where $f_c = 19$ kHz, and K is the amplitude of the pilot tone. The multiplexed signal $m(t)$ then frequency-modulates the main carrier to produce the transmitted signal. The pilot is allotted between 8 and 10 percent of the peak frequency deviation; the amplitude K in Eq. (4.58) is chosen to satisfy this requirement.

At a stereo receiver, the multiplexed signal $m(t)$ is recovered by frequency demodulating the incoming FM wave. Then $m(t)$ is applied to the *demultiplexing system* shown in Figure 4.15*b*. The individual components of the multiplexed signal $m(t)$ are separated by the use of three appropriate filters. The recovered pilot (using a narrowband filter tuned to 19 kHz) is frequency doubled to produce the desired 38-kHz subcarrier. The availability of this subcarrier enables the coherent detection of the DSB-SC modulated wave, thereby recovering the difference signal, $m_l(t) - m_r(t)$. The baseband low-pass filter in the top path of Figure 4.15*b* is designed to pass the sum signal, $m_l(t) + m_r(t)$. Finally, the simple matrixer reconstructs the left-hand signal $m_l(t)$ and right-hand signal $m_r(t)$ and applies them to their respective speakers.

4.4 PHASE-LOCKED LOOP

The *phase-locked loop* (PLL) is a negative feedback system, the operation of which is closely linked to frequency modulation. It can be used for synchronization, frequency division/multiplication, frequency modulation, and indirect frequency demodulation. The latter application is the subject of interest here.

Basically, the phase-locked loop consists of three major components: a *multiplier,* a *loop filter*, and a *voltage-controlled oscillator* (VCO) connected together in the form of a feedback loop, as in Figure 4.16. The VCO is a sinusoidal generator whose frequency is determined by a voltage applied to it from an external source. In effect, any frequency modulator may serve as a VCO.

We assume that initially we have adjusted the VCO so that when the control voltage is zero, two conditions are satisfied:

1. The frequency of the VCO in precisely set at the unmodulated carrier frequency f_c.
2. The VCO output has a 90-degree phase-shift with respect to the unmodulated carrier wave.

Suppose then that the input signal applied to the phase-locked loop is an FM signal defined by

$$s(t) = A_c \sin[2\pi f_c t + \phi_1(t)] \tag{4.59}$$

where A_c is the carrier amplitude. With a modulating signal $m(t)$, the angle $\phi_1(t)$ is related to $m(t)$ by the integral

$$\phi_1(t) = 2\pi k_f \int_0^t m(\tau)\, d\tau \tag{4.60}$$

where k_f is the frequency sensitivity of the frequency modulator. Let the VCO output in the phase-locked loop be defined by

$$r(t) = A_v \cos[2\pi f_c t + \phi_2(t)] \tag{4.61}$$

where A_v is the amplitude. With a control voltage $v(t)$ applied to the VCO input, the angle $\phi_2(t)$ is related to $v(t)$ by the integral

$$\phi_2(t) = 2\pi k_v \int_0^t v(t)\, dt \tag{4.62}$$

where k_v is the frequency sensitivity of the VCO, measured in Hertz per volt.

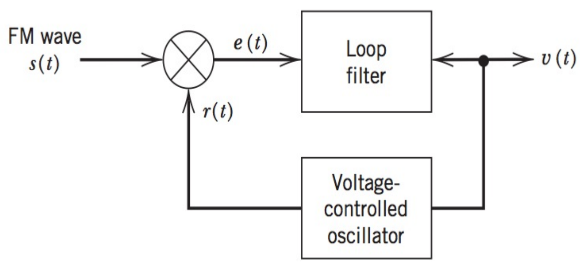

FIGURE 4.16 Phase-locked loop.

