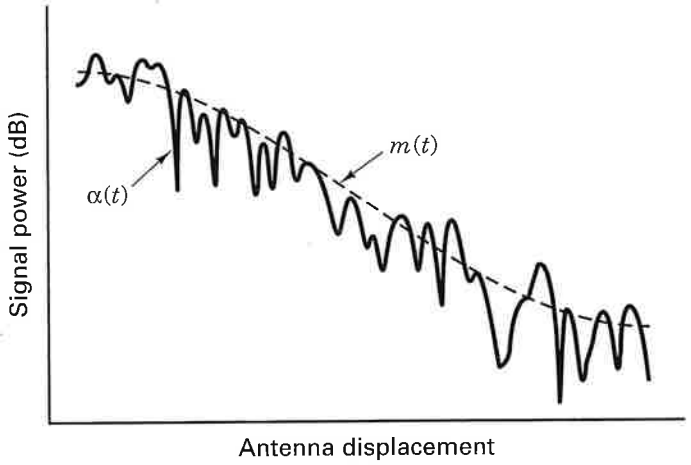
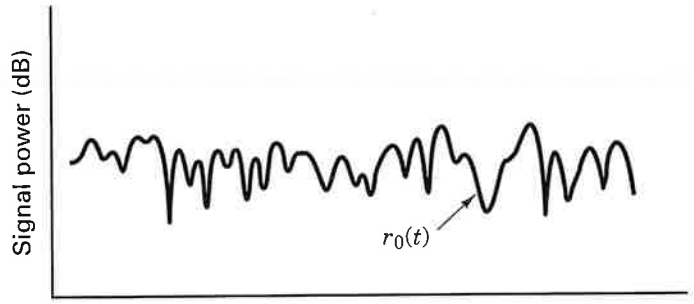


(a) Small-scale fading superimposed on large-scale fading



(b) Small-scale fading referred to an average constant power

Figure 15.3 Large-scale fading and small-scale fading.

tantamount to a function of time. In the selections that follow, some of the details regarding the statistics and mechanisms of large-scale and small-scale fading are enumerated.

15.2.1 Large-Scale Fading

For mobile radio applications, Okumura [4] made some of the earlier comprehensive path-loss measurements for a wide range of antenna heights and coverage distances. Hata [5] transformed Okumura's data into parametric formulas. In general, propagation models for both indoor and outdoor radio channels indicate that the mean path loss $\overline{L_p}(d)$, as a function of distance d between a transmitter and receiver

INTEL-1022, Part 9
9,584,262

is proportional to an nth-power of d relative to a reference distance d_0 [1], or, mathematically,

$$\overline{L_p}(d) \propto \left(\frac{d}{d_0}\right)^n \tag{15.5}$$

$\overline{L_p}(d)$ is often stated in decibels

$$\overline{L_p}(d)\,(\text{dB}) = L_s(d_0)\,(\text{dB}) + 10\,n \log\left(\frac{d}{d_0}\right) \tag{15.6}$$

The reference distance d_0 corresponds to a point located in the far field of the transmit antenna. Typically, the value of d_0 is taken to be 1 km for large cells, 100 m for microcells, and 1 m for indoor channels. Moreover, $L_s(d_0)$ is evaluated using Equation (15.1) or by conducting measurements. $\overline{L_p}(d)$ is the average path loss (over a multitude of different sites) for a given value of d. When plotted on a log-log scale, $\overline{L_p}(d)$ versus d (for distances greater than d_0) yields a straight line with a slope equal to 10 n dB/decade. The value of the path-loss exponent n depends on the frequency, antenna heights, and propagation environment. In free space where signal propagation follows an inverse-square law (as described in Section 5.3.1), n is equal to 2, as seen in Equation (15.1). In the presence of a very strong guided wave phenomenon (like urban streets), n can be lower than 2. When obstructions are present, n is larger. Figure 15.4 shows a scatter plot of path loss versus distance for measurements made at several sites in Germany [6]. Here, the path loss has been measured relative to a reference distance $d_0 = 100$ m. Also shown are straight-line fits to various exponent values.

The path loss versus distance expressed in Equation (15.6) is an average, and therefore not adequate to describe any particular setting or signal path. It is necessary to provide for variations about the mean because the environment of different sites may be quite different for similar transmitter-receiver (T-R) separations. Figure 15.4 illustrates that path-loss variations can be quite large. Measurements have shown that for any value of d, the path loss L_p is a random variable having a log-normal distribution about the mean distant-dependent value $\overline{L_p}(d)$ [7]. Thus, path loss $L_p(d)$ can be expressed in terms of $\overline{L_p}(d)$, in Equation (15.6), plus a random variable X_σ, as [1]

$$L_p(d)\,(\text{dB}) = L_s(d_0)\,(\text{dB}) + 10\,n \log_{10}(d/d_0) + X_\sigma\,(\text{dB}) \tag{15.7}$$

where X_σ denotes a zero-mean, Gaussian random variable (in decibels) with standard deviation σ (also in decibels). X_σ is site and distance dependent. Since X_σ and $L_p(d)$ are random variables, if Equation (15.7) is used as the basis for computing an estimate of path loss or link margin, some value for X_σ must first be chosen. The choice of the value is often based on measurements (made over a wide range of locations and T-R separations). It is not unusual for X_σ to take on values as high as 6 to 10 dB or greater. Thus, the parameters needed to statistically describe path loss due to large-scale fading, for an arbitrary location with a specific transmitter-receiver separation are (1) the reference distance, (2) the path-loss exponent, and (3) the standard deviation of X_σ. (There are several good references dealing with

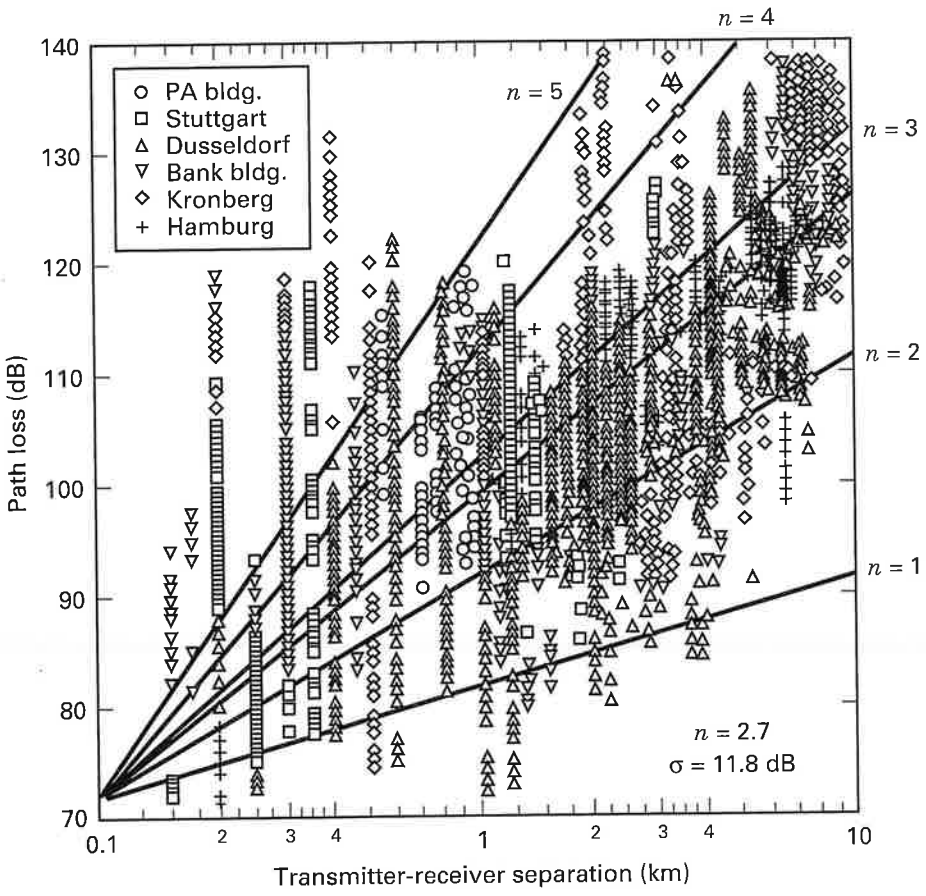


Figure 15.4 Path loss versus distance measured in several German cities.
Source: Seidel, S.Y., et. al., "Path Loss, Scattering and Multipath Delay Statistics in Four European Cities for Digital Cellular and Microcellular Radiotelephone," *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 4, pp. 721–730, November 1991.

the measurement and estimation of propagation path loss for many different applications and configurations [1, 5–9].)

15.2.2 Small-Scale Fading

Here we develop the small-scale fading component, $r_0(t)$. Analysis proceeds on the assumption that the antenna remains within a limited trajectory, so that the effect of large-scale fading $m(t)$ is a constant (assumed unity). Assume that the antenna is traveling, and that there are multiple scatterer paths, each associated with a time-variant propagation delay $\tau_n(t)$, and a time-variant multiplicative factor $\alpha_n(t)$. Neglecting noise, the received bandpass signal, can be written as

$$r(t) = \sum_n \alpha_n(t)\, s[t - \tau_n(t)] \tag{15.8}$$

Substituting Equation (15.2) into Equation (15.8), we write the received bandpass signal as follows:

$$\begin{aligned} r(t) &= \mathrm{Re}\left(\left\{ \sum_n \alpha_n(t) g[t - \tau_n(t)] \right\} e^{j 2\pi f_c [t - \tau_n(t)]} \right) \\ &= \mathrm{Re}\left(\left\{ \sum_n \alpha_n(t) e^{-j 2\pi f_c \tau_n(t)} g[t - \tau_n(t)] \right\} e^{j 2\pi f_c t} \right) \end{aligned} \tag{15.9}$$

From Equation (15.9), it follows that the equivalent received baseband signal is

$$z(t) = \sum_n \alpha_n(t) e^{-j 2\pi f_c \tau_n(t)} g[t - \tau_n(t)] \tag{15.10}$$

Consider the transmission of an *unmodulated* carrier at frequency f_c. In other words, for all time, $g(t) = 1$. Then, the received baseband signal, for this case of an unmodulated carrier and discrete multipath components given by Equation (15.10), reduces to

$$z(t) = \sum_n \alpha_n(t) e^{-j 2\pi f_c \tau_n(t)} = \sum_n \alpha_n(t)\, e^{-j \theta_n(t)} \tag{15.11}$$

where $\theta_n(t) = 2\pi f_c \tau_n(t)$. The baseband signal $z(t)$ consists of a sum of time-variant phasors having amplitudes $\alpha_n(t)$ and phases $\theta_n(t)$. Notice that $\theta_n(t)$ will change by 2π radians whenever τ_n changes by $1/f_c$ (typically, a very small delay). For a cellular radio operating at $f_c = 900$ Mhz, the delay $1/f_c = 1.1$ nanoseconds. In free space, this corresponds to a change in propagation distance of 33 cm. Thus, in Equation (15.11), $\theta_n(t)$ can change significantly with relatively small propagation delay changes. In this case, when two multipath components of a signal differ in path length by 16.5 cm, one signal will arrive 180 degrees out of phase with respect to the other signal. Sometimes the phasors add constructively and sometimes they add destructively, resulting in amplitude variations or fading of $z(t)$. Equation (15.11) can be expressed more compactly as the net received envelope, which is the summation over all the scatterers, as

$$z(t) = \alpha(t) e^{-j \theta(t)} \tag{15.12}$$

where $\alpha(t)$ is the resultant amplitude and $\theta(t)$ is the resultant phase. The right side of Equation (15.12) represents the same complex multiplicative factor that was described earlier in Section 15.2. Equation (15.12) is an important result because it tells us that even though a *bandpass* signal $s(t)$, as expressed in Equation (15.2), is the signal that experienced the fading effects and gave rise to the received signal $r(t)$, these effects can be described by analyzing $r(t)$ at the *baseband* level.

