

Handout 7: Phase and Frequency Modulation

Instructor: Wing-Kin Ma

Suggested Reading: Chapter 4 of Simon Haykin and Michael Moher, *Communication Systems (5th Edition)*, Wiley & Sons Ltd; or Chapter 5 of B. P. Lathi and Z. Ding, *Modern Digital and Analog Communication Systems (4th Edition)*, Oxford University Press.

1 Basic Concepts

We start out with introducing a general carrier-modulation approach called *angle modulation*. An angle-modulated wave is expressed as

$$s(t) = A_c \cos[\theta_i(t)] \quad (1)$$

where A_c is the carrier amplitude, $\theta_i(t)$ is called the *angle* of the modulated wave. In angle modulation, the angle $\theta_i(t)$ is used to carry information.

There is an important concept we must understand here—that of instantaneous frequency. The instantaneous frequency of the angle-modulated signal $s(t)$ is defined as

$$\begin{aligned} f_i(t) &= \lim_{\Delta t \rightarrow 0} \frac{\theta_i(t + \Delta t) - \theta_i(t)}{2\pi \Delta t} \\ &= \frac{1}{2\pi} \frac{d\theta_i(t)}{dt}. \end{aligned} \quad (2)$$

To get some insight, consider an example where the angle is given by

$$\theta_i(t) = 2\pi f_c t + \phi_c,$$

for some frequency f_c and phase ϕ_c ; i.e., the corresponding angle-modulated signal $s(t)$ is a pure sinusoidal wave. From (2), the instantaneous frequency is obtained as $f_i(t) = f_c$.

There are many different ways one can put information in $\theta_i(t)$. Here, we consider two representative schemes, namely, *phase modulation* and *frequency modulation*.

Phase Modulation (PM): PM has the angle taking the form

$$\theta_i(t) = 2\pi f_c t + k_p m(t), \quad (3)$$

where f_c is the carrier frequency, $m(t)$ is the message signal, and k_p is a constant and is called the *phase sensitivity* of the modulator. In words, the angle is used to carry information in a direct and linear manner. The phase-modulated signal is thus described by

