SECTION 12

MODULATORS, DEMODULATORS, AND CONVERTERS

Understanding modulation is important to understanding AM and FM broadcasting and receiving. The advantages of both are presented and methods of improving both are discussed. The key features of modulation and demodulation are discussed in this section.

Pulse modulation and demodulation is widely used in communications and control. Various techniques are discussed and compared.

Although optical radiation is just another form of electromagnetic radiation, the transmitters and receivers are considerably different from those used at radio frequencies. The use of these devices has increased rapidly necessitating a better understanding of the devices themselves.

We conclude the section with a discussion of frequency converters and detectors. Characteristics that lead to effective design are presented and discussed. C.A.

In This Section:

Section Bibliography:

Bach Andersen, J. (ed.), "Modern Radio Science 1990," International Union of Radio Science, Oxford University Press, 1990.

Bedrosian, E., "Amplitude and Phase Demodulation of Filtered AM/PM Signals," Rand Corporation, 1986.

Carden, F., "Telemetry Systems Design," Artech House, 1995.

Dixon, R. C., "Spread Spectrum Systems," Wiley-Interscience, 1976.

Dixon, R. C. (ed.), "Spread Spectrum Techniques," IEEE Press, 1976.

Engberg, J., and T. Larse, "Noise Theory of Linear and Nonlinear Circuits," Wiley, 1995.

Freeman, M. H., "Optics," 10th ed., Butterworth-Heinemann, 1995.

Gibson, J. D., and J. L. Melsa, "Introduction to Nonparametric Detection with Applications," IEEE Press, 1996.

Graf, R. F., "The Modern Converter and Filter Circuit Encyclopedia," TAB Books, 1993.

Green, D. C., "Radio Systems Technology," Longman Scientific & Technical, 1990.

Isermann, R., K. Lachman, and D. Matko, "Adaptive Control Systems," Prentice Hall, 1992.

Kahn, D., "Cryptology and the Origins of Spread Spectrum," *IEEE Spectrum*, Vol. 21, No. 9, pp. 70–80, September 1984.

Kaminow, I. P., and E. H. Turner, "Electo-Optic Light Modulators," *Proc IEEE*, Vol. 54, p. 1374, October 1966.

Mass, S. A., "Microwave Mixers," 2nd ed., Artech House, 1993.

Patterson, E. W., et al., "Frequency-Hopped Waveform Synthesis by Using S. A. W. Chirp Filters," *Electro. Lett.*, Vol. 13, No. 21, pp. 633–635, October 13, 1977.

Peterson, R. L., R. E. Ziemer, and D. E. Borth, "Introduction to Spread-Spectrum Communications," Prentice Hall, 1995.

Price, R., "Further Notes and Anecdotes on Spread-spectrum Origins," *IEEE Trans. Comm.*, Vol. Com 31, No. 1, pp. 85–97, January 1983.

Price, R., and P. E. Green, Jr., "A Communication Technique for Multipath Channels," *Proc. IRE*, Vol. 29, pp. 555–570, March 1953.

Purser, M., "Introduction to Error-Correcting Codes," Artech House, 1995.

Schoenbeck, R. J., "Electronic Communications: Modulation and Transmission," 2nd ed., Merrill, Maxwell Macmillan, Maxwell Macmillan International, 1992.

Scholtz, R. A., "The Origins of Spread-spectrum Communications," *IEEE Trans. Comm.*, Vol. Com 30, No. 5, pp. 822–854, May 1982.

Scholtz, R. A., "Notes on Spread-spectrum History," *IEEE Trans. Comm.*, Vol. Com 31, No. 1, pp. 82–84, January 1983.

Schwartz, M., "Information Transmission, Modulation, and Noise," 4th ed., McGraw-Hill, 1990.

Simon, M. K., et al., "Spread Spectrum Communications Handbook," Rev. ed., McGraw-Hill, 1994.

Titsworth, R. C., "Correlation Properties of Cyclic Sequence," *Calif. Inst. Technol. Jet Propulsion Lab. Tech. Rep.* 32–338, July 1963.

Van Trees, H. L., "Detection, Estimation, and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise," Kreiger Pub. Co., 1992.

Walker, B. H., "Optical Engineering Fundamentals," McGraw-Hill, 1994.

Webb, W., "Modern Quadrature Amplitude Modulation: Principles and Applications for Fixed and Wireless Communications," Pentech Press, London; IEEE Press, 1994.

Zlemer, R. E., and W. H. Tranter, "Principles of Communications: Systems, Modulation, and Noise," 4th ed., Houghton Mifflin, 1995.

CHAPTER 12.1
AMPLITUDE MODULATION/ DEMODULATION

Joseph P. Hesler

INTRODUCTION

Frequency translation is fundamental to radio communications in that it produces signal energy, in proportion to variations of an information source, at frequencies that have desirable transmission characteristics, such as antenna size, freedom of interference from similar information sources, line-of-sight to long-range propagation, and freedom of interference from particular noise sources. Frequency translation permits the efficient utilization of open and closed propagation media by many simultaneous users and/or signals.

One of the most used forms of frequency translation is linear modulation, the most common of which is amplitude modulation. In general, amplitude modulation consists in varying the magnitude of a carrier signal in direct correspondence to the instantaneous fluctuations of a modulating signal source, as illustrated in Fig. 12.1.1.