Figure 15.5 illustrates the primary mechanism that causes fading in multipath channels, as described by Equations (15.11) and (15.12). In the figure, a reflected signal has a phase delay (a function of additional path length) with respect to a

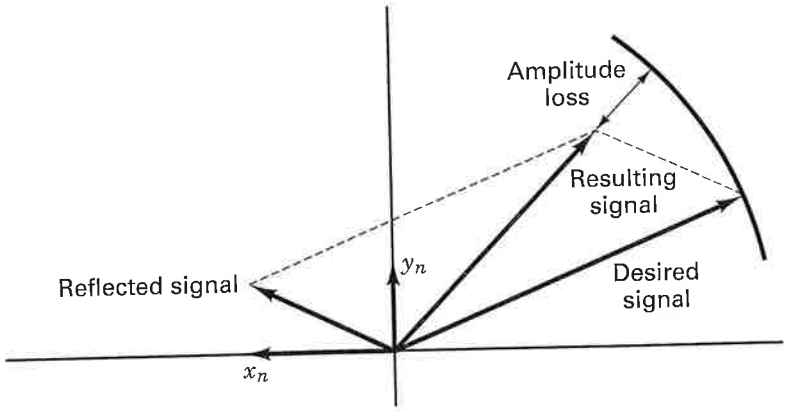


Figure 15.5 Effect of a multipath reflected signal on a desired signal.

desired signal. The reflected signal also has reduced amplitude (a function of the reflection coefficient of the obstruction). Reflected signals can be described in terms of orthogonal components $x_n(t)$ and $y_n(t)$, where $x_n(t) + jy_n(t) = \alpha_n(t)e^{-j\theta_n(t)}$. If the number of such stochastic components is large and none are dominant, then *at a fixed time,* the variables $x_r(t)$ and $y_r(t)$, resulting from their addition, will have a Gaussian pdf. These orthogonal components yield the small-scale (fading) magnitude $r_0(t)$, that was defined in Equation (15.4). For the case of an unmodulated carrier wave, as shown in Equation (15.12), $r_0(t)$ is the magnitude of $z(t)$, as follows:

$$r_0(t) = \sqrt{x_r^2(t) + y_r^2(t)} \tag{15.13}$$

When the received signal is made up of multiple reflective rays plus a significant line-of-sight (nonfaded) component, the received envelope amplitude has a Rician pdf as shown below, and the fading is referred to as *Rician Fading* [2]:

$$p(r_0) = \begin{cases} \dfrac{r_0}{\sigma^2} \exp\left[-\dfrac{(r_0^2 + A^2)}{2\sigma^2} \right] I_0\left(\dfrac{r_0 A}{\sigma^2}\right) & \text{for } r_0 \geq 0, A \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{15.14}$$

Although $r_0(t)$ varies dynamically with motion (time), at any *fixed time* it is a random variable, whose value stems from the ensemble of real positive numbers. Hence, in describing probability density functions, it is appropriate to drop the functional dependence on time. The parameter σ^2 is the predetection mean power of the multipath signal, A denotes the peak magnitude of the *non-faded* signal component (called the *specular component*) and $I_0(\cdot)$ is the modified Bessel function of the first kind and zero order [11]. The Rician distribution is often described in terms of a parameter K, which is defined as the ratio of the power in the specular component to the power in the multipath signal. It is given by $K = A^2/(2\sigma^2)$. As the

magnitude of the specular component approaches zero, the Rician pdf approaches a Rayleigh pdf, expressed as

$$p(r_0) = \begin{cases} \frac{r_0}{\sigma^2} \exp\left[-\frac{r_0^2}{2\sigma^2} \right] & \text{for } r_0 \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{15.15}$$

The Rayleigh faded component is sometimes called the *random* or *scatter* or *diffuse* component. The Rayleigh pdf results from having no specular signal component; thus, for a single link (no diversity), it represents the pdf associated with the worst case of fading per mean received signal power. For the remainder of this chapter, unless stated otherwise, it will be assumed that loss of signal-to-noise ratio (SNR) due to fading follows the Rayleigh model. It will also be assumed that the propagating signal is in the UHF band, encompassing cellular and personal communications services (PCS), with nominal frequency allocations of 1 GHz and 2 GHz, respectively.

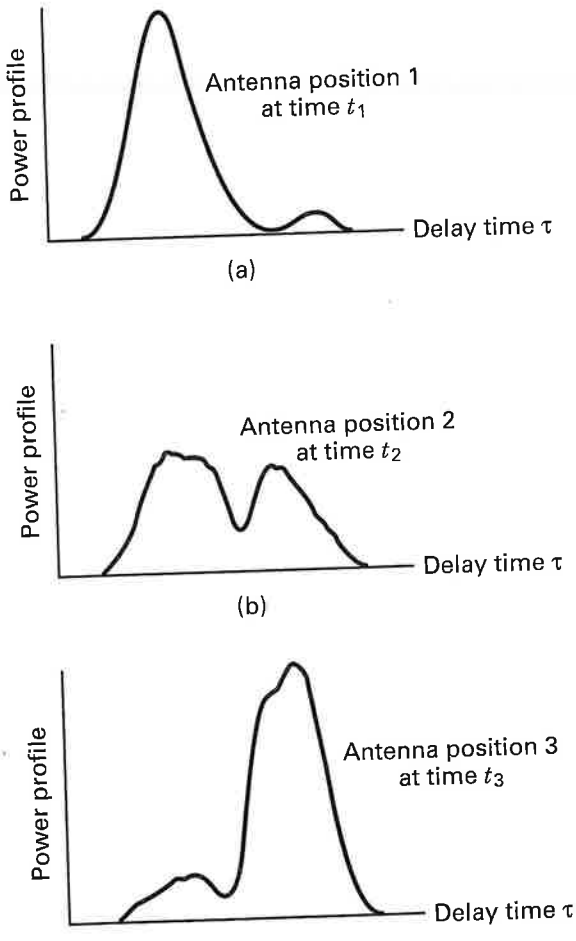


Figure 15.6 Response of a multipath channel to a narrow pulse versus delay, as a function of antenna position.

As indicated in Figure 15.1, blocks 4, 5, and 6, small-scale fading manifests itself in two mechanisms: (1) time-spreading of the underlying digital pulses within the signal, and (2) a time-variant behavior of the channel due to motion (e.g., a receive antenna on a moving platform). Figure 15.6 illustrates the consequences of both manifestations by showing the response of a multipath channel to a narrow pulse versus delay, as a function of antenna position (or time, assuming a mobile travelling at a constant velocity). In Figure 15.6, it is important to distinguish between two different time references—delay time τ and transmission or observation time t. Delay time refers to the time-spreading effect resulting from the fading channel's nonoptimum impulse response. The transmission time, however, is related to the antenna's motion or spatial changes, accounting for propagation path changes that are perceived as the channel's time-variant behavior. Note that for constant velocity, as is assumed in Figure 15.6, either antenna position or transmission time can be used to illustrate this time-variant behavior. Figures 15.6a–15.6c show the sequence of received pulse-power profiles as the antenna moves through a succession of equally spaced positions. Here, the interval between antenna positions is 0.4λ [12], where λ is the wavelength of the carrier frequency. For each of these three cases shown, the response pattern differs significantly in the delay time of the largest signal component, the number of signal copies, their magnitudes, and the total received power (area in each received power profile). Figure 15.7

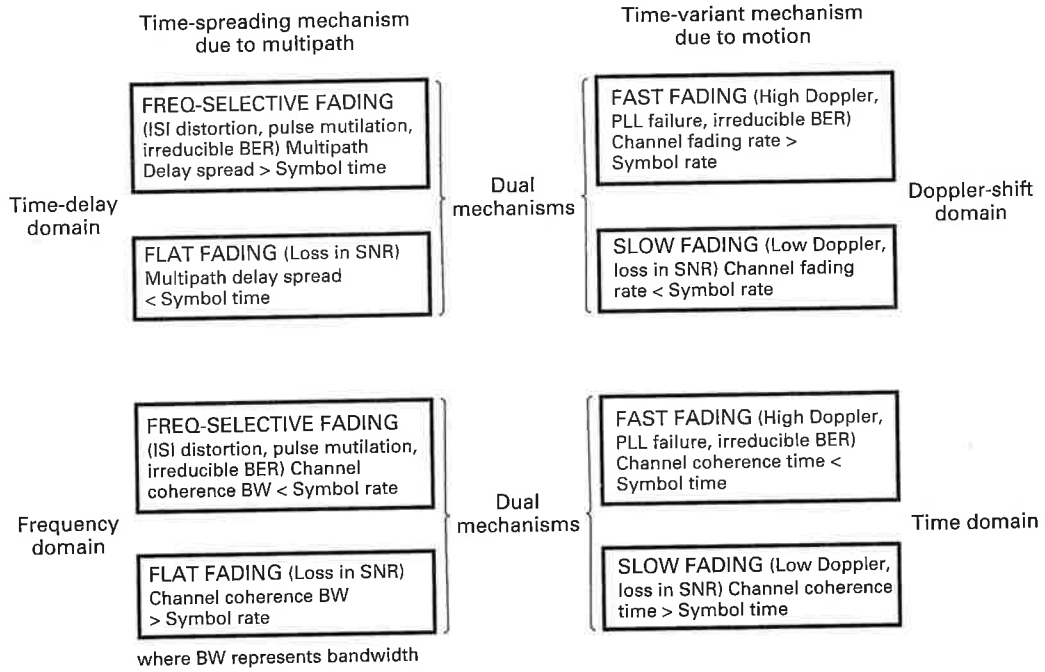


Figure 15.7 Small-scale fading: Mechanisms, degradation categories, and effects.

summarizes these two small-scale fading mechanisms, the two domains (time or time-delay and frequency or Doppler shift) for viewing each mechanism, and the degradation categories each mechanism can exhibit. Note that any mechanism characterized in the time domain can be characterized equally well in the frequency domain. Hence, as outlined in Figure 15.7, the time-spreading mechanism will be characterized in the time-delay domain as a multipath delay spread, and in the frequency domain it will be characterized as a channel coherence bandwidth. Similarly, the time-variant mechanism will be characterized in the time domain as a channel coherence time and in the Doppler-shift (frequency) domain as a channel fading rate or Doppler spread. These mechanisms and their associated degradation categories are examined in the sections that follow.

15.3 SIGNAL TIME-SPREADING

15.3.1 Signal Time-Spreading Viewed in the Time-Delay Domain

A simple way to model the fading phenomenon was introduced by Bello [13] in 1963; he proposed the notion of wide-sense stationary uncorrelated scattering (WSSUS). The model treats signals arriving at a receive antenna with different delays as uncorrelated. It can be shown [2, 13] that such a channel is effectively WSS in both the time and frequency domains. With such a model of a fading channel, Bello was able to define functions that apply for all time and all frequencies. For the mobile channel, Figure 15.8 contains four functions that make up this model [2, 10, 13–15]. These functions are examined here and in later sections, starting with Figure 15.8a and proceeding counter-clockwise toward Figure 15.8d.

In Figure 15.8a, a *multipath-intensity profile* is plotted ($S(\tau)$ versus time delay τ). Knowledge of $S(\tau)$ helps answer the question, "For a transmitted impulse, how does the average received power vary as a function of time delay τ?" The term "time delay" is used to refer to the excess delay. It represents the signal's propagation delay that exceeds the delay of the first signal arrival at the receiver. For a typical wireless channel, the received signal usually consists of several discrete multipath components, causing $S(\tau)$ to exhibit multiple isolated peaks, sometimes referred to as *fingers* or *returns*. For some channels, such as the tropospheric scatter channel, received signals are often seen as a continuum of multipath components [10, 15]. In such cases, $S(\tau)$ is a relatively smooth (continuous) function of τ. For making measurements of the multipath intensity profile, wideband signals (impulses or spread spectrum) need to be used [15]. For a single transmitted impulse, the time T_m between the first and last received component represents the *maximum excess delay*, after which the multipath signal power falls below some threshold level relative to the strongest component. The threshold level might be chosen at 10 dB or 20 dB below the level of the strongest component. Note that for an ideal system (zero excess delay) the function $S(\tau)$ would consist of an ideal impulse with weight equal to the total average received signal power.

