$$|\hat{D}_n| = |D_n| \quad \text{and} \quad \angle\hat{D}_n = \angle D_n - n\omega_0 T$$

This result shows that time shifting of a periodic signal by $T$ seconds merely changes the phase spectrum by $n\omega_0 T$. The amplitude spectrum is unchanged.
**(b)** Show that the exponential Fourier series for $\tilde{f}(t) = f(at)$ is given by

$$\tilde{f}(t) = \sum_{n=-\infty}^{\infty} D_n e^{jn(a\omega_0)t}$$

This result shows that time compression of a periodic signal by a factor $a$ expands its Fourier spectra by the same factor $a$. Similarly, time expansion of a periodic signal by a factor $a$ compresses its Fourier spectra by the factor $a$.

**3.5-5** **(a)** The Fourier series for the periodic signal in Fig. 3.10a is given in Exercise E3.6. Verify Parseval's theorem for this series, given that

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$$

**(b)** If $f(t)$ is approximated by the first $N$ terms in this series, find $N$ so that the power of the error signal is less than 1% of $P_f$.

**3.5-6** **(a)** The Fourier series for the periodic signal in Fig. 3.10b is given in Exercise E3.6. Verify Parseval's theorem for this series, given that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

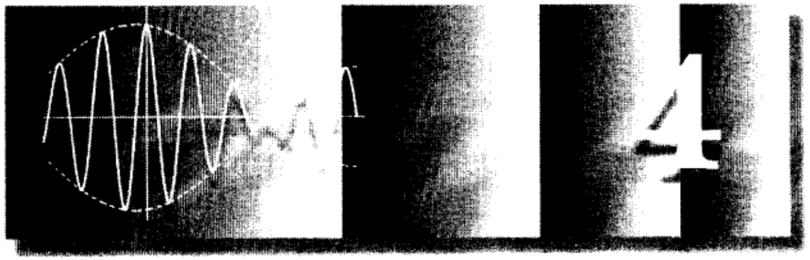
**(b)** If $f(t)$ is approximated by the first $N$ terms in this series, find $N$ so that the power of the error signal is less than 10% of $P_f$.

**3.5-7** The signal $f(t)$ in Fig. 3.18 is approximated by the first $2N+1$ terms (from $n = -N$ to $N$) in its exponential Fourier series given in Exercise E3.10. Determine the value of $N$ if this $(2N+1)$-term Fourier series power is to be no less than 99.75% of the power of $f(t)$.

**3.6-1** Find the response of an LTIC system with transfer function

$$H(s) = \frac{s}{s^2 + 2s + 3}$$

to the periodic input shown in Fig. 3.7b.




# Continuous-Time Signal Analysis: The Fourier Transform

In Chapter 3, we succeeded in representing periodic signals as a sum of (everlasting) sinusoids or exponentials. In this chapter we extend this spectral representation to aperiodic signals.

## 4.1 Aperiodic Signal Representation by Fourier Integral

Applying a limiting process, we now show that an aperiodic signal can be expressed as a continuous sum (integral) of everlasting exponentials. To represent an aperiodic signal $f(t)$ such as the one depicted in Fig. 4.1a by everlasting exponential signals, let us construct a new periodic signal $f_{T_0}(t)$ formed by repeating the signal $f(t)$ at intervals of $T_0$ seconds, as illustrated in Fig. 4.1b. The period $T_0$ is made long enough to avoid overlap between the repeating pulses. The periodic signal $f_{T_0}(t)$ can be represented by an exponential Fourier series. If we let $T_0 \to \infty$, the pulses in the periodic signal repeat after an infinite interval and, therefore

$$\lim_{T_0 \to \infty} f_{T_0}(t) = f(t)$$

Thus, the Fourier series representing $f_{T_0}(t)$ will also represent $f(t)$ in the limit $T_0 \to \infty$. The exponential Fourier series for $f_{T_0}(t)$ is given by

$$f_{T_0}(t) = \sum_{n=-\infty}^{\infty} D_n e^{jn\omega_0 t} \tag{4.1}$$

where

$$D_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} f_{T_0}(t) e^{-jn\omega_0 t}\, dt \tag{4.2a}$$

and

$$\omega_0 = \frac{2\pi}{T_0} \tag{4.2b}$$

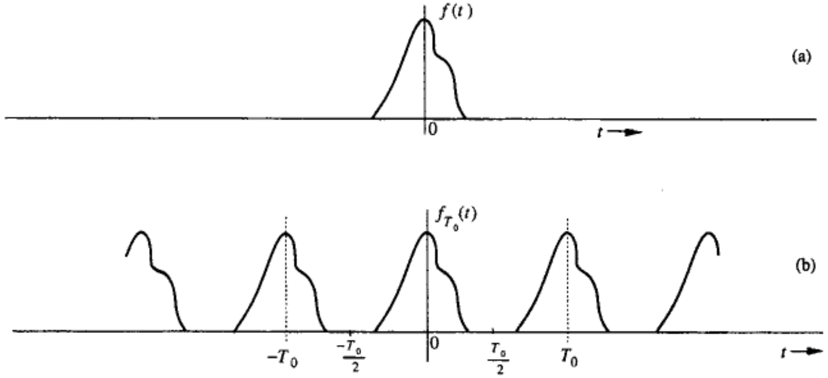



**Fig. 4.1** Construction of a periodic signal by periodic extension of $f(t)$.

Observe that integrating $f_{T_0}(t)$ over $(-\frac{T_0}{2}, \frac{T_0}{2})$ is the same as integrating $f(t)$ over $(-\infty, \infty)$. Therefore, Eq. (4.2a) can be expressed as

$$D_n = \frac{1}{T_0} \int_{-\infty}^{\infty} f(t) e^{-jn\omega_0 t}\, dt \tag{4.2c}$$

It is interesting to see how the nature of the spectrum changes as $T_0$ increases. To understand this fascinating behavior, let us define $F(\omega)$, a continuous function of $\omega$, as

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t}\, dt \tag{4.3}$$

A glance at Eqs. (4.2c) and (4.3) shows that

$$D_n = \frac{1}{T_0} F(n\omega_0) \tag{4.4}$$

This means that the Fourier coefficients $D_n$ are $(1/T_0)$ times the samples of $F(\omega)$ uniformly spaced at intervals of $\omega_0$, as depicted in Fig. 4.2a.† Therefore, $(1/T_0)F(\omega)$ is the envelope for the coefficients $D_n$. We now let $T_0 \to \infty$ by doubling $T_0$ repeatedly. Doubling $T_0$ halves the fundamental frequency $\omega_0$ [Eq. (4.2b)], so that there are now twice as many components (samples) in the spectrum. However, by doubling $T_0$, the envelope $(1/T_0)\ F(\omega)$ is halved, as shown in Fig. 4.2b. If we continue this process of doubling $T_0$ repeatedly, the spectrum progressively becomes denser while its magnitude becomes smaller. Note, however, that the relative shape of the envelope remains the same [proportional to $F(\omega)$ in Eq. (4.3)]. In the limit as $T_0 \to \infty$, $\omega_0 \to 0$ and $D_n \to 0$. This result means the spectrum is so dense that the spectral components are spaced at zero (infinitesimal) intervals. At the same time, the amplitude of each component is zero (infinitesimal). We have *nothing of everything, yet we have something!* This paradox sounds like *Alice in Wonderland*,

†For the sake of simplicity, we assume $D_n$, and therefore $F(\omega)$, in Fig. 4.2, to be real. The argument, however, is also valid for complex $D_n$ [or $F(\omega)$].



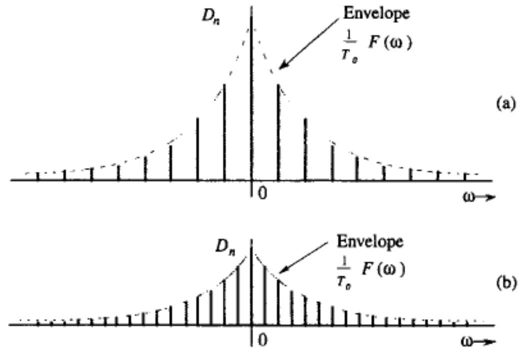



**Fig. 4.2** Change in the Fourier spectrum when the period $T_0$ in Fig. 4.1 is doubled.

but as we shall see, these are the classic characteristics of a very familiar phenomenon.†

Substitution of Eq. (4.4) in Eq. (4.1) yields

$$f_{T_0}(t) = \sum_{n=-\infty}^{\infty} \frac{F(n\omega_0)}{T_0} e^{jn\omega_0 t} \tag{4.5}$$

As $T_0 \to \infty$, $\omega_0$ becomes infinitesimal ($\omega_0 \to 0$). Hence, we shall replace $\omega_0$ by a more appropriate notation, $\Delta\omega$. In terms of this new notation, Eq. (4.2b) becomes

$$\Delta\omega = \frac{2\pi}{T_0}$$

and Eq. (4.5) becomes

$$f_{T_0}(t) = \sum_{n=-\infty}^{\infty} \left[\frac{F(n\Delta\omega)\Delta\omega}{2\pi}\right] e^{(jn\Delta\omega)t} \tag{4.6a}$$

Equation (4.6a) shows that $f_{T_0}(t)$ can be expressed as a sum of everlasting exponentials of frequencies $0, \pm\Delta\omega, \pm 2\Delta\omega, \pm 3\Delta\omega, \ldots$. (the Fourier series). The amount of the component of frequency $n\Delta\omega$ is $[F(n\Delta\omega)\Delta\omega]/2\pi$. In the limit as $T_0 \to \infty, \Delta\omega \to 0$ and $f_{T_0}(t) \to f(t)$. Therefore

$$f(t) = \lim_{T_0 \to \infty} f_{T_0}(t) = \lim_{\Delta\omega \to 0} \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(n\Delta\omega) e^{(jn\Delta\omega)t} \Delta\omega \tag{4.6b}$$

The sum on the right-hand side of Eq. (4.6b) can be viewed as the area under the function $F(\omega)e^{j\omega t}$, as illustrated in Fig. 4.3. Therefore

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \tag{4.7}$$

The integral on the right hand side is called the **Fourier integral**. We have now succeeded in representing an aperiodic signal $f(t)$ by a Fourier integral (rather

†If nothing else, the reader now has an irrefutable proof of the proposition that 0% ownership of everything is better than 100% ownership of nothing.

than a Fourier series).† This integral is basically a Fourier series (in the limit) with fundamental frequency $\Delta\omega \to 0$, as seen from Eq. (4.6). The amount of the exponential $e^{jn\Delta\omega t}$ is $F(n\Delta\omega)\Delta\omega/2\pi$. Thus, the function $F(\omega)$ given by Eq. (4.3) acts as a spectral function.

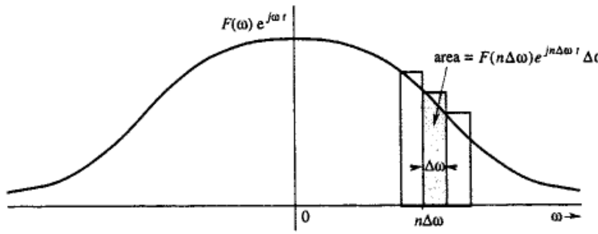



**Fig. 4.3** The Fourier series becomes the Fourier integral in the limit as $T_0 \to \infty$.

We call $F(\omega)$ the **direct** Fourier transform of $f(t)$, and $f(t)$ the **inverse** Fourier transform of $F(\omega)$. The same information is conveyed by the statement that $f(t)$ and $F(\omega)$ are a Fourier transform pair. Symbolically, this statement is expressed as

$$F(\omega) = \mathcal{F}[f(t)] \qquad \text{and} \qquad f(t) = \mathcal{F}^{-1}[F(\omega)]$$

or

$$f(t) \Longleftrightarrow F(\omega)$$

To recapitulate,

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}\,dt \tag{4.8a}$$

and

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}\,d\omega \tag{4.8b}$$

It is helpful to keep in mind that the Fourier integral in Eq. (4.8b) is of the nature of a Fourier series with fundamental frequency $\Delta\omega$ approaching zero [Eq. (4.6b)]. Therefore, most of the discussion and properties of Fourier series apply to the Fourier transform as well. We can plot the spectrum $F(\omega)$ as a function of $\omega$. Since $F(\omega)$ is complex, we have both amplitude and angle (or phase) spectra

$$F(\omega) = |F(\omega)|e^{j\angle F(\omega)} \tag{4.9}$$

in which $|F(\omega)|$ is the amplitude and $\angle F(\omega)$ is the angle (or phase) of $F(\omega)$. According to Eq. (4.8a),

$$F(-\omega) = \int_{-\infty}^{\infty} f(t)e^{j\omega t}\,dt$$

From this equation and Eq. (4.8a), it follows that if $f(t)$ is a real function of $t$, then $F(\omega)$ and $F(-\omega)$ are complex conjugates. Therefore

†This derivation should not be considered as a rigorous proof of Eq. (4.7). The situation is not as simple as we have made it appear.[1]



$$|F(-\omega)| = |F(\omega)| \tag{4.10a}$$

$$\angle F(-\omega) = -\angle F(\omega) \tag{4.10b}$$

Thus, for real $f(t)$, the amplitude spectrum $|F(\omega)|$ is an even function, and the phase spectrum $\angle F(\omega)$ is an odd function of $\omega$. This property the (**conjugate symmetry property**) is valid only for real $f(t)$. These results were derived earlier for the Fourier spectrum of a periodic signal [Eq. (3.77)] and should come as no surprise. *The transform $F(\omega)$ is the frequency-domain specification of $f(t)$.*

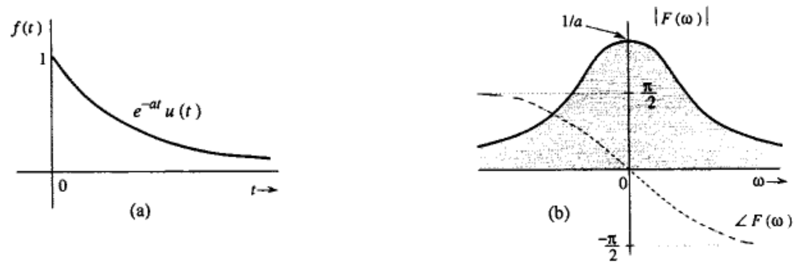



**Fig. 4.4** $e^{-at}u(t)$ and its Fourier spectra.

■ **Example 4.1**

Find the Fourier transform of $e^{-at}u(t)$.

By definition [Eq. (4.8a)],

$$F(\omega) = \int_{-\infty}^{\infty} e^{-at}u(t)e^{-j\omega t}\,dt = \int_{0}^{\infty} e^{-(a+j\omega)t}\,dt = \frac{-1}{a+j\omega}e^{-(a+j\omega)t}\Big|_0^{\infty}$$

But $|e^{-j\omega t}| = 1$. Therefore, as $t \to \infty$, $e^{-(a+j\omega)t} = e^{-at}e^{-j\omega t} = 0$ if $a > 0$. Therefore

$$F(\omega) = \frac{1}{a+j\omega} \qquad a > 0 \tag{4.11a}$$

Expressing $a + j\omega$ in the polar form as $\sqrt{a^2+\omega^2}\,e^{j\tan^{-1}(\frac{\omega}{a})}$, we obtain

$$F(\omega) = \frac{1}{\sqrt{a^2+\omega^2}}e^{-j\tan^{-1}(\frac{\omega}{a})} \tag{4.11b}$$

Therefore

$$|F(\omega)| = \frac{1}{\sqrt{a^2+\omega^2}} \qquad \text{and} \qquad \angle F(\omega) = -\tan^{-1}\left(\frac{\omega}{a}\right) \tag{4.12}$$

The amplitude spectrum $|F(\omega)|$ and the phase spectrum $\angle F(\omega)$ are depicted in Fig. 4.4b. Observe that $|F(\omega)|$ is an even function of $\omega$, and $\angle F(\omega)$ is an odd function of $\omega$, as expected. ■

### Existence of the Fourier Transform

In Example 4.1 we observed that when $a < 0$, the Fourier integral for $e^{-at}u(t)$ does not converge. Hence, the Fourier transform for $e^{-at}u(t)$ does not exist if $a < 0$ (growing exponential). Clearly, not all signals are Fourier-transformable. The existence of the Fourier transform is assured for any $f(t)$ satisfying the Dirichlet conditions mentioned on p. 194-195. The first of these conditions is†

$$\int_{-\infty}^{\infty} |f(t)|\, dt < \infty \tag{4.13}$$

Because $|e^{-j\omega t}| = 1$, from Eq. (4.8a), we obtain

$$|F(\omega)| \le \int_{-\infty}^{\infty} |f(t)|\, dt$$

This inequality shows that the existence of the Fourier transform is assured if condition (4.13) is satisfied. Otherwise, there is no guarantee. We have seen in Example 4.1 that the Fourier transform does not exist for an exponentially growing signal (which violates this condition). Although this condition is sufficient, it is not necessary for the existence of the Fourier transform of a signal. For example, the signal $(\sin at)/t$ violates condition (4.13), but does have a Fourier transform. Any signal that can be generated in practice satisfies the Dirichlet conditions and therefore has a Fourier transform. Thus, the physical existence of a signal is a sufficient condition for the existence of its transform.

#### Linearity of the Fourier Transform

The Fourier transform is linear; that is, if

$$f_1(t) \Longleftrightarrow F_1(\omega) \qquad \text{and} \qquad f_2(t) \Longleftrightarrow F_2(\omega)$$

then

$$a_1 f_1(t) + a_2 f_2(t) \Longleftrightarrow a_1 F_1(\omega) + a_2 F_2(\omega) \tag{4.14}$$

The proof is trivial and follows directly from Eq. (4.8a). This result can be extended to any finite number of terms.

### 4.1-1 Physical Appreciation of the Fourier Transform

In understanding any aspect of the Fourier transform, we should remember that Fourier representation is a way of expressing a signal in terms of everlasting sinusoids (or exponentials). The Fourier spectrum of a signal indicates the relative amplitudes and phases of the sinusoids that are required to synthesize that signal. A periodic signal Fourier spectrum has finite amplitudes and exists at discrete frequencies ($\omega_0$ and its multiples). Such a spectrum is easy to visualize, but the spectrum of an aperiodic signal is not easy to visualize because it has a continuous spectrum. The continuous spectrum concept can be appreciated by considering an analogous, more tangible phenomenon. One familiar example of a continuous distribution is the loading of a beam. Consider a beam loaded with weights $D_1, D_2, D_3, \ldots, D_n$ units at the uniformly spaced points $x_1, x_2, \ldots, x_n$, as shown in Fig. 4.5a. The total load $W_T$ on the beam is given by the sum of these loads at each of the $n$ points:

†The remaining Dirichlet conditions are as follows: in any finite interval, $f(t)$ may have only a finite number of maxima and minima and a finite number of finite discontinuities. When these conditions are satisfied, the Fourier integral on the right-hand side of Eq. (4.8b) converges to $f(t)$ at all points where $f(t)$ is continuous and converges to the average of the right-hand and left-hand limits of $f(t)$ at points where $f(t)$ is discontinuous.



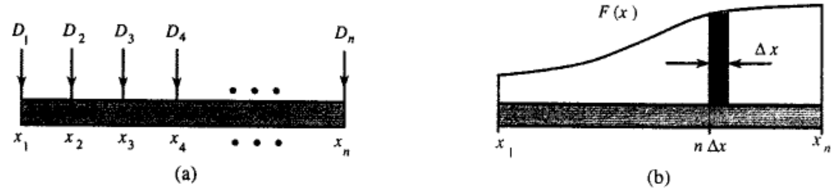



**Fig. 4.5** Analogy for Fourier transform.

$$W_T = \sum_{i=1}^{n} D_i$$

Consider now the case of a continuously loaded beam, as depicted in Fig. 4.5b. In this case, although there appears to be a load at every point, the load at any one point is zero. This does not mean that there is no load on the beam. A meaningful measure of load in this situation is not the load at a point, but rather the loading density per unit length at that point. Let $F(x)$ be the loading density per unit length of beam. It then follows that the load over a beam length $\Delta x$ ($\Delta x \to 0$), at some point $x$, is $F(x)\Delta x$. To find the total load on the beam, we divide the beam into segments of interval $\Delta x$ ($\Delta x \to 0$). The load over the $n$th such segment of length $\Delta x$ is $F(n\Delta x)\Delta x$. The total load $W_T$ is given by

$$\begin{aligned} W_T &= \lim_{\Delta x \to 0} \sum_{x_1}^{x_n} F(n\Delta x)\, \Delta x \\ &= \int_{x_1}^{x_n} F(x)\, dx \end{aligned}$$

In the case of discrete loading (Fig. 4.5a), the load exists only at the $n$ discrete points. At other points there is no load. On the other hand, in the continuously loaded case, the load exists at every point, but at any specific point $x$, the load is zero. The load over a small interval $\Delta x$, however, is $[F(n\Delta x)]\,\Delta x$ (Fig. 4.5b). Thus, even though the load at a point $x$ is zero, the relative load at that point is $F(x)$.

An exactly analogous situation exists in the case of a signal spectrum. When $f(t)$ is periodic, the spectrum is discrete, and $f(t)$ can be expressed as a sum of discrete exponentials with finite amplitudes:

$$f(t) = \sum_n D_n e^{jn\omega_0 t}$$

For an aperiodic signal, the spectrum becomes continuous; that is, the spectrum exists for every value of $\omega$, but the amplitude of each component in the spectrum is zero. The meaningful measure here is not the amplitude of a component of some frequency but the spectral density per unit bandwidth. From Eq. (4.6b) it is clear that $f(t)$ is synthesized by adding exponentials of the form $e^{jn\Delta\omega t}$, in which the contribution by any one exponential component is zero. But the contribution by exponentials in an infinitesimal band $\Delta\omega$ located at $\omega = n\Delta\omega$ is $\frac{1}{2\pi}F(n\Delta\omega)\Delta\omega$, and the addition of all these components yields $f(t)$ in the integral form:

$$f(t) = \lim_{\Delta\omega \to 0} \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(n\Delta\omega) e^{(jn\Delta\omega)t} \Delta\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t}\, d\omega \tag{4.15}$$

The contribution by components within a band $d\omega$ is $\frac{1}{2\pi}F(\omega)\,d\omega = F(\omega)\,d\mathcal{F}$, where $d\mathcal{F}$ is the bandwidth in hertz. Clearly, $F(\omega)$ is the **spectral density** per unit bandwidth (in hertz). It also follows that even if the amplitude of any one component is zero, the relative amount of a component of frequency $\omega$ is $F(\omega)$. Although $F(\omega)$ is a spectral density, in practice it is customarily called the **spectrum** of $f(t)$ rather than the spectral density of $f(t)$. Deferring to this convention, we shall call $F(\omega)$ the Fourier spectrum (or Fourier transform) of $f(t)$.

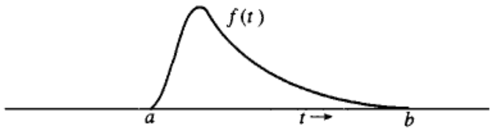



**Fig. 4.6** Marvel of the Fourier transform.

### A Marvelous Balancing Act

An important point to remember here is that $f(t)$ is represented (or synthesized) by exponentials or sinusoids that are everlasting (not causal). Such conceptualization leads to a rather fascinating picture when we try to visualize the synthesis of a timelimited pulse signal $f(t)$ (Fig. 4.6) by the sinusoidal components in its Fourier spectrum. The signal $f(t)$ exists only over an interval $(a, b)$ and is zero outside this interval. The spectrum of $f(t)$ contains an infinite number of exponentials (or sinusoids) which start at $t = -\infty$ and continue forever. The amplitudes and phases of these components are such that they add up exactly to $f(t)$ over the finite interval $(a, b)$ and add up to zero everywhere outside this interval. Juggling with amplitudes and phases of an infinite number of components to achieve such a perfect and delicate balance boggles the human imagination. Yet the Fourier transform accomplishes it routinely, without much thinking on our part. Indeed, we become so involved in mathematical manipulations that we fail to notice this marvel.

### A Word About Notation

In Chapter 2 [Eq. (2.48)], we defined the system transfer function $H(s)$ as

$$H(s) = \int_{-\infty}^{\infty} h(t) e^{-st}\, dt \tag{4.16}$$

Setting $s = j\omega$ in this equation yields

$$H(j\omega) = \int_{-\infty}^{\infty} h(t) e^{-j\omega t}\, dt \tag{4.17}$$

The right-hand side is the Fourier transform of $h(t)$, and according to our notation introduced in Eq. (4.3) this is $H(\omega)$, whereas the same entity is denoted by $H(j\omega)$



in the notation introduced in Chapter 2. Thus, to be consistent with the previous notation, we should have denoted the Fourier transform by $F(j\omega)$ rather than $F(\omega)$ in Eq. (4.3). In fact, the notation $F(j\omega)$ for the Fourier transform is often used in the literature. It is, however, a bit clumsy and does not lend itself so easily to manipulation as the notation $F(\omega)$. For this reason we shall continue with the dual notation, while remembering that both $F(\omega)$ and $F(j\omega)$ represent the same entity. This fact is particularly important when we discuss the Laplace transform and filtering in the future, and should be kept in mind throughout the rest of the book. In the same way, we must remember that $H(\omega)$ and $H(j\omega)$ represent the same entity.

## 4.1-2 LTIC System Response Using the Fourier Transform

We wanted to represent a signal $f(t)$ as a sum of (everlasting) exponentials so that we could find a system response to $f(t)$ as a sum of the system's responses to the exponential components of $f(t)$. Consider an asymptotically stable LTIC system with transfer function $H(s)$. The response of this system to everlasting exponential $e^{j\omega t}$ is $H(\omega)e^{j\omega t}$. Such an input-output pair will be denoted by the directed arrow representation as

$$e^{j\omega t} \Longrightarrow H(\omega)e^{j\omega t}$$

Therefore

$$e^{j(n\Delta\omega)t} \Longrightarrow H(n\Delta\omega)e^{j(n\Delta\omega)t}$$

and

$$\left[\frac{F(n\Delta\omega)\Delta\omega}{2\pi}\right] e^{(jn\Delta\omega)t} \Longrightarrow \left[\frac{F(n\Delta\omega)H(n\Delta\omega)\Delta\omega}{2\pi}\right] e^{j(n\Delta\omega)t}$$

Using the linearity property

$$\underbrace{\lim_{\Delta\omega \to 0} \sum_{n=-\infty}^{\infty} \left[\frac{F(n\Delta\omega)\Delta\omega}{2\pi}\right] e^{(jn\Delta\omega)t}}_{\text{input } f(t)} \Longrightarrow \underbrace{\lim_{\Delta\omega \to 0} \sum_{n=-\infty}^{\infty} \left[\frac{F(n\Delta\omega)H(n\Delta\omega)\Delta\omega}{2\pi}\right] e^{j(n\Delta\omega)t}}_{\text{output } y(t)}$$

The left-hand side is the input $f(t)$ [see Eqs. (4.6a) and (4.6b)], and the right-hand side is the response $y(t)$. Thus

$$\begin{aligned} y(t) &= \frac{1}{2\pi} \lim_{\Delta\omega \to 0} \sum_{n=-\infty}^{\infty} F(n\Delta\omega)H(n\Delta\omega)e^{j(n\Delta\omega)t}\Delta\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)H(\omega)e^{j\omega t}\, d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega)e^{j\omega t}\, d\omega \end{aligned} \tag{4.18}$$

where $Y(\omega)$, the Fourier transform of $y(t)$, is given by†

†The relationship (4.19) applies only to the asymptotically stable systems. The reason is that when $s = j\omega$, the integral on the right-hand side of Eq. (2.48) does not converge for unstable systems. Moreover, even for marginally stable systems, that integral does not converges in the ordinary sense, and $H(j\omega)$ [or $H(\omega)$] cannot be obtained by replacing $s$ in $H(s)$ with $j\omega$. As shown in Eq. (4.44b), Eq. (4.19) can be applied to marginally stable systems provided $H(\omega)$ is interpreted as the Fourier transform of $h(t)$ rather than as $H(s)$ with $s$ replaced by $j\omega$.

$$Y(\omega) = F(\omega)H(\omega) \tag{4.19}$$

In conclusion, we showed that for an LTIC system with transfer function $H(s)$, if the input and the output are $f(t)$ and $y(t)$, respectively, and if

$$f(t) \iff F(\omega) \qquad y(t) \iff Y(\omega)$$

then for asymptotically stable systems

$$Y(\omega) = F(\omega)H(\omega)$$

We shall derive this result again later in a more formal way.

The procedure of the frequency-domain method is identical to that of the time-domain method. In the time-domain case we express the input $f(t)$ as a sum of its impulse components; in the frequency-domain case, the input is expressed as a sum of everlasting exponentials (or sinusoids). In the former case, the response $y(t)$ obtained by summing the system's responses to impulse components results in the convolution integral; in the latter case, the response obtained by summing the system's response to everlasting exponential components results in the Fourier integral. These ideas can be expressed mathematically as follows:

1 For the time-domain case

$\delta(t) \implies h(t)$ the impulse response of the system is $h(t)$

$f(t) = \int_{-\infty}^{\infty} f(x)\delta(t-x)\,dx$ expresses $f(t)$ as a sum of impulse components

and

$y(t) = \int_{-\infty}^{\infty} f(x)h(t-x)\,dx$ expresses $y(t)$ as a sum of responses to impulse components

2 For the frequency-domain case

$e^{j\omega t} \implies H(\omega)e^{j\omega t}$ the system response to $e^{j\omega t}$ is $H(\omega)e^{j\omega t}$

$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}\,d\omega$ shows $f(t)$ as a sum of everlasting exponential components

and

$y(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)H(\omega)e^{j\omega t}\,d\omega$ $y(t)$ is a sum of responses to exponential components

The frequency-domain view sees a system in terms of its frequency response (system response to various sinusoidal components). It views a signal as a sum of various sinusoidal components. Transmission of a signal through a (linear) system is viewed as transmission of various sinusoidal components of the signal through the system.



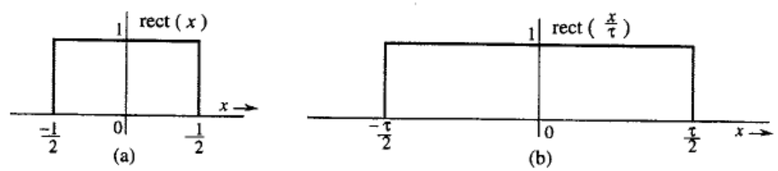



**Fig. 4.7** A gate pulse.

## 4.2 Transforms of Some Useful Functions

For convenience, we now introduce a compact notation for some useful functions such as gate, triangle, and interpolation functions.

### Unit Gate Function

We define a unit gate function $\text{rect}(x)$ as a gate pulse of unit height and unit width, centered at the origin, as illustrated in Fig. 4.7a:†

$$\text{rect}(x) = \begin{cases} 0 & |x| > \frac{1}{2} \\ \frac{1}{2} & |x| = \frac{1}{2} \\ 1 & |x| < \frac{1}{2} \end{cases} \tag{4.20}$$

The gate pulse in Fig. 4.7b is the unit gate pulse $\text{rect}(x)$ expanded by a factor $\tau$ and therefore can be expressed as $\text{rect}(\frac{x}{\tau})$ (see Sec. 1.3-2). Observe that $\tau$, the denominator of the argument of $\text{rect}(\frac{x}{\tau})$, indicates the width of the pulse.

### Unit Triangle Function

We define a unit triangle function $\Delta(x)$ as a triangular pulse of unit height and unit width, centered at the origin, as shown in Fig. 4.8a

$$\Delta(x) = \begin{cases} 0 & |x| \geq \frac{1}{2} \\ 1-2|x| & |x| < \frac{1}{2} \end{cases} \tag{4.21}$$

The pulse in Fig. 4.8b is $\Delta(\frac{x}{\tau})$. Observe that here, as for the gate pulse, the denominator $\tau$ of the argument of $\Delta(\frac{x}{\tau})$ indicates the pulse width.

