RF and Microwave Wireless Systems. Kai Chang

ISBNs: 0-471-35199-7 (Hardback); 0-471-22432-4 (Electronic)

CHAPTER NINE

Modulation and Demodulation

9.1 INTRODUCTION

Modulation is a technique of imposing information (analog or digital) contained in a lower frequency signal onto a higher frequency signal. The lower frequency is called the modulating signal, the higher frequency signal is called the carrier, and the output signal is called the modulated signal. The benefits of the modulation process are many, such as enabling communication systems to transmit many baseband channels simultaneously at different carrier frequencies without their interfering with each other. One example is that many users can use the same long-distance telephone line simultaneously without creating a jumbled mess or interference. The modulation technique also allows the system to operate at a higher frequency where the antenna is smaller.

Some form of modulation is always needed in an RF system to translate a baseband signal (e.g., audio, video, data) from its original frequency bandwidth to a specified RF frequency spectrum. Some simple modulation can be achieved by direct modulation through the control of the bias to the active device. A more common method is the use of an external modulator at the output of the oscillator or amplifier. Figure 9.1 explains the concept of modulation.

There are many modulation techniques, for example, AM, FM, amplitude shift keying (ASK), frequency shift keying (FSK), phase shift keying (PSK), biphase shift keying (BPSK), quadriphase shift keying (QPSK), 8-phase shift keying (8-PSK), 16-phase shift keying (16-PSK), minimum shift keying (MSK), and quadrature amplitude modulation (QAM). AM and FM are classified as analog modulation techniques, and the others are digital modulation techniques.

After modulation, the signal is amplified and radiated to free space by an antenna. The signal is then picked up by a receiver antenna at some distance away and is then

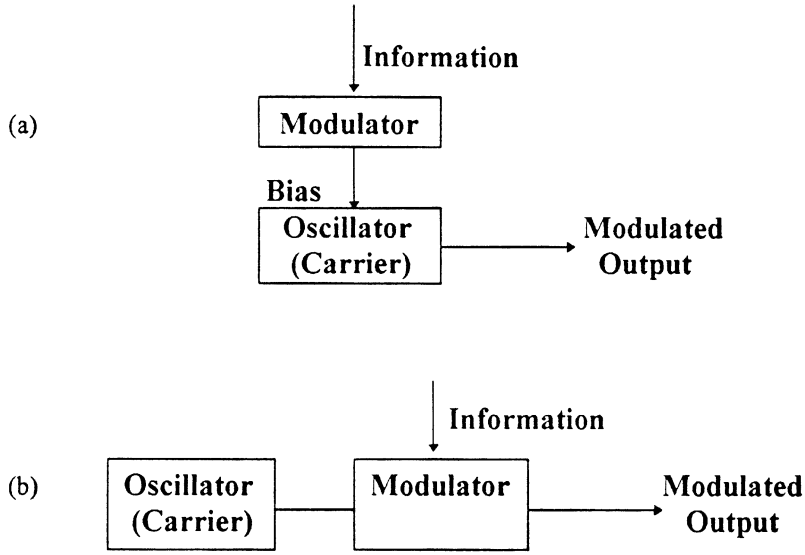

FIGURE 9.1 Different modulation schemes: (*a*) direct modulation; (*b*) external modulation.

amplified, downconverted, and demodulated to recover the original baseband signal (information).

9.2 AMPLITUDE MODULATION AND DEMODULATION

Amplitude and frequency modulation techniques are classified as analog modulation. They are old techniques, having been used for many years since the invention of the radio. Analog modulation uses the baseband signal (modulating signal) to vary one of three variables: amplitude A_c, frequency ω_c, or phase θ. The carrier signal is given by

$$v_c(t) = A_c \sin(\omega_c t + \theta) \tag{9.1}$$

The amplitude variation is AM, the frequency variation is FM, and the phase variation is PM. Phase modulation and FM are very similar processes and can be referred to as angle modulation.

The unique feature of AM is that the message of the modulated carrier has the same shape as the message waveform. Figure 9.2 illustrates the carrier, modulating, and modulated signals.

For simplicity, let a single audio tone be a modulating signal

$$v(t) = A_m \sin \omega_m t \tag{9.2}$$

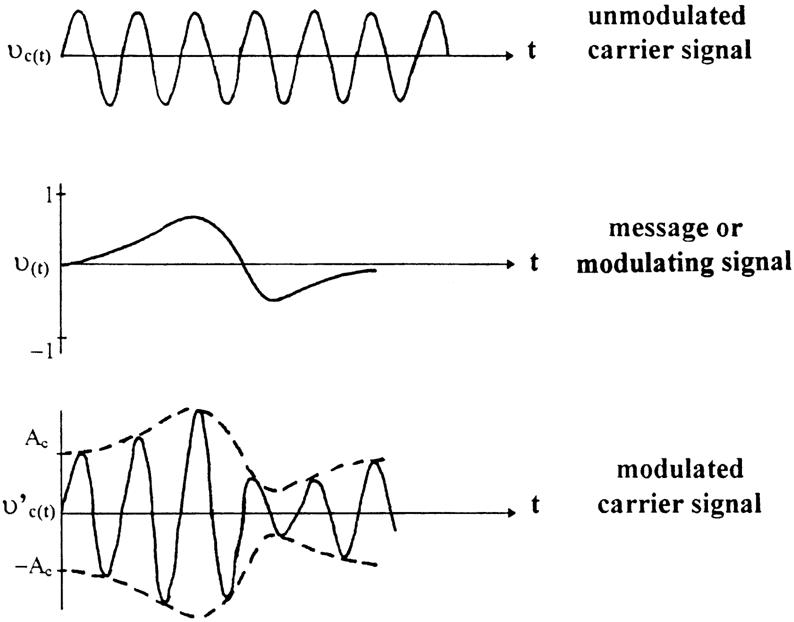

FIGURE 9.2 Signals in AM.

Although a sine wave is assumed, a more complex wave can be considered to be the sum of a set of pure sine waves.

The modulated signal can be written as

$$\begin{aligned} v_c'(t) &= (A_c + A_m \sin \omega_m t) \sin \omega_c t \\ &= A_c\left(1 + \frac{A_m}{A_c} \sin \omega_m t\right) \sin \omega_c t \\ &= A_c(1 + m \sin \omega_m t) \sin \omega_c t \end{aligned} \tag{9.3}$$

where

$$m = \frac{A_m}{A_c} = \frac{\text{peak value of modulating signal}}{\text{peak value of unmodulated carrier signal}}$$

where m is the modulation index, which sometimes is expressed in percentage as the percent of modulation. To preserve information without distortion would require m to be ≤ 1 or less than 100%. Figure 9.3 shows three cases of modulation: undermodulation ($m < 100\%$), 100% modulation, and overmodulation ($m > 100\%$).

