5

Frequency Modulation

Following the pattern set in Chapter 3, this chapter covers the theory of frequency modulation, and its generation. Both the theory and the generation of FM are a good deal more complex to think about and visualize than those of AM. This is mainly because FM involves minute frequency variations of the carrier, whereas AM results in large-scale amplitude variations of the carrier. FM is more difficult to determine mathematically and has sideband behavior that is equally complex.

Having studied this chapter, students will understand that FM is a form of angular modulation and that phase modulation is another similar form. The theory of both is discussed in detail, as are their similarity and important differences. It will be seen that frequency modulation is the preferred form for most applications. Frequency and amplitude modulation are then compared, on the basis that both are widely used practical systems.

Unlike amplitude modulation, FM is, or can be made, relatively immune to the effects of noise. This point is discussed at length. It will be seen that the effect of noise in FM depends on the noise sideband frequency, a point that is brought out under the heading of noise triangle. It will be shown that processing of the modulating signals, known as pre-emphasis and de-emphasis, plays an important part in making FM relatively immune to noise.

The final topic studied in this chapter is the generation of FM. It will be shown that two basic methods of generation exist. The first is direct generation, in which a voltage-dependent reactance varies the frequency of an oscillator. The second method is one in which basically phase modulation is generated, but circuitry is used to convert this to frequency modulation. Both methods are used in practice.

OBJECTIVES

Upon completing the material in Chapter 5, the student will be able to:

Understand the theory of frequency modulation (FM).

Draw an FM waveform.

Determine by calculation, the modulation index (MI).

Analyze the frequency spectrum using Bessel functions.

Understand the differences between AM, PM, and FM.

Explain the effect of noise on a frequency modulation wave.

Define and explain pre-emphasis and de-emphasis.

Understand the theory of stereo FM.

Identify the various methods of FM.

5-1 THEORY OF FREQUENCY AND PHASE MODULATION

Frequency modulation is a system in which the amplitude of the modulated carrier is kept constant, while its frequency and rate of change are varied by the modulating signal. The first practical system was put forward in 1936 as an alternative to AM in an effort to make radio transmissions more resistant to noise. *Phase modulation* is a similar system in which the phase of the carrier is varied instead of its frequency; as in FM, the amplitude of the carrier remains constant.

Let's assume for the moment that the carrier of the transmitter is at its resting frequency (no modulation) of 100 MHz and we apply a modulating signal. The amplitude of the modulating signal will cause the carrier to deviate (shift) from this resting frequency by a certain amount. If we increase the amplitude (loudness) of this signal (see Figure 5-16), we will increase the deviation to a maximum of 75 kHz as specified by the FCC. If we remove the modulation, the carrier frequency shifts back to its resting frequency (100 MHz).

We can see by this example that the deviation of the carrier is proportional to the amplitude of the modulating voltage. The shift in the carrier frequency from its resting point compared to the amplitude of the modulating voltage is called the *deviation ratio* (a deviation ratio of 5 is the maximum allowed in commercially broadcast FM).

The rate at which the carrier shifts from its resting point to a nonresting point is determined by the frequency of the modulating signal (the interaction between the amplitude and frequency of the modulating signal on the carrier is complex and requires the use of Bessel's functions to analyze the results).

If the modulating signal (AF) is 15 kHz at a certain amplitude and the carrier shift (because of the modulating voltage) is 75 kHz, the transmitter will produce *eight* significant sidebands (see Table 5-1). This is known as the *maximum deviation ratio:*

TABLE 5-1

MODULATION INDEX	SIDEBANDS
1	3
2	4
3	6
4	7
5	8 (maximum)

$$\text{Deviation ration} = \frac{f_{\text{dev}}\ (\text{max})}{f_{\text{AF}}\ (\text{max})}$$

If the frequency deviation of the carrier is known and the frequency of the modulating voltage (AF) is known, we can now establish the modulation index (MI):

$$\text{Modulation index} = \frac{f_{\text{dev}}}{f_{\text{AF}}}$$

Both these terms are important because of the bandwidth limitations placed on wideband FM transmitting stations by the regulating agencies throughout the world (this will be covered in detail later in this chapter).

