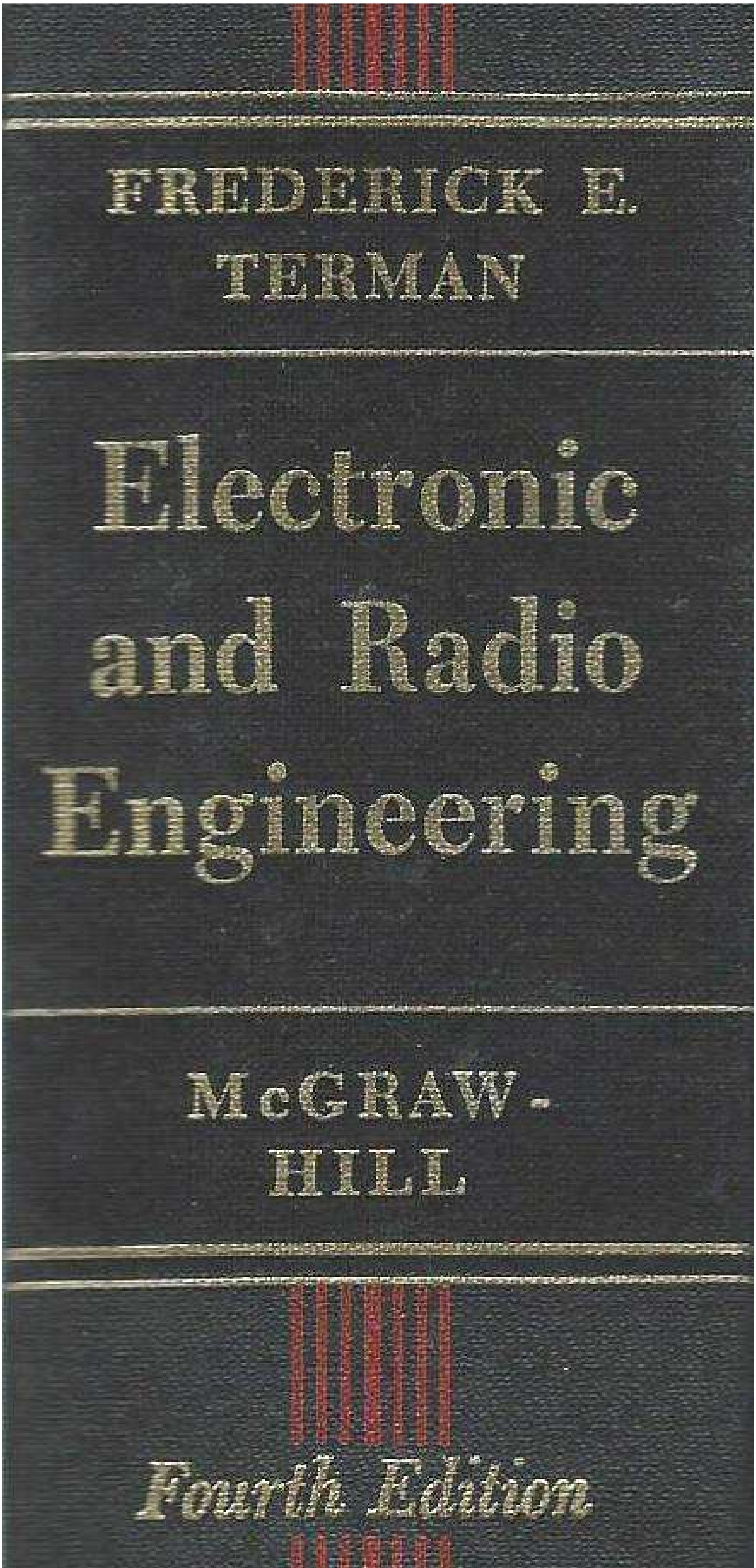

FREDERICK E.
TERMAN
Electronic
and Radio
Engineering
McGRAW-
HILL
Fourth Edition

from zero to infinity.[1] Although the relationship between the phase shift and the transconductance is not linear, neither is the relationship between the modulating voltage applied to the control grid and the resulting transconductance of the tube. By properly balancing these nonlinearities against each other, it is easily possible to obtain a relationship between phase shift and modulating voltage that is reasonably linear up to phase shifts of the order of ± 1.0 radian.

In phase-shifter modulators, the maximum modulation index obtainable can be readily increased by connecting several modulators in cascade. A two-stage system of this type is illustrated schematically in Fig. 17-13*b* and gives twice as large a modulation index as does a single-stage arrangement; it is possible, morever, to employ as many stages as desired.