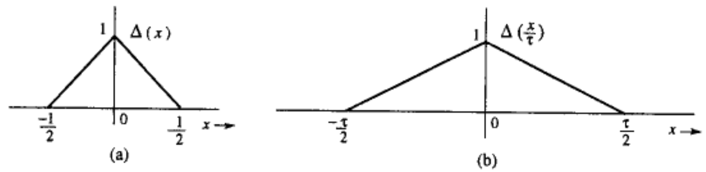



**Fig. 4.8** A triangle pulse.

†At $x = 0$, we require $\text{rect}(x) = 0.5$, because the inverse Fourier transform of a discontinuous signal converges to the mean of its two values at the discontinuity.

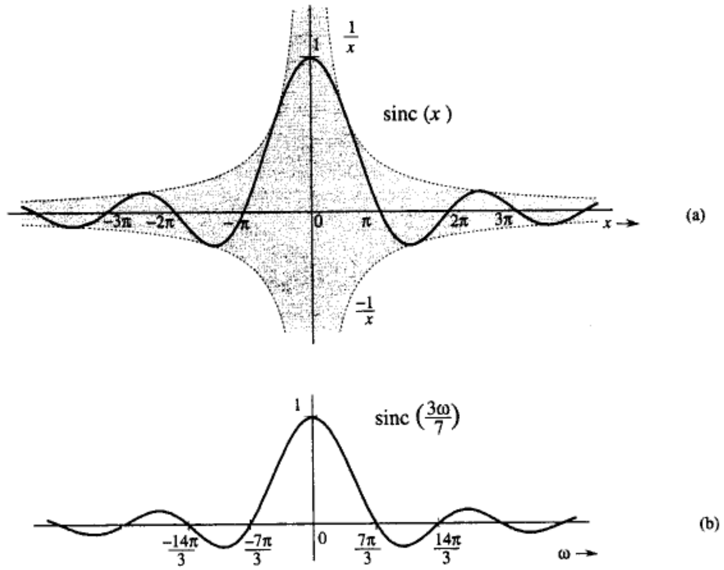



**Fig. 4.9** A sinc pulse.

### Interpolation Function sinc (x)

The function $\sin x/x$ is the "sine over argument" function denoted by $\text{sinc}\,(x)$.† This function plays an important role in signal processing. It is also known as the **filtering or interpolating function**. We define

$$\text{sinc}\,(x) = \frac{\sin x}{x} \tag{4.22}$$

Inspection of Eq. (4.22) shows that

1. $\text{sinc}\,(x)$ is an even function of $x$.
2. $\text{sinc}\,(x) = 0$ when $\sin x = 0$ except at $x = 0$, where it appears indeterminate. This means that $\text{sinc}\, x = 0$ for $x = \pm\pi, \pm 2\pi, \pm 3\pi, \ldots$
3. Using L'Hôpital's rule, we find $\text{sinc}\,(0) = 1$ .
4. $\text{sinc}\,(x)$ is the product of an oscillating signal $\sin x$ (of period $2\pi$) and a monotonically decreasing function $1/x$ . Therefore, $\text{sinc}\,(x)$ exhibits sinusoidal oscillations of period $2\pi$, with amplitude decreasing continuously as $1/x$.

Figure 4.9a shows $\text{sinc}\,(x)$. Observe that $\text{sinc}\,(x) = 0$ for values of $x$ that are positive and negative integral multiples of $\pi$. Figure 4.9b shows $\text{sinc}\,(\frac{3\omega}{7})$. The argument $\frac{3\omega}{7} = \pi$ when $\omega = \frac{7\pi}{3}$. Therefore, the first zero of this function occurs at $\omega = \frac{7\pi}{3}$.

†$\text{sinc}\,(x)$ is also denoted by $\text{Sa}\,(x)$ in the literature. Some authors define $\text{sinc}\,(x)$ as

$$\text{sinc}\,(x) = \frac{\sin \pi x}{\pi x}$$



△ **Exercise E4.1**

Sketch: **(a)** $\text{rect}\,(\frac{x}{8})$ **(b)** $\Delta(\frac{\omega}{10})$ **(c)** $\text{sinc}\,(\frac{3\pi\omega}{2})$ **(d)** $\text{sinc}\,(t)\text{rect}\,(\frac{t}{4\pi})$. ▽

■ **Example 4.2**

Find the Fourier transform of $f(t) = \text{rect}\,(\frac{t}{\tau})$ (Fig. 4.10a).

$$F(\omega) = \int_{-\infty}^{\infty} \text{rect}\left(\frac{t}{\tau}\right) e^{-j\omega t} dt$$

Since $\text{rect}\,(\frac{t}{\tau}) = 1$ for $|t| < \frac{\tau}{2}$, and since it is zero for $|t| > \frac{\tau}{2}$,

$$\begin{aligned} F(\omega) &= \int_{-\tau/2}^{\tau/2} e^{-j\omega t} dt \\ &= -\frac{1}{j\omega}(e^{-j\omega\tau/2} - e^{j\omega\tau/2}) = \frac{2\sin(\frac{\omega\tau}{2})}{\omega} \\ &= \tau \frac{\sin(\frac{\omega\tau}{2})}{(\frac{\omega\tau}{2})} = \tau \;\text{sinc}\left(\frac{\omega\tau}{2}\right) \end{aligned}$$

Therefore

$$\text{rect}\left(\frac{t}{\tau}\right) \Longleftrightarrow \tau \;\text{sinc}\left(\frac{\omega\tau}{2}\right) \tag{4.23}$$

Recall that $\text{sinc}\,(x) = 0$ when $x = \pm n\pi$. Hence, $\text{sinc}\,(\frac{\omega\tau}{2}) = 0$ when $\frac{\omega\tau}{2} = \pm n\pi$; that is, when $\omega = \pm\frac{2n\pi}{\tau}, (n = 1, 2, 3, \ldots)$, as depicted in Fig. 4.10b. The Fourier transform $F(\omega)$ shown in Fig. 4.10b exhibits positive and negative values. A negative amplitude can be considered as a positive amplitude with a phase of $-\pi$ or $\pi$. We use this observation to plot the amplitude spectrum $|F(\omega)| = |\text{sinc}\,(\frac{\omega\tau}{2})|$ (Fig. 4.10c) and the phase spectrum $\angle F(\omega)$ (Fig. 4.10d). The phase spectrum, which is required to be an odd function of $\omega$, may be drawn in several other ways because a negative sign can be accounted for by a phase of $\pm n\pi$, where $n$ is any odd integer. All such representations are equivalent. ■

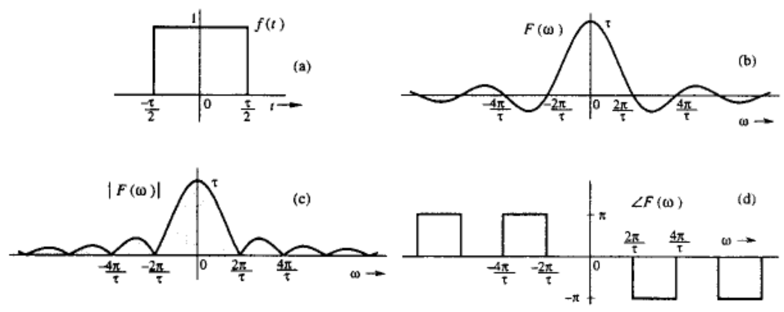



**Fig. 4.10** A gate pulse $f(t)$, its Fourier spectrum $F(\omega)$, amplitude spectrum $|F(\omega)|$, and phase spectrum $\angle F(\omega)$.

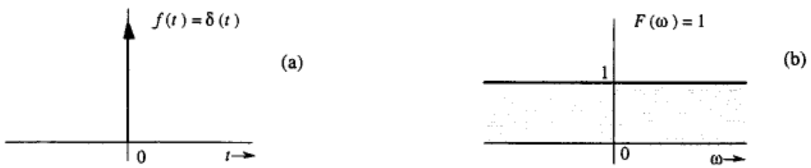



**Fig. 4.11** Unit impulse and its Fourier spectrum.

### Bandwidth of rect $(\frac{t}{\tau})$

The spectrum $F(\omega)$ in Fig. 4.10 peaks at $\omega = 0$ and decays at higher frequencies. Therefore, rect $(\frac{t}{\tau})$ is a lowpass signal with most of the signal energy in lower frequency components. Strictly speaking, because the spectrum extends from 0 to $\infty$, the bandwidth is $\infty$. However, much of the spectrum is concentrated within the first lobe (from $\omega = 0$ to $\omega = \frac{2\pi}{\tau}$). Therefore, a rough estimate of the bandwidth of a rectangular pulse of width $\tau$ seconds is $\frac{2\pi}{\tau}$ rad/s, or $\frac{1}{\tau}$ Hz.† Note the reciprocal relationship of the pulse width with its bandwidth. We shall observe later that this result is true in general.

■ **Example 4.3**

Find the Fourier transform of the unit impulse $\delta(t)$.

Using the sampling property of the impulse [Eq. (1.24)], we obtain

$$\mathcal{F}[\delta(t)] = \int_{-\infty}^{\infty} \delta(t)e^{-j\omega t}dt = 1 \tag{4.24a}$$

or

$$\delta(t) \Longleftrightarrow 1 \tag{4.24b}$$

Figure 4.11 shows $\delta(t)$ and its spectrum. ■

■ **Example 4.4**

Find the inverse Fourier transform of $\delta(\omega)$.

On the basis of Eq. (4.8b) and the sampling property of the impulse function,

$$\mathcal{F}^{-1}[\delta(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \delta(\omega)e^{j\omega t}\,d\omega = \frac{1}{2\pi}$$

Therefore

$$\frac{1}{2\pi} \Longleftrightarrow \delta(\omega) \tag{4.25a}$$

or

$$1 \Longleftrightarrow 2\pi\delta(\omega) \tag{4.25b}$$

This result shows that the spectrum of a constant signal $f(t) = 1$ is an impulse $2\pi\delta(\omega)$, as illustrated in Fig. 4.12.

The result [Eq. (4.25b)] also could have been anticipated on qualitative grounds. Recall that the Fourier transform of $f(t)$ is a spectral representation of $f(t)$ in terms of everlasting exponential components of the form $e^{j\omega t}$. Now, to represent a constant signal

†To compute bandwidth, we must consider the spectrum only for positive values of $\omega$. See discussion on p. 212.



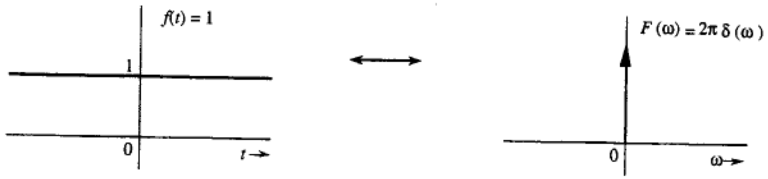



**Fig. 4.12** A constant (dc) signal and its Fourier spectrum.

$f(t) = 1$, we need a single everlasting exponential $e^{j\omega t}$ with $\omega = 0$.‡ This results in a spectrum at a single frequency $\omega = 0$. Another way of looking at the situation is that $f(t) = 1$ is a dc signal which has a single frequency $\omega = 0$ (dc). ■

If an impulse at $\omega = 0$ is a spectrum of a dc signal, what does an impulse at $\omega = \omega_0$ represent? We shall answer this question in the next example.

■ **Example 4.5**

Find the inverse Fourier transform of $\delta(\omega - \omega_0)$.

Using the sampling property of the impulse function, we obtain

$$\mathcal{F}^{-1}[\delta(\omega - \omega_0)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \delta(\omega - \omega_0)e^{j\omega t}\,d\omega = \frac{1}{2\pi}e^{j\omega_0 t}$$

Therefore

$$\frac{1}{2\pi}e^{j\omega_0 t} \Longleftrightarrow \delta(\omega - \omega_0)$$

or

$$e^{j\omega_0 t} \Longleftrightarrow 2\pi\delta(\omega - \omega_0) \tag{4.26a}$$

This result shows that the spectrum of an everlasting exponential $e^{j\omega_0 t}$ is a single impulse at $\omega = \omega_0$. We reach the same conclusion by qualitative reasoning. To represent the everlasting exponential $e^{j\omega_0 t}$, we need a single everlasting exponential $e^{j\omega t}$ with $\omega = \omega_0$. Therefore, the spectrum consists of a single component at frequency $\omega = \omega_0$.

From Eq. (4.26a) it follows that

$$e^{-j\omega_0 t} \Longleftrightarrow 2\pi\delta(\omega + \omega_0) \tag{4.26b}$$

■

■ **Example 4.6**

Find the Fourier transforms of the everlasting sinusoid $\cos \omega_0 t$.

Recall the Euler formula

$$\cos \omega_0 t = \frac{1}{2}(e^{j\omega_0 t} + e^{-j\omega_0 t})$$

‡The constant multiplier $2\pi$ in the spectrum $[F(\omega) = 2\pi\delta(\omega)]$ may be a bit puzzling. Since $1 = e^{j\omega t}$ with $\omega = 0$, it appears that the Fourier transform of $f(t) = 1$ should be an impulse of strength unity rather than $2\pi$. Recall, however, that in the Fourier transform $f(t)$ is synthesized not by exponentials of amplitude $F(n\Delta\omega)\Delta\omega$, but of amplitude $1/2\pi$ times $F(n\Delta\omega)\Delta\omega$, as seen from Eq. (4.6b). Had we used variable $\mathcal{F}$ (hertz) instead of $\omega$, the spectrum would have been a unit impulse.

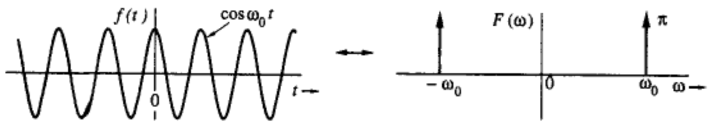



**Fig. 4.13** A cosine signal and its Fourier spectrum.

Adding Eqs. (4.26a) and (4.26b), and using the above formula, we obtain

$$\cos \omega_0 t \Longleftrightarrow \pi[\delta(\omega+\omega_0)+\delta(\omega-\omega_0)] \tag{4.27}$$

The spectrum of $\cos \omega_0 t$ consists of two impulses at $\omega_0$ and $-\omega_0$, as shown in Fig. 4.13. The result also follows from qualitative reasoning. An everlasting sinusoid $\cos \omega_0 t$ can be synthesized by two everlasting exponentials, $e^{j\omega_0 t}$ and $e^{-j\omega_0 t}$. Therefore the Fourier spectrum consists of only two components of frequencies $\omega_0$ and $-\omega_0$. ■

■ **Example 4.7**

Find the Fourier transform of the unit step function $u(t)$.

Trying to find the Fourier transform of $u(t)$ by direct integration leads to an indeterminate result, because

$$U(\omega)=\int_{-\infty}^{\infty} u(t)e^{-j\omega t}\,dt=\int_{0}^{\infty} e^{-j\omega t}\,dt=\frac{-1}{j\omega}e^{-j\omega t}\Big|_0^{\infty}$$

Observe that the upper limit of $e^{-j\omega t}$ as $t\to\infty$ yields an indeterminate answer. So we approach this problem by considering $u(t)$ as a decaying exponential $e^{-at}u(t)$ in the limit as $a\to 0$ (Fig. 4.14a). Thus

$$u(t)=\lim_{a\to 0} e^{-at}u(t)$$

and

$$U(\omega)=\lim_{a\to 0}\mathcal{F}\{e^{-at}u(t)\}=\lim_{a\to 0}\frac{1}{a+j\omega} \tag{4.28a}$$

Expressing the right-hand side in terms of its real and imaginary parts yields

$$\begin{aligned} U(\omega) &= \lim_{a\to 0}\left[\frac{a}{a^2+\omega^2}-j\frac{\omega}{a^2+\omega^2}\right] \\ &= \lim_{a\to 0}\left[\frac{a}{a^2+\omega^2}\right]+\frac{1}{j\omega} \end{aligned} \tag{4.28b}$$

The function $a/(a^2+\omega^2)$ has interesting properties. First, the area under this function (Fig. 4.14b) is $\pi$ regardless of the value of $a$

$$\int_{-\infty}^{\infty}\frac{a}{a^2+\omega^2}\,d\omega=\tan^{-1}\frac{\omega}{a}\Big|_{-\infty}^{\infty}=\pi$$

Second, when $a\to 0$, this function approaches zero for all $\omega\neq 0$, and all its area ($\pi$) is concentrated at a single point $\omega=0$. Clearly, as $a\to 0$, this function approaches an impulse of strength $\pi$.† Thus

$$U(\omega)=\pi\delta(\omega)+\frac{1}{j\omega} \tag{4.29}$$

†The second term on the right-hand side of Eq. (4.28b), being an odd function of $\omega$, has zero area regardless of the value of $a$. As $a\to 0$, the second term approaches $1/j\omega$.



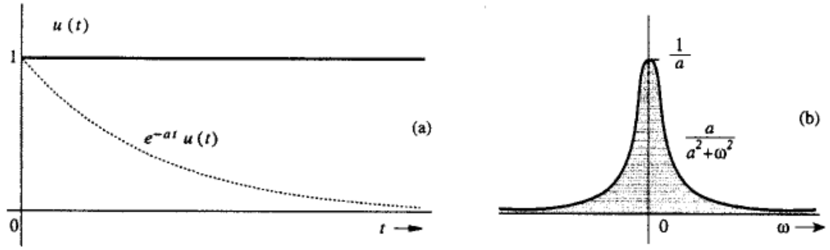



**Fig. 4.14** Derivation of the Fourier transform of the step function.

Note that $u(t)$ is not a "true" dc signal because it is not constant over the interval $-\infty$ to $\infty$. To synthesize a "true" dc, we require only one everlasting exponential with $\omega=0$ (impulse at $\omega=0$). The signal $u(t)$ has a jump discontinuity at $t=0$. It is impossible to synthesize such a signal with a single everlasting exponential $e^{j\omega t}$. To synthesize this signal from everlasting exponentials, we need, in addition to an impulse at $\omega=0$, all frequency components, as indicated by the term $1/j\omega$ in Eq. (4.29). ■

△ **Exercise E4.2**

Show that the Fourier transform of the sign function $\text{sgn}(t)$ depicted in Fig. 4.15a is $2/j\omega$. Hint: Note that $\text{sgn}(t)$ shifted vertically by 1 yields $2u(t)$ ▽.

△ **Exercise E4.3**

Show that the inverse Fourier transform of $F(\omega)$ illustrated in Fig. 4.15b is $f(t)=\frac{\omega_0}{\pi}\,\text{sinc}\,(\omega_0 t)$. Sketch $f(t)$. ▽.

△ **Exercise E4.4**

Show that $\cos(\omega_0 t+\theta)\Longleftrightarrow \pi[\delta(\omega+\omega_0)e^{-j\theta}+\delta(\omega-\omega_0)e^{j\theta}]$.

Hint: $\cos(\omega_0 t+\theta)=\frac{1}{2}[e^{j(\omega_0 t+\theta)}+e^{-j(\omega_0 t+\theta)}]$ ▽

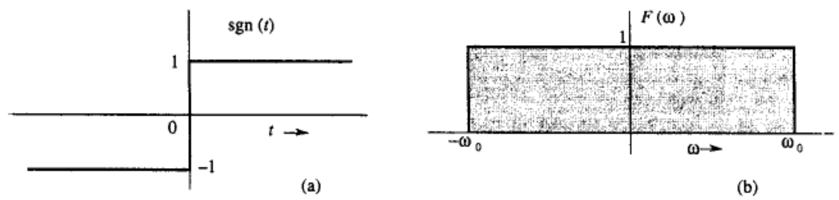



**Fig. 4.15**

## 4.3 Some properties of the Fourier Transform

We now study some of the important properties of the Fourier transform and their implications as well as applications. Before embarking on this study, we shall explain an important and pervasive aspect of the Fourier transform: the time-frequency duality.

**Table 4.1**

**A Short Table of Fourier Transforms**

| | $f(t)$ | $F(\omega)$ | |
|---|---|---|---|
| 1 | $e^{-at}u(t)$ | $\dfrac{1}{a+j\omega}$ | $a>0$ |
| 2 | $e^{at}u(-t)$ | $\dfrac{1}{a-j\omega}$ | $a>0$ |
| 3 | $e^{-a\|t\|}$ | $\dfrac{2a}{a^2+\omega^2}$ | $a>0$ |
| 4 | $te^{-at}u(t)$ | $\dfrac{1}{(a+j\omega)^2}$ | $a>0$ |
| 5 | $t^n e^{-at}u(t)$ | $\dfrac{n!}{(a+j\omega)^{n+1}}$ | $a>0$ |
| 6 | $\delta(t)$ | $1$ | |
| 7 | $1$ | $2\pi\delta(\omega)$ | |
| 8 | $e^{j\omega_0 t}$ | $2\pi\delta(\omega-\omega_0)$ | |
| 9 | $\cos\omega_0 t$ | $\pi[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]$ | |
| 10 | $\sin\omega_0 t$ | $j\pi[\delta(\omega+\omega_0)-\delta(\omega-\omega_0)]$ | |
| 11 | $u(t)$ | $\pi\delta(\omega)+\frac{1}{j\omega}$ | |
| 12 | $\operatorname{sgn} t$ | $\frac{2}{j\omega}$ | |
| 13 | $\cos\omega_0 t\, u(t)$ | $\frac{\pi}{2}[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]+\frac{j\omega}{\omega_0^2-\omega^2}$ | |
| 14 | $\sin\omega_0 t\, u(t)$ | $\frac{\pi}{2j}[\delta(\omega-\omega_0)-\delta(\omega+\omega_0)]+\frac{\omega_0}{\omega_0^2-\omega^2}$ | |
| 15 | $e^{-at}\sin\omega_0 t\, u(t)$ | $\frac{\omega_0}{(a+j\omega)^2+\omega_0^2}$ | $a>0$ |
| 16 | $e^{-at}\cos\omega_0 t\, u(t)$ | $\frac{a+j\omega}{(a+j\omega)^2+\omega_0^2}$ | $a>0$ |
| 17 | $\operatorname{rect}\left(\frac{t}{\tau}\right)$ | $\tau\operatorname{sinc}\left(\frac{\omega\tau}{2}\right)$ | |
| 18 | $\frac{W}{\pi}\operatorname{sinc}(Wt)$ | $\operatorname{rect}\left(\frac{\omega}{2W}\right)$ | |
| 19 | $\Delta\left(\frac{t}{\tau}\right)$ | $\frac{\tau}{2}\operatorname{sinc}^2\left(\frac{\omega\tau}{4}\right)$ | |
| 20 | $\frac{W}{2\pi}\operatorname{sinc}^2\left(\frac{Wt}{2}\right)$ | $\Delta\left(\frac{\omega}{2W}\right)$ | |
| 21 | $\sum_{n=-\infty}^{\infty}\delta(t-nT)$ | $\omega_0\sum_{n=-\infty}^{\infty}\delta(\omega-n\omega_0)$ | $\omega_0=\frac{2\pi}{T}$ |
| 22 | $e^{-t^2/2\sigma^2}$ | $\sigma\sqrt{2\pi}e^{-\sigma^2\omega^2/2}$ | |



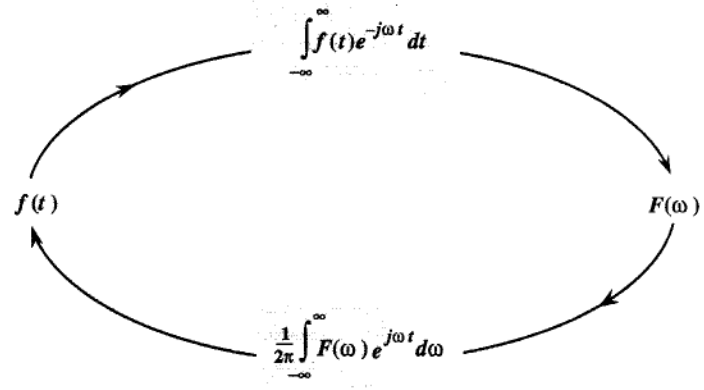



**Fig. 4.16** A near symmetry between the direct and the inverse Fourier transforms.

### 4.3-1 Symmetry of Direct and Inverse Transform Operations: Time-Frequency Duality

Equations (4.8) show an interesting fact: the direct and the inverse transform operations are remarkably similar. These operations, required to go from $f(t)$ to $F(\omega)$ and then from $F(\omega)$ to $f(t)$, are depicted graphically in Fig. 4.16. There are only two minor differences in these operations: the factor $2\pi$ appears only in the inverse operator, and the exponential indices in the two operations have opposite signs. Otherwise the two operations are symmetrical.† This observation has far-reaching consequences in the study of the Fourier transform. It is the basis of the so-called duality of time and frequency. *The duality principle may be compared with a photograph and its negative. A photograph can be obtained from its negative, and by using an identical procedure, the negative can be obtained from the photograph.* For any result or relationship between $f(t)$ and $F(\omega)$, there exists a dual result or relationship, obtained by interchanging the roles of $f(t)$ and $F(\omega)$ in the original result (along with some minor modifications arising because of the factor $2\pi$ and a sign change). For example, the time-shifting property, to be proved later, states that if $f(t) \Longleftrightarrow F(\omega)$, then

$$f(t-t_0) \Longleftrightarrow F(\omega)e^{-j\omega t_0} \tag{4.30a}$$

The dual of this property (the frequency-shifting property) states that

†Of the two differences, the former can be eliminated by change of variable from $\omega$ to $\mathcal{F}$ (in hertz). In this case

$$\omega = 2\pi\mathcal{F} \qquad \text{and} \qquad d\omega = 2\pi\, d\mathcal{F}$$

Therefore, the direct and the inverse transforms are given by

$$F(2\pi\mathcal{F}) = \int_{-\infty}^{\infty} f(t)e^{-j2\pi\mathcal{F}t}dt \qquad \text{and} \qquad f(t) = \int_{-\infty}^{\infty} F(2\pi\mathcal{F})e^{j2\pi\mathcal{F}t}d\mathcal{F}$$

This leaves only one significant difference, that of sign change in the exponential index. Otherwise the two operations are symmetrical.

$$f(t)e^{j\omega_0 t} \Longleftrightarrow F(\omega - \omega_0) \tag{4.30b}$$

Observe the role reversal of time and frequency in these two equations (with the minor difference of the sign change in the exponential index). The value of this principle lies in the fact that *whenever we derive any result, we can be sure that it has a dual.* This possibility can give valuable insights about many unsuspected properties or results in signal processing.

The properties of the Fourier transform are useful not only in deriving the direct and inverse transforms of many functions, but also in obtaining several valuable results in signal processing. The reader should not fail to observe the ever-present duality in this discussion. We begin with the symmetry property, which is one of the consequences of the duality principle discussed.

### 4.3-2 Symmetry Property

This property states that if

$$f(t) \Longleftrightarrow F(\omega)$$

then

$$F(t) \Longleftrightarrow 2\pi f(-\omega) \tag{4.31}$$

Proof: According to Eq. (4.8b)

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(x)e^{jxt}\,dx$$

Hence

$$2\pi f(-t) = \int_{-\infty}^{\infty} F(x)e^{-jxt}\,dx$$

Changing $t$ to $\omega$ yields Eq. (4.31).

■ **Example 4.8**

In this example we apply the symmetry property [Eq. (4.31)] to the pair in Fig. 4.17a. From Eq. (4.23) we have

$$\underbrace{\text{rect}\left(\frac{t}{\tau}\right)}_{f(t)} \Longleftrightarrow \underbrace{\tau\,\text{sinc}\left(\frac{\omega\tau}{2}\right)}_{F(\omega)} \tag{4.32}$$

Also, $F(t)$ is the same as $F(\omega)$ with $\omega$ replaced by $t$, and $f(-\omega)$ is the same as $f(t)$ with $t$ replaced by $-\omega$. Therefore, the symmetry property (4.31) yields

$$\underbrace{\tau\,\text{sinc}\left(\frac{\tau t}{2}\right)}_{F(t)} \Longleftrightarrow \underbrace{2\pi\,\text{rect}\left(\frac{-\omega}{\tau}\right)}_{2\pi f(-\omega)} = 2\pi\,\text{rect}\left(\frac{\omega}{\tau}\right) \tag{4.33}$$

In Eq. (4.33) we used the fact that $\text{rect}\,(-x) = \text{rect}\,(x)$ because rect is an even function. Figure 4.17b shows this pair graphically. Observe the interchange of the roles of $t$ and $\omega$



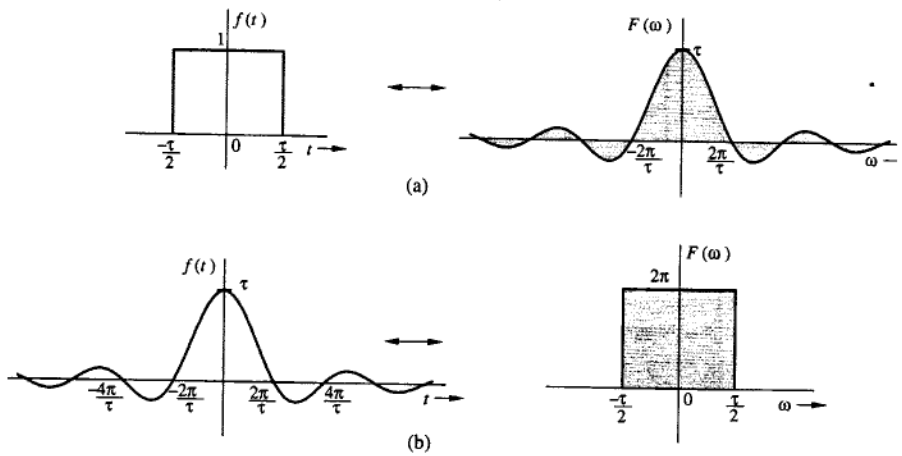


**Fig. 4.17** Symmetry property of the Fourier transform.

(with the minor adjustment of the factor $2\pi$). This result appears as pair 18 in Table 4.1 (with $\tau/2 = W$).

As an interesting exercise, the reader should generate the dual of every pair in Table 4.1 by applying the symmetry property. ■

△ **Exercise E4.5**

Apply symmetry property to pairs 1, 3, and 9 (Table 4.1) to show that

**(a)** $\frac{1}{jt+a} \Longleftrightarrow 2\pi e^{a\omega}u(-\omega)$ **(b)** $\frac{2a}{t^2+a^2} \Longleftrightarrow 2\pi e^{-a|\omega|}$

**(c)** $\delta(t+t_0) + \delta(t-t_0) \Longleftrightarrow 2\cos t_0\omega$ ▽

### 4.3-3 The Scaling Property

If

$$f(t) \Longleftrightarrow F(\omega)$$

then, for any real constant $a$,

$$f(at) \Longleftrightarrow \frac{1}{|a|}F\left(\frac{\omega}{a}\right) \tag{4.34}$$

Proof: For a positive real constant $a$,

$$\mathcal{F}[f(at)] = \int_{-\infty}^{\infty} f(at)e^{-j\omega t}\,dt = \frac{1}{a}\int_{-\infty}^{\infty} f(x)e^{(-j\omega/a)x}\,dx = \frac{1}{a}F\left(\frac{\omega}{a}\right)$$

Similarly, we can demonstrate that if $a < 0$,

$$f(at) \Longleftrightarrow \frac{-1}{a}F\left(\frac{\omega}{a}\right)$$

Hence follows Eq. (4.34).