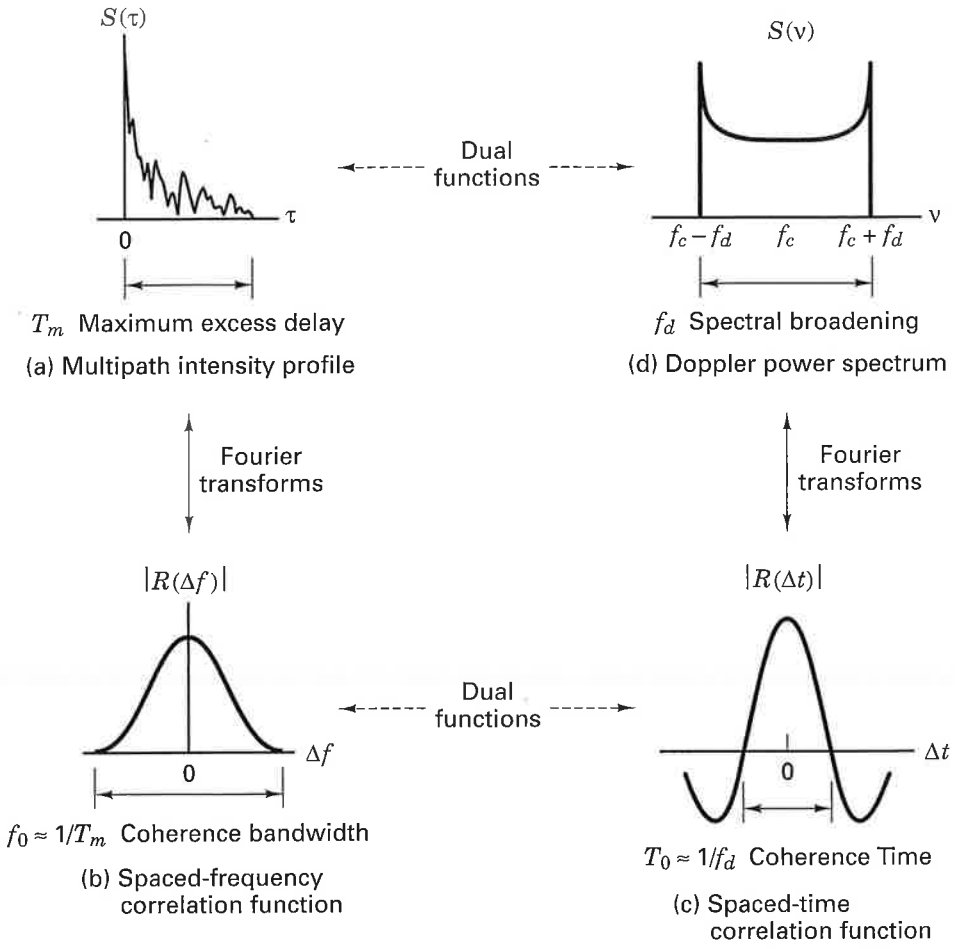


Figure 15.8 Relationships among the channel correlation functions and power density functions.

15.3.1.1 Degradation Categories due to Signal Time-Spreading Viewed in the Time-Delay Domain

In a fading channel, the relationship between maximum excess delay time T_m and symbol time T_s can be viewed in terms of two different degradation categories: *frequency-selective fading* and *frequency nonselective* or *flat fading*. (See Figure 15.1, blocks 8 and 9, and Figure 15.7.) A channel is said to exhibit frequency-selective fading if $T_m > T_s$. This condition occurs whenever the received multipath components of a symbol extend beyond the symbol's time duration. Such multipath dispersion of the signal yields the same kind of ISI distortion that is caused by an electronic filter. In fact, another name for this category of fading degradation is *channel-induced ISI*. In the case of frequency-selective fading, mitigating the

distortion is possible because many of the multipath components are resolvable by the receiver. (Several such mitigation techniques are described in later sections.)

A channel is said to exhibit *frequency nonselective* or *flat fading* if $T_m < T_s$. In this case, all of the received multipath components of a symbol arrive within the symbol time duration; hence, the components are not resolvable. Here, there is no channel-induced ISI distortion, since the signal time spreading does not result in significant overlap among neighboring received symbols. There is still performance degradation, since the unresolvable phasor components can add up destructively to yield a substantial reduction in SNR. Also, signals that are classified as exhibiting flat fading can sometimes experience the distortion effects of frequency-selective fading. This will be explained later when viewing degradation in the frequency domain, where the phenomenon is more easily described. For loss in SNR due to flat fading, the mitigation technique called for is to improve the received SNR (or to reduce the required SNR). For digital systems, introducing some form of signal diversity and using error-correction coding is the most efficient way to accomplish this.

15.3.2 Signal Time-Spreading Viewed in the Frequency Domain

A completely analogous characterization of signal dispersion can be specified in the frequency domain. In Figure 15.8b, the function $|R(\Delta f)|$, designated a *spaced-frequency* correlation function, can be seen; it is the Fourier transform of $S(\tau)$. The function $R(\Delta f)$ represents the correlation between the channel's response to two signals as a function of the frequency difference between the two signals. It can be thought of as the channel's frequency transfer function. Therefore, the time-spreading manifestation can be viewed as if it were the result of a filtering process. Knowledge of $R(\Delta f)$ helps answer the question, "What is the correlation between received signals that are spaced in frequency $\Delta f = f_1 - f_2$?" The function $R(\Delta f)$ can be measured by transmitting a pair of sinusoids separated in frequency by Δf, cross-correlating the complex spectra of the two separately received signals, and repeating the process many times with ever-larger separation Δf. Therefore, the measurement of $R(\Delta f)$ can be made with a sinusoid that is swept in frequency across the band of interest (a wideband signal). The *coherence bandwidth* f_0 is a statistical measure of the range of frequencies over which the channel passes all spectral components with approximately equal gain and linear phase. Thus, the coherence bandwidth represents a frequency range over which a signal's frequency components have a strong potential for amplitude correlation. That is, spectral components in that range are affected by the channel in a similar manner—as, for example, exhibiting fading or no fading. Note that f_0 and T_m are reciprocally related (within a multiplicative constant). As an approximation, we can say that

$$f_0 \approx 1/T_m \tag{15.16}$$

The maximum excess delay T_m is not necessarily the best indicator of how any given system will perform when signals propagate on a channel, because different

channels with the same value of T_m can exhibit very different signal-intensity profiles over the delay span. A more useful parameter is the delay spread, most often characterized in terms of its root-mean-squared (rms) value, called the rms delay spread, as

$$\sigma_\tau = \sqrt{\overline{\tau^2} - (\bar{\tau})^2} \tag{15.17}$$

where $\bar{\tau}$ is the mean excess delay, $(\bar{\tau})^2$ is the mean squared, $\overline{\tau^2}$ is the second moment, and σ_τ is the square root of the second central moment of $S(\tau)$ [1].

A universal relationship between coherence bandwidth and delay spread that would be useful for all applications does not exist. However, using Fourier transform techniques an approximation can be derived from actual signal dispersion measurements in various channels. Several approximate relationships have been developed. If coherence bandwidth is defined as the frequency interval over which the channel's complex frequency transfer function has a correlation of at least 0.9, the coherence bandwidth is approximately [16]

$$f_0 \approx \frac{1}{50\sigma_\tau} \tag{15.18}$$

For the case of a mobile radio, an array of radially uniformly spaced scatterers, all with equal-magnitude reflection coefficients but independent, randomly occurring reflection phase angles [17, 18], is generally accepted as a useful model for an urban propagation environment. This model is referred to as the *dense-scatterer* channel model. With the use of such a model, coherence bandwidth has similarly been defined [17], for a bandwidth interval over which the channel's complex frequency transfer function has a correlation of at least 0.5, to be

$$f_0 = \frac{0.276}{\sigma_\tau} \tag{15.19}$$

Studies involving ionospheric effects often employ the following definition [19]:

$$f_0 = \frac{1}{2\pi\sigma_\tau} \tag{15.20}$$

A more popular approximation of f_0, corresponding to a bandwidth interval having a correlation of at least 0.5, is [1]

$$f_0 \approx \frac{1}{5\sigma_\tau} \tag{15.21}$$

The delay spread and coherence bandwidth are related to a channel's multipath characteristics, differing for different propagation paths (such as metropolitan areas, suburbs, hilly terrain, indoors, etc.). It is important to note that neither of the parameters in Equation (15.21) depend on signaling speed. A system's signaling speed only influences its transmission bandwidth W.

15.3.2.1 Degradation Categories due to Signal Time-Spreading Viewed in the Frequency Domain

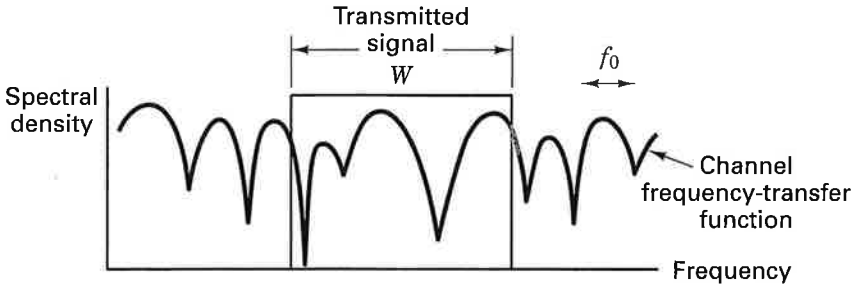
A channel is referred to as frequency-selective if $f_0 < 1/T_s \approx W$, where the symbol rate $1/T_s$ is nominally taken to be equal to the signaling rate or signal bandwidth W. In practice, W may differ from $1/T_s$, due to system filtering or data modulation type (e.g., QPSK, MSK, spread spectrum, etc.) [20]. Frequency-selective fading distortion occurs whenever a signal's spectral components are not all affected equally by the channel. Some of the signal's spectral components falling outside the coherence bandwidth will be affected differently (independently), compared with those components contained within the coherence bandwidth. Figure 15.9 contains three examples. Each one illustrates the spectral density versus frequency of a transmitted signal having a bandwidth of W Hz. Superimposed on the plot in Figure 15.9a is the frequency transfer function of a frequency-selective channel ($f_0 < W$). Figure 15.9a shows that various spectral components of the transmitted signal will be affected differently.

Frequency-nonselective of flat-fading degradation occurs whenever $f_0 > W$. Hence, all of the signal's spectral components will be affected by the channel in a similar manner (e.g., fading or no fading). This is illustrated in Figure 15.9b, which features the spectral density of the same transmitted signal having a bandwidth of W Hz. However, superimposed on this plot is the frequency transfer function of a flat-fading channel ($f_0 > W$). Figure 15.9b illustrates that all of the spectral components of the transmitted signal will be affected in approximately the same way. Flat fading does not introduce channel-induced ISI distortion, but performance degradation can still be expected due to the loss in SNR whenever the signal is fading. In order to avoid channel-induced ISI distortion, the channel is required to exhibit flat fading. This occurs, provided that

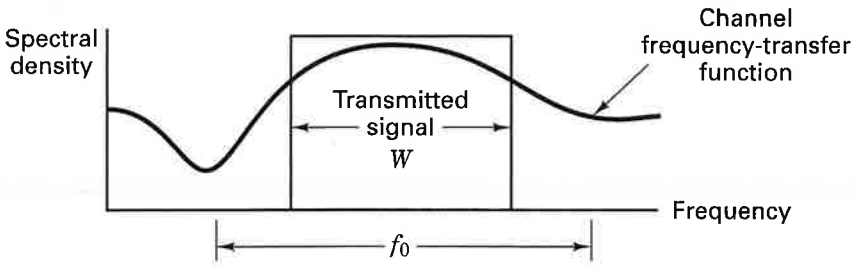
$$f_0 > W \approx \frac{1}{T_s} \tag{15.22}$$

Hence, the channel coherence bandwidth f_0 sets an upper limit on the transmission rate that can be used without incorporating an equalizer in the receiver.

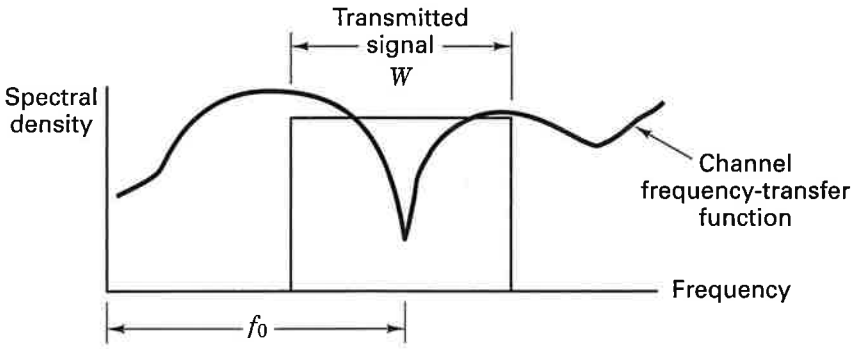
For the flat-fading case, where $f_0 > W$ (or $T_m < T_s$), Figure 15.9b shows the usual flat-fading pictorial representation. However, as a mobile radio changes its position, there will be times when the received signal experiences frequency-selective distortion even though $f_0 > W$. This is seen in Figure 15.9c, where the null of the channel's frequency transfer function occurs near the band center of the transmitted signal's spectral density. When this occurs, the baseband pulse can be especially mutilated by deprivation of its low-frequency components. One consequence of such loss is the absence of a reliable pulse peak on which to establish the timing synchronization or from which to sample the carrier phase carried by the pulse [17]. Thus, even through a channel is categorized as flat fading (based on rms relationships), it can still manifest frequency-selective fading on occasions. It is fair to say that a mobile radio channel, classified as exhibiting flat-fading degradation, cannot exhibit flat fading all of the time. As f_0 becomes much larger than W (or T_m becomes much smaller than T_s), less time will be spent exhibiting the type of



(a) Typical frequency-selective fading case ($f_0 < W$)



(b) Typical flat-fading case ($f_0 > W$)



(c) Null of channel frequency-transfer function occurs at signal band center ($f_0 > W$)

Figure 15.9 Relationships between the channel frequency-transfer function and a transmitted signal with bandwidth W.