Systems of frequency modulation based on phase modulators have the advantage that the carrier frequency can be obtained directly from a crystal oscillator. They have the disadvantage, however, that the largest modulation index that can be obtained is smaller than in the case of a frequency-modulated oscillator. In order to obtain the relatively large values of modulation index required when a large frequency deviation is desired at a low modulation frequency, it thus becomes necessary to depend upon frequency multiplication in order to increase m_p to the value called for by Eq. (17-16). The amount of multiplication required will depend upon the modulation index that is initially produced. It will, for example, be much less with a cascaded system using a number of phase shifters of the type illustrated in Fig. 17-13 than with a one-stage system of the type illustrated in Fig. 17-11, which gives a maximum value of m_p of the order of 0.25.

17-6. Detection of Frequency- and Phase-modulated Waves. Detection of a frequency- or phase-modulated wave is ordinarily carried out by modifying the frequency spectrum of the wave in such a manner that its envelope fluctuates in accordance with the intelligence involved. The resulting amplitude-modulated wave is then applied to an ordinary

[1] This is demonstrated as follows: Referring to Fig. 17-13*a*, one can write:

$$I = E_2 g_m \qquad (17\text{-}33a)$$

$$E_2 = (E_1 + E_3)/2 \qquad (17\text{-}33b)$$

$$E_1 - E_3 = -j2X_c I = -j2X_c E_2 g_m \qquad (17\text{-}33c)$$

Substituting Eq. (17-33*b*) into Eq. (17-33*c*) to eliminate E_2 and reducing the result gives

$$\frac{E_3}{E_1} = \frac{\text{output}}{\text{input}} = \frac{1 + jX_c g_m}{1 - jX_c g_m} \qquad (17\text{-}34)$$

Since $(1 + jX_c g_m)$ and $(1 - jX_c g_m)$ are conjugate quantities, the right-hand side of Eq. (17-34) is equal to unity irrespective of the value of g_m; however, the phase varies from 0 to 180° as g_m changes from zero to infinity.

amplitude-modulation detector.[1] The circuit arrangement that transforms the frequency-modulated signal into a wave possessing amplitude modulation is termed a *discriminator*.

The detuned resonant circuit illustrated in Fig. 17-7 represents a simple form of discriminator. Here variations in the instantaneous frequency of the applied wave produce corresponding variations in the amplitude of the response of the resonant circuit. Such an arrangement has the disadvantage, however, that the side of a resonance curve cannot be regarded as particularly linear except over a very limited frequency range; also the characteristics of such a discriminator depend rather critically on the amount of detuning of the resonant circuit.[2]

The Phase-shift Discriminator. The most widely used form of discriminator is the arrangement shown in Fig. 17-14*a*, in which the two tuned circuits P and S are resonant at the same frequency and are inductively coupled. This arrangement depends upon phase shift for its operation and so is commonly called a *phase-shift discriminator*.[3]

The action of the discriminator of Fig. 17-14*a* can be explained as follows: The center of the secondary is connected to the top (high-potential) side of the primary P by the capacitor C that blocks the d-c plate voltage from the secondary system; C serves as a by-pass to signal frequency but need be no larger than required to do this. Associated with C is radio-frequency choke L that provides a return path for the d-c component of the rectified current flowing through diodes T_1 and T_2. The inductance of L is effectively in shunt with inductance of P and is preferably considerably larger than L.

The radio-frequency voltages E_{a1} and E_{a2} applied to the two diodes are $E_3 + E_1$ and $E_3 - E_2$, respectively, where E_3 is the voltage across P and E_1 $(=E_2)$ is the vector voltage across half the secondary coil as indicated in Fig. 17-14.

The phase relations existing in the discriminator are shown in Fig. 17-14*b*. At the resonant frequency of the tuned secondary circuit, the

[1] Other methods of detecting a frequency-modulated wave are possible and find some use. In particular, when a very linear relation is required between the detector output and the variations in the instantaneous frequency, as in measurement work, a cycle-counting type of frequency meter as in Fig. 18-47 is used. Such an instrument will at any instant develop an output current exactly proportional to the instantaneous frequency; for further details see S. W. Seeley, C. N. Kimball, and A. Barco, Generation and Detection of Frequency-modulated Waves, *RCA Rev.*, vol. 6, p. 269, January, 1942; F. E. Terman and J. M. Pettit, "Electronic Measurements," p. 2[illegible], McGraw-Hill Book Company, Inc., New York, 1952.

[2] An analysis of this type of discriminator is given by A. R. Vallarino and Marvin Buyer, Harmonic Distortion in Frequency-Modulation Off-resonance Discriminators, *Elec. Commun.*, vol. 26, p. 167, June, 1949.