Using a trigonometric identity, Eq. (9.3) can be rewritten as

$$v_c'(t) = A_c \sin \omega_c t + \tfrac{1}{2}(mA_c) \cos(\omega_c - \omega_m)t - \tfrac{1}{2}(mA_c) \cos(\omega_c + \omega_m)t \tag{9.4}$$

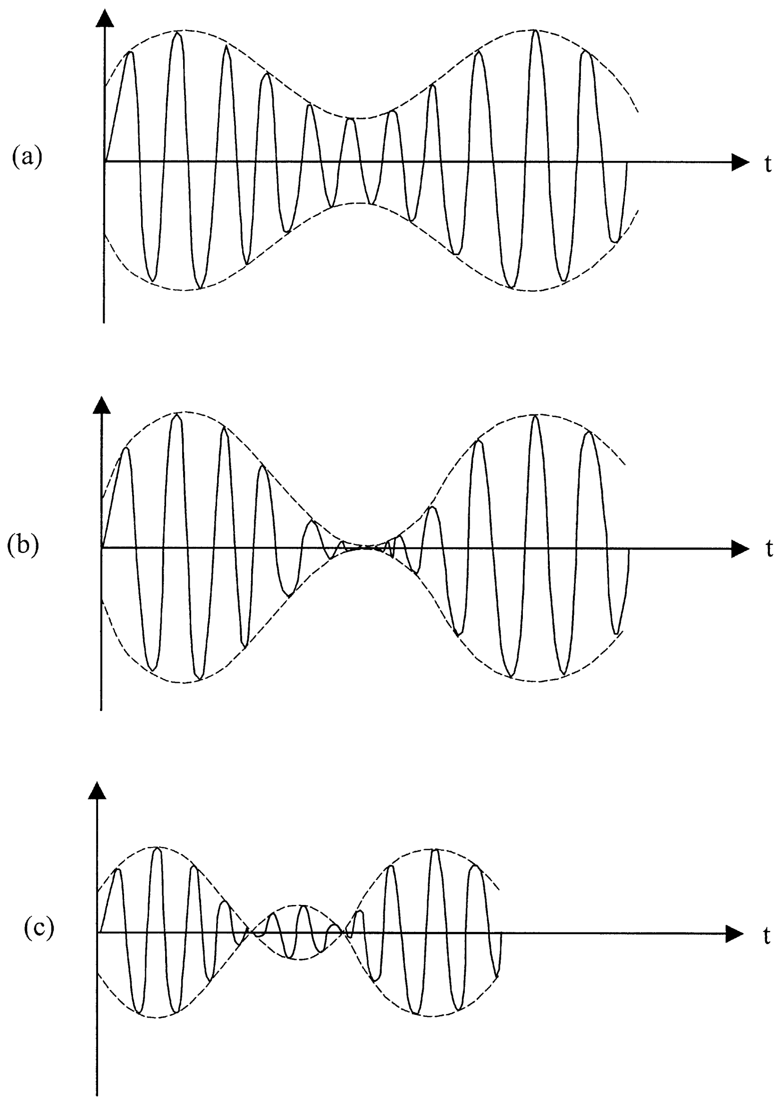

FIGURE 9.3 Degrees of modulatioin: (*a*) undermodulation; (*b*) 100% modulation; (*c*) overmodulation.

The modulated signal contains the carrier signal (ω_c), the upper sideband signal ($\omega_c + \omega_m$), and the lower sideband signal ($\omega_c - \omega_m$). This is quite similar to the output of a mixer.

A nonlinear device can be used to accomplish the amplitude modulation. Figure 9.4 shows examples using a modulated amplifier and a balanced diode modulator.

The demodulation can be achieved by using an envelope detector (described in Chapter 4) as a demodulator to recover the message [1].

Example 9.1 In an AM broadcast system, the total transmitted power is 2000 W. Assuming that the percent of modulation is 100%, calculate the transmitted power at the carrier frequency and at the upper and lower sidebands.

Solution From Equation (9.4)

$$P_T = P_c + \tfrac{1}{4}m^2P_c + \tfrac{1}{4}m^2P_c = 2000\ \text{W}$$

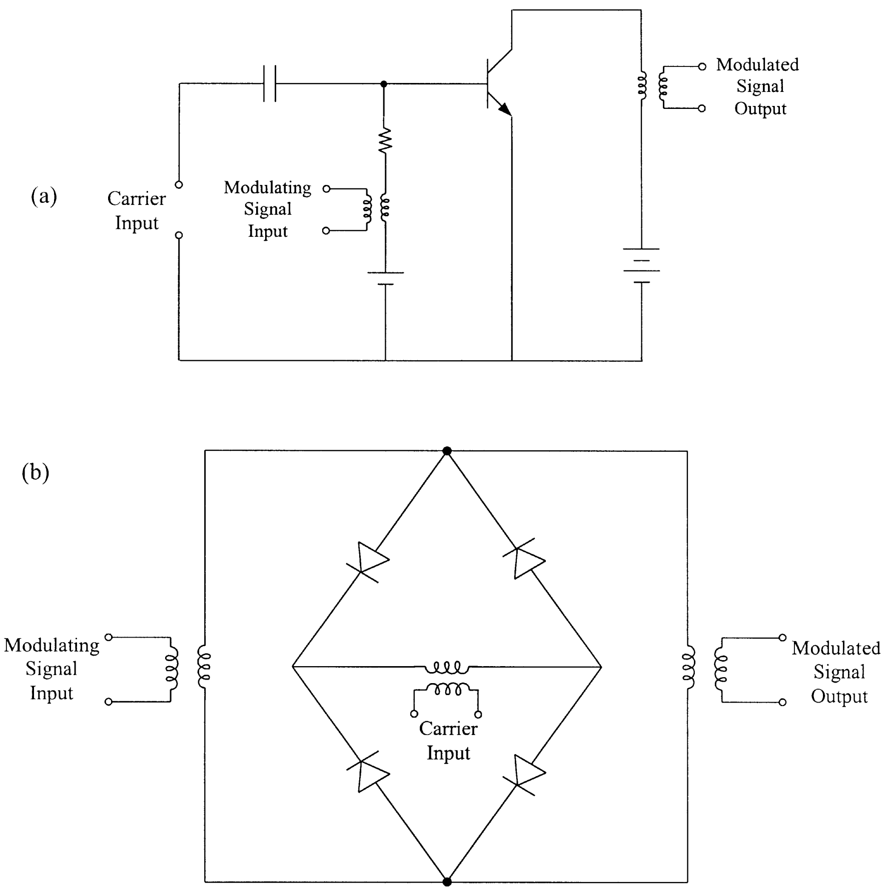

FIGURE 9.4 Amplitude modulation using (*a*) a modulated amplifier and (*b*) a balanced modulator.