The object of the phase-locked loop is to generate a VCO output $r(t)$ that has the same phase angle (except for the fixed difference of 90 degrees) as the input FM signal $s(t)$. The time-varying phase angle $\phi_1(t)$ characterizing $s(t)$ may be due to modulation by a message signal $m(t)$ as in Eq. (4.60), in which case we wish to recover $\phi_1(t)$ in order to estimate $m(t)$. In other applications of the phase-locked loop, the time-varying phase angle $\phi_1(t)$ of the incoming signal $s(t)$ may be an unwanted phase shift caused by fluctuations in the communication channel; in this latter case, we wish to *track* $\phi_1(t)$ so as to produce a signal with the same phase angle for the purpose of coherent detection (synchronous demodulation).

To develop an understanding of the phase-locked loop, it is desirable to have a *model* of the loop. In what follows, we first develop a nonlinear model, which is subsequently linearized to simplify the analysis.

NONLINEAR MODEL OF THE PHASE-LOCKED LOOP[2]

According to Figure 4.16, the incoming FM signal $s(t)$ and the VCO output $r(t)$ are applied to the multiplier, producing two components:

1. A high-frequency component, represented by the *double-frequency* term

$$k_m A_c A_v \sin[4\pi f_c t + \phi_1(t) + \phi_2(t)]$$

2. A low-frequency component represented by the *difference-frequency* term

$$k_m A_c A_v \sin[\phi_1(t) - \phi_2(t)]$$

where k_m is the *multiplier gain*, measured in volt^{-1}.

The loop filter in the phase-lock loop is a low-pass filter, and its response to the high-frequency component will be negligible. The VCO also contributes to the attenuation of this component. Therefore, discarding the high-frequency component (i.e., the double-frequency term), the input to the loop filter is reduced to

$$e(t) = k_m A_c A_v \sin[\phi_e(t)] \tag{4.63}$$

where $\phi_e(t)$ is the *phase error* defined by

$$\begin{aligned} \phi_e(t) &= \phi_1(t) - \phi_2(t) \\ &= \phi_1(t) - 2\pi k_v \int_0^t v(\tau)\, d\tau \end{aligned} \tag{4.64}$$

The loop filter operates on the input $e(t)$ to produce an output $v(t)$ defined by the convolution integral

$$v(t) = \int_{-\infty}^{\infty} e(\tau) h(t-\tau)\, d\tau \tag{4.65}$$

where $h(t)$ is the impulse response of the loop filter.

Using Eqs. (4.62) to (4.64) to relate $\phi_e(t)$ and $\phi_1(t)$, we obtain the following nonlinear integro-differential equation as the descriptor of the dynamic behavior of the phase-locked loop:

$$\frac{d\phi_e(t)}{dt} = \frac{d\phi_1(t)}{dt} - 2\pi K_0 \int_{-\infty}^{\infty} \sin[\phi_e(\tau)] h(t-\tau)\, d\tau \tag{4.66}$$

where K_0 is a *loop-gain parameter* defined by

$$K_0 = k_m k_v A_c A_v \tag{4.67}$$

The amplitudes A_c and A_v are both measured in volts, the multiplier gain k_m in volt^{-1} and the frequency sensitivity k_v in Hertz per volt. Hence, it follows from Eq. (4.67) that K_0

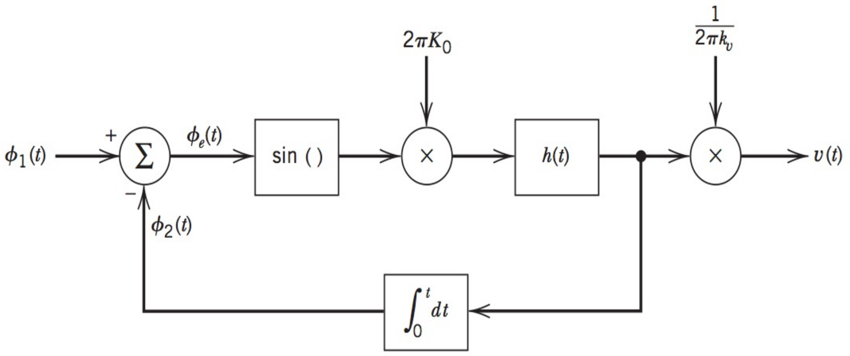

FIGURE 4.17 Nonlinear model of the phase-locked loop.

has the dimensions of frequency. Equation (4.66) suggests the model shown in Figure 4.17 for a phase-locked loop. In this model we have also included the relationship between $v(t)$ and $e(t)$ as represented by Eqs. (4.63) and (4.65). We see that the model resembles the block diagram of Figure 4.16. The multiplier at the input of the phase-locked loop is replaced by a subtracter and a sinusoidal nonlinearity, and the VCO by an integrator.