Variations of the basic amplitude-modulation process have been developed to achieve more efficient spectrum utilization and to reduce transmitter power requirements. These include suppressed-carrier systems such as vestigial-sideband, single-sideband, and double-sideband modulation systems. The companion form of frequency translation used in amplitude-modulation systems in *detection*. This is the process whereby the originally translated information is recovered as a baseband signal. Linear amplitude modulation has been the most widely used form of frequency translation in general communications for three reasons: relative ease of implementation, efficient utilization of bandwidth, and availability of devices to implement a simple detection procedure.

It can be argued that amplitude modulation includes such modulation methods as pulse-code keying, pulse-amplitude modulation, frequency-shift keying (the sequential keying of multiple carrier signals), and variations of them, such as pulse-position modulation and pulse-width modulation. This subsection, however, is restricted to the amplitude modulation by signal sources whose outputs are continuous time functions. The special types of modulation mentioned above are discussed in Chap. 12.3.

The general expression for the output of a linear amplitude modulator with a sinusoidal modulation input is

$$E = E_0(1 + m \sin \omega_m t) \sin (\omega_c t - \phi) \tag{1}$$

where E_0 = peak amplitude of carrier signal
ω_m = modulating signal frequency (rad/s)
ω_c = carrier frequency (rad/s)
m = modulation index
ϕ = arbitrary carrier phase angle (rad)
t = time (s)

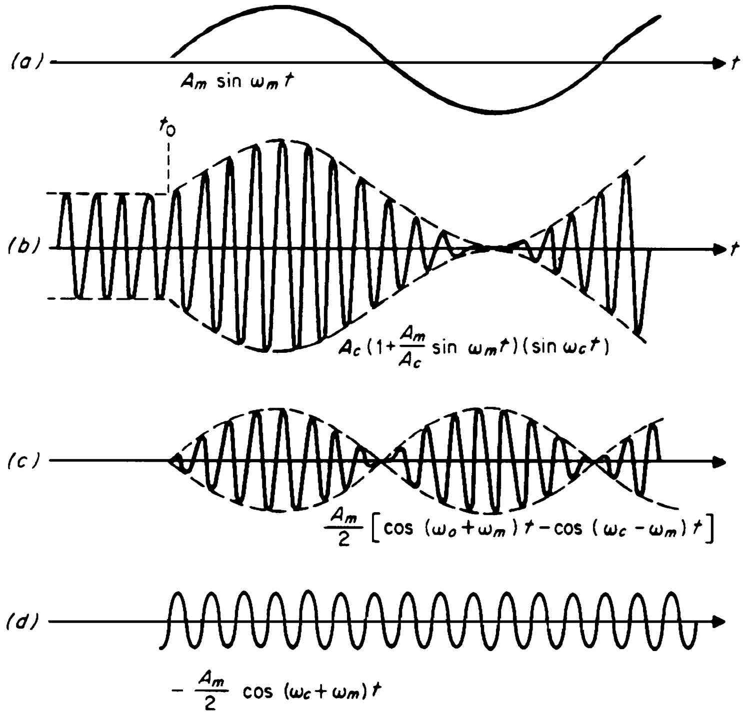

FIGURE 12.1.1 Amplitude modulation: (*a*) modulating signal; (*b*) double-sideband amplitude-modulated signal; (*c*) double-sideband suppressed-carrier amplitude-modulated signal; (*d*) single-sideband suppressed-carrier amplitude-modulated signal.

Expansion of Eq. (1) provides

$$E = E_0 \sin(\omega_c t + \phi) + \frac{mE_0}{2}\cos[(\omega_c - \omega_m)t + \phi] - \frac{mE_0}{2}\cos[(\omega_c + \omega_m)t + \phi] \tag{2}$$

Note that the carrier signal is reproduced exactly as if it carried no modulation. The carrier in itself does not carry any information. The second and third terms in Eq. (2) represented sideband signals produced in the modulation process. These signals are displaced from the carrier signal in the frequency spectrum, on each side of the carrier, by a frequency difference equal to the modulating-signal frequency. The magnitudes of the sideband signals are equal and are proportional to the modulating index m.

Both positive and negative amplitude modulation can be produced in an unsymmetrical manner. For each case the amplitude-modulation index m is defined as

$$m = \begin{cases} \dfrac{E_{max} - E_0}{E_0} & \text{positive modulation} \\[2ex] \dfrac{E_0 - E_{min}}{E_0} & \text{negative modulation} \end{cases} \tag{3}$$

where E_{max} is the peak amplitude of modulated carrier and E_0 is the peak amplitude of unmodulated carrier.

The maximum negative-modulation index of unity results from the reduction of the instantaneous amplitude modulation that can be produced corresponds to a modulation index of unity.

For a complex modulating signal $G(t)$ the modulated carrier spectrum is

$$\begin{aligned} F(\omega) &= \mathfrak{F}\{E_0[1+mG(t)] \sin(\omega_c t+\phi)\} \\ &= \frac{E_0}{2\pi}\int_{-\infty}^{\infty}[1+mG(t)] \sin(\omega_c t+\phi)e^{-j\omega t}dt \end{aligned} \tag{4}$$

where $F(\omega)$ is the Fourier transform of the time function $\varepsilon[f(t)]$.