Now $m = 1$, we have

$$1.5P_c = 2000 \text{ W} \qquad P_c = 1333.33 \text{ W}$$

$$\text{Power in the upper sideband} = P_{\text{USB}} = \tfrac{1}{4}m^2P_c = \tfrac{1}{4}P_c = 333.33 \text{ W}$$
$$\text{Power in the lower sideband} = P_{\text{LSB}} = \tfrac{1}{4}m^2P_c = 333.33 \text{ W}$$ ■

9.3 FREQUENCY MODULATION

Frequency modulation is accomplished if a sinusoidal carrier, shown in Eq. (9.1), has its instantaneous phase $\omega_c t + \theta$ varied by a modulating signal. There are two possibilities: Either the frequency $\omega_c/2\pi$ or the phase θ can be made to vary in direct proportion to the modulating signal. The difference between FM and PM is not obvious, since a change in frequency must inherently involve a change in phase. In FM, information is placed on the carrier by varying its frequency while its amplitude is fixed.

The carrier signal is given by

$$v_c(t) = A_c \sin \omega_c t \tag{9.5}$$

The modulating signal is described as

$$v(t) = A_m \sin \omega_m t \tag{9.6}$$

The modulated signal can be written as

$$v_c'(t) = A_c \sin[2\pi(f_c + \Delta f \sin 2\pi f_m t)t] \tag{9.7}$$

The maximum frequency swing occurs when $\sin 2\pi f_m t = \pm 1$. Here Δf is the frequency deviation, which is the maximum change in frequency the modulated signal undergoes. The amplitude remains the same. A modulation index is defined as

$$m_f = \frac{\Delta f}{f_m} \tag{9.8}$$

The total variation in frequency from the lowest to the highest is referred to as carrier swing, which is equal to 2 Δf.

In the transmitter, frequency modulation can be achieved by using VCOs. The message or modulating signal will control the VCO output frequencies. In the receiver, the demodulator is used to recover the information. One example is to use a frequency discriminator (frequency detector) that produces an output voltage that is dependent on input frequency. Figure 9.5 shows a block diagram, a circuit schematic, and the voltage–frequency characteristics of a balanced frequency

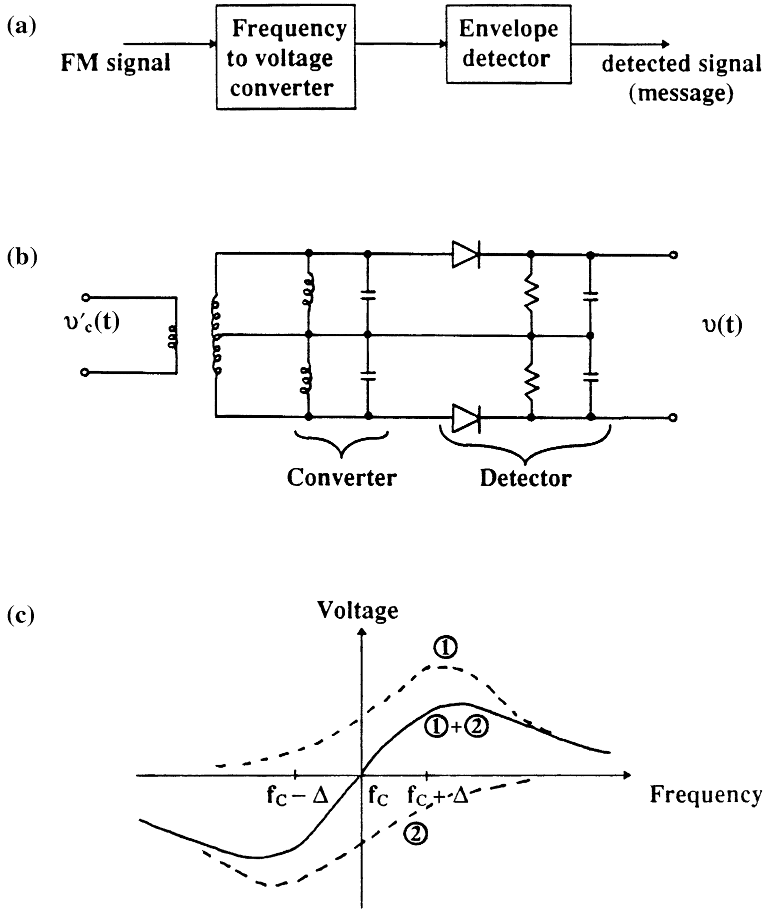

FIGURE 9.5 Balanced frequency discrimination: (*a*) block diagram; (*b*) circuit schematic; (*c*) voltage–frequency characteristics.

discriminator. The circuit consists of a frequency-to-voltage converter and an envelope detector. The balanced frequency-to-voltage converter has two resonant circuits, one tuned above f_c and the other below. Taking the difference of these gives the frequency-to-voltage characteristics of an S-shaped curve. The conversion curve is approximately linear around f_c. Direct current is automatically canceled, bypassing the need for a DC block.

9.4 DIGITAL SHIFT-KEYING MODULATION

Most modern wireless systems use digital modulation techniques. Digital modulation offers many advantages over analog modulation: increased channel capability, greater accuracy in the presence of noise and distortion, and ease of handling. In digital communication systems, bits are transmitted at a rate of kilobits, megabits, or gigabits per second. A certain number of bits represent a symbol or a numerical number. The receiver then estimates which symbol was originally sent from the transmitter. It is largely unimportant if the amplitude or shape of the received signal is distorted as long as the receiver can clearly distinguish one symbol from the other. Each bit is either 1 or 0. The addition of noise and distortion to the signal makes it harder to determine whether it is 1 or 0. If the distortion is under a certain limit, the receiver will make a correct estimate. If the distortion is too large, the receiver may give a wrong estimate. When this happens, a BER is generated. Most wireless systems can tolerate a BER of 10^{-3} (1 in 1000) before the performance is considered unacceptable.

Amplitude shift keying, FSK, BPSK, QPSK, 8-PSK, 16-PSK, MSK, Gaussian MSK (GMSK), and QAM are classified as digital modulation techniques. A brief description of these modulation methods is given below.

In ASK modulation, the amplitude of the transmitted signal is turned "on" and "off," which corresponds to 1 or 0. This can easily be done by bias modulating an oscillator; that is, the oscillator is switched on and off by DC bias. Alternatively, a single-pole, single-throw p–i–n or FET switch can be used as a modulator. Figure 9.6 shows the modulation arrangement for ASK. Demodulation can be obtained by a detector described in Chapter 4 [1, Ch. 6].

