```
end
Hq=Hq';
Yq=Gq.*Hq;
yk=ifft(Yq);
clf,stem(k,yk)
```

REFERENCES

1. R. V. Churchill and J. W. Brown, *Fourier Series and Boundary Value Problems*, 3rd ed., McGraw-Hill, New York, 1978.
2. R. N. Bracewell, *Fourier Transform and Its Applications*, rev. 2nd ed., McGraw-Hill, New York, 1986.
3. B. P. Lathi, *Signal Processing and Linear Systems*, Oxford University Press, 2000.
4. E. A. Guillemin, *Theory of Linear Physical Systems*, Wiley, New York, 1963.
5. F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proc. IEEE*, vol. 66, pp. 51–83, Jan. 1978.
6. J. W. Tukey and J. Cooley, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Mathematics of Computation*, Vol. 19, pp. 297–301, April 1965.

PROBLEMS

3.1-1 Show that the Fourier transform of $g(t)$ may be expressed as

$$G(f) = \int_{-\infty}^{\infty} g(t) \cos 2\pi ft \, dt - j \int_{-\infty}^{\infty} g(t) \sin 2\pi ft \, dt$$

Hence, show that if $g(t)$ is an even function of t, then

$$G(f) = 2 \int_{0}^{\infty} g(t) \cos 2\pi ft \, dt$$

and if $g(t)$ is an odd function of t, then

$$G(f) = -2j \int_{0}^{\infty} g(t) \sin 2\pi ft \, dt$$

Hence, prove that the following.

If $g(t)$ is:	*Then $G(f)$ is:*
a real and even function of t	a real and even function of f
a real and odd function of t	an imaginary and odd function of f
an imaginary and even function of t	an imaginary and even function of f
a complex and even function of t	a complex and even function of f
a complex and odd function of t	a complex and odd function of f

3.1-2 **(a)** Show that for a real $g(t)$, the inverse transform, Eq. (3.9b), can be expressed as

$$g(t) = 2 \int_{0}^{\infty} |G(f)| \cos[2\pi ft + \theta_g(2\pi f)] \, df$$

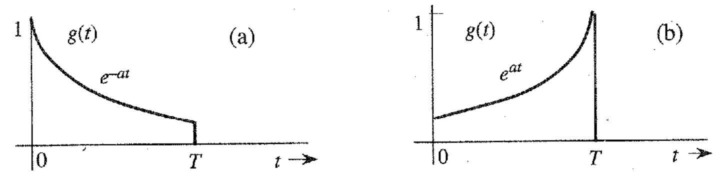
This is the trigonometric form of the (inverse) Fourier transform.

(b) Express the Fourier integral (inverse Fourier transform) for $g(t) = e^{-at}u(t)$ in the trigonometric form given in part **(a)**.

3.1-3 If $g(t) \Longleftrightarrow G(f)$, then show that $g^*(t) \Longleftrightarrow G^*(-f)$.

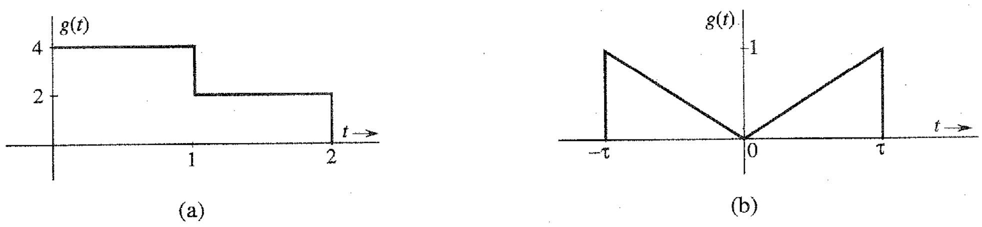
3.1-4 From definition (3.9a), find the Fourier transforms of the signals shown in Fig. P3.1-4.

Figure P.3.1-4



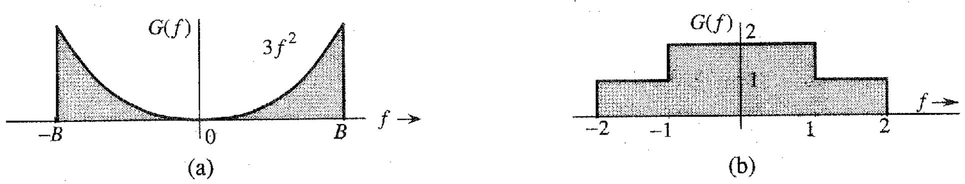
3.1-5 From definition (3.9a), find the Fourier transforms of the signals shown in Fig. P3.1-5.

Figure P.3.1-5



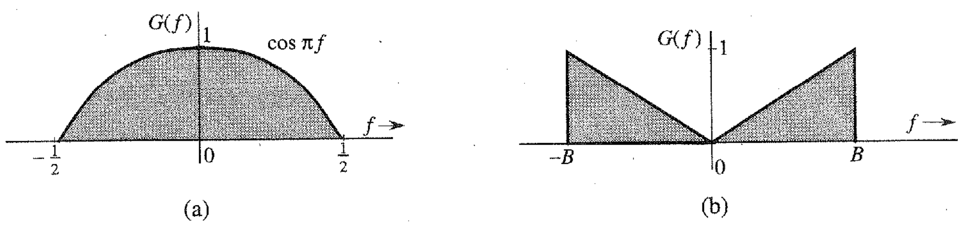
3.1-6 From definition (3.9b), find the inverse Fourier transforms of the spectra shown in Fig. P3.1-6.

Figure P.3.1-6



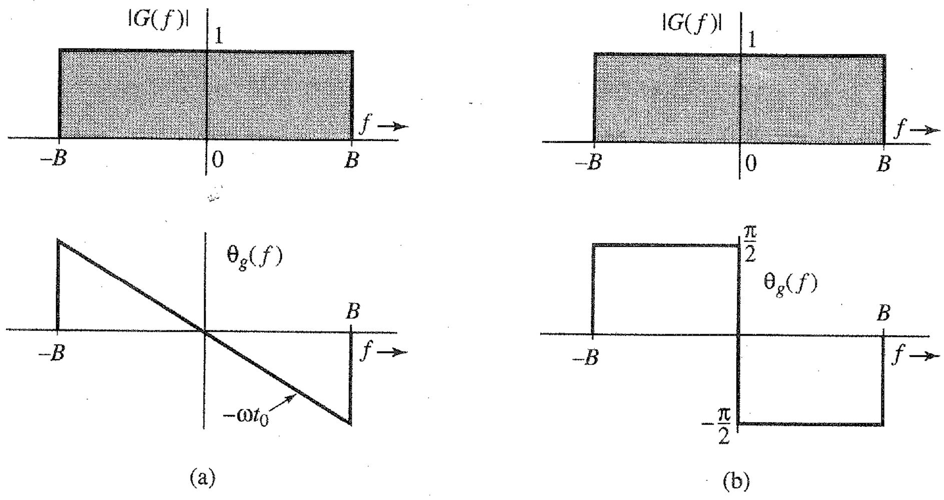
3.1-7 From definition (3.9b), find the inverse Fourier transforms of the spectra shown in Fig. P3.1-7.

Figure P.3.1-7



3.1-8 Show that the two signals in parts **(a)** and **(b)** of Fig. P3.1-8 are totally different in the time domain, despite their similarity.

Figure P.3.1-8



Hint: $G(f) = |G(f)|e^{j\theta_g(f)}$. For part **(a)**, $G(f) = 1 \cdot e^{-j2\pi f t_0}, \quad |f| \le B$, whereas for part **(b)**,

$$G(f) = \begin{cases} 1e^{-j\pi/2} = -j & 0 < f \le B \\ 1e^{j\pi/2} = j & 0 > f \ge -B \end{cases}$$

3.2-1 Sketch the following functions: **(a)** $\Pi(t/2)$; **(b)** $\Delta(3\omega/100)$; **(c)** $\Pi(t-10/8)$; **(d)** sinc $(\pi\omega//5)$; **(e)** sinc $[(\omega - 10\pi)/5]$; **(f)** sinc $(t/5)\,\Pi(t/10\pi)$.

Hint: $g(\frac{x-a}{b})$ is $g(\frac{x}{b})$ right-shifted by a.

3.2-2 From definition (3.9a), show that the Fourier transform of rect $(t-5)$ is sinc $(\pi f)e^{-j10\pi f}$.

3.2-3 From definition (3.9b), show that the inverse Fourier transform of rect $[(2\pi f - 10)/2\pi]$ is sinc $(\pi t)\, e^{j10t}$.

3.2-4 Using pairs 7 and 12 (Table 3.1) show that $u(t) \Longleftrightarrow 0.5\,\delta(f) + 1/j2\pi f$.

Hint: Add 1 to sgn (t), and see what signal comes out.

3.2-5 Show that $\cos(2\pi f_0 t + \theta) \Longleftrightarrow \frac{1}{2}[\delta(f + f_0)e^{-j\theta} + \delta(f - f_0)e^{j\theta}]$.
Hint: Express $\cos(2\pi f_0 t + \theta)$ in terms of exponentials using Euler's formula.

3.3-1 Apply the duality property to the appropriate pair in Table 3.1 to show that:

(a) $0.5[\delta(t) + (j/\pi t)] \Longleftrightarrow u(f)$

(b) $\delta(t+T) + \delta(t-T) \Longleftrightarrow 2\cos 2\pi f T$

(c) $\delta(t+T) - \delta(t-T) \Longleftrightarrow 2j \sin 2\pi f T$

Hint: $g(-t) \Longleftrightarrow G(-f)$ and $\delta(t) = \delta(-t)$.

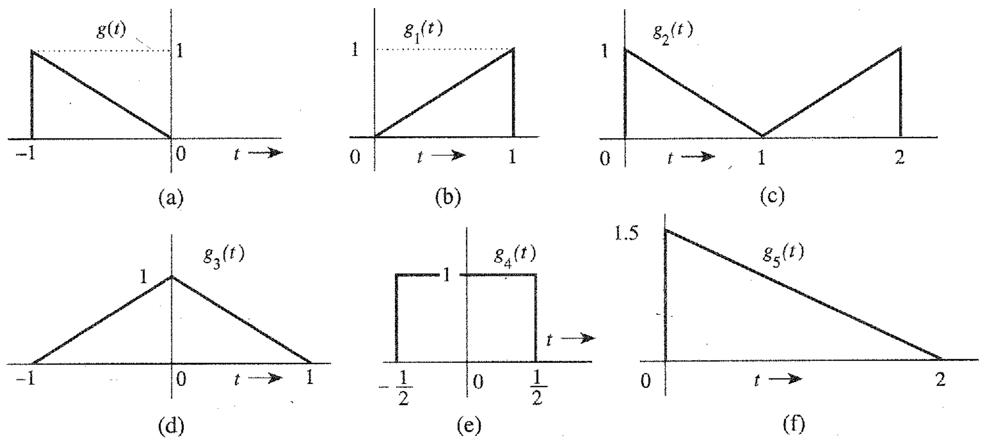
3.3-2 The Fourier transform of the triangular pulse $g(t)$ in Fig. P3.3-2a is given as

$$G(f) = \frac{1}{(2\pi f)^2}(e^{j2\pi f} - j2\pi f e^{j2\pi f} - 1)$$

Use this information, and the time-shifting and time-scaling properties, to find the Fourier transforms of the signals shown in Fig. P3.3-2b–f.

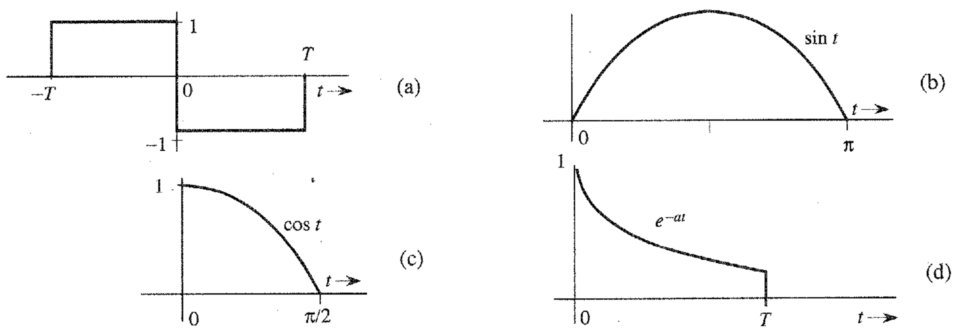
Hint: Time inversion in $g(t)$ results in the pulse $g_1(t)$ in Fig. P3.3-2b; consequently $g_1(t) = g(-t)$. The pulse in Fig. P3.3-2c can be expressed as $g(t-T)+g_1(t-T)$ [the sum of $g(t)$ and $g_1(t)$ both delayed by T]. Both pulses in Fig. P3.3-2d and e can be expressed as $g(t-T)+g_1(t+T)$ [the sum of $g(t)$ delayed by T and $g_1(t)$ advanced by T] for some suitable choice of T. The pulse in Fig. P3.3-2f can be obtained by time-expanding $g(t)$ by a factor of 2 and then delaying the resulting pulse by 2 seconds [or by first delaying $g(t)$ by 1 second and then time-expanding by a factor of 2].

Figure P.3.3-2



3.3-3 Using only the time-shifting property and Table 3.1, find the Fourier transforms of the signals shown in Fig. P3.3-3.

Figure P.3.3-3



Hint: The signal in Fig. P3.3-3a is a sum of two shifted rectangular pulses. The signal in Fig. P3.3-3b is $\sin t\,[u(t) - u(t-\pi)] = \sin t\,u(t) - \sin t\,u(t-\pi) = \sin t\,u(t) + \sin(t-\pi)$

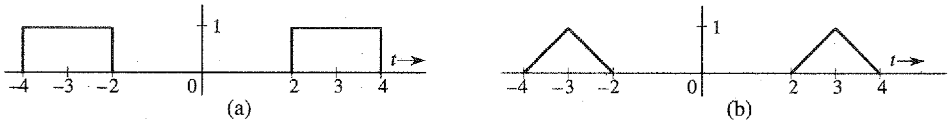
$u(t-\pi)$. The reader should verify that the addition of these two sinusoids indeed results in the pulse in Fig. P3.3-3b. In the same way, we can express the signal in Fig. P3.3-3c as $\cos t\, u(t) + \sin(t-\pi/2)u(t-\pi/2)$ (verify this by sketching these signals). The signal in Fig. P3.3-3d is $e^{-at}[u(t)-u(t-T)] = e^{-at}u(t) - e^{-aT}e^{-a(t-T)}u(t-T)$.

3.3-4 Use the time-shifting property to show that if $g(t) \Longleftrightarrow G(f)$, then

$$g(t+T) + g(t-T) \Longleftrightarrow 2G(f)\cos 2\pi fT$$

This is the dual of Eq. (3.36). Use this result and pairs 17 and 19 in Table 3.1 to find the Fourier transforms of the signals shown in Fig. P3.3-4.

