

10-5 AMPLIFIERS AND FEEDBACK

A system can be represented by a combination of blocks. Each block may have a single line input and a single line output. Each block may represent a single function. For example, the block shown in Fig. 10-40(a) represents an amplifier with a voltage gain of 250 times. With a 1-mV signal in, the amplifier will produce a 250-mV signal at its output. Figure 10-40(b) is a more general form with the input voltage, output voltage, and amplifier gain given as V_1, V_2, and G_1, respectively. It can be said that G_1 operates on the input V_1 to give V_2. If V_2 is divided by V_1, this result will be equal to the gain (G_1), provided that the amplifier stays within its linear range. This function can be expressed as

$$\frac{V_2}{V_1} = G_1 = \text{a constant}$$

Figure 10-41(a) shows a second amplifier connected in cascade with the first amplifier. The entire system can also be represented in reduced form, as shown in Fig. 10-41(b). It is important to note that G_1 is multiplied by G_2, which operates on V_1 to obtain V_3.

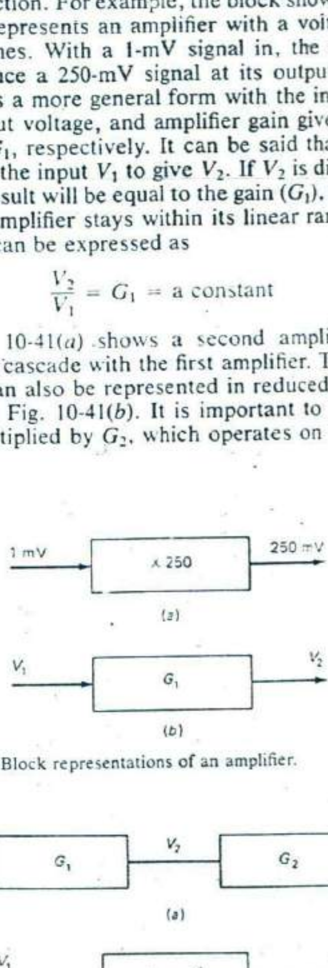


Fig. 10-40 Block representations of an amplifier.

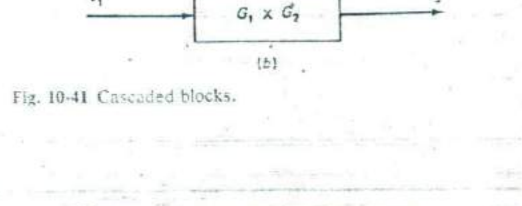


Fig. 10-41 Cascaded blocks.

EXAMPLE

If $V_1 = 1$ mV, $G_1 = 50$, and $G_2 = 50$, find V_3.

SOLUTION

$$V_3 = V_1 \times G_1 \times G_2$$
$$= 1 \text{ mV} \times 50 \times 50$$
$$= 2500 \text{ mV} = 2.5 \text{ V}$$

If G_2 were a potentiometer (voltage divider), its gain would be less than 1 (unity). At 50 percent rotation its gain (G_2) would be ½. The preceding total gain would now be 25 for the $G_1 \times G_2$ product. Since the gain of G_2 is 1/100 of the first example, the output will now be 25 mV.

The *summing junction* is used where signals are added or subtracted. In a block diagram, a circle is used as a summing junction as shown in Fig. 10-42. The Greek letter Σ may be used inside the circle to signify that a summing operation is to be performed. Summing junctions may perform addition or subtraction of two or more variables as shown in Fig. 10-42(*a* to *c*).

Figure 10-43(*a*) shows a gain block cascaded with a summing junction. The overall function can be found by

$$V_w = V_x - V_y + V_z$$
$$V_2 = V_w \times G_1$$
$$= G_1 (V_x - V_y + V_z)$$

if

$$G_1 = 20$$
$$V_x = -1 \text{ V}$$
$$V_y = -2 \text{ V}$$
$$V_z = +2 \text{ V}$$
$$V_w = -1 - (-2) + 2 = (-1 + 2 + 2) \text{ V}$$
$$= (3) \text{ V}$$
$$V_2 = G_1 \times V_w$$
$$= 20 \times 3$$
$$= 60 \text{ V}$$

An alternate representation is shown in Fig. 10-43(*b*). It illustrates that a junction can be viewed as several junctions. This may lead to simplified analysis in some cases.

Summing junctions are often based on operational amplifiers (op amps). Sum and difference amplifiers were covered in Chapter 6. Operational amplifiers may be used as noninverting or inverting amplifiers, depending on whether the signal is applied to the plus input or to the minus input. Figure 10-44 shows a summer (adder) circuit and a subtractor (difference) circuit. Other variations that are commonly used in feedback systems are shown in Fig. 10-45. Note that the summer (adder) is now of the noninverting type and that the difference amplifier output is now $V_2 - V_1$ as compared to $V_1 - V_2$ for the circuit of Fig. 10-44. These circuits are but a few examples of those available to fit the numerous mathematical needs of servomechanisms.

Servo amplifiers can be divided into two types. Most of the gain is contained in an early amplifier often called the *preamplifier*. The power gain is produced in the final amplifier. The power amplifier must supply the required load voltage and current, which can be substantial, as in the case of a large servomotor.

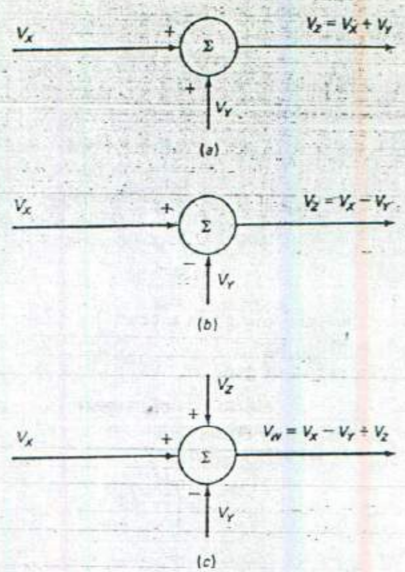


Fig. 10-42 Summing junctions.

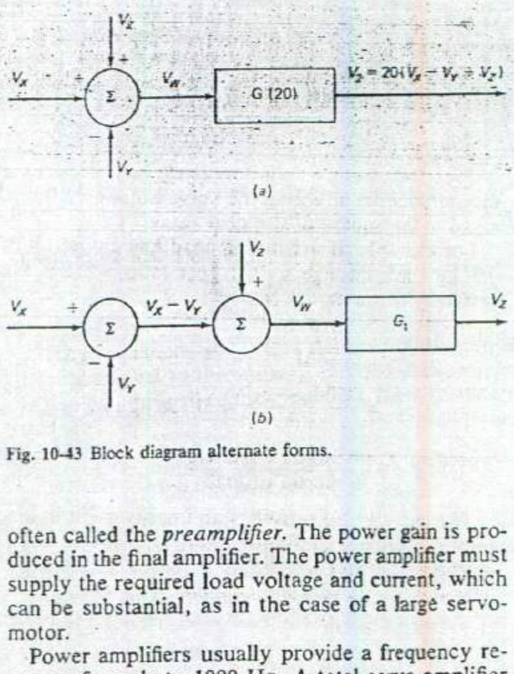


Fig. 10-43 Block diagram alternate forms.

Power amplifiers usually provide a frequency response from dc to 1000 Hz. A total servo amplifier (package) may appear as a solid-state unit. Figure 10-46 shows what is contained within such units. The amplifier gain can be adjusted as needed by resistor selection. The frequency response is adjusted during installation by selecting and connecting frequency-compensation components. In some low-power applications, the power amplifier stage may be a com-

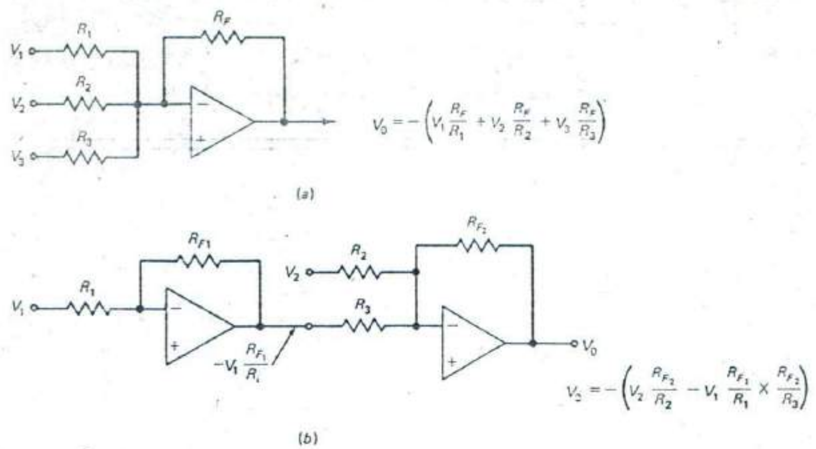


Fig. 10-44 Op Amp circuits. (a) Summer (adder). (b) Subtractor.

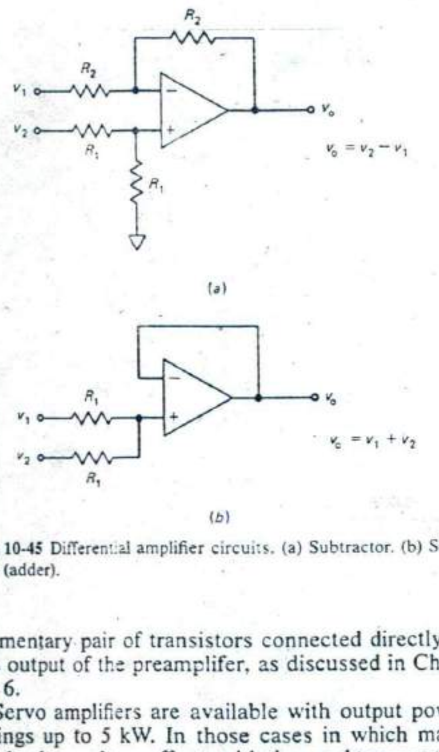


Fig. 10-45 Differential amplifier circuits. (a) Subtractor. (b) Summer (adder).

plementary pair of transistors connected directly to the output of the preamplifer, as discussed in Chapter 6.