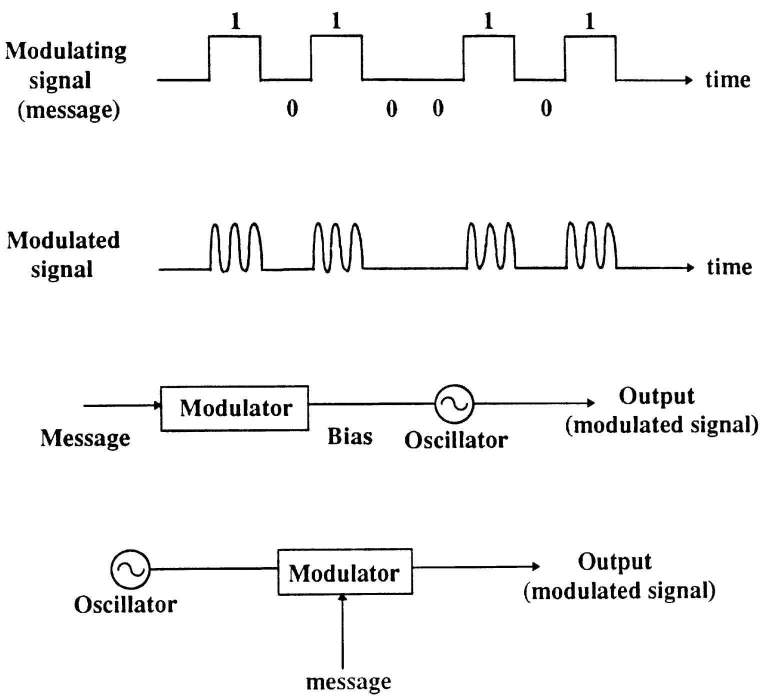

FIGURE 9.6 Amplitude shift keying modulation.

With FSK, when the modulating signal is 1, the transmitter transmits a carrier at frequency f_1; when the modulating signal is 0, the transmitting frequency is f_0. A VCO can be used to generate the transmitting signal modulated by the message. At the receiver, a frequency discriminator is used to distinguish these two frequencies and regenerate the original bit stream.

Minimum shift keying is the binary FSK with two frequencies selected to ensure that there is exactly an 180° phase shift difference between the two frequencies in a 1-bit interval. Therefore, MSK produces a maximum phase difference at the end of the bit interval using a minimum difference in frequencies and maintains good phase continuity at the bit transitions (see Fig. 9.7*a* [2]). Minimum shift keying is attractive

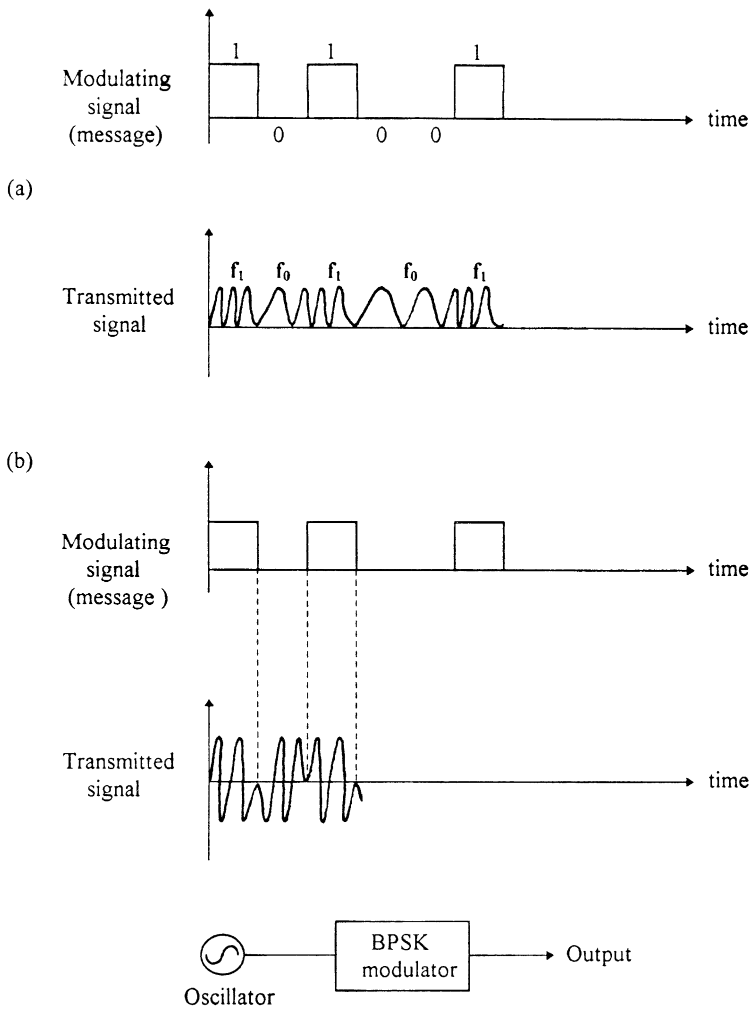

FIGURE 9.7 Modulation techniques: (*a*) MSK; (*b*) BPSK.

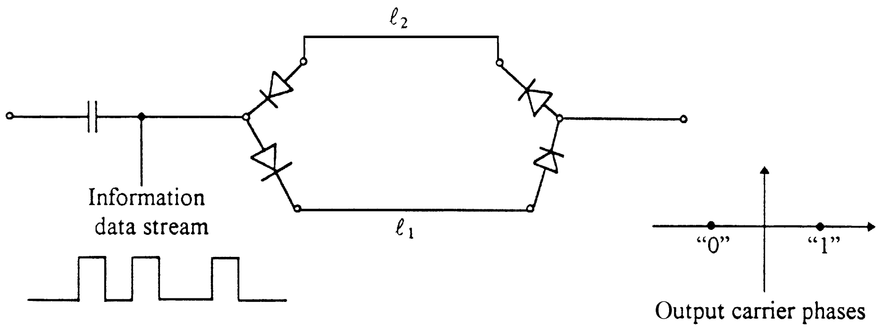

FIGURE 9.8 Biphase switch.

because it has a more compact spectrum and lower out-of-band emission as compared to FSK. Out-of-band emission can cause adjacent channel interference and can be further reduced by using filters. If a Gaussian-shaped filter is used, the modulation technique is called Gaussian MSK (GMSK).

In a PSK system, the carrier phase is switched between various discrete and equispaced values. For a BPSK system, the phase angles chosen are 0° and 180°. Figure 9.7 shows the MSK and BPSK system waveforms for comparison. A switch can be used as a BPSK modulator. Figure 9.8 shows an example circuit. When the data are positive or "1," the signal passes path 1 with a length l_1. When the data are negative or "0," the signal goes through path 2 with a length l_2. If the electrical phase difference for these two paths is set equal to 180°, we have a biphase switch/modulator. This is given by

$$\Delta\phi = \beta(l_1 - l_2) = \frac{2\pi}{\lambda_g}(l_1 - l_2) = 180^\circ \tag{9.9}$$

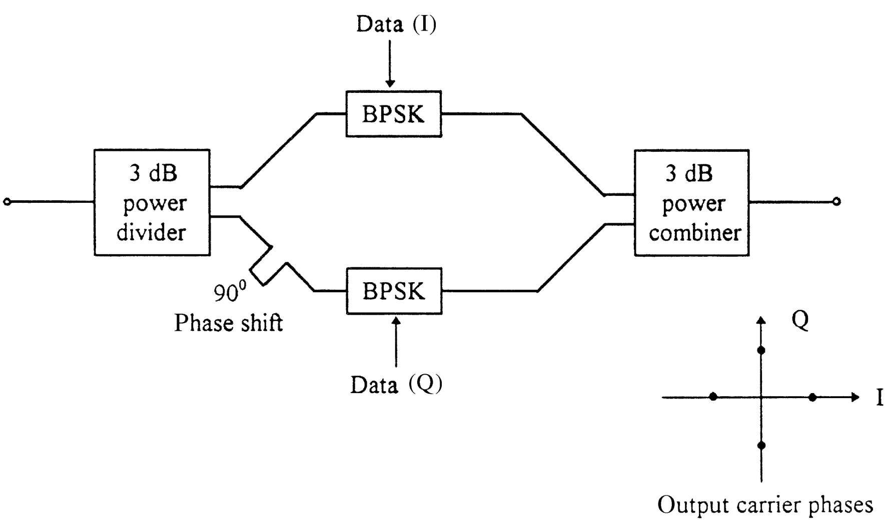

FIGURE 9.9 Quadriphase switch/modulator.