The sinusoidal nonlinearity in the model of Figure 4.17 greatly increases the difficulty of analyzing the behavior of the phase-locked loop. It would be helpful to *linearize* this model to simplify the analysis, yet give a good approximate description of the loop's behavior in certain modes of operation. This we do next.

LINEAR MODEL OF THE PHASE-LOCKED LOOP

When the phase error $\phi_e(t)$ is zero, the phase-locked loop is said to be in *phase-lock*. When $\phi_e(t)$ is at all times small compared with one radian, we may use the approximation

$$\sin[\phi_e(t)] \simeq \phi_e(t) \tag{4.68}$$

which is accurate to within 4 percent for $\phi_e(t)$ less than 0.5 radians. In this case, the loop is said to be *near phase-lock*, and the sinusoidal nonlinearity of Figure 4.17 may be disregarded. Thus, we may represent the phase-locked loop by the linearized model shown in Figure 4.18*a*. According to this model, the phase error $\phi_e(t)$ is

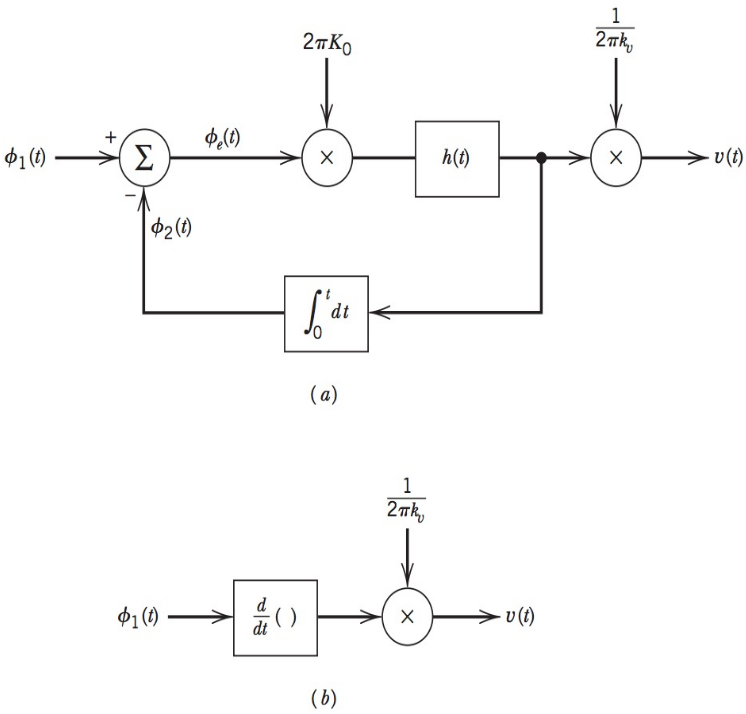

FIGURE 4.18 Models of the phase-locked loop. (*a*) Linearized model. (*b*) Simplified model when the loop gain is very large compared to unity.

related to the input phase $\phi_1(t)$ by the *linear integro-differential equation*

$$\frac{d\phi_e(t)}{dt} + 2\pi K_0 \int_{-\infty}^{\infty} \phi_e(\tau)h(t-\tau)\, d\tau = \frac{d\phi_1(t)}{dt} \tag{4.69}$$