Figure P.3.3-4



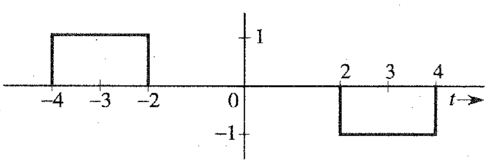
3.3-5 Prove the following results:

$$g(t)\sin 2\pi f_0 t \Longleftrightarrow \frac{1}{2j}[G(f-f_0) - G(f+f_0)]$$

$$\frac{1}{2j}[g(t+T) - g(t-T)] \Longleftrightarrow G(f)\sin 2\pi fT$$

Use the latter result and Table 3.1 to find the Fourier transform of the signal in Fig. P3.3-5.

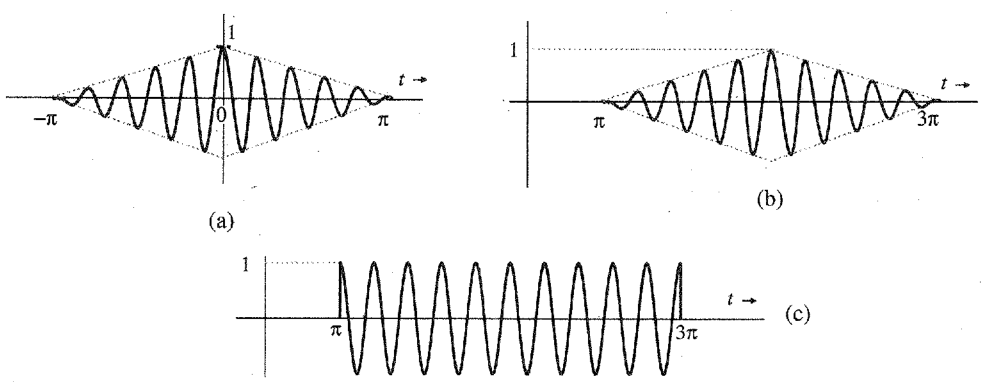
Figure P.3.3-5



3.3-6 The signals in Fig. P3.3-6 are modulated signals with carrier cos $10t$. Find the Fourier transforms of these signals by using the appropriate properties of the Fourier transform and Table 3.1. Sketch the amplitude and phase spectra for Fig. P3.3–6a and b.

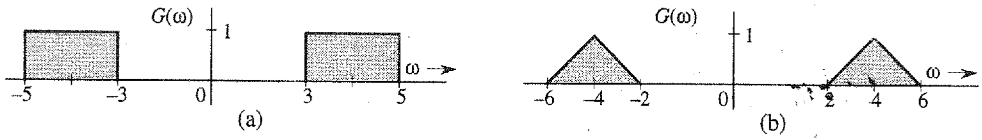
Hint: These functions can be expressed in the form $g(t)\cos 2\pi f_0 t$.

Figure P.3.3-6



3.3-7 Use the frequency shift property and Table 3.1 to find the inverse Fourier transform of the spectra shown in Fig. P3.3-7. Notice that this time, the Fourier transform is in the ω domain.

Figure P.3.3-7



3.3-8 A signal $g(t)$ is band-limited to B Hz. Show that the signal $g^n(t)$ is band-limited to nB Hz.

Hint: $g^2(t) \Longleftrightarrow [G(f) * G(f)]$, and so on. Use the width property of convolution.

3.3-9 Find the Fourier transform of the signal in Fig. P3.3-3a by three different methods:

(a) By direct integration using the definition (3.9a).

(b) Using only pair 17 Table 3.1 and the time-shifting property.

(c) Using the time differentiation and time-shifting properties, along with the fact that $\delta(t) \Longleftrightarrow 1$.

Hint: $1 - \cos 2x = 2\sin^2 x$.

3.3-10 The process of recovering a signal $g(t)$ from the modulated signal $g(t)\cos 2\pi f_0 t$ is called **demodulation**. Show that the signal $g(t)\cos 2\pi f_0 t$ can be demodulated by multiplying it by $2\cos 2\pi f_0 t$ and passing the product through a low-pass filter of bandwidth B Hz [the bandwidth of $g(t)$]. Assume $B < f_0$. *Hint*: $2\cos^2 2\pi f_0 t = 1 + \cos 4\pi f_0 t$. Recognize that the spectrum of $g(t)\cos 4\pi f_0 t$ is centered at $2f_0$ and will be suppressed by a low-pass filter of bandwidth B Hz.

3.4-1 Signals $g_1(t) = 10^4 \Pi(10^4 t)$ and $g_2(t) = \delta(t)$ are applied at the inputs of the ideal low-pass filters $H_1(f) = \Pi(f/20{,}000)$ and $H_2(f) = \Pi(f/10{,}000)$ (Fig. P3.4-1). The outputs $y_1(t)$ and $y_2(t)$ of these filters are multiplied to obtain the signal $y(t) = y_1(t)y_2(t)$.

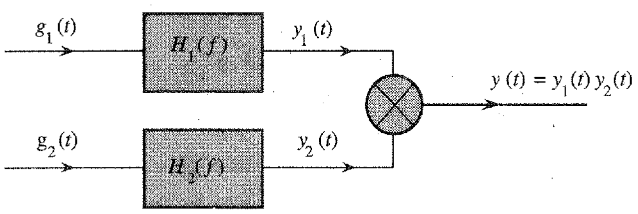
(a) Sketch $G_1(f)$ and $G_2(f)$.

(b) Sketch $H_1(f)$ and $H_2(f)$.

(c) Sketch $Y_1(f)$ and $Y_2(f)$.

(d) Find the bandwidths of $y_1(t), y_2(t)$, and $y(t)$.

Figure P.3.4-1



3.5-1 For systems with the following impulse responses, which system is causal?

(a) $h(t) = e^{-at}u(t), \quad a > 0$

(b) $h(t) = e^{-a|t|}, \quad a > 0$

(c) $h(t) = e^{-a(t-t_0)}u(t - t_0), \quad a > 0$

(d) $h(t) = \text{sinc}(at), \quad a > 0$

(e) $h(t) = \text{sinc}[a(t - t_0)], \quad a > 0.$

3.5-2 Consider a filter with the transfer function

$$H(f) = e^{-k(2\pi kf)^2 - j2\pi f t_0}$$

Show that this filter is physically unrealizable by using the time domain criterion [noncausal $h(t)$] and the frequency domain (Paley-Wiener) criterion. Can this filter be made approximately realizable by choosing a sufficiently large t_0? Use your own (reasonable) criterion of approximate realizability to determine t_0.

Hint: Use pair 22 in Table 3.1.

3.5-3 Show that a filter with transfer function

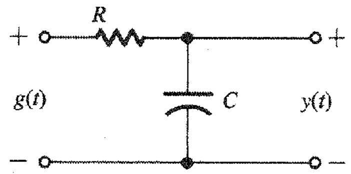
$$H(f) = \frac{2(10^5)}{(2\pi f)^2 + 10^{10}} e^{-j2\pi f t_0}$$

is unrealizable. Can this filter be made approximately realizable by choosing a sufficiently large t_0? Use your own (reasonable) criterion of approximate realizability to determine t_0.

Hint: Show that the impulse response is noncausal.

3.5-4 Determine the maximum bandwidth of a signal that can be transmitted through the low-pass RC filter in Fig. P3.5-4 with $R = 1000$ and $C = 10^{-9}$ if, over this bandwidth, the amplitude response (gain) variation is to be within 5% and the time delay variation is to be within 2%.

Figure P.3.5-4



3.5-5 A bandpass signal $g(t)$ of bandwidth $B = 2000$ Hz centered at $f = 10^5$ Hz is passed through the RC filter in Fig. P3.5-4 with $RC = 10^{-3}$. If over the passband, a variation of less than 2% in amplitude response and less than 1% in time delay is considered distortionless transmission, would $g(t)$ be transmitted without distortion? Find the approximate expression for the output signal.

3.6-1 A certain channel has ideal amplitude, but nonideal phase response (Fig. P3.6-1), given by

$$|H(f)| = 1$$
$$\theta_h(f) = -2\pi f t_0 - k \sin 2\pi f T \qquad k \ll 1$$

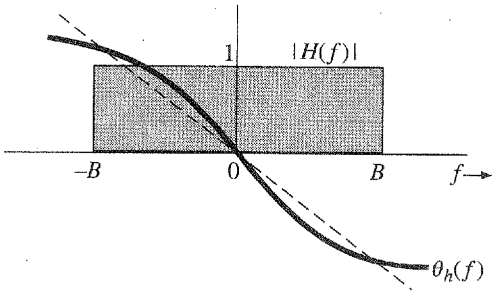
(a) Show that $y(t)$, the channel response to an input pulse $g(t)$ band-limited to B Hz, is

$$y(t) = g(t - t_0) + \frac{k}{2}[g(t - t_0 - T) - g(t - t_0 + T)]$$

Hint: Use $e^{-jk \sin 2\pi f T} \approx 1 - jk \sin 2\pi f T$.

(b) Discuss how this channel will affect TDM and FDM systems from the viewpoint of interference among the multiplexed signals.

Figure P.3.6-1

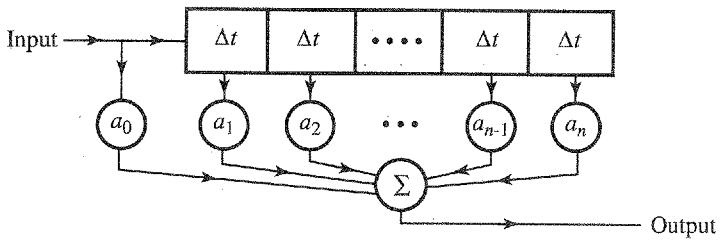


3.6-2 The distortion caused by multipath transmission can be partly corrected by a tapped delay-line equalizer. Show that if $\alpha \ll 1$, the distortion in the multipath system in Fig. 3.31a can be approximately corrected if the received signal in Fig. 3.31a is passed through the tapped delay-line equalizer shown in Fig. P3.6-2.

Hint: From Eq. (3.64a), it is clear that the equalizer filter transfer function should be $H_{eq}(f) = 1/(1+\alpha\, e^{-j2\pi f \Delta t})$. Use the fact that $1/(1-x) = 1+x+x^2+x^3+\cdots$ if $x \ll 1$ to show what should be the tap parameters a_i to make the resulting transfer function

$$H(f)H_{eq}(f) \approx e^{-j2\pi f t_d}$$

Figure P.3.6-2



3.7-1 Show that the energy of the Gaussian pulse

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$$

from direct integration is $1/2\sigma\sqrt{\pi}$. Verify this result by using Parseval's theorem to derive the energy E_g from $G(f)$. *Hint:* See pair 22 in Table 3.1. Use the fact that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2}\,dx\,dy = \pi \Rightarrow \int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

3.7-2 Show that

$$\int_{-\infty}^{\infty} \text{sinc}^2\,(kt)dt = \frac{\pi}{4}$$

Hint: Recognize that the integral is the energy of $g(t) = \text{sinc}\,(kt)$. Use Parseval's theorem to find this energy.

3.7-3 Generalize Parseval's theorem to show that for real Fourier transformable signals $g_1(t)$ and $g_2(t)$,

$$\int_{-\infty}^{\infty} g_1(t)g_2(t)\,dt = \int_{-\infty}^{\infty} G_1(-f)G_2(f)\,df = \int_{-\infty}^{\infty} G_1(f)G_2(-f)\,df$$

3.7-4 Show that

$$\int_{-\infty}^{\infty} \text{sinc}\,(2\pi Bt - m\pi)\,\text{sinc}\,(2\pi Bt - n\pi)\,dt = \begin{cases} 0 & m \neq n \\ \frac{1}{2B} & m = n \end{cases}$$

Hint: Recognize that

$$\text{sinc}\,(2\pi Bt - k\pi) = \text{sinc}\left[2\pi B\left(t - \frac{k}{2B}\right)\right] \Longleftrightarrow \frac{1}{2B}\Pi,\left(\frac{f}{2B}\right)e^{-j\pi fk/B}$$

Use this fact and the result in Prob. 3.7-2 to show that

$$\int_{-\infty}^{\infty} \text{sinc}\,(2\pi Bt - m\pi)\,\text{sinc}\,(2\pi Bt - n\pi)\,dt = \frac{1}{4B^2}\int_{-B}^{B} e^{j[(n-m)/2B]2\pi f}\,df$$

The desired result follows from this integral.

3.7-5 For the signal

$$g(t) = \frac{2a}{t^2 + a^2}$$

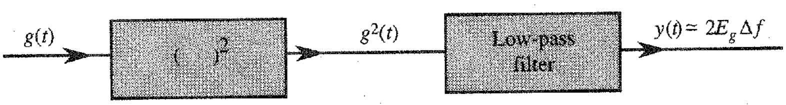
determine the essential bandwidth B Hz of $g(t)$ such that the energy contained in the spectral components of $g(t)$ of frequencies below B Hz is 99% of the signal energy E_g.

Hint: Determine $G(f)$ by applying the duality property [Eq. (3.26)] to pair 3 of Table 3.1.

3.7-6 A low-pass signal $g(t)$ is applied to a squaring device. The squarer output $g^2(t)$ is applied to a unity gain ideal low-pass filter of bandwidth Δf Hz (Fig. P3.7-6). Show that if Δf is very small ($\Delta f \to 0$), the filter output is a dc signal of amplitude $2E_g\Delta f$, where E_g is the energy of $g(t)$.

Hint: The output $y(t)$ is a dc signal because its spectrum $Y(f)$ is concentrated at $f = 0$ from $-\Delta f$ to Δf with $\Delta f \to 0$ (impulse at the origin). If $g^2(t) \Longleftrightarrow A(f)$, and $y(t) \Longleftrightarrow Y(f)$, then $Y(f) \approx [2A(0)\Delta f]\delta(f)$. Now, show that $E_g = A(0)$.

Figure P.3.7-6



3.8-1 Show that the autocorrelation function of $g(t) = C\cos(2\pi f_0 t + \theta_0)$ is given by $\mathcal{R}_g(\tau) = (C^2/2)\cos 2\pi f_0\tau$, and the corresponding PSD is $S_g(f) = (C^2/4)[\delta(f - f_0) + \delta(f + f_0)]$. Hence, show that for a signal $y(t)$ given by

$$y(t) = C_0 + \sum_{n=1}^{\infty} C_n \cos(n2\pi f_0 t + \theta_n)$$

the autocorrelation function and the PSD are given by

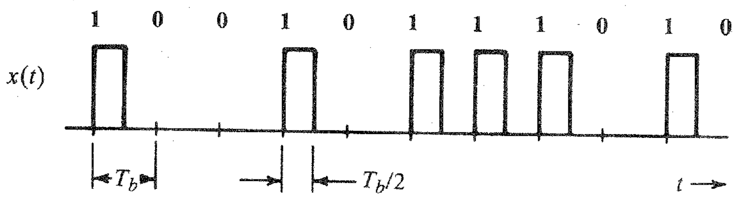
$$\mathcal{R}_y(\tau) = C_0{}^2 + \frac{1}{2}\sum_{n=1}^{\infty} C_n{}^2 \cos n2\pi f_0\tau$$

$$S_y(f) = C_0{}^2\delta(f) + \frac{1}{4}\sum_{n=1}^{\infty} C_n{}^2[\delta(f - nf_0) + \delta(f + nf_0)]$$

Hint: Show that if $g(t) = g_1(t)+g_2(t)$, then $\mathcal{R}_g(\tau) = \mathcal{R}_{g_1}(\tau)+\mathcal{R}_{g_2}(\tau)+\mathcal{R}_{g_1g_2}(\tau)+\mathcal{R}_{g_2g_1}(\tau)$, where $\mathcal{R}_{g_1g_2}(\tau) = \lim_{T\to\infty}(1/T)\int_{-T/2}^{T/2} g_1(t)g_2(t+\tau)\,dt$. If $g_1(t)$ and $g_2(t)$ represent any two of the infinite terms in $y(t)$, then show that $\mathcal{R}_{g_1g_2}(\tau) = \mathcal{R}_{g_2g_1}(\tau) = 0$. To show this, use the fact that the area under any sinusoid over a very large time interval is at most equal to the area of the half-cycle of the sinusoid.