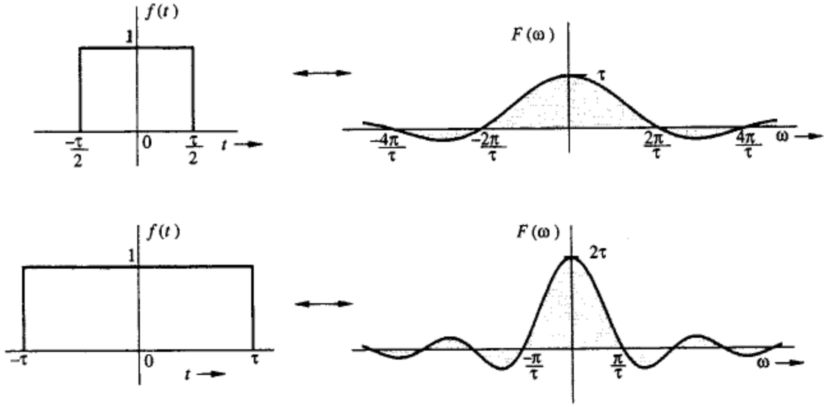



**Fig. 4.18** Scaling property of the Fourier transform.

### Significance of the Scaling Property

The function $f(at)$ represents the function $f(t)$ compressed in time by a factor $a$ (see Sec. 1.3-2). Similarly, a function $F(\frac{\omega}{a})$ represents the function $F(\omega)$ expanded in frequency by the same factor $a$. *The scaling property states that time compression of a signal results in its spectral expansion, and time expansion of the signal results in its spectral compression.* Intuitively, compression in time by a factor $a$ means that the signal is varying rapidly by the same factor. To synthesize such a signal, the frequencies of its sinusoidal components must be increased by the factor $a$, implying that its frequency spectrum is expanded by the factor $a$. Similarly, a signal expanded in time varies more slowly; hence the frequencies of its components are lowered, implying that its frequency spectrum is compressed. For instance, the signal $\cos 2\omega_0 t$ is the same as the signal $\cos \omega_0 t$ time-compressed by a factor of 2. Clearly, the spectrum of the former (impulse at $\pm 2\omega_0$) is an expanded version of the spectrum of the latter (impulse at $\pm\omega_0$). The effect of this scaling is demonstrated in Fig. 4.18.

### Reciprocity of Signal Duration and Its Bandwidth

The scaling property implies that if $f(t)$ is wider, its spectrum is narrower, and vice versa. Doubling the signal duration halves its bandwidth and vice versa. This suggests that the bandwidth of a signal is inversely proportional to the signal duration or width (in seconds). We have already verified this fact for the gate pulse, where we found that the bandwidth of a gate pulse of width $\tau$ seconds is $\frac{1}{\tau}$ Hz. More discussion of this interesting topic appears in the literature.[2]

By letting $a = -1$ in Eq. (4.34), we obtain the **time and frequency inversion property**

$$f(-t) \Longleftrightarrow F(-\omega) \tag{4.35}$$



■ **Example 4.9**

Find the Fourier transforms of $e^{at}u(-t)$ and $e^{-a|t|}$.

Application of Eq. (4.35) to pair 1 (Table 4.1) yields

$$e^{at}u(-t) \Longleftrightarrow \frac{1}{a - j\omega}$$

Also

$$e^{-a|t|} = e^{-at}u(t) + e^{at}u(-t)$$

Therefore

$$e^{-a|t|} \Longleftrightarrow \frac{1}{a+j\omega} + \frac{1}{a-j\omega} = \frac{2a}{a^2+\omega^2} \tag{4.36}$$

The signal $e^{-a|t|}$ and its spectrum are illustrated in Fig. 4.19. ■

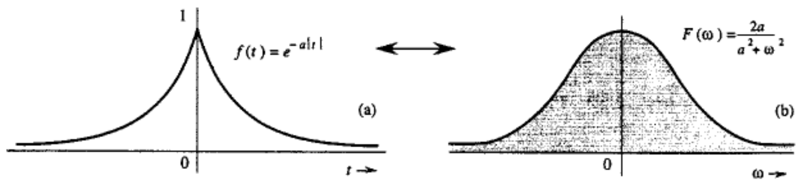



**Fig. 4.19** $e^{-a|t|}$ and its Fourier spectrum.

## 4.3-4 The Time-Shifting Property

If

$$f(t) \Longleftrightarrow F(\omega)$$

then

$$f(t - t_0) \Longleftrightarrow F(\omega)e^{-j\omega t_0} \tag{4.37a}$$

Proof: By definition,

$$\mathcal{F}[f(t-t_0)] = \int_{-\infty}^{\infty} f(t-t_0)e^{-j\omega t}\,dt$$

Letting $t - t_0 = x$, we have

$$\mathcal{F}[f(t-t_0)] = \int_{-\infty}^{\infty} f(x)e^{-j\omega(x+t_0)}\,dx$$

$$= e^{-j\omega t_0}\int_{-\infty}^{\infty} f(x)e^{-j\omega x}\,dx = F(\omega)e^{-j\omega t_0} \tag{4.37b}$$

This result shows that *delaying a signal by $t_0$ seconds does not change its amplitude spectrum. The phase spectrum, however, is changed by $-\omega t_0$.*

### Physical Explanation of the Linear Phase

Time delay in a signal causes a linear phase shift in its spectrum. This result can also be derived by heuristic reasoning. Imagine $f(t)$ being synthesized by its Fourier components, which are sinusoids of certain amplitudes and phases. The delayed signal $f(t - t_0)$ can be synthesized by the same sinusoidal components,

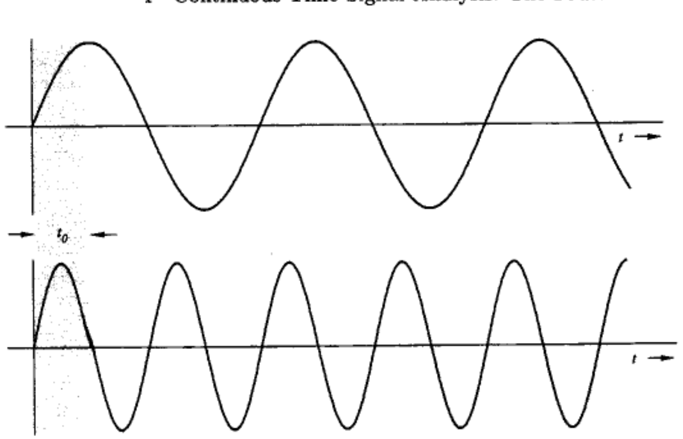



**Fig. 4.20** Physical explanation of the time-shifting property.

each delayed by $t_0$ seconds. The amplitudes of the components remain unchanged. Therefore, the amplitude spectrum of $f(t - t_0)$ is identical to that of $f(t)$. The time delay of $t_0$ in each sinusoid, however, does change the phase of each component. Now, a sinusoid cos $\omega t$ delayed by $t_0$ is given by

$$\cos \omega(t - t_0) = \cos(\omega t - \omega t_0)$$

Therefore a time delay $t_0$ in a sinusoid of frequency $\omega$ manifests as a phase delay of $\omega t_0$. This is a linear function of $\omega$, meaning that higher-frequency components must undergo proportionately higher phase shifts to achieve the same time delay. This effect is depicted in Fig. 4.20 with two sinusoids, the frequency of the lower sinusoid being twice that of the upper. The same time delay $t_0$ amounts to a phase shift of $\pi/2$ in the upper sinusoid and a phase shift of $\pi$ in the lower sinusoid. This verifies the fact that *to achieve the same time delay, higher-frequency sinusoids must undergo proportionately higher phase shifts.* The principle of linear phase shift is very important and we shall encounter it again in distortionless signal transmission and filtering applications.

■ **Example 4.10**

Find the Fourier transform of $e^{-a|t-t_0|}$.

This function, shown in Fig. 4.21a, is a time-shifted version of $e^{-a|t|}$ (depicted in Fig. 4.19a). From Eqs. (4.36) and (4.37) we have

$$e^{-a|t-t_0|} \Longleftrightarrow \frac{2a}{a^2 + \omega^2} e^{-j\omega t_0} \tag{4.38}$$

The spectrum of $e^{-a|t-t_0|}$ (Fig. 4.21b) is the same as that of $e^{-a|t|}$ (Fig. 4.19b), except for an added phase shift of $-\omega t_0$.

Observe that the time delay $t_0$ causes a linear phase spectrum $-\omega t_0$. This example clearly demonstrates the effect of time shift. ■



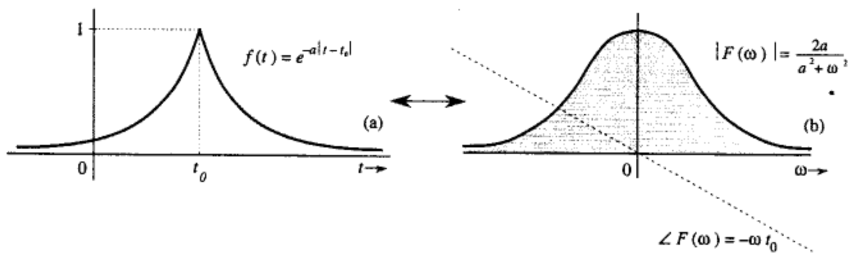



**Fig. 4.21** Effect of time-shifting on the Fourier spectrum of a signal.

■ **Example 4.11**

Find the Fourier transform of the gate pulse $f(t)$ illustrated in Fig. 4.22a.

The pulse $f(t)$ is the gate pulse rect $\left(\frac{t}{\tau}\right)$ in Fig. 4.10a delayed by $\tau/2$ seconds. Hence, according to Eq. (4.37a), its Fourier transform is the Fourier transform of rect $\left(\frac{t}{\tau}\right)$ multiplied by $e^{-j\omega\frac{\tau}{2}}$. Therefore

$$F(\omega) = \tau \operatorname{sinc}\left(\frac{\omega\tau}{2}\right) e^{-j\omega\frac{\tau}{2}}$$

The amplitude spectrum $|F(\omega)|$ (depicted in Fig. 4.22b) of this pulse is the same as that indicated in Fig. 4.10c. But the phase spectrum has an added linear term $-\omega\tau/2$. Hence, the phase spectrum of $f(t)$ is identical to that in Fig. 4.10b plus a linear term $-\omega\tau/2$, as indicated in Fig. 4.22c. ■

△ **Exercise E4.6**

Using pair 18 and the time-shifting property, show that the Fourier transform of sinc $[\omega_0(t - T)]$ is $\frac{\pi}{\omega_0}$ rect $\left(\frac{\omega}{2\omega_0}\right)e^{-j\omega T}$. Sketch the amplitude and phase spectra of the Fourier transform. ▽

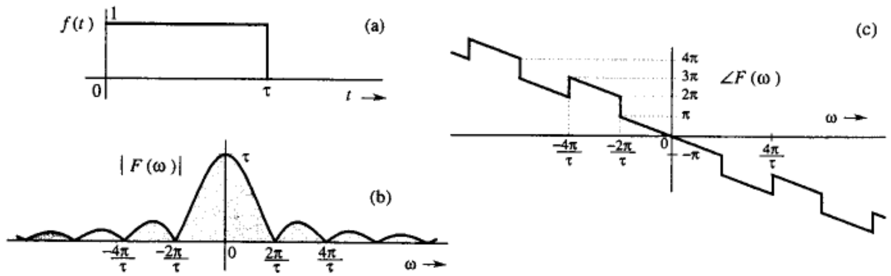



**Fig. 4.22** Another example of time-shifting and its effect on the Fourier spectrum of a signal.

### 4.3-5 The Frequency-Shifting Property

If

$$f(t) \Longleftrightarrow F(\omega)$$

then

$$f(t)e^{j\omega_0 t} \Longleftrightarrow F(\omega - \omega_0) \tag{4.39}$$

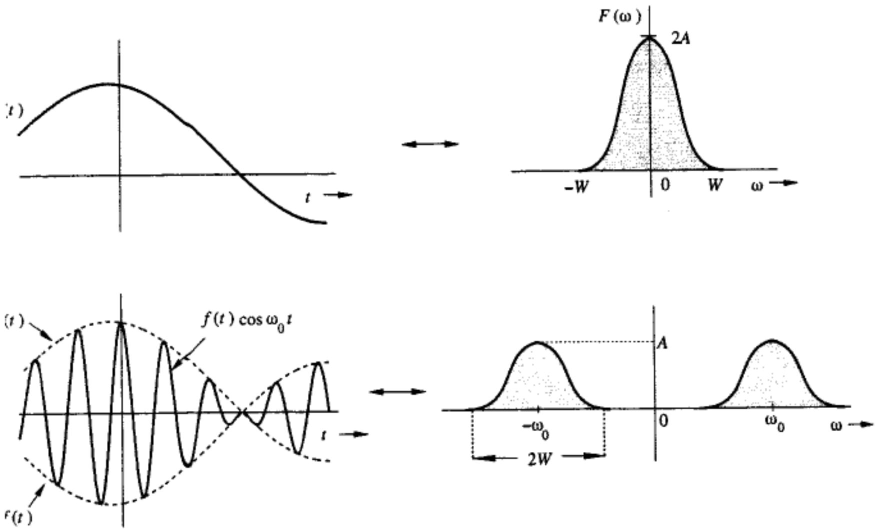



**Fig. 4.23** Amplitude modulation of a signal causes spectral shifting.

Proof: By definition,

$$\mathcal{F}[f(t)e^{j\omega_0 t}] = \int_{-\infty}^{\infty} f(t)e^{j\omega_0 t}e^{-j\omega t}dt = \int_{-\infty}^{\infty} f(t)e^{-j(\omega-\omega_0)t}dt = F(\omega - \omega_0)$$

According to this property, the multiplication of a signal by a factor $e^{j\omega_0 t}$ shifts the spectrum of that signal by $\omega = \omega_0$. Note the duality between the time-shifting and the frequency-shifting properties.

Changing $\omega_0$ to $-\omega_0$ in Eq. (4.39) yields

$$f(t)e^{-j\omega_0 t} \Longleftrightarrow F(\omega + \omega_0) \tag{4.40}$$

Because $e^{j\omega_0 t}$ is not a real function that can be generated, frequency shifting in practice is achieved by multiplying $f(t)$ by a sinusoid. This assertion follows from the fact that

$$f(t)\cos\omega_0 t = \tfrac{1}{2}[f(t)e^{j\omega_0 t} + f(t)e^{-j\omega_0 t}]$$

From Eqs. (4.39) and (4.40), it follows that

$$f(t)\cos\omega_0 t \Longleftrightarrow \tfrac{1}{2}[F(\omega - \omega_0) + F(\omega + \omega_0)] \tag{4.41}$$

This result shows that the multiplication of a signal $f(t)$ by a sinusoid of frequency $\omega_0$ shifts the spectrum $F(\omega)$ by $\pm\omega_0$, as depicted in Fig. 4.23.

Multiplication of a sinusoid $\cos\omega_0 t$ by $f(t)$ amounts to modulating the sinusoid amplitude. This type of modulation is known as **amplitude modulation**. The sinusoid $\cos\omega_0 t$ is called the **carrier**, the signal $f(t)$ is the **modulating signal**, and the signal $f(t)\cos\omega_0 t$ is the **modulated signal**. Further discussion of modulation



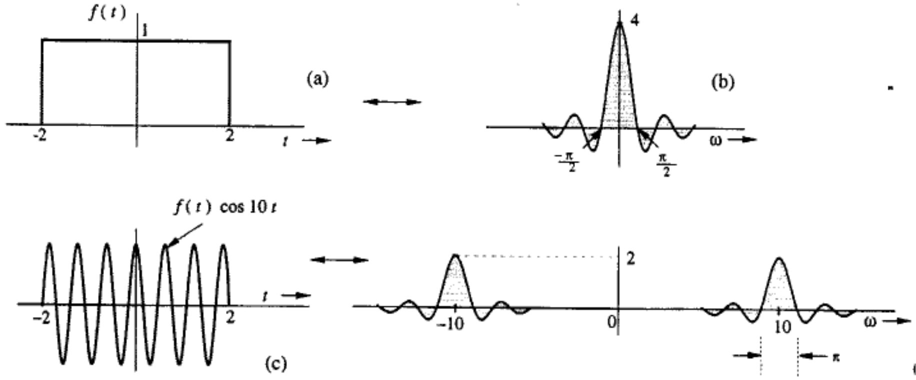



**Fig. 4.24** An example of spectral shifting by amplitude modulation.

and demodulation appears in Secs. 4.7 and 4.8.

To sketch a signal $f(t)\cos\omega_0 t$, we observe that

$$f(t)\cos\omega_0 t = \begin{cases} f(t) & \text{when } \cos\omega_0 t = 1 \\ -f(t) & \text{when } \cos\omega_0 t = -1 \end{cases}$$

Therefore, $f(t)\cos\omega_0 t$ touches $f(t)$ when the sinusoid $\cos\omega_0 t$ is at its positive peaks and touches $-f(t)$ when $\cos\omega_0 t$ is at its negative peaks. This means that $f(t)$ and $-f(t)$ act as envelopes for the signal $f(t)\cos\omega_0 t$ (see Fig. 4.23). The signal $-f(t)$ is a mirror image of $f(t)$ about the horizontal axis. Figure 4.23 shows the signal $f(t)$, $f(t)\cos\omega_0 t$ and their spectra.

■ **Example 4.12**

Find and sketch the Fourier transform of the modulated signal $f(t)\cos 10t$ in which $f(t)$ is a gate pulse $\text{rect}\left(\frac{t}{4}\right)$ as illustrated in Figure 4.24a.

From pair 17 (Table 4.1) we find $\text{rect}\left(\frac{t}{4}\right) \Longleftrightarrow 4\,\text{sinc}\,(2\omega)$, which is depicted in Fig. 4.24b. From Eq. (4.41) it follows that

$$f(t)\cos 10t \Longleftrightarrow \tfrac{1}{2}[F(\omega + 10) + F(\omega - 10)]$$

In this case, $F(\omega) = 4\,\text{sinc}\,(2\omega)$. Therefore

$$f(t)\cos 10t \Longleftrightarrow 2\,\text{sinc}\,[2(\omega + 10)] + 2\,\text{sinc}\,[2(\omega - 10)]$$

The spectrum of $f(t)\cos 10t$ is obtained by shifting $F(\omega)$ in Fig. 4.24b to the left by 10 and also to the right by 10, and then multiplying it by one-half, as depicted in Fig. 4.24d. ■

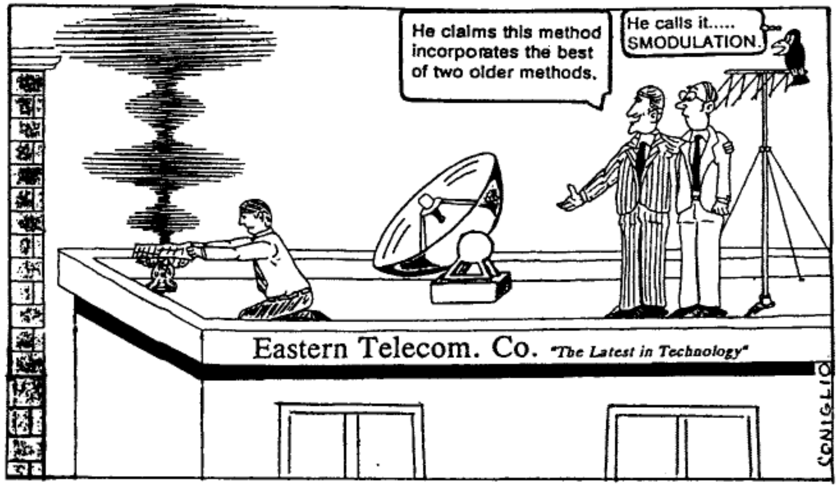
△ **Exercise E4.7**

Sketch signal $e^{-|t|}\cos 10t$. Find the Fourier transform of this signal and sketch its spectrum.

Answer: $F(\omega) = \frac{1}{(\omega-10)^2+1} + \frac{1}{(\omega+10)^2+1}$. The spectrum is that in Fig. 4.19b (with $a = 1$), shifted to $\pm 10$ and multiplied by one-half. ▽

### Application to Modulation

Modulation is used to shift signal spectra. Some of the situations where spectrum shifting is necessary are presented next.




1. If several signals, each occupying the same frequency band, are transmitted simultaneously over the same transmission medium, they will all interfere; it will be impossible to separate or retrieve them at a receiver. For example, if all radio stations decide to broadcast audio signals simultaneously, the receiver will not be able to separate them. This problem is solved by using modulation, whereby each radio station is assigned a distinct carrier frequency. Each station transmits a modulated signal. This procedure shifts the signal spectrum to its allocated band, which is not occupied by any other station. A radio receiver can pick up any station by tuning to the band of the desired station. The receiver must now demodulate the received signal (undo the effect of modulation). Demodulation therefore consists of another spectral shift required to restore the signal to its original band. Note that both modulation and demodulation implement spectral shifting; consequently, demodulation operation is similar to modulation (see Sec. 4.7).

   This method of transmitting several signals simultaneously over a channel by sharing its frequency band is known as **frequency-division multiplexing (FDM)**.
2. For effective radiation of power over a radio link, the antenna size must be of the order of the wavelength of the signal to be radiated. Audio signal frequencies are so low (wavelengths are so large) that impracticably large antennas will be required for radiation. Here, shifting the spectrum to a higher frequency (a smaller wavelength) by modulation solves the problem.

### 4.3-6 Convolution

The time convolution property and its dual, the frequency convolution property, state that if

$$f_1(t) \Longleftrightarrow F_1(\omega) \qquad \text{and} \qquad f_2(t) \Longleftrightarrow F_2(\omega)$$



then **(time convolution)**

$$f_1(t) * f_2(t) \Longleftrightarrow F_1(\omega)F_2(\omega) \tag{4.42}$$

and **(frequency convolution)**

$$f_1(t)f_2(t) \Longleftrightarrow \frac{1}{2\pi}F_1(\omega) * F_2(\omega) \tag{4.43}$$

Proof: By definition

$$\begin{aligned}\mathcal{F}|f_1(t) * f_2(t)| &= \int_{-\infty}^{\infty} e^{-j\omega t}\left[\int_{-\infty}^{\infty} f_1(\tau)f_2(t-\tau)d\tau\right] dt \\ &= \int_{-\infty}^{\infty} f_1(\tau)\left[\int_{-\infty}^{\infty} e^{-j\omega t}f_2(t-\tau)dt\right] d\tau\end{aligned}$$

The inner integral is the Fourier transform of $f_2(t-\tau)$, given by [time-shifting property in Eq. (4.37)] $F_2(\omega)e^{-j\omega\tau}$. Hence

$$\mathcal{F}[f_1(t)*f_2(t)] = \int_{-\infty}^{\infty} f_1(\tau)e^{-j\omega\tau}F_2(\omega)d\tau = F_2(\omega)\int_{-\infty}^{\infty} f_1(\tau)e^{-j\omega\tau}d\tau = F_1(\omega)F_2(\omega)$$

We have demonstrated in Eq. (2.48) that the transfer function $H(\omega)$ is the Fourier transform of the unit impulse response $h(t)$. Thus

$$h(t) \Longleftrightarrow H(\omega) \tag{4.44a}$$

Application of the time convolution property to $y(t) = f(t) * h(t)$ yields (assuming both $f(t)$ and $h(t)$ are Fourier transformable)

$$Y(\omega) = F(\omega)H(\omega) \tag{4.44b}$$

This is precisely the result proved earlier in Eq. (4.19).†

The frequency convolution property (4.43) can be proved in exactly the same way by reversing the roles of $f(t)$ and $F(\omega)$.

■ **Example 4.13**

Using the time convolution property, show that if

$$f(t) \Longleftrightarrow F(\omega)$$

then

$$\int_{-\infty}^{t} f(\tau)d\tau \Longleftrightarrow \frac{F(\omega)}{j\omega} + \pi F(0)\delta(\omega) \tag{4.45}$$

Because

$$u(t-\tau) = \begin{cases} 1 & \tau \le t \\ 0 & \tau > t \end{cases}$$

it follows that

$$f(t) * u(t) = \int_{-\infty}^{\infty} f(\tau)u(t-\tau)d\tau = \int_{-\infty}^{t} f(\tau)\,d\tau$$

†In Eq. (4.44b), $h(t) \Longleftrightarrow H(\omega)$. To understand finer points of Eq. (4.44b), see footnote on p. 243.

Now, from the time convolution property [Eq. 4.42], it follows that

$$f(t) * u(t) = \int_{-\infty}^{t} f(\tau)\, d\tau \Longleftrightarrow F(\omega)\left[\frac{1}{j\omega} + \pi\delta(\omega)\right]$$
$$= \frac{F(\omega)}{j\omega} + \pi F(0)\delta(\omega)$$

In deriving the last result, we used Eq. (1.23a) ■

△ **Exercise E4.8**
Using the time convolution property, show that $f(t) * \delta(t) = f(t)$ ▽

△ **Exercise E4.9**
Using the time convolution property, show that

$$e^{-at}u(t) * e^{-bt}u(t) = \frac{1}{b-a}\left[e^{-at} - e^{-bt}\right] u(t)$$

Hint: Use property (4.42) to find the Fourier transform of $e^{-at}u(t) * e^{-bt}u(t)$. Then use partial fraction expansion to find its inverse Fourier transform. ▽

### 4.3-7 Time Differentiation and Time Integration

If

$$f(t) \Longleftrightarrow F(\omega)$$

then **(time differentiation)**†

$$\frac{df}{dt} \Longleftrightarrow j\omega F(\omega) \tag{4.46}$$

and **(time integration)**

$$\int_{-\infty}^{t} f(\tau) d\tau \Longleftrightarrow \frac{F(\omega)}{j\omega} + \pi F(0)\delta(\omega) \tag{4.47}$$

Proof: Differentiation of both sides of Eq. (4.8b) yields

$$\frac{df}{dt} = \frac{1}{2\pi}\int_{-\infty}^{\infty} j\omega F(\omega) e^{j\omega t}\, d\omega$$

This result shows that

$$\frac{df}{dt} \Longleftrightarrow j\omega F(\omega)$$

†Valid only if the transform of $df/dt$ exists.



**Table 4.2**

**Fourier Transform Operations**

| Operation | $f(t)$ | $F(\omega)$ |
|---|---|---|
| Addition | $f_1(t) + f_2(t)$ | $F_1(\omega) + F_2(\omega)$ |
| Scalar multiplication | $kf(t)$ | $kF(\omega)$ |
| Symmetry | $F(t)$ | $2\pi f(-\omega)$ |
| Scaling ($a$ real) | $f(at)$ | $\frac{1}{\lvert a\rvert}F\left(\frac{\omega}{a}\right)$ |
| Time shift | $f(t-t_0)$ | $F(\omega)e^{-j\omega t_0}$ |
| Frequency shift ($\omega_0$ real) | $f(t)e^{j\omega_0 t}$ | $F(\omega - \omega_0)$ |
| Time convolution | $f_1(t) * f_2(t)$ | $F_1(\omega)F_2(\omega)$ |
| Frequency convolution | $f_1(t)f_2(t)$ | $\frac{1}{2\pi}F_1(\omega) * F_2(\omega)$ |
| Time differentiation | $\frac{d^n f}{dt^n}$ | $(j\omega)^n F(\omega)$ |
| Time integration | $\int_{-\infty}^{t} f(x)\, dx$ | $\frac{F(\omega)}{j\omega} + \pi F(0)\delta(\omega)$ |

Repeated application of this property yields

$$\frac{d^n f}{dt^n} \Longleftrightarrow (j\omega)^n F(\omega) \tag{4.48}$$

The time-integration property [Eq. (4.47)] has already been proved in Example 4.13.

■ **Example 4.14**

Using the time-differentiation property, find the Fourier transform of the triangle pulse $\Delta(\frac{t}{\tau})$ illustrated in Fig. 4.25a.

To find the Fourier transform of this pulse we differentiate the pulse successively, as illustrated in Fig. 4.25b and c. Because $df/dt$ is constant everywhere, its derivative, $d^2f/dt^2$, is zero everywhere. But $df/dt$ has jump discontinuities with a positive jump of $2/\tau$ at $t = \pm\frac{\tau}{2}$, and a negative jump of $4/\tau$ at $t = 0$. Recall that the derivative of a signal at a jump discontinuity is an impulse at that point of strength equal to the amount of jump. Hence, $d^2f/dt^2$, the derivative of $df/dt$, consists of a sequence of impulses, as depicted in Fig. 4.25c; that is,

$$\frac{d^2 f}{dt^2} = \frac{2}{\tau}[\delta(t + \tfrac{\tau}{2}) - 2\delta(t) + \delta(t - \tfrac{\tau}{2})] \tag{4.49}$$

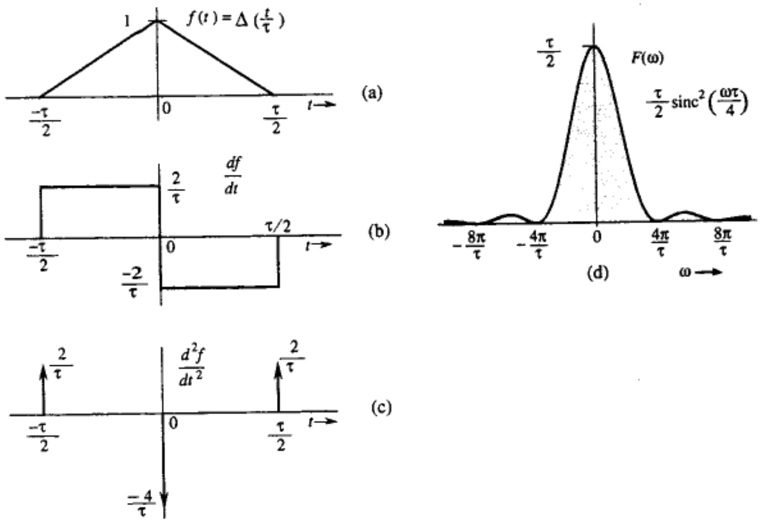



**Fig. 4.25** Finding the Fourier transform of a piecewise-linear signal using the time-differentiation property.

From the time-differentiation property (4.48)

$$\frac{d^2 f}{dt^2} \Longleftrightarrow (j\omega)^2 F(\omega) = -\omega^2 F(\omega) \tag{4.50a}$$

Also, from the time-shifting property (4.37)

$$\delta(t - t_0) \Longleftrightarrow e^{-j\omega t_0} \tag{4.50b}$$

Taking the Fourier transform of Eq. (4.49) and using the results in Eqs. (4.50), we obtain

$$-\omega^2 F(\omega) = \tfrac{2}{\tau}[e^{j\frac{\omega\tau}{2}} - 2 + e^{-j\frac{\omega\tau}{2}}] = \tfrac{4}{\tau}\left(\cos\tfrac{\omega\tau}{2} - 1\right) = -\tfrac{8}{\tau}\sin^2\left(\tfrac{\omega\tau}{4}\right)$$

and

$$F(\omega) = \frac{8}{\omega^2\tau}\sin^2\left(\frac{\omega\tau}{4}\right) = \frac{\tau}{2}\left[\frac{\sin\left(\frac{\omega\tau}{4}\right)}{\frac{\omega\tau}{4}}\right]^2 = \frac{\tau}{2}\,\mathrm{sinc}^2\left(\frac{\omega\tau}{4}\right) \tag{4.51}$$

The spectrum $F(\omega)$ is depicted in Fig. 4.25d. This procedure of finding the Fourier trans- form can be applied to any function $f(t)$ made up of straight-line segments with $f(t) \to 0$ as $|t| \to \infty$. The second derivative of such a signal yields a sequence of impulses whose Fourier transform can be found by inspection. This example suggests a numerical method of finding the Fourier transform of an arbitrary signal $f(t)$ by approximating the signal by straight-line segments. ■

△ **Exercise E4.10**

Find the Fourier transform of $\mathrm{rect}\left(\frac{t}{\tau}\right)$, using the time-differentiation property. ▽



## 4.4 Signal Transmission through LTIC Systems

If $f(t)$ and $y(t)$ are the input and output of an LTIC system with transfer function $H(\omega)$, then, as demonstrated in Eq. (4.44b)

$$Y(\omega) = H(\omega)F(\omega) \tag{4.52}$$

This result applies only to asymptotically (and marginally) stable systems because of the reasons discussed in the footnote of p. 243. Moreover, $f(t)$ has to be Fourier transformable. Consequently, exponentially growing inputs cannot be handled by this method.