Transforming Eq. (4.69) into the frequency domain and solving for $\Phi_e(f)$, the Fourier transform of $\phi_e(t)$, in terms of $\Phi_1(f)$, the Fourier transform of $\phi_1(t)$, we get

$$\Phi_e(f) = \frac{1}{1+L(f)}\Phi_1(f) \tag{4.70}$$

The function $L(f)$ in Eq. (4.70) is defined by

$$L(f) = K_0 \frac{H(f)}{jf} \tag{4.71}$$

where $H(f)$ is the transfer function of the loop filter. The quantity $L(f)$ is called the *open-loop transfer function* of the phase-locked loop. Suppose that for all values of f inside the baseband we make the magnitude of $L(f)$ very large compared with unity. Then from Eq. (4.70) we find that $\Phi_e(f)$ approaches zero. That is, the phase of the VCO becomes asymptotically equal to the phase of the incoming signal. Under this condition, phase-lock is established, and the objective of the phase-locked loop is thereby satisfied.

From Figure 4.18*a* we see that $V(f)$, the Fourier transform of the phase-locked loop output $v(t)$, is related to $\Phi_e(f)$ by

$$V(f) = \frac{K_0}{k_v}H(f)\Phi_e(f) \tag{4.72}$$

Equivalently, in light of Eq. (4.71), we may write

$$V(f) = \frac{jf}{k_v}L(f)\Phi_e(f) \tag{4.73}$$

Therefore, substituting Eq. (4.70) in (4.73), we get

$$V(f) = \frac{(jf/k_v)L(f)}{1+L(f)}\Phi_1(f) \tag{4.74}$$

Again, when we make $|L(f)| \gg 1$ for the frequency band of interest, we may approximate Eq. (4.74) as follows:

$$V(f) \simeq \frac{jf}{k_v}\Phi_1(f) \tag{4.75}$$

The corresponding time-domain relation is

$$v(t) \simeq \frac{1}{2\pi k_v}\frac{d\phi_1(t)}{dt} \tag{4.76}$$

Thus, provided that the magnitude of the open-loop transfer function $L(f)$ is very large for all frequencies of interest, the phase-locked loop may be modeled as a *differentiator* with its output scaled by the factor $1/2\pi k_v$, as in Figure 4.18*b*.

The simplified model of Figure 4.18*b* provides an indirect method of using the phase-locked loop as a frequency demodulator. When the input is an FM signal as in Eq. (4.59), the angle $\phi_1(t)$ is related to the message signal $m(t)$ as in Eq. (4.60). Therefore, substituting Eq. (4.60) in (4.76), we find that the resulting output signal of the phase-locked loop is approximately

$$v(t) \simeq \frac{k_f}{k_v}m(t) \tag{4.77}$$

Equation (4.77) states that when the loop operates in its phase-locked mode, the output $v(t)$ of the phase-locked loop is approximately the same, except for the scale factor k_f/k_v, as the original message signal $m(t)$; frequency demodulation of the incoming FM signal $s(t)$ is thereby accomplished.

A significant feature of the phase-locked loop acting as a demodulator is that the bandwidth of the incoming FM signal can be much wider than that of the loop filter characterized by $H(f)$. The transfer function $H(f)$ can and should be restricted to the baseband. Then the control signal of the VCO has the bandwidth of the baseband (message) signal $m(t)$, whereas the VCO output is a wideband frequency-modulated signal whose instantaneous frequency tracks that of the incoming FM signal. Here we are merely restating the fact that the bandwidth of a wide-band FM signal is much larger than the bandwidth of the message signal responsible for its generation.

The complexity of the phase-locked loop is determined by the transfer function $H(f)$ of the loop filter. The simplest form of a phase-locked loop is obtained when $H(f) = 1$; that is, there is no loop filter, and the resulting phase-locked loop is referred to as a *first-order phase-locked loop*. For higher-order loops, the transfer function $H(f)$ assumes a more complex form. The order of the phase-locked loop is determined by the order of the denominator polynomial of the *closed-loop transfer function*, which defines the output transform $V(f)$ in terms of the input transform $\Phi_1(f)$, as shown in Eq. (4.74).