TYPES OF AMPLITUDE MODULATION

Generation of an amplitude-modulated waveform requires the multiplication of signals, $f_1(t)f_2(t)$. Both signals need not be time-variant, e.g., a carbon microphone modulates a dc potential with voice signals. The main applications, however, concern the modulation of *carrier* signals to exploit desirable transmission characteristics, that is, $f_1(t)[A \sin(\omega_c t+\phi)]$. Two classes of circuits are used as modulators, *square-law* devices and *linear* modulators.

Square-Law Modulation

Any device having a nonlinear transfer function can be expressed in a power series form, e.g., the current in a diode, $i(e) = a_0 + a_1 e + a_2 e^2 + a_3 e^3 + \cdots$.

When the diode characteristics and bias conditions are chosen so as to enhance the coefficient a_2 with respect to the other coefficients, the device is considered to be a square-law nonlinear element.

Under these conditions, when two signal inputs $f_1(t)$ and $f_2(t)$ are summed and used as the driving function e, a significant portion of the output will exist as a_2e^2

$$e = f_1(t) + f_2(t) \tag{5}$$

$$e^2 = 2f_1(t)f_2(t) + [f_1(t)]^2 + [f_2(t)]^2 \tag{6}$$

Suitable nonlinear characteristics are exhibited by various types of rectifiers, triodes, and transistors.

Linear Modulation

Linear modulators are devices with transfer functions that are linearly related to a control parameter. Examples include the outputs of class C rf amplifiers as a function of the B-plus supply voltage, the transconductance gain of transistor differential amplifiers versus emitter current-source magnitude, and Hall effect devices whose transconductance is proportional to the applied magnetic field.

METHODS OF AMPLITUDE MODULATION

Square-Law Amplitude Modulators

A square-law modulator requires three features: a method of summing the two input signals $f_1(t)$ and $f_2(t)$, a device with a nonlinear transfer function, and a tuned circuit and coupling network for extracting the desired modulation products. Voltage summing or current summing are used depending on the transfer characteristic of interest. The nonlinear device is biased in a region that enhances the second-order coefficient of the power series that represents the nonlinear transfer function. The most common devices used for this type of modulation are semiconductor diodes and vacuum-tube triodes. An example of a typical circuit is shown in Fig. 12.1.2.

The efficiency of this type of amplitude modulation is generally low, and all the output energy is supplied by the driving functions.

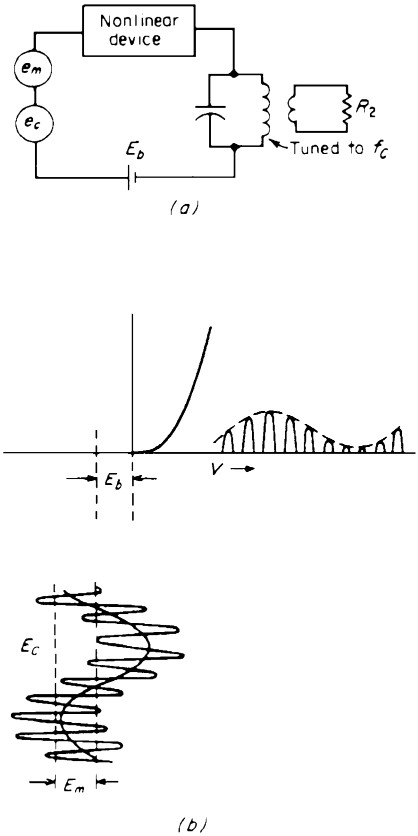

FIGURE 12.1.2 Square-law modulator.

Consider two input signals:

$$e_m = E_m \cos \omega_m t \quad \text{modulating signal} \tag{7}$$

$$e_c = E_c \cos \omega_c t \quad \text{carrier signal} \tag{8}$$

The input applied to the nonlinear device is

$$e_s = E_m \cos \omega_m t + E_c \cos \omega_c t \tag{9}$$

If the transfer function is represented by the two terms of interest from a Taylor series,

$$e_0 = a_1 e_s + a_2 e_s^2 \tag{10}$$

The output components resulting are

$$\begin{array}{ll} e_0 = (a_2/2)(E_m^2 - e_c^2) & \text{dc rectified component} \\ -a_1 E_m \cos \omega_m t & \text{modulating signal} \\ -a_1 E_c \cos \omega_c t & \text{carrier} \\ -(a_2/2)E_m^2 \cos^2 2\omega_m t & \text{second harmonic of modulation} \\ -(a_2/2)E_c^2 \cos^2 2\omega_c t & \text{second harmonic of carrier} \\ -a_2 E_c E_m \cos (\omega_c - \omega_m)t & \text{lower sideband} \\ -a_2 E_c E_m \cos (\omega_c + \omega_m)t & \text{upper sideband} \end{array} \tag{11}$$

There would be other terms, also, from the higher-order coefficients of the Taylor series. The degree of modulation is expressed as

$$\text{Modulation index} = 2(a_2/a_1)E_m \tag{12}$$

The desired outputs for double-sideband amplitude modulation are

$$e_0' = a_1 E_c \cos \omega_c t + a_2 E_c E_m \cos (\omega_c \pm \omega_m)t \tag{13}$$

The square-law devices are reciprocal in that the modulating frequency will appear as an output if a modulated signal is applied as the input. Thus the square-law device can also be used as a demodulator or detector.