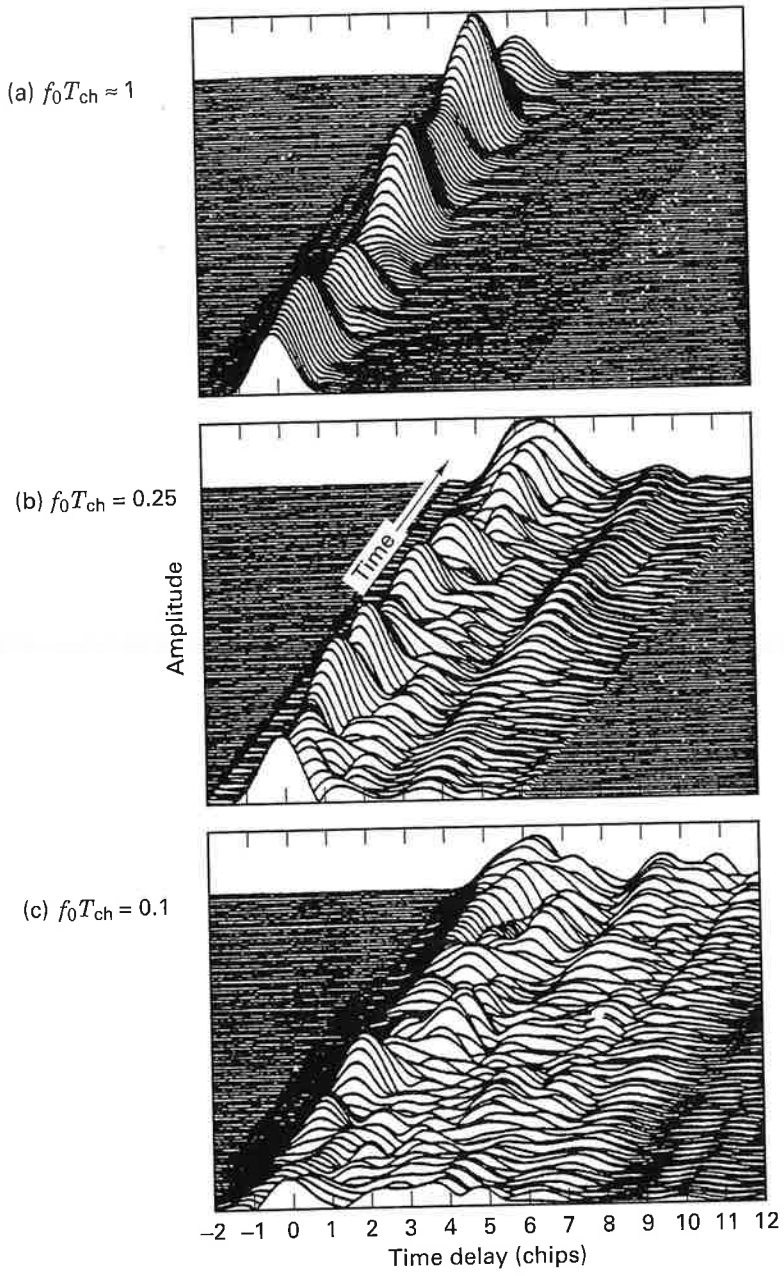


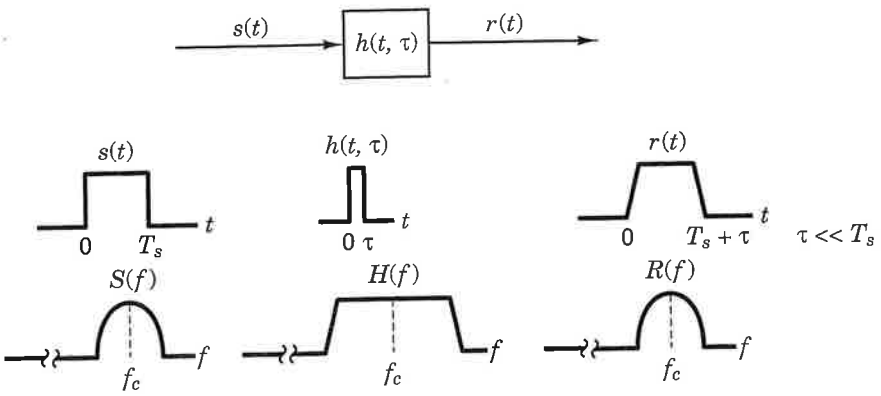
Figure 15.10 DS/SS matched-filter output time-history examples for 3 levels of channel conditions, where T_{ch} is the time duration of a chip.
Source: Bogusch, R.L., *Digital Communications in Fading Channels: Modulation and Coding*, Mission Research Corp., Santa Barbara, California, Report no. MRC-R-1043, March 11, 1987.

condition shown in Figure 15.9c. By comparison, it should be clear that in Figure 15.9a the fading is independent of the position of the signal band, and frequency-selective fading occurs all the time and not just occasionally.

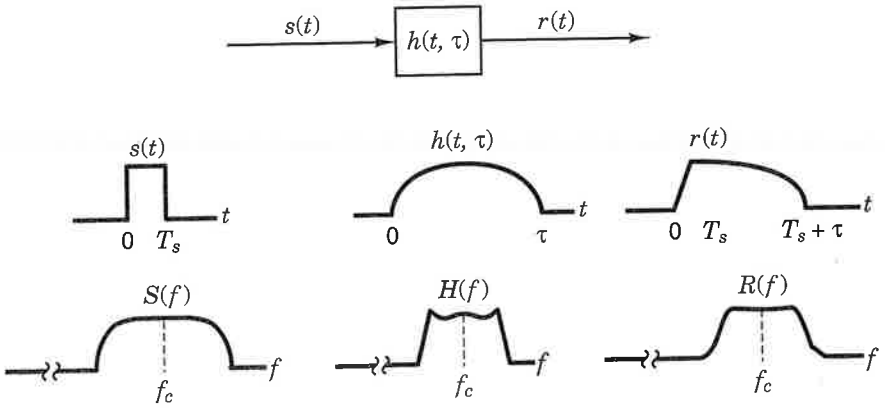
15.3.3 Examples of Flat Fading and Frequency-Selective Fading

Figure 15.10 shows some examples of flat fading and frequency-selective fading for a direct-sequence spread-spectrum (DS/SS) system [19, 20]. In Figure 15.10, there are three plots of the output of a pseudonoise (PN) code correlator versus delay as a function of time (transmission or observation time). Each amplitude-versus-delay plot is akin to $S(\tau)$ versus τ, as shown in Figure 15.8a. The key difference is that the amplitudes shown in Figure 15.10 represent the output of a correlator; hence, the waveshapes are a function not only of the impulse response of the channel but also of the impulse response of the correlator. The delay time is expressed in units of chip durations (chips), where the chip is defined as the spread-spectrum, minimal-duration keying element. For each plot, the observation time is shown on an axis perpendicular to the amplitude-versus-time-delay plane. Figure 15.10 is drawn from a satellite-to-ground communications link exhibiting scintillation because of atmospheric disturbances. However, Figure 15.10 is still a useful illustration of three different channel conditions that might apply to a mobile radio situation. A mobile radio that moves along the observation-time axis is affected by changing multipath profiles along the route, as seen in the figure. The scale along the observation-time axis is also in units of chips. In Figure 15.10a, the signal dispersion (one "finger" of return) is on the order of a chip time duration, T_{ch}. In a typical DS/SS system, the spread-spectrum signal bandwidth is approximately equal to $1/T_{ch}$; hence, the normalized coherence bandwidth f_0T_{ch} of approximately unity in Figure 15.10a, implies that the coherence bandwidth is about equal to the spread-spectrum bandwidth. This describes a channel that can be called frequency-nonselective or slightly frequency-selective. In Figure 15.10b, where $f_0T_{ch} = 0.25$, the signal dispersion is more pronounced. There is definite interchip interference, due to the coherence bandwidth being approximately 25 percent of the spread-spectrum bandwidth. In Figure 15.10c, where $f_0T_{ch} = 0.1$, the signal dispersion is even more pronounced, with greater interchip-interference effects, due to the coherence bandwidth being approximately 10 percent of the spread-spectrum bandwidth. The coherence bandwidths (relative to the spread-spectrum signaling speed) shown in Figures 15.10b and 15.10c depict channels that can be categorized, respectively, as moderately and highly frequency selective. Later, it is shown that a DS/SS system operating over a frequency-selective channel at the chip level does not necessarily experience frequency-selective distortion at the symbol level.

The signal dispersion manifestation of a fading channel is analogous to the signal spreading that characterizes an electronic filter. Figure 15.11a depicts a wideband filter (narrow impulse response) and its effect on a signal in both the time domain and the frequency domain. This filter resembles a flat-fading channel yielding an output that is relatively free of distortion. Figure 15.11b shows a narrowband filter (wide impulse response). The output signal suffers much distortion, as shown in both time and frequency. Here, the process resembles a frequency-selective channel.



(a) Flat fading channel characteristics



(b) Frequency selective fading channel characteristics

Figure 15.11 Flat-fading and frequency-selective fading characteristics.
Source: Rappaport, T.S., *Wireless Communications,* Prentice-Hall, Upper Saddle River, New Jersey, 1996.

15.4 TIME VARIANCE OF THE CHANNEL CAUSED BY MOTION

15.4.1 Time Variance Viewed in the Time Domain

Signal dispersion and coherence bandwidth characterize the channel's time-spreading properties in a local area. However, they do not offer information about the time-varying nature of the channel caused by relative motion between a transmitter and receiver or by movement of objects within the channel. For mobile-radio applications, the channel is time variant because motion between the

transmitter and receiver results in propagation-path changes. For a transmitted continuous wave (CW) signal, such changes cause variations in the signal's amplitude and phase at the receiver. If all scatterers making up the channel are stationary, whenever motion ceases the amplitude and phase of the received signal remains constant—that is, the channel appears to be time invariant. Whenever motion begins again, the channel appears time variant. Since the channel characteristics are dependent on the positions of the transmitter and receiver, time variance in this case is equivalent to spatial variance.

Figure 15.8c shows the function $R(\Delta t)$, designated the *spaced-time* correlation function; it is the autocorrelation function of the channel's response to a sinusoid. This function specifies the extent to which there is correlation between the channel's response to a sinusoid sent at time t_1 and the response to a similar sinusoid sent at time t_2, where $\Delta t = t_2 - t_1$. The *coherence time* T_0 is a measure of the expected time duration over which the channel's response is essentially invariant. Earlier, measurements of signal dispersion and coherence bandwidth were made by using wideband signals. Now, to measure the time-variant nature of the channel, a narrowband signal is used [15]. To measure $R(\Delta t)$, a single sinusoid ($\Delta f = 0$) can be transmitted at times t_1 and t_2, and the cross-correlation function of the received signals is determined. The function $R(\Delta t)$ and the parameter T_0 provide knowledge about the fading rapidity of the channel. Note that for an ideal *time-invariant* channel (e.g., transmitter and receiver exhibiting no motion at all), the channel's response would be highly correlated for all values of Δt; thus, $R(\Delta t)$ as a function of Δt would be a constant. For example, if a stationary user's location is characterized by a multipath null, then that null remains unchanged until there is some movement (either by the transmitter or receiver or by objects within the propagation path). When using the dense-scatterer channel model described earlier, with constant mobile velocity V and an unmodulated CW signal having wavelength λ, the normalized $R(\Delta t)$ is described as [18]

$$R(\Delta t) = J_0(kV\Delta t) \tag{15.23}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind [11], $V\Delta t$ is distance traversed, and $k = 2\pi/\lambda$ is the free-space phase constant (transforming distance to radians of phase). Coherence time can be measured in terms of either time or distance traversed (assuming some fixed velocity). Amoroso described such a measurement using a CW signal and a dense-scatterer channel model [17]. He measured the statistical correlation between the combination of received magnitude and phase sampled at a particular antenna location x_0 and the corresponding combination sampled at some displaced location $x_0 + \zeta$, with displacement measured in units of wavelength λ. For a displacement ζ of 0.4λ between two antenna locations, the combined magnitudes and phases of the received CW are statistically uncorrelated. In other words, the signal observation at x_0 provides no information about the signal at $x_0 + \zeta$. For a given velocity, this displacement is readily transformed into units of time (coherence time).