A major limitation of a first-order phase-locked loop is that the loop gain parameter K_0 controls both the loop bandwidth as well as the hold-in frequency range of the loop; the *hold-in frequency range* refers to the range of frequencies for which the loop remains phase-locked to the input signal. It is for this reason that a first-order phase-locked loop is seldom used in practice. Accordingly, in the remainder of this section we deal only with a second-order phase-locked loop.

SECOND-ORDER PHASE-LOCKED LOOP

To be specific, consider a *second-order phase-locked loop* using a loop filter with the transfer function

$$H(f) = 1 + \frac{a}{jf} \tag{4.78}$$

where a is a constant. The filter consists of an integrator (using an operational amplifier) and a direct connection, as shown in Figure 4.19. For this phase-locked loop, the use of Eqs. (4.70) and (4.78) yields

$$\Phi_e(f) = \frac{(jf)^2/aK_0}{1 + [(jf)/a] + [(jf)^2/aK_0]}\Phi_1(f) \tag{4.79}$$

Define the *natural frequency* of the loop:

$$f_n = \sqrt{aK_0} \tag{4.80}$$

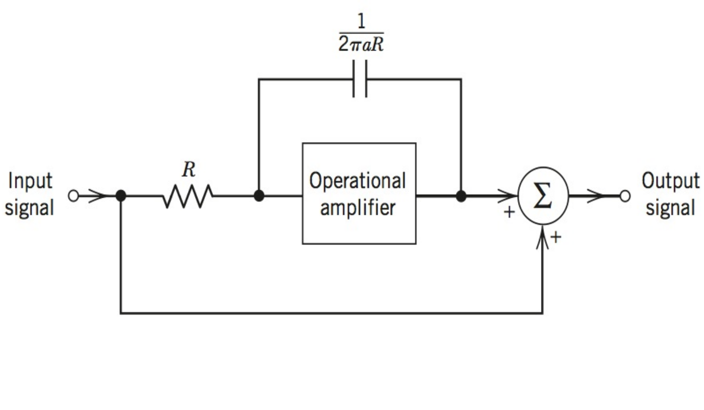

FIGURE 4.19 Loop filter for second-order phase-locked loop.

and the *damping factor*:

$$\zeta = \sqrt{\frac{K_0}{4a}} \tag{4.81}$$

Then we may recast Eq. (4.79) in terms of the parameters f_n and ζ as follows:

$$\Phi_e(f) = \left(\frac{(jf/f_n)^2}{1 + 2\zeta(jf/f_n) + (jf/f_n)^2} \right) \Phi_1(f) \tag{4.82}$$

Assume that the incoming FM signal is produced by a single-tone modulating wave, for which the phase input is

$$\phi_1(t) = \beta \sin(2\pi f_m t) \tag{4.83}$$

Hence, from Eq. (4.82) we find that the corresponding phase error is

$$\phi_e(t) = \phi_{e0} \cos(2\pi f_m t + \psi) \tag{4.84}$$

where the amplitude ϕ_{e0} and phase ψ are, respectively, defined by

$$\phi_{e0} = \frac{(\Delta f/f_n)(f_m/f_n)}{\{[1 - (f_m/f_n)^2]^2 + 4\zeta^2 (f_m/f_n)^2\}^{1/2}} \tag{4.85}$$

and

$$\psi = \frac{\pi}{2} - \tan^{-1}\left[\frac{2\zeta f_m/f_n}{1 - (f_m/f_n)^2} \right] \tag{4.86}$$

In Figure 4.20 we have plotted the phase error amplitude ϕ_{e0}, normalized with respect to $\Delta f/f_n$, versus f_m/f_n for different values of ζ. It is apparent that for all values of the damping factor ζ, and assuming a fixed frequency deviation Δf, the phase error is small at low modulation frequencies, rises to a maximum at $f_m = f_n$, and then falls off at higher modulation frequencies. Note also that the maximum value of phase error amplitude decreases with increasing ζ.