5-1.1 Description of Systems

The general equation of an unmodulated wave, or carrier, may be written as

$$x = A \sin (\omega t + \phi) \tag{5-1}$$

where x = instantaneous value (of voltage or current)

A = (maximum) amplitude

ω = angular velocity, radians per second (rad/s)

ϕ = phase angle, rad

Note that ωt represents an angle in radians.

If any one of these three parameters is varied in accordance with another signal, normally of a lower frequency, then the second signal is called the *modulation,* and the first is said to be *modulated* by the second. Amplitude modulation, already discussed, is achieved when the amplitude A is varied. Alteration of the phase angle ϕ will yield phase modulation. If the frequency of the carrier is made to vary, frequency-modulated waves are obtained.

It is assumed that the modulating signal is sinusoidal. This signal has two important parameters which must be represented by the modulation process without distortion, specifically, its amplitude and frequency. It is understood that the phase relations of a complex modulation signal will be preserved. By the definition of frequency modulation, the amount by which the carrier frequency is varied from its unmodulated value, called the *deviation,* is made *proportional to* the *instantaneous amplitude* of the *modulating voltage.* The *rate* at which this frequency variation changes or takes place is equal to the modulating frequency.

The situation is illustrated in Figure 5-1, which shows the modulating voltage and the resulting frequency-modulated wave. Figure 5-1 also shows the frequency variation with time, which can be seen to be identical to the variation with time of the modulating voltage. The result of using that modulating voltage to produce AM is also shown for comparison. As in FM, all signals having the same amplitude will deviate the carrier frequency by the same amount, for example, 45 kHz, no matter what their frequencies. All signals of the same frequency, for example, 2 kHz, will deviate the

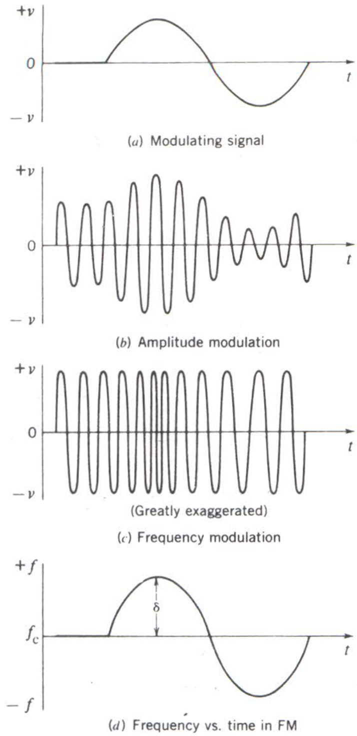

FIGURE 5-1 Basic modulation waveforms.

carrier at the same rate of 2000 times per second, no matter what their individual amplitudes. *The amplitude of the frequency-modulated wave remains constant at all times*. This is the greatest single advantage of FM.

5-1.2 Mathematical Representation of FM

From Figure 5-1*d*, it is seen that the instantaneous frequency f of the frequency-modulated wave is given by

$$f = f_c\,(1 + kV_m \cos \omega_m t) \tag{5-2}$$

where f_c = unmodulated (or average) carrier frequency
k = proportionality constant
$V_m \cos \omega_m t$ = instantaneous modulating voltage (cosine being preferred for simplicity in calculations)

The maximum deviation for this particular signal will occur when the cosine term has its maximum value, ± 1. Under these conditions, the instantaneous frequency will be

$$f = f_c\,(1 \pm kV_m) \tag{5-3}$$

so that the maximum deviation δ will be given by

$$\delta = kV_m f_c \tag{5-4}$$

The instantaneous amplitude of the FM signal will be given by a formula of the form

$$v = A \sin\,[F\,(\omega_c,\omega_m)] = A \sin \theta \tag{5-5}$$

where $F\,(\omega_c,\omega_m)$ is some function of the carrier and modulating frequencies. This function represents an angle and will be called θ for convenience. The problem now is to determine the instantaneous value (i.e., formula) for this angle.