Servo amplifiers are available with output power ratings up to 5 kW. In those cases in which maintaining low-voltage offsets with time and temperature is essential or when external set adjustments are not practical, a chopper-stabilized amplifier is used to achieve drifts as low as 0.1 μV/°C. The chopper amplifier is a high-gain feedback amplifier, containing a MOSFET chopping transistor. The chopper converts the difference between the dc or low-frequency input voltage and the feedback voltage to a high-frequency square wave and amplifies it with no drift. The high-frequency square wave is then rectified and filtered to produce an output waveform that is an amplified version of the input.

Figure 10-47 presents a detailed block diagram of a chopper-stabilized amplifier. In a system diagram, it may be simplied and represented by a single OP-AMP symbol. The chopper-stabilized amplifier is best understood by looking at the waveforms in Fig. 10-47. The input signal, V_{in}, is split into two components by high-pass network $C_1 - R_1$ and by the low-pass network R_2C_2. The low-frequency signals are applied to the chopper, along with a square wave that gates the chopper transistor off and on at a high frequency rate. The ac amplifier can be designed for high gain and drift will not be a problem since its frequency response does not extend down to dc (it will utilize coupling capacitors between stages). The amplified output is then peak-detected to recover the low-frequency and dc components. Capacitor C_3 filters the detected signal to remove any component of the chopper signal. The output signal V_{out} is formed by summing the detected and filtered signal with the high-frequency signal.

In cases in which overloads are unavoidable or may create very long recovery time, an overload circuit may be incorporated into the amplifier. Such a circuit is illustrated in Fig. 10-48. This overload recovery circuit will prevent the amplifier circuitry from saturating. The input circuit is protected from large signals by a diode clipper. The feedback circuit will drop in impedance with a large swing in V_{out} due to the zener's becoming forward-biased. This will lower the amplifier gain and allow the amplifier to recover in 1 μs. Without this overload protection, it may take up to 10 s to recover.

A *closed-loop system* is one in which the output of a process affects the input. Adjustments are made by the control system until the difference between

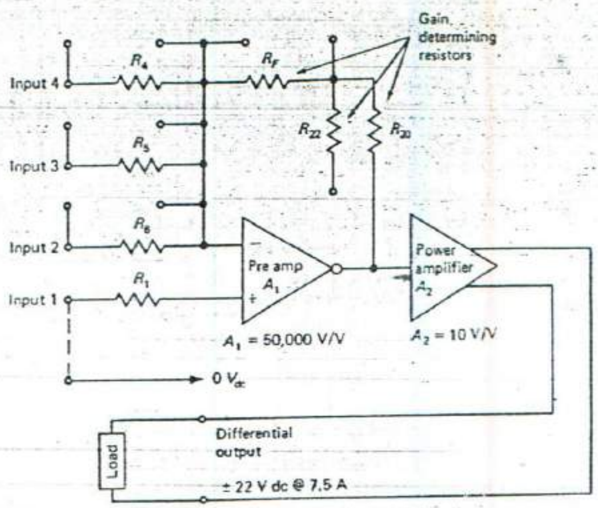


Fig. 10-46 Servo amplifier block diagram.

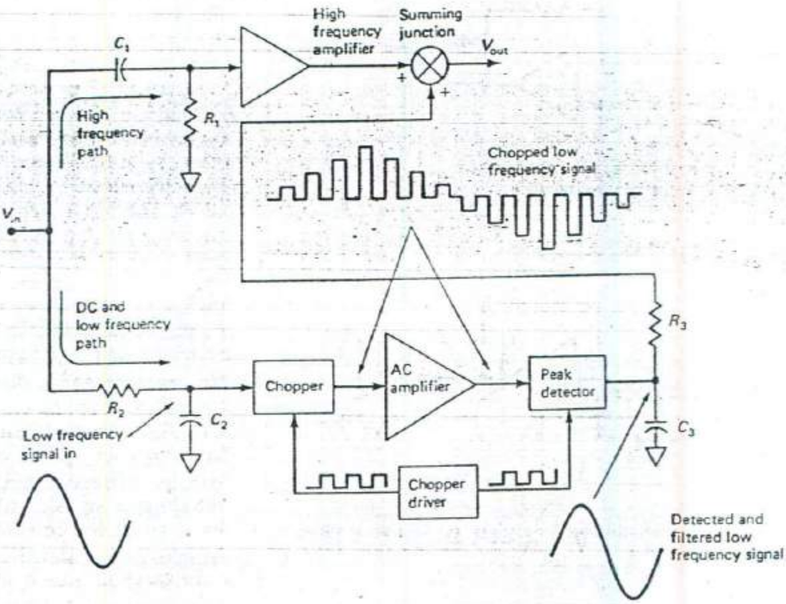


Fig. 10-47 Chopper-stabilized amplifier block diagram.

the desired and actual output is as small as practical. In other words, the controlled parameter (output), whether position, angle, or speed, is sampled and fed back to the input, where it can be compared with the desired condition. Look at Fig. 10-49, in which a human operator is trying to maintain a speed of a trolley at 50 mph (80 kph). The operator observes the speedometer and decides (compares) whether to increase or decrease the speed control, depending upon whether the indicator is above or below the desired 50 mph (80 kph). Note the two signal paths: (1) a forward path from the operator handle controlling the speed (speed is the output), (2) a backward (feedback) path via the speedometer to the operator, serving as the comparator to the speed controller. The input to the speed-control handle is the difference between that indicated by the speedometer and the desired speed computed by the operator. It would be almost impossible to maintain a constant speed without feedback. If we had no way of knowing the speed we would have to guess.

Block diagram representation is a technique commonly used in control system analysis. Consider the illustration in Fig. 10-50(*a*), which shows a simple

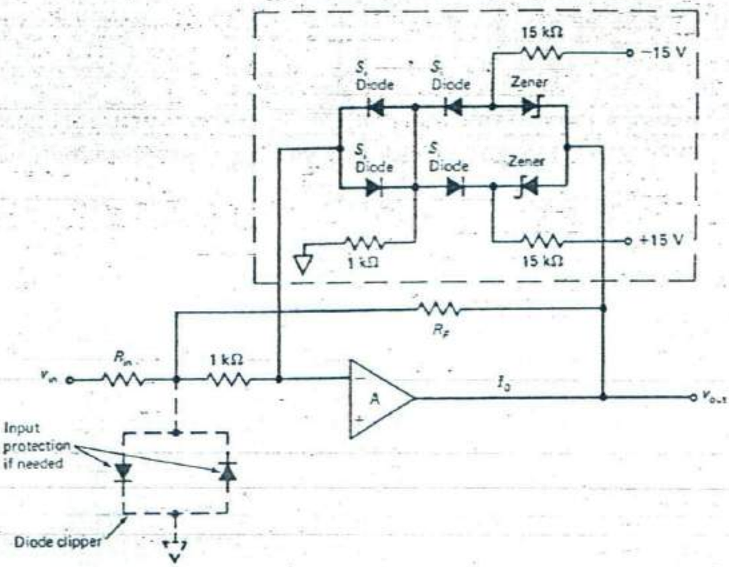


Fig. 10-48 Overload recovery circuit.

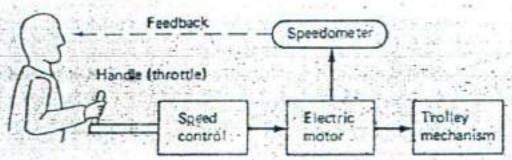


Fig. 10-49 Simple closed loop feedback system.

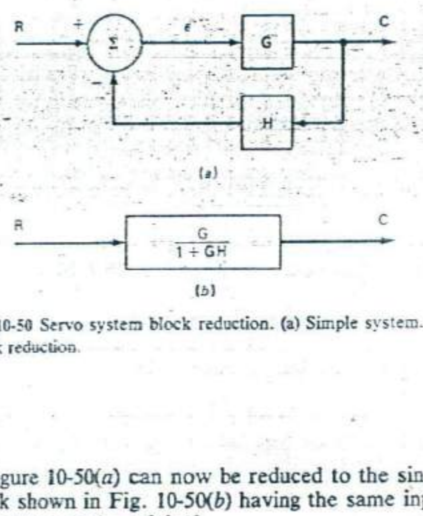


Fig. 10-50 Servo system block reduction. (a) Simple system. (b) Block reduction.

form of feedback control system. The block diagram can be reduced by block diagram algebra into a form that is commonly used and should be remembered. From Fig. 10-50(*a*)

$$\epsilon G = C$$

also,

$$R - CH = \epsilon$$

Substituting for ϵ yields

$$(R - CH)G = C$$

or,

$$RG - CGH = C$$

Combining

$$RG = C + CGH$$

then,

$$RG = C(1 + GH)$$

The control ratio is defined as the output C over the command signal R and is equal to

$$\frac{C}{R} = \frac{G}{1 + GH}$$

Figure 10-50(*a*) can now be reduced to the single block shown in Fig. 10-50(*b*) having the same input and output as the original.

The unity feedback system shown in Fig. 10-51(*a*) can also be reduced to one block and the control ratio C/R derived. The feedback loop can be considered to have a value of $H = 1$. All the output C is fed back to the summing junction. The reduction will therefore be the same as previously shown in Fig. 10-50(*b*):

$$\frac{C}{R} = \frac{G}{1 + G \times 1}$$

$$\frac{C}{R} = \frac{G}{1 + G}$$

The reduction is shown in block form in Fig. 10-51(*b*).