In Chapter 6, we shall see that the Laplace transform, which is a generalized Fourier transform, is more versatile and capable of analyzing all kinds of LTIC systems whether stable, unstable, or marginally stable. Laplace transform can also handle exponentially growing inputs. Compared to the Laplace transform, the Fourier transform in system analysis is clumsier. Hence, the Laplace transform is preferable to the Fourier transform in LTIC system analysis, and we shall not belabor the application of the Fourier transform to LTIC system analysis. We consider just one example here.

■ **Example 4.15**

Find the zero-state response of a stable LTIC system with transfer function†

$$H(s) = \frac{1}{s+2} \tag{4.53}$$

and the input $f(t) = e^{-t}u(t)$. In this case,

$$F(\omega) = \frac{1}{j\omega + 1} \quad \text{and} \quad H(\omega) = H(s)|_{s=j\omega} = \frac{1}{j\omega + 2}$$

Therefore

$$\begin{aligned} Y(\omega) &= H(\omega)F(\omega) \\ &= \frac{1}{(j\omega+2)(j\omega+1)} \end{aligned}$$

Expanding the right-hand side in partial fractions (Sec. B.5)

$$Y(\omega) = \frac{1}{j\omega+1} - \frac{1}{j\omega+2} \tag{4.54}$$

and

$$y(t) = (e^{-t} - e^{-2t})u(t) \quad ■$$

△ **Exercise E4.11**

For the system in Example 4.15, show that the zero-input response to the input $e^t u(-t)$ is $y(t) = \frac{1}{3}[e^t u(-t) + e^{-2t}u(t)]$.

Hint: Use pair 2 (Table 4.1) to find the Fourier transform of $e^t u(-t)$. ▽

†Stability implies that the region of convergence of $H(s)$ includes the $j\omega$ axis.

### 4.4-1 Signal Distortion during Transmission

For a system with transfer function $H(\omega)$, if $F(\omega)$ and $Y(\omega)$ are the spectra of the input and the output signals, respectively, then

$$Y(\omega) = F(\omega)H(\omega) \tag{4.55}$$

The transmission of the input signal $f(t)$ through the system changes it into the output signal $y(t)$. Equation (4.55) shows the nature of this change or modification. Here $F(\omega)$ and $Y(\omega)$ are the spectra of the input and the output, respectively. Therefore, $H(\omega)$ is the spectral response of the system. The output spectrum is obtained by the input spectrum multiplied by the spectral response of the system. Equation (4.55), which clearly brings out the spectral shaping (or modification) of the signal by the system, can be expressed in polar form as

$$|Y(\omega)|e^{j\angle Y(\omega)} = |F(\omega)||H(\omega)|e^{j[\angle F(\omega)+\angle H(j\omega)]}$$

Therefore

$$|Y(\omega)| = |F(\omega)|\,|H(\omega)| \tag{4.56a}$$

$$\angle Y(\omega) = \angle F(\omega) + \angle H(\omega) \tag{4.56b}$$

During transmission, the input signal amplitude spectrum $|F(\omega)|$ is changed to $|F(\omega)|\,|H(\omega)|$. Similarly, the input signal phase spectrum $\angle F(\omega)$ is changed to $\angle F(\omega) + \angle H(\omega)$. An input signal spectral component of frequency $\omega$ is modified in amplitude by a factor $|H(\omega)|$ and is shifted in phase by an angle $\angle H(\omega)$. Clearly, $|H(\omega)|$ is the amplitude response, and $\angle H(\omega)$ is the phase response of the system. The plots of $|H(\omega)|$ and $\angle H(\omega)$ as functions of $\omega$ show at a glance how the system modifies the amplitudes and phases of various sinusoidal inputs. For this reason, $H(\omega)$ is called the **frequency response** of the system. During transmission through the system, some frequency components may be boosted in amplitude, while others may be attenuated. The relative phases of the various components also change. In general, the output waveform will be different from the input waveform.

#### Distortionless Transmission

In several applications, such as signal amplification or message signal transmission over a communication channel, we require that the output waveform be a replica of the input waveform. In such cases we need to minimize the distortion caused by the amplifier or the communication channel. It is, therefore, of practical interest to determine the characteristics of a system that allows a signal to pass without distortion (**distortionless transmission**).

Transmission is said to be distortionless if the input and the output have identical waveshapes within a multiplicative constant. A delayed output that retains the input waveform is also considered distortionless. Thus, in distortionless transmission, the input $f(t)$ and the output $y(t)$ satisfy the condition

$$y(t) = kf(t - t_d) \tag{4.57}$$



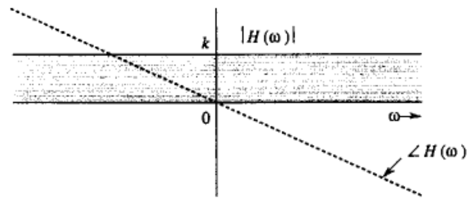



**Fig. 4.26** LTIC system Frequency response for distortionless transmission.

The Fourier transform of this equation yields

$$Y(\omega) = kF(\omega)e^{-j\omega t_d}$$

But

$$Y(\omega) = F(\omega)H(\omega)$$

Therefore

$$H(\omega) = k\,e^{-j\omega t_d}$$

This is the transfer function required for distortionless transmission. From this equation it follows that

$$|H(\omega)| = k \tag{4.58a}$$

$$\angle H(\omega) = -\omega t_d \tag{4.58b}$$

This result shows that for distortionless transmission, the amplitude response $|H(\omega)|$ must be a constant, and the phase response $\angle H(\omega)$ must be a linear function of $\omega$ with slope $-t_d$, where $t_d$ is the delay of the output with respect to input (Fig. 4.26).

#### Intuitive Explanation of the Distortionless Transmission Conditions

It is instructive to derive the conditions for distortionless transmission heuristically. Once again, imagine $f(t)$ to be composed of various sinusoids (its spectral components), which are being passed through a distortionless system. For the distortionless case, the output signal is the input signal multiplied by $k$ and delayed by $t_d$. To synthesize such a signal, we need exactly the same components as those of $f(t)$, with each component multiplied by $k$ and delayed by $t_d$. Thus, the system transfer function $H(\omega)$ should be such that each sinusoidal component suffers the same attenuation $k$ and each component undergoes the same time delay of $t_d$ seconds. The first condition requires that

$$|H(\omega)| = k$$

We have seen in our discussion on p. 258 that to achieve the same time delay $t_d$ for every frequency component requires a linear phase delay $\omega t_d$ (Fig. 4.20). Therefore

$$\angle H(\omega) = -\omega t_d$$

This equation shows that the time delay resulting from signal transmission through a system is the negative of the slope of the system phase response $\angle H(\omega)$; that is,

$$t_d(\omega) = -\frac{d}{d\omega}\angle H(\omega) \tag{4.59}$$

If the slope of $\angle H(\omega)$ is constant (that is, if $\angle H(\omega)$ is linear with $\omega$), all the components are delayed by the same time interval $t_d$. But if the slope is not constant, the time delay $t_d$ varies with frequency. This variation means that different frequency components undergo different amounts of time delay, and consequently the output waveform will not be a replica of the input waveform. A good way to judge phase linearity is to plot $t_d$ as a function of frequency. For a distortionless system, $t_d$ should be constant over the band of interest.

It is often thought (erroneously) that flatness of amplitude response $|H(\omega)|$ alone can guarantee signal quality. However, a system may have a flat amplitude response and yet distort a signal beyond recognition if the phase response is not linear ($t_d$ not constant).

### The Nature of Distortion in Audio and Video Signals

Generally speaking, a human ear can readily perceive amplitude distortion, although it is relatively insensitive to phase distortion. For the phase distortion to become noticeable, the variation in delay [variation in the slope of $\angle H(\omega)$] should be comparable to the signal duration (or the physically perceptible duration, in case the signal itself is long). In the case of audio signals, each spoken syllable can be considered as an individual signal. The average duration of a spoken syllable is of a magnitude of the order of 0.01 to 0.1 seconds. The audio systems may have nonlinear phases, yet no noticeable signal distortion results because in practical audio systems, maximum variation in the slope of $\angle H(\omega)$ is only a small fraction of a millisecond. This is the real truth underlying the statement that "the human ear is relatively insensitive to phase distortion."[3] As a result, the manufacturers of audio equipment make available only $|H(\omega)|$, the amplitude response characteristic of their systems.

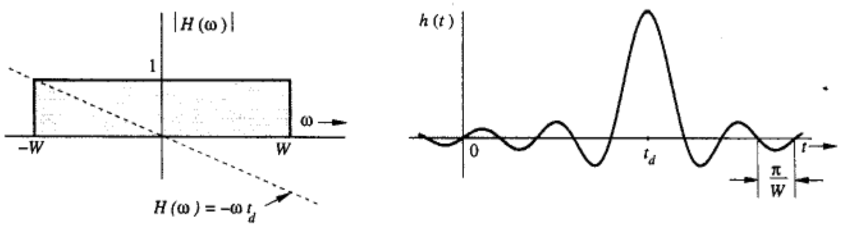
For video signals, in contrast, the situation is exactly the opposite. The human eye is sensitive to phase distortion but is relatively insensitive to amplitude distortion. The amplitude distortion in television signals manifests itself as a partial destruction of the relative half-tone values of the resulting picture, which is not readily apparent to the human eye. The phase distortion (nonlinear phase), on the other hand, causes different time delays in different picture elements. The result is a smeared picture, which is readily perceived by the human eye. Phase distortion is also very important in digital communication systems because the nonlinear phase characteristic of a channel causes pulse dispersion (spreading out), which in turn causes pulses to interfere with neighboring pulses. This interference can cause an error in the pulse amplitude at the receiver: a binary **1** may read as **0**, and vice versa.

## 4.5 Ideal and practical filters

Ideal filters allow distortionless transmission of a certain band of frequencies and suppress all the remaining frequencies. The ideal lowpass filter (Fig. 4.27), for example, allows all components below $\omega = W$ rad/s to pass without distortion and suppresses all components above $\omega = W$. Figure 4.28 illustrates ideal highpass and bandpass filter characteristics.

The ideal lowpass filter in Fig. 4.27a has a linear phase of slope $-t_d$, which results in a time delay of $t_d$ seconds for all its input components of frequencies






(a) (b)

**Fig. 4.27** Ideal lowpass filter: its frequency response and impulse response.

below $W$ rad/s. Therefore, if the input is a signal $f(t)$ bandlimited to $W$ rad/s, the output $y(t)$ is $f(t)$ delayed by $t_d$; that is,

$$y(t) = f(t - t_d)$$

The signal $f(t)$ is transmitted by this system without distortion, but with time delay $t_d$. For this filter $|H(\omega)| = \text{rect}\left(\frac{\omega}{2W}\right)$ and $\angle H(\omega) = e^{-j\omega t_d}$, so that

$$H(\omega) = \text{rect}\left(\frac{\omega}{2W}\right) e^{-j\omega t_d} \tag{4.60a}$$

The unit impulse response $h(t)$ of this filter is obtained from pair 18 (Table 4.1) and the time-shifting property

$$\begin{aligned} h(t) &= \mathcal{F}^{-1}\left[\text{rect}\left(\frac{\omega}{2W}\right) e^{-j\omega t_d}\right] \\ &= \frac{W}{\pi}\,\text{sinc}\,[W(t - t_d)] \end{aligned} \tag{4.60b}$$

Recall that $h(t)$ is the system response to impulse input $\delta(t)$, which is applied at $t = 0$. Figure 4.27b shows a curious fact: the response $h(t)$ begins even before the input is applied (at $t = 0$). Clearly, the filter is noncausal and therefore physically unrealizable. Similarly, one can show that other ideal filters (such as the ideal highpass or the ideal bandpass filters depicted in Fig. 4.28) are also physically unrealizable.

For a physically realizable system, $h(t)$ must be causal; that is,

$$h(t) = 0 \qquad \text{for } t < 0$$

In the frequency domain, this condition is equivalent to the well-known **Paley-Wiener criterion**, which states that the necessary and sufficient condition for the amplitude response $|H(\omega)|$ to be realizable is

$$\int_{-\infty}^{\infty} \frac{|\ln|H(\omega)||}{1 + \omega^2}\,d\omega < \infty \tag{4.61}$$

If $H(\omega)$ does not satisfy this condition, it is unrealizable. Note that if $|H(\omega)| = 0$ over any finite band, $|\ln|H(\omega)|| = \infty$ over that band, and the condition (4.61) is violated. If, however, $H(\omega) = 0$ at a single frequency (or a set of discrete frequencies),

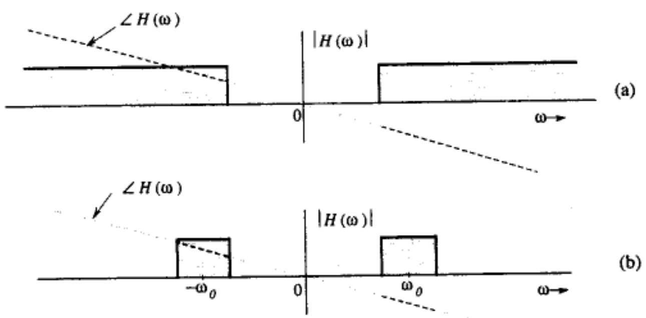



**Fig. 4.28** Ideal highpass and bandpass filter frequency response.


**Fig. 4.29** Approximate realization of an ideal lowpass filter by truncation of its impulse response.

the integral in Eq. (4.61) may still be finite even though the integrand is infinite. Therefore, for a physically realizable system, $H(\omega)$ may be zero at some discrete frequencies, but it cannot be zero over any finite band. According to this criterion, ideal filter characteristics (Figs. 4.27 and 4.28) are clearly unrealizable.†

The impulse response $h(t)$ in Fig. 4.27 is not realizable. One practical approach to filter design is to cut off the tail of $h(t)$ for $t < 0$. The resulting causal impulse response $\widehat{h}(t)$, where

$$\widehat{h}(t) = h(t)u(t)$$

is physically realizable because it is causal (Fig. 4.29). If $t_d$ is sufficiently large, $\widehat{h}(t)$ will be a close approximation of $h(t)$, and the resulting filter $\widehat{H}(\omega)$ will be a good approximation of an ideal filter. This close realization of the ideal filter is achieved because of the increased value of time-delay $t_d$. This observation means that the price of close realization is higher delay in the output; this situation is common in noncausal systems. Of course, theoretically, a delay $t_d = \infty$ is needed to realize the ideal characteristics. But a glance at Fig. 4.27b shows that a delay $t_d$ of three or four times $\frac{\pi}{W}$ will make $\widehat{h}(t)$ a reasonably close version of $h(t-t_d)$. For instance, an audio

†$|H(\omega)|$ is assumed to be square-integrable, that is,

$$\int_{-\infty}^{\infty} |H(\omega)|^2 \, d\omega$$

is finite. Note that the Paley-Wiener criterion is a criterion for the realizability of the amplitude response $|H(\omega)|$.



filter is required to handle frequencies of up to 20 kHz ($W = 40,000\pi$). In this case, a $t_d$ of about $10^{-4}$ (0.1 ms) would be a reasonable choice. The truncation operation (cutting the tail of $h(t)$ to make it causal), however, creates some unsuspected problems. We discuss these problems and their cure in Sec. 4.9.

In practice, we can realize a variety of filter characteristics to approach ideal characteristics. Practical (realizable) filter characteristics are gradual, without jump discontinuities in amplitude response. We shall study such filter families (Butterworth and Chebyshev) in Secs. 7.4 and 7.5. Figure 7.17 illustrates the amplitude response of lowpass Butterworth filters.

△ **Exercise E4.12**

Show that a filter with Gaussian transfer function $H(\omega) = e^{-\alpha\omega^2}$ is unrealizable. Demonstrate this fact in two ways: first by showing that its impulse response is noncausal, and then by showing that $|H(\omega)|$ violates the Paley-Wiener criterion.

Hint: Use pair 22 in Table 4.1 ▽

### Thinking in Time- and Frequency-Domains: A Two Dimensional View of Signals and Systems

Both signals and systems have dual personalities; the time domain and the frequency domain. For a deeper perspective, we should examine and understand both these identities because they offer complementary insights. An exponential signal, for instance, can be specified by its time domain description such as $e^{-2t}u(t)$ or by its Fourier transform (its frequency domain description) $\frac{1}{j\omega+2}$. The time-domain description depicts the waveform of a signal. The frequency-domain description portrays its spectral composition (relative amplitudes of its sinusoidal (or exponential) components and their phases). For the signal $e^{-2t}$, for instance, the time-domain description portrays the exponentially decaying signal with a time constant 0.5. The frequency-domain description characterizes it as a lowpass signal, which can be synthesized by sinusoids with amplitudes decaying with frequency roughly as $1/\omega$.

An LTIC system can also be described or specified in the time domain by its impulse response $h(t)$ or in the frequency domain by its transfer function $H(\omega)$. In Sec. 2.7, we studied intuitive insights in the system behavior offered by the impulse response, which consists of characteristic modes of the system. By purely qualitative reasoning, we saw that the system responds well to signals that are similar to the characteristic modes and responds poorly to signals which are very different from those modes. We also saw that the shape of the impulse response $h(t)$ determines the system time constant (speed of response), and pulse dispersion (spreading), which, in turn, determines the rate of pulse transmission.

The transfer function $H(\omega)$ specifies the frequency response; that is, the system response to exponential or sinusoidal input of various frequencies. This is precisely the filtering characteristic of the system.

Experienced electrical engineers instinctively think in both domains (the time and frequency) whenever possible. When they look at a signal, they consider, its waveform, the signal width (duration), and the rate at which the waveform decays. This is basically a time-domain perspective. They also think of the signal in terms of its frequency spectrum–that is, in terms of its sinusoidal components and their relative amplitudes and phases. This is the frequency-domain perspective.

When they think of a system, they think of its impulse response $h(t)$. The width of $h(t)$ indicates the time constant (response time); that is, how fast the system is capable of responding to an input, and how much dispersion (spreading) it will cause. This is the time-domain perspective. From the frequency-domain perspective, these engineers view a system as a filter, which selectively transmits certain frequency components and suppresses the others [frequency response $H(\omega)$]. Knowing the input signal spectrum and the frequency response of the system, they create a mental image of the output signal spectrum. This concept is precisely expressed by $Y(\omega) = F(\omega)H(\omega)$.

We can analyze LTI systems by time-domain techniques or by frequency-domain techniques. Then why learn both? The reason is that the two domains offer complementary insights into system behavior. Some aspects are easily grasped in one domain; other aspects may be easier to see in the other domain. Both the time-domain and the frequency-domain methods are as essential for the study of signals and systems as two eyes are essential to a human being for correct visual perception of reality. A person can see with either eye, but for proper perception of three dimensional-reality, both eyes are essential.

It is important to keep the two domains separate, and not to mix the entities in the two domains. If we are using the frequency domain to determine the system response, we must deal with all signals in terms of their spectra (Fourier transforms) and all systems in terms of their transfer functions. For example, to determine the system response $y(t)$ to an input $f(t)$, we must first convert the input signal into its frequency domain description $F(\omega)$. The system description also must be in the frequency-domain; that is, the transfer function $H(\omega)$. The output signal spectrum $Y(\omega) = F(\omega)H(\omega)$. Thus, the result (output) is also in the frequency domain. To determine the final answer $y(t)$, we must take the inverse transform of $Y(\omega)$.

## 4.6 Signal Energy

The signal energy $E_f$ of a signal $f(t)$ was defined in Chapter 1 as

$$E_f = \int_{-\infty}^{\infty} |f(t)|^2 \, dt \tag{4.62}$$

Signal energy can be related to the signal spectrum $F(\omega)$ by substituting Eq. (4.8b) in the above equation:

$$E_f = \int_{-\infty}^{\infty} f(t)f^*(t)\, dt = \int_{-\infty}^{\infty} f(t) \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} F^*(\omega)e^{-j\omega t}\, d\omega\right] dt$$

Here we used the fact that $f^*(t)$, being the conjugate of $f(t)$, can be expressed as the conjugate of the right-hand side of Eq. (4.8b). Now, interchanging the order of integration yields

$$\begin{aligned} E_f &= \frac{1}{2\pi}\int_{-\infty}^{\infty} F^*(\omega)\left[\int_{-\infty}^{\infty} f(t)e^{-j\omega t}\, dt\right] d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)F^*(\omega)\, d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2\, d\omega \end{aligned} \tag{4.63}$$



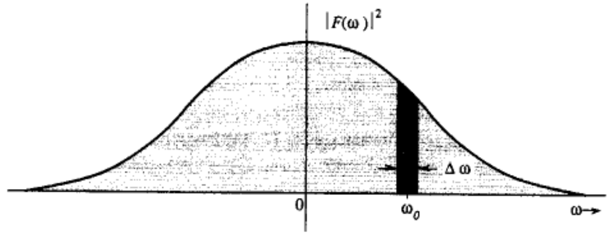



**Fig. 4.30** Interpretation of Energy spectral density of a signal.

Consequently,

$$E_f = \int_{-\infty}^{\infty} |f(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2\, d\omega \tag{4.64}$$

This is the statement of the well-known **Parseval's theorem** (for Fourier transform). A similar result was obtained in Eqs. (3.42)and (3.82) for a periodic signal and its Fourier series. This result allows us to determine the signal energy from either the time-domain specification $f(t)$ or the frequency-domain specification $F(\omega)$ of the same signal.

Equation (4.63) can be interpreted to mean that the energy of a signal $f(t)$ results from energies contributed by all the spectral components of the signal $f(t)$. The total signal energy is the area under $|F(\omega)^2|$ (divided by $2\pi$). If we consider a small band $\Delta\omega$ ($\Delta\omega \to 0$), as illustrated in Fig. 4.30, the energy $\Delta E_f$ of the spectral components in this band is the area of $|F(\omega)|^2$ under this band (divided by $2\pi$):

$$\Delta E_f = \frac{1}{2\pi}|F(\omega)|^2\, \Delta\omega = |F(\omega)|^2\, \Delta\mathcal{F} \qquad \frac{\Delta\omega}{2\pi} = \Delta\mathcal{F} \text{ Hz} \tag{4.65}$$

Therefore, the energy contributed by the components in this band of $\Delta\mathcal{F}$ (in hertz) is $|F(\omega)|^2\Delta\mathcal{F}$. The total signal energy is the sum of energies of all such bands and is indicated by the area under $|F(\omega)|^2$ as in Eq. (4.63). Therefore, $|F(\omega)|^2$ is the **energy spectral density** (per unit bandwidth in hertz).

For real signals, $F(\omega)$ and $F(-\omega)$ are conjugates, and $|F(\omega)|^2$ is an even function of $\omega$ because

$$|F(\omega)|^2 = F(\omega)F^*(\omega) = F(\omega)F(-\omega)$$

Therefore, Eq. (4.63) can be expressed as†

$$E_f = \frac{1}{\pi}\int_0^{\infty} |F(\omega)|^2\, d\omega \tag{4.66}$$

The signal energy $E_f$, which results from contributions from all the frequency components from $\omega = 0$ to $\infty$, is given by ($1/\pi$ times) the area under $|F(\omega)|^2$ from $\omega = 0$ to $\infty$. It follows that the energy contributed by spectral components of frequencies between $\omega_1$ and $\omega_2$ is

$$\Delta E_f = \frac{1}{\pi}\int_{\omega_1}^{\omega_2} |F(\omega)|^2\, d\omega \tag{4.67}$$

†In Eq. (4.66) it is assumed that $F(\omega)$ does not contain an impulse at $\omega = 0$. If such an impulse exists, it should be integrated separately with a multiplying factor of $1/2\pi$ rather than $1/\pi$.

■ **Example 4.16**

Find the energy of signal $f(t) = e^{-at}u(t)$. Determine the frequency $W$ (rad/s) so that the energy contributed by the spectral components of all the frequencies below $W$ is 95% of the signal energy $E_f$.

We have

$$E_f = \int_{-\infty}^{\infty} f^2(t)dt = \int_0^{\infty} e^{-2at}dt = \frac{1}{2a}$$

We can verify this result by Parseval's theorem. For this signal

$$F(\omega) = \frac{1}{j\omega + a}$$

and

$$E_f = \frac{1}{\pi}\int_0^{\infty} |F(\omega)|^2 d\omega = \frac{1}{\pi}\int_0^{\infty} \frac{1}{\omega^2 + a^2}\,d\omega = \frac{1}{\pi a}\tan^{-1}\frac{\omega}{a}\bigg|_0^{\infty} = \frac{1}{2a}$$

The band $\omega = 0$ to $\omega = W$ contains 95% of the signal energy, that is, $0.95/2a$. Therefore, from Eq. (4.67) with $\omega_1 = 0$ and $\omega_2 = W$, we obtain

$$\frac{0.95}{2a} = \frac{1}{\pi}\int_0^{W} \frac{d\omega}{\omega^2 + a^2} = \frac{1}{\pi a}\tan^{-1}\frac{\omega}{a}\bigg|_0^{W} = \frac{1}{\pi a}\tan^{-1}\frac{W}{a}$$

or

$$\frac{0.95\pi}{2} = \tan^{-1}\frac{W}{a} \Longrightarrow W = 12.706a \text{ rad/s}$$

This result indicates that the spectral components of $f(t)$ in the band from 0 (dc) to $12.706a$ rad/s ($2.02a$ Hz) contribute 95% of the total signal energy; all the remaining spectral components (in the band from $12.706a$ rad/s to $\infty$) contribute only 5% of the signal energy. ■

△ **Exercise E4.13**

Use Parseval's theorem to show that the energy of the signal

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

is $\frac{2\pi}{a}$. Hint: Find $F(\omega)$ using pair 3 and the symmetry property. ▽

### The Essential Bandwidth of a Signal

Spectra of most of the signals extend to infinity. However, because the energy of any practical signal is finite, the signal spectrum must approach 0 as $\omega \to \infty$. Most of the signal energy is contained within a certain band of $B$ Hz, and the energy contributed by the components beyond $B$ Hz is negligible. We can therefore suppress the signal spectrum beyond $B$ Hz with little effect on the signal shape and energy. The bandwidth $B$ is called the **essential bandwidth** of the signal. The criterion for selecting $B$ depends on the error tolerance in a particular application. We may, for example, select $B$ to be that band which contains 95% of the signal energy.† This figure may be higher or lower than 95%, depending on the precision needed. Using such a criterion, we can determine the essential bandwidth of a signal. The essential bandwidth $B$ for the signal $e^{-at}u(t)$, using 95% energy criterion, was determined in Example 4.16 to be $2.02a$ Hz.

†For lowpass signals, the essential bandwidth may also be defined as a frequency at which the value of the amplitude spectrum is a small fraction (about 1%) of its peak value. In Example 4.16, for instance, the peak value, which occurs at $\omega = 0$, is $1/a$.



Suppression of all the spectral components of $f(t)$ beyond the essential bandwidth results in a signal $\hat{f}(t)$, which is a close approximation of $f(t)$. If we use the 95% criterion for the essential bandwidth, the energy of the error (the difference) $f(t) - \hat{f}(t)$ is 5% of $E_f$.

### Energy Spectral Density From Autocorrelation Function

Correlation of a function $f(t)$ with itself is its **autocorrelation function** $\psi_f(t)$, which, for a real $f(t)$, is given by [see Eq. (3.32)]

$$\psi_f(t) = \int_{-\infty}^{\infty} f(x)f(x-t)\,dx \tag{4.68a}$$

Also, from Eq. (3.31) with $g(t) = f(t)$, it follows that

$$\psi_f(t) = f(t) * f(-t) \tag{4.68b}$$

From Eq. (4.68b) it is clear that

$$\psi_f(-t) = f(-t) * f(t) = \psi_f(t)$$

Therefore, for real $f(t)$, autocorrelation function $\psi_f(t)$ is an even function of $t$. The Fourier transform of Eq. (4.68b) yields

$$\psi_f(t) \Longleftrightarrow |F(\omega)|^2 \tag{4.69}$$

Therefore, the Fourier transform of the autocorrelation function is its energy spectral density $|F(\omega)|^2$. It is clear that $\psi_f(t)$ provides the spectral information of $f(t)$ directly.

The direct link of the autocorrelation function to the spectral information can be explained intuitively as follows. The autocorrelation function $\psi_f(t)$ is the correlation of a signal with itself delayed by $t$ seconds. A signal $f(t)$ correlates perfectly with itself with zero delay. But as the delay increases, the similarity decreases. Thus, the autocorrelation function $\psi_f(t)$ is a nonincreasing function of $t$. If $f(t)$ is a slowly varying signal (low frequency signal), it changes slowly with $t$. Consequently such a signal will show considerable similarity or correlation with itself even for relatively large delay. The autocorrelation function $\psi_f(t)$ decays slowly with $t$ and has a larger width. On the other hand, for a rapidly varying signal, the signal similarity will decrease rapidly with delay $t$ and $\psi_f(t)$ has a smaller width. Thus, the shape of $\psi_f(t)$ has a direct link to spectral information of $f(t)$.

## 4.7 Application to Communications: Amplitude Modulation

**Modulation** causes a spectral shift in a signal and is used to gain certain advantages mentioned in Sec. 4.3-5. Broadly speaking, there are two classes of modulation: amplitude (linear) modulation and angle (nonlinear) modulation, which are the subject of the next two sections. In this section, we shall discuss some practical forms of amplitude modulation.

### 4.7-1 Double Sideband, Suppressed Carrier (DSB-SC) Modulation

In amplitude modulation, the amplitude $A$ of the carrier $A\cos(\omega_c t + \theta_c)$ is varied in some manner with the **baseband** (message)† signal $m(t)$ (known as the

†The term baseband is used to designate the band of frequencies of the signal delivered by the source or the input transducer.

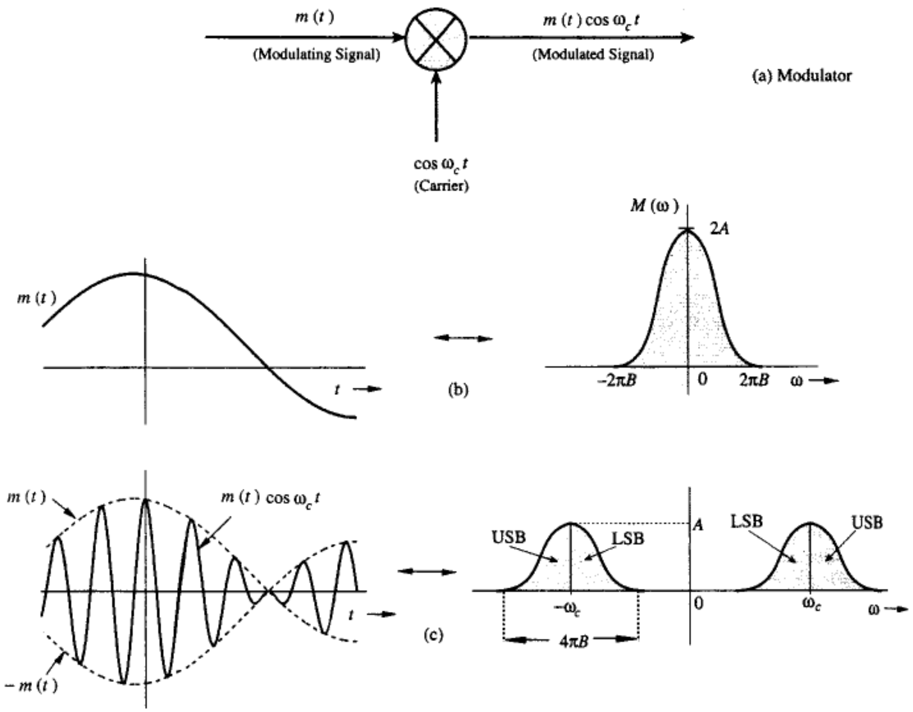



**Fig. 4.31** DSB-SC modulation.

**modulating signal**). The frequency $\omega_c$ and the phase $\theta_c$ are constant. We can assume $\theta_c = 0$ without a loss of generality. If the carrier amplitude $A$ is made directly proportional to the modulating signal $m(t)$, the modulated signal is $m(t) \cos \omega_c t$ (Fig. 4.31a). As was indicated earlier [Eq. (4.41)], this type of modulation simply shifts the spectrum of $m(t)$ to the carrier frequency (Fig. 4.31c). Thus, if

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

then

$$m(t) \cos \omega_c t \Longleftrightarrow \frac{1}{2}[M(\omega + \omega_c) + M(\omega - \omega_c)] \tag{4.70}$$

Recall that $M(\omega - \omega_c)$ is $M(\omega)$ shifted to the right by $\omega_c$ and $M(\omega + \omega_c)$ is $M(\omega)$ shifted to the left by $\omega_c$. Thus, the process of modulation shifts the spectrum of the modulating signal to the left and the right by $\omega_c$. Note also that if the bandwidth of $m(t)$ is $B$ Hz, then, as indicated in Fig. 4.31c, the bandwidth of the modulated signal is $2B$ Hz. We also observe that the modulated signal spectrum centered at $\omega_c$ is composed of two parts: a portion that lies above $\omega_c$, known as the **upper sideband (USB)**, and a portion that lies below $\omega_c$, known as the **lower sideband (LSB)**. Similarly, the spectrum centered at $-\omega_c$ has upper and lower sidebands. This form of modulation is called **double sideband (DSB)** modulation for obvious reason.