[3] The phase-shift discriminator was originally developed as a means of obtaining automatic frequency control; see D. E. Foster and S. W. Seeley, Automatic Tuning, Simplified Circuits and Design Practice, *Proc. IRE*, vol. 25, p. 289, March, 1937.

secondary voltages E_1 and E_2 are in quadrature with the voltage E_3 existing across the primary inductance.[1] However when the applied frequency is either higher or lower than the resonant frequency of the secondary, the phase position of E_1 and E_2 relative to E_3 will differ from 90°.

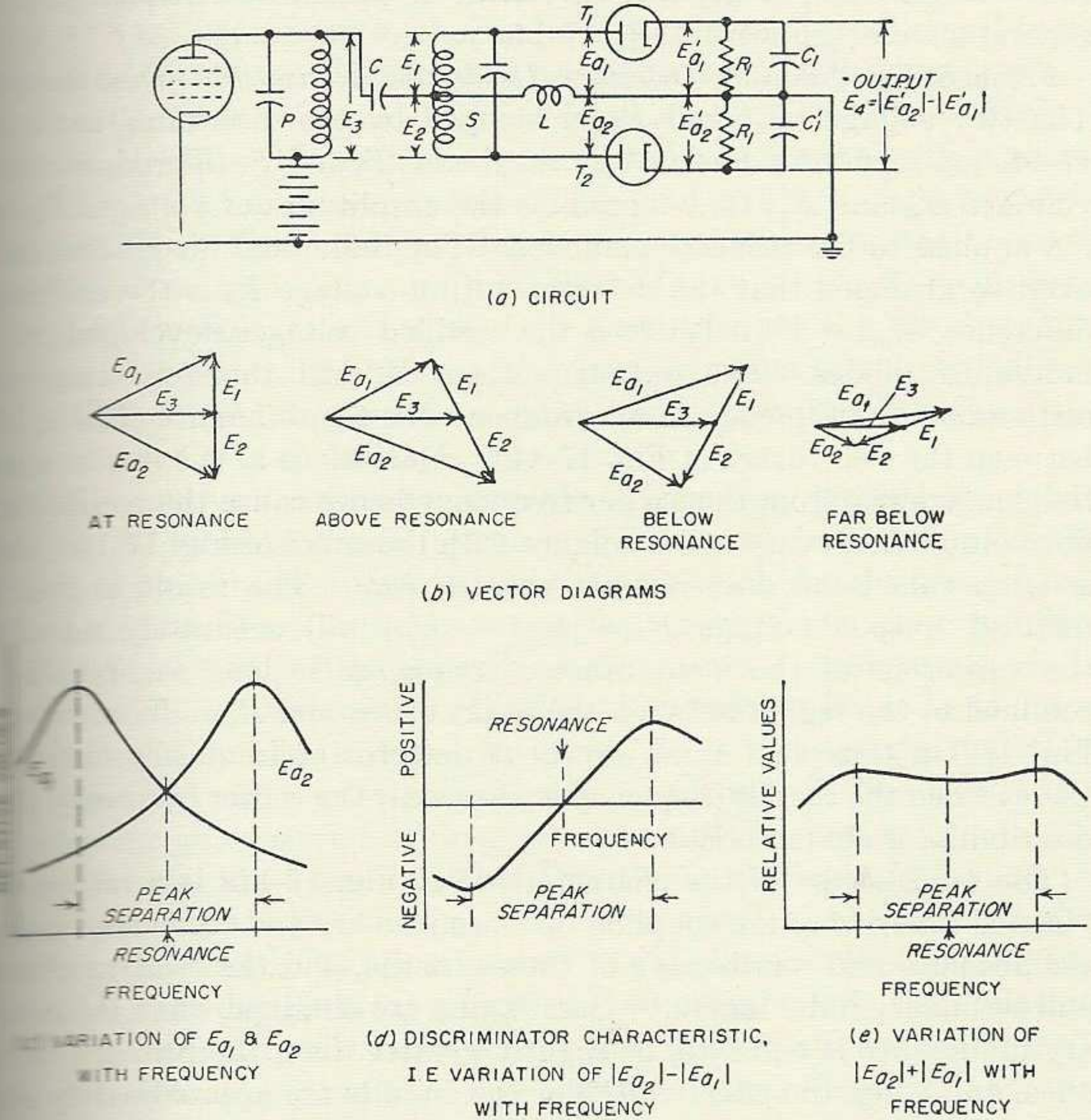


Fig. 17-14. Frequency-modulation detector employing phase-shift discriminator.