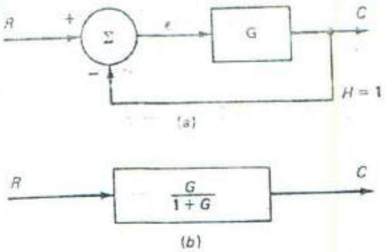


Fig. 10-51 Unity feedback system block reduction. (a) System. (b) Block reduction.

There are five important characteristics for a closed loop system:

1. Accuracy
2. Sensitivity
3. Resolution
4. Linearity
5. Frequency response

Together, all five characteristics provide a complete description of the system.

The *accuracy* is an indication of how closely a system meets the desired control parameter. If the normal speed of a system is 77 kph and the actual varies between 70 and 84 kph, the system is said to be accurate to ± 7 kph. Accuracy may also be defined as a percentage of deviation. Instead of ± 7 kph we could specify ± 9 percent deviation. The percentage of error may be based on a maximum value. For example, if the maximum value were 210 kph, the maximum error of ± 7 kph would become

$$\text{max error, } \% = \frac{7 \times 100}{210} = 3.33\%$$

which is less than a third of the nominal percentage error.

The *sensitivity* of a control system specifies the level of input required to obtain a desired output. A speed-control system may have a sensitivity of 1000 rpm/V; a temperature-control system may have a 100°/V sensitivity. Therefore, sensitivity has various interpretations, depending on the specific system used as reference.

Resolution is defined as the smallest quantity recognizable by the system. Resolution can be specified in percentage of maximum or in absolute units. If a thermometer has 1° markings, 1° is the smallest recognizable quantity.

Linearity is defined as the amount by which a relationship between two quantities deviates from a straight line. Linearity is usually expressed as a percentage of a value or as a percentage of maximum. Operation within a linear working range is obtained only if the system signals are restricted in magnitude to avoid nonlinear regions of the components. For example, if an amplifier is driven into saturation by a large input signal, linearity is lost.

Amplifier frequency response was covered in Chapter 6. The same definition applies to a control system. The output of the system may be plotted on a frequency-response curve and the −3-dB points obtained. The *response time* of a system is usually used, rather than the frequency response, in control systems. This is the time a system takes to respond to an input signal. By definition it is the time needed for the output to change from 10 to 90 percent of its final value when a step signal is applied to the input as shown in Fig. 10-52(*a*).

This is referred to as *rise time* and it is also the response time of a servomechanism. In Fig. 10-52(*b*), curve (*a*) shows an underdamped response. The output overshoots the value dictated by the systems input, then undershoots this value, and finally settles to a value close to the input value. This type of response has an oscillating or ringing effect. In the overdamped response (*b*) the output does not overshoot the desired value but takes a very long time to reach its final value. The third response (*c*) is that of a critically damped system, in which the output reaches its final value in the minimum possible time without overshooting the desired final value.

Both the mechanical and electrical components of a control system have response-time characteristics, and both determine the overall system performance. After a transient period, a final steady-state value is

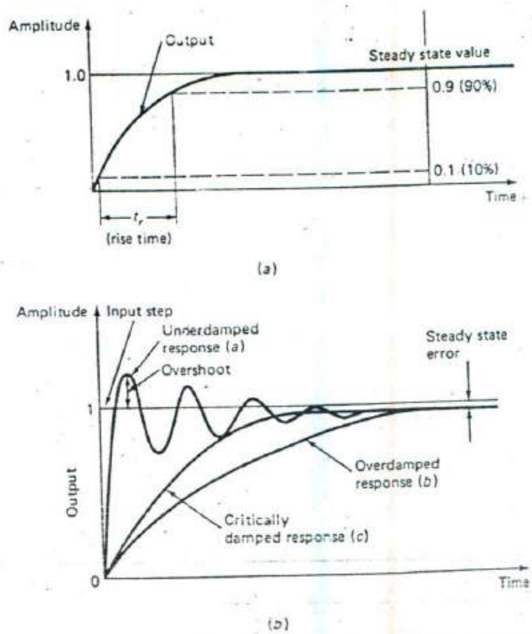


Fig. 10-52 Transient response curves. (a) Rise time (response time). (b) Response curves.

obtained. The difference between the final steady-state output reached and the value called for by the input is called the *steady-state error.* The steady-state error is shown in Fig. 10-52(*b*). The gain of the system can be increased to reduce the steady-state error. However, with increased gain, a situation in which the oscillations obtain a fixed amplitude and the system never settles or reaches steady state may arise. This is an *oscillatory system,* and it is said to be unstable. Open-loop systems are never unstable, but when feedback is introduced, the control system can become unstable. Any feedback system may become unstable if the feedback is in phase with the input. Amplifiers and feedback networks exhibit phase errors that are especially pronounced at the frequency limits. It is possible for negative feedback to become positive feedback at some frequency extreme. For this reason servomechanisms tend to become unstable when the gain is increased in an effort to reduce the steady-state error. In order for oscillations to occur, the magnitude of G must be greater than 1 when the phase error is $-180°$. Various techniques, such as Bode and Nyquist plots are used to ensure stability when the system is designed. It is important to understand that system stability can be lost if gain or phase-shift networks are altered.

There are three types of servomechanisms:

1. Type 0: A constant input signal (x) will result in a constant position at the controlled output (y).
2. Type 1: A constant input signal (x) results in a constant velocity at the output ($\dot{y}$).
3. Type 2: A constant input signal (x) results in a constant acceleration at the output (y).

The control system type is determined mathematically by examining the G and H factors of the loop which determine its transfer function. A mathematical analysis of a control system's transfer function yields two key factors. The first is the steady-state response of the system to three types of inputs. The second is the steady-state error, which is either zero, finite and constant, or infinite.

Figure 10-53(*a*) shows a *type 0 servomechanism* (often called a *position* or *follow-up system*). The output shaft (either driven directly or through a gear box) is to follow the angular setting of the input potentiometer. If a step input is applied to the type 0 system, the steady-state error, E_{ss}, for a step (also called a *position* or *setpoint*) with a value of P is

$$E_{ss} = \frac{P}{1 + K}$$

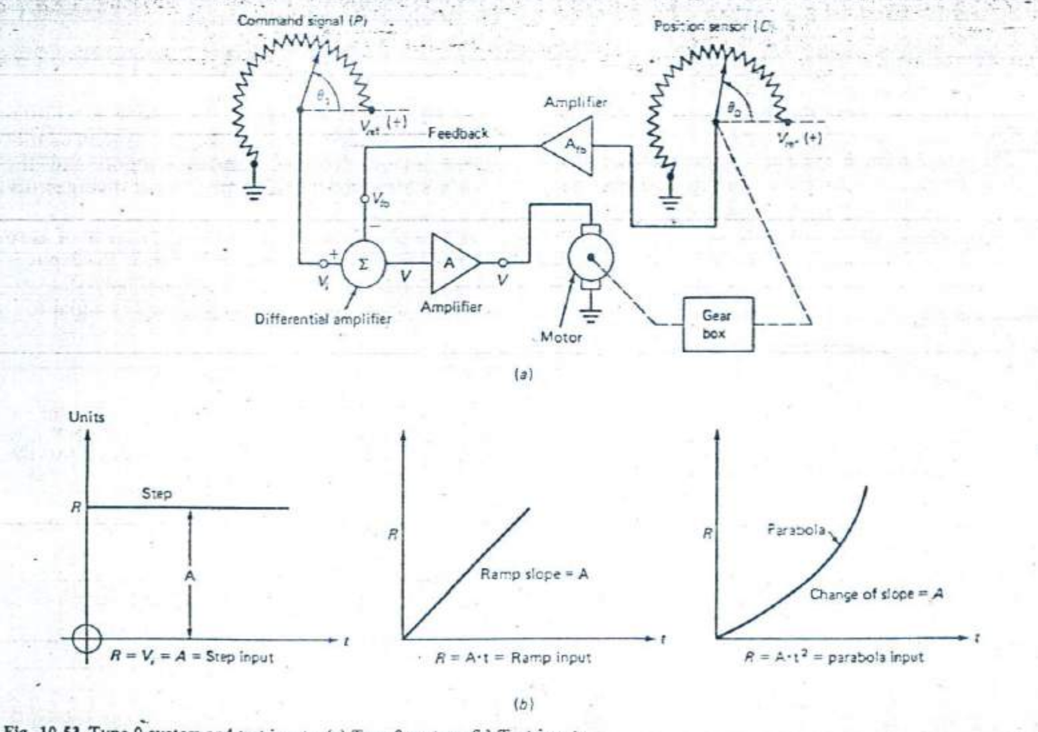


Fig. 10-53 Type 0 system and test inputs. (a) Type 0 system. (b) Test inputs.

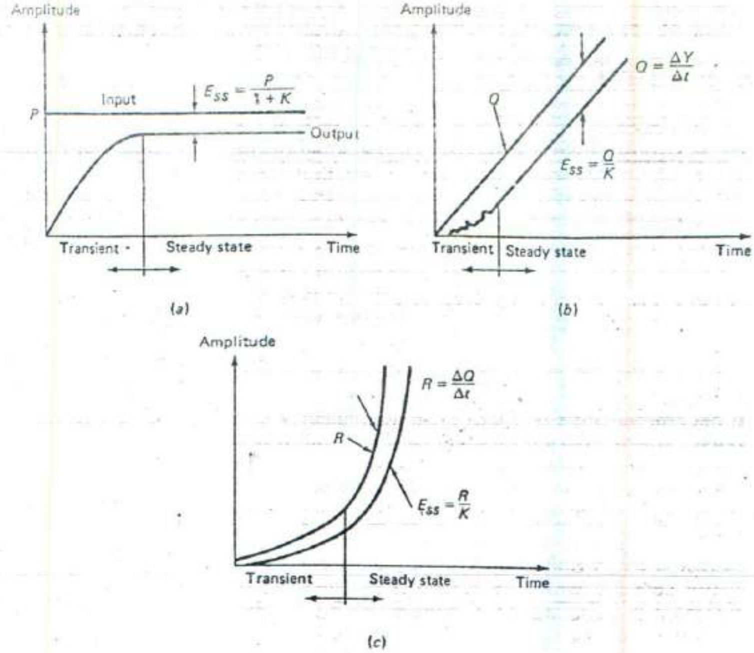


Fig. 10-54 Steady-state error for different system types. **(a) Type 0** system. (b) Type 1 system. (c) **Type 2** system.

