Two properties of *linear* CW modulation bear repetition at the outset of this chapter: the modulated spectrum is basically the translated message spectrum and the transmission bandwidth never exceeds twice the message bandwidth. A third property, derived in Chap. 10, is that the destination signal-to-noise ratio $(S/N)_D$ is no better than baseband transmission and can be improved only by increasing the transmitted power. **Exponential** modulation differs on all three counts.

In contrast to linear modulation, exponential modulation is a *nonlinear* process; therefore, it should come as no surprise that the modulated spectrum is not related in a simple fashion to the message spectrum. Moreover, it turns out that the transmission bandwidth is usually much greater than twice the message bandwidth. Compensating for the bandwidth liability is the fact that exponential modulation can provide increased signal-to-noise ratios without increased transmitted power. Exponential modulation thus allows you to *trade bandwidth for power* in the design of a communication system.

We begin our study of exponential modulation by defining the two basic types, **phase modulation** (PM) and **frequency modulation** (FM). We'll examine signals and spectra, investigate the transmission bandwidth and distortion problem, and describe typical hardware for generation and detection. The analysis of interference at the end of the chapter brings out the value of FM for radio broadcasting and sets the stage for our consideration of noise in Chap. 10.

OBJECTIVES

After studying this chapter and working the exercises, you should be able to do each of the following:

1. Find the instantaneous phase and frequency of a signal with exponential modulation (Sect. 5.1).
2. Construct the line spectrum and phasor diagram for FM or PM with tone modulation (Sect. 5.1).
3. Estimate the bandwidth required for FM or PM transmission (Sect. 5.2).
4. Identify the effects of distortion, limiting, and frequency multiplication on an FM or PM signal (Sect. 5.2).
5. Design an FM generator and detector appropriate for an application (Sect. 5.3).
6. Use a phasor diagram to analyze interference in AM, FM, and PM (Sect. 5.4).

5.1 PHASE AND FREQUENCY MODULATION

This section introduces the concepts of instantaneous phase and frequency for the definition of PM and FM signals. Then, since the nonlinear nature of exponential modulation precludes spectral analysis in general terms, we must work instead with the spectra resulting from particular cases such as narrowband modulation and tone modulation.

PM and FM Signals

Consider a CW signal with constant envelope but time-varying phase, so

$$x_c(t) = A_c \cos [\omega_c t + \phi(t)] \tag{1}$$

Upon defining the **total instantaneous angle**

$$\theta_c(t) \triangleq \omega_c t + \phi(t)$$

we can express $x_c(t)$ as

$$x_c(t) = A_c \cos \theta_c(t) = A_c \operatorname{Re}\left[e^{j\theta_c(t)}\right]$$

Hence, if $\theta_c(t)$ contains the message information $x(t)$, we have a process that may be termed either **angle** modulation or **exponential** modulation. We'll use the latter name because it emphasizes the nonlinear relationship between $x_c(t)$ and $x(t)$.

As to the specific dependence of $\theta_c(t)$ on $x(t)$, **phase modulation** (PM) is defined by

$$\phi(t) \triangleq \phi_\Delta x(t) \qquad \phi_\Delta \leq 180° \tag{2}$$

so that

$$x_c(t) = A_c \cos\left[\omega_c t + \phi_\Delta x(t)\right] \tag{3}$$

These equations state that the instantaneous phase varies directly with the modulating signal. The constant ϕ_Δ represents the **maximum phase shift** produced by $x(t)$, since we're still keeping our normalization convention $|x(t)| \leq 1$. The upper bound $\phi_\Delta \leq 180°$ (or π radians) limits $\phi(t)$ to the range $\pm 180°$ and prevents phase ambiguities—after all, there's no physical distinction between angles of $+270$ and $-90°$, for instance. The bound on ϕ_Δ is analogous to the restriction $\mu \leq 1$ in AM, and ϕ_Δ can justly be called the **phase modulation index,** or the **phase deviation.**