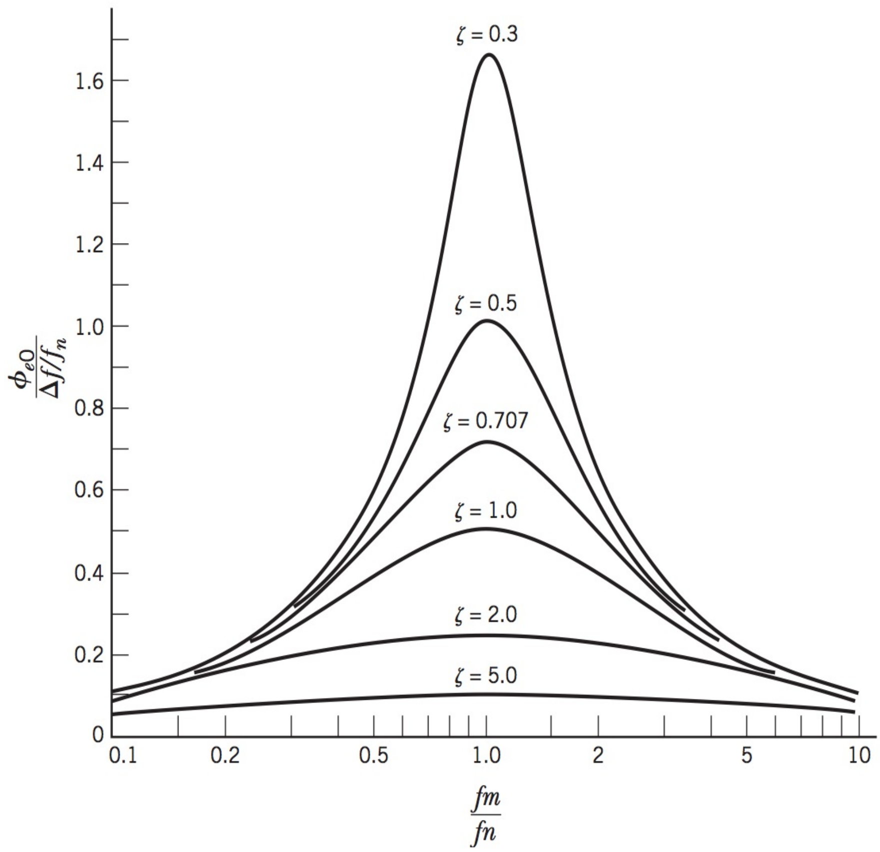

FIGURE 4.20 Phase-error amplitude characteristic of second-order phase-locked loop.

The Fourier transform of the loop output is related to $\Phi_e(f)$ by Eq. (4.72); hence, with $H(f)$ as defined in Eq. (4.78), we get

$$V(f) = \frac{K_0}{k_v}\left(1 + \frac{a}{jf}\right)\Phi_e(f) \tag{4.87}$$

In light of the definitions given in Eqs. (4.80) and (4.81), we have

$$V(f) = \left(\frac{f_n^2}{jfk_v}\right)\left[1 + 2\zeta\left(\frac{jf}{f_n}\right)\right]\Phi_e(f) \tag{4.88}$$

Substituting Eq. (4.82) in (4.88), we get

$$V(f) = \left(\frac{(jf/k_v)[1 + 2\zeta(jf/f_n)]}{1 + 2\zeta(jf/f_n) + (jf/f_n)^2}\right)\Phi_1(f) \tag{4.89}$$

Therefore, for the phase input $\phi_1(t)$ of Eq. (4.83), we find that the corresponding loop output is

$$v(t) = A_0 \cos(2\pi f_m t + \alpha) \tag{4.90}$$

where the amplitude A_0 and phase α are, respectively, defined by

$$A_0 = \frac{(\Delta f/k_v)[1 + 4\zeta^2(f_m/f_n)^2]^{1/2}}{\{[1 - (f_m/f_n)^2]^2 + 4\zeta^2(f_m/f_n)^2\}^{1/2}} \tag{4.91}$$

and

$$\alpha = \tan^{-1}\left[2\zeta\left(\frac{f_m}{f_n}\right)\right] - \tan^{-1}\left[\frac{2\zeta(f_m/f_n)}{1 - (f_m/f_n)^2}\right] \tag{4.92}$$

From Eq. (4.91), we see that the amplitude A_0 attains its maximum value of $\Delta f/k_v$ at $(f_m/f_n) = 0$; it decreases with increasing f_m/f_n, dropping to zero at $(f_m/f_n) = \infty$.