15.4.1.1 The Basic Fading Manifestations are Independent of One Another

For a moving antenna, the fading of a received carrier wave is usually regarded as a random process, even though the fading record may be completely predetermined from the disposition of scatterers and the propagation geometry from the transmitter to the receiving antenna. This is because the same waveform received by two antennas that are displaced by at least 0.4λ are statistically uncorrelated [17, 18]. Since such a small distance (about 13 cm for a carrier wave at 900 MHz) corresponds to statistical decorrelation in received signals, the basic fading manifestations of signal dispersion and fading rapidity can be considered to be independent of each other. Any of the cases in Figure 15.10 can provide some insight here. At each instant of time (corresponding to a spatial location) we see a multipath intensity profile $S(\tau)$ as a function of delay τ. The multipath profiles are primarily determined by the surrounding terrain (buildings, vegetation, etc.). Consider Figure 15.10b, where the direction of motion through regions of differing multipath profiles is indicated by an arrow labeled *time* (it might also be labeled *antenna displacement*). As the mobile moves to a new spatial location characterized by a different profile, there will be changes in the fading state of the channel as characterized by the profile at the new location. However, because one profile is decorrelated with another profile at a distance as short as 13 cm (for a carrier at 900 MHz), the rapidity of such changes only depends on the speed of movement, not on the underlying geometry of the terrain.

15.4.1.2 The Concept of Duality

The mathematical concept of duality can be defined as follows: Two processes (functions, elements, or systems) are *dual* to each other if their mathematical relationships are the same even though they are described in terms of different parameters. In this chapter, it is interesting to note duality when examining time-domain versus frequency-domain relationships.

In Figure 15.8, we can identify functions that exhibit similar behavior across domains. For the purpose of understanding the fading-channel model, it is useful to refer to such functions as *duals*. For example, the phenomenon of signal dispersion can be characterized in the frequency domain by $R(\Delta f)$, as shown in Figure 15.8b. It yields knowledge about the range of frequencies over which two spectral components of a received signal have a strong potential for amplitude and phase correlation. Fading rapidity is characterized in the time domain by $R(\Delta t)$, as shown in Figure 15.8c. It yields knowledge about the span of time over which two received signals have a strong potential for amplitude and phase correlation. These two correlation functions, $R(\Delta f)$ and $R(\Delta t)$, have been labeled as duals. This is also noted in Figure 15.1 as the duality between blocks 10 and 13, and in Figure 15.7 as the duality between the time-spreading mechanism in the frequency domain and the time-variant mechanism in the time domain.

15.4.1.3 Degradation Categories due to Time Variance, Viewed in the Time Domain

The time-variant nature or fading rapidity mechanism of the channel can be viewed in terms of two degradation categories, as listed in Figure 15.7: *fast fading* and *slow fading*. The term "fast fading" is used for describing channels in which $T_0 < T_s$, where T_0 is the channel coherence time, and T_s is the time duration of a transmission symbol. Fast fading describes a condition where the time duration in which the channel behaves in a correlated manner is short compared with the time duration of a symbol. Therefore, it can be expected that the fading character of the channel will change several times during the time span of a symbol, leading to distortion of the baseband pulse shape. Analogous to the distortion previously described as channel-induced ISI, here distortion takes place because the received signal's components are not all highly correlated throughout time. Hence, fast fading can cause the baseband pulse to be distorted, often resulting in an irreducible error rate. Such distorted pulses cause synchronization problems (failure of phase-locked-loop receivers), in addition to difficulties in adequately designing a matched filter.

A channel is generally referred to as introducing slow fading if $T_0 > T_s$. Here, the time duration in which the channel behaves in a correlated manner is long compared with the time duration of a transmission symbol. Thus, one can expect the channel state to virtually remain unchanged during the time in which a symbol is transmitted. The propagating symbols will likely not suffer from the pulse distortion described earlier. The primary degradation in a slow-fading channel, as with flat fading, is loss in SNR.

15.4.2 Time Variance Viewed in the Doppler-Shift Domain

A completely analogous characterization of the time-variant nature of the channel can be presented in the Doppler-shift (frequency) domain. Figure 15.8d shows a *Doppler power spectral density* (or Doppler spectrum) $S(\nu)$, plotted as a function of Doppler-frequency shift, ν. For the case of the dense-scatterer model, a vertical receive antenna with constant azimuthal gain, a uniform distribution of signals arriving at all arrival angles throughout the range $(0, 2\pi)$, and an unmodulated CW signal, the signal spectrum at the antenna terminals is [18]

$$S(\nu) = \frac{1}{\pi f_d \sqrt{1 - \left(\dfrac{\nu - f_c}{f_d}\right)^2}} \tag{15.24}$$

The equality holds for frequency shifts of ν that are in the range $\pm f_d$ about the carrier frequency f_c, and would be zero outside that range. The shape of the RF Doppler spectrum described by Equation (15.24) is classically bowl-shaped, as seen in Figure 15.8d. Note that the spectral shape is a result of the dense-scatterer channel model. Equation (15.24) has been shown to match experimental data gathered for mobile radio channels [22]; however, different applications yield different

spectral shapes. For example, the dense-scatterer model does not hold for the indoor radio channel; the channel model for an indoor area assumes $S(\nu)$ to be a flat spectrum [23].

In Figure 15.8d, the sharpness and steepness of the boundaries of the Doppler spectrum are due to the sharp upper limit on the Doppler shift produced by a vehicular antenna travelling among the stationary scatterers of the dense scatterer model. The largest magnitude (infinite) of $S(\nu)$ occurs when the scatterer is directly ahead of the moving antenna platform or directly behind it. In that case, the magnitude of the frequency shift is given by

$$f_d = \frac{V}{\lambda} \tag{15.25}$$

where V is relative velocity and λ is the signal wavelength. When the transmitter and receiver move toward each other, f_d is positive, and when they move away from each other, f_d is negative. For scatterers directly broadside of the moving platform, the magnitude of the frequency shift is zero. The fact that Doppler components arriving at exactly 0° and 180° have an infinite power spectral density is not a problem, since the angle of arrival is continuously distributed and the probability of components arriving at exactly these angles is zero [1, 18].

$S(\nu)$ is the Fourier transform of $R(\Delta t)$. It is known that the Fourier transform of the autocorrelation function of a time series equals the magnitude squared of the Fourier transform of the original time series. Therefore, measurements can be made by simply transmitting a sinusoid (narrowband signal) and using Fourier analysis to generate the power spectrum of the received amplitude [15]. This Doppler power spectrum of the channel yields knowledge about the spectral spreading of a transmitted sinusoid (impulse in frequency) in the Doppler-shift domain. As indicated in Figure 15.8, $S(\nu)$ can be regarded as the dual of the multipath intensity profile $S(\tau)$, since the latter yields knowledge about the time spreading of a transmitted impulse in the time-delay domain. This is also noted in Figure 15.1 as the duality between blocks 7 and 16, and in Figure 15.7 as the duality between the time-spreading mechanism in the time-delay domain and the time-variant mechanism in the Doppler-shift domain.

Knowledge of $S(\nu)$ makes it possible to estimate how much spectral broadening is imposed on the signal as a function of the rate of change in the channel state. The width of the Doppler power spectrum (denoted f_d) is referred to in the literature by several different names: *Doppler spread, fading rate, fading bandwidth*, or *spectral broadening*. Equation (15.24) describes the Doppler frequency shift. In a typical multipath environment, the received signal travels over several reflected paths, each with a different distance and a different angle of arrival. The Doppler shift of each arriving path is generally different from that of other paths. The effect on the received signal manifests itself as a Doppler spreading of the transmitted signal frequency, rather than as a shift. Note that the Doppler spread f_d and the coherence time T_0 are reciprocally related (within a multiplicative constant), resulting in an approximate relationship between the two parameters given by

$$T_0 \approx \frac{1}{f_d} \tag{15.26}$$

Hence, the Doppler spread f_d (or $1/T_0$) is regarded as the typical *fading rate* of the channel. Earlier, T_0 was described as the expected time duration over which the channel's response to a sinusoid is essentially invariant. When T_0 is defined more precisely as the time duration over which the channel's response to sinusoids yields a correlation between them of at least 0.5, the relationship between T_0 and f_d is approximately [2]

$$T_0 \approx \frac{9}{16\pi f_d} \tag{15.27}$$

A popular rule of thumb is to define T_0 as the geometric mean of Equations (15.26) and (15.27). This yields

$$T_0 = \sqrt{\frac{9}{16\pi f_d^2}} = \frac{0.423}{f_d} \tag{15.28}$$

For the case of a 900 MHz mobile radio, Figure 15.12 illustrates the typical effect of Rayleigh fading on a signal's envelope amplitude versus time [1]. The figure shows that the distance traveled by the mobile in a time interval corresponding to two adjacent nulls (small-scale fades) is on the order of a half-wavelength ($\lambda/2$). Thus, from Figure 15.12 and Equation (15.25), the time required to traverse a distance $\lambda/2$ (approximately the coherence time), when traveling at a constant velocity V, is

$$T_0 \approx \frac{\lambda/2}{V} = \frac{0.5}{f_d} \tag{15.29}$$

Thus, when the interval between fades is approximately $\lambda/2$, as shown in Figure 15.12, the resulting expression for T_0 in Equation (15.29) is quite close to the geometric mean shown in Equation (15.28). From Equation (15.29), and using the parameters shown in Figure 15.12 (velocity = 120 km/hr, and carrier frequency = 900 MHz), it is straightforward to determine that the channel coherence time is approximately 5 ms and the Doppler spread (channel fading rate) is approximately 100 Hz. Therefore, if this example represents a channel over which digitized speech signals are transmitted with a typical rate of 10^4 symbols/s, the fading rate is considerably less than the symbol rate. Under such conditions, the channel would manifest slow-fading effects. Note that if the abscissa of Figure 15.12 were labeled in units of wavelength instead of time, the plotted fading characteristics would look the same for any radio frequency and any antenna speed.

15.4.2.1 Analogy for Spectral Broadening in Fading Channels

Let us discuss the reason why a signal experiences spectral broadening as it propagates from or is received by a moving platform, and why this spectral broadening (also called the fading rate of the channel) is a function of the speed of motion. An analogy can be used to explain this phenomenon. Figure 15.13 shows the keying of a digital signal (such as amplitude-shift-keying or frequency-shift-keying)

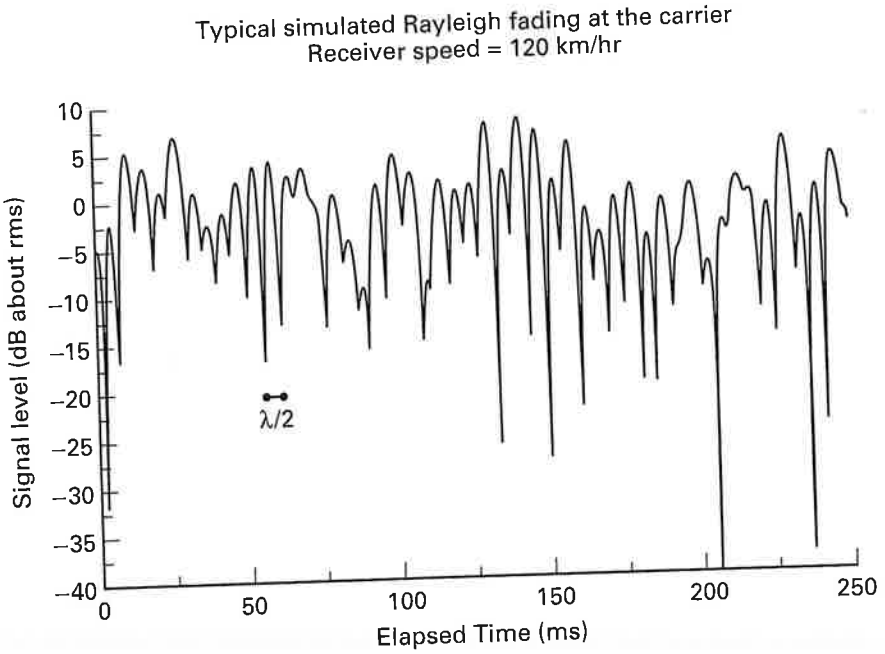


Figure 15.12 A typical Rayleigh fading envelope at 900 MHz.
Rappaport, T.S., *Wireless Communications,* Chapter 4, Prentice Hall, Upper Saddle River, New Jersey, 1996.

where a single tone $\cos 2\pi f_c t$ defined for $-\infty < t < \infty$, is characterized in the frequency domain in terms of impulses (at $\pm f_c$). This frequency domain representation is ideal (i.e., zero bandwidth), since the tone is a single frequency with infinite time duration. In practical applications, digital signaling involves switching (keying) signals on and off at a required rate. The keying operation can be viewed as multiplying the infinite-duration tone in Figure 15.13a by an ideal rectangular on-off (switching) function in Figure 15.13b. The frequency-domain description of this switching function is of the form $\text{sinc}\, fT$. (See Appendix A, Table A.1.)