3.8-2 The random binary signal $x(t)$ shown in Fig. P3.8-2 transmits one digit every T_b seconds. A binary **1** is transmitted by a pulse $p(t)$ of width $T_b/2$ and amplitude A; a binary **0** is transmitted by no pulse. The digits **1** and **0** are equally likely and occur randomly. Determine the autocorrelation function $\mathcal{R}_x(\tau)$ and the PSD $S_x(f)$.

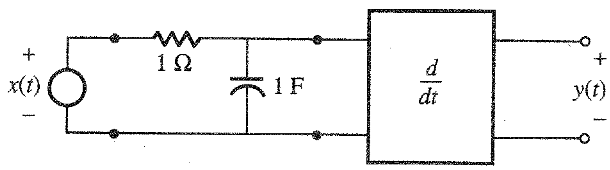
Figure P.3.8-2



3.8-3 Find the mean square value (or power) of the output voltage $y(t)$ of the RC network shown in Fig. P3.5-4 with $RC = 2\pi$ if the input voltage PSD $S_x(f)$ is given by **(a)** K; **(b)** $\Pi(\pi f)$; **(c)** $[\delta(f+1)+\delta(f-1)]$. In each case calculate the power (mean square value) of the input signal $x(t)$.

3.8-4 Find the mean square value (or power) of the output voltage $y(t)$ of the system shown in Fig. P3.8-4 if the input voltage PSD $S_x(f) = \Pi(\pi f)$. Calculate the power (mean square value) of the input signal $x(t)$.

Figure P.3.8-4



4 AMPLITUDE MODULATIONS AND DEMODULATIONS

Modulation often refers to a process that moves the message signal into a specific frequency band that is dictated by the physical channel (e.g. voiceband telephone modems). Modulation provides a number of advantages mentioned in Chapter 1 including ease of RF transmission and frequency division multiplexing. Modulations can be analog or digital. Though traditional communication systems such as AM/FM radios and NTSC television signals are based on analog modulations, more recent systems such as 2G and 3G cellphones, HDTV, and DSL are all digital.

In this chapter and the next, we will focus on the classic analog modulations: amplitude modulation and angle modulation. Before we begin our discussion of different analog modulations, it is important to distinguish between communication systems that do not use modulation (**baseband communications**) and systems that use modulation (**carrier communications**).

4.1 BASEBAND VERSUS CARRIER COMMUNICATIONS

The term **baseband** is used to designate the frequency band of the original message signal from the source or the input transducer (see Fig. 1.2). In telephony, the baseband is the audio band (band of voice signals) of 0 to 3.5 kHz. In NTSC television, the video baseband is the video band occupying 0 to 4.3 MHz. For digital data or pulse code modulation (PCM) that uses bipolar signaling at a rate of R_b pulses per second, the baseband is approximately 0 to R_b Hz.

Baseband Communications

In baseband communication, message signals are directly transmitted without any modification. Because most baseband signals such as audio and video contain significant low-frequency content, they cannot be effectively transmitted over radio (wireless) links. Instead, dedicated user channels such as twisted pairs of copper wires and coaxial cables are assigned to each user for long-distance communications. Because baseband signals have overlapping bands, they would interfere severely if sharing a common channel. Thus, baseband communications leave much of the channel spectrum unused. By modulating several baseband signals and shifting their spectra to nonoverlapping bands, many users can share one channel by utilizing most of the available bandwidth through frequency division multiplexing (FDM). Long-haul communication over a radio link also requires modulation to shift the signal spectrum to higher frequencies in order to enable efficient power radiation using antennas of reasonable dimensions. Yet another use of modulation is to exchange transmission bandwidth for better performance against interferences.

Carrier Modulations

Communication that uses modulation to shift the frequency spectrum of a signal is known as **carrier communication**. In terms of analog modulation, one of the basic parameters (amplitude, frequency, or phase) of a **sinusoidal carrier** of high frequency f_c Hz (or $\omega_c = 2\pi f_c$ rad/s) is varied linearly with the baseband signal $m(t)$. This results in amplitude modulation (AM), frequency modulation (FM), or phase modulation (PM), respectively. Amplitude modulation is linear, while the latter two types of carrier modulation are similar and nonlinear, often known collectively as **angle modulation**.

A comment about pulse-modulated signals [pulse amplitude modulation (PAM), pulse width modulation (PWM), pulse position modulation (PPM), pulse code modulation (PCM), and delta modulation (DM)] is in order here. Despite the term *modulation*, these signals are baseband digital signals. "Modulation" is used here not in the sense of frequency or band shifting. Rather, in these cases it is in fact describing digital pulse coding schemes used to represent the original analog signals. In other words, the analog message signal is modulating parameters of a digital pulse train. These signals can still modulate a carrier in order to shift their spectra.

Amplitude Modulations and Angle Modulations

We denote as $m(t)$ the source message signal that is to be transmitted by the sender to its receivers; its Fourier transform is denoted as $M(f)$. To move the frequency response of $m(t)$ to a new frequency band centered at f_c Hz, we begin by noting that the Fourier transform has already revealed a very strong property known as the *frequency shifting* property to achieve this goal. In other words, all we need to do is to multiply $m(t)$ by a sinusoid of frequency f_c such that

$$s_1(t) = m(t) \cos 2\pi f_c t$$

This immediately achieves the basic aim of modulation by moving the signal frequency content to be centered at $\pm f_c$ via

$$S_1(f) = \frac{1}{2}M(f - f_c) + \frac{1}{2}M(f + f_c)$$

This simple multiplication is in fact allowing changes in the amplitude of the sinusoid $s_1(t)$ to be proportional to the message signal. This method is indeed a very valuable modulation known as amplitude modulation.

More broadly, consider a sinusoidal signal

$$s(t) = A(t) \cos [\omega_c t + \phi(t)]$$

There are three variables in a sinusoid: amplitude, (instantaneous) frequency, and phase. Indeed, the message signal can be used to modulate any one of these three parameters to allow $s(t)$ to carry the information from the transmitter to the receiver:

Amplitude $A(t)$ linearly varies with $m(t)$	$\Longleftrightarrow$	amplitude modulation
Frequency linearly varies with $m(t)$	$\Longleftrightarrow$	frequency modulation
Phase $\phi(t)$ linearly varies with $m(t)$	$\Longleftrightarrow$	phase modulation

These are known, respectively, as amplitude modulation, frequency modulation, and phase modulation. In this chapter, we describe various forms of amplitude modulation in practical communication systems. Amplitude modulations are linear, and their analysis in the time and frequency domains is simpler. In Chapter 5, we will separately discuss nonlinear angle modulations.

The Interchangeable Use of f and ω

In Chapter 3, we noted the equivalence of frequency response denoted by frequency f with angular frequency ω. Each of these two notations has its own advantages and disadvantages. After the examples and problems of Chapter 3, readers should be familiar and comfortable with the use of either notation. Thus, from this point on, we will use the two different notations interchangeably, selecting one or the other on the basis of notational or graphical simplicity.

4.2 DOUBLE-SIDEBAND AMPLITUDE MODULATION

Amplitude modulation is characterized by an information-bearing **carrier** amplitude $A(t)$ that is a linear function of the baseband (message) signal $m(t)$. At the same time, the carrier frequency ω_c and the phase θ_c remain constant. We can assume $\theta_c = 0$ without loss of generality. If the carrier amplitude A is made directly proportional to the modulating signal $m(t)$, then *modulated signal* is $m(t)\cos\ \omega_c t$ (Fig. 4.1). As we saw earlier [Eq. (3.36)], this type of modulation simply shifts the spectrum of $m(t)$ to the carrier frequency (Fig. 4.1a). Thus, if

$$m(t) \Longleftrightarrow M(f)$$

then

$$m(t)\cos\ 2\pi f_c t \Longleftrightarrow \frac{1}{2}[M(f+f_c)+M(f-f_c)] \tag{4.1}$$

Recall that $M(f-f_c)$ is $M(f)$ shifted to the right by f_c, and $M(f+f_c)$ is $M(f)$ shifted to the left by f_c. Thus, the process of modulation shifts the spectrum of the modulating signal to the left and to the right by f_c. Note also that if the bandwidth of $m(t)$ is B Hz, then, as seen from Fig. 4.1c, the modulated signal now has bandwidth of $2B$ Hz. We also observe that the modulated signal spectrum centered at $\pm f_c$ (or $\pm\omega_c$ in rad/s) consists of two parts: a portion that lies outside $\pm f_c$, known as the *upper sideband (USB)*, and a portion that lies inside $\pm f_c$, known as the *lower sideband (LSB)*. We can also see from Fig. 4.1c that, unless the message signal $M(f)$ has an impulse at zero frequency, the modulated signal in this scheme does not contain a discrete component of the carrier frequency f_c. In other words, the modulation process does not introduce a sinusoid at f_c. For this reason it is called **double-sideband suppressed carrier (DSB-SC) modulation.***

The relationship of B to f_c is of interest. Figure 4.1c shows that $f_c \geq B$, thus avoiding overlap of the modulated spectra centered at f_c and $-f_c$. If $f_c < B$, then the two copies of message spectra overlap and the information of $m(t)$ is lost during modulation, which makes it impossible to get back $m(t)$ from the modulated signal $m(t)\cos\ \omega_c t$.

Note that practical factors may impose additional restrictions on f_c. For instance, in broadcast applications, a transmit antenna can radiate only a narrow band without distortion. This means that to avoid distortion caused by the transmit antenna, we must have $f_c/B \gg 1$. The

---

* The term **suppressed carrier** does not necessarily mean absence of the spectrum at the carrier frequency f_c. It means that there is no discrete component of the carrier frequency. This implies that the spectrum of the DSB-SC does not have impulses at $\pm f_c$, which also implies that the modulated signal $m(t)\cos\ 2\pi f_c t$ does not contain a term of the form $k\cos\ 2\pi f_c t$ [assuming that $m(t)$ has a zero mean value].

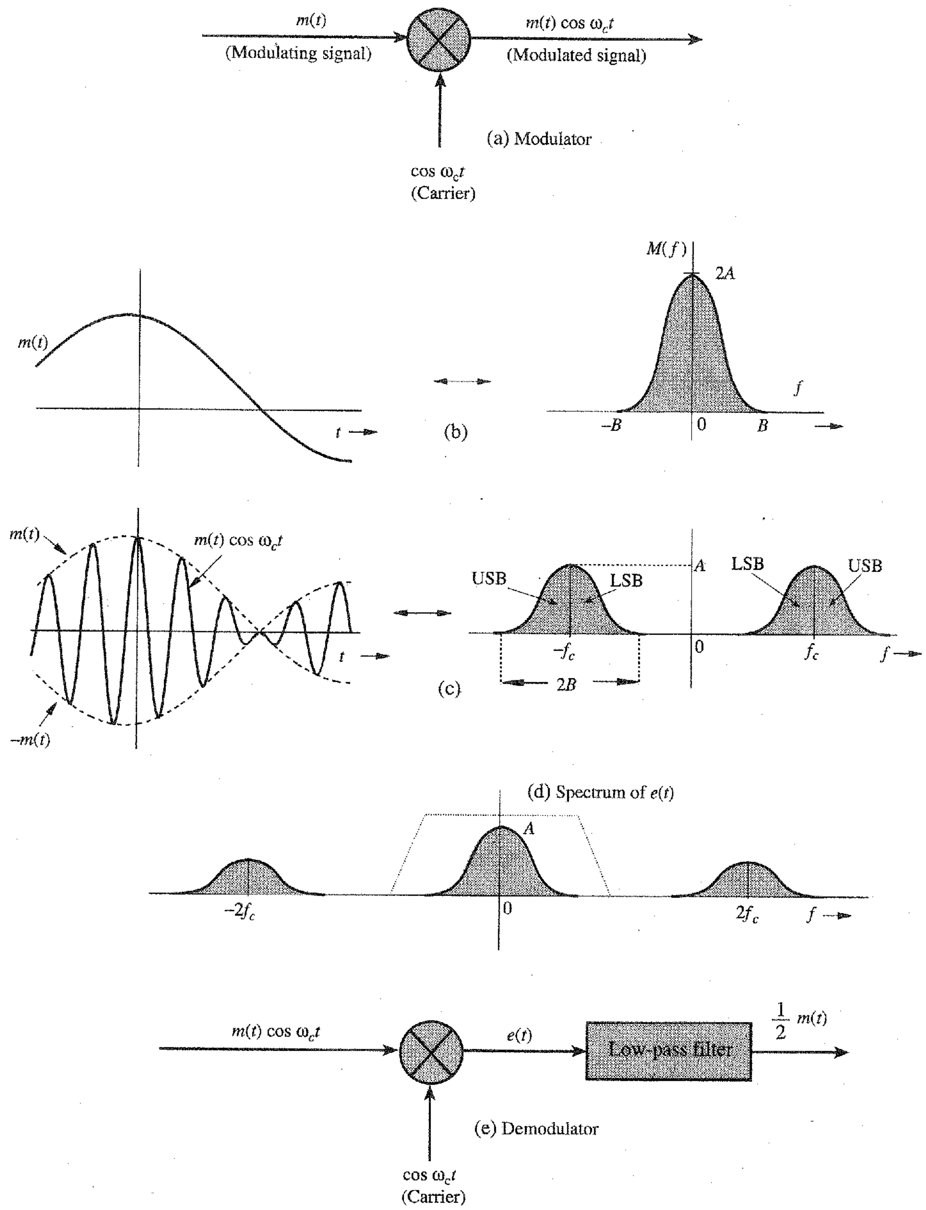


Figure 4.1 DSB-SC modulation and demodulation.

broadcast band AM radio, for instance, with $B = 5$ kHz and the band of 550 to 1600 kHz for the carrier frequencies, gives a ratio of f_c/B roughly in the range of 100 to 300.