The larger the value of K (gain) the smaller the error, but as mentioned earlier, large values of gain can make a system unstable. This response for a type 0 system is shown in Fig. 10-54(a). If a ramp or acceleration (parabolic) input is applied to a type 0 system, the output cannot follow it, and the steady-state error increases with time and approaches a value of infinity. These inputs are shown in Fig. 10-53(b).

The *type 1 control system*, which is known as a *rate (velocity) servo*, is shown in Fig. 10-55. As defined earlier, the output shaft will run at a constant velocity (speed) for a constant input. The steady-state error of the type 1 system to the step input is zero, which is the desired circumstance. The steady-state error of a type 1 system due to a ramp input Q is shown in Fig. 10-54(b). If the system gain is K, the steady-state error is

$$E_{ss} = \frac{Q}{K}$$

As with the type 0 system, increasing K will decrease the steady-state error. A type 1 system cannot follow an acceleration (parabolic) input, and the steady-state error for this type of input diverges (increases) as time increases.

A *type 2 system* has a steady-state error of zero for both position and velocity inputs. If the input is an acceleration of value R, then the steady-state error E_{ss} is as shown in Fig. 10-54(c) and is equal to

$$E_{ss} = \frac{R}{K}$$

The type 2 control system is seldom used industrially and is more commonly used in missile and guidance systems. The steady-state errors for the three types of systems are summarized in Table 10-2.

If a control system is found to be marginally or inherently unstable, a compensating network may be added to improve the system's gain and phase margin. These networks may be based on integrators or differentiators, and their applications in servo damping are shown in Fig. 10-56. These stabilizing networks are usually found either preceding or incorporated within the amplifier (gain stage) of the system. They are critical for proper system response

Fig. 10-55 Type 1 (rate/velocity) system.

TABLE 10-2 STEADY-STATE ERRORS FOR VARIOUS INPUTS TO THREE TYPES OF SERVOMECHANISMS

	System Type		
Input Type	0	1	2
Position, P	$\frac{P}{1+K}$	0	0
Velocity, Q	Infinity	$\frac{Q}{K}$	0
Acceleration, R	Infinity	Infinity	$\frac{R}{K}$

and should not be modified; if they are modified, oscillations may occur.

Steady-state error can be reduced by increasing loop gain, but stability can be lost if the gain is too high. Compensating networks improve the phase margin and allow greater gain for improved performance. Integral damping is a related technique to reduce steady-state error. Integrators are presented in Chapter 6. The output of an integrator will ramp in response to a steady signal applied to its input. In the case of a type 0 servo, as in Fig. 10-56(a), suppose a step command signal is applied to position the output to a new location. An immediate and relatively large error signal results, driving the input of the summing amplifier. The summing amplifier output drives the controller in a direction that eliminates the error. When the step signal is first applied, the integrator output is zero, and the system responds as if the integrator were not in the circuit. As time passes, the error signal decreases, but the integrator output increases. Without the integrator, a small residual error signal would be present when the controller finally stopped. However, with the integrator the error signal will eventually be removed because the integrator output continues to ramp as long as any residual error voltage remains at its input.

Integral damping may also be added to a type 1 servo. Suppose there is a sudden change in the command set point. The immediate reponse will be a relatively large error signal that will be amplified and will act on the output device, thus reducing the error to a small value as the output speed comes close to matching the command signal. However, because of friction and loading torque on the output, there will be some steady-state error and some residual error voltage. If this error voltage is integrated and applied to the input of the amplifier, it will eventually be removed as the integrator output continues to ramp in response to the steady-state error signal.

The underdamped response that was shown in Fig. 10-52 is undesirable. Because of inertia (both mechanical and electrical) the controlled parameter overshoots the set point value, then undershoots it, and so on. Gain can be reduced to control overshoot but at the expense of response time and accuracy. Derivative damping is a technique based on a second

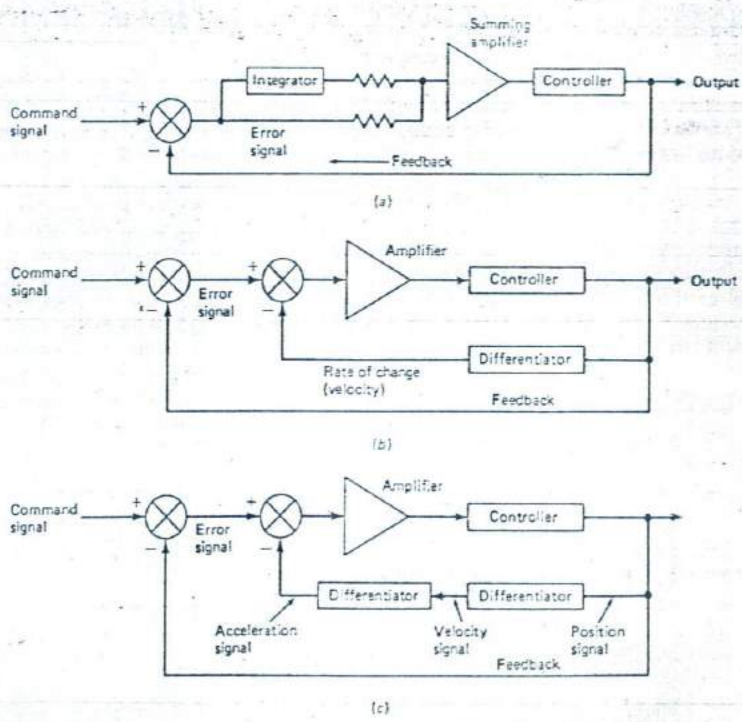


Fig. 10-56 Compensation/damping techniques. (a) Integral damping. (b) Derivative damping. (c) Second-derivative damping.

feedback loop that contains a differentiator. (Differentiators are covered in Chapter 6.) The output of a differentiator is proportional to the rate of change at its input. When derivative damping is added to a servomechanism, overshoot and undershoot can be reduced without sacrificing loop gain. Suppose a change in command signal is applied to the damped system shown in Fig. 10-56(b). Because of system inertia, the immediate response is slow; the output of the differentiator is also small, so it has little impact on the input to the amplifier. As the system continues to respond, the error signal begins to decrease, and the rate of change is now increasing, so the output of the differentiator is increasing. Eventually, the error signal and the rate signal (velocity) will be equal, resulting in no input to the amplifier. However, the system continues to respond because of its inertia. Now, the differentiator output is greater than the error signal which reverses the polarity of the amplifier output. The controller responds by reversing, and the overshoot is reduced (damped).

Everything in electronic systems has an analog in mechanical systems, and the reverse is also true. For example, mechanical damping (the analog of the electronic damping just discussed) is also possible. Friction forces that oppose oscillatory motion can be developed. Simple brakes are usually not used because they degrade response time and accuracy. Rate brakes that respond to the rate of change are used. These are usually called *viscous damping systems* and include devices such as fluid turbines, magnetic particle brakes, and eddy current types.

Derivative damping can cause an error in type 0 servomechanisms if there is any output from the differentiator under steady-state conditions. The first derivative of motion is velocity, and the second derivative of motion is acceleration [Fig. 10-56(*c*)]. Therefore, a second-derivative damping system provides damping correction only when the output is accelerating or decelerating.

Servomechanism characteristics can also be controlled by digital techniques. With the advent of the microprocessor, there is an increased tendency to calculate the correct input to the servomotor at any given time to provide the desired response. Digital techniques and microprocessors are covered in the next two chapters.

REVIEW PROBLEMS

27. If in Fig. 10-42(*c*) V_x is +3, V_y is +3, and V_z is −4, then V_w is ______.

28. If V_w of question 27 is fed into a block with G_1 of 0.25 V/V the output is ______ V.

29. In Fig. 10-44(*b*), V_1 is +4, V_2 is −2, and R_F equals R_1 equals R_2 equals R_3, V_{out} equals ______ V.

30. The amplifier that drives a servomotor is a ______.

31. In special high-stability cases, a ______ amplifier may be employed.

32. To prevent lock-up, an amplifier may incorporate an ______ circuit.

33. A type 1 system will have an error equal to ______ for a ramp input as shown in Fig. 10-54.

10-6
ROTATING AMPLIFIERS

When the control of a large power load is required, the choice is usually limited to hydraulic systems or electrical systems involving motors, mostly dc. If a dc motor is chosen for the output, a rotating power amplifier may be used to excite it. Rotating amplifiers are being replaced with solid-state devices as the power capabilities of these devices continues to improve. Most new systems employ solid-state amplifiers. But rotating amplifiers are still in use at many industrial sites, and familiarity with their characteristics is still important.

Rotating amplifiers fall into one of three basic categories. They are the Ward Leonard system, the Regulex or Rototrol generator, and the Amplidyne generator. Each system has a generator driven by an ac induction motor at a constant speed, 3600 rpm for small to medium systems and 1800 rpm for large systems. They can be considered as amplifiers because a small change in field current will produce a large change in armature current.

The Ward Leonard system employs a simple motor-generator set. The circuit shown in Fig. 10-57 is a Ward-Leonard system in its simplest form. The dc motor in the circuit is fed directly from a dc generator, which is operated at a constant speed. The generator may be viewed as a power amplifier since the power required to excite the field is much lower than the power output from the armature. The dc field to the generator (F_g), is adjustable in magnitude and polarity by means of the field rheostat and the reversing switch. The motor armature is supplied by the generator, which provides a smoothly varying voltage from zero to some full-load value. The dc motor field (F_m) is fed from a constant source derived from the ac line. The generator is driven by an ac induction motor of constant speed. This motor may be either single- or three-phase, depending upon the size of the application. The Ward Leonard system allows a small variation in field current to provide a smooth, reversible, flexible, and stable control of a large dc motor. Such systems are employed in hoists and elevators and in large machining centers.