The relationship of $B$ to $\omega_c$ is of interest. Figure 4.31c shows that $\omega_c \geq 2\pi B$ in order to avoid the overlap of the spectra centered at $\pm\omega_c$. If $\omega_c < 2\pi B$, the spectra overlap and the information of $m(t)$ is lost in the process of modulation, a loss which makes it impossible to get back $m(t)$ from the modulated signal $m(t) \cos \omega_c t$.†

### ■ Example 4.17

For a baseband signal $m(t) = \cos \omega_m t$, find the DSB signal, and sketch its spectrum. Identify the upper and lower sidebands.

We shall work this problem in the frequency-domain as well as the time-domain in order to clarify the basic concepts of DSB-SC modulation. 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

$$M(\omega) = \pi[\delta(\omega - \omega_m) + \delta(\omega + \omega_m)]$$

The spectrum consists of two impulses located at $\pm\omega_m$, as depicted in Fig. 4.32a. The DSB-SC (modulated) spectrum, as indicated by Eq. (4.70), is the baseband spectrum in Fig. 4.32a shifted to the right and the left by $\omega_c$ (times one-half), as depicted in Fig. 4.32b. This spectrum consists of impulses at $\pm(\omega_c - \omega_m)$ and $\pm(\omega_c + \omega_m)$. The spectrum beyond $\omega_c$ is the upper sideband (USB), and the one below $\omega_c$ is the lower sideband (LSB). Observe that the DSB-SC spectrum does not have the component of the carrier frequency $\omega_c$. This is why it is called **double sideband-suppressed carrier (DSB-SC)** .

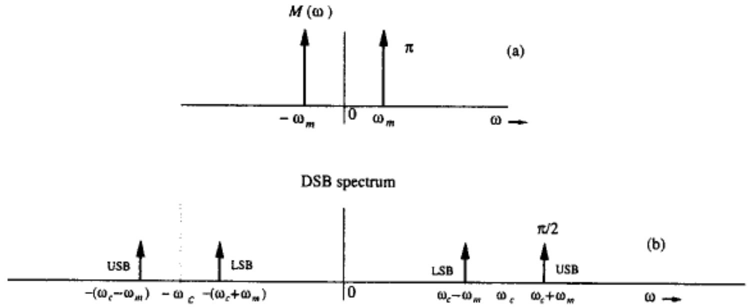



**Fig. 4.32** An example of DSB-SC modulation.

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

$$\begin{aligned}\varphi_{\text{DSB-SC}}(t) &= m(t) \cos \omega_c t \\ &= \cos \omega_m t \cos \omega_c t \\ & \frac{1}{2}[\cos(\omega_c + \omega_m)t + \cos(\omega_c - \omega_m)t]\end{aligned} \tag{4.71}$$

†Practical factors may impose additional restrictions on $\omega_c$. For instance, in broadcast applications, a radiating antenna can radiate only a narrowband without distortion. This restriction implies that to avoid distortion caused by the radiating antenna, $\omega_c/2\pi B \gg 1$. The broadcast band AM radio, for instance, with $B = 5$ kHz and the band of 550 to 1600 kHz for carrier frequency give a ratio of $\omega_c/2\pi B$ roughly in the range of 100 to 300.

This result shows that when the baseband (message) signal is a single sinusoid of frequency $\omega_m$, the modulated signal consists of two sinusoids: the component of frequency $\omega_c + \omega_m$ (the upper sideband), and the component of frequency $\omega_c - \omega_m$ (the lower sideband). Figure 4.32b illustrates precisely the spectrum of $\varphi_{\text{DSB-SC}}(t)$. Thus, each component of frequency $\omega_m$ in the modulating signal results into two components of frequencies $\omega_c + \omega_m$ and $\omega_c - \omega_m$ in the modulated signal. This being a DSB-SC (suppressed carrier) modulation, there is no component of the carrier frequency $\omega_c$ on the right-hand side of the above equation as expected.† ■

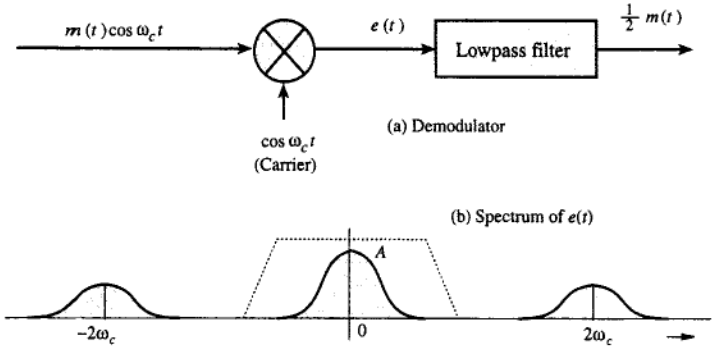



**Fig. 4.33** Demodulation of DSB-SC.

### Demodulation of DSB-SC Signals

The DSB-SC modulation translates or shifts the frequency spectrum to the left and the right by $\omega_c$ (that is, at $+\omega_c$ and $-\omega_c$), as seen from Eq. (4.70). To recover the original signal $m(t)$ from the modulated signal, we must 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**, or **detection**. Observe that if the modulated signal spectrum in Fig. 4.31c is shifted to the left and to the right by $\omega_c$ (and halved), we obtain the spectrum illustrated in Fig. 4.33b, which contains the desired baseband spectrum in addition to an unwanted spectrum at $\pm 2\omega_c$. The latter can be suppressed by a lowpass 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 lowpass filter, as depicted in Fig. 4.33a. We can verify this conclusion directly in the time-domain by observing that the signal $e(t)$ in Fig. 4.33a is

$$\begin{aligned} e(t) &= m(t)\cos^2\omega_c t \\ &= \frac{1}{2}[m(t) + m(t)\cos 2\omega_c t] \end{aligned} \tag{4.72a}$$

†The term suppressed carrier does not necessarily mean absence of the spectrum at the carrier frequency. The term "suppressed carrier" merely implies that there is no discrete component of the carrier frequency. Since no discrete component exists, the spectrum of DSB-SC does not have impulses at $\pm\omega_c$, a fact which further implies that the modulated signal $m(t)\cos\omega_c t$ does not contain a term of the form $k\cos\omega_c t$ (assuming that $m(t)$ has a zero mean value).



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

$$E(\omega) = \frac{1}{2}M(\omega) + \frac{1}{4}[M(\omega + 2\omega_c) + M(\omega - 2\omega_c)] \tag{4.72b}$$

Hence, $e(t)$ consists of two components $\frac{1}{2}m(t)$ and $\frac{1}{2}m(t)\cos 2\omega_c t$, with their spectra, as illustrated in Fig. 4.33b. The spectrum of the second component, being a modulated signal with carrier frequency $2\omega_c$, is centered at $\pm 2\omega_c$. Hence, this component is suppressed by the lowpass filter in Fig. 4.33a. The desired component $\frac{1}{2}M(\omega)$, being a lowpass spectrum (centered at $\omega = 0$), passes through the filter unharmed, resulting in the output $\frac{1}{2}m(t)$.

A possible form of lowpass filter characteristics is depicted (dotted) in Fig. 4.33b. 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. We shall demonstrate in Example 4.18 that both, the phase and frequency synchronism, are extremely critical.

#### ■ Example 4.18

Discuss the effect of lack of frequency and phase coherence (synchronism) between the carriers at the modulator (transmitter) and the demodulator (receiver) in DSB-SC.

Let the modulator carrier be $\cos\omega_c t$ (Fig. 4.31a). For the demodulator in Fig. 4.33a, we shall consider two cases: (1) the first case with carrier $\cos(\omega_c t + \theta)$ (phase error of $\theta$) and (2) the second case with carrier $\cos(\omega_c + \Delta\omega)t$ (frequency error $\Delta\omega$).

(a) With the demodulator carrier $\cos(\omega_c t + \theta)$ (instead of $\cos\omega_c t$) in Fig. 4.33a, the multiplier output is $e(t) = m(t)\cos\omega_c t\cos(\omega_c t + \theta)$ instead of $m(t)\cos^2\omega_c t$. Using the trigonometric identity, we obtain

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

The spectrum of the component $\frac{1}{2}m(t)\cos(2\omega_c t + \theta)$ is centered at $\pm 2\omega_c$. Consequently, it will be filtered out by the lowpass filter at the output. The component $\frac{1}{2}m(t)\cos\theta$ is the signal $m(t)$ multiplied by a constant $\frac{1}{2}\cos\theta$. The spectrum of this component is centered at $\omega = 0$ (lowpass spectrum), and will pass through the lowpass filter at the output, yielding the output $\frac{1}{2}m(t)\cos\theta$.

If $\theta$ is constant, the phase asynchronism merely yields an attenuated output (by a factor $\cos\theta$). Unfortunately, in practice, $\theta$ is often the phase difference between the carriers generated by two distant generators, and varies randomly with time. This variation would result in an output whose gain varies randomly with time.

(b) In the case of frequency error, the demodulator carrier is $\cos(\omega_c + \Delta\omega)t$. This situation is very similar to the phase error case in (a) with $\theta$ replaced by $(\Delta\omega)t$. Following the analysis in part (a), we can express the demodulator product $e(t)$ as

$$\begin{aligned} e(t) &= m(t)\cos\omega_c t\,\cos(\omega_c + \Delta\omega)t \\ &= \frac{1}{2}m(t)[\cos(\Delta\omega)t + \cos(2\omega_c + \Delta\omega)t] \end{aligned}$$

The spectrum of the component $\frac{1}{2}m(t)\cos(2\omega_c + \Delta\omega)t$ is centered at $\pm(2\omega_c + \Delta\omega)$. Consequently, this component will be filtered out by the lowpass filter at the output. The component $\frac{1}{2}m(t)\cos(\Delta\omega)t$ is the signal $m(t)$ multiplied by a low frequency carrier of frequency $\Delta\omega$. The spectrum of this component is centered at $\pm\Delta\omega$. In practice, the frequency error $(\Delta\omega)$ is usually very small. Hence, the signal $\frac{1}{2}m(t)\cos(\Delta\omega)t$ (whose spectrum is centered at $\pm\Delta\omega$) is a lowpass signal and passes through the lowpass filter at the output, resulting in the output $\frac{1}{2}m(t)\cos(\Delta\omega)t$. The output is the desired signal $m(t)$ multiplied by a very low frequency sinusoid $\cos(\Delta\omega)t$. Clearly, the output in this case is not merely an attenuated replica of the desired signal $m(t)$, but represents $m(t)$ multiplied by a time-varying gain $\cos(\Delta\omega)t$. If, for instance, the transmitter and the receiver carrier frequencies differ just by 1 Hz, the output will be the desired signal $m(t)$ multiplied by a time-varying signal whose gain goes from the maximum to 0 every half second. This is like some restless child fiddling with the volume control knob of a receiver, going from maximum volume to zero volume every half second. This kind of distortion (called the **beat effect**) is beyond repair. ■

### 4.7-2 Amplitude Modulation (AM)

For the suppressed carrier scheme just discussed, a receiver must generate a carrier in frequency and phase synchronism with the carrier at the transmitter that may be located hundreds or thousands of miles away. This situation calls for a sophisticated receiver, which could be quite costly. The other alternative is for the transmitter to transmit 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 much larger power, a rather expensive procedure. In point-to-point communications, where there is one transmitter for each receiver, substantial complexity in the receiver system can be justified, provided there is a large enough saving in expensive high-power transmitting equipment. On the other hand, for a broadcast system with a multitude of receivers for each transmitter, it is more economical to have one expensive high-power transmitter and simpler, less expensive receivers. The second option (transmitting a carrier along with the modulated signal) is the obvious choice in this case. This is the so-called AM (amplitude modulation), in which the transmitted signal $\varphi_{AM}(t)$ is given by

$$\varphi_{AM}(t) = A\cos\omega_c t + m(t)\cos\omega_c t \tag{4.73a}$$

$$= [A + m(t)]\cos\omega_c t \tag{4.73b}$$

Recall that the DSB-SC signal is $m(t)\cos\omega_c t$. From Eq. (4.73b) it follows that the AM signal is identical to the DSB-SC signal with $A + m(t)$ as the modulating signal [instead of $m(t)$]. Therefore, to sketch $\varphi_{AM}(t)$, we sketch $A + m(t)$ and $-[A + m(t)]$ and fill in between with the sinusoid of the carrier frequency. Two cases are considered in Fig. 4.34. In the first case, $A$ is large enough so that $A + m(t) \geq 0$ (is nonnegative) for all values of $t$. In the second case, $A$ is not large enough to satisfy this condition. In the first case, the envelope (Fig. 4.34d) has the same shape as $m(t)$ (although riding on a dc of magnitude $A$). In the second case the envelope shape is not $m(t)$, for some parts get rectified (Fig. 4.34e). Thus, we can detect the desired signal $m(t)$ by detecting the envelope in the first case. In the second case, such a detection is not possible. We shall see that the envelope detection is



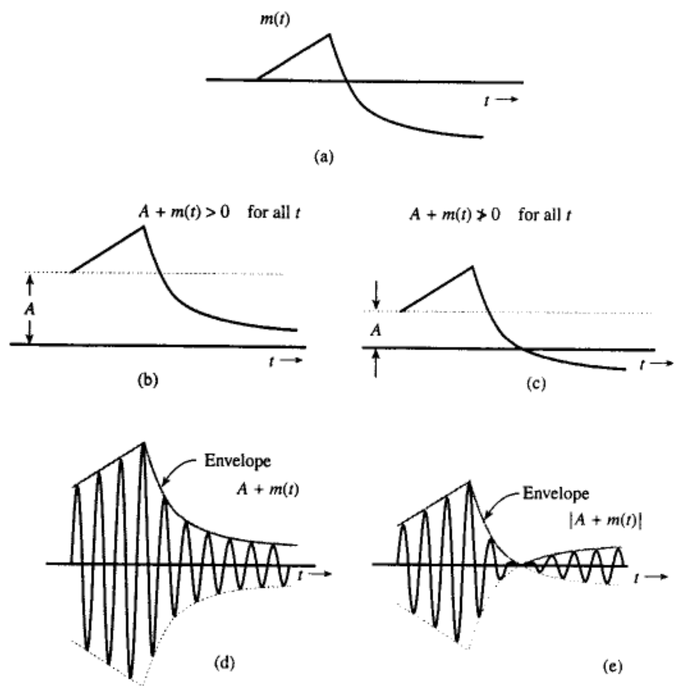



**Fig. 4.34** AM signal and its envelope.

an extremely simple and inexpensive operation, which does not require generation of a local carrier for the demodulation. But as seen above the envelope of AM 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 $t$. Thus, the condition for envelope detection of an AM signal is

$$A + m(t) \geq 0 \qquad \text{for all } t \tag{4.74}$$

If $m_p$ is the peak amplitude (positive or negative) of $m(t)$ (see Fig. 4.34), then $m(t) \geq -m_p$. Hence, the condition (4.74) is equivalent to†

$$A \geq m_p \tag{4.75}$$

Thus the minimum carrier amplitude required for the viability of envelope detection is $m_p$. This point is clearly illustrated in Fig. 4.34.

We define the modulation index $\mu$ as

$$\mu = \frac{m_p}{A} \tag{4.76}$$

†In case the negative and the positive peak amplitudes are not identical, $m_p$ in condition (4.75) is the absolute negative peak amplitude.

where $A$ is the carrier amplitude. Note that $m_p$ is a constant of the signal $m(t)$. Because $A \geq m_p$ and because there is no upper bound on $A$, it follows that

$$0 \leq \mu \leq 1 \tag{4.77}$$

as the required condition for the viability of demodulation of AM by an envelope detector.

When $A < m_p$, Eq. (4.76) 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 $\mu$ (see Prob. 4.7-4). The envelope detector, which is considerably simpler and less expensive than the synchronous detector, can be used only for $\mu \leq 1$.

■ **Example 4.19**

Sketch $\varphi_{AM}(t)$ for modulation indices of $\mu = 0.5$ (50% modulation) and $\mu = 1$ (100% modulation), 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, $m_p = B$ and the modulation index according to Eq. (4.76) is

$$\mu = \frac{B}{A}$$

Hence, $B = \mu A$ and

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

Therefore

$$\varphi_{AM}(t) = [A + m(t)]\cos\omega_c t = A[1 + \mu\cos\omega_m t]\cos\omega_c t \tag{4.78}$$

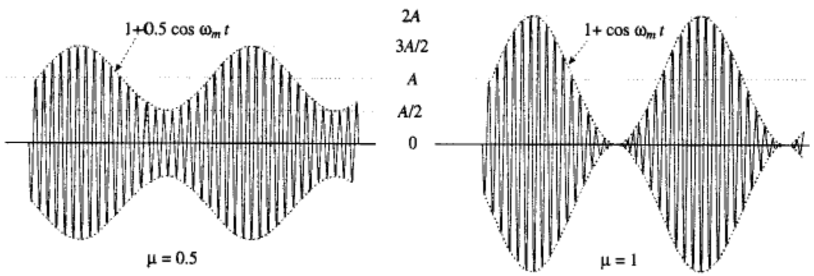



**Fig. 4.35** Tone-modulated AM (a) $\mu = 0.5$. (b) $\mu = 1$.

Figures 4.35a and b show the modulated signals corresponding to $\mu = 0.5$ and $\mu = 1$, respectively. ■

### Demodulation of AM: The Envelope Detector

The AM signal can be demodulated coherently by a locally generated carrier (see Prob. 4.7-4). However, coherent, or synchronous, demodulation of AM (with $\mu \leq 1$) will defeat the very purpose of AM and, hence, is rarely used in practice.



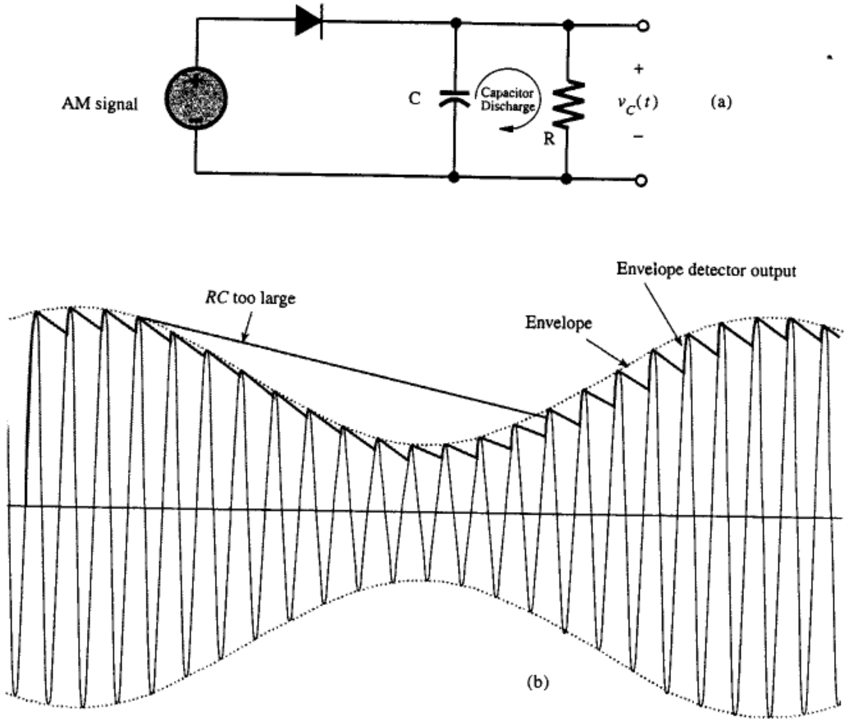



**Fig. 4.36** Envelope detector.

We shall consider here one of the noncoherent methods of AM demodulation, the **envelope detection.**†

In an envelope detector, the output of the detector follows the envelope of the (modulated) input signal. The circuit illustrated in Fig. 4.36a functions as an envelope detector. During the positive cycle of the input signal, the diode conducts and the capacitor $C$ charges up to the peak voltage of the input signal (Fig. 4.36b). As the input signal falls below this peak value, the diode is cut off, because the capacitor voltage (which is very nearly the peak voltage) is greater than the input signal voltage, a circumstance 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. When the input signal becomes greater than the capacitor voltage, the diode conducts again. The capacitor again charges to the peak value of this (new) cycle. As the input voltage falls below the new peak value, the diode cuts off again and the capacitor discharges slowly during the cutoff period, a process that changes the capacitor voltage very slightly.

†There are also other methods of noncoherent detection. The rectifier detector consists of a rectifier followed by a lowpass filter. This method is also simple and almost as inexpensive as the envelope detector[4]. The nonlinear detector, although simple and inexpensive, results in a distorted output.

In this manner, 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. Thus, the output voltage $v_C(t)$ follows the envelope of the input. The capacitor discharge between positive peaks, however, causes a ripple signal of frequency $\omega_c$ in the output. This ripple can be reduced by increasing the time constant $RC$ so that the capacitor discharges very little between the positive peaks ($RC \leq 1/\omega_c$). Making $RC$ too large, however, would make it impossible for the capacitor voltage to follow the envelope (see Fig. 4.36b). Thus, $RC$ should be large compared to $1/\omega_c$ but should be small compared to $1/2\pi B$, where $B$ is the highest frequency in $m(t)$. Incidentally, these two conditions also require that $\omega_c \gg 2\pi B$, a condition necessary for a well-defined envelope.

The envelope-detector output $v_C(t)$ is $A + m(t)$ plus a ripple of frequency $\omega_c$. The dc term $A$ can be blocked out by a capacitor or a simple $RC$ highpass filter. The ripple may be reduced further by another (lowpass) $RC$ filter. In the case of audio signals, the speakers cannot respond to the high frequency ripple, and therefore, they act as lowpass filters themselves.

### 4.7-3 Single Sideband Modulation (SSB)

Figures 4.37a and 4.37b show the baseband spectrum $M(\omega)$, and the spectrum of the DSB-SC modulated signal $m(t)\cos\omega_c t$. The DSB spectrum in Fig. 4.37b has two sidebands: the upper sideband (USB) and the lower sideband (LSB), both containing complete information of $M(\omega)$ [see Eq. (4.10)]. Clearly, it is redundant to transmit both sidebands, a process which requires twice the bandwidth of the baseband signal. A scheme where only one sideband is transmitted is known as **single sideband (SSB) transmission**, which requires only one-half the bandwidth of the DSB signal. Thus, we transmit only the upper sidebands (Figures 4.37c) or only the lower sidebands (Fig. 4.37d).

An SSB signal can be coherently (synchronously) demodulated. For example, multiplication of a USB signal (Fig. 4.37c) by $\cos\omega_c t$ shifts its spectrum to the left and to the right by $\omega_c$, yielding the spectrum in Fig. 4.37e. Lowpass filtering of this signal yields the desired baseband signal. The case is similar with LSB signal. Hence, demodulation of SSB signals is identical to that of DSB-SC signals, and the synchronous demodulator in Fig. 4.33a can demodulate SSB signals. Note that we are talking of SSB signals without an additional carrier. Hence, they are suppressed carrier signals (SSB-SC).

**■ Example 4.20**

Find the USB (the upper sideband) and LSB (the lower sideband) signals when $m(t) = \cos\omega_m t$. Sketch their spectra, and show that these SSB signals can be demodulated using the synchronous demodulator in Fig. 4.33a.

The DSB-SC signal for this case is

$$\begin{aligned}\varphi_{\text{DSB-SC}}(t) &= m(t)\cos\omega_c t \\ &= \cos\omega_m t \cos\omega_c t \\ &\frac{1}{2}[\cos(\omega_c - \omega_m)t + \cos(\omega_c + \omega_m)t]\end{aligned} \tag{4.79}$$

As pointed out in Example 4.17, the terms $\frac{1}{2}\cos(\omega_c + \omega_m)t$ and $\frac{1}{2}\cos(\omega_c - \omega_m)t$ represent the upper and lower sidebands, respectively. Figure 4.38a and b show the spectra



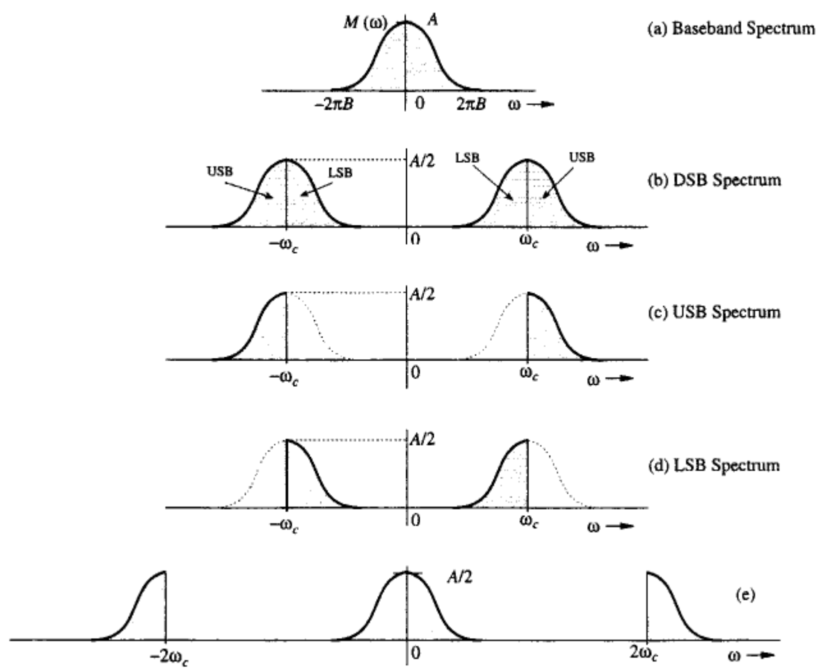



**Fig. 4.37** Single sideband transmission.

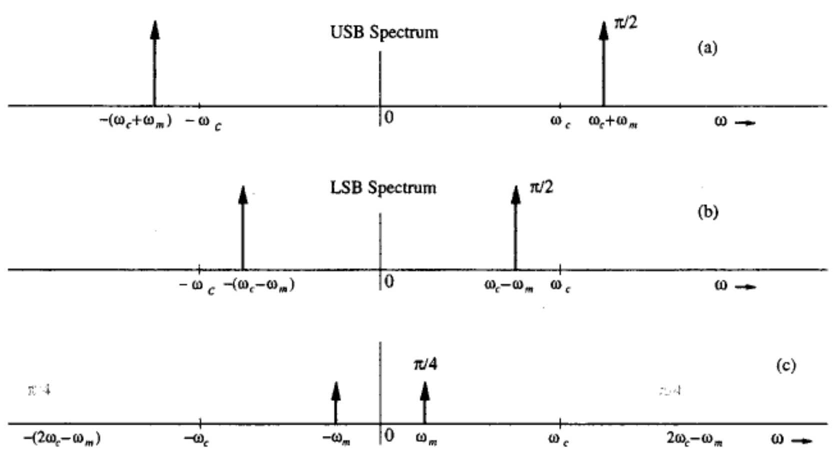



**Fig. 4.38** Single sideband spectra for $m(t) = \cos\omega_m t$.

of the upper and lower sidebands. Observe that these spectra can be obtained from the DSB-SC spectrum in Fig. 4.32b by suppressing the undesired sidebands using a proper filter. For instance, the USB signal in Fig. 4.38a can be obtained by passing the DSB-SC signal (Fig. 4.32b) through a highpass filter of cutoff frequency $\omega_c$. Similarly, the LSB signal in Fig. 4.38b can be obtained by passing the DSB-SC signal through a lowpass filter of cutoff frequency $\omega_c$.

If we apply the LSB signal $\frac{1}{2}\cos(\omega_c - \omega_m)t$ to the synchronous demodulator in fig. 4.33a, the multiplier output is

$$e(t) = \frac{1}{2}\cos(\omega_c - \omega_m)t \cos \omega_c t$$
$$= \frac{1}{4}[\cos \omega_m t + \cos(2\omega_c - \omega_m)t]$$

The term $\frac{1}{4}\cos(2\omega_c - \omega_m)t$ is suppressed by the lowpass filter, a fact which results in the desired output $\frac{1}{4}\cos\omega_m t$ (which is $m(t)/4$). The spectrum of this term is $\pi[\delta(\omega+\omega_0) + \delta(\omega-\omega_0)]/4$, as depicted in Fig. 4.38c. In the same way we can show that the USB signal can be demodulated by the synchronous demodulator.

In frequency-domain, demodulation (multiplication by $\cos\omega_c t$ amounts to shifting the LSB spectrum (Fig. 4.38b) to the left and the right by $\omega_c$ (times one-half) and then suppressing the high frequency, as illustrated in Fig. 4.38c. The resulting spectrum represents the desired signal $\frac{1}{4}m(t)$. ■

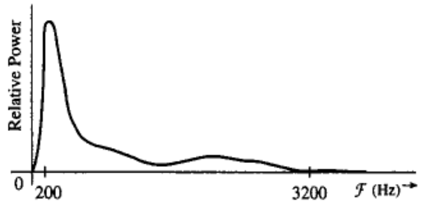



**Fig. 4.39** Voice Spectrum.

### Generation of SSB Signals

Two methods are commonly used to generate SSB signals. The first method, the **selective-filtering method** uses sharp cutoff filters to eliminate the undesired sideband, and the second method uses phase-shifting networks[4] to achieve the same goal.† We shall consider here only the first method.

The selective-filtering Method is the most commonly used method of generating SSB signals. In this method, 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 $\omega_c$ unattenuated and completely suppress all components below $\omega_c$. Such an operation requires an ideal filter, which is unrealizable. It can, however, be realized closely if there is some separation between the passband and the stopband. Fortunately, the voice signal provides this condition, because its spectrum shows little power content at the origin (Fig. 4.39). Moreover, articulation tests show that for speech signals, frequency components below 300 Hz are not important. In other words, we may suppress all speech components below 300 Hz without affecting the intelligibility

†Yet another method, known as Weaver's method, is also used to generate SSB signals.



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 $\omega_c$. For signals, which have considerable power at low frequencies (around $\omega = 0$), SSB techniques cause considerable distortion. Such is the case with video signals. Consequently, for video signals, instead of SSB, we use another technique, the **vestigial sideband (VSB)**, which is a compromise between SSB and DSB. It inherits the advantages of SSB and DSB but avoids their disadvantages. VSB signals are relatively easy to generate, and their bandwidth is only slightly (typically 25%) greater than that of the SSB signals. In VSB signals, instead of rejecting one sideband completely (as in SSB), we accept a gradual cutoff off of one sideband[4].