In Figure 15.13c, the result of the multiplication yields a tone, $\cos 2\pi f_c t$, that is time-duration limited. The resulting spectrum is obtained by convolving the spectral impulses shown in part (a) of Figure 15.13 with the $\text{sinc}\, fT$ function of part (b), yielding the broadened spectrum depicted in part (c). It is further seen that, if the signaling occurs at a faster rate characterized by the rectangle of shorter duration in part (d), the resulting signal spectrum in part (e) exhibits greater spectral broadening. The changing state of a fading channel is somewhat analogous to the on-off keying of digital signals. The channel behaves like a switch, turning the signal "on and off." The greater the rapidity of the change in the channel state, the greater spectral broadening experienced by signals propagating over such a channel. The analogy is not exact because the on and off switching of signals may result in phase discontinuities, while the typical multipath-scatterer environment induces phase-continuous effects.

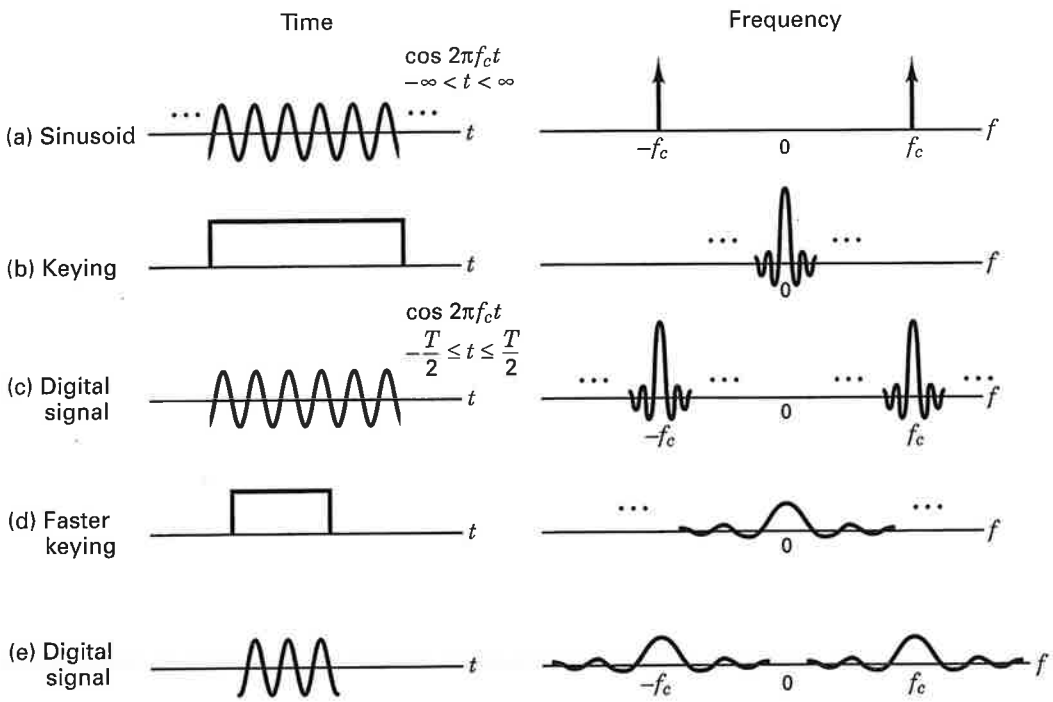


Figure 15.13 Analogy between spectral broadening in fading and spectral broadening in keying a digital signal.

15.4.2.2 Degradation Categories due to Time Variance, Viewed in the Doppler-Shift Domain

A channel is said to be fast fading if the symbol rate $1/T_s$ (approximately equal to the signaling rate or bandwidth W) is less than the fading rate $1/T_0$ (approximately equal to f_d); that is, fast fading is characterized by

$$W < f_d \tag{15.30}$$

or

$$T_s > T_0 \tag{15.31}$$

Conversely, the channel is referred to as slow fading if the signaling rate is greater than the fading rate. Thus, in order to avoid signal distortion caused by fast fading, the channel must be made to exhibit slow fading characteristics by insuring that the signaling rate exceeds the channel fading rate. That is,

$$W > f_d \tag{15.32}$$

or

$$T_s < T_0 \tag{15.33}$$

In Equation (15.22), it was shown that due to signal dispersion, the coherence bandwidth f_0 sets an *upper limit* on the signaling rate that can be used without suffering frequency-selective distortion. Similarly, Equation (15.32) shows that due to Doppler spreading, the channel fading rate f_d sets a *lower limit* on the signaling rate that can be used without suffering fast-fading distortion. For HF communication systems, when teletype or Morse-coded messages were transmitted at low data rates, the channels often exhibited fast-fading characteristics. However, most present-day terrestrial mobile-radio channels can generally be characterized as slow fading.

Equations (15.32) and (15.33) do not go far enough in describing the desirable behavior of the channel. A better way to state the requirement for mitigating the effects of fast fading would be that we desire $W >> f_d$ (or $T_s << T_0$). If this condition is not satisfied, the random frequency modulation (FM) due to varying Doppler shifts will degrade system performance significantly. The Doppler effect yields an irreducible error rate that cannot be overcome by simply increasing E_b/N_0 [24]. This irreducible error rate is most pronounced for any transmission scheme that involves modulating the carrier phase. A single specular Doppler path, without scatterers, registers an instantaneous frequency shift, classically calculated as $f_d = V/\lambda$. However, a combination of specular and multipath components yields a rather complex time dependence of instantaneous frequency, which can cause frequency swings much larger than $\pm V/\lambda$ when the information is recovered by an instantaneous frequency detector (a nonlinear device) [25]. Figure 15.14 illustrates how this can happen. Owing to vehicle motion, at time t_1 the specular phasor has rotated through an angle θ, while the net phasor has rotated through an angle ϕ, which is about four times greater than θ. The rate of change of phase at a time near this particular fade is about 4 times that of the specular Doppler alone. Therefore, the instantaneous frequency shift $d\phi/dt$ would be about 4 times that of the specular doppler shift. The peaking of instantaneous frequency shifts at a time near deep fades is akin to the phenomenon of FM "clicks" or "spikes." Figure 15.15 illustrates the seriousness of this problem. The figure shows bit-error rate versus E_b/N_0 performance plots for $\pi/4$ DQPSK signaling at $f_c = 850$ MHz for various simulated mobile speeds [26]. It should be clear that at high speeds, the performance curve bottoms out at an error-rate level that may be unacceptably high. Ideally, coherent demodulators that lock onto and track the information signal should suppress the effect of this FM noise and thus cancel the impact of Doppler shift. However, for large values of f_d, carrier recovery is difficult to implement because very wideband (relative to the data rate) phase-lock loops (PLLs) need to be designed. For voice-grade applications with bit-errors rates of 10^{-3} to 10^{-4}, a large value of Doppler shift is considered to be on the order of $0.01 \times W$. Therefore, to avoid fast-fading distortion and the Doppler-induced irreducible error rate, the signaling rate should exceed the fading rate by a factor of 100 to 200 [27]. The exact factor depends on the signal modulation, receiver design, and required error rate [1, 25–29]. Davarian [29] showed that a frequency-tracking loop can help lower but not completely remove the irreducible error rate in a mobile system by using differential minimum-shift keying (DMSK) modulation.

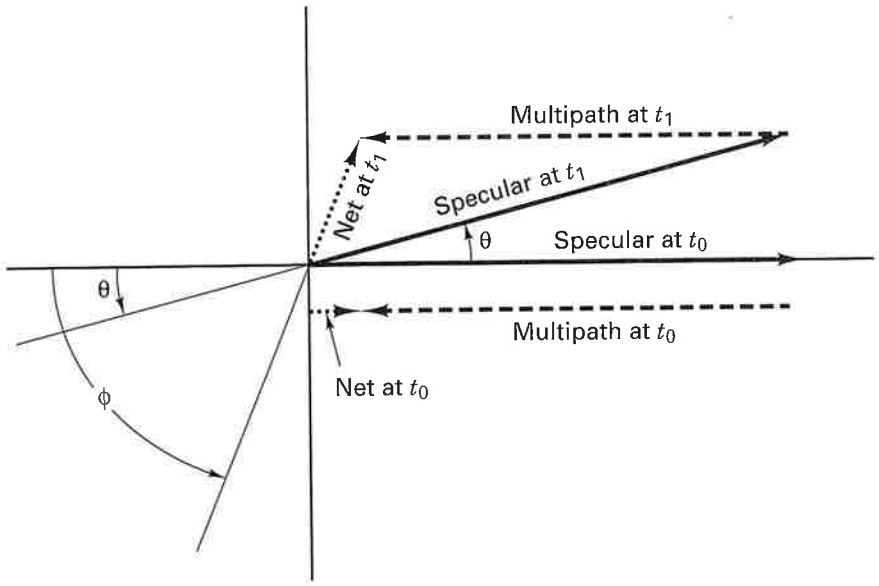


Figure 15.14 A combination of specular and multipath components can register much larger frequency swings than $\pm v/\lambda$.
Source: Amoroso, F. "Instantaneous Frequency Effects in a Doppler Scattering Environment", *IEEE International Conference on Communications*, June 7–10, 1987, pages 1458–1466

15.4.3 Performance over a Slow- and Flat-Fading Rayleigh Channel

For the case of a discrete multipath channel, with a complex envelope $g(t)$ described by Equation (15.3), a demodulated signal (neglecting noise) is described by Equation (15.10), which is rewritten as follows:

$$z(t) = \sum_n \alpha_n(t)\, e^{-j2\pi f_c \tau_n(t)}\, R[t - \tau_n(t)]\, e^{j\phi(t-\tau_n)} \tag{15.34a}$$

where $R(t) = |g(t)|$ is the envelope magnitude, and $\phi(t)$ is its phase. Assume that the channel exhibits flat fading, so that the multipath components are not resolvable. Then the $\{\alpha_n(t)\}$ terms in Equation 15.34a, in one signaling interval T, need to be expressed as a resultant amplitude $\alpha(T)$ of all the n phasors received in that interval. Similarly, the phase terms in Equation 15.34a, in one signaling interval, need to be expressed as the resultant phase $\theta(T)$ of all the n fading phasors plus the information phase received in that interval. Assume also that the channel exhibits slow fading, so that the phase can be estimated from the received signal without significant error using phase-lock loop (PLL) circuitry or some other appropriate techniques. Therefore, for a slow- and flat-fading channel, we can express a received

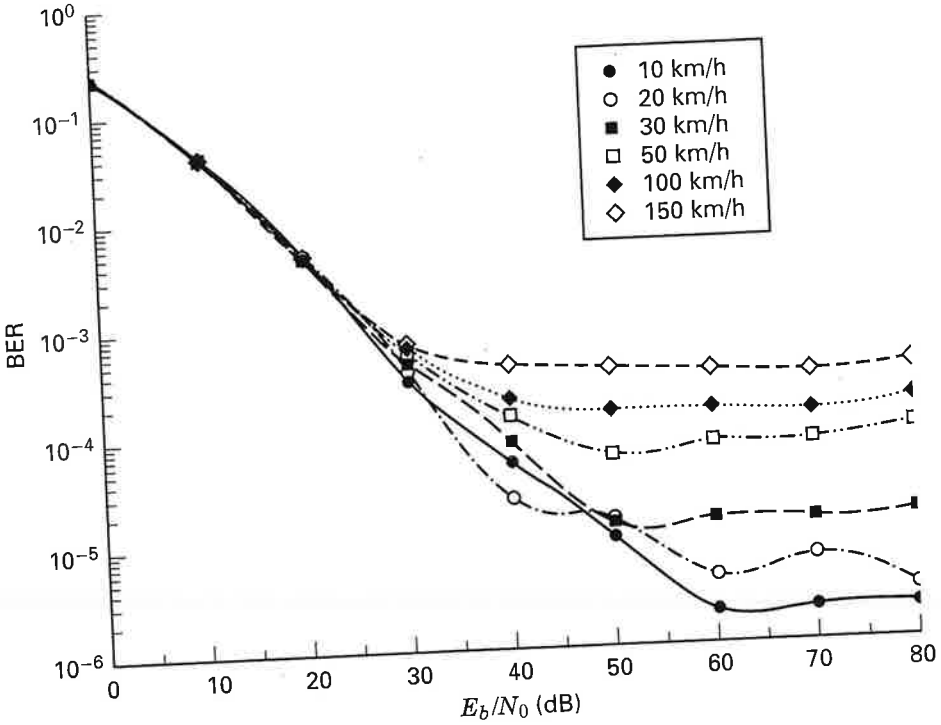


Figure 15.15 Error performance versus E_b/N_0 for $\pi/4$ DQPSK for various mobile speeds: $f_c = 850$ MHz, $R_s = 24$ ksymbol/s.
Source: Fung, V., Rappaport T.S., and Thoma, B., "Bit-Error Simulation for $\pi/4$ DQPSK Mobile Radio Communication Using Two-Ray and Measurement-Based Impulse Response Models," *IEEE Journal on Selected Areas in Communication*, Vol. 11, No. 3, April 1993, pp. 393–405.