The system shown in Fig. 10-57 is open loop and only applicable where an operator must have complete control at all times. This system can be modified for closed loop operation by providing a feedback signal to be compared to the control input. The feedback can be proportional to rate (type 1) or to position (type 0). Figure 10-58(*a*) shows how the system can be connected with tachometer feedback to provide motor-speed regulation. The motor speed is measured by the output from the dc tachometer (rate

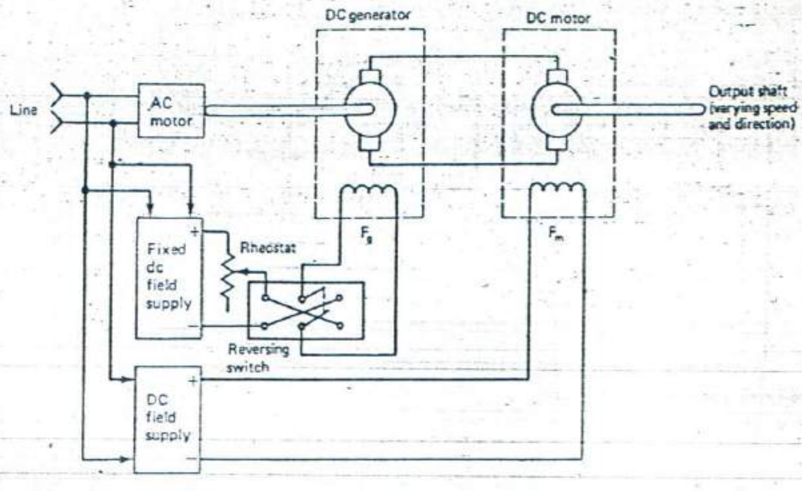


Fig. 10-57 Simple Ward-Leonard system.

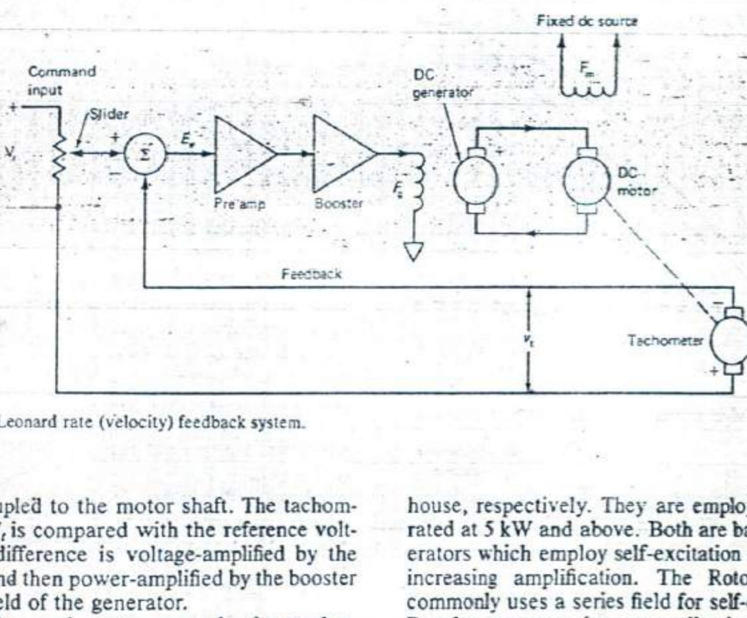


Fig. 10-58 Ward-Leonard rate (velocity) feedback system.

generator) coupled to the motor shaft. The tachometer voltage V_t is compared with the reference voltage V_r. The difference is voltage-amplified by the preamplifier and then power-amplified by the booster to drive the field of the generator.

The Ward Leonard system can also be used to control the angular position of an output shaft at a high power level by following an input signal at a lower power level. This type of system is shown in Fig. 10-59. The output angle θ_o is converted into an electrical signal, which is compared with the command signal. Any error is amplified and applied to the field circuit of the dc generator. The slider/wiper of the position-sensing potentiometer is driven by the output shaft through a gear box. Some systems use synchros and control transformers in place of potentiometers for input and feedback devices.

The Regulex generator and the Rototrol generator are trade names of Allis-Chalmers and Westinghouse, respectively. They are employed in systems rated at 5 kW and above. Both are basically dc generators which employ self-excitation as a method of increasing amplification. The Rototrol generator commonly uses a series field for self-excitation. The Regulex generator, in most applications, uses a shunt field. The typical magnetization curve of the dc shunt generator shown in Fig. 10-60(*b*) is shown in Fig. 10-60(*a*). The simple shunt generator is driven at a constant speed with the field switch (*S*) open. A small residual magnetism is assumed to generate some value of emf at zero excitation. The straight line (θ_a) is called the *field-resistance line* and is a plot of

$$V_f = R_f \cdot I_f$$

The slope of the field-resistance line is determined by the field rheostat. This slope (O_a) is less than the air-gap line (a straight-line projection of the magnetization curve through the origin). Point (*a*) is the

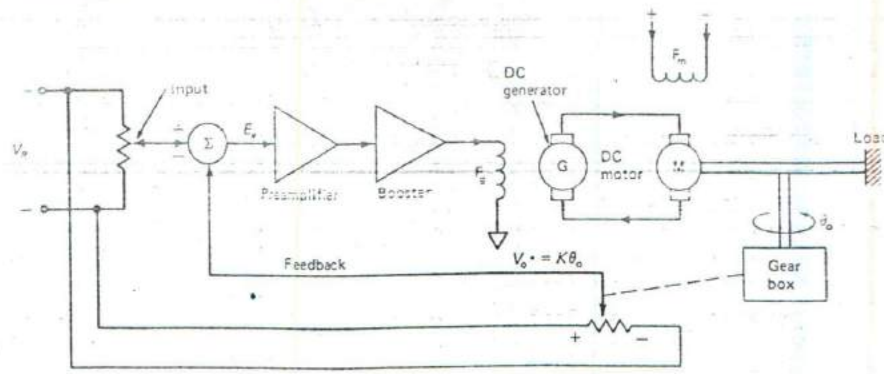
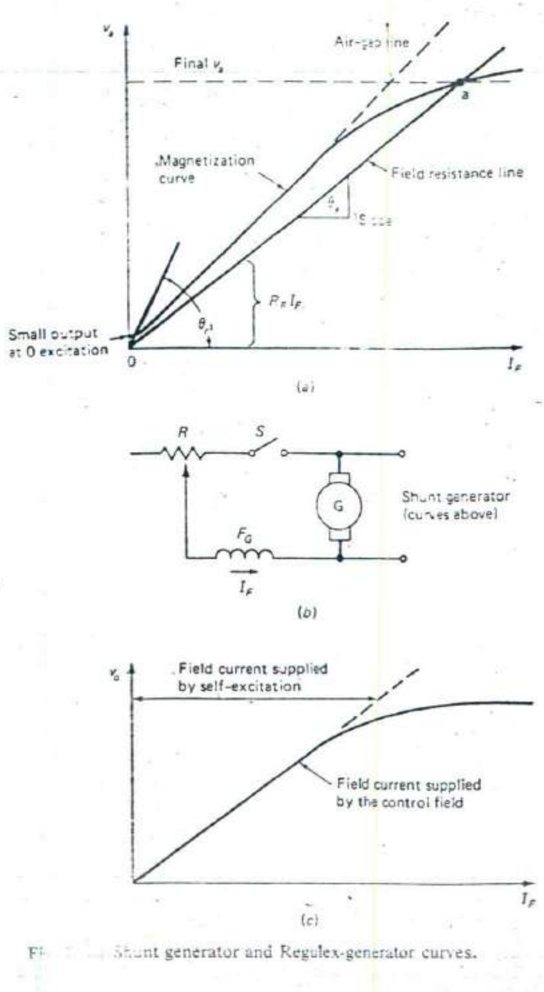


Fig. 10-59 Ward-Leonard position control system.



Fi[illegible] Shunt generator and Regulex-generator curves.

intersection of the field-resistance line with the generator's magnetization curve. If the switch (S) is closed, the residual output will start a build-up of the shunt-field current. If the connections are such that positive feedback results, the generator voltage builds up until limited by magnetic saturation at point (a). The generated voltage will just satisfy the field current required to sustain itself.

If the field resistance is large, the resulting field-resistance line will be small (θ_1' in Fig. 10-60[a]), and very little output voltage build-up will occur. The value of R_f corresponding to the slope of the air-gap line is called the *critical field resistance*. An adjustment to this value is referred to as *tuning*. Suppose the resistance of the self-excited field is increased until it has the slope shown in Fig. 10-60(c), and the necessary additional field current required is supplied by an auxiliary control field. Up to the saturation of the iron, the output will be proportional to the applied control current, which is only a small percentage of the total field required.

The self-excited generators have a number of field windings. Up to eight separate windings can be put on a machine. For a given power output, these added windings will increase the physical size of the generator compared to that of a conventional model. Figure 10-61 shows two basic self-exciting generators. Both have critical values of resistance, and both can be tuned. The main differences show up in the magnitudes of field currents involved. The shunt generator will have a small field current flowing in a high-impedance winding of many turns. The series generator shows a large current flowing in a low-impedance winding of few turns. Whether shunt or series, the field is *tuned* to the air-gap line, and operation is on the linear part of the magnetization curve.