## 4.8 Angle Modulation

A sinusoid is characterized by its amplitude and angle (which includes its frequency and phase). In amplitude modulated signals, the information content of the baseband (message) signal $m(t)$ appears in the amplitude variations of the carrier. In **angle modulation** discussed in this section, the information content of $m(t)$ is carried by the angle of the carrier. Angle modulation also goes by the name **exponential modulation**.

The generalized angle modulated (or exponentially modulated) carrier can be described as

$$\varphi_{\text{EM}}(t) = A\cos[\omega_c t + k\psi(t)] \tag{4.80}$$

where $k$ is an arbitrary constant and $\psi(t)$, which is a measure of $m(t)$, is obtained by an invertible linear operation on $m(t)$. In other words, $\psi(t)$ is the output of some linear system with a suitable transfer function $H(s)$ when the input is $m(t)$, as depicted in Fig. 4.40.† If $h(t)$ is the unit impulse response of this system; that is, if $h(t) \Longleftrightarrow H(s)$, then

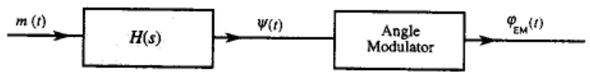



**Fig. 4.40** Generation of angle modulated signal.

$$\psi(t) = \int_{-\infty}^{t} m(\alpha)h(t-\alpha)\,d\alpha \tag{4.81}$$

By selecting suitable $h(t)$, we can obtain a variety of subclasses of angle modulation. For instance, if we select $h(t) = u(t)$, the resulting form is the well-known **frequency modulation (FM)**. In contrast, use of $h(t) = \delta(t)$ leads to **phase modulation (PM)**. These are but two of the infinite possibilities. Although in

‡Similarly, suppression of speech-signal components above 3500 Hz causes no appreciable change in intelligibility.

†Because $H(s)$ is required to be invertible, we can obtain $m(t)$ by passing $\psi(t)$ through the inverse linear system, which has the transfer function $1/H(s)$.

digital communication, use of phase and frequency modulation is common, the so-called broadcast FM is not FM in the classical sense, but is a generalized angle modulation because of the inclusion of the **preemphasis** filter used to improve its noise suppressing abilities. It is called FM for historical reason in the sense that angle modulation was first conceived and introduced in the form of frequency modulation. The broadcast FM, although originating as true FM in the laboratories, was modified in broadcasting for better performance. Yet, the term FM continued to be used to describe this scheme.

In amplitude modulation, the carrier frequency is constant, but the amplitude changes with $m(t)$. In contrast, in angle modulation, the carrier amplitude is always constant, but the carrier frequency varies continuously with the message $m(t)$. By definition, a sinusoidal signal is expected to have a constant frequency; hence, the variation of frequency with time appears to be contradictory to the conventional definition of a sinusoidal signal frequency. Therefore, we must generalize the notion of a sinusoid so as to make allowance for variation of frequency with time. This generalization leads us to a new concept of **instantaneous frequency**.

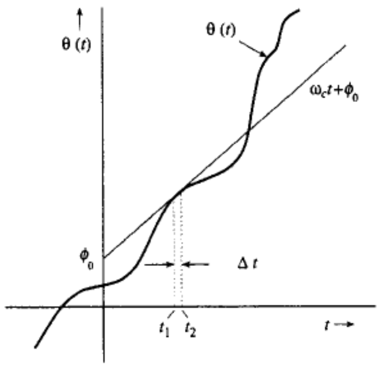



**Fig. 4.41** Concept of instantaneous frequency.

### 4.8-1 The Concept of Instantaneous Frequency

As seen above, the carrier frequency is changing continuously every instant in FM. Prima facie, this does not make much sense because to define a frequency, we must have a sinusoidal signal at least over one cycle with the same frequency. We cannot imagine a sinusoid whose frequency is different at every instant. This problem reminds us of our first encounter with the concept of **instantaneous velocity** in our beginning mechanics course. Until that time, we were used to thinking of velocity as being constant over a time interval and were incapable of even imagining that velocity could vary at each instant. But after some mental struggle, the idea gradually sinks in. We never forget, however, the wonder and amazement that was caused by the idea when it was first introduced. A similar experience awaits the reader with the concept of instantaneous frequency.

Let us consider a generalized sinusoidal signal $\varphi(t)$ given by



$$\varphi(t) = A \cos \theta(t) \tag{4.82}$$

where $\theta(t)$, the **generalized angle**, is a function of $t$. Figure 4.41 illustrates a hypothetical case of $\theta(t)$. The generalized angle for a conventional sinusoid $A \cos(\omega_c t + \phi_o)$ is $\omega_c t + \phi_o$. This plot, a straight line with a slope $\omega_c$ and intercept $\phi_o$, is also illustrated in Fig. 4.41. The plot of $\theta(t)$ for the hypothetical case happens to be tangential to the angle $(\omega_c t + \phi_o)$ at some instant $t$. The crucial point is that over a small interval $\Delta t \to 0$, the signal $\varphi(t) = A \cos \theta(t)$ and the sinusoid $A \cos(\omega_c t + \phi_o)$ are identical; that is,

$$\varphi(t) = A \cos(\omega_c t + \phi_o) \qquad t_1 < t < t_2$$

We are certainly justified in saying that over this small interval $\Delta t$, the frequency of $\varphi(t)$, is $\omega_c$. Because $(\omega_c t + \phi_o)$ is tangential to $\theta(t)$, the frequency of $\varphi(t)$ is the slope of its angle $\theta(t)$ over this small interval. We can generalize this concept at every instant and say that the instantaneous frequency $\omega_i$ at any instant $t$ is the slope of $\theta(t)$ at $t$. Thus, for $\varphi(t)$ in Eq. (4.82), the instantaneous frequency $\omega_i(t)$ is given by

$$\omega_i(t) = \frac{d\theta}{dt} \tag{4.83a}$$

$$\theta(t) = \int_{-\infty}^{t} \omega_i(\alpha)\, d\alpha \tag{4.83b}$$

For a conventional sinusoid $A \cos(\omega_c t + \phi_o)$, we have $\theta(t) = \omega_c t + \phi_o$, and $\omega_i(t) = d\theta(t)/dt = \omega_c$, a constant, as desired. Clearly, the generalized definition of instantaneous frequency does not conflict with our old notion of frequency.

Now we can see the possibility of transmitting the information of $m(t)$ by varying the angle $\theta$ of a carrier. Two simple possibilities are **phase modulation (PM)** and **frequency modulation (FM)**. In PM, the angle $\theta(t)$ is varied linearly with $m(t)$ :

$$\theta(t) = \omega_c t + k_p m(t) \tag{4.84a}$$

where $k_p$ is a constant and $\omega_c$ is the carrier frequency. The resulting PM wave is

$$\varphi_{\text{PM}}(t) = A \cos\left[\omega_c t + k_p m(t)\right] \tag{4.84b}$$

The instantaneous frequency $\omega_i(t)$ in this case is given by

$$\omega_i(t) = \frac{d\theta}{dt} = \omega_c + k_p \dot{m}(t) \tag{4.84c}$$

Hence in phase modulation, the instantaneous frequency $\omega_i$ varies linearly with the derivative of the modulating signal. If the instantaneous frequency $\omega_i$ is varied linearly with the modulating signal, we have frequency modulation. Thus, in FM, the instantaneous frequency $\omega_i$ is

$$\omega_i(t) = \omega_c + k_f m(t) \tag{4.85a}$$

where $k_f$ is a constant. From Eq. (4.83b), we find the angle $\theta(t)$ as

$$\theta(t) = \int_{-\infty}^{t} [\omega_c + k_f m(\alpha)]\, d\alpha$$
$$= \omega_c t + k_f \int_{-\infty}^{t} m(\alpha)\, d\alpha \tag{4.85b}$$

Here we have assumed the constant term in $\theta(t)$ to be zero without loss of generality. Thus, the FM wave is

$$\varphi_{\text{FM}}(t) = A \cos \left[\omega_c t + k_f \int_{-\infty}^{t} m(\alpha)\, d\alpha\right] \tag{4.85c}$$

Observe that both PM and FM are special cases of the exponentially modulated signal $\varphi_{\text{EM}}(t)$ in Eq. (4.80). If $h(t) = \delta(t)$ in Eq. (4.81), then use of the sampling property of the impulse in Eq. (4.81) yields $\psi(t) = m(t)$, and Eq. (4.80) reduces to PM in Eq. (4.84b). Similarly, if $h(t) = u(t)$, then the fact that $u(t-\alpha) = 1$ over $-\infty < \alpha \le t$ yields $\int m(\alpha)h(t-\alpha)\, d\alpha = \int m(\alpha)\, d\alpha$, and Eq. (4.80) reduces to FM in Eq. (4.85c).

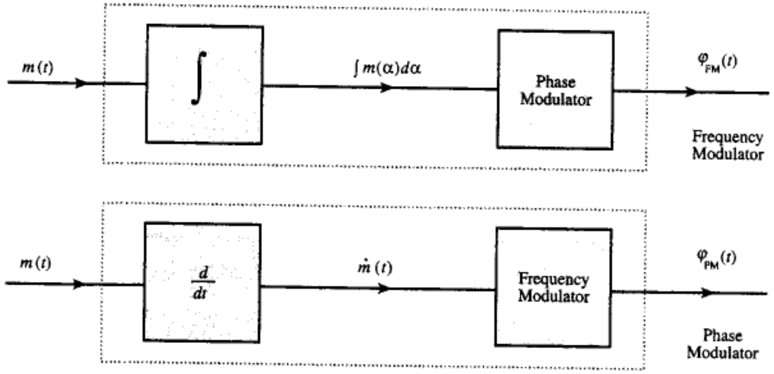



**Fig. 4.42** Phase and frequency modulation are inseparable.

### All In The Family

Equations (4.84b) and (4.85c) indicate that PM and FM are not only very similar but are inseparable. Replacing $m(t)$ in Eq. (4.84b) with $\int^t m(\alpha)\, d\alpha$ changes PM into FM. Thus, a signal that is an FM wave corresponding to $m(t)$ is also the PM wave corresponding to $\int^t m(\alpha)\, d\alpha$ (Fig. 4.42a). Similarly, a PM wave corresponding to $m(t)$ is the FM wave corresponding to $\dot{m}(t)$ (Fig. 4.42b).

We conclude that just by looking at an angle-modulated carrier, we cannot tell whether it is FM or PM. In fact, it is meaningless to enquire if a certain angle modulated wave is FM or PM. An analogous situation would be to ask a person (who is married, with children), whether he is a father or a son. The person would be puzzled because he is both, a father (of his child) and a son (of his father).



We have seen that PM and FM are not different kind of modulation, but two special cases of generalized angle modulation. Such a view is very fruitful because it shows the convertibility of one type of angle modulation (such as PM) to another (such as FM). This convertibility is quite clear in Fig. 4.42. For instance, we show later that the bandwidth of FM is approximately $2k_f m_p$, where $m_p$ is the peak amplitude of $m(t)$. We can derive the equivalent result for PM by referring to Fig. 4.42b, which shows that PM is actually the FM when the modulating signal is $\dot{m}(t)$. Clearly, the bandwidth of PM is approximately $2k_p m_p'$, where $m_p'$ is the peak amplitude of $\dot{m}(t)$. This argument shows that if we analyze one type of angle modulation (such as FM), we could readily extend those results to any other kind. Historically, the angle modulation concept began with FM. Hence, it is customary to analyze FM and then modify those results for other forms, such as PM. But this does not imply that FM is superior to other kinds of angle modulation. On the contrary, PM is superior to FM for most analog signals such as audio and video. Actually, the optimum performance is realized neither by PM nor FM, but by some other form, depending on the nature of the baseband (message) signal.

This discussion also shows that we need not discuss methods of generation and demodulation of each type of modulation. Figure 4.42 clearly indicates that the PM can be generated by an FM generator, and the FM can be generated by a PM generator. One of the methods of generating FM in practice (the Armstrong indirect-FM system) actually integrates $m(t)$ and uses it to phase-modulate a carrier. Similar remarks apply to demodulation of FM and PM.

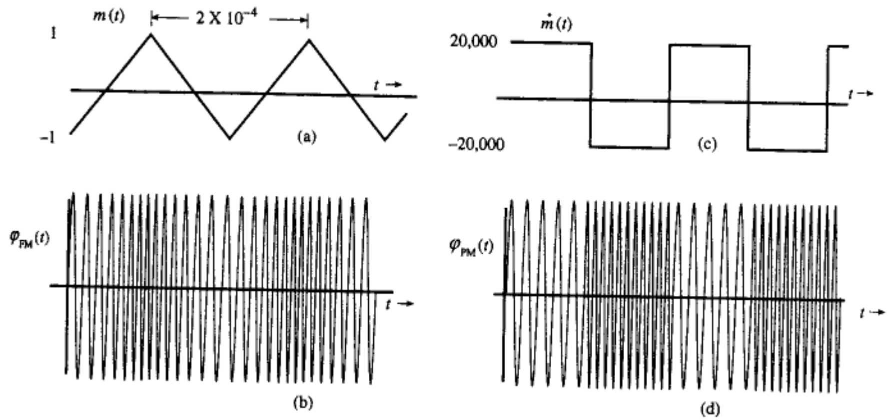



**Fig. 4.43** FM and PM waveforms.

■ **Example 4.21**

Sketch FM and PM waves for the modulating signal $m(t)$ depicted in Fig. 4.43a. The constants $k_f$ and $k_p$ are $2\pi(10^5)$ and $10\pi$, respectively, and the carrier frequency $\mathcal{F}_c$ is 100 MHz.

**For FM** [see Eq. (4.85a)]
$\omega_i = \omega_c + k_f m(t)$. Dividing throughout by $2\pi$, we obtain the equation in terms of the variable $\mathcal{F}$ (frequency in Hz). The instantaneous frequency $\mathcal{F}_i$ is

$$\mathcal{F}_i = \mathcal{F}_c + \frac{k_f}{2\pi} m(t) = 10^8 + 10^5 m(t)$$
$$(\mathcal{F}_i)_{\min} = 10^8 - 10^5\,|[m(t)]_{\min}| = 99.9 \text{ MHz}$$
$$(\mathcal{F}_i)_{\max} = 10^8 + 10^5 [m(t)]_{\max} = 100.1 \text{ MHz}$$

Because $m(t)$ increases and decreases linearly with time, the instantaneous frequency increases linearly from 99.9 to 100.1 MHz over a half-cycle and decreases linearly from 100.1 to 99.9 MHz over the remaining half-cycle of the modulating signal (Fig. 4.43b).

**For PM**
PM for $m(t)$ is FM for $\dot{m}(t)$. This assertion also follows from Eq. (4.84c) or Fig. 4.42b.

$$\mathcal{F}_i = \mathcal{F}_c + \frac{k_p}{2\pi}\dot{m}(t) = 10^8 + 5\,\dot{m}(t)$$
$$(\mathcal{F}_i)_{\min} = 10^8 - 5\,|[\dot{m}(t)]_{\min}| = 10^8 - 10^5 = 99.9 \text{ MHz}$$
$$(\mathcal{F}_i)_{\max} = 10^8 + 5\,[\dot{m}(t)]_{\max} = 100.1 \text{ MHz}$$

Because $\dot{m}(t)$ switches back and forth from a value of $-20{,}000$ to $20{,}000$, the carrier frequency switches back and forth from 99.9 to 100.1 MHz every half-cycle of $\dot{m}(t)$, as illustrated in Fig. 4.43d.

This indirect method of sketching PM (using $\dot{m}(t)$ to frequency-modulate a carrier) works as long as $m(t)$ is a continuous signal. If $m(t)$ is discontinuous, $\dot{m}(t)$ contains impulses, and this method is not so convenient. In such a case, a direct approach should be used. This is demonstrated in the next example. ■

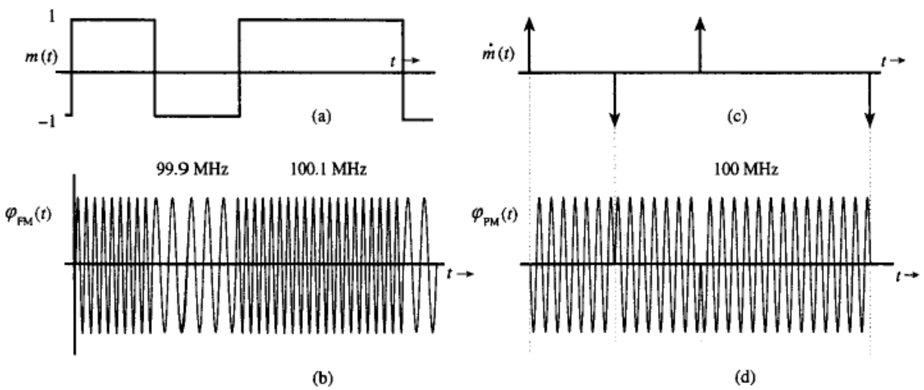



**Fig. 4.44** FM and PM waveforms.

■ **Example 4.22**
Sketch FM and PM waves for the digital modulating signal $m(t)$ depicted in Fig. 4.44a. The constants $k_f$ and $k_p$ are $2\pi(10^5)$ and $\pi/2$, respectively, and $\mathcal{F}_c = 100$ MHz.



**For FM**

$$\mathcal{F}_i = \mathcal{F}_c + \frac{k_f}{2\pi} m(t) = 10^8 + 10^5 m(t)$$

Because $m(t)$ switches back and forth from 1 to $-1$ and vice versa, the FM wave frequency switches back and forth from 99.9 MHz to 100.1 MHz and vice versa, as shown in Fig. 4.44b. This scheme of a carrier frequency modulation by a digital signal is known as **frequency-shift keying (FSK)**, because the information digits are transmitted by shifting the carrier frequency.

**For PM**

$$\mathcal{F}_i = \mathcal{F}_c + \frac{k_f}{2\pi}\dot{m}(t) = 10^8 + \frac{1}{4}\dot{m}(t)$$

The derivative $\dot{m}(t) = 0$ everywhere except for impulses of strength $\pm 2$ at the points of discontinuities of $m(t)$ (Fig. 4.44c). This fact means the carrier frequency is $\mathcal{F}_c = 100$ MHz everywhere except at the points of discontinuities, where it changes momentarily by infinite amount. It is not immediately apparent how an instantaneous frequency can be changed by an infinite amount and then changed back to the original frequency in zero time. Let us consider the direct approach.

$$\begin{aligned}\varphi_{\text{PM}}(t) &= A\cos\left[\omega_c t + k_p m(t)\right] \\ &= A\cos\left[\omega_c t + \frac{\pi}{2} m(t)\right] \\ &= \begin{cases} A\ \sin\omega_c t & \text{when } m(t) = -1 \\ -A\ \sin\omega_c t & \text{when } m(t) = 1 \end{cases}\end{aligned}$$

This PM wave, illustrated in Fig. 4.44c, has the same frequency $\mathcal{F}_c = 100$ MHz everywhere. However, there are phase discontinuities of $\pi$ radians at the instants where impulses of $\dot{m}(t)$ are located. At these instants, the carrier phase shifts by $\pi$ instantaneously. A finite phase shift in zero time implies infinite instantaneous frequency ($d\theta/dt = \infty$) at these instants. This conclusion agrees with our observation about $\dot{m}(t)$.

This scheme of carrier phase modulation by a digital signal is called **phase-shifting keying (PSK)**, because information digits are transmitted by shifting the carrier phase. Note that PSK may also be viewed as a DSB-SC modulation by $m(t)$.

The amount of phase discontinuity in $\varphi_{\text{PM}}(t)$ at the instant where $m(t)$ is discontinuous is $k_p m_d$, where $m_d$ is the amount of discontinuity in $m(t)$ at that instant. In the present example, the amplitude of $m(t)$ changes by 2 (from $-1$ to 1) at the discontinuity. Hence, the phase discontinuity in $\varphi_{\text{PM}}(t)$ is $k_p m_d = \frac{\pi}{2}(2) = \pi$ radians, which confirms our earlier result.

When $m(t)$ is a digital signal (as in Fig. 4.44a), $\varphi_{\text{PM}}(t)$ shows a phase discontinuity where $m(t)$ has a jump discontinuity. In such a case the phase deviation $k_p m(t)$ must be restricted to a range $(-\pi,\ \pi)$ in order to avoid ambiguity in demodulation. For example, if $k_p$ were $3\pi/2$ in the present example, then

$$\varphi_{\text{PM}}(t) = A\cos\left[\omega_c t + \frac{3\pi}{2} m(t)\right]$$

In this case $\varphi_{\text{PM}}(t) = A\sin\omega_c t$ when $m(t) = 1$ or $-1/3$. This will certainly cause ambiguity at the receiver when $A\sin\omega_c t$ is received. Such ambiguity never arises if $k_p m(t)$ is restricted to the range $(-\pi,\ \pi)$.

The ambiguity arises only when $m(t)$ has jump discontinuities. In such a case, the phase of $\varphi_{\text{PM}}(t)$ changes instantaneously. Because a phase $\varphi_o + 2n\pi$ is indistinguishable from the phase $\varphi_o$, ambiguities will be inherent in the demodulator unless the phase variations are limited to the range $(-\pi, \pi)$. For this reason $k_p$ should be small enough to restrict the phase change $k_p m(t)$ to the range $(-\pi, \pi)$.

No such restriction on $k_p$ is required if $m(t)$ is continuous. In this case the phase change is not instantaneous, but gradual over a time, and a phase $\varphi_o + 2n\pi$ will exhibit $n$ additional carrier cycles over the case of phase of only $\varphi_o$. This conclusion can also be verified from Example 4.21, where the maximum phase change $\Delta\varphi = 10\pi$.

Because a bandlimited signal cannot have jump discontinuities, we can say that when $m(t)$ is bandlimited, $k_p$ has no restrictions. ■

### 4.8-2 Bandwidth of Angle-Modulated Signals

Unlike amplitude modulation, there is no simple relationship between the baseband signal waveform and the corresponding angle modulated waveform. The same is true of their spectra. Because of nonlinear nature of angle modulation, derivation of $\Phi_{EM}(\omega)$, the frequency spectrum of the modulated signal is extremely complicated and can be obtained only for few special cases. Generally, the bandwidth of an angle modulated signal is infinite even when the baseband signal bandwidth is finite. However, most of the signal power (or energy) resides in a finite band. We shall now try to estimate this essential bandwidth of an angle modulated signal.

Let us start with the angle modulated signal in Eq. (4.80), and consider first the case of small $k$ $(k \to 0)$.

$$\begin{aligned}\varphi_{EM}(t) &= A\cos[\omega_c t + k\psi(t)] \\ &= A\cos\omega_c t\cos[k\psi(t)] - A\sin\omega_c t\sin[k\psi(t)] \\ &\approx A\cos\omega_c t - Ak\psi(t)\sin\omega_c t \qquad k \to 0\end{aligned} \tag{4.86}$$

Comparison of the right-hand side expression with $\varphi_{AM}(t)$ in Eq. (4.73a) shows that the two expressions are very similar. The first term is the carrier, and the second term, representing the sidebands, has the same form as the DSB-SC signal corresponding to the baseband signal $Ak\psi(t)$. The only difference is that the carrier is sine instead of cosine. This is just a matter of carrier phase difference of $\pi/2$. Hence, the bandwidth of the angle modulated signal is the same as that of AM signal corresponding to the baseband signal $\psi(t)$. If $m(t)$ is bandlimited to $B$ Hz, then the bandwidth of $\psi(t)$ is also $B$ Hz.† Hence, the bandwidth of $\varphi_{EM}(t)$ is $2B$ Hz, the same as that of AM. But this true only when $k \to 0$. Let us now consider the general case.

In angle modulation, the carrier frequency is varied from its quiescent value $\omega_c$. Let the maximum deviation of the carrier frequency be $\Delta\omega$. In other words, the carrier frequency varies in the range from $\omega_c - \Delta\omega$ to $\omega_c + \Delta\omega$. Because the carrier frequency always remains in this band of width $2\Delta\omega$ radians/s, could we say that the resulting spectrum also remains within this band and the bandwidth of the angle modulated signal is $2\Delta\omega$? This assertion implies that if a sinusoid takes an instantaneous frequency $\omega_x$, the resulting spectrum is concentrated only at $\omega_x$. This is true only if the carrier has infinite duration. For a finite duration sinusoid of

†Because $\psi(t)$ is the output of a linear system when the input is $m(t)$, the bandwidth of $\psi(t)$ cannot be greater than $B$ Hz. Moreover, the filter is invertible. Hence, the filter bandwidth cannot be less than $B$ Hz, or some of the components of $m(t)$ will be lost. So, the bandwidth of $\psi(t)$ cannot be less than $B$ Hz. Hence, it is exactly equal to $B$ Hz.



frequency $\omega_x$, the spectrum is not concentrated at $\omega_x$, but spreads out on both sides of $\omega_x$, as can be seen from Fig. 4.24d in Example 4.12. In a typical angle modulated signal, the carrier frequency is directly proportional to $m(t)$, which changes with $t$. Hence, the instantaneous frequency will also change with $t$ continuously. Such continuous shift in frequency will cause the spectral spread beyond the band $2\Delta\omega$. Clearly, the bandwidth of the angle modulated signal is somewhat larger than $2\Delta\omega$ rads/s. How much larger? This missing link can be found by looking at the results derived earlier for the case of $k \to 0$. Let us first determine $\Delta\omega$.

From Eq. (4.80), it follows that

$$\omega_i(t) = \omega_c + k\dot{\psi}(t) \tag{4.87}$$

if the peak amplitude of $\dot{\psi}(t)$ is denoted by $\psi'_p$, then the carrier frequency varies in the range from $\omega_c - k\psi'_p$ to $\omega_c + k\psi'_p$† Therefore

$$\Delta\omega = k\psi'_p \tag{4.88a}$$

The carrier frequency deviation $\Delta\mathcal{F}$ in Hz is

$$\Delta\mathcal{F} = \frac{\Delta\omega}{2\pi} = \frac{k}{2\pi}\psi'_p \tag{4.88b}$$

As demonstrated earlier, because of spectral spreading, the angle modulated signal bandwidth is somewhat larger than $2\Delta\mathcal{F}$. Let the actual bandwidth $B_{EM}$ in Hz be

$$\begin{aligned}B_{EM} &= 2\Delta\mathcal{F} + X \\ &= \frac{k}{\pi}\psi'_p + X\end{aligned} \tag{4.89}$$

where $X$ is unknown. To determine $X$, recall that for the case $k \to 0$, we found the bandwidth to be $2B$. But as Eq. (4.89) indicates, this bandwidth is $X$ when $k \to 0$. Therefore, $X = 2B$, and

$$B_{EM} = 2(\Delta\mathcal{F} + B)\text{ Hz} \tag{4.90}$$

A more rigorous derivation of this result appears in reference 4. Note that when $k \to 0$, $\Delta\mathcal{F} \to 0$ and $\Delta\mathcal{F} \ll B$. On the other hand when $k$ is very large, $\Delta\mathcal{F} \gg B$. The former case is known as the **narrowband** angle modulation and the latter is known as the **wideband** angle modulation.

Recall that for FM, $\dot{\psi}(t) = m(t)$, and $\psi'_p = m_p$, where $m_p$ is the peak amplitude of $m(t)$. Similarly, for PM, $\psi(t) = m(t)$. Hence, $\psi'_p = m'_p$, where $m'_p$ is the peak amplitude of $\dot{m}(t)$. Thus

$$(\Delta\mathcal{F})_{FM} = \frac{k_f}{\pi}m_p \qquad \text{and} \qquad (\Delta\mathcal{F})_{PM} = \frac{k_p}{\pi}m'_p \tag{4.91}$$

We observe an interesting fact in angle modulation. The bandwidth of the modulated signal is adjustable by choosing suitable value of $\Delta\mathcal{F}$ or the constant $k$ ($k_f$ in FM or $k_p$ in PM). Amplitude modulation lacks this feature. The bandwidth of each AM scheme is fixed. It is a general principle in communication theory

†This assertion implies an assumption $\dot{\psi}(t)|_{max} = \left|\dot{\psi}(t)|_{min}\right|$

that widening a signal bandwidth makes the signal more immune to noise during transmission. Thus, widening the transmission bandwidth makes angle modulated signals can be made more immune to noise. Moreover, this very property allows us to reduce the signal power required to achieve the same quality of transmission. Thus, angle modulation allows us to exchange signal power for bandwidth.

Also, because of its constant amplitude, angle modulation has a major advantage over amplitude modulation. This feature makes angle modulation less susceptible to nonlinear distortion. We shall see in the following section (Sec. 4.8-3) that no distortion results when we pass an angle modulated signal through a nonlinear device whose output $y(t)$ and the input $x(t)$ are related by $y(t) = x^2(t)$ [in general $y(t) = \sum a_n x^n(t)$]. Such a nonlinearity can be disastrous in amplitude modulated systems. This is the primary reason why angle modulation is used in microwave relay systems, where nonlinear operation of amplifiers and other devices has thus far been unavoidable at the required high power levels. In addition, the constant amplitude of FM gives it a kind of immunity against rapid fading. The effect of amplitude variations caused by rapid fading can be eliminated by using automatic gain control and bandpass limiting[4]. Angle modulation is also less vulnerable than amplitude modulation to small interference from adjacent channels. But the price for all these advantages is paid in terms of increased bandwidth. We can demonstrate that for the same bandwidth, the pulse code modulation (PCM), discussed in Chapter 5, is superior to angle modulation[4].

### 4.8-3 Generation and Demodulation of Angle Modulated Signals.

In Eq. (4.86), we see that a narrowband angle (or exponential) modulated signal (NBEM) consists of a carrier term and a DSB-SC term whose carrier has a $\pi/2$ phase shift. Hence, we can readily generate this signal using the procedure discussed in Sec. 4.7. Wideband modulation (WBEM) can be obtained from NBEM by passing the NBEM signal through a nonlinear device. Consider, for example, a nonlinear device whose input $x(t)$ and the output $y(t)$ are related by $y(t) = x^2(t)$. If the input is an angle modulated signal $\cos[\omega_c t + k\psi(t)]$, then the output $y(t)$ is given by

$$\begin{aligned} y(t) &= \cos^2[\omega_c t + k\psi(t)] \\ &= \frac{1}{2} + \frac{1}{2}\cos[2\omega_c t + 2k\psi(t)] \end{aligned}$$

If we pass this signal through a bandpass filter centered at $2\omega_c$, the output is

$$z(t) = \frac{1}{2}\cos[2\omega_c t + 2k\psi(t)]$$

Observe that the second order of nonlinearity has doubled the carrier frequency as well as the effective value of $k$ without causing any distortion. In a similar way, we can show that an $n$th-order of nonlinearity increases $n$-fold the carrier frequency as well as the effective value of $k$. This fact allows us the convert the NBEM into WBEM. This is the indirect method of generating angle modulated signal.

We can also generate angle modulated signal by a direct method, which uses a **voltage controlled oscillator (VCO)**. The output of a VCO is a constant amplitude sinusoid, whose instantaneous frequency is directly proportional to an input voltage $m(t)$. Clearly, a VCO is an FM generator. As demonstrated earlier, FM generator, with minor modification, can be used to generate any other form of angle modulation.