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)]. \quad (4)$$

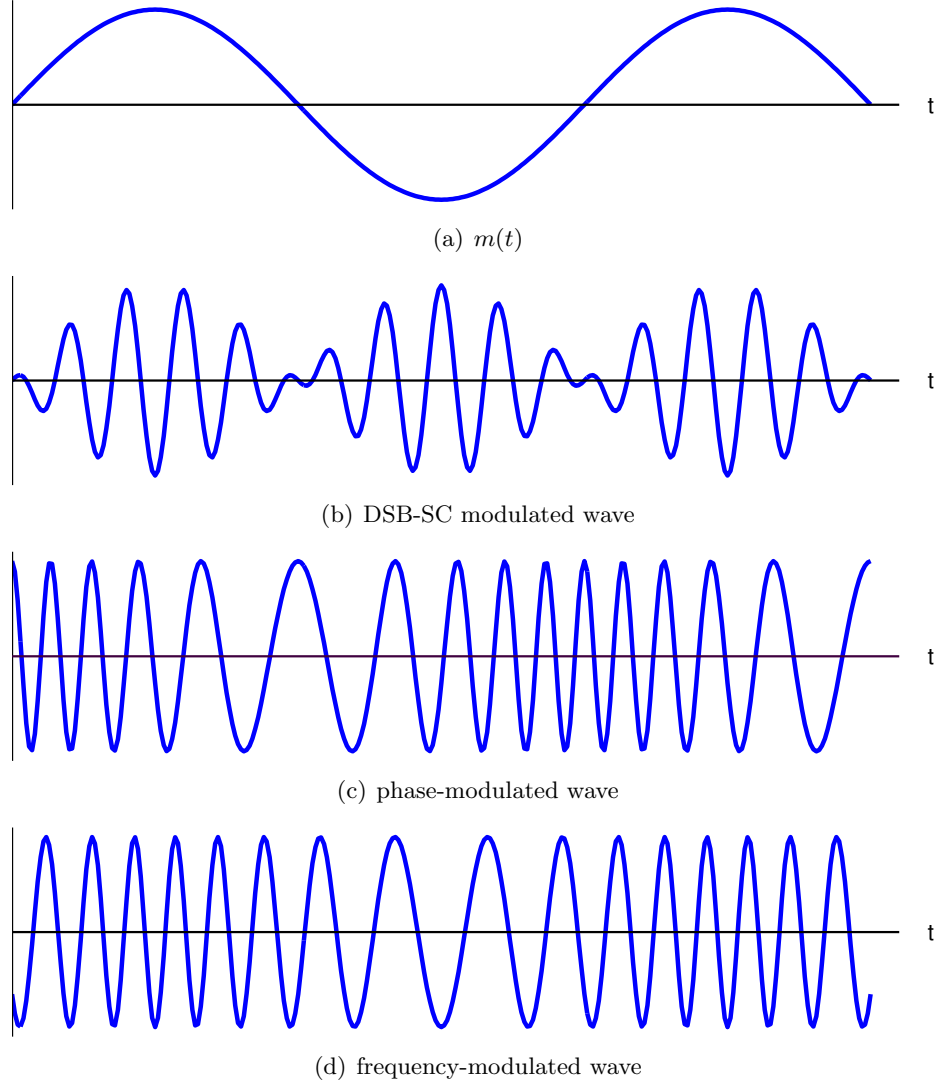


Figure 1: An illustration of phase-modulated and frequency-modulated signals.

Frequency Modulation (FM): The principle of FM is to use the instantaneous frequency to carry information in a linear manner. Specifically, FM aims at having the instantaneous frequency taking the form

$$f_i(t) = f_c + k_f m(t), \quad (5)$$

where f_c and $m(t)$ are again the carrier frequency and message signal, respectively; k_f is a constant and is called the *frequency sensitivity* of the modulator. Eq. (5) may be achieved by choosing the angle as

$$\theta_i(t) = 2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau, \quad (6)$$

where we should note that in the second term of the right-hand side of (6), indefinite integrals is

employed. By substituting (6) into (2), the key FM expression in (5) is shown to be satisfied. The frequency-modulated signal can be written as

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau \right]. \quad (7)$$

Figure 1 gives an illustration of the phase-modulated and frequency-modulated signals, wherein we also show the DBS-SC modulated signal for the purpose of comparison. From the figure we see that the PM and FM signals both exhibit constant envelope. In fact, the transmitted power of PM and FM signals is constant for a given A_c , and does not vary with the message signal. Moreover, the FM wave appears to vary faster when the value of the message signal $m(t)$ is large, and slower when the value of the message signal $m(t)$ is small.

It is interesting to compare the AM-based and angle modulation schemes. AM-based schemes have non-constant envelope, while angle modulation always has constant envelope. In AM-based schemes, the modulated signal $s(t)$ usually has a linear relationship with respect to (w.r.t.) the message signal $m(t)$. This enables us to use simple Fourier transform properties to derive the Fourier transforms of the modulated signals and perform spectral analysis, e.g., determining the transmission bandwidth. Angle modulation schemes has a nonlinear relationship between $s(t)$ and $m(t)$. The nonlinear nature of angle modulation makes spectral analysis difficult.

2 A Case Study Based on a Single-Tone Message Signal

Analyzing how FM or PM works for a general message signal can be both difficult and complicated. A logical engineering approach would therefore be to narrow down attention to special cases, thereby attempting to extract useful insights from a simpler problem setting. Specifically, we consider the special case of single-tone modulating signals

$$m(t) = A_m \cos(2\pi f_m t),$$

where f_m denotes the frequency of the tone, and A_m is the tone amplitude. Also, we are interested in FM only, although it is possible to extend the study to PM. The instantaneous frequency in this special case is given by

$$\begin{aligned} f_i(t) &= f_c + k_f A_m \cos(2\pi f_m t) \\ &= f_c + \Delta f \cos(2\pi f_m t), \end{aligned} \quad (8)$$