The rotating-phasor diagram in Fig. 5.1–1 helps interpret phase modulation and leads to the definition of frequency modulation. The total angle $\theta_c(t)$ consists of the constant rotational term $\omega_c t$ plus $\phi(t)$, which corresponds to angular shifts relative to the dashed line. Consequently, the phasor's instantaneous rate of rotation in cycles per second will be

$$f(t) \triangleq \frac{1}{2\pi}\dot{\theta}_c(t) = f_c + \frac{1}{2\pi}\dot{\phi}(t) \tag{4}$$

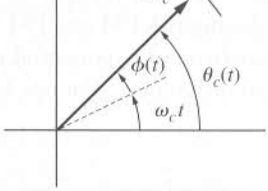

Figure 5.1–1 Rotating-phasor representation of exponential modulation.

in which the dot notation stands for the time derivative, that is, $\dot{\phi}(t) = d\phi(t)/dt$, and so on. We call $f(t)$ the **instantaneous frequency** of $x_c(t)$. Although $f(t)$ is measured in *hertz*, it should *not* be equated with spectral frequency. Spectral frequency f is the independent variable of the frequency domain, whereas instantaneous frequency $f(t)$ is a time-dependent property of waveforms with exponential modulation.

In the case of **frequency modulation** (FM), the instantaneous frequency of the modulated wave is defined to be

$$f(t) \triangleq f_c + f_\Delta x(t) \qquad f_\Delta < f_c \tag{5}$$

so $f(t)$ varies in proportion with the modulating signal. The proportionality constant f_Δ, called the **frequency deviation,** represents the maximum shift of $f(t)$ relative to the carrier frequency f_c. The upper bound $f_\Delta < f_c$ simply ensures that $f(t) > 0$. However, we usually want $f_\Delta \ll f_c$ in order to preserve the bandpass nature of $x_c(t)$.

Equations (4) and (5) show that an FM wave has $\dot{\phi}(t) = 2\pi f_\Delta\, x(t)$, and integration yields the phase modulation

$$\phi(t) = 2\pi f_\Delta \int_{t_0}^{t} x(\lambda)\, d\lambda + \phi(t_0) \qquad t \geq t_0 \tag{6a}$$

If t_0 is taken such that $\phi(t_0) = 0$, we can drop the lower limit of integration and use the informal expression

$$\phi(t) = 2\pi f_\Delta \int^{t} x(\lambda)\, d\lambda \tag{6b}$$

The FM waveform is then written as

$$x_c(t) = A_c \cos\left[\omega_c t + 2\pi f_\Delta \int^{t} x(\lambda)\, d\lambda\right] \tag{7}$$

But it must be assumed that the message has no *dc component* so the above integrals do not diverge when $t \to \infty$. Physically, a dc term in $x(t)$ would produce a constant carrier-frequency shift equal to $f_\Delta\langle x(t)\rangle$.

A comparison of Eqs. (3) and (7) implies little difference between PM and FM, the essential distinction being the integration of the message in FM. Moreover, nomenclature notwithstanding, both FM and PM have both time-varying phase and frequency, as underscored by Table 5.1–1. These relations clearly indicate that, with the help of integrating and differentiating networks, a phase modulator can produce frequency modulation and vice versa. In fact, in the case of tone modulation it's nearly impossible visually to distinguish FM and PM waves.

On the other hand, a comparison of exponential modulation with linear modulation reveals some pronounced differences. For one thing,

The amplitude of an exponentially modulated wave is *constant*.

Table 5.1–1 Comparison of PM and FM

	Instantaneous phase $\phi(t)$	Instantaneous frequency $f(t)$
PM	$\phi_\Delta x(t)$	$f_c + \frac{1}{2\pi}\phi_\Delta \dot{x}(t)$
FM	$2\pi f_\Delta \int^t x(\lambda)\,d\lambda$	$f_c + f_\Delta x(t)$

Therefore, regardless of the message $x(t)$, the average transmitted power is

$$S_T = \tfrac{1}{2}A_c^2 \qquad [8]$$

For another, the zero crossings of an exponentially modulated wave are *not periodic*, whereas they are always periodic in linear modulation. Indeed, because of the constant-amplitude property of FM and PM, it can be said that

The message resides in the *zero crossings alone*, providing the carrier frequency is large.