A QPSK modulator consists of two BPSK modulators, connected as shown in Fig. 9.9. A 90° phase shift made of a transmission line is used to introduce the 90° rotation between the outputs of the two BPSK switches. An output phase error of less than 3° and maximum amplitude error of 0.5 dB have been reported at 60 GHz using this circuit arrangement [3]. Quadrature PSK can transmit higher data rates, since two data streams can be transmitted simultaneously. Therefore, the theoretical bandwidth efficiency for QPSK is 2 bits per second per hertz (bps/Hz) instead of 1 bps/Hz for BPSK. Quadrature PSK transmits four (2^2) phases of 0°, 90°, 180°, and 270°. Two data streams can be transmitted simultaneously. The in-phase (I) data stream transmits 0° or 180° depending on whether the data are 1 or 0. The quadrature-phase (Q) data stream transmits 90° and 270°.

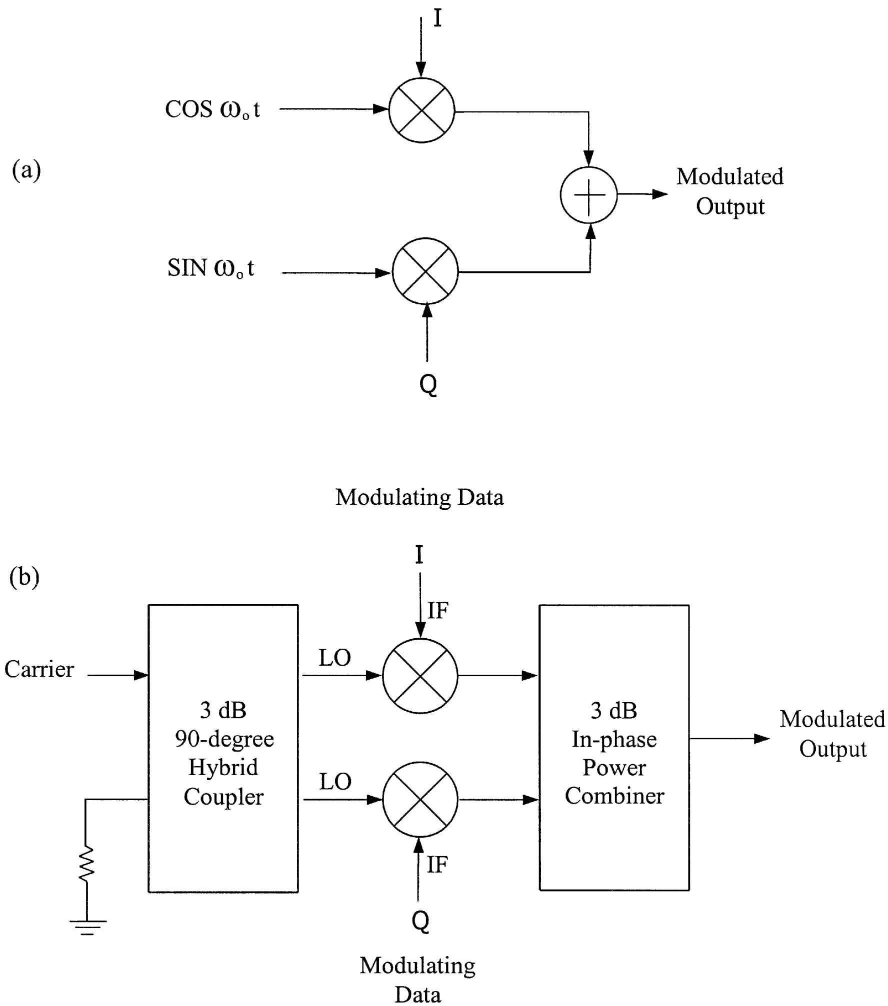

FIGURE 9.10 I/Q modulator: (*a*) simplified block diagram; (*b*) circuit realization.

In-phase (I)/quadrature-phase (Q) modulators are extensively used in communication systems for QPSK modulation. As shown in Fig. 9.10, the modulator is comprised of two double-balanced mixers. The mixers are fed at the LO ports by a carrier phase-shifted through a 3-dB 90° hybrid coupler. The carrier signal thus has a relative phase of 0° to one mixer and 90° to the other mixer. Modulation signals are fed externally in phase quadrature to the IF ports of the two mixers. The output modulated signals from the two mixers are combined through a two-way 3-dB in-phase power divider/combiner.

The 8-PSK consists of eight (2^3) phase states and a theoretical bandwidth efficiency of 3 bps/Hz. It transmits eight phases of 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°. The 16-PSK transmits 16 phases. However, it is not used very much due to the small phase separation, which is difficult to maintain accurately. Instead, a modulation having both PSK and AM has evolved, called quadrature amplitude modulation (QAM). Figure 9.11 shows the output signal diagrams for 8-PSK, 16-PSK, and 16-QAM for comparison. Higher levels of QAM (64-, 256-, 1024-QAM)