test statistic $z(T)$ out of the demodulator in each signaling interval, including the noise $n_0(T)$, as

$$z(T) = \alpha(T)\, R(T) e^{-j[\theta(T) - \phi(T)]} + n_0(T) \tag{15.34b}$$

For simplicity, we now replace $\alpha(T)$ with α. For binary signaling over an AWGN channel with a fixed attenuation of $\alpha = 1$, the bit-error probabilities for the basic coherent and noncoherent PSK and orthogonal FSK are given in Chapter 4, Table 4.1. The plots of bit-error probability versus E_b/N_0 for these signaling schemes each manifest the classical exponential relationship (a waterfall-shape associated with AWGN performance). However, for multipath conditions, if there is no specular signal component, α is a Rayleigh distributed random variable; or equivalently, α^2 is described by a chi-square pdf. Under these Rayleigh fading conditions, Figure 15.16 depicts the performance curves. When $(E_b/N_0)\ \mathbf{E}(\alpha^2) >> 1$, where $\mathbf{E}(\cdot)$ represents

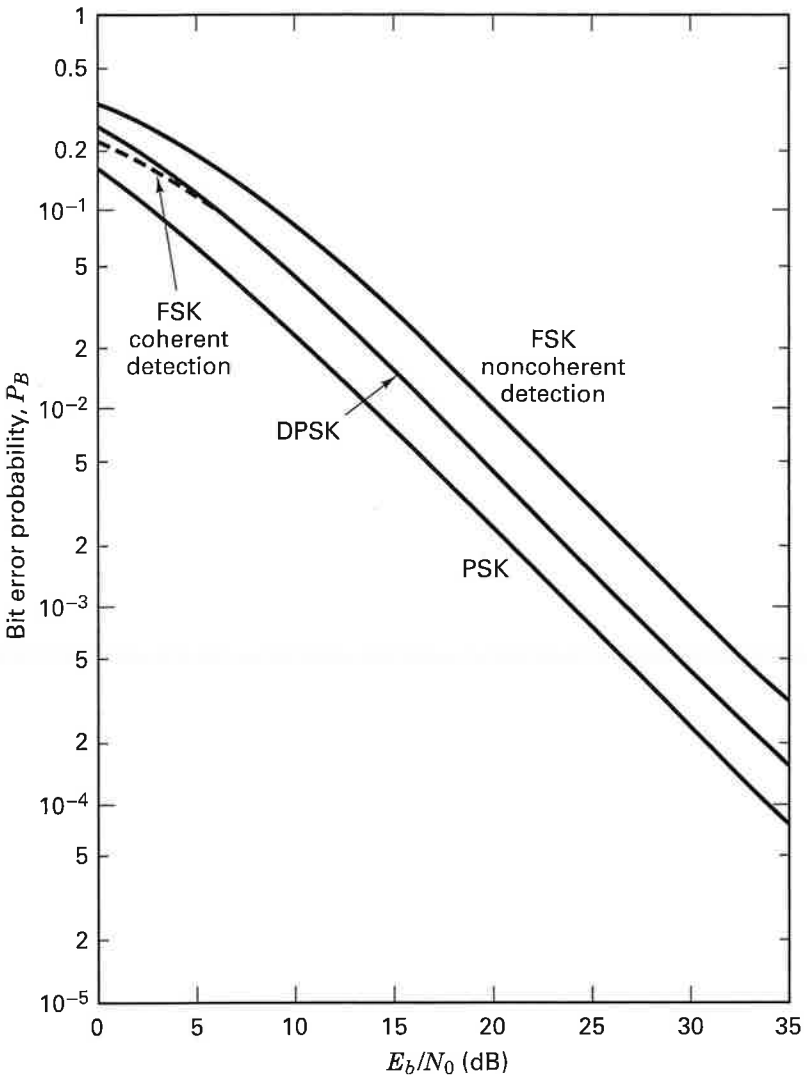


Figure 15.16 Performance of binary signaling over a slow Rayleigh fading channel.
Source: Proakis, J.G., *Digital Communications,* McGraw-Hill Book Company, New York, 1983.

statistical expectation, then the bit-error probability expressions for the basic binary signaling schemes shown in Figure 15.16 are given in Table 15.1. Each of the signaling schemes that had manifested a waterfall-shaped performance plot under AWGN interference in Figure 4.25 of Chapter 4 now exhibits performance that takes the form of an inverse linear function, as a result of the Rayleigh fading.

TABLE 15.1 Rayleigh-Limit Bit-Error Performance (where $(E_b/N_0)\,\mathrm{E}(\alpha^2) >> 1$)

Modulation	P_B
PSK (Coherent)	$\dfrac{1}{4(E_b/N_0)\,\mathbf{E}(\alpha^2)}$
DPSK (Differentially Coherent)	$\dfrac{1}{2(E_b/N_0)\,\mathbf{E}(\alpha^2)}$
Orthogonal FSK (Coherent)	$\dfrac{1}{2(E_b/N_0)\,\mathbf{E}(\alpha^2)}$
Orthogonal FSK (Noncoherent)	$\dfrac{1}{(E_b/N_0)\,\mathbf{E}(\alpha^2)}$

Proakis, J.G., *Digital Communications,* McGraw-Hill, New York, 1983.

15.5 MITIGATING THE DEGRADATION EFFECTS OF FADING

Figure 15.17 subtitled "The Good, The Bad, and The Awful," highlights three major performance categories in terms of bit-error probability P_B versus E_b/N_0. The leftmost exponentially shaped curve highlights the performance that can be expected when using any nominal modulation scheme in AWGN interference. Observe that at a reasonable E_b/N_0 level, good performance can be expected. The middle curve, referred to as the *Rayleigh limit,* shows the performance degradation resulting from a loss in E_b/N_0 that is characteristic of flat fading or slow fading when there is no line-of-sight signal component present. The curve is a function of the reciprocal of E_b/N_0 (an inverse-linear function), so for practical values of E_b/N_0, performance will generally be "bad." In the case of Rayleigh fading, parameters with overbars are often introduced to indicate that an average is being taken over the "ups" and "downs" of the fading experience. Therefore, one often sees such bit-error probability plots with averaged parameters denoted by $\overline{P_B}$ and $\overline{E_b/N_0}$. This notation emphasizes the fact that the fading channel has memory; thus, received samples of the signal are correlated to one another in time. Therefore, when producing such error-probability plots for a fading channel, one needs to examine the process over a window of time that is much larger than the channel coherence time. The curve that reaches an irreducible error-rate level, sometimes called an *error floor,* represents "awful" performance, where the bit-error probability can level off at values nearly equal to 0.5. This shows the severe performance degrading effects that are possible with frequency-selective fading or fast fading.

If the channel introduces signal distortion as a result of fading, the system performance can exhibit an irreducible error rate at a level higher than the desired error rate. In such cases, no amount of E_b/N_0 will help achieve the desired level of performance, and the only approach available for improving performance is to use some other form of mitigation to remove or reduce the signal distortion. The mitigation method depends on whether the distortion is caused by frequency-selective fading or fast fading. Once the signal distortion has been mitigated, the

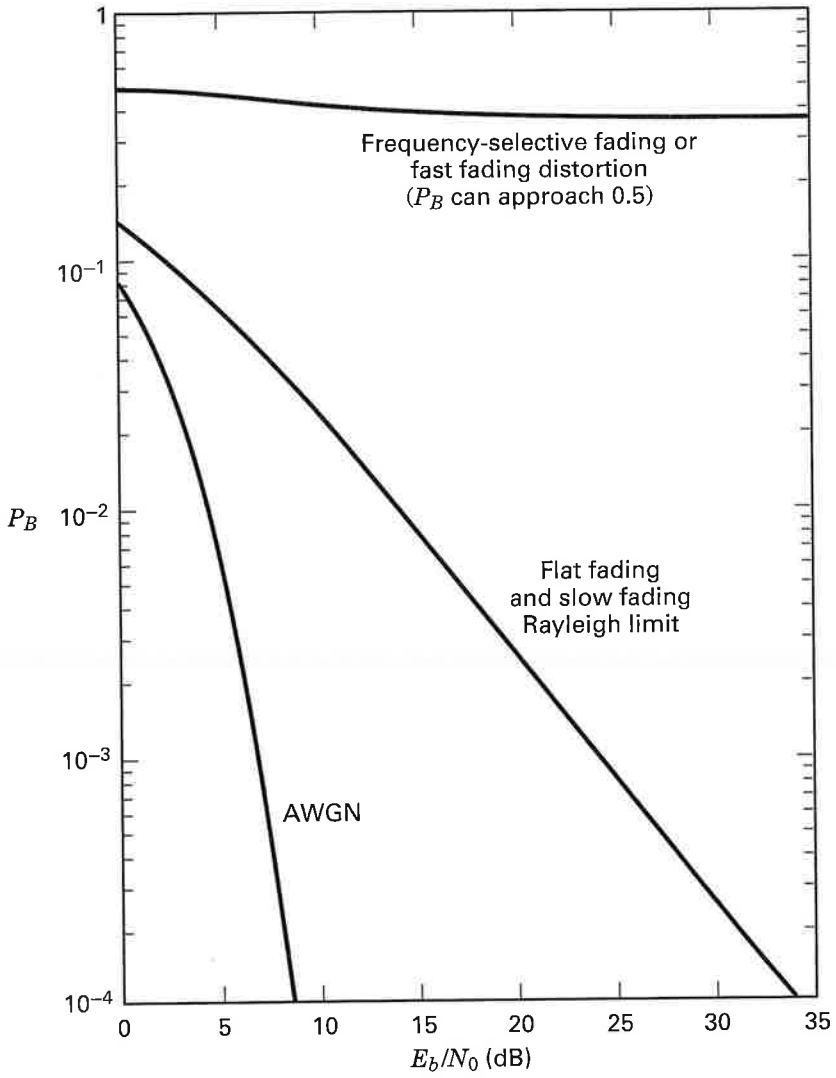


Figure 15.17 Error performance: The good, the bad, and the awful.

P_B-versus-E_b/N_0 performance can transition from the "awful" category to the merely "bad" Rayleigh-limit curve. Next, it is possible to further ameliorate the effects of fading and strive to approach AWGN system performance by using some form of diversity to provide the receiver with a collection of uncorrelated replicas of the signal, and by using a powerful error-correction code.

In Figure 15.18, several mitigation techniques for combating the effects of both signal distortion and loss in SNR are listed. Just as Figures 15.1 and 15.7 serve as a guide for characterizing fading phenomena and their effects, Figure 15.18 can

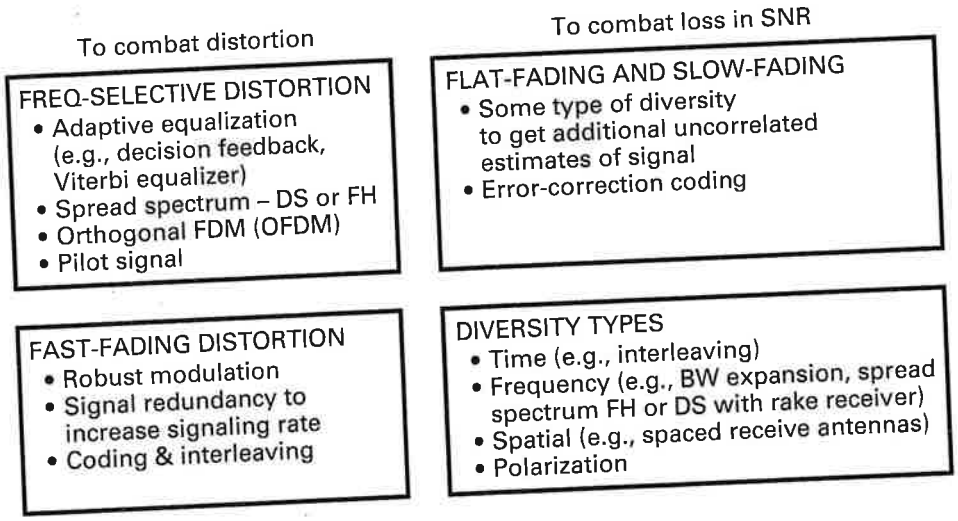


Figure 15.18 Basic mitigation types.

similarly serve to describe mitigation methods that can be used to ameliorate the effects of fading. The mitigation approaches to be used when designing a system should be considered in two basic steps: first, choose the type of mitigation to reduce or remove any distortion degradation; next, choose a diversity type that can best approach AWGN system performance.