The important feature of the second-order phase-locked loop is that with an incoming FM signal produced by a modulating sinusoidal wave of fixed amplitude (corresponding to a fixed frequency deviation) and varying frequency, the frequency response that defines the phase error $\phi_e(t)$ is representative of a band-pass filter [see Eq. (4.85)], but the frequency response that defines the loop output $v(t)$ is representative of a low-pass filter [see Eq. (4.91)]. Therefore, by appropriately choosing the parameters ζ and f_n, which determine the frequency response of the loop, it is possible to restrain the phase error to always remain small and thereby lie within the linear range of the loop, whereas at the same time the modulating (message) signal is reproduced at the loop output with minimum distortion. This restraint is, however, conservative with respect to the hold-in capabilities of the loop. As a reasonable rule of thumb, the loop should remain locked if the maximum value of the phase error ϕ_{e0} (which occurs when the modulation frequency f_m is equal to the loop's natural frequency f_n) is always less than 90 degrees.

Phase-locked loop performance is explored experimentally in Problem 4.29.

4.5 NONLINEAR EFFECTS IN FM SYSTEMS

In the preceding three sections, we studied frequency modulation theory and methods for its generation and demodulation. We complete the discussion of frequency modulation by considering nonlinear effects in FM systems.

Nonlinearities, in one form or another, are present in all electrical networks. There are two basic forms of nonlinearity to consider:

1. The nonlinearity is said to be *strong* when it is introduced intentionally and in a controlled manner for some specific application. Examples of strong nonlinearity include square-law modulators, limiters, and frequency multipliers.
2. The nonlinearity is said to be *weak* when a linear performance is desired, but nonlinearities of a parasitic nature arise due to imperfections. The effect of such weak nonlinearities is to limit the useful signal levels in a system and thereby become an important design consideration.

In this section we examine the effects of weak nonlinearities on frequency modulation.[3]

Consider a communications channel, the transfer characteristic of which is defined by the nonlinear input–output relation

$$v_o(t) = a_1 v_i(t) + a_2 v_i^2(t) + a_3 v_i^3(t) \tag{4.93}$$

where $v_i(t)$ and $v_o(t)$ are the input and output signals, respectively, and a_1, a_2 and a_3 are constants. The channel described in Eq. (4.93) is said to be *memoryless* in that the output signal $v_o(t)$ is an instantaneous function of the input signal $v_i(t)$ (i.e., there is no energy storage involved in the description). We wish to determine the effect of transmitting a frequency-modulated wave through such a channel. The FM signal is defined by

$$v_i(t) = A_c \cos[2\pi f_c t + \phi(t)]$$

where

$$\phi(t) = 2\pi k_f \int_0^t m(\tau)\, d\tau$$

For this input signal, the use of Eq. (4.93) yields

$$\begin{aligned} v_o(t) = {} & a_1 A_c \cos[2\pi f_c t + \phi(t)] + a_2 A_c^2 \cos^2[2\pi f_c t + \phi(t)] \\ & + a_3 A_c^3 \cos^3[2\pi f_c t + \phi(t)] \end{aligned} \tag{4.94}$$

Expanding the squared and cubed cosine terms in Eq. (4.94) and then collecting common terms, we get

$$\begin{aligned} v_o(t) = {} & \frac{1}{2} a_2 A_c^2 + \left(a_1 A_c + \frac{3}{4} a_3 A_c^3\right) \cos[2\pi f_c t + \phi(t)] \\ & + \frac{1}{2} a_2 A_c^2 \cos[4\pi f_c t + 2\phi(t)] \\ & + \frac{1}{4} a_3 A_c^3 \cos[6\pi f_c t + 3\phi(t)] \end{aligned} \tag{4.95}$$

Thus the channel output consists of a dc component and three frequency-modulated signals with carrier frequencies f_c, $2f_c$, and $3f_c$; the sinusoidal components are contributed by the linear, second-order, and third-order terms of Eq. (4.93), respectively.