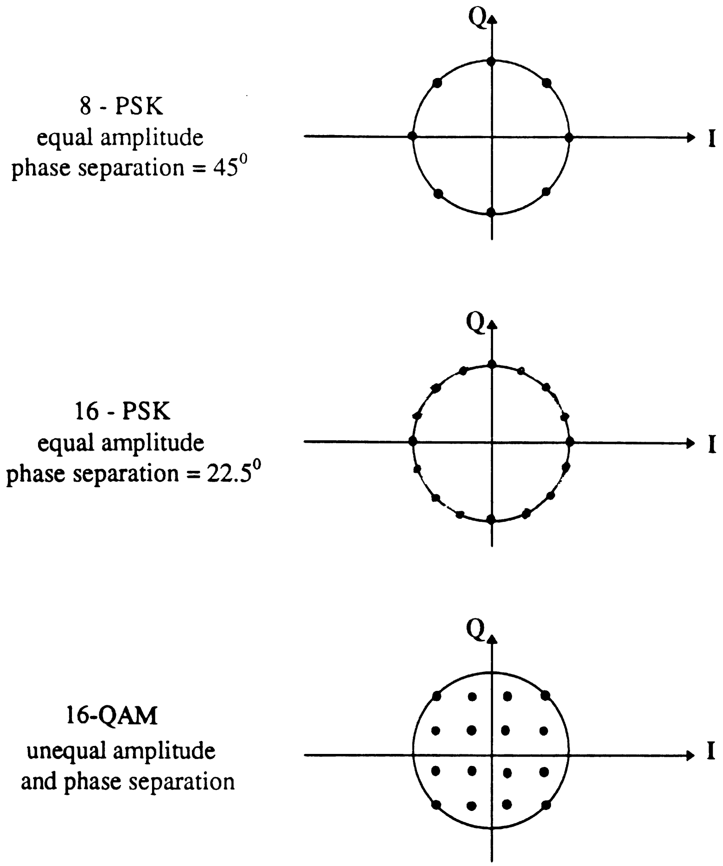

FIGURE 9.11 Constellation diagrams of signals for multilevel modulation.

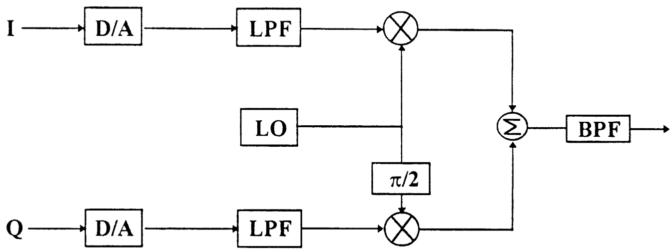

FIGURE 9.12 Quadrature amplitude modulator.

can be used for higher bandwidth efficiency. Figure 9.12 shows the typical QAM modulator block diagram. Two bit streams (I and Q) are provided from the pulse amplitude modulation process.

Some variations of QPSK are also in use. Offset-keyed or staggered quadriphase shift keying (OQPSK or SQPSK) modulation is used with only 90° phase transitions occurring in the modulator output signals. A $\frac{1}{4}\pi$-shifted, differentially encoded quadrature phase shift keying ($\frac{1}{4}\pi$-DQPSK) has been used for the U.S. and Japanese digital cellular time division multiple access (TDMA) radio standard; it has high power efficiency and spectral efficiency. In power-efficient, nonlinearly amplified (NLA) environments, where fully saturated class C amplifiers are used, the instantaneous 180° phase shift of conventional QPSK systems leads to a significant spectral regrowth and thus a low spectral efficiency. The OQPSK has 0° and ±90° phase transitions instead of 0°, ±90°, and ±180° for conventional QPSK. The compromise between conventional QPSK and OQPSK is $\frac{1}{4}\pi$-DQPSK, with 0°, ±45°, and ±135° phase transitions.

9.5 BIT ERROR RATE AND BANDWIDTH EFFICIENCY

A binary digital modulation system transmits a stream of data with binary symbols 0 and 1. The baseband information could be voice, music, fax, computer, or telemetry data. For analog information such as voice and music, an analog-to-digital (A/D) converter is used to convert the analog information into a digital form. The receiver will recover the data stream information.

In the ideal case, a receiver will recover the same binary digital stream that is transmitted, but the presence of noise in a communication system (e.g., transmitter, propagation, receiver) introduces the probability of errors that will be generated in the detection process. The likelihood that a bit is received incorrectly is called the bit error rate or the probability of error, defined as

$$\text{BER} = \frac{\text{false bits}}{\text{received bits}} \tag{9.10}$$

Example 9.2 A communication system transmits data at a rate of 2.048 Mb/sec. Two false bits are generated in each second. What is the BER?

Solution

$$\text{BER} = \frac{2\ \text{b/sec}}{2.048\ \text{Mb/sec}} = 9.76 \times 10^{-7} \qquad \blacksquare$$

The BER can be reduced if the system's SNR is increased. Figure 9.13 shows the BERs as a function of SNR for various modulation levels [4]. It can be seen that the BER drops rapidly as the SNR increases. The higher levels of modulation give better bandwidth efficiency but would require higher values of SNR to achieve a given BER. This is a trade-off between bandwidth efficiency and signal (carrier) power. The desired signal power must exceed the combined noise and interference power by an amount specified by the SNR ratio. The lower the SNR, the higher the BER and the more difficult it is to reconstruct the desired data information.

Coherent communication systems can improve the BER. The LO in the receiver is in synchronism with the incoming carrier. Synchronism means that both frequency and phase are identical. This can be accomplished in two ways: transmitting a pilot carrier or using a carrier-recovery circuit. The transmitting carrier used as a reference will pass through the same propagation delays as the modulated signal and will arrive at the receiver with the same phase and frequency. The carrier can be used to phase-lock the receiver's LO. Figure 9.14 shows the improvement in BER for coherent systems as compared to noncoherent systems [5].

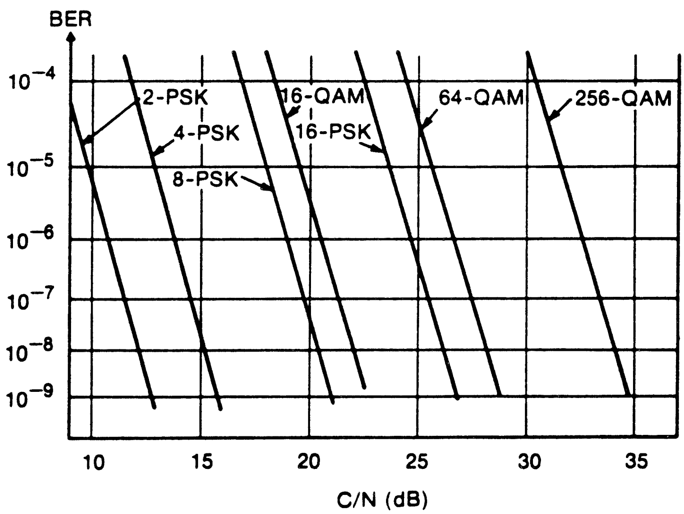

FIGURE 9.13 Bit error rates vs. carrier-to-noise ratio for different modulation schemes. (From reference [4], with permission from IEEE.)