Thus when the instantaneous frequency differs from the resonant frequency f_0 by $f_0/2Q_s$ cycles, the phase shift will be 45° (or 135°). The result of this situation is that at resonance the two resultant voltages E_{a1}

[1] This assumes that the impedance $(\omega M)^2/R_s$ that the secondary couples into the primary at resonance is small compared with the inductive reactance of the primary, which will always be true in practice. With this simplification a voltage E_3 across the primary inductance induces a voltage in series with the secondary circuit that is in phase with E_3. However, when the secondary is tuned to resonance, the voltage developed across the secondary inductance (or capacitance) is 90° out of phase with the voltage induced in series with the secondary circuit. Thus the secondary voltage $E_1 - E_2$ is 90° out of phase with the primary voltage E_3 when the applied frequency coincides with the secondary resonant frequency.

and E_{a2} are equal in amplitude, but at frequencies slightly below resonance the amplitude of one of these voltages is decreased while that of the other becomes larger. Above resonance the situation is reversed. This is illustrated by the vector diagrams in Fig. 17-14*b*. The amplitudes of the voltages E_{a1} and E_{a2} will vary with instantaneous frequency in the general manner[1] shown in Fig. 17-14*c*.

Frequency-modulation Detectors Using the Phase-shift Discriminator. The two voltages E_{a1} and E_{a2} developed by the discriminator in Fig. 17-14*a* are separately rectified by the diodes T_1 and T_2 to produce output voltages E'_{a1} and E'_{a2} that reproduce the amplitudes of voltages E_{a1} and E_{a2} applied to the respective anodes. The individual diodes are, moreover, so arranged that the detector output voltage E_4 is the arithmetic difference $|E'_{a2}| - |E'_{a1}|$ between the rectified voltages developed by the individual diodes. The output voltage E_4 will therefore vary with instantaneous frequency in accordance with the difference $|E_{a2}| - |E_{a1}|$ between the two curves of Fig. 17-14*c*. Deviations in the instantaneous frequency away from the carrier frequency hence cause the rectified output voltage E_4 to vary in accordance with the curve of Fig. 17-14*d*, which is often called the *discriminator characteristic.* The result is that the rectified output voltage $E_4 = |E'_{a2}| - |E'_{a1}|$ will accurately reproduce the variations of the instantaneous frequency as long as operation is confined to the region between the peaks of E_{a1} and E_{a2}. The system of Fig. 17-14*a* thus acts as an excellent detector of frequency-modulated waves when the carrier frequency is at or near the center frequency of the discriminator characteristic.

The exact shape of the characteristic of Fig. 17-14*d* is a rather complicated function of the coupling between primary and secondary circuits, the absolute and relative Q's of these circuits, and the relative primary and secondary inductances.[2] Best results are obtained when the secondary inductance is equal to, or slightly greater than, the primary inductance, and when the effective Q's of the circuits are approximately equal.

[1] The exact details are complicated by the fact that the voltages E_3 and $E_1 - E_2$ change in magnitude as well as relative phase as the frequency varies. Thus when the two resonant circuits are overcoupled, as is customary, E_3 and $E_1 + E_2$ have a double-peaked characteristic (See Sec. 3-5). Under these circumstances, increasing deviation of the signal frequency first causes these voltages to become larger until the frequency of the resonant peak is reached, after which both voltages rapidly diminish in amplitude. This is shown in Fig. 17-14*b*, where the second and third vector diagrams are for instantaneous frequencies closer to resonance than the coupling peak, whereas the final diagram applies to an instantaneous frequency somewhat below the low-frequency coupling peak.

[2] Quantitative analyses of discriminator behavior are given by K. R. Sturley, The Phase Discriminator, *Wireless Eng.*, vol. 21, p. 72, February, 1944; W. G. T... and T. P. Cheatham, Jr., Adjustable Bandwidth FM Discriminator, *E...*, vol. 20, p. 117, September, 1947.

when the loading of the diodes is taken into account.[1] The coupling should simultaneously be of the order of twice the critical value.

When properly designed, the phase-shift discriminator will give a very linear relation over a range of instantaneous frequencies only slightly less than the frequency separation of the peaks of the individual curves of Fig. 17-14c. This peak separation, which therefore must exceed twice the peak deviation Δf, is determined primarily by the Q's of the resonant circuits of the discriminator, but is affected somewhat by the coefficient of coupling. When the coefficient of coupling is twice the critical value, the peak separation approximates $2f_0/Q$, and will be proportionately greater if the coefficient of coupling is higher. Here f_0 is the resonant frequency of the secondary, and it is assumed that the Q's of the primary and secondary circuits are equal. In proportioning the discriminator circuits, the Q that should be used is hence determined by the peak deviation Δf of the instantaneous frequency; for twice critical coupling, the proper value of Q is slightly less than $f_0/\Delta f$.