where

$$\Delta f = k_f A_m \quad (9)$$

is called the *frequency deviation*. The frequency deviation describes the maximum difference of the instantaneous frequency $f_i(t)$ and the carrier frequency f_c . By (6), the angle of the FM signal is obtained as

$$\begin{aligned} \theta_i(t) &= 2\pi f_c t + 2\pi k_f \int_{-\infty}^t A_m \cos(2\pi f_m \tau) d\tau \\ &= 2\pi f_c t + \frac{\Delta f}{f_m} \sin(2\pi f_m t). \end{aligned} \quad (10)$$

Let us denote

$$\beta = \frac{\Delta f}{f_m}. \quad (11)$$

The quantity β is called the *modulation index* of the FM signal. Consequently, we can write

$$\theta_i(t) = 2\pi f_c t + \beta \sin(2\pi f_m t), \quad (12)$$

and hence the corresponding FM signal is given by

$$s(t) = A_c \cos[2\pi f_c t + \beta \sin(2\pi f_m t)]. \quad (13)$$

We examine two cases, namely, narrowband FM and wideband FM, respectively.

2.1 Narrowband FM

Suppose that the modulation index β is small compared to one radian. To see how the FM signal behaves, we first apply basic trigonometry results to rewrite (13) as

$$s(t) = A_c \cos[\beta \sin(2\pi f_m t)] \cos(2\pi f_c t) - A_c \sin[\beta \sin(2\pi f_m t)] \sin(2\pi f_c t). \quad (14)$$

Next, we apply approximations, namely, that for $|\alpha| \ll 1$ we have $\sin(\alpha) \approx \alpha$ and $\cos(\alpha) \approx 1$. By using the above mentioned approximations, we may simplify (14) to

$$s(t) \approx A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_m t) \sin(2\pi f_c t) \quad (15)$$

when β is small. Now, we argue that (15) is *AM-like*. To elaborate on this, the AM signal under the same message signal is

$$\begin{aligned} s_{\text{AM}}(t) &= A_c [1 + k_a A_m \cos(2\pi f_m t)] \cos(2\pi f_c t) \\ &= A_c \cos(2\pi f_c t) + \mu A_c \cos(2\pi f_m t) \cos(2\pi f_c t), \end{aligned} \quad (16)$$

where $\mu = k_a A_m$ is the modulation index of the AM signal. Comparing (15)-(16), we observe this: If we replace $-\sin(2\pi f_m t)$ in (15) by $\cos(2\pi f_m t)$, replace $\sin(2\pi f_c t)$ in (15) by $\cos(2\pi f_c t)$, and set $\beta = \mu$, then the resulting modulated signal is identical to that in (16). In fact, it can be easily shown that under the setting $\beta = \mu$, Eqs. (15) and (16) have the same amplitude spectrum.

The observation suggests that for small β , the FM signal takes a form similar to the AM signal. This case is commonly referred to as *narrowband FM*. By the AM-like nature of FM signals in the narrowband FM case, we conclude that the transmission bandwidth of narrowband FM is $2f_m$ Hz.

2.2 Wideband FM

Suppose that the modulation index β is large compared to one radian. This case is called *wideband FM*, and is a (much) more interesting case. To study the wideband FM case, let us rewrite (13) as

$$\begin{aligned} s(t) &= \text{Re}[A_c e^{j2\pi f_c t + j\beta \sin(2\pi f_m t)}] \\ &= \text{Re}[\tilde{s}(t) e^{j2\pi f_c t}], \end{aligned} \quad (17)$$

where we denote

$$\tilde{s}(t) = A_c e^{j\beta \sin(2\pi f_m t)}. \quad (18)$$

Our attention is turned to analysis of $\tilde{s}(t)$. Since $\sin(2\pi f_m t)$ is a periodic signal with period $T_0 = 1/f_m$, $\tilde{s}(t)$ is also a periodic signal with the same period. As such, we can apply the Fourier series expansion on $\tilde{s}(t)$

$$\tilde{s}(t) = \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_m t}, \quad (19)$$

where the Fourier coefficients c_n are to be determined. The coefficients c_n are given by

$$\begin{aligned} c_n &= \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} \tilde{s}(t) e^{-j\frac{2\pi n t}{T_0}} dt \\ &= f_m A_c \int_{-\frac{1}{2f_m}}^{\frac{1}{2f_m}} e^{j[\beta \sin(2\pi f_m t) - 2\pi n f_m t]} dt. \end{aligned} \quad (20)$$