LOW- AND MEDIUM-POWER LINEAR MODULATORS

Many applications exist in mobile equipment for amplitude modulators with output powers from milliwatts to tens of watts. Transistor circuits are used almost exclusively for these circuits. Carrier frequencies above 1 GHz can be used in the lower-power transistor circuits. The most common methods of amplitude modulation used are class C collector-modulated stages with the rf applied in the common-emitter or common-base configuration, as shown in Fig. 12.1.3. The common-emitter configuration provides the maximum power gain and excellent efficiency. Common-base stages are used to increase the upward modulation capabilities, where the maximum modulation indices are important. Increased linearity can be achieved at the expense of efficiency by biasing the amplifier class B so that a nominal collector current flows under no-modulation conditions.

For class C operation the transistor transfer characteristics of the modulated amplifier are determined from the large-signal input and output parallel equivalent-impedance data. These are determined experimentally or provided on device specification sheets. The transistors are operated in a very nonlinear manner as class C amplifiers; therefore the experimental data should be representative of the expected operating point, because the small-signal transistor parameters are not adequate.

POWER RELATIONSHIPS

The instantaneous rms output voltage E varies about the unmodulated carrier rms level E_0. The maximum rms output is $E_0(1+m)$, where m is the modulation index that can vary from zero to unity for symmetrical modulation. The minimum rms output is $E_0(1-m)$. The unmodulated power into the load R is

$$P_o = E_0^2/R \tag{14}$$

The peak power into the load is

$$P_{\max} = [E_0(1+m)]^2/R \tag{15}$$

The minimum is

$$P_{\min} = [E_0(1-m)]^2/R \tag{16}$$

The average power into the load for sinusoidal modulation is

$$P_{\mathrm{av}} = P_o(1+m^2/2) = (E_0^2/R)(1+m^2/2) \tag{17}$$

The unmodulated output power E_0^2/R is supplied by the class C amplifier, and the sideband energy $(E_0^2/R)(m^2/2)$ is supplied by the modulator. The class C amplifier can be biased very close to the peak modulated-output envelope swing, $V_{CC} \approx \sqrt{2}mE_0$. For 100 percent modulation, $V_{CC} \approx \sqrt{2}E_0$. The dissipation in the output voltage stage is the difference between the total input power, consisting of the dc collector bias and the input rf drive, and the output rf power. The input rf drive for an amplifier with power gain A is

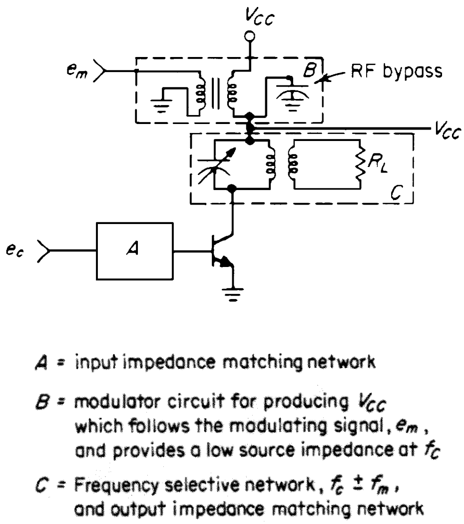

FIGURE 12.1.3 Collector-modulated transistor (class C) rf amplifier.

$$P_{\mathrm{drive}} = E_0^2/RA \tag{18}$$

The input dc bias, unmodulated, is slightly greater than

$$P_{\mathrm{dc}} = \sqrt{2}E_0\bar{I}_c \tag{19}$$

where $\bar{I}_c$ is the average dc collector current.

The output transistor dissipation unmodulated is

$$P_{TR_O} \approx \sqrt{2}E_0\bar{I}_c + P_{\mathrm{drive}} - E_0^2/R \tag{20}$$

$$P_{TR_O} \approx \sqrt{2}E_0\bar{I}_c - \frac{E_0^2}{R}\frac{A-1}{A} \tag{21}$$

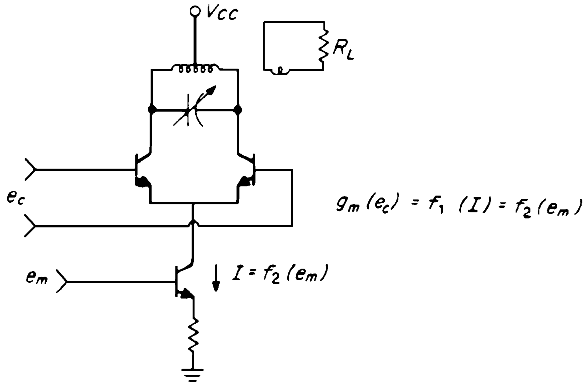

FIGURE 12.1.4 Differential amplifier amplitude modulator (without dc bias details).

In higher-power systems the output transistors can be paralleled. At higher frequencies the gain of the output transistors may be such that the output power is limited by the dissipation in the driver. In these instances the driver may also be collector-modulated to achieve adequate upward modulation and to reduce the power dissipation in the driver.