As Figure 5-2 shows, θ is the angle traced out by the vector A in time t. If A were rotating with a constant angular velocity, for example, p, this angle θ would be given by pt (in radians). In this instance the angular velocity is anything but constant. It is governed by the formula for ω obtained from Equation (5-2), that is, $\omega = \omega_c\,(1 + kV_m \cos \omega_m t)$. In order to find θ, ω must be integrated with respect to time. Thus

$$\begin{aligned}
\theta &= \int \omega\, dt = \int \omega_c\,(1 + kV_m \cos \omega_m t)\, dt = \omega_c \int (1 + kV_m \cos \omega_m t)\, dt \\
&= \omega_c \left(t + \frac{kV_m \sin \omega_m t}{\omega_m}\right) = \omega_c t + \frac{kV_m \omega_c \sin \omega_m t}{\omega_m} \\
&= \omega_c t + \frac{kV_m f_c \sin \omega_m t}{f_m} \\
&= \omega_c t + \frac{\delta}{f_m} \sin \omega_m t
\end{aligned} \tag{5-6}$$

The derivation utilized, in turn, the fact that ω_c is constant, the formula $\int \cos nx\, dx = (\sin nx)/n$ and Equation (5-4), which had shown that $kV_m f_c = \delta$. Equation (5-6) may now be substituted into Equation (5-5) to give the instantaneous value of the FM voltage; therefore

$$v = A \sin \left(\omega_c t + \frac{\delta}{f_m} \sin \omega_m t\right) \tag{5-7}$$

The modulation index for FM, m_f, is defined as

$$m_f = \frac{\text{(maximum) frequency deviation}}{\text{modulating frequency}} = \frac{\delta}{f_m} \tag{5-8}$$

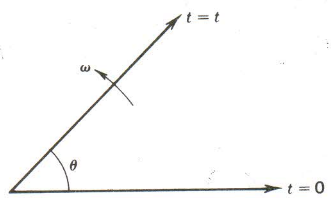

FIGURE 5-2 Frequency-modulated vectors.

Substituting Equation (5-8) into (5-7), we obtain

$$v = A \sin (\omega_c t + m_f \sin \omega_m t) \tag{5-9}$$

It is important to note that as the modulating frequency decreases and the modulating voltage amplitude (δ) remains constant, the modulation index increases. This will be the basis for distinguishing frequency modulation from phase modulation. Note that m_f, which is the ratio of two frequencies, is measured in radians.

EXAMPLE 5-1 In an FM system, when the audio frequency (AF) is 500 Hz and the AF voltage is 2.4 V, the deviation is 4.8 kHz. If the AF voltage is now increased to 7.2 V, what is the new deviation? If the AF voltage is raised to 10 V while the AF is dropped to 200 Hz, what is the deviation? Find the modulation index in each case.

SOLUTION

As $\delta \propto V_m$, we may write

$$\frac{\delta}{V} = \frac{4.8}{2.4} = 2 \text{ kHz/V} \qquad \text{(of modulating signal)}$$

When $V_m = 7.2$ V,

$$\delta = 2 \times 7.2 = 14.4 \text{ kHz}$$

Similarly, when $V_m = 10$V,

$$\delta = 2 \times 10 = 20 \text{ kHz}$$

Note that the change in modulating frequency made no difference to the deviation since it is independent of the modulating frequency. Calculation of the modulation indices gives

$$m_{f1} = \frac{\delta_1}{f_{m1}} = \frac{4.8}{0.5} = 9.6$$

$$m_{f2} = \frac{\delta_2}{f_{m1}} = \frac{14.4}{0.5} = 28.8$$

$$m_{f3} = \frac{\delta_3}{f_{m2}} = \frac{20}{0.2} = 100$$

The modulating frequency change did have to be taken into account in the modulation index calculation.