Many detailed variations are possible in frequency-modulation detectors using the phase-shift discriminator. The circuit shown in Fig. 17-14a is particularly suitable for explaining the principles of operation. Practical arrangements are, however, more likely to resemble the form illustrated in Fig. 17-15. Here the ground connection is made to one of the output terminals, which causes capacitor C to perform the same function as capacitor C'_1 in Fig. 17-14a. In Fig. 17-15 it is necessary that C be small enough to offer a high impedance to modulation frequencies and yet large enough to serve as a by-pass to the radio-frequency signal. By rearranging the capacitor C_1 of Fig. 17-14a as shown in Fig. 17-15, the choke L of Fig. 17-14a is no longer necessary. The circuit of Fig. 17-15 functions in exactly the same way as does the circuit of Fig. 17-14a, but has one output terminal grounded, and requires one less capacitor and no radio-frequency choke. As an aid in tracing out the correspondence between these two circuits, the rectified voltages E'_{a1} and E'_{a2} appearing in different parts of the respective output systems are indicated in each circuit.

In practical arrangements, it is also customary to obtain the output through a resistance-capacitance combination R_cC_c, as shown dotted in Fig. 17-15. In this way, there is no d-c voltage transmitted to the out-

[1] The input resistance of each diode in Fig. 17-14a is $R_1/2\eta$, where η is the efficiency of rectification. With respect to the voltage $E_1 + E_2$ existing between the terminals of the secondary, the input resistances of the individual diodes are in series so that the diodes place a load R_1/η across the full secondary. However, to the voltage E_3 developed across the primary circuit P the diode inputs are in parallel. The equivalent load resistance that the diodes place on the primary is hence $R_1/4\eta$, a much lower value than the load on the secondary. To achieve equality of Q's, it is therefore often found necessary to shunt the secondary with an additional resistance.

put terminals when the average or carrier frequency of the signal differs from the center frequency of the discriminator.

Frequency-modulation detectors based on the circuit of Fig. 17-14a (or Fig. 17-15) have the disadvantage that variations in the amplitude of the applied voltage produce a proportional change in the amplitude of the discriminator characteristic; this is shown in Fig. 17-16. Thus when

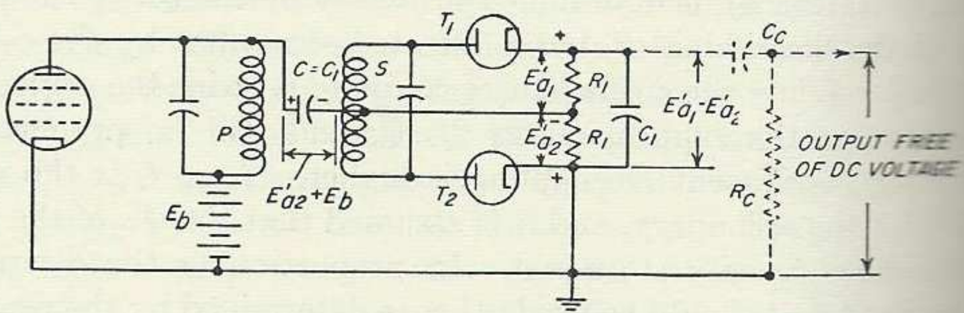


FIG. 17-15. Practical form of phase-shift discriminator of Fig. 17-14a.

the applied signal consists of a frequency-modulated wave that also varies in amplitude, the detector output will contain undesired components corresponding to the amplitude variations as well as the output representing the frequency modulation. This situation is generally to be avoided, since amplitude variations are commonly the result of interfering effects, such as noise. This limitation of detectors of the type shown in Figs. 17-14a and 17-15 is overcome in a modification known as the ratio detector.

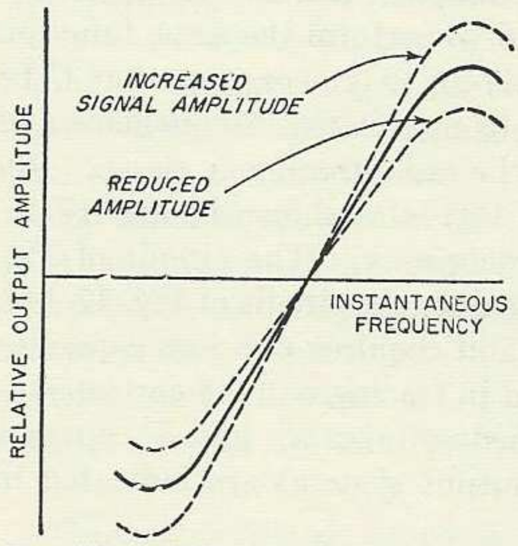


FIG. 17-16. Effect produced on the discriminator characteristic of Fig. 17-14d by changes in the amplitude of the applied frequency-modulated signal.

17-7. The Ratio Detector.[1] The ratio detector is a modification of the phase-shift discriminator detector of Fig. 17-14a, which can be so designed as to be unresponsive to amplitude modulation while behaving toward frequency modulation in the same way as the detector of Fig. 17-14a.