When a change in the output voltage is necessitated by a change in the load requirements, the control fields are used to establish the new operating point. The power required by the control winding is very small because it only has to initiate change or

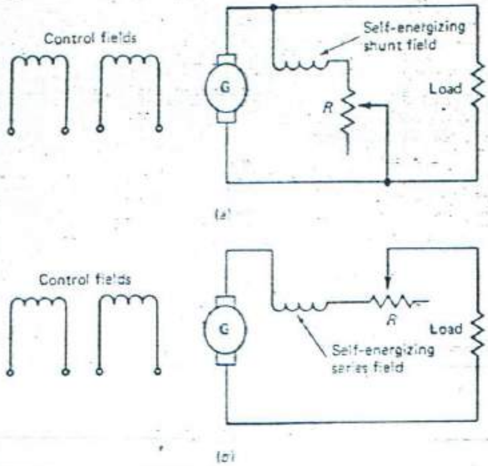


Fig. 10-61 Basic self-energizing generators.

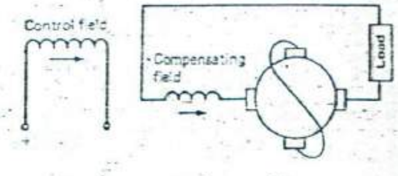


Fig. 10-63 Schematic diagram of an amplidyne.

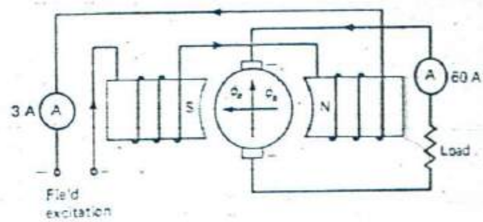


Fig. 10-64 Magnetic fields and currents in a conventional dc generator.

stabilize steady operating conditions. Figure 10-62 shows a voltage control system with an exciter (*E*), control winding (*C*), and a tuned shunt-field winding (S_f). The control field voltage V_c is obtained by feedback comparison of the generator voltage V_g with the reference voltage (V_r). The exciter is driven at a constant speed. At steady-state conditions, the shunt field provides all the required excitation, and error voltage V_c is zero. Any tendency to drift from this point results in a correction voltage V_c across the control field. Suppose a load change decreases the output voltage. This error voltage $V_c = V_r - V_g$ is applied to the control field, and the exciter output builds up. This output will continue to change as long as the error exists. The final operating point will shift to the point at which the self-energizing shunt field again alone supplies all the excitation required by the new load requirements. At this new operating point, the steady-state error V_c is again zero. The series-excited generator will act in essentially the same way as the shunt generator just discussed.

The amplidyne motor generator consists of a constant-speed ac drive motor and a two-stage electromechanical power amplifier contained in a single housing. As the symbol in Fig. 10-63 shows, the amplidyne generator has two sets of quadrature brushes with one set shorted. The drive motor, usually a squirrel cage type, has its rotor shaft coupled to the armature of the generator section. Since this motor drive is similar to previously discussed systems, it need not be covered again. The amplidyne is radically different from the conventional generator because of the unusual method employed to obtain high-power amplification.

Figure 10-64 shows a dc generator with a 60-A load on the armature. To meet this demand, the armature must have induced in it sufficient voltage to force the required current to the load. Therefore, the armature conductors must cut a magnetic field of a certain flux density to provide the required output. A field current of 3 A may be necessary in this case. The generator can now be considered as a current amplifier with a gain of 20.

Figure 10-65 is the same as Fig. 10-64, except that the load has been removed and the armature leads short-circuited. Since the load resistance is gone, the only significant opposition to current flow is the resistance of the armature windings. This condition would produce abnormally high armature current and

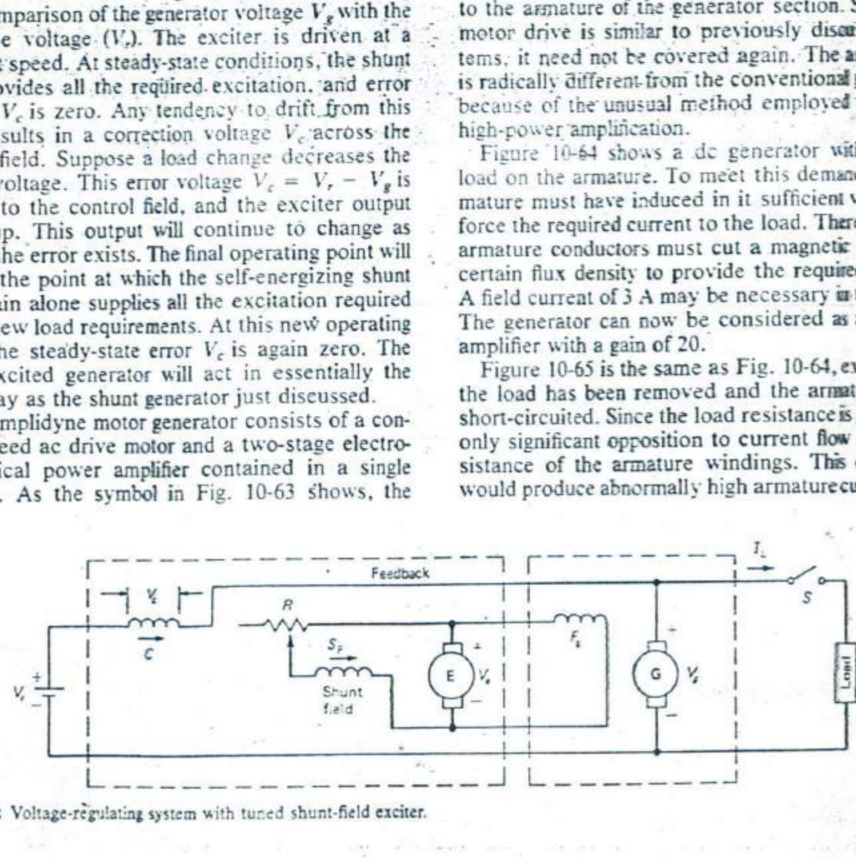


Fig. 10-62 Voltage-regulating system with tuned shunt-field exciter.

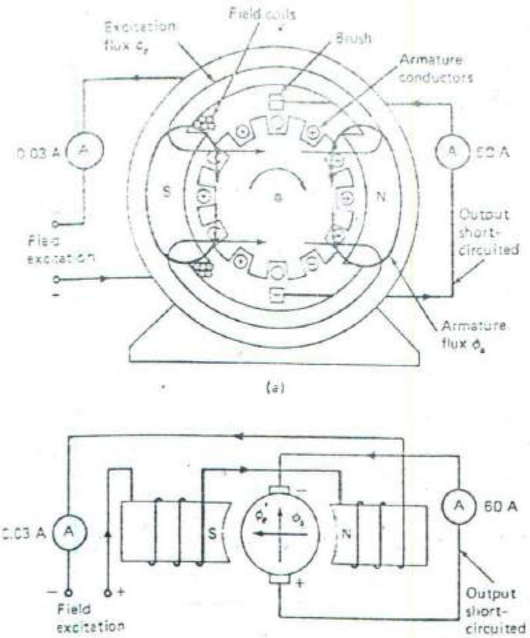


Fig. 10-65 Magnetic fields and currents in a short-circuited dc generator.

quickly lead to a burned-out armature. However, one way to reduce the enormous current is to reduce the excitation flux to a much lower level. It is possible to reduce the short-circuit current in the armature to 60 A by reducing the flux. Since the armature handled a 60-A load before, a short circuit of the same current will not cause any damage. A reduction in the field excitation current will weaken the flux to the proper level, in this case, perhaps 0.03 A (30 mA). It can be seen that 0.03 A now controls a short-circuited current of 60 A. Before, 3 A with a load applied was required. The generator gain has increased to 2000. The problem as to how this increased power gain can be put to use now arises. Obviously, the load cannot be put in series with the short circuit, since this would just be a return to the original circuit. The short circuit must remain. It can be seen in Fig. 10-65(*b*) that two flux circuits exist: ϕ_e, a weak excitation flux, and ϕ_a, a strong armature flux due to the 60 A. The cross section of Fig. 10-65(*a*) shows that the armature conductors are evenly spaced around the core. They will cut across the heavy armature flux, ϕ_a, at the same rate that they will cut the excitation flux ϕ_e. The maximum voltage induced in the conductors as they cut the armature flux will be at right angles to the voltage induced by the [illegible] flux. To take advantage of this second [illegible], a second set of brushes, shown [illegible], is added to the commutator at right

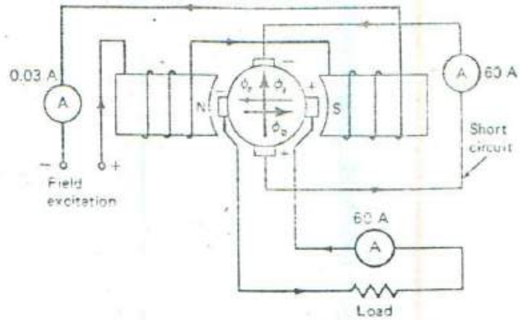


Fig. 10-66 Short-circuited generator supplied with additional brushes.

angles to the short-circuited brushes and connected to the load. The voltage developed across the second set of brushes is sufficient to supply a 60-A current to the load.

Another problem arises, as can be seen in Fig. 10-66. As the armature current in the short-circuited section creates a flux at 90° to the excitation flux, so will the load current set up a flux at 90° to the armature flux. This new reaction flux, ϕ_b, is 180° from the original excitation flux, ϕ_e. The reaction flux is much stronger than the excitation flux, and because it opposes it, the excitation flux no longer has control of the output. To overcome this condition, a compensating winding is placed on the pole pieces and is connected in series with the load. The number of turns is adjusted so that the compensating flux ϕ_c will exactly cancel the load armature reaction flux for all values of load current within the operating range. The equivalent circuit is shown in Fig. 10-67.

Since any residual magnetism along the axis of the control field would considerably affect the Amplidyne output, it is necessary to demagnetize the core material. A small ac generator is used to eliminate any residual magnetism. This generator has a permanent magnet attached to the end of the armature.