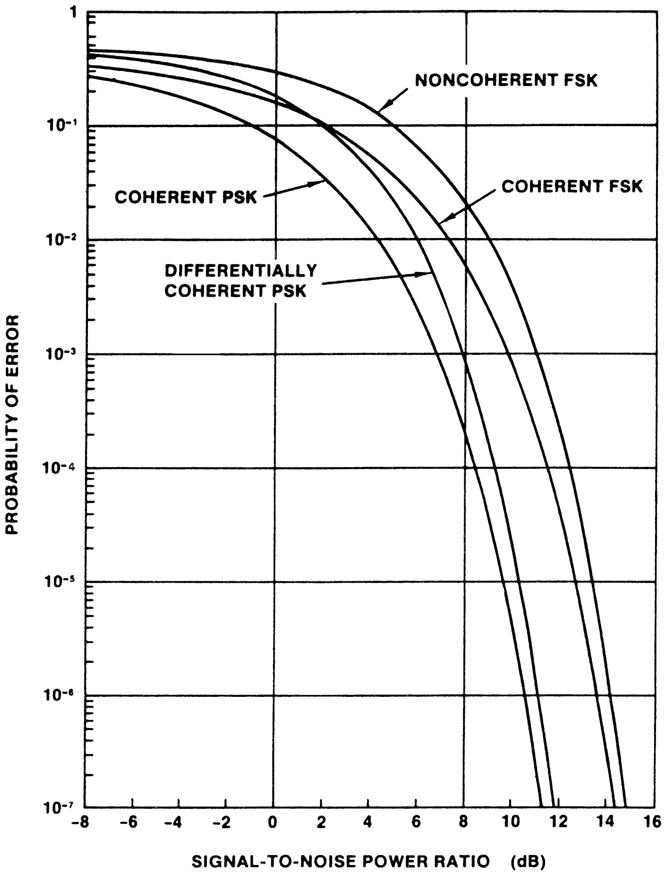

FIGURE 9.14 Comparison between coherent and noncoherent systems [5].

In binary digital modulation systems, if the system transmits 1 bit during each bit period, the system has a bandwidth efficiency of 1 bps/Hz. A bandwidth of 30 kHz can transmit (30 kHz)(1 bps/Hz) = 30 Kbps data rate. For the n-PSK and n-QAM modulation, the total number of states (or phases) is given by

$$n = 2^M \tag{9.11}$$

The theoretical bandwidth efficiency (η) is equal to M bps/Hz. The bandwidth efficiencies for BPSK, QPSK, 8-PSK, and 16-PSK are 1, 2, 3, and 4 bps/Hz, respectively. In other words, one can transmit 2, 4, 8, and 16 phases per second for these different modulation levels. In practice, when the number of states is increased, the separation between two neighboring states becomes smaller. This will cause uncertainty and increase the BER. Nonideal filtering characteristics also limit the

TABLE 9.1 Modulation Bandwidth Efficiencies

M	Modulation	Theoretical Bandwidth Efficiencies (bps/Hz)	Actual Bandwidth Efficiencies (bps/Hz)
1	BPSK	1	1
2	QPSK	2	2
3	8-PSK	3	2.5
4	16-PSK/16-QAM	4	3
6	64-QAM	6	4.5
8	256-QAM	8	6

bandwidth efficiency [2]. Therefore, the actual bandwidth efficiency becomes smaller, given by

$$\eta = 0.75M \tag{9.12}$$

Equation (9.12) is normally used for high-level or multilevel modulation with $M \geq 4$. Table 9.1 summarizes the bandwidth efficiencies for different modulation schemes.

9.6 SAMPLING AND PULSE CODE MODULATION

For a continuous signal (analog signal), the signal will be sampled and pulse encoded before the digital modulation takes place. If the discrete sampling points have sufficiently close spacing, a smooth curve can be drawn through them. It can be therefore said that the continuous curve is adequately described by the sampling points alone. One needs to transmit the sampling points instead of the entire continuous signal.

Figure 9.15 illustrates the sampling process and the results after sampling. The samples are then quantized, encoded, and modulated, as shown in the block diagram in Fig. 9.16. This process is called pulse code modulation (PCM). The advantages of this process are many:

1. The transmitted power can be concentrated into short bursts rather than delivered in CW. Usually, the pulses are quite short compared to the time between them, so the source is "off" most of the time. Peak power higher than the CW power can thus be transmitted.
2. The time intervals between pulses can be filled with sample values from other messages, thereby permitting the transmission of many messages on one communication system. This time-sharing transmission is called time division multiplexing.
3. The message is represented by a coded group of digital pulses. The effects of random noise can be virtually eliminated.

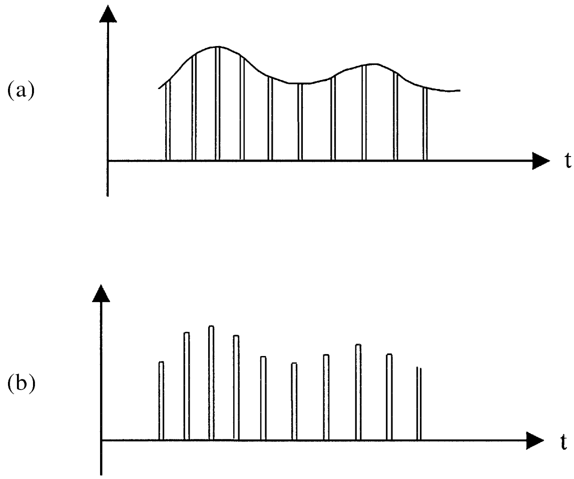

FIGURE 9.15 Sampling of a continuous signal: (*a*) continuous signal and sampling points; (*b*) sampled signal.

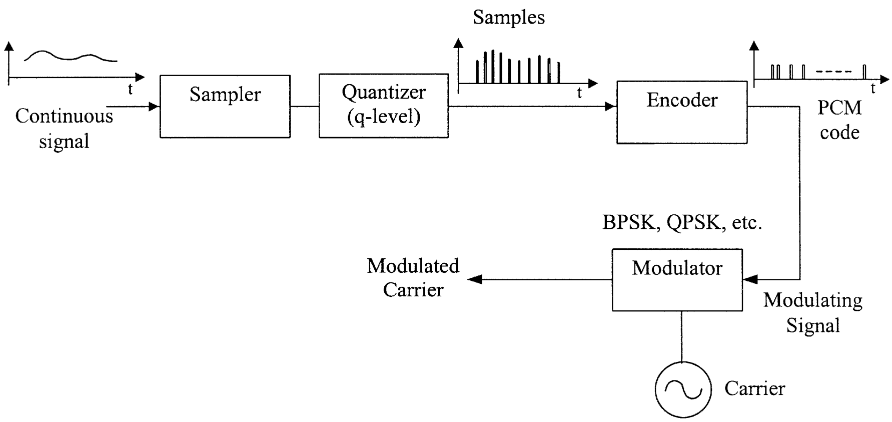

FIGURE 9.16 Pulse code modulation block diagram.

As shown in Fig. 9.16, the continuous signal is first sampled. The sample values are then rounded off to the nearest predetermined discrete value (quantum value). The encoder then converts the quantized samples into appropriate code groups, one group for each sample, and generates the corresponding digital pulses forming the baseband. This is basically an A/D converter.

The quantization is done depending on the number of pulses used. If N is the number of pulses used for each sample, the quantized levels q for a binary system are given by

$$q = 2^N \tag{9.13}$$

Figure 9.17 shows an example of sampling and encoding with $N = 3$. The more quantized levels used, the more accurate the sample data can be represented. However, it will require more bits (pulses) per sample transmission. Table 9.2 shows the number of bits per sample versus the number of quantizing steps. The PCM codes will be used as the modulating signal (information) to a digital modulator.