To extract the desired FM signal from the channel output $v_o(t)$, that is, the particular component with carrier frequency f_c, it is necessary to separate the FM signal with this carrier frequency from the one with the closest carrier frequency: $2f_c$. Let Δf denote the frequency deviation of the incoming FM signal $v_i(t)$, and W denote the highest frequency component of the message signal $m(t)$. Then, applying Carson's rule and noting that the frequency deviation about the second harmonic of the carrier frequency is doubled, we find that the necessary condition for separating the desired FM signal with the carrier frequency f_c from that with the carrier frequency $2f_c$ is

$$2f_c - (2\Delta f + W) > f_c + \Delta f + W$$

or

$$f_c > 3\Delta f + 2W \tag{4.96}$$

Thus, by using a band-pass filter of mid-band frequency f_c and bandwidth $2\Delta f + 2W$, the channel output is reduced to

$$v_o'(t) = \left(a_1 A_c + \frac{3}{4} a_3 A_c^3\right) \cos[2\pi f_c t + \phi(t)] \tag{4.97}$$

We see therefore that the only effect of passing an FM signal through a channel with amplitude nonlinearities, followed by appropriate filtering, is simply to modify its amplitude. That is, unlike amplitude modulation, frequency modulation is not affected by distortion produced by transmission through a channel with amplitude nonlinearities. It is for this reason that we find frequency modulation widely used in microwave radio and satellite communication systems: It permits the use of highly nonlinear amplifiers and power transmitters, which are particularly important to producing a maximum power output at radio frequencies.

An FM system is extremely sensitive to *phase nonlinearities*, however, as we would intuitively expect. A common type of phase nonlinearity that is encountered in microwave radio systems is known as *AM-to-PM conversion.* This is the result of the phase characteristic of repeaters or amplifiers used in the system being dependent on the instantaneous amplitude of the input signal. In practice, AM-to-PM conversion is characterized by a constant K, which is measured in degrees per dB and may be interpreted as the peak phase change at the output for a 1-dB change in envelope at the input. When an FM wave is transmitted through a microwave radio link, it picks up spurious amplitude variations due to noise and interference during the course of transmission, and when such an FM wave is passed through a repeater with AM-to-PM conversion, the output will contain unwanted phase modulation and resultant distortion. It is therefore important to keep the AM-to-PM conversion at a low level. For example, for a good microwave repeater, the AM-to-PM conversion constant K is less than 2 degrees per dB.

4.6 THE SUPERHETERODYNE RECEIVER

In a communication system, irrespective of whether it is based on amplitude modulation or frequency modulation, the receiver not only has the task of demodulating the incoming modulated signal, but it is also required to perform some other system functions:

- *Carrier-frequency tuning*, the purpose of which is to select the desired signal (i.e., desired radio or TV station).
- *Filtering*, which is required to separate the desired signal from other modulated signals that may be picked up along the way.
- *Amplification*, which is intended to compensate for the loss of signal power incurred in the course of transmission.

The *superheterodyne receiver*, or *superhet* as it is often referred to, is a special type of receiver that fulfils all three functions, particularly the first two, in an elegant and practical fashion. Specifically, it overcomes the difficulty of having to build a tunable high- (and variable-) Q filter. Indeed, practically all analog radio and TV receivers are of the superheterodyne type.

Basically, the receiver consists of a radio-frequency (RF) section, a mixer and local oscillator, an intermediate frequency (IF) section, demodulator, and power amplifier. Typical frequency parameters of commercial AM and FM radio receivers are listed in Table 4.2. Figure 4.21 shows the block diagram of a superheterodyne receiver for amplitude modulation using an envelope detector for demodulation.