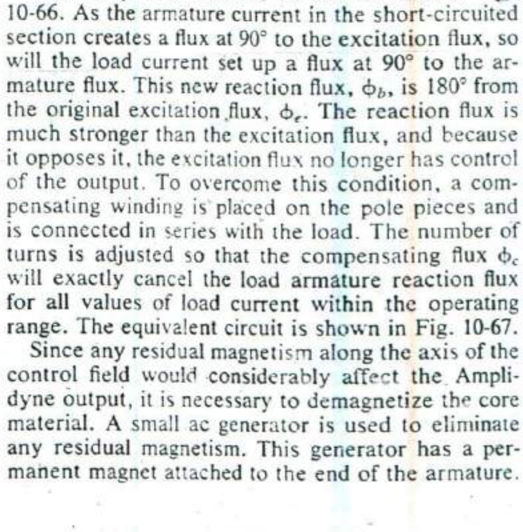


Fig. 10-67 Amplidyne generator equivalent circuit, showing magnetic fields.

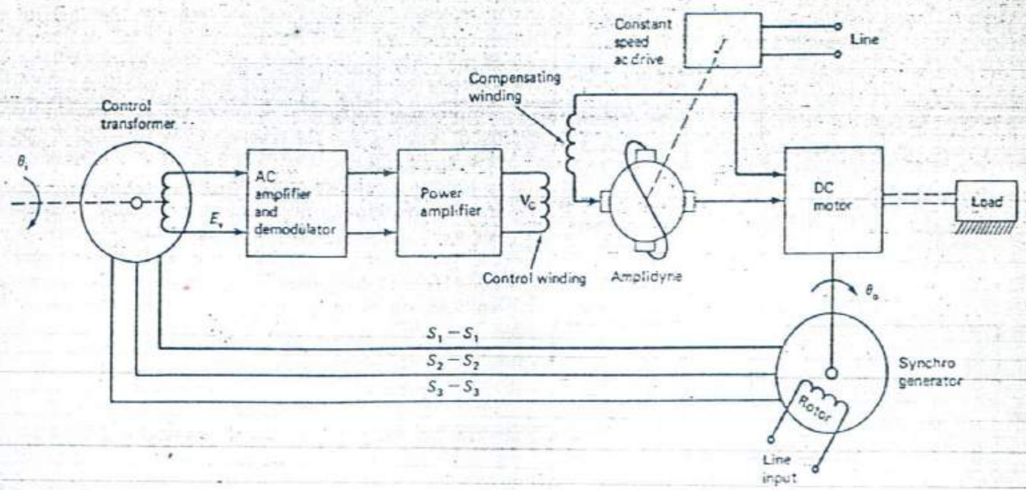


Fig. 10-68 Hydrid position control system.

The magnet revolves within a separate field winding and generates a small ac voltage, which is applied to the two sets of opposed windings on the field pole pieces. They are sometimes called *killer windings*. The generated alternating current neutralizes any residual magnetism when the control field is zero.

The power gain obtainable from an Amplidyne varies from 700 to 100,000. For example, typical gains are 5000 for a 500-W unit and 25,000 for an 8-kW unit. Figure 10-68 shows a type 0 position control system utilizing an Amplidyne with synchros as the feedback control elements.

REVIEW QUESTIONS

34. The Ward Leonard system is always an open loop system. (true or false)

35. The Ward Leonard system in Fig. 10-59 is ______-directional.

36. The Regulex generator typically uses a ______ field for self-excitation.

37. The adjustment of the field resistance in a Regulex generator is called ______.

38. In Fig. 10-62, voltage V_c is also called the ______ voltage.

10-7 TROUBLESHOOTING AND MAINTENANCE

Servomechanisms are much like other electronic systems in that the fundamental measurements of voltage, current, and resistance are of primary importance. The technician must test and analyze the system to ensure that it is performing properly and within specified tolerances. The best way to maintain a system is to be aware of its performance, so that if there is a tendency toward error it can be discovered and corrected before it becomes detrimental.

During their useful life, potentiometers will undergo certain changes in their characteristics. The two most common characteristics that change are linearity and noise. Linearity is checked with a suitable master potentiometer whose accuracy is known to be at least ten times that of the unit to be tested. The method of testing is shown in Fig. 10-69(*a*), where the master and the potentiometer under test are connected in parallel across a dc power supply. The wipers (sliders) of the two potentiometers are connected to the vertical (dc) input of an oscilloscope. The shafts of the two units are mechanically coupled so they rotate simultaneously. The oscilloscope deflection may be calibrated by means of a DVM for allowable linearity error. As the shafts are slowly rotated throughout the specified angular travel, any deviation can be observed in the oscilloscope. The horizontal sweep can be off during this test. In some test units, a third potentiometer is driven along with the other potentiometers. Its output can be used to deflect the oscilloscope beam horizontally, and the horizontal position of the beam will correspond to the shaft angle.

Noise is a common characteristic of potentiometers and tends to increase with the life of the unit. When a circuit is affected by noise, the potentiometer should be checked. Some commercial instruments to perform this measurement are available. If such a unit is not available, the set-up of Fig. 10-69(*b*) gives satisfactory results. The dc voltage is applied across the short-circuited ends of the potentiometer and the wiper arm through a 300 kΩ resistor. With the shaft slowly rotated throughout its range, the ac oscilloscope is monitored for any vertical deflection. The ac noise generated may be calibrated in *equivalent*

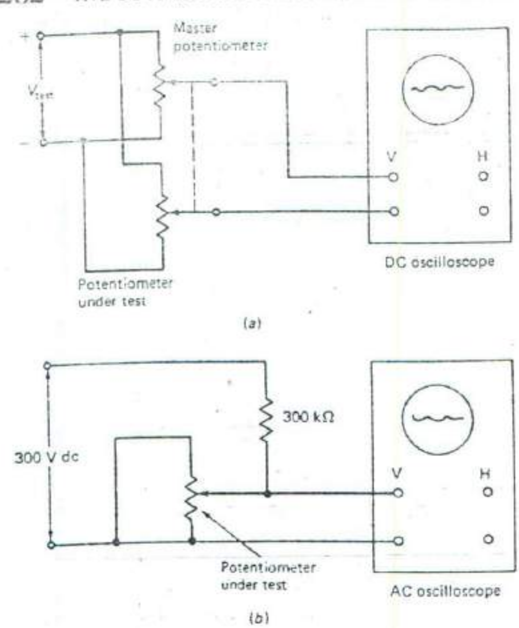


Fig. 10-69 Potentiometer test set-ups.

noise resistance (ENR) by alternately bridging the resistor in series with the wiper to obtain a given deflection in terms of the added resistance.

Most shaft encoders do not require periodic adjustments. A common failure is the internal light source, which may be solid-state or incandescent, with the latter more susceptible to failure. If the output of the encoder is missing on all channels, the light source should first be verified. Many units have modular construction for the electronics. Spare units should always be kept on site. As with any device that is connected to a shaft, loose connections are not uncommon and should not be discounted as a source of problems, especially when operation is erratic or intermittent.

Since the different units of a synchro system may be located at some distance from each other, they are connected by cables. Whenever a synchro system operates improperly, it is advisable to check the wiring of the unit (especially for loose or dirty connections) before looking for trouble in the synchros themselves. This is particularly important when working with systems which may have been under repair or overhaul. Should the symptoms indicate wiring trouble, it is suggested that all wiring be disconnected and checked for continuity.

Troubles involving open and short-circuited wiring, with the associated symptoms are listed for easy reference in Table 10-3. Should the symptoms indicate that the trouble might be in both the rotor and stator circuits, the rotor circuit should be checked first. To avoid electrical shock, all safety precautions should be observed at all times.

The nominal output of a resolver is known explicitly when the signal input is normal and the angle of the rotor is known. However, accurate measurement of the rotor angle in the field is often extremely difficult. A method of resolver testing, which does not require measurement of the rotor angle, is frequently used. Assume that the resolver to be tested is of the type having both sine and cosine outputs. If the input is 10 V, it is apparent that for normal operation the output voltages are $10 \cdot \cos\theta$ and $10 \cdot \sin\theta$. The rotor must be locked (clamped) so it does not turn. If the two voltmeter readings are first squared and added, the result should equal the square of the input voltage, since

$$(10\cos\theta)^2 = 100\cos^2\theta$$
$$(10\sin\theta)^2 = 100\sin^2\theta$$

and

$$(100\cos^2\theta + 100\sin^2\theta) = 100(\sin^2\theta + \cos^2\theta) = 100$$

This result is based on the trigonometric identity

$$\sin^2\theta + \cos^2\theta = 1$$

If the input voltage is 10 V and the rotor angle is 30°, then

$$(10\sin 30)^2 = (5)^2 = 25$$
$$(10\cos 30)^2 = (8.67)^2 = 75$$

The sum of these squares is seen to be equal to 100 V, which is equal to the input voltage squared. Thus, voltmeter readings of 5V and 8.67V indicate a properly functioning resolver in this example.

The typical ac motor used in servomechanisms is a two-phase induction motor. The important aspects of the motor are that the two stator windings (control and reference) are operated 90 degrees electrically from each other. The voltages fed to the stator windings must be 90° out of phase. Since a two-phase supply is rarely available, it is common to operate the two-phase motor with the phase-splitting capacitor. If the motor runs slow or is sluggish, the capacitor is suspect and should be tested. Improper phase shift between the windings is a common clue. An increase in bearing friction can have a detrimental effect on a motor, especially at low speeds, and should not be overlooked. Any suspected leakage or short circuit in the stator windings should be tested as specified in earlier chapters.

Tachometers that are used in accurate velocity servomechanisms must meet specific accuracies. The signal-to-noise ratio of a tachometer is usually specified by the manufacturer, and the unit must always meet these requirements. In dc types, high output noise is often an indication of bad brushes or a dirty commutator.