(a)

Quantized Samples	Binary Code
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

Binary code
4 2 1
2^2 2^1 2^0

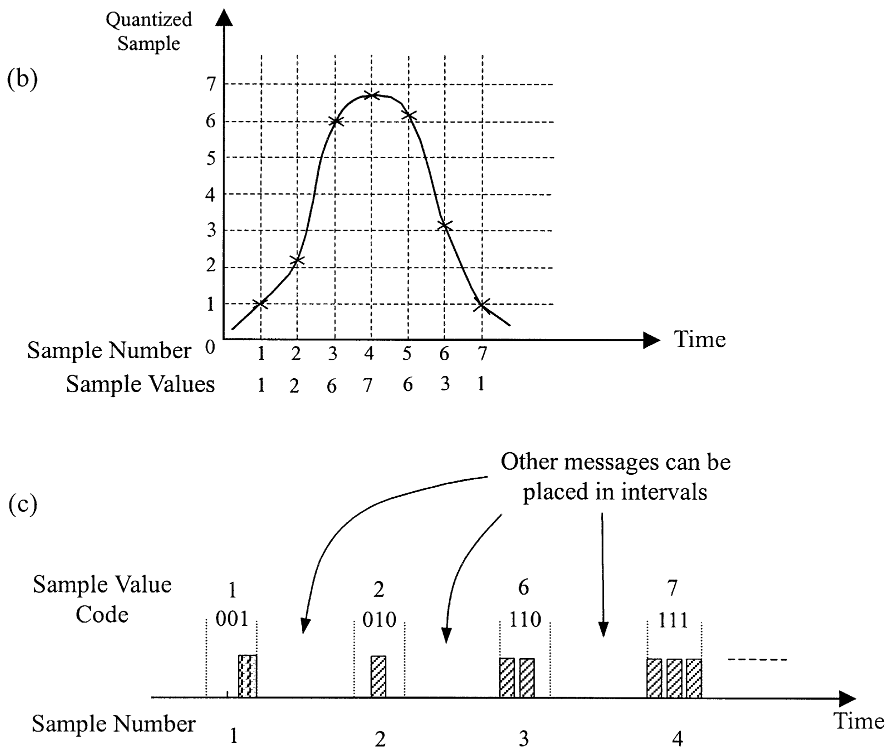

FIGURE 9.17 Sampling and PCM encoding for $N = 3$: (*a*) quantized samples and binary codes; (*b*) sampling and quantization; (*c*) PCM codes.

TABLE 9.2 Quantizing Steps

Number of Bits, N	Number of Quantizing Steps, q
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024
11	2048
12	4096

PROBLEMS

9.1 In an AM broadcast system, the total power transmitted is 1000 W. Calculate the transmitted power at the carrier frequency and at the upper and lower sidebands for an 80% modulation.

9.2 The power content for the carrier in an AM modulation is 1 kW. Determine the power content of each of the sidebands and the total transmitted power when the carrier is modulated at 75%.

9.3 Explain how the balanced modulator shown in Fig. P9.3 works.

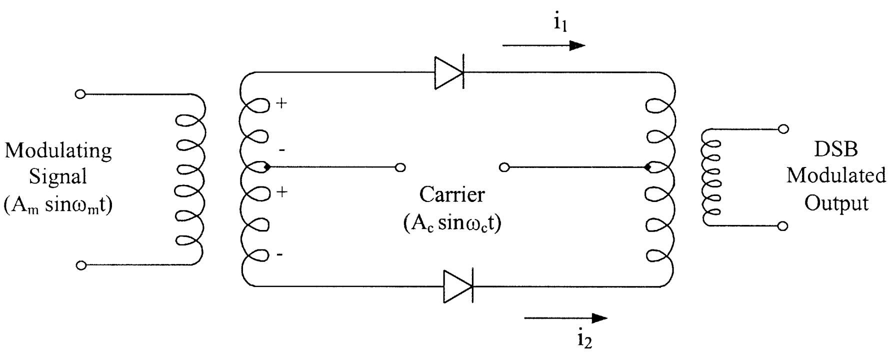

FIGURE P9.3

9.4 A 107.6-MHz carrier is frequency modulated by a 5-kHz sine wave. The frequency deviation is 50 kHz.

(a) Determine the carrier swing of the FM signal.

(b) Determine the highest and lowest frequencies attained by the modulated signal.

(c) What is the modulation index?

9.5 A QPSK communication link has a BER of 10^{-6}. The system data rate is 200 Mb/sec. Calculate the bandwidth requirement and the number of false bits generated per second. What is the bandwidth requirement if one uses 64-QAM modulation for the same data rate?

9.6 A QPSK mobile communication system has a maximum range of 20 km for a receiver output SNR of 10 dB. What is the SNR if the system is operated at a distance of 10 km? At the maximum range, if the system has 10 false bits per second for the transmission of 1 Mb/sec, what is the BER of this system?

9.7 For the signal shown in Fig. P9.7, a sample is generated every 1 msec, and four pulses are used for each sample (i.e., $N = 4$). Determine the quantized samples and binary codes used. What are the sampled values at each sampling time?

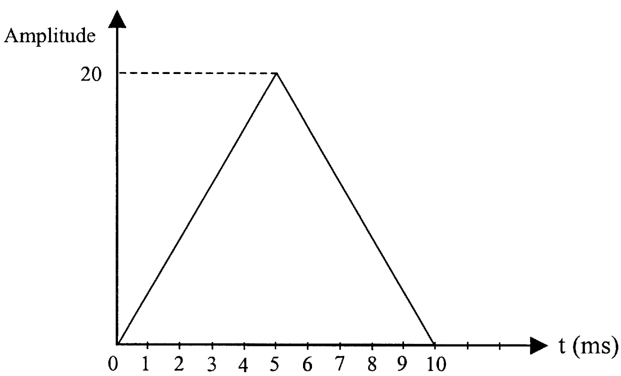

FIGURE P9.7

REFERENCES

1. K. Chang, *Microwave Solid-State Circuits and Applications*, John Wiley & Sons, New York, 1994.
2. R. G. Winch, *Telecommunication Transmission Systems*, McGraw-Hill, New York, 1993.
3. A. Grote and K. Chang, "60-GHz Integrated-Circuit High Data Rate Quadriphase Shift Keying Exciter and Modulator," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-32, No. 12, pp. 1663–1667, 1984.
4. T. Noguchi, Y. Daido, and J. A. Nossek, "Modulation Techniques for Microwave Digital Radio," *IEEE Commun. Mag.*, Vol. 24, No. 10, pp. 21–30, 1986.
5. E. A. Wolff and R. Kaul, *Microwave Engineering and Systems Applications*, John Wiley & Sons, New York, 1988.