The circuit of a simple form of the ratio detector is shown in Fig. 17-17a. Neglecting the capacitor C_2 for the moment, this arrangement is seen to differ from the detector of Fig. 17-14a only in that (1) diode T_2 has been reversed in polarity, and (2) the output voltage is obtained between ground and the center tap on the high resistance R_2 that shunts the load impedance of the two diodes.[2]

[1] For further discussion of this subject see S. W. Seeley and J. Avins, The Ratio Detector, *RCA Rev.*, vol. 8, p. 201, June, 1947.

[2] The radio-frequency choke and the capacitor C in Fig. 17-17a must meet the same requirements as in Fig. 17-14a.

It will now be shown that the output voltage in Fig. 17-17*a* varies with instantaneous frequency in exactly the same way as it does in the circuit of Fig. 17-14*a*, but is only half as great. To do this, it is to be noted first that the individual output voltages E'_{a1} and E'_{a2} developed by diodes T_1 and T_2 have the same magnitude as before. However, E'_{a2} is now reversed in polarity, so that the voltage E_4, instead of being $|E'_{a2}| - |E'_{a1}|$, as in Fig. 17-14*a*, is now $|E'_{a1}| + |E'_{a2}|$. The output voltage in Fig. 17-17*a* is the potential between the midpoint of R_2 and ground; its value is the potential E'_{a2} at the lower end of R_2, minus half the total voltage E_4 developed across R_2. Thus

$$\left.\begin{array}{l}\text{Output voltage}\\ \text{in Fig. 17-17}a\end{array}\right\} = |E'_{a2}| - \frac{|E'_{a1}| + |E'_{a2}|}{2} = \frac{|E'_{a2}| - |E'_{a1}|}{2} \qquad (17\text{-}35)$$

This is exactly half the magnitude of the output obtained from the system of Fig. 17-14*a*. Thus, the ratio detector responds to variations in instantaneous frequency in exactly the same way as does the system of Fig 17-14*a*.

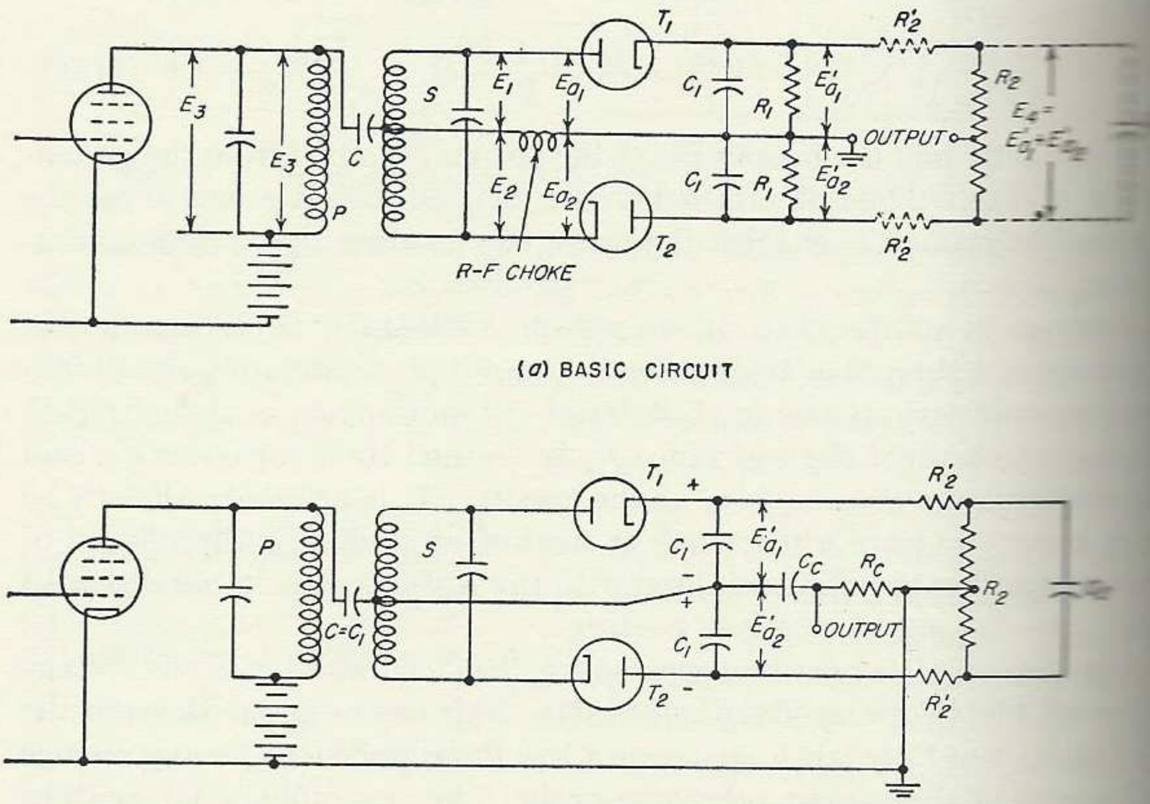
Suppression of Response to Amplitude Modulation Occurring Simultaneously with Frequency Modulation.[1] Amplitude modulation simultaneously present with frequency modulation will not appear at the output of the ratio detector if the resistance R_2 is shunted by a capacitor C_2, and suitable resistors R'_2 are added to the circuit. It is necessary that C_2 be large enough to have a reactance at the lowest modulation frequency of importance that is small compared with the resistance R_2 in parallel with $2R'_2$.