Demodulation

The DSB-SC modulation translates or shifts the frequency spectrum to the left and the right by f_c (i.e., at $+f_c$ and $-f_c$), as seen from Eq. (4.1). To recover the original signal $m(t)$ from the modulated signal, it is necessary to retranslate the spectrum to its original position. The process of recovering the signal from the modulated signal (retranslating the spectrum to its original position) is referred to as **demodulation**. Observe that if the modulated signal spectrum in Fig. 4.1c is shifted to the left and to the right by f_c (and multiplied by one-half), we obtain the spectrum shown in Fig. 4.1d, which contains the desired baseband spectrum plus an unwanted spectrum at $\pm 2f_c$. The latter can be suppressed by a low-pass filter. Thus, demodulation, which is almost identical to modulation, consists of multiplication of the incoming modulated signal $m(t)\cos\,\omega_c t$ by a carrier $\cos\,\omega_c t$ followed by a low-pass filter, as shown in Fig. 4.1e. We can verify this conclusion directly in the time domain by observing that the signal $e(t)$ in Fig. 4.1e is

$$\begin{aligned} e(t) &= m(t)\cos^2 \omega_c t \\ &= \frac{1}{2}[m(t) + m(t)\cos\, 2\omega_c t] \end{aligned} \tag{4.2a}$$

Therefore, the Fourier transform of the signal $e(t)$ is

$$E(f) = \frac{1}{2}M(f) + \frac{1}{4}[M(f + 2f_c) + M(f - 2f_c)] \tag{4.2b}$$

This analysis shows that the signal $e(t)$ consists of two components $(1/2)m(t)$ and $(1/2)m(t)\cos\,2\omega_c t$, with their nonoverlapping spectra as shown in Fig. 4.1d. The spectrum of the second component, being a modulated signal with carrier frequency $2f_c$, is centered at $\pm 2f_c$. Hence, this component is suppressed by the low-pass filter in Fig. 4.1e. The desired component $(1/2)M(f)$, being a low-pass spectrum (centered at $f = 0$), passes through the filter unharmed, resulting in the output $(1/2)m(t)$. A possible form of low pass filter characteristics is shown (under the dotted line) in Fig. 4.1d. The filter leads to a distortionless demodulation of the message signal $m(t)$ from the DSB-SC signal. We can get rid of the inconvenient fraction $1/2$ in the output by using a carrier $2\cos\,\omega_c t$ instead of $\cos\,\omega_c t$. In fact, later on, we shall often use this strategy, which does not affect general conclusions.

This method of recovering the baseband signal is called *synchronous detection*, or *coherent detection*, where we use a carrier of exactly the same frequency (and phase) as the carrier used for modulation. Thus, for demodulation, we need to generate a local carrier at the receiver in frequency and phase coherence (synchronism) with the carrier used at the modulator.

Example 4.1 For a baseband signal

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

find the DSB-SC signal, and sketch its spectrum. Identify the USB and LSB. Verify that the DSB-SC modulated signal can be demodulated by the demodulator in Fig. 4.1e.

The case in this example is referred to as *tone modulation* because the modulating signal is a pure sinusoid, or tone, $\cos\,\omega_m t$. To clarify the basic concepts of DSB-SC modulation, we shall work this problem in the frequency domain as well as the time domain. In the frequency domain approach, we work with the signal spectra. The spectrum of the baseband signal $m(t) = \cos\,\omega_m t$ is given by

$$\begin{aligned} M(f) &= \frac{1}{2}[\delta(f - f_m) + \delta(f + f_m)] \\ &= \pi[\delta(\omega - \omega_m) + \delta(\omega + \omega_m)] \end{aligned}$$

The message spectrum consists of two impulses located at $\pm f_m$, as shown in Fig. 4.2a. The DSB-SC (modulated) spectrum, as seen from Eq. (4.1), is the baseband spectrum in Fig. 4.2a shifted to the right and the left by f_c (times one-half), as shown in Fig. 4.2b. This spectrum consists of impulses at angular frequencies $\pm(f_c - f_m)$ and $\pm(f_c + f_m)$. The spectrum beyond f_c is the USB, and the one below f_c is the LSB. Observe that the DSB-SC spectrum does not have the component of the carrier frequency f_c. This is why it is called *suppressed carrier.*

In the time domain approach, we work directly with signals in the time domain. For the baseband signal $m(t) = \cos\,\omega_m t$, the DSB-SC signal $\varphi_{\text{DSB-SC}}(t)$ is

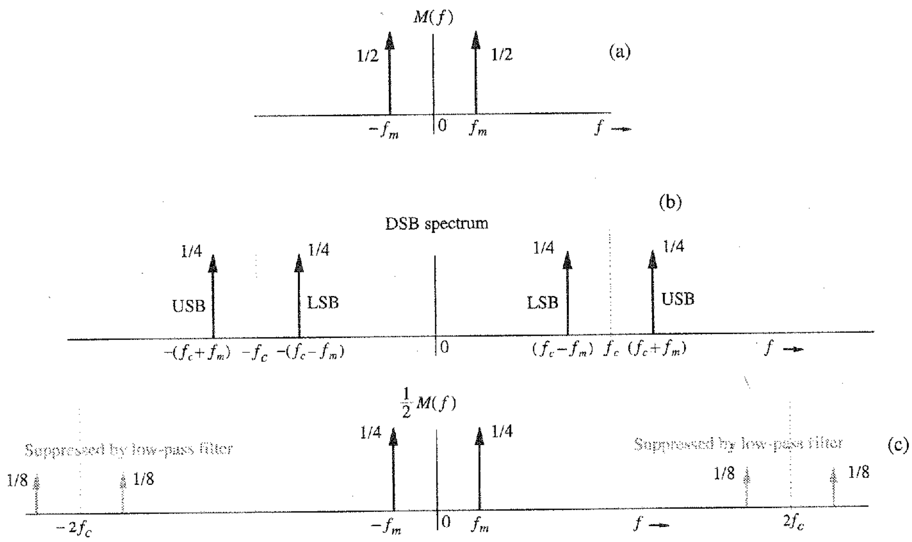
$$\begin{aligned}\varphi_{\text{DSB-SC}}(t) &= m(t)\cos\,\omega_c t\\ &= \cos\,\omega_m t\cos\,\omega_c t\\ &= \frac{1}{2}[\cos\,(\omega_c + \omega_m)t + \cos\,(\omega_c - \omega_m)t]\end{aligned}$$

This shows that when the baseband (message) signal is a single sinusoid of frequency f_m, the modulated signal consists of two sinusoids: the component of frequency $f_c + f_m$ (the USB) and the component of frequency $f_c - f_m$ (the LSB). Figure 4.2b shows precisely the spectrum of $\varphi_{\text{DSB-SC}}(t)$. Thus, each component of frequency f_m in the modulating signal turns into two components of frequencies $f_c + f_m$ and $f_c - f_m$ in the modulated signal. Note the curious fact that there is no component of the carrier frequency f_c on the right-hand side of the preceding equation. As mentioned, this is why it is called double-sideband suppressed carrier (DSB-SC) modulation.

We now verify that the modulated signal $\varphi_{\text{DSB-SC}}(t) = \cos\,\omega_m t\cos\,\omega_c t$, when applied to the input of the demodulator in Fig. 4.1e, yields the output proportional to the desired baseband signal $\cos\,\omega_m t$. The signal $e(t)$ in Fig. 4.1e is given by

$$\begin{aligned}e(t) &= \cos\,\omega_m t\cos^2\omega_c t\\ &= \frac{1}{2}\cos\,\omega_m t\,(1 + \cos\,2\omega_c t)\end{aligned}$$

Figure 4.2 Example of DSB-SC modulation.



The spectrum of the term $\cos \omega_m t \cos 2\omega_c t$ is centered at $2\omega_c$ and will be suppressed by the low-pass filter, yielding $\frac{1}{2} \cos \omega_m t$ as the output. We can also derive this result in the frequency domain. Demodulation causes the spectrum in Fig. 4.2b to shift left and right by ω_c (and to be multiplied by one-half). This results in the spectrum shown in Fig. 4.2c. The low-pass filter suppresses the spectrum centered at $\pm 2\omega_c$, yielding the spectrum $\frac{1}{2}M(f)$.

Modulators

Modulation can be achieved in several ways. We shall discuss some important categories of modulators.

Multiplier Modulators: Here modulation is achieved directly by multiplying $m(t)$ with $\cos \omega_c t$, using an analog multiplier whose output is proportional to the product of two input signals. Typically, such a multiplier may be obtained from a variable-gain amplifier in which the gain parameter (such as the β of a transistor) is controlled by one of the signals, say, $m(t)$. When the signal $\cos \omega_c t$ is applied at the input of this amplifier, the output is proportional to $m(t) \cos \omega_c t$.

In the early days, multiplication of two signals over a sizable dynamic range was a challenge to circuit designers. However, as semiconductor technologies continued to advance, signal multiplication ceased to be a major concern. Still, we will present several classical modulators that avoid the use of multipliers. Studying these modulators can provide unique insight and an excellent opportunity to pick up some new signal analysis skills.

Nonlinear Modulators: Modulation can also be achieved by using nonlinear devices, such as a semiconductor diode or a transistor. Figure 4.3 shows one possible scheme, which uses two identical nonlinear elements (boxes marked NL).

Let the input-output characteristics of either of the nonlinear elements be approximated by a power series

$$y(t) = ax(t) + bx^2(t) \tag{4.3}$$

where $x(t)$ and $y(t)$ are the input and the output, respectively, of the nonlinear element. The summer output $z(t)$ in Fig. 4.3 is given by

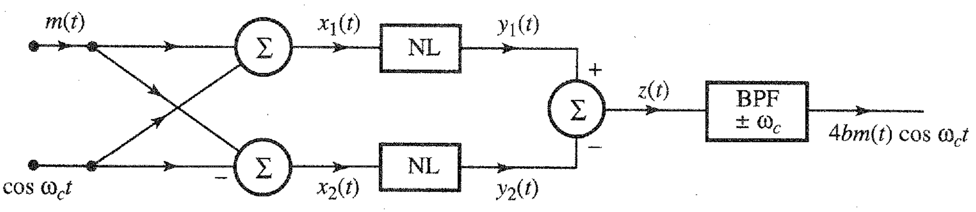
$$z(t) = y_1(t) - y_2(t) = [ax_1(t) + bx_1{}^2(t)] - [ax_2(t) + bx_2{}^2(t)]$$

Substituting the two inputs $x_1(t) = \cos \omega_c t + m(t)$ and $x_2(t) = \cos \omega_c t - m(t)$ in this equation yields

$$z(t) = 2a \cdot m(t) + 4b \cdot m(t) \cos \omega_c t$$

The spectrum of $m(t)$ is centered at the origin, whereas the spectrum of $m(t) \cos \omega_c t$ is centered at $\pm\omega_c$. Consequently, when $z(t)$ is passed through a bandpass filter tuned to ω_c, the signal $am(t)$ is suppressed and the desired modulated signal $4bm(t) \cos \omega_c t$ can pass through the system without distortion.

Figure 4.3 Nonlinear DSB-SC modulator.



In this circuit there are two inputs: $m(t)$ and $\cos\ \omega_c t$. The output of the last summer, $z(t)$, no longer contains one of the inputs, the carrier signal $\cos\ \omega_c t$. Consequently, the carrier signal does not appear at the input of the final bandpass filter. The circuit acts as a balanced bridge for one of the inputs (the carrier). Circuits that have this characteristic are called *balanced circuits*. The nonlinear modulator in Fig. 4.3 is an example of a class of modulators known as *balanced modulators*. This circuit is balanced with respect to only one input (the carrier); the other input $m(t)$ still appears at the final bandpass filter, which must reject it. For this reason, it is called a *single balanced modulator*. A circuit balanced with respect to both inputs is called a *double balanced modulator*, of which the ring modulator (see later: Fig. 4.6) is an example.

Switching Modulators: The multiplication operation required for modulation can be replaced by a simpler switching operation if we realize that a modulated signal can be obtained by multiplying $m(t)$ not only by a pure sinusoid but by any periodic signal $\phi(t)$ of the fundamental radian frequency ω_c. Such a periodic signal can be expressed by a trigonometric Fourier series as

$$\phi(t) = \sum_{n=0}^{\infty} C_n \cos\ (n\omega_c t + \theta_n) \tag{4.4a}$$

Hence,

$$m(t)\phi(t) = \sum_{n=0}^{\infty} C_n\ m(t) \cos\ (n\omega_c t + \theta_n) \tag{4.4b}$$

This shows that the spectrum of the product $m(t)\phi(t)$ is the spectrum $M(\omega)$ shifted to $\pm\omega_c,\ \pm 2\omega_c, \ldots, \pm n\omega_c, \ldots$. If this signal is passed through a bandpass filter of bandwidth $2B$ Hz and tuned to ω_c, then we get the desired modulated signal $c_1 m(t) \cos\ (\omega_c t + \theta_1)$.*

The square pulse train $w(t)$ in Fig. 4.4b is a periodic signal whose Fourier series was found earlier (by rewriting the results of Example 2.4) as

$$w(t) = \frac{1}{2} + \frac{2}{\pi}\left(\cos\ \omega_c t - \frac{1}{3}\cos\ 3\omega_c t + \frac{1}{5}\cos\ 5\omega_c t - \cdots\right) \tag{4.5}$$

The signal $m(t)w(t)$ is given by

$$m(t)w(t) = \frac{1}{2}m(t) + \frac{2}{\pi}\left[m(t)\cos\ \omega_c t - \frac{1}{3}m(t)\cos\ 3\omega_c t + \frac{1}{5}m(t)\cos\ 5\omega_c t - \cdots\right] \tag{4.6}$$

The signal $m(t)w(t)$ consists not only of the component $m(t)$ but also of an infinite number of modulated signals with carrier frequencies ω_c, $3\omega_c$, $5\omega_c$, ... Therefore, the spectrum of $m(t)w(t)$ consists of multiple copies of the message spectrum $M(f)$, shifted to 0, $\pm f_c$, $\pm 3f_c$, $\pm 5f_c$, ... (with decreasing relative weights), as shown in Fig. 4.4c.

For modulation, we are interested in extracting the modulated component $m(t) \cos\ \omega_c t$ only. To separate this component from the rest of the crowd, we pass the signal $m(t)w(t)$ through a bandpass filter of bandwidth $2B$ Hz (or $4\pi B$ rad/s), centered at the frequency $\pm f_c$. Provided the carrier frequency $f_c \geq 2B$ (or $\omega_c \geq 4\pi B$), this will suppress all the spectral components not centered at $\pm f_c$ to yield the desired modulated signal $(2/\pi)m(t) \cos\ \omega_c t$ (Fig. 4.4d).

We now see the real payoff of this method. Multiplication of a signal by a square pulse train is *in reality* a switching operation in which the signal $m(t)$ is switched on and off periodically; it

* The phase θ_1 is not important.