#### Demodulation

We shall discuss here demodulation of FM waves. As explained earlier, FM demodulator, with some minor modification, can be used for demodulation of any other form of angle modulation. Because the instantaneous frequency of FM wave is proportional to the baseband signal $m(t)$, an FM demodulator is a device whose output is proportional to frequency of the input signal. Thus, the gain $H(\omega)$ of an ideal FM demodulator is of the form $c_1\omega + c_2$. An ideal differentiator has this property. If the input to an ideal differentiator is an angle modulated signal $x(t) = \cos[\omega_c t + k\psi(t)]$, the output $y(t)$ is given by

$$\begin{aligned} y(t) = \frac{dx(t)}{dt} &= -[\omega_c + k\dot{\psi}(t)]\sin[\omega_c t + k\psi(t)] \\ &= [\omega_c + k\dot{\psi}(t)]\sin[\omega_c t + k\psi(t) + \pi] \end{aligned}$$

The output is also an angle modulated signal, whose envelope is $\omega_c + k\dot{\psi}(t)$. Hence, an ideal differentiator followed by an envelope detector will result in the output $\omega_c + k\dot{\psi}(t)$. After blocking the dc, we obtain the desired output $k\dot{\psi}(t)$. Recall that for FM, $\psi(t) = \int^t m(\alpha)d\alpha$. Hence, $\dot{\psi}(t) = m(t)$. Another device that can be used as an FM demodulator is a tuned circuit, whose resonant frequency is selected either above or below the carrier frequency of the FM signal to be demodulated. The frequency response of a tuned circuit (below the resonant frequency) is approximately linear with the input frequency (at least over a small band). This scheme suffers from the fact that the slope of $H(\omega)$ of a tuned circuit is linear only over a small band, and therefore causes considerable distortion in the output. This fault can partially be corrected by a balanced discriminator that uses two resonant circuits, one tuned above and the other tuned below $\omega_c$.

These days, a **phase-locked loop (PLL)**, whose performance is superior to any of the methods discussed here (especially in the large noise environment) has become very popular as a demodulator of angle modulated signals because of its reasonable cost. More discussion about modulation and demodulation of angle modulated signals appears in reference 4.

#### A Historical Note

In the twenties, broadcasting was in its infancy. However, there was a constant search for techniques that would reduce noise (static). Now, since the noise power is proportional to the modulated signal bandwidth (sidebands), attempts were focused on finding a modulation scheme to reduce the bandwidth. It was rumored that a new method had been discovered for eliminating sidebands (no sidebands, no bandwidth!). The concept of FM, where the carrier frequency would be varied in proportion to the message $m(t)$, appeared quite intriguing. The carrier frequency $\omega(t)$ would be varied with time so that $\omega(t) = \omega_c + km(t)$, where $k$ is an arbitrary constant. The carrier frequency will remain within the band from $\omega_c - km_p$ to

$\omega_c + km_p$. The spectrum, centered at $\omega_c$, would have a bandwidth $2km_p$, which is controlled by the arbitrary constant $k$. By using an arbitrarily small $k$, we could make the information bandwidth arbitrarily small. This was a passport to communication heaven. Unfortunately, the experimental results showed that something was seriously wrong somewhere. The FM bandwidth was found to be always greater than (at best equal to) the AM bandwidth. In some cases, its bandwidth was several times that of AM.

Careful analysis by Carson showed that the FM bandwidth could never be smaller than that of AM; at best equal to that of AM. Unfortunately, Carson did not recognize the compensating advantage of FM in its ability to suppress noise. Without any justification, he states, "Thus, FM introduces inherent distortion and has no compensating advantages whatsoever."[5] In his later paper he says: "In fact, as more and more schemes are analyzed and tested, and as the essential nature of the problem is clearly perceivable, we are unavoidably forced to the conclusion that static, like the poor, will always be with us."[6] This opinion of one of the ablest mathematicians of the day in the communication industry set back the development of FM. The noise-suppressing advantage of FM was later proved by Major Edwin H. Armstrong[7], a brilliant engineer whose contributions to the field of radio systems are comparable to those of Hertz and Marconi. It was largely the work of Armstrong that was responsible for rekindling the interest in FM. Lamentably, Armstrong, who became despondent over the lengthy, most acrimonious, and expensive court battles with some titans of the communication industry over his patent rights, committed suicide in 1954 by walking out of a window 13 stories above the street.

### 4.8-4 Frequency-Division Multiplexing

Signal multiplexing allows transmission of several signals on the same channel. In Chapter 5, we shall discuss time-division multiplexing (TDM), where several signals time-share the same channel, such as a cable or an optical fiber. In frequency-division multiplexing (FDM), the use of modulation, as illustrated in Fig. 4.45, makes several signals share the band of the same channel. Each signal is modulated by a different carrier frequency. The various carriers are adequately separated to avoid overlap (or interference) between the spectra of various modulated signals. These carriers are referred to as **subcarriers**. Each signal may use a different kind of modulation (for example, DSB-SC, AM, SSB-SC, VSB-SC, or even FM or PM). The modulated-signal spectra may be separated by a small guard band to avoid interference and facilitate signal separation at the receiver.

When all of the modulated spectra are added, we have a composite signal that may be considered as a baseband signal. Sometimes, this composite baseband signal may be used to further modulate a high-frequency (radio frequency, or RF) carrier for the purpose of transmission.

At the receiver, the incoming signal is first demodulated by the RF carrier to retrieve the composite baseband, which is then bandpass filtered to separate each modulated signal. Then each modulated signal is individually demodulated by an appropriate subcarrier to obtain all the basic baseband signals.

## 4.9 Data Truncation: Window Functions

We often need to truncate data in diverse situations from numerical computa-



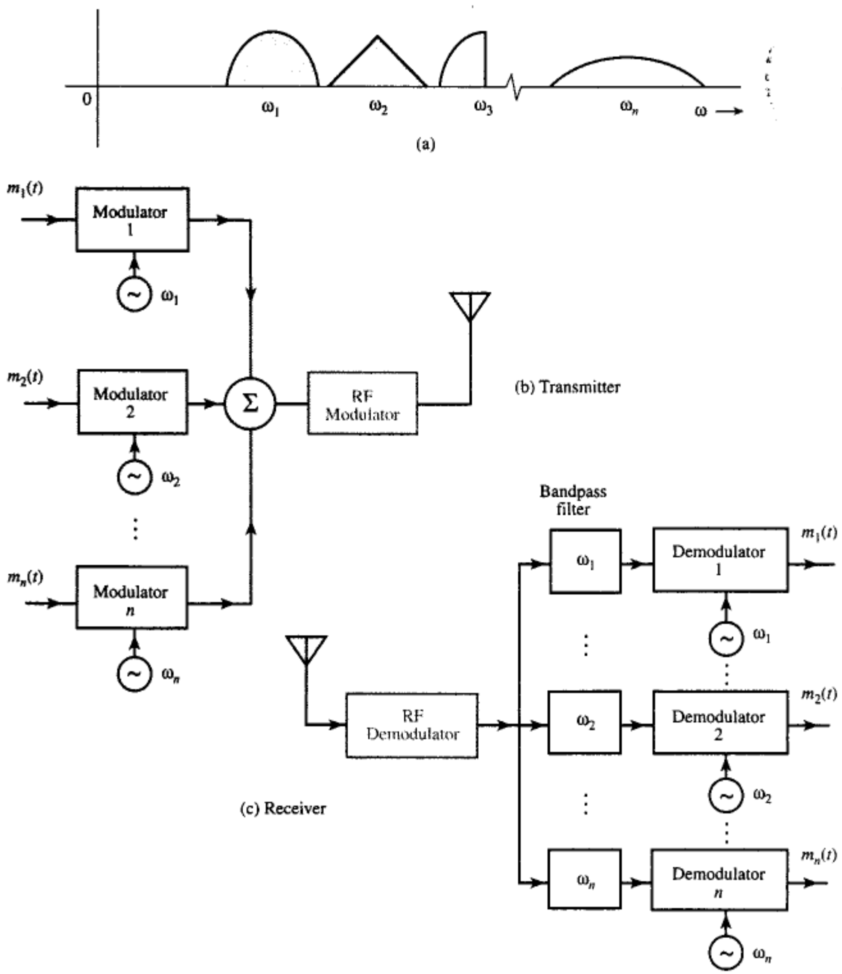



**Fig. 4.45** Frequency division multiplexing.

tions to filter design. For example, if we need to compute numerically the Fourier transform of some signal, say $e^{-t}u(t)$, on a computer, we will have to truncate the signal $e^{-t}u(t)$ beyond a sufficiently large value of $t$ (typically five time constants and above). The reason is that in numerical computations, we have to deal with data of finite duration. Similarly, the impulse response $h(t)$ of an ideal lowpass filter is noncausal, and approaches zero asymptotically as $|t| \to \infty$. For a practical design, we may want to truncate $h(t)$ beyond a sufficiently large value of $|t|$ to make $h(t)$ causal. In signal sampling, to eliminate aliasing, we need to truncate the signal spectrum beyond the half sampling frequency $\omega_s/2$, using an anti-aliasing filter.

Again, we may want to synthesize a periodic signal by adding the first $n$ harmonics and truncating all the higher harmonics. These examples show that data truncation can occur in both time and frequency domain. On the surface, truncation appears to be a simple problem of cutting off the data at a point where it is deemed to be sufficiently small. Unfortunately, this is not the case. Simple truncation can cause some unsuspected problems.

### Window Functions

Truncation operation may be regarded as multiplying a signal of a large width by a window function of a smaller (finite) width. Simple truncation amounts to using a **rectangular window** $w_R(t)$ (Fig. 4.48a) in which we assign unit weight to all the data within the window width ($|t| < \frac{T}{2}$), and assign zero weight to all the data lying outside the window ($|t| > \frac{T}{2}$). It is also possible to use a window in which the weight assigned to the data within the window may not be constant. In a **triangular window** $w_T(t)$, for example, the weight assigned to data decreases linearly over the window width (Fig. 4.48b).

Consider a signal $f(t)$ and a window function $w(t)$. If $f(t) \Longleftrightarrow F(\omega)$ and $w(t) \Longleftrightarrow W(\omega)$, and if the windowed function $f_w(t) \Longleftrightarrow F_w(\omega)$, then

$$f_w(t) = f(t)w(t) \qquad \text{and} \qquad F_w(\omega) = \frac{1}{2\pi}F(\omega) * W(\omega)$$

According to the width property of convolution, it follows that the width of $F_w(\omega)$ equals the sum of the widths of $F(\omega)$ and $W(\omega)$. Thus, truncation of a signal increases its bandwidth by the amount of bandwidth of $w(t)$. Clearly, the truncation of a signal causes its spectrum to spread (or smear) by the amount of the bandwidth of $w(t)$. Recall that the signal bandwidth is inversely proportional to the signal duration (width). Hence, the wider the window, the smaller is its bandwidth, and the smaller is the **spectral spreading**. This result is predictable because a wider window means we are accepting more data (closer approximation), which should cause smaller distortion (smaller spectral spreading). Smaller window width (poorer approximation) causes more spectral spreading (more distortion). There are also other effects produced by the fact that $W(\omega)$ is really not strictly bandlimited, and its spectrum $\rightarrow 0$ only asymptotically. This causes the spectrum of $F_w(\omega) \rightarrow 0$ asymptotically also at the same rate as that of $W(\omega)$, even though the $F(\omega)$ may be strictly bandlimited. Thus, windowing causes the spectrum of $F(\omega)$ to leak in the band where it is supposed to be zero. This effect is called **leakage**. These twin effects, the spectral spreading and the leakage, will now be clarified by an example.

For an example, let us take $f(t) = \cos \omega_0 t$ and a rectangular window $w_R(t) = \text{rect}(\frac{t}{T})$, illustrated in Fig. 4.46b. The reason for selecting a sinusoid for $f(t)$ is that its spectrum consists of spectral lines of zero width (Fig. 4.46a). This choice will make the effect of spectral spreading and leakage clearly visible. The spectrum of the truncated signal $f_w(t)$ is the convolution of the two impulses of $F(\omega)$ with the sinc spectrum of the window function. Because the convolution of any function with an impulse is the function itself (shifted at the location of the impulse), the resulting spectrum of the truncated signal is ($1/2\pi$ times) the two sinc pulses at $\pm\omega_0$, as depicted in Fig. 4.46c. Comparison of spectra $F(\omega)$ and $F_w(\omega)$ reveals the effects of truncation. These are:

1 The spectral lines of $F(\omega)$ have zero width. But the truncated signal is spread out by $4\pi/T$ about each spectral line. The amount of spread is equal to the



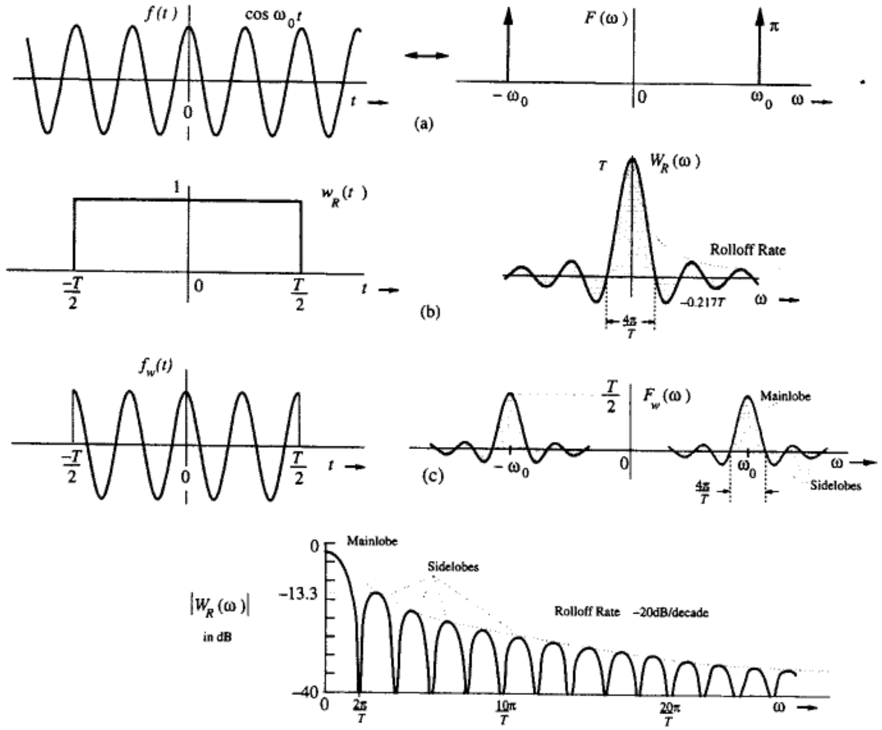



**Fig. 4.46** Windowing and its effects.

width of the mainlobe of the window spectrum. One effect of this **spectral spreading** (or smearing) is that if $f(t)$ has two spectral components of frequencies differing by less than $4\pi/T$ rad/s ($2/T$ Hz), they will be indistinguishable in the truncated signal. The result is loss of spectral resolution. We would like the spectral spreading (mainlobe width) to be as small as possible.

2 In addition to the mainlobe spreading, the truncated signal also has sidelobes, which decay slowly with frequency. The spectrum of $f(t)$ is zero everywhere except at $\pm\omega_0$. On the other hand, the truncated signal spectrum $F_w(\omega)$ is zero nowhere because of sidelobes. These sidelobes decay asymptotically as $1/\omega$. Thus, the truncation causes spectral **leakage** in the band where the spectrum of the signal $f(t)$ is zero. The peak **sidelobe** magnitude is 0.217 times the mainlobe magnitude (13.3 dB below the peak mainlobe magnitude). Also, the sidelobes decay at a rate $1/\omega$, which is $-6$ dB/octave (or $-20$ dB/decade). This is the **rolloff rate** of sidelobes. We want smaller sidelobes with a faster rate of decay (high rolloff rate). Figure 4.46d shows $|W_R(\omega)|$ (in dB) as a function of $\omega$. This plot clearly shows the mainlobe and sidelobe features, with the first sidelobe amplitude $-13.3$ dB below the mainlobe amplitude, and the sidelobes decaying at a rate of $-6$ dB/octave (or $-20$ dB per decade).

So far, we have discussed the effect of signal truncation (truncation in time domain) on the signal spectrum. Because of time-frequency duality, the effect of spectral truncation (truncation in frequency domain) on the signal shape is similar.

**Remedies for Side Effects of Truncation**

For better results, we must try to minimize the truncation's twin side effects, the spectral spreading (mainlobe width) and leakage (sidelobe). Let us consider each of these ills.

1 The spectral spread (mainlobe width) of the truncated signal is equal to the bandwidth of the window function $w(t)$. We know that the signal bandwidth is inversely proportional to the signal width (duration). Hence, to reduce the spectral spread (mainlobe width), we need to increase the window width.

2 To improve the leakage behavior, we must search for the cause of the slow decay of sidelobes. In Chapter 3, we saw that the Fourier spectrum decays as $1/\omega$ for a signal with jump discontinuity, and decays as $1/\omega^2$ for a continuous signal whose first derivative is discontinuous, and so on.† Smoothness of a signal is measured by the number of continuous derivatives it possesses. The smoother the signal, the faster the decay of its spectrum. Thus, we can achieve a given leakage behavior by selecting a suitably smooth window.

3 For a given window width, the remedies for the two effects are incompatible. If we try to improve one, the other deteriorates. For instance, among all the windows of a given width, the rectangular window has the smallest spectral spread (mainlobe width), but has high level sidelobes, which decay slowly. A tapered (smooth) window of the same width has smaller and faster decaying sidelobes, but it has a wider mainlobe.‡ But we can compensate for the increased mainlobe width by widening the window. Thus, we can remedy both the side effects of truncation by selecting a suitably smooth window of sufficient width.

There are several well-known tapered-window functions, such as Bartlett (triangular), Hanning (von Hann), Hamming, Blackman, and Kaiser, which truncate the data gradually. These windows offer different tradeoffs with respect to spectral spread (mainlobe width), the peak sidelobe magnitude, and the leakage rolloff rate as indicated in Table 4.3.[8,9] Observe that all windows are symmetrical about the origin (even functions of $t$). Because of this feature, $W(\omega)$ is a real function of $\omega$; that is, $\angle W(\omega)$ is either 0 or $\pi$. Hence, the phase function of the truncated signal has a minimal amount of distortion.

Figure 4.47 shows two well-known tapered-window functions, the von Hann (or Hanning) window $w_{\text{HAN}}(x)$ and the Hamming window $w_{\text{HAM}}(x)$. We have intentionally used the independent variable $x$ because windowing can be performed in time domain as well as in frequency domain; so $x$ could be $t$ or $\omega$, depending on the application.

†This result was demonstrated for periodic signals. However, it applies to aperiodic signals also. This is because we showed in Chapter 4 that if $f_{T_0}(t)$ is a periodic signal formed by periodic extension of an aperiodic signal $f(t)$, then the spectrum of $f_{T_0}(t)$ is ($1/T_0$ times) the samples of $F(\omega)$. Thus, what is true of the decay rate of the spectrum of $f_{T_0}(t)$ is also true of the rate of decay of $F(\omega)$.

‡A tapered window yields a higher mainlobe width because the effective width of a tapered window is smaller than that of the rectangular window (see Sec. 2.7-2 [Eq. (2.67)] for the definition of effective width). Therefore, from the reciprocity of the signal width and its bandwidth, it follows that the rectangular window mainlobe width is smaller than that of a tapered window.



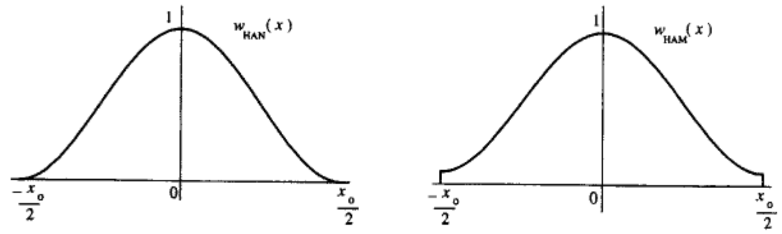



**Fig. 4.47** Hanning and Hamming windows.

There are hundreds of windows, each with differing characteristics. But the choice depends on a particular application. The rectangular window has the narrowest mainlobe. The Bartlett (triangle) window (also called the Fejer or Cesaro) is inferior in all respects to the Hanning window. For this reason it is rarely used in practice. Hanning is preferred over Hamming in spectral analysis because it has faster sidelobe decay. For filtering applications, on the other hand, the Hamming window is the choice because it has the smallest sidelobe magnitude for a given mainlobe width. The Hamming window is the most widely used, general purpose window. The Kaiser window, which uses $I_0(\alpha)$, the Bessel function of the order 0, is more versatile and adjustable. Selecting a proper value of $\alpha$ $(0 \leq \alpha \leq 10)$ allows the designer to tailor the window to suit a particular application. The parameter $\alpha$ controls the mainlobe and sidelobe trade-off. When $\alpha = 0$, the Kaiser window is the rectangular window. For $\alpha = 5.4414$, it is the Hamming window, and when $\alpha = 8.885$, it is the Blackman window. As $\alpha$ increases, the mainlobe width increases and the sidelobe level decreases.

**Table 4.3**

**Some Window Functions and Their Characteristics**

| | Window $w(t)$ | Mainlobe Width | Rolloff Rate dB/oct | Peak Sidelobe Level in dB |
|---|---|---|---|---|
| 1 | Rectangular: $\text{rect}(\frac{t}{T})$ | $\frac{4\pi}{T}$ | $-6$ | $-13.3$ |
| 2 | Bartlett: $\Delta(\frac{t}{2T})$ | $\frac{8\pi}{T}$ | $-12$ | $-26.5$ |
| 3 | Hanning: $0.5\left[1+\cos\left(\frac{2\pi t}{T}\right)\right]$ | $\frac{8\pi}{T}$ | $-18$ | $-31.5$ |
| 4 | Hamming: $0.54+0.46\cos\left(\frac{2\pi t}{T}\right)$ | $\frac{8\pi}{T}$ | $-6$ | $-42.7$ |
| 5 | Blackman: $0.42+0.5\cos\left(\frac{2\pi t}{T}\right)+0.08\cos\left(\frac{4\pi t}{T}\right)$ | $\frac{12\pi}{T}$ | $-18$ | $-58.1$ |
| 6 | Kaiser: $\frac{I_0\left[\alpha\sqrt{1-4\left(\frac{t}{T}\right)^2}\right]}{I_0(\alpha)}$ $\quad 1 \leq \alpha \leq 10$ | $\frac{11.2\pi}{T}$ | $-6$ | $-59.9$ $(\alpha = 8.168)$ |

### 4.9-1 Filter Design Using Windows

We shall design an ideal lowpass filter of bandwidth $W$ rad/s. For this filter, the impulse response $h(t) = \frac{W}{\pi}\text{sinc}\,(Wt)$ (Fig. 4.48c) is noncausal and, therefore, unrealizable. Truncation of $h(t)$ by a suitable window (Fig. 4.48a) makes it realizable, although the resulting filter is now an approximation to the desired ideal filter.† We shall use a rectangular window $w_R(t)$ and a triangular (Bartlett) window $w_T(t)$ to truncate $h(t)$, and then examine the resulting filters. The truncated impulse responses $h_R(t)$ and $h_T(t)$ for the two cases are depicted in Fig. (4.48d).

$$h_R(t) = h(t)w_R(t) \qquad \text{and} \qquad h_T(t) = h(t)w_T(t)$$

Hence, the windowed filter transfer function is the convolution of $H(\omega)$ with the Fourier transform of the window, as illustrated in Fig. 4.48e and f. We make the following observations.

1. The windowed filter spectra show **spectral spreading** at the edges, and instead of a sudden switch there is a gradual transition from the passband to the stopband of the filter. The transition band is smaller ($2\pi/T$ rad/s) for the rectangular case compared to the triangular case ($4\pi/T$ rad/s).
2. Although $H(\omega)$ is bandlimited, the windowed filters are not. But the stopband behavior of the triangular case is superior to that of the rectangular case. For the rectangular window, the leakage in the stopband decreases slowly (as $1/\omega$) compared to that of the triangular window (as $1/\omega^2$). Moreover, the rectangular case has a higher peak sidelobe amplitude compared to that of the triangular window.

## 4.10 Summary

In Chapter 3 we represented periodic signals as a sum of (everlasting) sinusoids or exponentials (Fourier series). In this chapter we extended this result to aperiodic signals, which are represented by the Fourier integral (instead of the Fourier series). An aperiodic signal $f(t)$ may be regarded as a periodic signal with period $T_0 \to \infty$, so that the Fourier integral is basically a Fourier series with a fundamental frequency approaching zero. Therefore, for aperiodic signals, the Fourier spectra are continuous. This continuity means that a signal is represented as a sum of sinusoids (or exponentials) of all frequencies over a continuous frequency interval. The Fourier transform $F(\omega)$, therefore, is the spectral density (per unit bandwidth in Hz).

An ever-present aspect of the Fourier transform is the duality between time and frequency, which also implies duality between the signal $f(t)$ and its transform $F(\omega)$. This duality arises because of near-symmetrical equations for direct and inverse Fourier transforms. The duality principle has far-reaching consequences and yields many valuable insights into signal analysis.

The scaling property of the Fourier transform leads to the conclusion that the signal bandwidth is inversely proportional to signal duration (signal width). Time

†In addition to truncation, we also need to delay the truncated function by $\frac{T}{2}$ in order to render it causal. However, the time delay only adds a linear phase to the spectrum without changing the amplitude spectrum. For this reason, we shall ignore the delay in order to simplify our discussion.



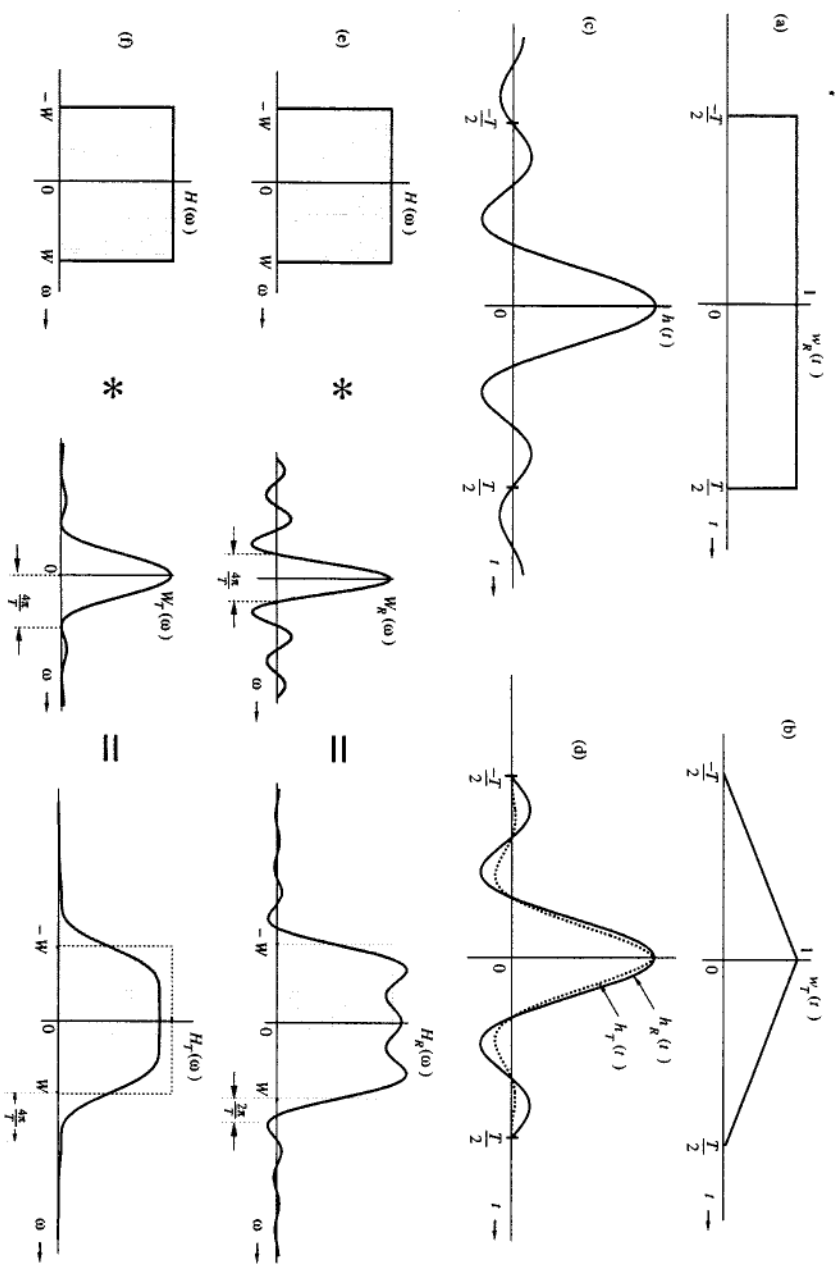



Fig. 4.48 Filter design using windows.

shifting of a signal does not change its amplitude spectrum, but adds a linear phase spectrum. Multiplication of a signal by an exponential $e^{j\omega_0 t}$ results in shifting the spectrum to the right by $\omega_0$. In practice, spectral shifting is achieved by multiplying a signal with a sinusoid such as $\cos \omega_0 t$ (rather than the exponential $e^{j\omega_0 t}$). This process is known as amplitude modulation. Multiplication of two signals results in convolution of their spectra, whereas convolution of two signals results in multiplication of their spectra.

For an LTIC system with the transfer function $H(\omega)$, the input and output spectra $F(\omega)$ and $Y(\omega)$ are related by the equation $Y(\omega) = F(\omega)H(\omega)$. This is valid only for asymptotically stable systems. For distortionless transmission of a signal through an LTIC system, the amplitude response $|H(\omega)|$ of the system must be constant, and the phase response $\angle H(\omega)$ should be a linear function of $\omega$ over a band of interest. Ideal filters, which allow distortionless transmission of a certain band of frequencies and suppress all the remaining frequencies, are physically unrealizable (noncausal). In fact, it is impossible to build a physical system with zero gain $[H(\omega) = 0]$ over a finite band of frequencies. Such systems (which include ideal filters) can be realized only with infinite time delay in the response.

The energy of a signal $f(t)$ is equal to $1/2\pi$ times the area under $|F(\omega)^2|$ (Parseval's theorem). The energy contributed by spectral components within a band $\Delta\mathcal{F}$ (in Hz) is given by $|F(\omega)|^2\Delta\mathcal{F}$. Therefore, $|F(\omega)|^2$ is the energy spectral density per unit bandwidth (in Hz). The energy spectral density $|F(\omega)|^2$ of a signal $f(t)$ is the Fourier transform of the autocorrelation function $\psi_f(t)$ of the signal $f(t)$. Thus, a signal autocorrelation function has a direct link to its spectral information.

The process of modulation shifts the signal spectrum to different frequencies. Modulation is used for many reasons: to transmit several messages simultaneously over the same channel to utilize channel's high bandwidth, to effectively radiate power over a radio link, to shift signal spectrum at higher frequencies to overcome the difficulties associated with signal processing at lower frequencies, to effect the exchange of transmission bandwidth and transmission power required to transmit data at a certain rate. Broadly speaking there are two types of modulation; amplitude and angle modulation. Each of these two classes has several subclasses. Amplitude modulation bandwidth is generally fixed. The bandwidth in angle modulation, however, is controllable. The higher the bandwidth, the more immune is the scheme to noise.

In practice, we often need to truncate data. Truncating data is like viewing it through a window, which permits a view of only certain portions of the data and hides (suppresses) the remainder. Abrupt truncation of data amounts to a rectangular window, which assigns a unit weight to data seen from the window and assigns zero weight to the remaining data. Tapered windows, on the other hand, reduce the weight gradually from 1 to 0. Data truncation can cause some unsuspected problems. For example, in computation of the Fourier transform, windowing (data truncation) causes spectral spreading (spectral smearing) that is characteristic of the window function used. A rectangular window results in the least spreading, but it does so at the cost of a high and oscillatory spectral leakage outside the signal band which decays slowly as $1/\omega$. Compared to a rectangular window, tapered windows in general have larger spectral spreading (smearing), but the spectral leakage is smaller and decays faster with frequency. If we try to reduce spectral leakage by using a smoother window, the spectral spreading increases. Fortunately, the



spectral spreading can be reduced by increasing the window width. Therefore, we can achieve a given combination of spectral spread (transition bandwidth) and leakage characteristics by choosing a suitable tapered window function of a sufficiently longer width $T$.