The effect of C_2 is to reduce greatly amplitude fluctuation in the voltage $E_4 = |E'_{a1}| + |E'_{a2}|$ appearing across R_2. This comes about through the fact that when C_2 is large, it acts as a low impedance load to any change in amplitude that might otherwise occur. For example, a momentary increase in amplitude of E_4 causes a large charging current to flow through the diodes into C_2. This represents power absorbed by the diodes from the resonant circuits P and S, and so causes the voltages that these circuits apply to the diodes to be reduced in magnitude. Conversely, if the amplitude of the incoming signal attempts momentarily to drop below the average amplitude, then C_2 attempts to prevent the voltage E_4 from dropping by supplying current that flows from C_2 into R_2 and R_1. This relieves the diode tubes of the necessity of supplying as much rectified current as before, thereby increasing their input impedance and reducing the loading on the resonant circuits P and S, with corresponding increase in the voltages they apply to the diodes. It is thus seen that the presence of C_2 reduces (but does not entirely eliminate) the amplitude variations that

[1] A quantitative analysis of amplitude-modulation rejection in the ratio detector is given by B. D. Loughlin, The Theory of Amplitude-modulation Rejection in the Ratio Detector, *Proc. IRE*, vol. 40, p. 289, March, 1952.

would otherwise occur in the voltage E_4, and likewise in voltages E_1, E_2, and E_3. It will also be noted that these variations in input impedance with amplitude modulation that are present also produce corresponding variations in the Q of the primary and secondary circuits P and S. Specifically, the effective Q's of the circuits will be reduced during the periods when the amplitude is greater than the carrier value, and will be increased when it is less.

Consider now the situation that exists when the incoming signal is a pure frequency-modulated wave. When C_2 is disconnected, the voltage



(b) PRACTICAL FORM OF RATIO DETECTOR CIRCUIT

Fig. 17-17. Ratio-detector circuits.

$|E_{a1}'| + |E_{a2}'|$ across resistance R_2 is proportional to $|E_{a1}| + |E_{a2}|$. This voltage varies with instantaneous frequency in the manner illustrated in Fig. 17-14e, and is seen to be substantially constant in the range of frequencies between the peaks of E_{a1} and E_{a2}. When C_2 is connected, it tends to make $|E_{a1}'| + |E_{a2}'|$ even more nearly constant than in Fig. 17-14e. This modifies the output voltage slightly in a way that is equivalent to increasing the discriminator output for instantaneous frequencies near the center frequency of the discriminator characteristic of Fig. 17-14d, and decreasing it slightly for instantaneous frequencies near the peaks. These effects are trivial in magnitude, however, and for all practical purposes the presence of C_2 can be regarded as having negligible influence on the detection of a wave possessing pure frequency modulation.

Next examine what happens when the amplitude of the frequency-modulated signal varies. The voltages E_1, E_2, and E_3 developed across the discriminator in Fig. 17-17*a* show corresponding but lesser variations in magnitude, as discussed above. At the same time, the effective Q of the resonant circuit S is altered in such a manner that Q becomes less when the amplitude increases above the average, and vice versa, as explained above. This change in secondary Q affects the phase relations between voltages E_3 and $E_1 + E_2$ in Fig. 17-14*b*. The consequences of these two actions have opposite effects on each other, because an increase in the amplitude of $E_1 + E_2$ and E_3 tends to make the detector output greater, while the reduction in Q and hence of phase shift that goes with increased amplitude of the signal tends to make the detector output less. Thus by properly controlling the relative magnitudes of these two effects, they can be balanced against each other. When this is done, the output will be determined only by the variations in instantaneous frequency and the average amplitude of the incoming signal, and will be unaffected by amplitude variations at any modulation frequency for which C_2 is an effective by-pass.[1]