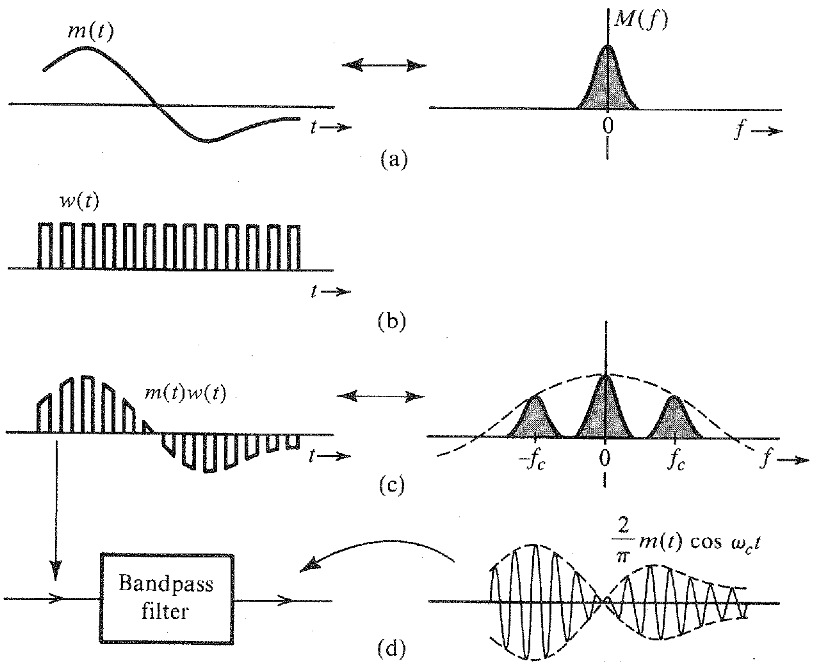


Figure 4.4 Switching modulator for DSB-SC.

can be accomplished by simple switching elements controlled by $w(t)$. Figure 4.5a shows one such electronic switch, the **diode-bridge modulator**, driven by a sinusoid $A \cos \omega_c t$ to produce the switching action. Diodes D_1, D_2 and D_3, D_4 are matched pairs. When the signal $\cos \omega_c t$ is of a polarity that will make terminal c positive with respect to d, all the diodes conduct. Because diodes D_1 and D_2 are matched, terminals a and b have the same potential and are effectively shorted. During the next half-cycle, terminal d is positive with respect to c, and all four diodes open, thus opening terminals a and b. The diode bridge in Fig. 4.5a, therefore, serves as a desired electronic switch, where terminals a and b open and close periodically with carrier frequency f_c when a sinusoid $A \cos \omega_c t$ is applied across terminals c and d. To obtain the signal $m(t)w(t)$, we may place this electronic switch (terminals a and b) in series (Fig. 4.5b) or across (in parallel) $m(t)$, as shown in Fig. 4.5c. These modulators are known as the **series-bridge diode modulator** and the **shunt-bridge diode modulator**, respectively. This switching on and off of $m(t)$ repeats for each cycle of the carrier, resulting in the switched signal $m(t)w(t)$, which when bandpass-filtered, yields the desired modulated signal $(2/\pi)m(t) \cos \omega_c t$.

Another switching modulator, known as the **ring modulator**, is shown in Fig. 4.6a. During the positive half-cycles of the carrier, diodes D_1 and D_3 conduct, and D_2 and D_4 are open. Hence, terminal a is connected to c, and terminal b is connected to d. During the negative half-cycles of the carrier, diodes D_1 and D_3 are open, and D_2 and D_4 are conducting, thus connecting terminal a to d and terminal b to c. Hence, the output is proportional to $m(t)$ during the positive half-cycle and to $-m(t)$ during the negative half-cycle. In effect, $m(t)$ is multiplied by a square pulse train $w_0(t)$, as shown in Fig. 4.6b. The Fourier series for $w_0(t)$ can be found by using the signal $w(t)$ of Eq. (4.5) to yield $w_0(t) = 2w(t) - 1$. Therefore, we can use the Fourier series of $w(t)$ [Eq. (4.5)] to determine the Fourier series of $w_0(t)$ as

$$w_0(t) = \frac{4}{\pi}\left(\cos \omega_c t - \frac{1}{3}\cos 3\omega_c t + \frac{1}{5}\cos 5\omega_c t - \cdots\right) \tag{4.7a}$$

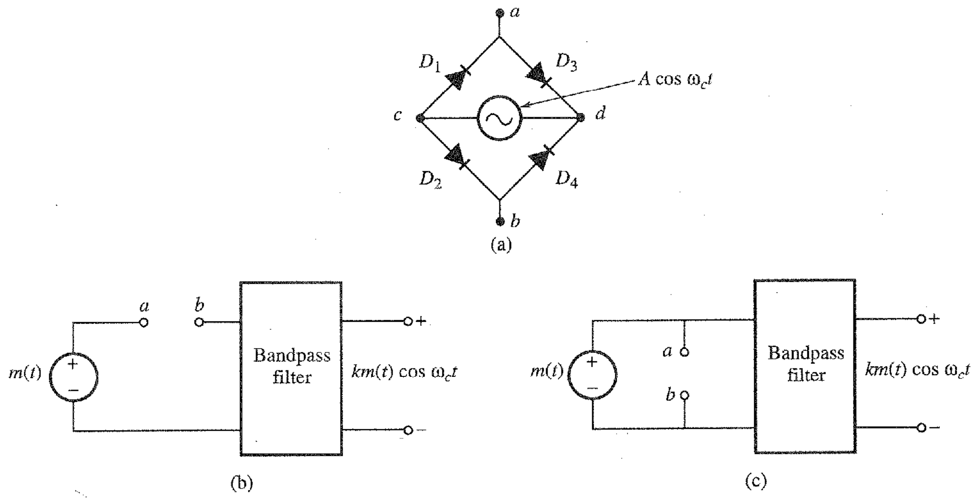


Figure 4.5
(a) Diode-bridge electronic switch.
(b) Series-bridge diode modulator.
(c) Shunt-bridge diode modulator.

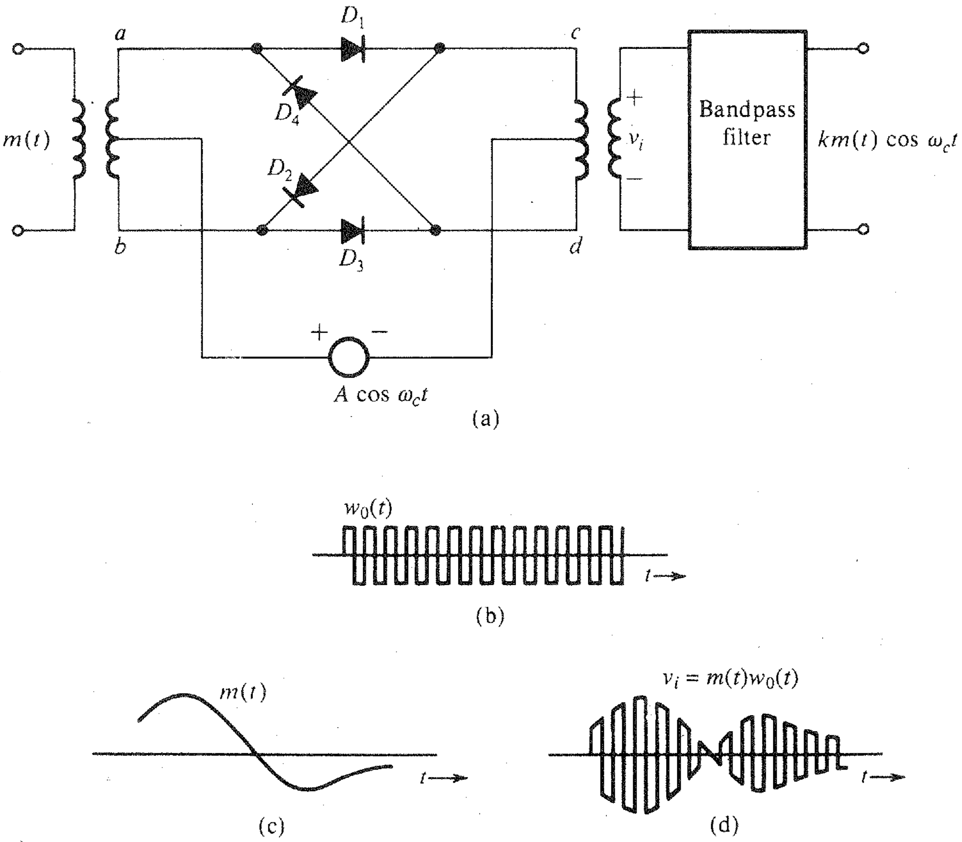


Figure 4.6
Ring modulator.

Hence, we have

$$v_i(t) = m(t)w_0(t) = \frac{4}{\pi}\left[m(t)\cos\ \omega_c t - \frac{1}{3}m(t)\cos\ 3\omega_c t + \frac{1}{5}m(t)\cos\ 5\omega_c t - \cdots\right] \quad (4.7b)$$

The signal $m(t)w_0(t)$ is shown in Fig. 4.6d. When this waveform is passed through a bandpass filter tuned to ω_c (Fig. 4.6a), the filter output will be the desired signal $(4/\pi)m(t)\cos\ \omega_c t$.

In this circuit there are two inputs: $m(t)$ and $\cos\ \omega_c t$. The input to the final bandpass filter does not contain either of these inputs. Consequently, this circuit is an example of a **double balanced modulator**.

Example 4.2 Frequency Mixer or Converter

We shall analyze a frequency mixer, or frequency converter, used to change the carrier frequency of a modulated signal $m(t)\cos\ \omega_c t$ from ω_c to another frequency ω_I.

This can be done by multiplying $m(t)\cos\ \omega_c t$ by $2\cos\ \omega_{\text{mix}} t$, where $\omega_{\text{mix}} = \omega_c + \omega_I$ or $\omega_c - \omega_I$, and then bandpass-filtering the product, as shown in Fig. 4.7a.

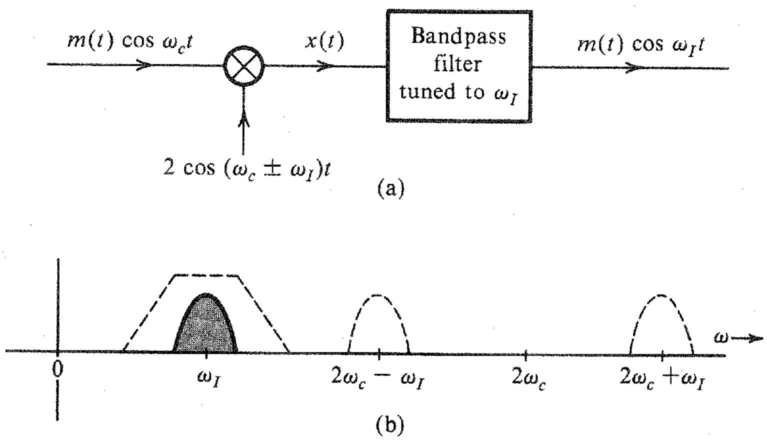


Figure 4.7 Frequency mixer or converter.

The product $x(t)$ is

$$x(t) = 2m(t)\cos\ \omega_c t \cos\ \omega_{\text{mix}} t$$
$$= m(t)[\cos\ (\omega_c - \omega_{\text{mix}})t + \cos\ (\omega_c + \omega_{\text{mix}})t]$$

If we select $\omega_{\text{mix}} = \omega_c - \omega_I$, then

$$x(t) = m(t)[\cos\ \omega_I t + \cos\ (2\omega_c - \omega_I)t]$$

If we select $\omega_{\text{mix}} = \omega_c + \omega_I$, then

$$x(t) = m(t)[\cos\ \omega_I t + \cos\ (2\omega_c + \omega_I)t]$$

In either case, as long as $\omega_c - \omega_I \geq 2\pi B$ and $\omega_I \geq 2\pi B$, the various spectra in Fig. 4.7b will not overlap. Consequently, a bandpass filter at the output, tuned to ω_I, will pass the term $m(t)\cos\ \omega_I t$ and suppress the other term, yielding the output $m(t)\cos\ \omega_I t$. Thus, the carrier frequency has been translated to ω_I from ω_c.

The operation of frequency mixing/conversion (also known as heterodyning) is basically a shifting of spectra by an additional ω_{mix}. This is equivalent to the operation of modulation with a modulating carrier frequency (the mixer oscillator frequency ω_{mix}) that differs from the incoming carrier frequency by ω_I. Any one of the modulators discussed earlier can be used for frequency mixing. When we select the local carrier frequency $\omega_{\text{mix}} = \omega_c + \omega_I$, the operation is called **upconversion**, and when we select $\omega_{\text{mix}} = \omega_c - \omega_I$, the operation is **downconversion**.

Demodulation of DSB-SC Signals

As discussed earlier, demodulation of a DSB-SC signal essentially involves multiplication by the carrier signal and is identical to modulation (see Fig. 4.1). At the receiver, we multiply the incoming signal by a local carrier of frequency and phase in synchronism with the incoming carrier. The product is then passed through a low-pass filter. The **only difference** between the modulator and the demodulator lies in the input signal and the output filter. In the modulator, message $m(t)$ is the input while the multiplier output is passed through a bandpass filter tuned to ω_c, whereas in the demodulator, the DSB-SC signal is the input while the multiplier output is passed through a low-pass filter. Therefore, all the modulators discussed earlier without multipliers can also be used as demodulators, provided the bandpass filters at the output are replaced by low-pass filters of bandwidth B.

For demodulation, the receiver must generate a carrier in phase and frequency synchronism with the incoming carrier. These demodulators are synonymously called **synchronous** or **coherent** (also **homodyne**) demodulators.

Example 4.3 Analyze the switching demodulator that uses the electronic switch (diode bridge) in Fig. 4.5a as a switch (either in series or in parallel).

The input signal is $m(t) \cos \omega_c t$. The carrier causes the periodic switching on and off of the input signal. Therefore, the output is $m(t) \cos \omega_c t \times w(t)$. Using the identity $\cos x \cos y = 0.5[\cos (x+y) + \cos (x-y)]$, we obtain

$$\begin{aligned} m(t) \cos \omega_c t \times w(t) &= m(t) \cos \omega_c t \left[\frac{1}{2} + \frac{2}{\pi}\left(\cos \omega_c t - \frac{1}{3}\cos 3\omega_c t + \cdots\right)\right] \\ &= \frac{2}{\pi} m(t) \cos^2 \omega_c t + \text{terms of the form } m(t) \cos n\omega_c t \\ &= \frac{1}{\pi} m(t) + \frac{1}{\pi} m(t) \cos 2\omega_c t + \text{terms of the form } m(t) \cos n\omega_c t \end{aligned}$$

Spectra of the terms of the form $m(t) \cos n\omega_c t$ are centered at $\pm n\omega_c$ and are filtered out by the low-pass filter, yielding the output $(1/\pi)m(t)$. It is left as an exercise for the reader to show that the output of the ring circuit in Fig. 4.6a operating as a demodulator (with the low-pass filter at the output) is $(2/\pi)m(t)$ (twice that of the switching demodulator in this example).