The input and output impedances of the transistor class C amplifiers are characteristically low. Typical parallel equivalent input impedances are

$$2 < R_{\text{in}} < 50\ \Omega \qquad 30 < C_{\text{in}} < 5000\ \text{pF}$$

The collector load impedance-resistance component is determined from the bias voltage and output power

$$R'_L = (V_{CC})^2/2P_o \tag{22}$$

where V_{CC} is the dc collector bias voltage and the output voltage swing $2V_{CC}$ peak to peak and P_o is the unmodulated output power.

The collector parallel output capacitance depends on the device geometry and may range from a few picofarads for very-high-frequency lower-power devices to a few hundred picofarads for large-geometry high-power devices.

Interstage and output networks are used to obtain conjugate matches for maximum power gain.

Other types of linear transistor modulators can be used where efficiency is less critical or where very-wide-band operation precludes effective output filtering for the elimination of harmonics. A differential amplifier circuit, as shown in Fig. 12.1.4, will have a transconductance gain that is very nearly proportional to the emitter current-source magnitude. Modulation of the current source will produce a nearly ideal multiplication of the rf input signal and the modulation signal for a wide range of low current levels. The differential amplifier must be biased class A with minimum dc offset at the base inputs. A single transistor multiplier can also be used, as shown in Fig. 12.1.5 for very-low-level outputs. An emitter bypass capacitor for the rf signal is used in place of the additional transistor for the rf return.

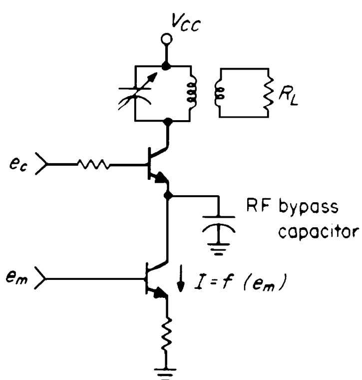

FIGURE 12.1.5 Two-transistor transconductance modulator.

HIGH-POWER LINEAR MODULATORS

High-power linear modulators, 50 W and up, are generally constructed using class C plate-modulation vacuum-circuits. Some of the intermediate power and frequency applications use paralleled transistor configurations with class C collector modulation.

Triodes, tetrodes, and pentodes are used in the vacuum-tube circuits. Multigrid tubes require screen-grid modulation in conjunction with the control-grid modulation to achieve space-charge modulation and to minimize screen current and screen dissipation. The two methods of screen modulation commonly used are: (1) self-bias of the screen grid with a bypassed dropping resistor or inductor from the screen to the plate supply or the screen-grid supply and (2) screen modulation via a separate winding on the modulation transformer.

GRID MODULATION

The amplitude of a class C rf amplifier output can also be modulated by changing the grid bias with the modulating signal. The modulating signal is added to the rf input signal. The effect is to change the magnitude of the plate-current pulses, and hence the fundamental component of the plate current.

The disadvantage associated with grid modulation is that the fixed-plate supply voltage must be twice the peak rf voltage without modulation. This causes high plate dissipation and lowers the plate efficiency to the range of 35 to 45 percent when unmodulated.

The carrier power obtained from a plate-modulated class C amplifier is about three times that available from a grid-modulated circuit using the same tube.

CATHODE MODULATION OF CLASS C RF AMPLIFIERS

Cathode modulation can be used with a class C rf amplifier. The modulation transformer output varies the grid-cathode as well as the plate-cathode voltages. The ratio of grid and plate modulation can be selected by varying the tap: thus the circuit provides a means of producing varying combinations of grid and plate modulation. Some grid-leak bias is normally used to improve linearity.

MODIFIED AM METHODS

The information transmitted by an amplitude-modulated carrier is contained wholly in the modulation sidebands. The transmission of the carrier energy simplifies the receiver-detector implementation but adds no information. In addition, each sideband contains the same information, and only one is required to transmit the intelligence. Elimination of the carrier and/or one sideband can affect a substantial transmitter power saving. For 100 percent amplitude modulation the carrier power is two-thirds of the transmitter power and each sideband one-sixth. Elimination of the carrier will result in *double-sideband suppressed-carrier modulation*. Elimination of one sideband while retaining the carrier or a substantial portion of the carrier results in *vestigial-sideband transmission*. Elimination of the carrier and one sideband is called *single-sideband suppressed-carrier modulation*.

The easiest of these modulation schemes to implement is the vestigial-sideband transmission, both from a transmitter and a receiver viewpoint. The unwanted sideband is generally filtered out at low levels in the transmitter chain and is known as *transmitter attenuation* (TA). One sideband can also be eliminated in the receiver by selective filtering and is called *receiver attenuation* (RA). The latter scheme is not practical from a spectrum-utilization sense; hence it is normally used only as in television broadcast to complete unwanted sideband rejection performed primarily at the transmitter.

A vector notation can be used to illustrate the phenomena of various types of amplitude modulation. The vector is a complex function. The sinusoidal function of time that is of interest is the real part or the vector projection on the real axis of the complex plane. Thus the real part of $Ae^{j\omega t}$ is $A \cos \omega t$, as shown in Fig. 12.1.6*a*.