EXAMPLE 5-2 Find the carrier and modulating frequencies, the modulation index, and the maximum deviation of the FM wave represented by the voltage equation $v = 12 \sin (6 \times 10^8 t + 5 \sin 1250t)$. What power will this FM wave dissipate in a 10-Ω resistor?

SOLUTION

Comparing the numerical equation with Equation (5-9), we have

$$f_c = \frac{6 \times 10^8}{2\pi} = 95.5 \text{ MHz} \quad f_m = \frac{1250}{2\pi} = 199 \text{ Hz}$$

$m_f = 5$ (as given)

$\delta = m_f f_m = 5 \times 199 = 995$ Hz

$$P = \frac{V_{rms}^2}{R} = \frac{(12/\sqrt{2})^2}{10} = \frac{72}{10} = 7.2 \text{ W}$$

5-1.3 Frequency Spectrum of the FM Wave

When a comparable stage was reached with AM theory, i.e., when Equation (3-7) had been derived, it was possible to tell at a glance what frequencies were present in the modulated wave. Unfortunately, the situation is far more complex, mathematically speaking, for FM. Since Equation (5-9) is the sine of a sine, the only solution involves the use of *Bessel functions*. Using these, it may then be shown that Equation (5-9) may be expanded to yield

$$\begin{aligned} v = A\{&J_0(m_f) \sin \omega_c t \\ &+ J_1(m_f) [\sin (\omega_c + \omega_m)t - \sin (\omega_c - \omega_m)t] \\ &+ J_2(m_f) [\sin (\omega_c + 2\omega_m)t + \sin (\omega_c - 2\omega_m)t] \\ &+ J_3(m_f) [\sin (\omega_c + 3\omega_m)t - \sin (\omega_c - 3\omega_m)t] \\ &+ J_4(m_f) [\sin (\omega_c + 4\omega_m)t + \sin (\omega_c - 4\omega_m)t] \cdots\} \end{aligned} \tag{5-10}$$

It can be shown that the output consists of a carrier and an apparently infinite number of pairs of sidebands, each preceded by J coefficients. These are Bessel functions. Here they happen to be of the first kind and of the order denoted by the subscript, with the argument m_f. $J_n(m_f)$ may be shown to be a solution of an equation of the form

$$(m_f)^2 \frac{d^2y}{dm_f^2} + m_f \frac{dy}{dm_f} + (m_f^2 - n^2)y = 0 \tag{5-11}$$

This solution, i.e., the formula for the Bessel function, is

$$J_n(m_f) = \left(\frac{m_f}{2}\right)^n \left[\frac{1}{n!} - \frac{(m_f/2)^2}{1!(n+1)!} + \frac{(m_f/2)^4}{2!(n+2)!} - \frac{(m_f/2)^6}{3!(n+1)!} + \cdots\right] \tag{5-12}$$

In order to evaluate the value of a given pair of sidebands or the value of the carrier, it is necessary to know the value of the corresponding Bessel function. Separate calculation from Equation (5-12) for each case is not required since information of this type is freely available in table form, as in Table 5-2, or graphical form, as in Figure 5-3.

Observations The mathematics of the previous discussion may be reviewed in a series of observations as follows:

1. Unlike AM, where there are only three frequencies (the carrier and the first two sidebands), *FM has an infinite number of sidebands,* as well as the carrier. They are separated from the carrier by f_m, $2f_m$, $3f_m$, . . . , and thus have a recurrence frequency of f_m.