4.3 AMPLITUDE MODULATION (AM)

In the last section, we began our discussion of amplitude modulation by introducing the DSB-SC amplitude modulation because it is easy to understand and to analyze in both the time and frequency domains. However, analytical simplicity does not always equate to simplicity in practical implementation. The (coherent) demodulation of a DSB-SC signal requires the receiver to possess a carrier signal that is synchronized with the incoming carrier. This requirement is not easy to achieve in practice. Because the modulated signal may have traveled hundreds of miles and could even suffer from some unknown frequency shift, the bandpass received signal in fact has the form of

$$r(t) = A_c m(t - t_0) \cos [(\omega_c + \Delta\omega)(t - t_0)] = A_c m(t - t_0) \cos [(\omega_c + \Delta\omega)t - \theta_d)]$$

in which $\Delta\omega$ represents the Doppler effect while

$$\theta_d = (\omega_c + \Delta\omega)t_d$$

comes from the unknown delay t_0. To utilize the coherent demodulator, the receiver must be sophisticated enough to generate a local oscillator $\cos [(\omega_c + \Delta\omega)t - \theta_d)]$ purely from the received signal $r(t)$. Such a receiver would be harder to implement and could be quite costly. This cost is particularly to be avoided in broadcasting systems, which have many receivers for every transmitter.

The alternative to a coherent demodulator is for the transmitter to send a carrier $A \cos \omega_c t$ [along with the modulated signal $m(t) \cos \omega_c t$] so that there is no need to generate a carrier at the receiver. In this case the transmitter needs to transmit at a much higher power level, which increases its cost as a trade-off. In point-to-point communications, where there is one transmitter for every receiver, substantial complexity in the receiver system can be justified, provided its cost is offset by a less expensive transmitter. On the other hand, for a broadcast system with a huge number of receivers for each transmitter, it is more economical to have one expensive high-power transmitter and simpler, less expensive receivers because any cost saving at the receiver is multiplied by the number of receiver units. For this reason, broadcasting systems tend to favor the trade-off by migrating cost from the (many) receivers to the (fewer) transmitters.

The second option of transmitting a carrier along with the modulated signal is the obvious choice in broadcasting because of its desirable trade-offs. This leads to the so-called AM (amplitude modulation), in which the transmitted signal $\varphi_{AM}(t)$ is given by

$$\varphi_{AM}(t) = A \cos \omega_c t + m(t) \cos \omega_c t \tag{4.8a}$$

$$= [A + m(t)] \cos \omega_c t \tag{4.8b}$$

The spectrum of $\varphi_{AM}(t)$ is basically the same as that of $\varphi_{DSB-SC}(t) = m(t) \cos \omega_c t$ except for the two additional impulses at $\pm f_c$,

$$\varphi_{AM}(t) \iff \frac{1}{2}[M(f + f_c) + M(f - f_c)] + \frac{A}{2}[\delta(f + f_c) + \delta(f - f_c)] \tag{4.8c}$$

Upon comparing $\varphi_{AM}(t)$ with $\varphi_{DSB-SC}(t) = m(t) \cos \omega_c t$, it is clear that the AM signal is identical to the DSB-SC signal with $A + m(t)$ as the modulating signal [instead of $m(t)$]. The value of A is always chosen to be positive. Therefore, to sketch $\varphi_{AM}(t)$, we sketch the envelope $|A + m(t)|$ and its mirror image $-|A + m(t)|$ and fill in between with the sinusoid of the carrier frequency f_c. The size of A affects the time domain envelope of the modulated signal.

Two cases are considered in Fig. 4.8. In the first case, A is large enough that $A + m(t) \geq 0$ is always nonnegative. In the second case, A is not large enough to satisfy this condition. In the first case, the envelope has the same shape as $m(t)$ (although riding on a dc of magnitude A). In the second case, the envelope shape differs from the shape of $m(t)$ because the negative

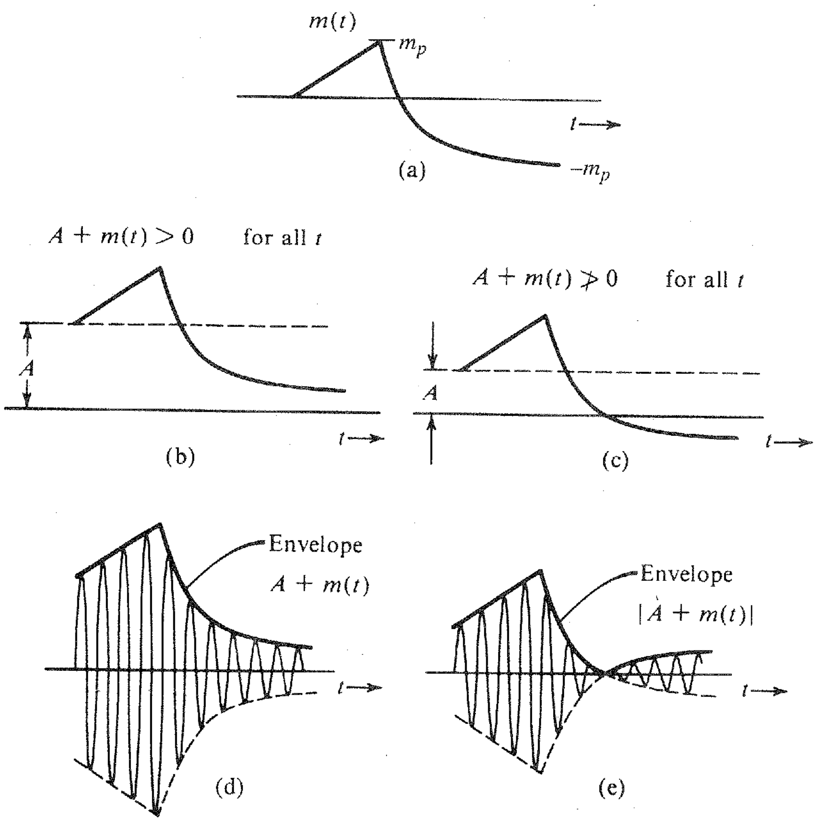


Figure 4.8 AM signal and its envelope.

part of $A + m(t)$ is rectified. This means we can detect the desired signal $m(t)$ by detecting the envelope in the first case when $A + m(t) > 0$. Such detection is not possible in the second case. We shall see that envelope detection is an extremely simple and inexpensive operation, which does not require generation of a local carrier for the demodulation. But as seen earlier, the AM envelope has the information about $m(t)$ only if the AM signal $[A + m(t)] \cos \omega_c t$ satisfies the condition $A + m(t) > 0$ for all t.

Let us now be more precise about the definition of "envelope." Consider a signal $E(t) \cos \omega_c t$. If $E(t)$ varies slowly in comparison with the sinusoidal carrier $\cos \omega_c t$, then the **envelope** of $E(t) \cos \omega_c t$ is $|E(t)|$. This means [see Eq. (4.8b)] that if and only if $A + m(t) \geq 0$ for all t, the envelope of $\varphi_{AM}(t)$ is

$$|A + m(t)| = A + m(t)$$

In other words, for envelope detection to properly detect $m(t)$, two conditions must be met:

(a) $f_c \gg$ bandwidth of $m(t)$

(b) $A + m(t) \geq 0$

This conclusion is readily verified from Fig. 4.8d and e. In Fig. 4.8d, where $A + m(t) \geq 0$, $A + m(t)$ is indeed the envelope, and $m(t)$ can be recovered from this envelope. In Fig. 4.8e, where $A + m(t)$ is not always positive, the envelope $|A + m(t)|$ is rectified from $A + m(t)$, and $m(t)$ cannot be recovered from the envelope. Consequently, demodulation of $\varphi_{AM}(t)$ in Fig. 4.8d amounts to simple envelope detection. Thus, the **condition for envelope detection** of an AM signal is

$$A + m(t) \geq 0 \qquad \text{for all } t \tag{4.9a}$$

If $m(t) \geq 0$ for all t, then $A = 0$ already satisfies condition (4.9a). In this case there is no need to add any carrier because the envelope of the DSB-SC signal $m(t) \cos \omega_c t$ is $m(t)$ and such a DSB-SC signal can be detected by envelope detection. In the following discussion we assume that $m(t) \not\geq 0$ for all t; that is, $m(t)$ can be negative over some range of t.

Message Signals $m(t)$ with Zero Offset: Let $\pm m_p$ be the maximum and the minimum values of $m(t)$, respectively (see Fig. 4.8). This means that $m(t) \geq -m_p$. Hence, the condition of envelope detection (4.9a) is equivalent to

$$A \geq -m_{\text{min}} \tag{4.9b}$$

Thus, the minimum carrier amplitude required for the viability of envelope detection is m_p. This is quite clear from Fig. 4.8. We define the modulation index μ as

$$\mu = \frac{m_p}{A} \tag{4.10a}$$

For envelope detection to be distortionless, the condition is $A \geq m_p$. Hence, it follows that

$$0 \leq \mu \leq 1 \tag{4.10b}$$

is the required condition for the distortionless demodulation of AM by an envelope detector.

When $A < m_p$, Eq. (4.10a) shows that $\mu > 1$ (overmodulation). In this case, the option of envelope detection is no longer viable. We then need to use synchronous demodulation. Note that synchronous demodulation can be used for any value of μ, since the demodulator will recover signal $A + m(t)$. Only an additional dc block is needed to remove the DC voltage A. The envelope detector, which is considerably simpler and less expensive than the synchronous detector, can be used only for $\mu \leq 1$.

Message Signals $m(t)$ with Nonzero Offset: On rare occasions, the message signal $m(t)$ will have a nonzero offset such that its maximum m_{max} and its minimum m_{min} are not symmetric, that is,

$$m_{\text{min}} \neq -m_{\text{max}}$$

In this case, it can be recognized that any offset to the envelope does not change the shape of the envelope detector output. In fact, one should note that constant offset does not carry any fresh information.

In this case, envelope detection would still remain distortionless if

$$0 \leq \mu \leq 1 \tag{4.11a}$$

with a modified modulation index definition of

$$\mu = \frac{m_{\text{max}} - m_{\text{min}}}{2A + m_{\text{max}} + m_{\text{min}}} \tag{4.11b}$$

Example 4.4 Sketch $\varphi_{\text{AM}}(t)$ for modulation indices of $\mu = 0.5$ and $\mu = 1$, when $m(t) = b \cos \omega_m t$. This case is referred to as **tone modulation** because the modulating signal is a pure sinusoid (or tone).

In this case, $m_{\max} = b$ and $m_{\min} = -b$. Hence the modulation index according to Eq. (4.10a) is

$$\mu = \frac{b - (-b)}{2A + b + (-b)} = \frac{b}{A}$$

Hence, $b = \mu A$ and

$$m(t) = b \cos \omega_m t = \mu A \cos \omega_m t$$

Therefore,

$$\varphi_{\text{AM}}(t) = [A + m(t)] \cos \omega_c t = A[1 + \mu \cos \omega_m t] \cos \omega_c t$$

Figure 4.9 shows the modulated signals corresponding to $\mu = 0.5$ and $\mu = 1$, respectively.

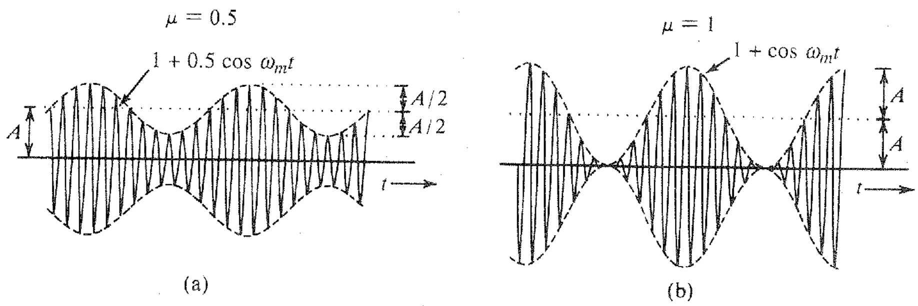


Figure 4.9 Tone-modulated AM. (a) $\mu = 0.5$. (b) $\mu = 1$.

Sideband and Carrier Power

The advantage of envelope detection in AM comes at a price. In AM, the carrier term does not carry any information, and hence, the carrier power is wasteful from this point of view:

$$\varphi_{\text{AM}}(t) = \underbrace{A \cos \omega_c t}_{\text{carrier}} + \underbrace{m(t) \cos \omega_c t}_{\text{sidebands}}$$

The carrier power P_c is the mean square value of $A \cos \omega_c t$, which is $A^2/2$. The sideband power P_s is the power of $m(t) \cos \omega_c t$, which is $0.5\ \widetilde{m^2(t)}$ [see Eq. (3.93)]. Hence,

$$P_c = \frac{A^2}{2} \quad \text{and} \quad P_s = \frac{1}{2}\ \widetilde{m^2(t)}$$

The useful message information resides in the sideband power, whereas the carrier power is the used for convenience in modulation and demodulation. The total power is the sum of the carrier (wasted) power and the sideband (useful) power. Hence, η, the power efficiency, is

$$\eta = \frac{\text{useful power}}{\text{total power}} = \frac{P_s}{P_c + P_s} = \frac{\widetilde{m^2(t)}}{A^2 + \widetilde{m^2(t)}} 100\%$$

For the special case of tone modulation,

$$m(t) = \mu A \cos \omega_m t \quad \text{and} \quad \widetilde{m^2(t)} = \frac{(\mu A)^2}{2}$$

Hence

$$\eta = \frac{\mu^2}{2+\mu^2} 100\%$$

with the condition that $0 \leq \mu \leq 1$. It can be seen that η increases monotonically with μ, and η_{max} occurs at $\mu = 1$, for which

$$\eta_{max} = 33\%$$

Thus, for tone modulation, under the best conditions ($\mu = 1$), only one-third of the transmitted power is used for carrying messages. For practical signals, the efficiency is even worse—on the order of 25% or lower—compared with the DSB-SC case. The best condition implies $\mu = 1$. Smaller values of μ degrade efficiency further. For this reason, volume compression and peak limiting are commonly used in AM to ensure that full modulation ($\mu = 1$) is maintained most of the time.

Example 4.5 Determine η and the percentage of the total power carried by the sidebands of the AM wave for tone modulation when $\mu = 0.5$ and when $\mu = 0.3$.

For $\mu = 0.5$,

$$\eta = \frac{\mu^2}{2+\mu^2} 100\% = \frac{(0.5)^2}{2+(0.5)^2} 100\% = 11.11\%$$

Hence, only about 11% of the total power is in the sidebands. For $\mu = 0.3$,

$$\eta = \frac{(0.3)^2}{2+(0.3)^2} 100\% = 4.3\%$$

Hence, only 4.3% of the total power is in the sidebands that contain the message signal.