## References


1. Churchill, R.V., and J.W. Brown, *Fourier Series and Boundary Value Problems, 3d ed.*, McGraw-Hill, New York, 1978.
2. Bracewell, R.N., *Fourier Transform and Its Applications, revised 2nd Ed.*, McGraw-Hill, New York, 1986.
3. Guillemin, E.A., *Theory of Linear Physical Systems*, Wiley, New York, 1963.
4. Lathi, B.P., *Modern Digital and Analog Communication Systems, 3rd ed.*, Oxford University Press, New York, 1998.
5. J. Carson, "Notes on Theory of Modulation" *Proc. IRE*, vol 10, February 1922, pp. 57-64.
6. J. Carson, "The Reduction of Atmospheric Disturbances" *Proc. IRE*, vol 16, July 1928, pp. 966-975.
7. Armstrong E.H., "A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation", *Proc. IRE*, vol. 24, May 1936, pp. 689-740.
8. Hamming, R.W., *Digital Filters, 2nd Ed.*, Prentice-Hall, Englewood Cliffs, N.J. 1983.
9. Harris, F.J., "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", *Proc. IEEE*, vol. 66, No. 1, January 1978, pp. 51-83.


## Problems

**4.1-1** Show that if $f(t)$ is an even function of $t$, then

$$F(\omega) = 2\int_0^\infty f(t)\cos \omega t\, dt$$

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

$$F(\omega) = -2j\int_0^\infty f(t)\sin \omega t\, dt$$

Hence, prove that if $f(t)$ is a real and even function of $t$, then $F(\omega)$ is a real and even function of $\omega$. In addition, if $f(t)$ is a real and odd function of $t$, then $F(\omega)$ is an imaginary and odd function of $\omega$.

**4.1-2** Show that for a real $f(t)$, Eq. (4.8b) can be expressed as

$$f(t) = \frac{1}{\pi}\int_0^\infty |F(\omega)|\cos[\omega t + \angle F(\omega)]\, d\omega$$

This is the trigonometric form of the Fourier integral. Compare this with the compact trigonometric Fourier series.

**4.1-3** A signal $f(t)$ can be expressed as the sum of even and odd components (see Sec. 1.5-2):

$$f(t) = f_e(t) + f_o(t)$$

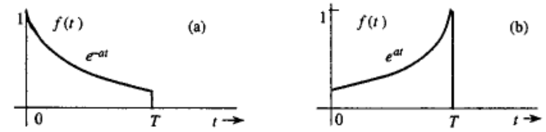



**Fig. P4.1-4**


**Fig. P4.1-5**

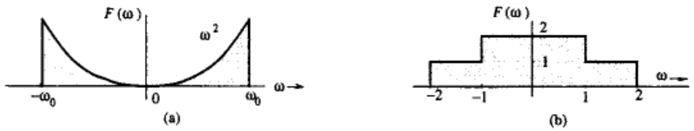



**Fig. P4.1-6**

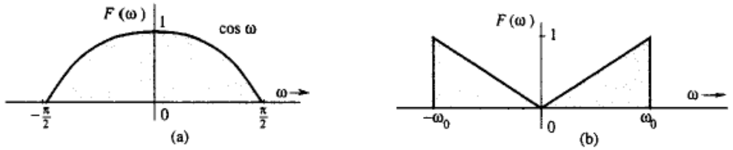



**Fig. P4.1-7**

**(a)** If $f(t) \Longleftrightarrow F(\omega)$, show that for real $f(t)$,

$$f_e(t) \Longleftrightarrow \text{Re}[F(\omega)] \qquad \text{and} \qquad f_o(t) \Longleftrightarrow j\ \text{Im}[F(\omega)]$$

**(b)** Verify these results by finding the Fourier transforms of the even and odd components of the following signals: **(i)** $u(t)$ **(ii)** $e^{-at}u(t)$.

**4.1-4** From definition (4.8a), find the Fourier transforms of the signals $f(t)$ in Fig. P4.1-4.

**4.1-5** From definition (4.8a), find the Fourier transforms of the signals depicted in Fig. P4.1-5.

**4.1-6** Using Eq. (4.8b), find the inverse Fourier transforms of the spectra in Fig. P4.1-6

**4.1-7** Using Eq. (4.8b), find the inverse Fourier transforms of the spectra in Fig. P4.1-7.

**4.2-1** Sketch the following functions:
**(a)** $\text{rect}\left(\frac{t}{2}\right)$ **(b)** $\Delta\left(\frac{3\omega}{100}\right)$ **(c)** $\text{rect}\left(\frac{t-10}{8}\right)$ **(d)** $\text{sinc}\left(\frac{\pi\omega}{5}\right)$ **(e)** $\text{sinc}\left(\frac{\omega-10\pi}{5}\right)$
**(f)** $\text{sinc}\left(\frac{t}{5}\right)\text{rect}\left(\frac{t}{10\pi}\right)$. Hint: $f\left(\frac{x-a}{b}\right)$ is $f\left(\frac{x}{b}\right)$ right-shifted by $a$.

**4.2-2** From definition (4.8a), show that the Fourier transform of $\text{rect}\,(t-5)$ is $\text{sinc}\left(\frac{\omega}{2}\right)e^{-j5\omega}$. Sketch the resulting amplitude and phase spectra.



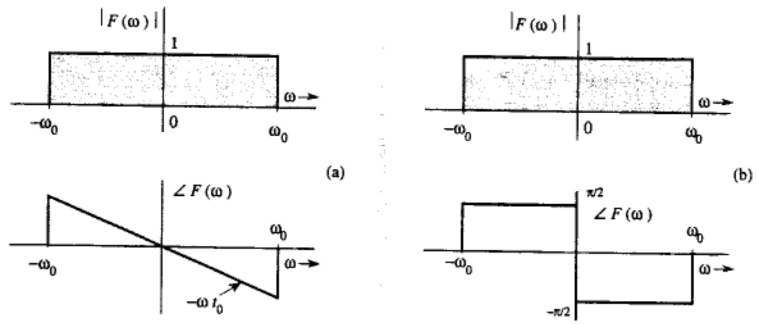



**Fig. P4.2-4**

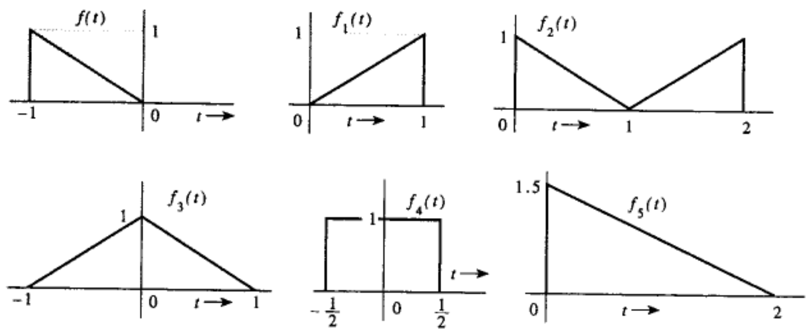



**Fig. P4.3-2**

**4.2-3** From definition (4.8b), show that the inverse Fourier transform of $\text{rect}\left(\frac{\omega-10}{2\pi}\right)$ is $\text{sinc}\,(\pi t)\,e^{j10t}$.

**4.2-4** Find the inverse Fourier transform of $F(\omega)$ for the spectra illustrated in Figs. P4.2-4a and b.
Hint: $F(\omega) = |F(\omega)|e^{j\angle F(\omega)}$. This problem illustrates how different phase spectra (both with the same amplitude spectrum) represent entirely different signals.

**4.3-1** Apply the symmetry property to the appropriate pair in Table 4.1 to show that
**(a)** $\frac{1}{2}[\delta(t) + \frac{j}{\pi t}] \Longleftrightarrow u(\omega)$ **(b)** $\delta(t+T) + \delta(t-T) \Longleftrightarrow 2\cos T\omega$
**(c)** $\delta(t+T) - \delta(t-T) \Longleftrightarrow 2j\sin T\omega$.

**4.3-2** The Fourier transform of the triangular pulse $f(t)$ in Fig. P4.3-2a is expressed as

$$F(\omega) = \frac{1}{\omega^2}(e^{j\omega} - j\omega e^{j\omega} - 1)$$

Using this information, and the time-shifting and time-scaling properties, find the Fourier transforms of the signals $f_i(t)$ $(i = 1, 2, 3, 4, 5)$ shown in Fig. P4.3-2.
Hint: See Sec. 1.3 for explanation of various signal operations. Pulses $f_i(t)$ $(i = 2, 3, 4$ can be expressed as a combination of $f(t)$ and $f_1(t)$ with suitable time shift (which may be positive or negative).

**4.3-3** Using only the time-shifting property and Table 4.1, find the Fourier transforms of the signals depicted in Fig. P4.3-3.

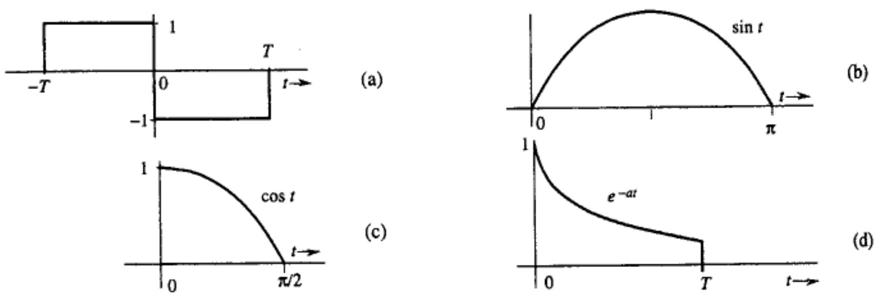



**Fig. P4.3-3**

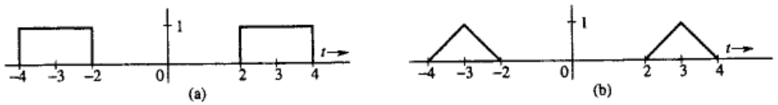



**Fig. P4.3-4**

**Fig. P4.3-5**

Hint: Signals in Figs. b, c, and d can be expressed in the form $f(t)[u(t) - u(t-a)]$.

**4.3-4** Using the time-shifting property, show that if $f(t) \Longleftrightarrow F(\omega)$, then

$$f(t+T) + f(t-T) \Longleftrightarrow 2F(\omega)\cos T\omega$$

This is the dual of Eq. (4.41). Using this result and pairs 17 and 19 in Table 4.1, find the Fourier transforms of the signals shown in Fig. P4.3-4.

**4.3-5** Prove the following results, which are duals of each other:

$$f(t)\sin\omega_0 t \Longleftrightarrow \tfrac{1}{2j}[F(\omega-\omega_0) - F(\omega+\omega_0)]$$
$$\tfrac{1}{2j}[f(t+T) - f(t-T)] \Longleftrightarrow F(\omega)\sin T\omega$$

Using the latter result and Table 4.1, find the Fourier transform of the signal in Fig. P4.3-5.

**4.3-6** The signals in Fig. P4.3-6 are modulated signals with carrier $\cos 10t$. Find the Fourier transforms of these signals using the appropriate properties of the Fourier transform and Table 4.1. Sketch the amplitude and phase spectra for parts **(a)** and **(b)**.

**4.3-7** Using the frequency-shifting property and Table 4.1, find the inverse Fourier transform of the spectra depicted in Fig. P4.3-7.

**4.3-8** Using the time convolution property, prove pairs 2, 4, 13 and 14 in Table 2.1 (assume $\lambda < 0$ in pair 2, $\lambda_1$ and $\lambda_2 < 0$ in pair 4, $\lambda_1 < 0$ and $\lambda_2 > 0$ in pair 13, and $\lambda_1$ and $\lambda_2 > 0$ in pair 14). Hint: You will need partial fraction expansion. For pair 2, you need to apply the result in Eq. (1.23).



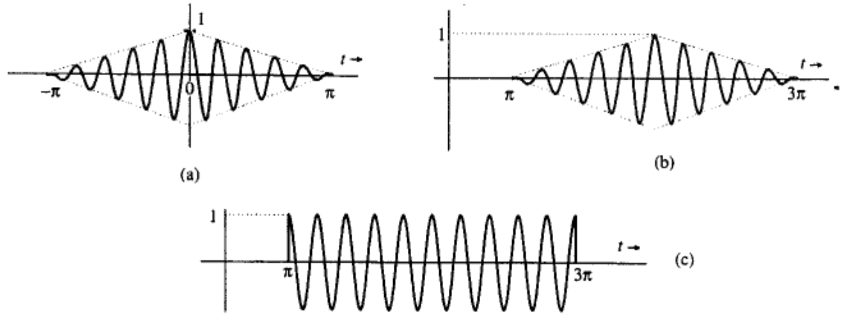



**Fig. P4.3-6**

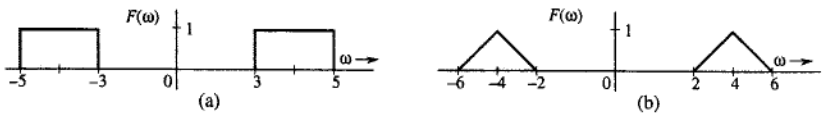



**Fig. P4.3-7**

**4.3-9** A signal $f(t)$ is bandlimited to $B$ Hz. Show that the signal $f^n(t)$ is bandlimited to $nB$ Hz. Hint: Start with $n = 2$. Use frequency convolution property and the width property of convolution.

**4.3-10** Find the Fourier transform of the signal in Fig. P4.3-3a by three different methods:
**(a)** By direct integration using the definition (4.8a).
**(b)** Using only pair 17 Table 4.1 and the time-shifting property.
**(c)** Using the time-differentiation and time-shifting properties, along with the fact that $\delta(t) \Longleftrightarrow 1$.
Hint: $1 - \cos 2x = 2\sin^2 x$.

**4.3-11** **(a)** Prove the frequency differentiation property (dual of the time differentiation):

$$-jtf(t) \Longleftrightarrow \frac{d}{d\omega}F(\omega)$$

**(b)** Using this property and pair 1 (Table 4.1), determine the Fourier transform of $te^{-at}u(t)$.

**4.4-1** For an LTIC system with transfer function

$$H(s) = \frac{1}{s+1}$$

find the (zero-state) response if the input $f(t)$ is **(a)** $e^{-2t}u(t)$ **(b)** $e^{-t}u(t)$ **(c)** $e^{t}u(-t)$ **(d)** $u(t)$
Hint: For part **(d)**, you need to apply the result in Eq. (1.23).

**4.4-2** A stable LTIC system is specified by the transfer function

$$H(\omega) = \frac{-1}{j\omega - 2}$$

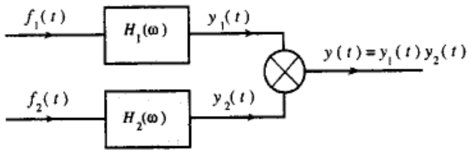



**Fig. P4.4-3**

Find the impulse response of this system and show that this is a noncausal system. Find the (zero-state) response of this system if the input $f(t)$ is
(a) $e^{-t}u(t)$ (b)$e^{t}u(-t)$.

**4.4-3** Signals $f_1(t) = 10^4\text{rect}\,(10^4 t)$ and $f_2(t) = \delta(t)$ are applied at the inputs of the ideal lowpass filters $H_1(\omega) = \text{rect}\,(\frac{\omega}{40,000\pi})$ and $H_2(\omega) = \text{rect}\,(\frac{\omega}{20,000\pi})$ (Fig. P4.4-3). 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 $F_1(\omega)$ and $F_2(\omega)$.
(b) Sketch $H_1(\omega)$ and $H_2(\omega)$.
(c) Sketch $Y_1(\omega)$ and $Y_2(\omega)$.
(d) Find the bandwidths of $y_1(t), y_2(t)$, and $y(t)$.
Hint for part (d): Use the convolution property and the width property of convolution to determine the bandwidth of $y_1(t)y_2(t)$.

**4.4-4** A lowpass system time constant is often defined as the width of its unit impulse response $h(t)$ (see Sec. 2.7-2). An input pulse $p(t)$ to this system acts like an impulse of strength equal to the area of $p(t)$ if the width of $p(t)$ is much smaller than the system time constant. Assume $p(t)$ to be a lowpass pulse, that is its spectrum is concentrated at low frequencies. Verify this behavior by considering a system whose unit impulse response is $h(t) = \text{rect}\,(\frac{t}{10^{-3}})$. The input pulse is a triangle pulse $p(t) = \Delta(\frac{t}{10^{-6}})$. The area under this pulse is $A = 0.5 \times 10^{-6}$. Show that the system response to this pulse is very nearly the system response to the input $A\delta(t)$.

**4.4-5** A lowpass system time constant is often defined as the width of its unit impulse response $h(t)$ (see Sec. 2.7-2). An input pulse $p(t)$ to this system passes practically without distortion, if the width of $p(t)$ is much greater than the system time constant. Assume $p(t)$ to be a lowpass pulse, that is its spectrum is concentrated at low frequencies. Verify this behavior by considering a system whose unit impulse response is $h(t) = \text{rect}\,(\frac{t}{10^{-3}})$. The input pulse is a triangle pulse $p(t) = \Delta(t)$. Show that the system output to this pulse is very nearly $k\,p(t)$, where $k$ is the system gain to a dc signal, that is, $k = H(0)$.

**4.4-6** A causal signal $h(t)$ has a Fourier transform $H(\omega)$. If $R(\omega)$ and $X(\omega)$ are the real and the imaginary parts of $H(\omega)$, that is , $H(\omega) = R(\omega) + jX(\omega)$, then show that

$$R(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{X(\omega)}{\omega - y} \qquad \text{and} \qquad X(\omega) = -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{R(\omega)}{\omega - y}$$

assuming that $h(t)$ has no impulse at the origin. This pair of integrals defines the **Hilbert transform**.
Hint: Let $h_e(t)$ and $h_o(t)$ be the even and odd components of $h(t)$. Use results in Prob. 4.1-3. See Fig. 1.24 for the relationship between $h_e(t)$ and $h_o(t)$. Recall that $\text{sgn}(t) \Longleftrightarrow 2/j\omega$. Use convolution property.
This problem states one of the important properties of causal systems: that the real and imaginary parts of the transfer function of a causal system are related. If one



specifies the real part, the imaginary part cannot be specified independently. The imaginary part is predetermined by the real part, and vice versa. This result also leads to the conclusion that the magnitude and angle of $H(\omega)$ are related provided all the poles and zeros of $H(\omega)$ lie in the LHP.

**4.5-1** Consider a filter with the transfer function

$$H(\omega) = e^{-(k\omega^2 + j\omega t_0)}$$

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 4.1.

**4.5-2** Show that a filter with transfer function

$$H(\omega) = \frac{2(10^5)}{\omega^2 + 10^{10}}e^{-j\omega t_0}$$

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.

**4.5-3** Determine if the filters with the following transfer functions are physically realizable. If they are not realizable, can they be realized exactly or approximately by allowing a finite time delay in the response?

$$H(\omega) = \text{(a) } 10^{-6}\,\text{sinc}\,(10^{-6}\omega) \;\; \text{(b) } 10^{-4}\,\Delta\left(\frac{\omega}{40,000\pi}\right) \;\; \text{(c) } 2\pi\,\delta(\omega)$$

**4.6-1** Show that the energy of a Gaussian pulse

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

is $\frac{1}{2\sigma\sqrt{\pi}}$. Verify this result by deriving the energy $E_f$ from $F(\omega)$ using Parseval's theorem.
Hint: See pair 22 in Table 4.1. Use the fact that

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}$$

**4.6-2** Show that

$$\int_{-\infty}^{\infty} \text{sinc}^2\,(kx)\,dx = \frac{\pi}{k}$$

Hint: Recognize that the integral is the energy of $f(t) = \text{sinc}\,(kt)$. Find this energy by using Parseval's theorem.

**4.6-3** A lowpass signal $f(t)$ is applied to a squaring device. The squarer output $f^2(t)$ is applied to a lowpass filter of bandwidth $\Delta\mathcal{F}$ Hz (Fig. P4.6-3). Show that if $\Delta\mathcal{F}$ is very small ($\Delta\mathcal{F} \to 0$), then the filter output is a dc signal $y(t) \approx 2E_f\Delta\mathcal{F}$.

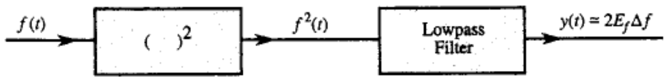



Fig. P4.6-3

Hint: If $f^2(t) \Longleftrightarrow A(\omega)$, then show that $Y(\omega) \approx [4\pi A(0)\Delta\mathcal{F}]\delta(\omega)$ if $\Delta\mathcal{F} \to 0$. Now, show that $A(0) = E_f$.

**4.6-4** Generalize Parseval's theorem to show that for real, Fourier transformable signals $f_1(t)$ and $f_2(t)$

$$\int_{-\infty}^{\infty} f_1(t)f_2(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_1(-\omega)F_2(\omega)\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_1(\omega)F_2(-\omega)\,d\omega$$

**4.6-5** For the signal

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

determine the essential bandwidth $B$ Hz of $f(t)$ such that the energy contained in the spectral components of $f(t)$ of frequencies below $B$ Hz is 99% of the signal energy $E_f$. Hint: See Exercise E4.5b.

**4.7-1** For each of the following 3 baseband signals (i) $m(t) = \cos 1000t$ (ii) $m(t) = 2\cos 1000t + \cos 2000t$ (iii) $m(t) = \cos 1000t \cos 3000t$

**(a)** Sketch the spectrum of $m(t)$.
**(b)** Sketch the spectrum of the DSB-SC signal $m(t)\cos 10{,}000t$.
**(c)** Identify the upper sideband (USB) and the lower sideband (LSB) spectra.
**(d)** Identify the frequencies in the baseband, and the corresponding frequencies in the DSB-SC, USB and LSB spectra. Explain the nature of frequency shifting in each case.

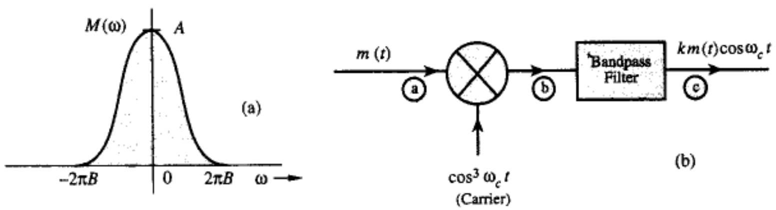



Fig. P4.7-2

**4.7-2** You are asked to design a DSB-SC modulator to generate a modulated signal $km(t)\cos\omega_c t$, where $m(t)$ is a signal bandlimited to $B$ Hz (Fig. P4.7-2a). Figure P4.7-2b shows a DSB-SC modulator available in the stock room. The bandpass filter is tuned to $\omega_c$. The carrier generator available generates not $\cos\omega_c t$, but $\cos^3\omega_c t$.
**(a)** Explain whether you would be able to generate the desired signal using only this equipment. If so, what is the value of $k$?
**(b)** Determine the signal spectra at points b and c, and indicate the frequency bands occupied by these spectra.
**(c)** What is the minimum usable value of $\omega_c$?
**(d)** Would this scheme work if the carrier generator output were $\cos^2\omega_c t$? Explain.
**(e)** Would this scheme work if the carrier generator output were $\cos^n\omega_c t$ for any integer $n \geq 2$?



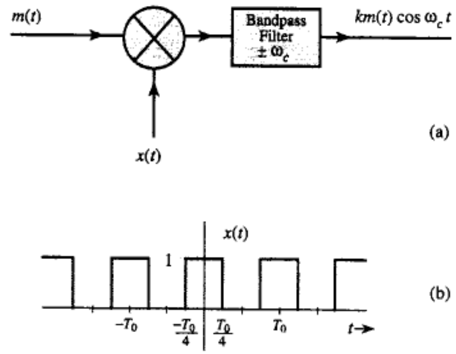



Fig. P4.7-3

**4.7-3** In practice, the analog multiplication operation is difficult and expensive. For this reason, in amplitude modulators, it is necessary to find some alternative to multiplication of $m(t)$ with $\cos\omega_c t$. Fortunately, for this purpose, we can replace multiplication with switching operation. A similar observation applies to demodulators. In the scheme depicted in Fig. P4.7-3a, the period of the rectangular periodic pulse $x(t)$ shown in Fig. P4.7-3b is $T_0 = 2\pi/\omega_c$. The bandpass filter is centered at $\pm\omega_c$. Note that multiplication by a square periodic pulse $x(t)$ in Fig. P4.7-3b amounts to periodic on-off switching of $m(t)$. This is a relatively simple and inexpensive operation.
Show that this scheme can generate amplitude modulated signal $k\cos\omega_c t$. Determine the value of $k$. Show that the same scheme can also be used for demodulation provided the bandpass filter in Fig. P4.7-3a is replaced by a lowpass (or baseband) filter.

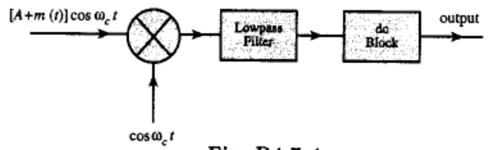



Fig. P4.7-4

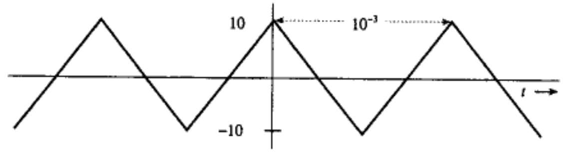



Fig. P4.7-5

**4.7-4** Figure P4.7-4 presents a scheme for coherent (synchronous) demodulation. Show that this scheme can demodulate the AM signal $[A + m(t)]\cos\omega_c t$ regardless of the value of $A$.

**4.7-5** Sketch the AM signal $[A+m(t)]\cos\omega_c t$ for the periodic triangle signal $m(t)$ illustrated in Fig. P4.7-5 corresponding to the modulation index: **(a)** $\mu = 0.5$, **(b)** $\mu = 1$, **(c)** $\mu = 2$, and **(d)** $\mu = \infty$. How do you interpret the case $\mu = \infty$?

**4.7-6** For each of the following three baseband signals **(a)** $m(t) = \cos 100t$ **(b)** $m(t) = \cos 100t + 2\cos 300t$ **(c)** $m(t) = \cos 100t \cos 500t$

(i) Sketch the spectrum of $m(t)$.
(ii) Find and sketch the spectrum of the DSB-SC signal $2m(t)\cos 1000t$.
(iii) From the spectrum obtained in (ii), suppress the LSB spectrum to obtain the USB spectrum.
(iv) Knowing the USB spectrum in (ii), write the expression $\varphi_{\text{USB}}(t)$ for the USB signal.
(v) Repeat (iii) and (iv) to obtain the LSB signal $\varphi_{\text{LSB}}(t)$.

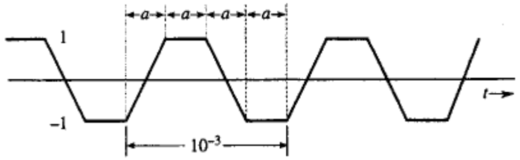


**Fig. P4.8-1**

**4.8-1** Sketch $\varphi_{\text{FM}}(t)$ and $\varphi_{\text{PM}}(t)$ for the modulating signal $m(t)$ depicted in Fig. P4.8-1, given $\omega_c = 2\pi \times 10^7$, $k_f = 2\pi \times 10^5$, and $k_p = 50\pi$.

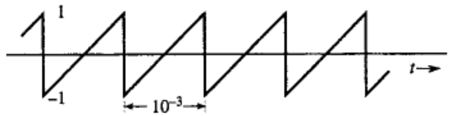


**Fig. P4.8-2**

**4.8-2** A baseband signal $m(t)$ is a periodic saw tooth signal shown in Fig. P4.8-2. Sketch $\varphi_{\text{FM}}(t)$ and $\varphi_{\text{PM}}(t)$ for this $m(t)$ if $\omega_c = 2\pi \times 10^6$, $k_f = 20{,}000\pi$, and $k_p = \pi/2$. Explain why it is necessary to use $k_p < \pi$ in this case.

**4.8-3** For a modulating signal

$$m(t) = 2\cos 100t + 18\cos 2000\pi t$$

Determine the bandwidths of the corresponding $\varphi_{\text{FM}}(t)$ and $\varphi_{\text{PM}}(t)$ if $k_f = 1000\pi$ and $k_p = 1$.

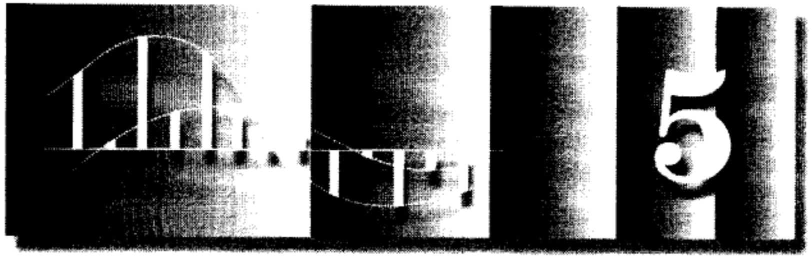
**4.8-4** An angle-modulated signal is described by the equation

$$\varphi_{\text{EM}}(t) = 10\cos(\omega_c t + 0.1\sin 2000\pi t)$$

(a) Find the frequency deviation $\Delta\mathcal{F}$ (b) Estimate the bandwidth of $\varphi_{\text{EM}}(t)$.

**4.8-5** Repeat Prob. 4.8-4 if

$$\varphi_{\text{EM}}(t) = 5\cos(\omega_c t + 20\sin 1000\pi t + 10\sin 2000\pi t)$$




# Sampling

A continuous-time signal can be processed by processing its samples through a discrete-time system. For this purpose, it is important to maintain the signal sampling rate sufficiently high so that the original signal can be reconstructed from these samples without error (or with an error within a given tolerance). The necessary quantitative framework for this purpose is provided by the sampling theorem derived in the following section.

## 5.1 The Sampling Theorem

We now show that a real signal whose spectrum is bandlimited to $B$ Hz $[F(\omega) = 0$ for $|\omega| > 2\pi B]$ can be reconstructed exactly (without any error) from its samples taken uniformly at a rate $\mathcal{F}_s > 2B$ samples per second. In other words, the minimum sampling frequency is $\mathcal{F}_s = 2B$ Hz.†

To prove the sampling theorem, consider a signal $f(t)$ (Fig. 5.1a) whose spectrum is bandlimited to $B$ Hz (Fig. 5.1b).‡ For convenience, spectra are shown as functions of $\omega$ as well as of $\mathcal{F}$ (Hz). Sampling $f(t)$ at a rate of $\mathcal{F}_s$ Hz ($\mathcal{F}_s$ samples per second) can be accomplished by multiplying $f(t)$ by an impulse train $\delta_T(t)$(Fig. 5.1c), consisting of unit impulses repeating periodically every $T$ seconds, where $T = 1/\mathcal{F}_s$. The result is the sampled signal $\overline{f}(t)$ resented in Fig. 5.1d. The sampled signal consists of impulses spaced every $T$ seconds (the sampling interval). The $n$th impulse, located at $t = nT$, has a strength $f(nT)$, the value of $f(t)$ at $t = nT$.

$$\overline{f}(t) = f(t)\delta_T(t) = \sum_n f(nT)\delta(t - nT) \tag{5.1}$$

†The theorem stated here (and proved subsequently) applies to lowpass signals. A bandpass signal whose spectrum exists over a frequency band $\mathcal{F}_c - \frac{B}{2} < |\mathcal{F}| < \mathcal{F}_c + \frac{B}{2}$ has a bandwidth of $B$ Hz. Such a signal is uniquely determined by $2B$ samples per second. In general, the sampling scheme is a bit more complex in this case. It uses two interlaced sampling trains, each at a rate of $B$ samples per second (known as second-order sampling). See, for example, the references.[1,2]

‡The spectrum $F(\omega)$ in Fig. 5.1b is shown as real, for convenience. However, our arguments are valid for complex $F(\omega)$ as well.