By the change of variable $x = 2\pi f_m t$, we can rewrite (20) as

$$c_n = \frac{A_c}{2\pi} \int_{-\pi}^{\pi} e^{j[\beta \sin(x) - nx]} dx. \quad (21)$$

At this point, we should mention that the integral in (21) does not have an explicit expression—at least not in a simple manner. The integral in (21) is known as the *n*th order Bessel function of the first kind, which is commonly denoted as

$$J_n(\beta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j[\beta \sin(x) - nx]} dx. \quad (22)$$

While the Bessel function of the first kind does not have a simple closed form, its values can be evaluated numerically by computers¹. Note that $J_n(\beta)$ can be shown to be real-valued, and that $J_n(\beta) = J_{-n}(\beta)$. Figure 2 shows numerically computed values of $J_n(\beta)$ w.r.t. n . By substituting (19), (21) and (22) into (17), we obtain

$$s(t) = A_c \cdot \text{Re} \left[\sum_{n=-\infty}^{\infty} J_n(\beta) e^{j2\pi(f_c + n f_m)t} \right] \quad (23)$$

$$= A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos[2\pi(f_c + n f_m)t]. \quad (24)$$

With the FM signal expression in (25), we are now ready to examine the spectrum of the FM signal. The Fourier transform of (25) is

$$S(f) = \frac{A_c}{2} \sum_{n=-\infty}^{\infty} J_n(\beta) [\delta(f - f_c - n f_m) + \delta(f + f_c + n f_m)]. \quad (25)$$

In Figure 3 we illustrate the corresponding amplitude spectra for various values of β . It can be seen that the spectral content, in terms of significant components, occupies a larger frequency band as the modulation index β increases.

It is important to point out that the FM spectrum in (25) is *not* bandlimited, since it is composed of an infinite number of frequency components. In fact, it can be shown that for an arbitrary

¹For example, in MATLAB, we use the command `besselj` to obtain the values of $J_n(\beta)$.

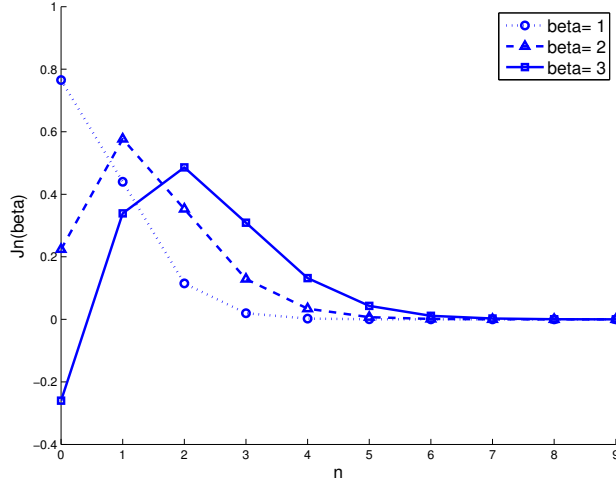


Figure 2: A plot of the Bessel function of the first kind.

message signal $m(t)$, the Fourier transform of the corresponding FM signal is not bandlimited in general (the proof is more difficult than that of the single tone case above). By the same spirit as the discussion of bandwidth in Handout 3, we may evaluate the FM transmission bandwidth by finding a frequency interval within which a significant portion of the frequency components lies. For the single tone case, a common approximate rule of the FM transmission bandwidth is

$$B_T \simeq 2\Delta f + 2f_m = 2f_m (\beta + 1). \quad (26)$$

We may at least observe from Figure 3 that the bandwidth estimate in (26) appears to be reasonable.