Generation of AM Signals

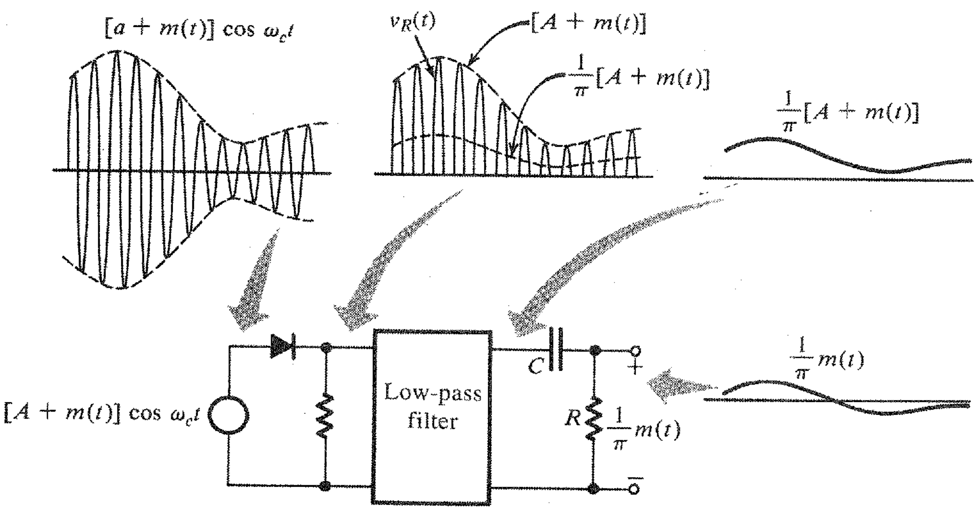
In principle, the generation of AM signals is identical to that of the DSB-SC modulations discussed in Sec. 4.2 except that an additional carrier component $A \cos \omega_c t$ needs to be added to the DSB-SC signal.

Demodulation of AM Signals

Like DSB-SC signals, the AM signal can be demodulated coherently by a locally generated carrier. Coherent, or synchronous, demodulation of AM, however, defeats the purpose of AM because it does not take advantage of the additional carrier component $A \cos \omega_c t$. As we have seen earlier, in the case of $\mu \leq 1$, the envelope of the AM signal follows the message signal $m(t)$. Hence, we shall consider here two noncoherent methods of AM demodulation under the condition of $0 < \mu \leq 1$: rectifier detection and envelope detection.

Rectifier Detector: If an AM signal is applied to a diode and a resistor circuit (Fig. 4.10), the negative part of the AM wave will be removed. The output across the resistor is a half-wave-rectified version of the AM signal. Visually, the diode acts like a pair of scissors by cutting off any negative half-cycle of the modulated sinusoid. In essence, at the rectifier output, the AM

Figure 4.10 Rectifier detector for AM.



signal is multiplied by $w(t)$. Hence, the half-wave-rectified output $v_R(t)$ is

$$v_R(t) = \{[A + m(t)] \cos \omega_c t\}\, w(t) \tag{4.12}$$

$$= [A + m(t)] \cos \omega_c t \left[\frac{1}{2} + \frac{2}{\pi}\left(\cos \omega_c t - \frac{1}{3} \cos 3\omega_c t + \frac{1}{5} \cos 5\omega_c t - \cdots\right)\right] \tag{4.13}$$

$$= \frac{1}{\pi}[A + m(t)] + \text{other terms of higher frequencies} \tag{4.14}$$

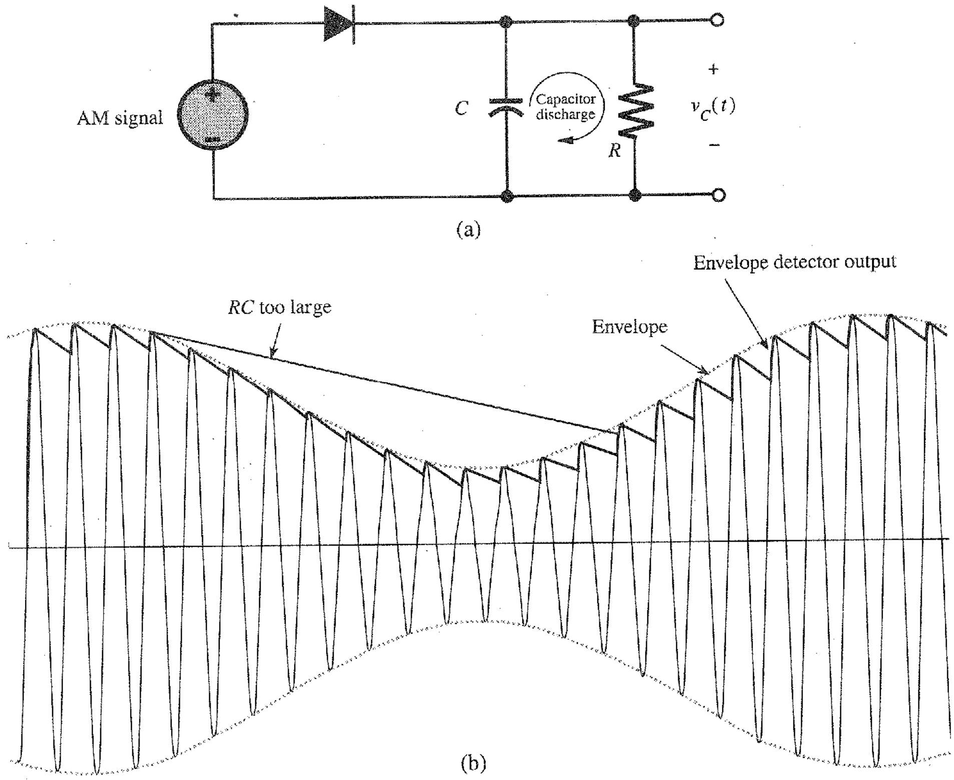
When $v_R(t)$ is applied to a low-pass filter of cutoff B Hz, the output is $[A + m(t)]/\pi$, and all the other terms in v_R of frequencies higher than B Hz are suppressed. The dc term A/π may be blocked by a capacitor (Fig. 4.10) to give the desired output $m(t)/\pi$. The output can be doubled by using a full-wave rectifier.

It is interesting to note that because of the multiplication with $w(t)$, rectifier detection is in effect synchronous detection performed without using a local carrier. The high carrier content in AM ensures that its zero crossings are periodic and the information about the frequency and phase of the carrier at the transmitter is built in to the AM signal itself.

Envelope Detector: The output of an envelope detector follows the envelope of the modulated signal. The simple circuit shown in Fig. 4.11a functions as an envelope detector. On the positive cycle of the input signal, the input grows and may exceed the charged voltage on the capacity $v_C(t)$, turning on the diode and allowing the capacitor C to charge up to the peak voltage of the input signal cycle. As the input signal falls below this peak value, it falls quickly below the capacitor voltage (which is very nearly the peak voltage), thus causing the diode to open. The capacitor now discharges through the resistor R at a slow rate (with a time constant RC). During the next positive cycle, the same drama repeats. As the input signal rises above the capacitor voltage, the diode conducts again. The capacitor again charges to the peak value of this (new) cycle. The capacitor discharges slowly during the cutoff period.

During each positive cycle, the capacitor charges up to the peak voltage of the input signal and then decays slowly until the next positive cycle as shown in Fig. 4.11b. The output voltage $v_C(t)$, thus, closely follows the (rising) envelope of the input AM signal. Equally important, the slow capacity discharge via the resistor R allows the capacity voltage to follow a declining

Figure 4.11
Envelope detector for AM.



envelope. Capacitor discharge between positive peaks causes a ripple signal of frequency ω_c in the output. This ripple can be reduced by choosing a larger time constant RC so that the capacitor discharges very little between the positive peaks ($RC \gg 1/\omega_c$). Picking RC too large, however, would make it impossible for the capacitor voltage to follow a fast-declining envelope (see Fig. 4.11b). Because the maximum rate of AM envelope decline is dominated by the bandwidth B of the message signal $m(t)$, the design criterion of RC should be

$$1/\omega_c \ll RC < 1/(2\pi B) \qquad \text{or} \qquad 2\pi B < \frac{1}{RC} \ll \omega_c$$

The envelope detector output is $v_C(t) = A + m(t)$ with a ripple of frequency ω_c. The dc term A can be blocked out by a capacitor or a simple RC high-pass filter. The ripple may be reduced further by another (low-pass) RC filter.

4.4 BANDWIDTH-EFFICIENT AMPLITUDE MODULATIONS

As seen from Fig. 4.12, the DSB spectrum (including suppressed carrier and AM) has two sidebands: the upper sideband (USB) and the lower sideband (LSB), each containing the complete information of the baseband signal $m(t)$. As a result, for a baseband signal $m(t)$ with bandwidth B Hz, DSB modulations require twice the radio-frequency bandwidth to transmit. To improve the spectral efficiency of amplitude modulation, there exist two basic schemes to either utilize or remove the 100% spectral redundancy:

- Single-sideband (SSB) modulation, which removes either the LSB or the USB that uses only bandwidth of B Hz for one message signal $m(t)$;
- Quadrature amplitude modulation (QAM), which utilizes the spectral redundancy by sending two messages over the same bandwidth of $2B$ Hz.

Amplitude Modulation: Single Sideband (SSB)

As shown in Fig. 4.13, either the LSB or the USB can be suppressed from the DSB signal via bandpass filtering. Such a scheme in which only one sideband is transmitted is known as **single-sideband (SSB) transmission**, and requires only one-half the bandwidth of the DSB signal.

Figure 4.12 (a) Original message spectrum. (b) The redundant bandwidth consumption in DSB modulations.

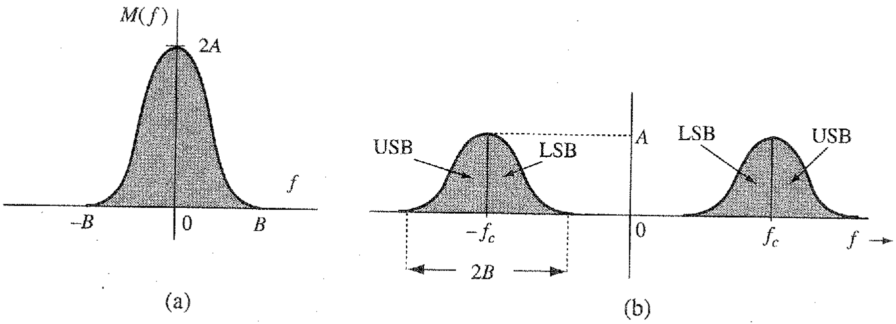
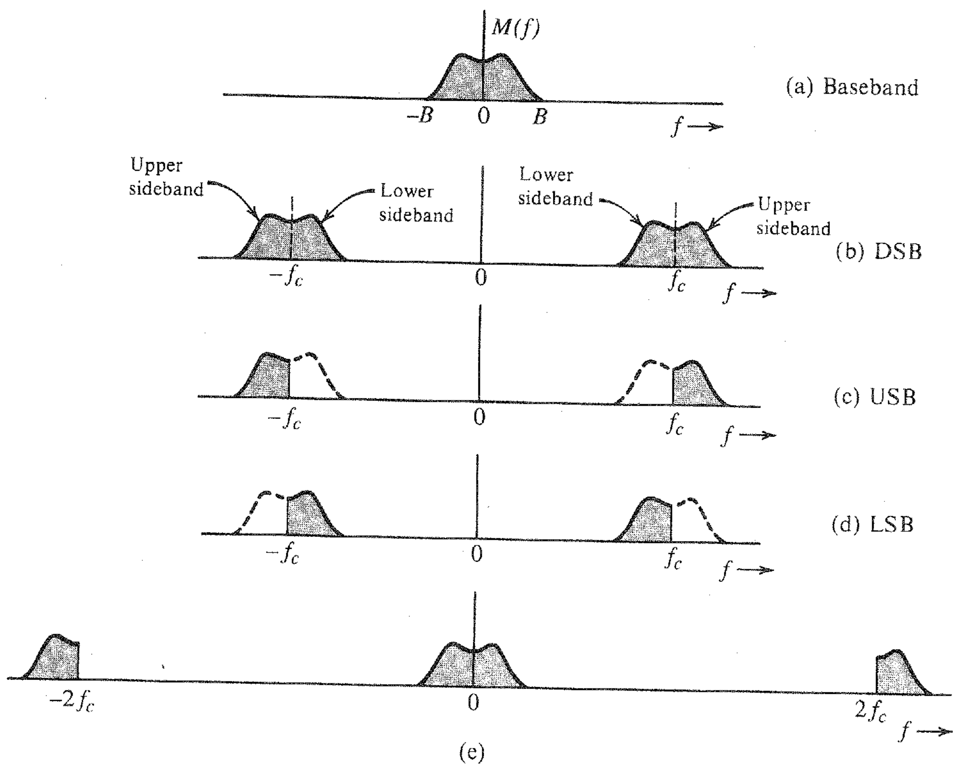


Figure 4.13 SSB spectra from suppressing one DSB sideband.



An SSB signal can be coherently (synchronously) demodulated just like DSB-SC signals. For example, multiplication of a USB signal (Fig. 4.13c) by cos $\omega_c t$ shifts its spectrum to the left and right by ω_c, yielding the spectrum in Fig. 4.13e. Low-pass filtering of this signal yields the desired baseband signal. The case is similar with LSB signals. Since the demodulation of SSB signals is identical to that of DSB-SC signals, the transmitters can now utilize only half the DSB-SC signal bandwidth without any additional cost to the receivers. Since no additional carrier accompanies the modulated SSB signal, the resulting modulator outputs are known as suppressed carrier signals (SSB-SC).

Hilbert Transform

We now introduce for later use a new tool known as the **Hilbert transform**. We use $x_h(t)$ and $\mathcal{H}\{x(t)\}$ to denote the Hilbert transform of signal $x(t)$

$$x_h(t) = \mathcal{H}\{x(t)\} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\alpha)}{t-\alpha}\, d\alpha \tag{4.15}$$

Observe that the right-hand side of Eq. (4.15) has the form of a convolution

$$x(t) * \frac{1}{\pi t}$$

Now, application of the duality property to pair 12 of Table 3.1 yields $1/\pi t \Longleftrightarrow -j\,\text{sgn}\,(f)$. Hence, application of the time convolution property to the convolution (of Eq. (4.15) yields

$$X_h(f) = -jX(f)\,\text{sgn}\,(f) \tag{4.16}$$

From Eq. (4.16), it follows that if $m(t)$ passes through a transfer function $H(f) = -j\,\text{sgn}\,(f)$, then the output is $m_h(t)$, the Hilbert transform of $m(t)$. Because

$$H(f) = -j\,\text{sgn}\,(f) \tag{4.17}$$

$$= \begin{cases} -j = 1 \cdot e^{-j\pi/2} & f > 0 \\ j = 1 \cdot e^{j\pi/2} & f < 0 \end{cases} \tag{4.18}$$

it follows that $|H(f)| = 1$ and that $\theta_h(f) = -\pi/2$ for $f > 0$ and $\pi/2$ for $f < 0$, as shown in Fig. 4.14. Thus, if we change the phase of every component of $m(t)$ by $\pi/2$ (without changing its amplitude), the resulting signal is $m_h(t)$, the Hilbert transform of $m(t)$. Therefore, a Hilbert transformer is an ideal phase shifter that shifts the phase of every spectral component by $-\pi/2$.