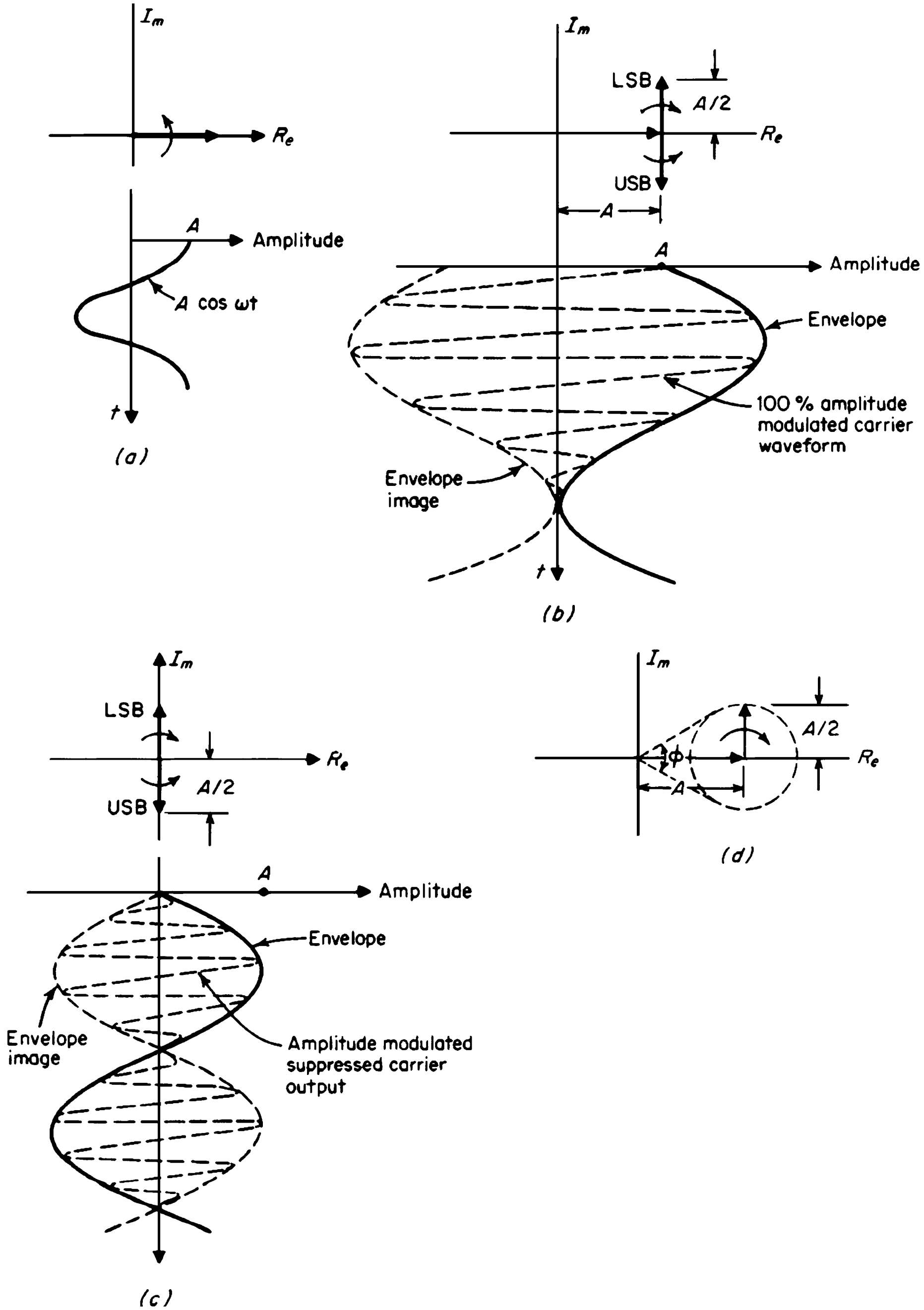

FIGURE 12.1.6 Vector representation of amplitude modulation, (*a*) Rotating carrier vector and its real projection. (*b*) Vector representation with sideband vectors resulting from 100 percent amplitude modulation. Envelope function is locus of projection on the real axis with a nonrotating carrier vector. (*c*) Double-sideband suppressed-carrier vector representation. Envelope function is locus of projection on the real axis with coordinate rotation as in (*b*). (*d*) Vector representation of a single-sideband amplitude-modulated waveform showing peak-to-peak carrier phase modulation ϕ resulting from the absence of the other sideband signal.

AMPLITUDE MODULATION/DEMODULATION

The projection on the real axis of the vector $Ae^{j\omega t}$ can be considered as the carrier signal in subsequent amplitude-modulation (AM) discussions. Since the unmodulated carrier signal in amplitude-modulation processes is a fixed peak amplitude and fixed frequency function of time, a modified vector representation can be used to describe the envelope of the modulated waveform. The modified vector diagram maintains the carrier vector as a fixed, nonrotating vector. By this means subsequent illustrations are referenced to a complex plane rotating at the carrier angular rate, and the projection on the real axis corresponds to the *modulation envelope variations with time*.

For an amplitude-modulation system using a sinusoidal modulating function at 100 percent modulation, $m = 1$, the addition of two vectors to the basic vector diagram is required. The two additional vectors represent the sideband signals produced in the modulation process. The two sideband signals are displaced on either side of the carrier signal in the modulated signal spectrum by a frequency equal to the modulating frequency. Therefore, with respect to the modified reference system, the lower sideband vector will rotate clockwise at the modulation-signal angular rate, and the upper sideband will rotate counterclockwise at the same rate.