TABLE 5-2 Bessel Functions of the First Kind

x	*n* or Order																
(m_f)	J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}	J_{11}	J_{12}	J_{13}	J_{14}	J_{15}	J_{16}
0.00	1.00	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.25	0.98	0.12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.5	0.94	0.24	0.03	—	—	—	—	—	—	—	—	—	—	—	—	—	—
1.0	0.77	0.44	0.11	0.02	—	—	—	—	—	—	—	—	—	—	—	—	—
1.5	0.51	0.56	0.23	0.06	0.01	—	—	—	—	—	—	—	—	—	—	—	—
2.0	0.22	0.58	0.35	0.13	0.03	—	—	—	—	—	—	—	—	—	—	—	—
2.5	−0.05	0.50	0.45	0.22	0.07	0.02	—	—	—	—	—	—	—	—	—	—	—
3.0	−0.26	0.34	0.49	0.31	0.13	0.04	0.01	—	—	—	—	—	—	—	—	—	—
4.0	−0.40	−0.07	0.36	0.43	0.28	0.13	0.05	0.02	—	—	—	—	—	—	—	—	—
5.0	−0.18	−0.33	0.05	0.36	0.39	0.26	0.13	0.05	0.02	—	—	—	—	—	—	—	—
6.0	0.15	−0.28	−0.24	0.11	0.36	0.36	0.25	0.13	0.06	0.02	—	—	—	—	—	—	—
7.0	0.30	0.00	−0.30	−0.17	0.16	0.35	0.34	0.23	0.13	0.06	0.02	—	—	—	—	—	—
8.0	0.17	0.23	−0.11	−0.29	−0.10	0.19	0.34	0.32	0.22	0.13	0.06	0.03	—	—	—	—	—
9.0	−0.09	0.24	0.14	−0.18	−0.27	−0.06	0.20	0.33	0.30	0.21	0.12	0.06	0.03	0.01	—	—	—
10.0	−0.25	0.04	0.25	0.06	−0.22	−0.23	−0.01	0.22	0.31	0.29	0.20	0.12	0.06	0.03	0.01	—	—
12.0	0.05	−0.22	−0.08	0.20	0.18	−0.07	−0.24	−0.17	0.05	0.23	0.30	0.27	0.20	0.12	0.07	0.03	0.01
15.0	−0.01	0.21	0.04	−0.19	−0.12	0.13	0.21	0.03	−0.17	−0.22	−0.09	0.10	0.24	0.28	0.25	0.18	0.12

2. The J coefficients eventually decrease in value as n increases, but not in any simple manner. As seen in Figure 5-3, the value fluctuates on either side of zero, gradually diminishing. Since each J coefficient represents the amplitude of a particular pair of sidebands, these also eventually decrease, but only past a certain value of n. *The modulation index determines how many sideband components have significant amplitudes.*
3. The sidebands at equal distances from f_c have equal amplitudes, so that the sideband distribution is symmetrical about the carrier frequency. The J coefficients occasionally have negative values, signifying a 180° phase change for that particular pair of sidebands.

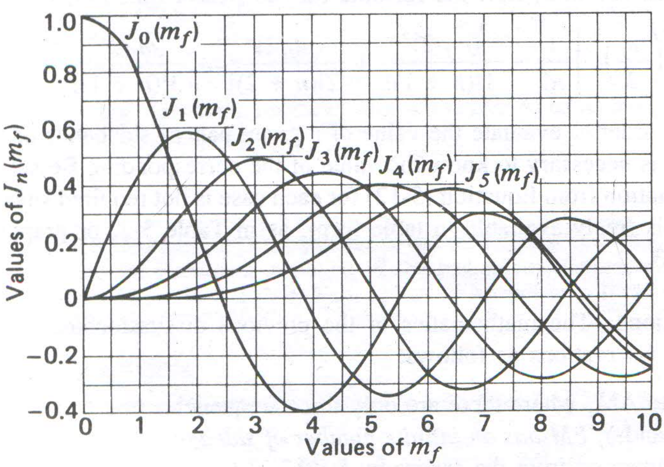

FIGURE 5-3 Bessel functions.