Finally, since exponential modulation is a nonlinear process,

The modulated wave does not look at all like the message waveform.

Figure 5.1–2 illustrates some of these points by showing typical AM, FM, and PM waves. As a mental exercise you may wish to check these waveforms against the corresponding modulating signals. For FM and PM this is most easily done by considering the instantaneous frequency rather than by substituting $x(t)$ in Eqs. (3) and (7).

Despite the many similarities of PM and FM, frequency modulation turns out to have superior noise-reduction properties and thus will receive most of our attention. To gain a qualitative appreciation of FM noise reduction, suppose a demodulator simply extracts the instantaneous frequency $f(t) = f_c + f_\Delta x(t)$ from $x_c(t)$. The demodulated output is then proportional to the frequency deviation f_Δ, which can be increased without increasing the transmitted power S_T. If the noise level remains constant, increased signal output is equivalent to reduced noise. However, noise reduction does require increased transmission bandwidth to accommodate large frequency deviations.

Ironically, frequency modulation was first conceived as a means of **bandwidth reduction,** the argument going somewhat as follows: If, instead of modulating the

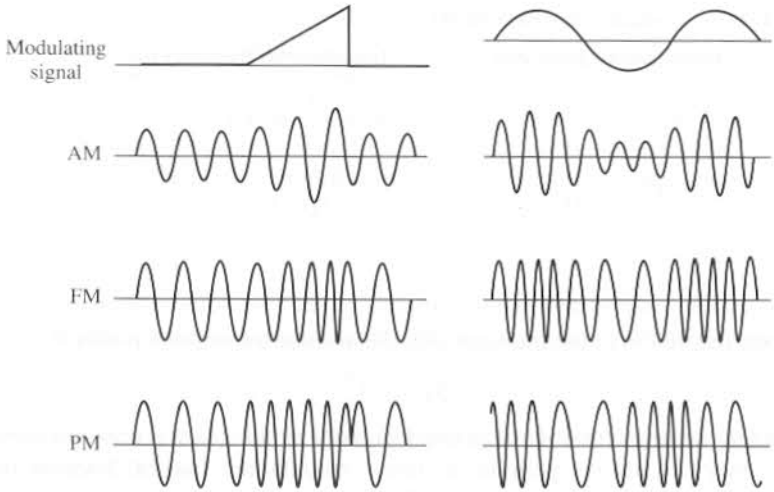

Figure 5.1–2 Illustrative AM, FM, and PM waveforms.

carrier amplitude, we modulate the frequency by swinging it over a range of, say, ± 50 Hz, then the transmission bandwidth will be 100 Hz regardless of the message bandwidth. As we'll soon see, this argument has a serious flaw, for it ignores the distinction between **instantaneous** and **spectral frequency.** Carson (1922) recognized the fallacy of the bandwidth-reduction notion and cleared the air on that score. Unfortunately, he and many others also felt that exponential modulation had no advantages over linear modulation with respect to noise. It took some time to overcome this belief but, thanks to Armstrong (1936), the merits of exponential modulation were finally appreciated. Before we can understand them quantitatively, we must address the problem of spectral analysis.

EXERCISE 5.1–1 Suppose FM had been defined in direct analogy to AM by writing $x_c(t) = A_c \cos \omega_c(t)\, t$ with $\omega_c(t) = \omega_c[1 + \mu x(t)]$. Demonstrate the physical impossibility of this definition by finding $f(t)$ when $x(t) = \cos \omega_m t$.

Narrowband PM and FM

Our spectral analysis of exponential modulation starts with the quadrature-carrier version of Eq. (1), namely

$$x_c(t) = x_{ci}(t) \cos \omega_c t - x_{cq}(t) \sin \omega_c t \qquad [9]$$

where

$$x_{ci}(t) = A_c \cos \phi(t) = A_c\left[1 - \frac{1}{2!}\phi^2(t) + \cdots\right] \qquad [10]$$