The relative phases of the two sideband vectors are such that no change in carrier phase angle is introduced when they are both present, i.e., the imaginary parts of the sideband contributions cancel. The initial angles are set by the phase of the modulating function. As the sideband vectors rotate in the modified complex plane, the vector sum of the carrier plus sidebands describes sinusoidal projection on the real axis at the modulating frequency rate.

Note that the unmodulated carrier projection is constant and equal to the peak carrier magnitude. The 100 percent modulated carrier projection is nonnegative. If the original complex plane were used, the entire vector system would rotate at the carrier angular rate, and the projection on the real axis would be the actual time function produced by the modulator. This function would have negative portions and would fill the envelope function and its image with amplitude-modified sinusoids at the carrier frequency, as shown in Fig. 12.1.6*b* with the dashed lines within the envelope.

VESTIGIAL-SIDEBAND SYSTEMS

Vestigial-sideband transmission introduces angle modulation on the carrier because the symmetry of the contrarotating sideband vectors is lost, as illustrated in Fig. 12.1.6*d*. A standard envelope detector can still be used to recover the modulating signal, however. The primary objective in the application of vestigial-sideband transmission is to conserve spectrum in the transmission medium.

SUPPRESSED-CARRIER SYSTEMS

Suppressed-carrier systems for AM transmission and reception require modifications to the receiver. The double-sideband suppressed-carrier signal cannot be envelope-detected without the reinsertion of a carrier signal. The frequency and the phase of the reinserted carrier are critical. This type of transmission is used with special phase-locked receivers or with the transmission of low-level carrier to permit the reconstitution of the carrier frequency and phase at the receiver. A double-sideband suppressed-carrier signal can be generated with a balanced modulator, as shown in Fig. 12.1.7*a*.

The nonlinear devices can be diodes or modulated class C rf amplifiers. The balanced modulator simplifies the tuned-output-circuit design because all even harmonics of the modulating process tend to be canceled along with the carrier.

Single-sideband suppressed-carrier modulation simplifies the receiver design in some applications. The phase of the reinserted carrier at the receiver is not critical as in double-sideband suppressed-carrier systems. Also, a frequency error in the reconstituted carrier will result only in a corresponding shift in the demodulated signal frequencies, which may be tolerable. If accurate modulation-frequency preservation is required, a low-level carrier can be transmitted to aid in the reconstruction of the carrier signal at the receiver. The reinserted carrier amplitude with respect to the received sideband signal is generally made large to minimize angle modulation of the carrier at the detector.

Two methods are usually employed to generate suppressed-carrier single-sideband transmissions. The most direct method is to filter out the undesired sideband and carrier at low levels in the transmitter chain.

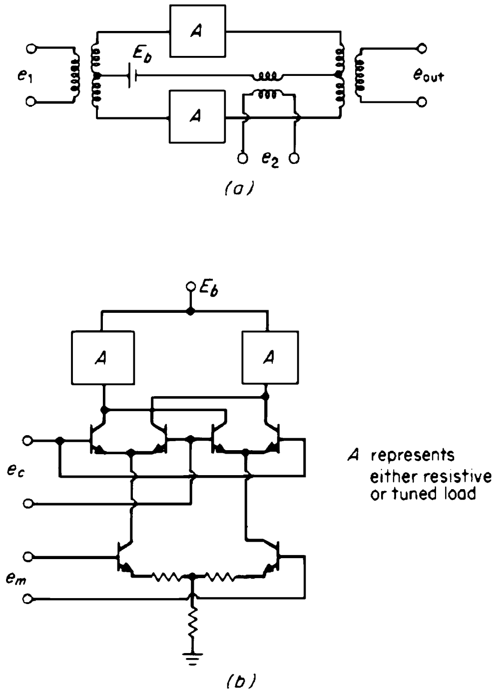

FIGURE 12.1.7 Double-sideband suppressed-carrier modulators: (*a*) general form of a balanced modulator for suppressed-carrier output using nonlinear elements *A* and bias voltage E_m, (*b*) four-quadrant multiplier.

This filtering problem is difficult when the modulation sidebands of interest are close to the carrier and the carrier frequency is high.

An alternative method of suppressed-carrier single-sideband modulation is by unwanted sideband cancellation. Two balanced modulators are used, with their outputs combined in push-pull.

MODULATED OSCILLATORS

A direct modulated class C oscillator can be used as an AM transmitter. The linearity of such circuits is generally as good as or better than the plate-modulated class C amplifiers. The main disadvantage is the tendency for carrier-frequency pulling that results from the changes in the oscillator operating point as the modulation signal varies.

A very useful circuit for the generation of low-level double-sideband suppressed-carrier signals is the four-quadrant multiplier. The same circuit will also operate as a double-sideband modulator with carrier and as low-level demodulator in phase-locked receiver systems. Multipliers are available in integrated-circuit form, and they can be used at frequencies from dc to beyond 100 MHz (Fig. 12.1.7*b*).

DETECTORS

The most common AM detector or demodulator is a diode rectifier. The ideal diode detector passes current in only one direction and will essentially follow the envelope of an amplitude-modulated waveform when used in a circuit as shown in Fig. 12.1.8.