4. Looking down Table 5-2, as m_f increases, so does the value of a particular J coefficient, such as J_{12}. Bearing in mind that m_f is inversely proportional to the modulating frequency, we see that the relative amplitude of distant sidebands increases when the modulation frequency is lowered. The previous statement assumes that deviation (i.e., the modulating voltage) has remained constant.
5. In AM, increased depth of modulation increases the sideband power and therefore the total transmitted power. In FM, the total transmitted power always remains constant, but with increased depth of modulation the required bandwidth is increased. To be quite specific, what increases is the bandwidth required to transmit a relatively undistorted signal. This is true because increased depth of modulation means increased deviation, and therefore an increased modulation index, so that more distant sidebands acquire significant amplitudes.
6. As evidenced by Equation (5-10), the theoretical bandwidth required in FM is infinite. In practice, the bandwidth used is one that has been calculated to allow for all significant amplitudes of sideband components under the most exacting conditions. This really means ensuring that, with maximum deviation by the highest modulating frequency, no significant sideband components are lopped off.
7. In FM, unlike in AM, the amplitude of the carrier component does not remain constant. Its J coefficient is J_0, which is a function of m_f. This may sound somewhat confusing but keeping the overall amplitude of the FM wave constant would be very difficult if the amplitude of the carrier were not reduced when the amplitude of the various sidebands increased.
8. It is possible for the carrier component of the FM wave to disappear completely. This happens for certain values of the modulation index, called *eigenvalues*. Figure 5-3 shows that these are approximately 2.4, 5.5, 8.6, 11.8, and so on. These disappearances of the carrier for specific values of m_f form a handy basis for measuring deviation.

Bandwidth and required spectra Using Table 5-2, it is possible to evaluate the size of the carrier and each sideband for each specific or value of the modulation index. When this is done, the frequency spectrum of the FM wave for that particular value of m_f may be plotted. This is done in Figure 5-4, which shows these spectrograms first for increasing deviation (f_m constant), and then for decreasing modulating frequency (δ constant). Both the table and the spectrograms illustrate the observations, especially points 2, 3, 4, and 5. It can be seen that as modulation depth increases, so does bandwidth (Figure 5-4*a*), and also that reduction in modulation frequency increases the number of sidebands, though not necessarily the bandwidth (Figure 5-4*b*). Another point shown very clearly is that although the number of sideband components is theoretically infinite, in practice a lot of the higher sidebands have insignificant relative amplitudes, and this is why they are not shown in the spectrograms. Their exclusion in a practical system will not distort the modulated wave unduly.

In order to calculate the required bandwidth accurately, the student need only look at the table to see which is the last J coefficient shown for that value of modulation index.