3 Carson's Rule

Carson's rule is a rule for approximating the FM transmission bandwidth of a *general* message signal. The result is somehow reminiscent of the bandwidth estimate for the single tone case in (26), although the proof is much more complex than that of the latter and takes a different set of assumptions to obtain the bandwidth formula. We state Carson's rule without giving the proof (see the Lathi-Ding textbook for a description of how Carson's rule is proven). Let

$$\Delta f = k_f \max_t |m(t)| \quad (27)$$

be the frequency deviation of a general message signal $m(t)$. Also, let

$$D = \frac{\Delta f}{W}, \quad (28)$$

which is called the *deviation ratio*. Note that the deviation ratio is somehow similar to the modulation index β in the single tone case. The approximate FM transmission bandwidth by Carson's rule is

$$B_T \simeq 2\Delta f + 2W = 2W(D + 1). \quad (29)$$

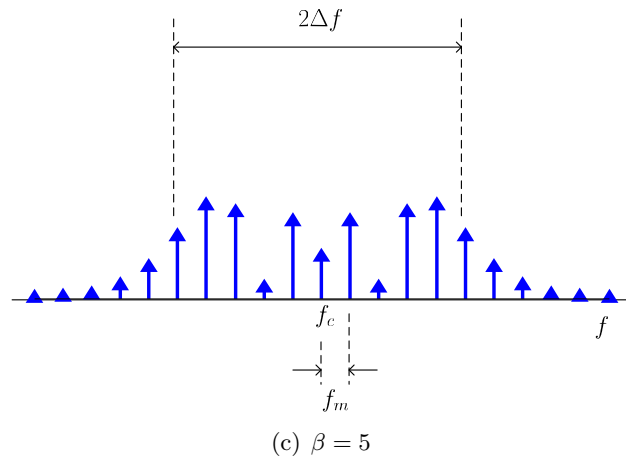
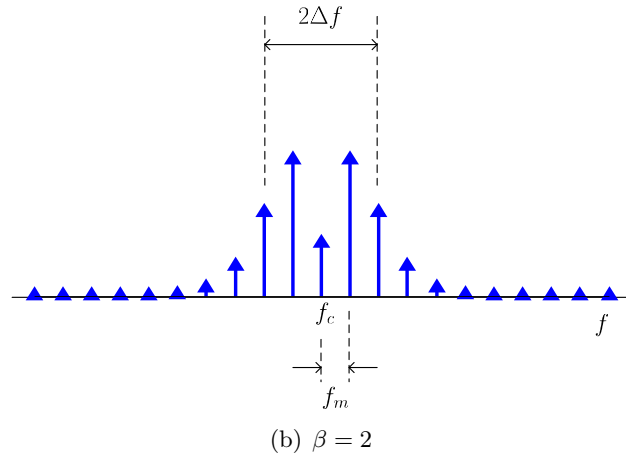
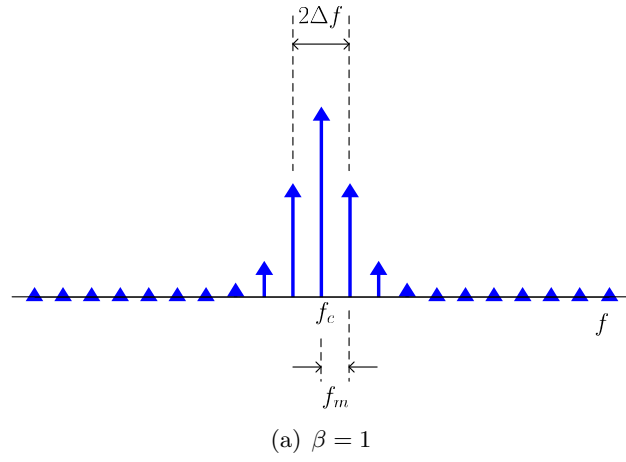


Figure 3: Amplitude spectra of FM signals of a single-tone message signal.