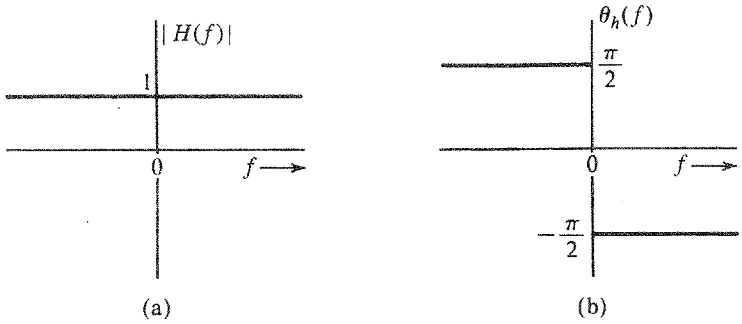
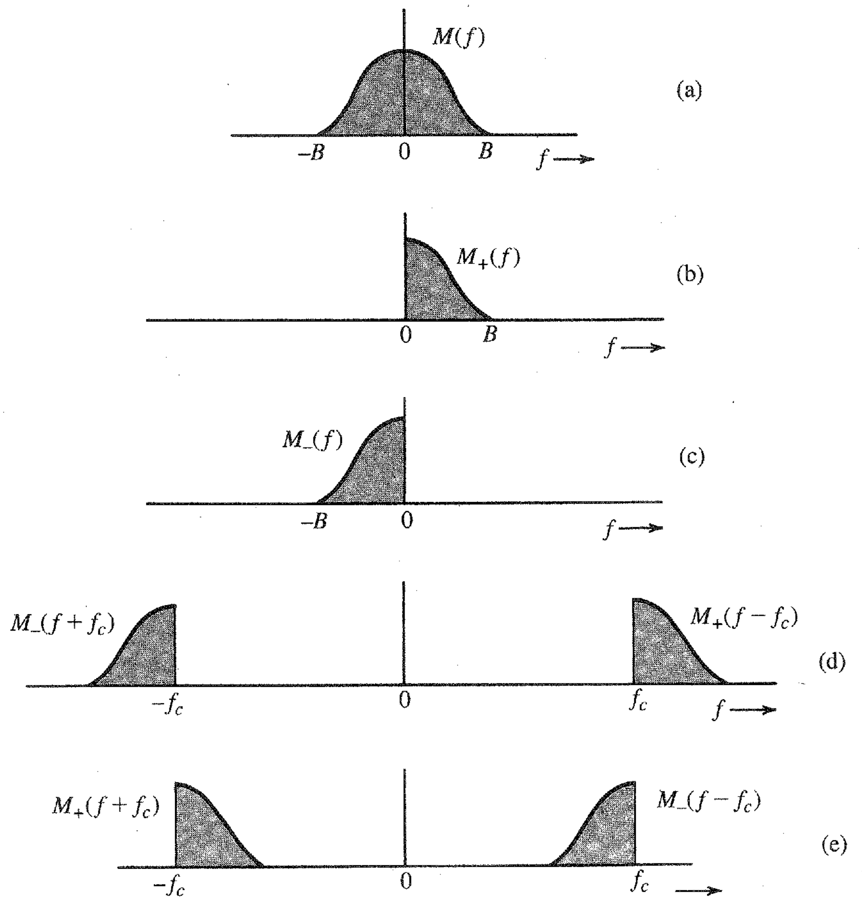


Figure 4.14 Transfer function of an ideal $\pi/2$ phase shifter (Hilbert transformer).

Figure 4.15 Expressing SSB spectra in terms of $M_+(f)$ and $M_-(f)$.



Time Domain Representation of SSB Signals

Because the building blocks of an SSB signal are the sidebands, we shall first obtain a time domain expression for each sideband.

Figure 4.15a shows the message spectrum $M(f)$. Figure 4.15b shows its right half $M_+(f)$, and Fig. 4.15c shows its left half $M_-(f)$. From Fig. 4.15b and c, we observe that

$$M_+(f) = M(f) \cdot u(f) = M(f)\frac{1}{2}[1 + \text{sgn}(f)] = \frac{1}{2}[M(f) + jM_h(f)] \tag{4.19a}$$

$$M_-(f) = M(f)u(-f) = M(f)\frac{1}{2}[1 - \text{sgn}(f)] = \frac{1}{2}[M(f) - jM_h(f)] \tag{4.19b}$$

We can now express the SSB signal in terms of $m(t)$ and $m_h(t)$. From Fig. 4.15d it is clear that the USB spectrum $\Phi_{\text{USB}}(f)$ can be expressed as

$$\begin{aligned}\Phi_{\text{USB}}(f) &= M_+(f - f_c) + M_-(f + f_c)\\ &= \frac{1}{2}[M(f - f_c) + M(f + f_c)] - \frac{1}{2j}[M(f - f_c) - M(f + f_c)]\end{aligned}$$

From the frequency-shifting property, the inverse transform of this equation yields

$$\varphi_{\text{USB}}(t) = m(t)\cos\,\omega_c t - m_h(t)\sin\,\omega_c t \tag{4.20a}$$

Similarly, we can show that

$$\varphi_{\text{LSB}}(t) = m(t)\cos\,\omega_c t + m_h(t)\sin\,\omega_c t \tag{4.20b}$$

Hence, a general SSB signal $\varphi_{\text{SSB}}(t)$ can be expressed as

$$\varphi_{\text{SSB}}(t) = m(t)\cos\,\omega_c t \mp m_h(t)\sin\omega_c t \tag{4.20c}$$

where the minus sign applies to USB and the plus sign applies to LSB.

Given the time domain expression of SSB-SC signals, we can now confirm analytically (instead of graphically) that SSB-SC signals can be coherently demodulated:

$$\begin{aligned}\varphi_{\text{SSB}}(t)\cos\,\omega_c t &= [m(t)\cos\,\omega_c t \mp m_h(t)\sin\omega_c t]\,2\cos\,\omega_c t \\ &= m(t)[1+\cos\,2\omega_c t] \mp m_h(t)\sin\,2\omega_c t \\ &= m(t) + \underbrace{[m(t)\cos\,2\omega_c t \mp m_h(t)\sin\,2\omega_c t]}_{\text{SSB-SC signal with carrier } 2\omega_c}\end{aligned}$$

Thus, the product $\varphi_{\text{SSB}}(t)\cdot 2\cos\,\omega_c t$ yields the baseband signal and another SSB signal with a carrier $2\omega_c$. The spectrum in Fig. 4.13e shows precisely this result. A low-pass filter will suppress the unwanted SSB terms, giving the desired baseband signal $m(t)$. Hence, the demodulator is identical to the synchronous demodulator used for DSB-SC. Thus, any one of the synchronous DSB-SC demodulators discussed earlier in Sec. 4.2 can be used to demodulate an SSB-SC signal.

Example 4.6 Tone Modulation: SSB

Find $\varphi_{\text{SSB}}(t)$ for a simple case of a tone modulation, that is, when the modulating signal is a sinusoid $m(t) = \cos\,\omega_m t$. Also demonstrate the coherent demodulation of this SSB signal.

Recall that the Hilbert transform delays the phase of each spectral component by $\pi/2$. In the present case, there is only one spectral component of frequency ω_m. Delaying the phase of $m(t)$ by $\pi/2$ yields

$$m_h(t) = \cos\left(\omega_m t - \frac{\pi}{2}\right) = \sin\,\omega_m t$$

Hence, from Eq. (4.20c),

$$\begin{aligned}\varphi_{\text{SSB}}(t) &= \cos\,\omega_m t\cos\,\omega_c t \mp \sin\,\omega_m t\sin\,\omega_c t \\ &= \cos\,(\omega_c \pm \omega_m)t\end{aligned}$$

Thus,

$$\varphi_{\text{USB}}(t) = \cos\,(\omega_c + \omega_m)t \qquad \text{and} \qquad \varphi_{\text{LSB}}(t) = \cos\,(\omega_c - \omega_m)t$$

To verify these results, consider the spectrum of $m(t)$ (Fig. 4.16a) and its DSB-SC (Fig. 4.16b), USB (Fig. 4.16c), and LSB (Fig. 4.16d) spectra. It is evident that the spectra in Fig. 4.16c and d do indeed correspond to the $\varphi_{\text{USB}}(t)$ and $\varphi_{\text{LSB}}(t)$ derived earlier.

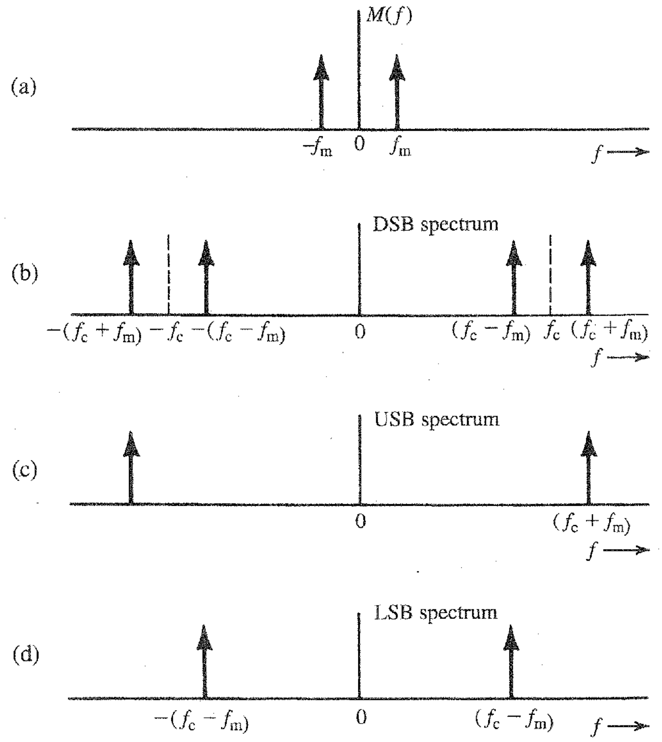


Figure 4.16 SSB spectra for tone modulation.

Finally, the coherent demodulation of the SSB tone modulation is can be achieved by

$$\begin{aligned}\varphi_{\rm SSB}(t)2\cos\,\omega_c t &= 2\cos\,(\omega_c \pm \omega_m)t\cos\,\omega_c t\\ &= \cos\,\omega_m t + \cos\,(\omega_c + \omega_m)t\end{aligned}$$

which can be sent to a lowpass filter to retrieve the message tone $\cos\,\omega_m t$.

SSB Modulation Systems

Three methods are commonly used to generate SSB signals: phase shifting, selective filtering, and the Weaver method.[1] None of these modulation methods are precise, and all generally require that the baseband signal spectrum have little power near the origin.

The **phase shift method** directly uses Eq. (4.20) as its basis. Figure 4.17 shows its implementation. The box marked "$-\pi/2$" is a phase shifter, which delays the phase of every positive spectral component by $\pi/2$. Hence, it is a Hilbert transformer. Note that an ideal Hilbert phase shifter is unrealizable. This is because the Hilbert phase shifter requires an abrupt phase change of π at zero frequency. When the message $m(t)$ has a dc null and very little low-frequency content, the practical approximation of this ideal phase shifter has almost no real effect and does not affect the accuracy of SSB modulation.

In the **selective-filtering method**, the most commonly used method of generating SSB signals, a DSB-SC signal is passed through a sharp cutoff filter to eliminate the undesired sideband. To obtain the USB, the filter should pass all components above frequency f_c unattenuated and completely suppress all components below f_c. Such an operation requires an ideal filter, which is unrealizable. It can, however, be approximated closely if there is some separation

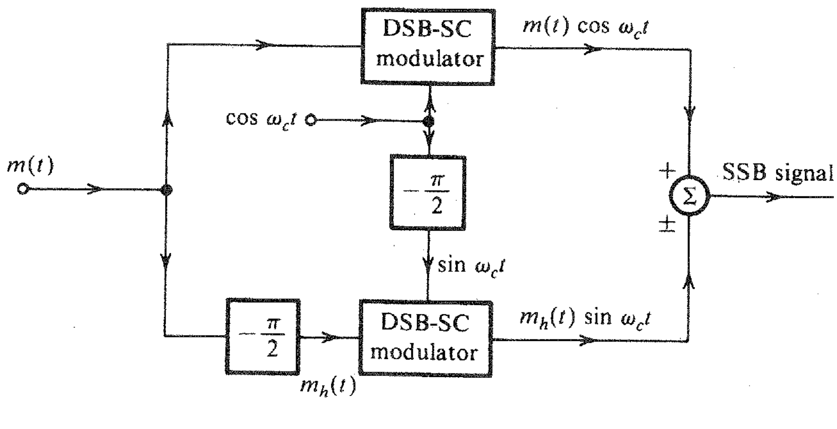


Figure 4.17 Generating SSB using the phase shift method.

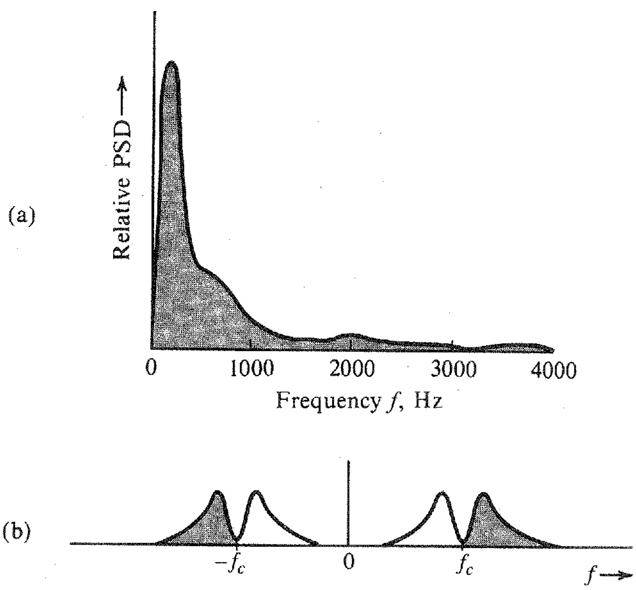


Figure 4.18 (a) Relative power spectrum of speech signal and (b) corresponding USB spectrum.

between the passband and the stopband. Fortunately, the voice signal provides this condition, because its spectrum shows little power content at the origin (Fig. 4.18a). In addition, articulation tests have shown that for speech signals, frequency components below 300 Hz are not important. In other words, we may suppress all speech components below 300 Hz (and above 3500 Hz) without affecting intelligibility appreciably. Thus, filtering of the unwanted sideband becomes relatively easy for speech signals because we have a 600 Hz transition region around the cutoff frequency f_c. To minimize adjacent channel interference, the undesired sideband should be attenuated at least 40 dB.

For very high carrier frequency f_c, the ratio of the gap band (600 Hz) to the carrier frequency may be too small, and, thus, a transition of 40 dB in amplitude over 600 Hz may be difficult. In such a case, a third method, known as **Weaver's method**,[1] utilizes two stages of SSB amplitude modulation. First, the modulation is carried out by using a smaller carrier frequency (f_{c_1}). The resulting SSB signal effectively widens the gap to $2f_{c_1}$ (see shaded spectra in Fig. 4.18b). Now by treating this signal as the new baseband signal, it is possible to achieve SSB-modulation at a higher carrier frequency.