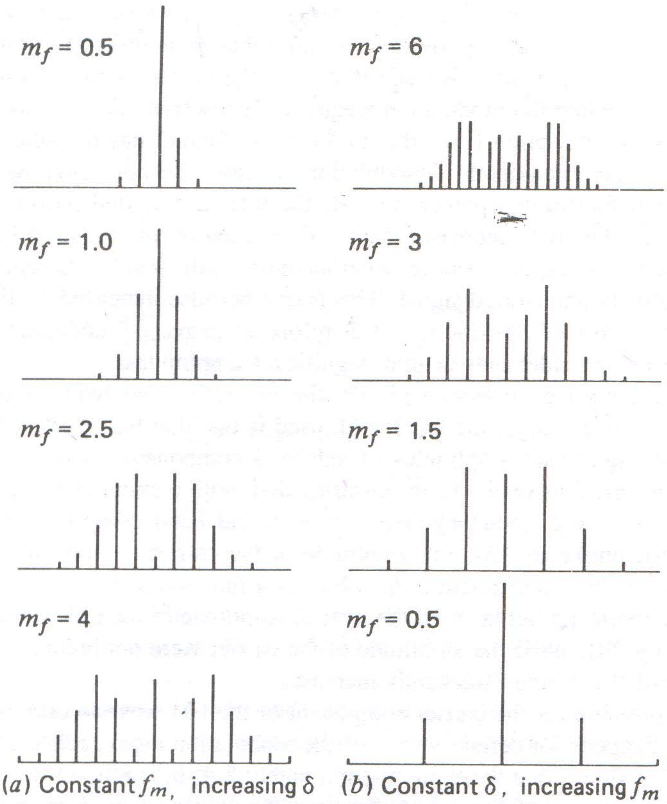

FIGURE 5-4 FM spectrograms. ***(After K. R. Sturley, Frequency-Modulated Radio, 2d ed., George Newnes Ltd., London, 1958, by permission of the publisher.)***

EXAMPLE 5-3 What is the bandwidth required for an FM signal in which the modulating frequency is 2 kHz and the maximum deviation is 10 kHz?

SOLUTION

$$m_f = \frac{\delta}{f_m} = \frac{10}{2} = 5$$

From Table 5-1, it is seen that the highest J coefficient included for this value of m_f is J_8. This means that all higher values of Bessel functions for that modulation index have values less than 0.01 and may therefore be ignored. *The eighth pair of sidebands is the furthest from the carrier to be included in this instance.* This gives

$$\Delta = f_m \times \text{highest needed sideband} \times 2$$
$$= 2 \text{ kHz} \times 8 \times 2 = 32 \text{ kHz}$$

A rule of thumb (Carson's rule) states that (as a good approximation) the bandwidth required to pass an FM wave is twice the sum of the deviation and the highest modulating frequency, but it must be remembered that this is only an approximation. Actually, it does give a fairly accurate result if the modulation index is in excess of about 6.

5-1.4 Phase Modulation

Strictly speaking, there are two types of continuous-wave modulation; amplitude modulation and angle modulation. Angle modulation may be subdivided into two distinct types; frequency modulation and phase modulation (PM). Thus, PM and FM are closely allied, and this is the first reason for considering PM here. The second reason is somewhat more practical. It is possible to obtain frequency modulation from phase modulation by the so-called *Armstrong system*. Phase modulation is not used in practical analog transmission systems.

If the phase ϕ in the equation $v = A \sin (\omega_c t + \phi)$ is varied so that its magnitude is proportional to the instantaneous amplitude of the modulating voltage, the resulting wave is phase-modulated. The expression for a PM wave is

$$v = A \sin (\omega_c t + \phi_m \sin \omega_m t) \tag{5-13}$$

where ϕ_m is the maximum value of phase change introduced by this particular modulating signal and is proportional to the maximum amplitude of this modulation. For the sake of uniformity, this is rewritten as

$$v = A \sin (\omega_c t + m_p \sin \omega_m t) \tag{5-14}$$

where $m_p = \phi_m$ = modulation index for phase modulation.

To visualize phase modulation, consider a horizontal metronome or pendulum placed on a rotating record turntable. As well as rotating, the arm of this metronome is swinging sinusoidally back and forth about its mean point. If the maximum displacement of this swing can be made proportional to the size of the "push" applied to the metronome, and if the frequency of swing can be made equal to the number of "pushes" per second, then the motion of the arm is *exactly* the same as that of a phase-modulated vector. Actually, PM seems easier to visualize than FM.

Equation (5-13) was obtained directly, without recourse to the derivation required for the corresponding expression for FM, Equation (5-9). This occurs because in FM an equation for the *angular velocity* was postulated, from which the phase angle for $v = A \sin (\theta)$ had to be derived, whereas in PM the phase relationship is defined and may be substituted directly. Comparison of Equations (5-14) and (5-9) shows them to be identical, except for the different definitions of the modulation index. It is obvious that these two forms of angle modulation are indeed similar. They will now be compared and contrasted.

5-1.5 Intersystem Comparisons

Frequency and phase modulation From the purely theoretical point of view, the difference between FM and PM is quite simple—the modulation index is defined differently in each system. However, this is not nearly as obvious as the difference