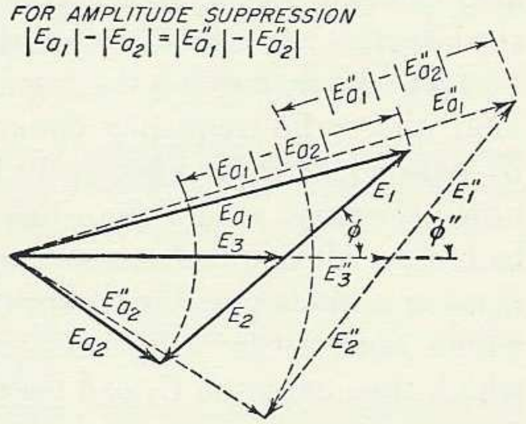


Fig. 17-18. Vector diagrams applying to a ratio detector, showing how a change in the amplitude of the applied voltages E_1, E_2, and E_3 is prevented from changing the discriminator voltage $|E_{a2}| - |E_{a1}|$ by the fact that as the vectors change in length, the angle ϕ simultaneously changes in magnitude.

A vector diagram illustrating this behavior is shown in Fig. 17-18. Here the solid vectors correspond to the voltages E_1, E_2, and E_3 in the absence of amplitude modulation when the instantaneous frequency is such as to cause the secondary circuit to shift the phase of the secondary voltages E_1 and E_2 by an amount ϕ. The dotted vectors illustrate what happens when the amplitude of the incoming waves is momentarily increased while maintaining the same frequency deviation. The corresponding vectors E''_1, E''_2, and E''_3 are now longer than before, but the phase shift ϕ'' produced by the same frequency deviation is less. With the conditions shown in Fig. 17-18, this change in phase angle is just enough to compensate for the increased length of the vectors, and results in the difference $|E''_{a1}| - |E''_{a2}|$ being the same for the dotted system as for the difference $|E_{a1}| - |E_{a2}|$ of the solid-line system.

The relative magnitudes of the effects that are thus to be balanced against each other can be controlled practically by means of the resistances R'_2 shown in Fig. 17-17*a*. Increasing these resistances will decrease the variations in diode input resistance caused by a given amount of amplitude modulation; this causes the voltages E_1, E_2, and E_3 that actu-

[1] The name *ratio detector* arises from the fact that the variations in detector output that occur as the instantaneous frequency changes arise as a result of variations in the ratio $|E'_{a1}/E'_{a2}|$, while the sum $|E'_{a1}| + |E|'_{a2}$ remains substantially constant.

ally get applied to the diode to have their magnitude changed more and their phase position changed less by the amplitude modulation. Decreasing R_2' will have the opposite effect, accentuating shifts in the phase of $E_1 + E_2$ relative to E_3, while minimizing amplitude change in these voltages. Thus there is some particular value of R_2' for which amplitude modulation of the incoming signal will not affect the output of the ratio detector.

Practical Ratio-detector Circuits. Many variations in the circuit details of the ratio detector are possible. While the circuit of Fig. 17-17*a* is the arrangement best adapted to explain the principles involved, more practical forms are usually employed. An example is illustrated in Fig. 17-17*b*. Here, moving the ground to the center of R_2 makes it possible to omit the radio-frequency choke. Since this places the radio-frequency by-pass capacitor C effectively in shunt with the output terminals, it is now necessary, rather than being merely permissible, for this condenser to have a high impedance to modulation frequencies. A further simplification is made possible by omitting the two resistances R_1 of Fig. 17-17*a*, which were never really needed anyway, since R_2 provides a means by which the charge on C_1 can leak off in the circuits of both (*a*) and (*b*).

PROBLEMS AND EXERCISES

17-1. In Eq. (17-8) explain how the mathematics shows that the time required for an oscillation to go through one cycle is greatest when $m_f \sin \omega_m t$ is zero going negative.

17-2. Complete the detailed steps whereby Eq. (17-10) is obtained from Eqs. (17-11) and (17-12).

17-3. *a.* The following mathematical relation can be shown to be true for all values of x:

$$J_0^2(x) + 2 \sum_{n=1}^{n=\infty} J_n^2(x) = 1$$

Demonstrate that this relation proves that the energy contained in a sinusoidally modulated frequency-modulated wave is constant, irrespective of the frequency deviation or modulating frequency.

b. Explain how this relation shows that the sideband energy in a frequency-modulated wave is exactly the difference between the carrier energy of the unmodulated wave and the carrier energy of the modulated wave.

c. Verify the above equation by numerical values taken from Fig. 17-3 for a modulation index of 2.

17-4. A carrier wave is frequency modulated at 3000 cycles. What is the lowest value of frequency deviation for which all of the energy of the wave will be in the sidebands?

17-5. A carrier wave, having a crest amplitude of 10 volts and a frequency of 60 Mc, is modulated at 5000 cycles with a frequency deviation of 15 kc. From this information determine the amplitude of the carrier, and of the first-, second-, third-, fourth-, and fifth-order sideband components.

17-6. *a.* A frequency-modulated wave having a frequency deviation of [illegible]