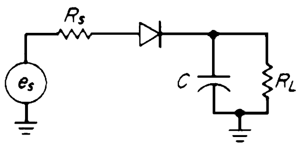

FIGURE 12.1.8 Diode envelope detector.

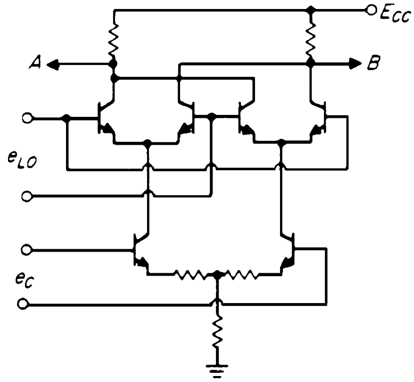

FIGURE 12.1.9 Four-quadrant multiplier used as a product detector. The local-oscillator input e_{LO} must be phase-coherent with the carrier of the modulated input signal e_C. Complementary outputs are obtained at *A* and *B*.

The charging time constant R_sC must be short, so that the capacitor voltage follows the input signal E_{rf} when the diode is forward-biased or conducting. Conversely, the discharge time constant R_LC must be long enough to retain most of the rectified voltage between cycles of the carrier signal but not so large that the capacitor voltage will not discharge at the maximum rate of change of the input signal envelope. The envelope detector is essentially insensitive to residual angle modulation of the carrier, and it is therefore usable in single-sideband receivers.

Practical diode rectifiers have nonlinear resistance characteristics in the conduction bias region and are therefore operated at fairly high input signal levels, on the order of 2 to 10 V peak for semiconductor diodes.

PRODUCT DETECTORS

Another type of amplitude demodulator is the product detector, or multiplier circuit (Fig. 12.1.9). This circuit has distinct advantages and disadvantages. The advantages include the ability to detect lower-level signals with a linear response; the ability to differentiate phase reversals in the modulated waveform, resulting from balanced amplitude modulation with suppressed carrier; and the ability to produce, in some designs, error signals for automatic-frequency-control systems in receivers.

The analytical expression for the output of a product detector is

$$e(t) = E\ [1 + m \underbrace{\sin(\omega_m t + \phi_m)]\ \cos(\omega_c t + \phi_c)}_{\text{amplitude-modulated signal}}\ \underbrace{\cos(\omega_c t + \phi_d)}_{\text{local-oscillator signal}} \tag{23}$$

Expansion of Eq. (23) shows that the product of the incoming carrier signal, $\cos(\omega_c t + \phi_c)$, and the local-oscillator signal, $\cos(\omega_c t + \phi_d)$, produces a dc term except when these two inputs are in quadrature phase. The output dc term is proportional to the cosine of the relative phase angle of the carrier and local-oscillator signals. A four-quadrant multiplier circuit capable of performing this type of detection is shown in Fig. 12.1.9.

The details of the dc bias network are not included in the elementary schematic of the multiplier circuit. The input signals, rf and local-oscillator, are applied to either input port. Balanced-differential or single-ended inputs can be used, although the balanced inputs give the added performance of increased linearity and common-mode signal rejection. Two outputs of opposite polarity are available at the collectors of the upper-rank transistors. Design options are available to increase the efficiency and linearity of the circuit. Normally, an overdrive is applied to the local-oscillator port to make that section of the multiplier operate in a switching mode. The effect is to multiply the rf signal with a square wave at the carrier frequency instead of a sine wave. This type of operation also produces additional outputs at the higher harmonics of the carrier frequency.

For balanced linear in-phase inputs at both ports, the outputs consist of

$$\begin{aligned} e_0 &= {}^1\!/_2(A_{LO} + A_{LO}\cos 2\omega_c t)E[1 + m\sin(\omega_m t + \phi_m)] \\ &= {}^1\!/_2 A_{LO}E[1 + m\sin(\omega_m t + \phi_m)] + {}^1\!/_2 A_{LO}E\cos 2\omega_c t \end{aligned} \tag{24}$$

The disadvantages of the product detector are relative circuit complexity and the need for a phase-coherent local-oscillator signal. The circuit complexity can be circumvented for input signal frequencies up to 100 MHz by using integrated-circuit multipliers. This approach also minimizes the possibility of serious dc offset problems because of device mismatch. The generation of the coherent local-oscillator signal can be achieved in two ways. For simple double-sideband (DSB) amplitude-modulated signals, the carrier can be stripped from the input rf signal, with a parallel limiting amplifier with narrow bandwidth. This approach cannot be used in suppressed-carrier AM systems, where the carrier phase reversals are introduced in the modulation process.

The more general approach is to use an additional multiplier circuit as a phase detector. When the rf and local-oscillator signals are equal in frequency and in phase quadrature, the multiplier output has no dc component. Any relative phase shift from quadrature will produce an odd-function-error signal at dc. This signal can be used with a voltage-controlled oscillator to correct the phase of the local oscillator or input rf signal. The system described is a phase-locked loop. The bandwidth of this loop can be controlled independently from the rf bandwidth for optimum acquisition and noise-suppression characteristics. This type of system is capable of producing a stable and noise-free local-oscillator signal that tracks any variations in the frequency and phase of the input rf signal.

With additional modifications the phase-locked receiver system is capable of reinserting the desired carrier in suppressed-carrier double- and vestigial-sideband systems.