With ac tachometers, variation in amplitude of the residual voltage may have a modulation effect on a servomechanism. This measurement can be made by using the potentiometer method with an oscilloscope as a null detector. The output phase of the ac tachometer should remain constant. It is specified by

TABLE 10-3 SYNCHRO SYMPTOMS AND REMEDIES

Symptoms	Possible Cause of Trouble	Remedy
Receiver rotor either in correspondence with transmitter of 180° displaced, but follows in proper direction. Stator voltages vary from 0 to 90 V. Both rotor voltages are 115 V.	Rotor winding open, connection to slip ring open, or brush not making contact.	If ascertained trouble is not in ring connection or brush, unit must be replaced.
Same as above except that one rotor voltage is 115 V and the other is 90 V.	Supply line is open to the rotor reading 90 V, the 90 V appears across the rotor by virtue of transformer action.	Locate open in supply line; repair.
Voltage between one pair of stator wires is zero for all transmitter positions. Other stator-lead voltages read from 0 to 90 V. Both rotor voltages are 115 V.	The pair of stator leads which read 0 V is short-circuited.	Remove short circuit from wiring or interconnecting switches. If it is internal, unit may require replacement.
Both transmitter and receiver units hum and heat excessively. Receiver either does not follow or may spin.	All three stator wires are short-circuited together.	Locate defective wiring or switches; repair.
Sudden change in transmitter rotor position causes oscillation at receiver or a spinning effect.	Inertia damper jammed tight on receiver rotor shaft. Absence of damper indicates that transmitter unit is being used.	Free damper if it is jammed. If transmitter has been used, replace with a receiver unit.
Intermittent operation	Corroded rings, defective brushes, loose connections.	Respectively, clean rings, install new brushes, tighten loose terminals, etc.
Torque normal. Receiver lags or leads the transmitter or may turn in proper direction or reverse direction.	Stator wiring incorrect.	Correct stator wiring.
Torque normal. Receiver follows transmitter, but is displaced 180° from it.	Rotor connections reversed.	Correct wiring at proper unit.
Receiver shows large error and lags transmitter. Connections normal, but excessive current flows, producing overload indication.	Bearings frozen or partially frozen because of improper lubrication.	Replace unit, since bearing trouble usually damages other parts of the unit.

the manufacturer, and its value should be checked to ensure that the unit is within the limits specified. It is often easy to couple a rate generator to a variable-speed drill. Use a calibrated strobe to ascertain its output magnitude and direction. Be sure to check both directions of rotation.

Most servomotors operate most efficiently at higher speeds than are actually necessary to drive the load. Gearing is quite common to convert the low-torque high-speed motor output to a high-torque low-speed output for the load. The main problems encountered in gearing are backlash and friction. Backlash introduced by the gear train on the order of only a fraction of a degree can have a serious effect on the stability of a servomechanism. Backlash introduces a time delay between the servomotor and the input command signal. As the amount of backlash increases, so will the oscillations of the servomechanism. Even small oscillations caused by backlash will eventually cause excessive equipment wear.

Static friction, due to the tightness of the gearing, is probably the best deterrent to backlash oscillations. But friction problems will occur if the meshing of gears is not concentric with their supporting shafts. When gears are rotated, eccentricities combine in the gear box to produce excessive friction at one point of the revolution and backlash at another point.

Clutches, by their nature, are susceptible to wear more than any other mechanical part of a system. If the clutch is adjustable, the adjustment mechanism can work loose, especially if the system was or is oscillatory. The clutch surface must be clean and free of any contaminants, especially oil or grease. If light sanding with a recommended abrasive does not rectify the problem, a new clutch should be considered.

The typical servo amplifier is a summer (adder) or difference (subtractor) amplifier. In many cases, the reference or feedback signals may come from remote locations; noise pickup can be a problem, especially if a shield connection is loose or broken. Dirt or grease can cause electrical leakage and cause a ground loop which makes the system noisy or unstable. An oscilloscope is recommended for close inspection of the signals for the presence of noise. The low-voltage power supplies should always be verified as being within tolerances.

The components used in control systems have high precision in most cases and will be adversely affected if overheated by improper ventilation. Loss of feedback will cause the amplifier to lock up to a power supply rail, unless an overload circuit is incorporated. In some instances, the feedback resistance can be shunted with a value to make the gain approximately unity. The decreased gain will allow the system to be analyzed under more reasonable conditions.

Loss of feedback can also cause a servomechanism to go to the extreme end (limit) if it is a position type, or turn at its maximum rate (velocity) if it is a rate type. In either case, a potentiometer or an adjustable power supply may be substituted to simulate the closing of the loop. Most systems have either static or dynamic tests that can be performed to aid in localizing any problems. Refer to the manufacturer's service manuals for the necessary details. Remember that any significant gain or feedback change can cause a system to become unstable. This can be dangerous in some instances. You must know and thoroughly understand all tests before performing them.

Block-level understanding is necessary to localize a problem to a particular block. If the system is dead to any input command, signal tracing through the blocks should lead to the malfunctioning block. Final fault isolation is done with conventional voltage or resistance tests.

Loose or slipping follow-up potentiometers or rate generators are not uncommon. Most of these units have specific alignments and if not properly set can cause position or rate feedback errors and imbalances in the system. If a servomechanism is at its limit and no clutch is used, shut it down immediately to prevent damage to the mechanical components (gears, couplings, etc.). A good stock of replacement modules is a must to minimize down time. When a replacement is made, be sure the malfunctioning unit is repaired or returned to the manufacturer for prompt return to the spares stock or returned to the manufacturer.

Rotating amplifiers, as do other components with brushes and commutators, require careful maintenance procedures. Precautions with lubrication, according to the manufacturer's recommendations, are essential. The Amplidyne has multiple sets of brushes, and proper attention is a must. In most cases, the output from a large servomotor, either position-type or rate-type, is fed back to close the systems loop. This factor must be taken into account when energizing the field of a rotating amplifier with an auxiliary source. Remember that these units can have current gains in the thousands. A few milliamperes of Amplidyne field current can produce amperes of output current for a servomotor or load.

The loss of the killer winding will leave a residual field in the Amplidyne, thereby causing an imbalance in the system. Leakage resistance to the frame is unacceptable, as with any type of generator, and should not be overlooked if overheating occurs. Routine preventive maintenance is essential for all rotating devices and will enhance the operation of the units and extend the operating life while minimizing down time.

REVIEW QUESTIONS

39. The ______ source is a cause of frequent failures in a shaft encoder.

40. If a synchro receiver follows the transmitter but with 180° of error, this is an indication of rotor connection ______.

41. The voltmeter reading of a resolver winding is the square of the applied voltage. (true or false)

42. The phase difference between the reference and control windings of the ac servomotor is ______ degrees.

43. Loss of ______ can cause a type 0 servo to travel to one of its limits.

CHAPTER REVIEW QUESTIONS

10-1. What type of encoder retains position information during a power outage?

10-2. Incremental encoders can provide position and ______ information.

10-3. A linear motion optical encoder is a type of absolute encoder. (true or false)

10-4. The resolver solves the unknowns of a ______ triangle.

10-5. If $V \sin \omega t$ is 100 V and θ is 45°, VR_3 to VR_4 is ______?

10-6. The output voltages from a resolver represent cartesian or rectangular coordinates. (true or false)

10-7. What happens to the output of an ac servomotor when the control voltage is shifted by 180°?

10-8. A dc rate generator is free of any zero speed ______ voltage.

10-9. Alternating current rate generators produce an output frequency proportional to their speed. (true or false)

10-10. The mutual coupling of the stator windings in an ac servomotor is due to the induced ______ currents in the rotor.

10-11. Direct current servomotors require standby power. (true or false)

10-12. Most servomechanisms use ______ type gears.

10-13. Mechanical addition and subtraction with gearing is accomplished by the use of a ______.

10-14. A motor with an inertia of 0.2 $kg \cdot m^2$ that is coupled through a gear box of 20:1 can handle a load inertia of ______ $kg \cdot m^2$.

10-15. In a simple feedback system, the *H* block represents the feedback. (true or false)

10-16. A servosystem is usually ______ damped for best performance.

10-17. A type 0 servosystem is a rate (velocity) system. (true or false)

10-18. A type 0 servosystem is a position (follow-up) system. (true or false)

10-19. Feedback in a type 1 servosystem is usually from a ______.

10-20. Stability of a closed loop system may be improved by addition of an integrator or a ______.

10-21. Double differentiating a position signal will produce a velocity signal for damping purposes. (true or false)

10-22. The ______ windings correct for the effect of load flux on the control excitation in an Amplidyne.

10-23. The brushes of the Amplidyne are ______ degrees apart.

10-24. An Amplidyne needs a residual flux to establish its excitation flux. (true or false)

10-25. The Amplidyne is a type of ______ amplifier.

10-26. The typical servo amplifier is usually a summer or a ______ amplifier.

10-27. The Amplidyne requires periodic checking of its two sets of ______.

10-28. The Amplidyne's control field is a high-current, low-voltage winding. (true or false)

10-29. The maximum torque of a two-phase servomotor occurs at ______ rpm.

10-30. The output phase of the ac tachometer should change as the shaft angle changes speed. (true or false)

10-31. Lost motion, or play, in a mechanism is known as ______.

ANSWERS TO REVIEW QUESTIONS

1. cermet, wire-wound, and conductive plastic **2.** slider **3.** 7.71 kΩ **4.** linearity **5.** hunt **6.** false **7.** 90 **8.** false **9.** 52 V **10.** false **11.** clockwise **12.** 160° **13.** control **14.** phase **15.** 90 **16.** capacitor **17.** false **18.** drag-cup **19.** 24,000 **20.** 16.7 **21.** false **22.** bore (inside) **23.** false **24.** universal joints **25.** clutch **26.** speed **27.** −4 **28.** −1 V **29.** 6 **30.** power **31.** chopper **32.** overload recovery **33.** *Q/K* **34.** false **35.** uni **36.** shunt **37.** tuning **38.** error **39.** light **40.** reversal **41.** false **42.** 90 **43.** feedback