15.5.1 Mitigation to Combat Frequency-Selective Distortion

Equalization can mitigate the effects of channel-induced ISI that is brought on by frequency-selective fading. That is, it can help modify system performance described by the curve that is "awful" in Figure 15.17 to the one that is merely "bad." The process of equalizing for mitigating ISI effects involves using methods to gather the dispersed symbol energy back into its original time interval. In effect, an equalizer is an inverse filter of the channel. If the channel is frequency selective, the equalizer enhances the frequency components with small amplitudes and attenuates those with large amplitudes. The goal is for the combination of channel and equalizer filter to provide a flat composite received frequency response and linear phase [30]. Because in a mobile system the channel response varies with time, the equalizer filter must also change or adapt to the time-varying channel characteristics. Such equalizer filters are therefore adaptive devices that accomplish more than distortion mitigation; they also provide diversity. Since distortion mitigation is achieved by gathering the dispersed symbol's energy back into the symbol's original time interval (so that it doesn't hamper the detection of other symbols), the equalizer is simultaneously providing the receiver with symbol energy that would otherwise be lost.

The decision feedback equalizer (DFE) has a feedforward section that is a linear transversal filter [30] whose stage length and tap weights are selected to co-

herently combine virtually all of the current symbol's energy. The DFE also has a feedback section, which removes energy remaining from previously detected symbols [10, 30–32]. The basic idea behind the DFE is that once an information symbol has been detected, the ISI that it induces on future symbols can be estimated and subtracted before the detection of subsequent symbols. (See Section 3.4.3.2.)

A maximum-likelihood sequence estimation (MLSE) equalizer tests all possible data sequences (rather than detecting each received symbol by itself) and chooses the data sequence that is the most probable of all the candidates. The MLSE equalizer was first proposed by Forney [33] and implemented by using the Viterbi decoding algorithm [34]. The MLSE is optimal in the sense that it minimizes the probability of a sequence error. Because the Viterbi decoding algorithm is typically used in the implementation of the MLSE equalizer, this device is often referred to as the *Viterbi equalizer.* Later in this chapter, we illustrate the adaptive equalization performed in the Global System for Mobile (GSM) Communications using the Viterbi equalizer.

Direct-sequence spread-spectrum (DS/SS) techniques can be used to mitigate frequency-selective ISI distortion because the hallmark of spread-spectrum systems is their capability to reject interference, and ISI is a type of interference. Consider a DS/SS binary phase-shift keying (PSK) communication channel comprising one direct path and one reflected path. Assume that the propagation from transmitter to receiver results in a multipath wave that is delayed by τ compared with the direct wave. Neglecting noise, the received signal can be expressed as

$$r(t) = Ax(t)p(t)\cos(2\pi f_c t) + \alpha Ax(t-\tau)p(t-\tau)\cos(2\pi f_c t + \theta) \qquad (15.35)$$

where $x(t)$ is the data signal, $p(t)$ is the pseudonoise (PN) spread code, and τ is the differential time delay between the two paths. The angle θ is a random phase, assumed to be uniformly distributed in the range $(0, 2\pi)$, and α is the attenuation of the multipath signal relative to the direct path signal. The receiver multiples the incoming $r(t)$ by the code $p(t)$. If the receiver is synchronized to the direct path signal, multiplication by the code signal yields

$$r(t)p(t) = Ax(t)p^2(t)\cos(2\pi f_c t) + \alpha Ax(t-\tau)p(t)p(t-\tau)\cos(2\pi f_c t + \theta) \qquad (15.36)$$

where $p^2(t) = 1$. If τ is greater than the PN chip duration, then

$$\left|\int p(t)p(t-\tau)dt\right| \ll \left|\int p^2(t)dt\right| \qquad (15.37)$$

over some appropriate interval of integration (correlation). Thus, the spread spectrum system effectively eliminates the multipath interference by virtue of its code-correlation receiver. Even though channel-induced ISI is typically transparent to DS/SS systems, such systems suffer from the loss in energy contained in the multipath components rejected by the receiver. The need to gather up this lost energy belonging to such multipath chips was the motivation for developing the Rake receiver [35–37]. The Rake receiver dedicates a separate correlator to each multipath component (finger) and coherently adds the energy from each finger by selectively delaying

each (the earliest component gets the longest delay), so that they can all the coherently combined.

Earlier, we described a channel that could be classified as flat fading but occasionally exhibits frequency-selective distortion when the null of the channel's frequency transfer function occurs at the center of the signal band. The use of DS/SS is a practical way of mitigating such distortion because the wideband SS signal can span many lobes of the selectively faded channel frequency response. Hence, a great deal of pulse energy is passed by the scatterer medium, in contrast to the channel nulling effect on a relatively narrowband signal [17]. (See Figure 15.9c.) The ability of the signal spectrum to span over many lobes of the frequency-selective channel transfer function is the key to how DS/SS signaling can overcome the degrading effects of a multipath environment. This requires that the spread-spectrum bandwidth W_{ss} (or the chip rate, R_{ch}) be greater than the coherence bandwidth f_0. The larger the ratio of W_{ss} to f_0, the more effective will be the mitigation. The time-domain view of such mitigation was expressed in Equations (15.36) and (15.37). That is, to resolve multipath components (and either reject them or, with a Rake receiver, exploit them) requires that the spread-spectrum signal dispersion be greater than a chip time.

Frequency hopping spread spectrum (FH/SS) can be used as a technique to mitigate the distortion caused by frequency-selective fading, provided the hopping rate is at least equal to the symbol rate. Compared with DS/SS, mitigation takes place through a different mechanism. FH systems avoid the degradation effects due to multipath by rapidly changing the carrier-frequency band in the transmitter. Interference is avoided by similarly changing the band position in the receiver before the arrival of the multipath signal.

Orthogonal frequency-division multiplexing (OFDM) can be used for signal transmission in frequency-selective fading channels to avoid the use of an equalizer by lengthening the symbol duration. The approach is to partition (demultiplex) a high symbol-rate sequence into N symbol groups, so that each group contains a sequence of a lower symbol rate (by the factor $1/N$) than the original sequence. The signal band is made up of N orthogonal carrier waves, and each one is modulated by a different symbol group. The goal is to reduce the symbol rate (signaling rate) $W \approx 1/T_s$ on each carrier such that it is less than the channel's coherence bandwidth f_0. OFDM, originally referred to as *Kineplex*, is a technique that has been implemented in the United States in mobile radio systems [38] and has been chosen by the European community, under the name Coded OFDM (COFDM), for high-definition television (HDTV) broadcasting [39].

Pilot Signal is the name given to a signal intended to facilitate the coherent detection of waveforms. Pilot signals can be implemented in the frequency domain as in-band tones [40], or they can be implemented in the time domain as digital sequences, which can also provide information about the channel state and thus improve performance in fading conditions [41].

15.5.2 Mitigation to Combat Fast-Fading Distortion

Fast-fading distortion calls for the use of a robust modulation (noncoherent or differentially coherent) scheme that does not require phase tracking, and that reduces the detector integration time [19]. Another technique is to increase the symbol rate

$W \approx 1/T_s$, such that it is greater than the fading rate $f_d \approx 1/T_0$, by adding signal redundancy. Error-correction coding can also provide mitigation; instead of providing more signal energy, a code reduces the required E_b/N_0 for a desired error performance. For a given E_b/N_0, with coding present, the error floor out of the demodulator will not be lowered, but a lower error rate out of the decoder can be achieved [19]. Thus, with coding, one can get acceptable error performance and in effect withstand a large error floor from the demodulator that might have otherwise been unacceptable. To realize these coding benefits, errors out of the demodulator should be uncorrelated (which will generally be the case in a fast-fading environment), or an interleaver must be incorporated into the system design.

An interesting filtering technique can provide mitigation when fast-fading distortion and frequency-selective distortion occur simultaneously. The frequency-selective distortion can be mitigated by the use of an OFDM signal set. Fast fading, however, will typically degrade conventional OFDM because the Doppler spreading corrupts the orthogonality of the OFDM subcarriers. A polyphase filtering technique [42] is used to provide time-domain shaping and partial-response coding (see Section 2.9) to reduce the spectral sidelobes of the signal set and thus help preserve its orthogonality. The process introduces known ISI and adjacent channel interference (ACI), which are then removed by a post-processing equalizer and canceling filter [43].

15.5.3 Mitigation to Combat Loss in SNR

After implementing some mitigation technique to combat signal distortion due to frequency-selective or fast fading, the next step is to use diversity methods to move the system operating point from the error-performance curve labeled as "bad" in Figure 15.17 to a curve that approaches AWGN performance. The term "diversity" is used to denote the various methods available for providing the receiver with uncorrelated renditions of the signal of interest. Uncorrelated is the important feature here, since it would not help the receiver to have additional copies of a signal if the copies were all equally poor. Following are some of the ways in which diversity methods can be implemented:

- *Time diversity* can be provided by transmitting the signal on L different time slots with time separation of at least T_0. Interleaving, when used along with error-correction coding, is a form of time diversity.
- Frequency diversity can be provided by transmitting the signal on L different carriers with frequency separation of at least f_0. Bandwidth expansion is a form of frequency diversity. The signal bandwidth W is expanded so that it is greater than f_0, thus providing the receiver with several independently fading signal replicas. This achieves frequency diversity of the order $L = W/f_0$. Whenever W is made larger than f_0, there is the potential for frequency-selective distortion unless mitigation in the form of equalization is provided. Thus, an expanded bandwidth can improve system performance (via diversity) only if the frequency-selective distortion that the diversity may have introduced is mitigated.

- *Spread Spectrum Systems* are systems that excel at rejecting interfering signals by the use of bandwidth expansion techniques. In the case of direct-sequence spread-spectrum (DS/SS), it was demonstrated earlier that multipath components are rejected if they are time delayed by more than the duration of one chip. However, in order to approach AWGN performance, it is necessary to compensate for the loss in energy contained in those rejected components. The Rake receiver (described later) makes it possible to coherently combine the energy from several of the multipath components arriving along different paths (with sufficient differential delay). Thus, used with a Rake receiver, DS/SS modulation can be said to achieve path diversity. The Rake receiver is needed in phase-coherent reception; but in differentially coherent bit detection, a simple delay—equivalent to the duration of one bit with complex conjugation—can be implemented [44].
- *Frequency-hopping spread-spectrum* (FH/SS) is sometimes used as a diversity mechanism. The GSM system uses slow FH (217 hops/s) to compensate for those cases where the mobile unit is moving very slowly (or not at all) and experiencing deep fading due to a spectral null.
- *Spatial diversity* is usually accomplished through the use of multiple receive antennas, separated by a distance of at least 10 wavelengths when located at a base station (and less when located at a mobile unit). Signal processing techniques must be employed to choose the best antenna output or to coherently combine all the outputs. Systems have also been implemented with multiple transmitters, each at a different location, as in the Global Positioning System (GPS).
- *Polarization diversity* [45] is yet another way to achieve additional uncorrelated samples of the signal.
- Any diversity scheme may be viewed as a trivial form of repetition coding in space or time. However, techniques exist for improving the loss in SNR in a fading channel that are more efficient and more powerful than repetition coding. Error-correction coding represents a unique mitigation technique, because, instead of providing more signal energy, it reduces the required E_b/N_0 needed to achieve a desired performance level. Error-correction coding, coupled with interleaving [19, 46–51], is probably the most prevalent of the mitigation schemes used to provide improved system performance in a fading environment. Note that the time diversity mechanism obtained through interleaving relies on the vehicle motion to spread the errors during the fading. The faster the speed of the mobile unit, the more effective is the interleaver. The interleaver is less effective at slow speeds. (This vehicle speed-versus-interleaver performance is demonstrated in Section 15.5.6.)

15.5.4 Diversity Techniques

The goal in implementing diversity techniques is to utilize additional independent (or at least uncorrelated) signal paths to improve the received SNR. Diversity can provide improved system performance at relatively low cost; unlike equalization, diversity requires no training overhead. In this section, we show the error-

performance improvements that can be obtained with the use of diversity techniques. The bit-error-probability, $\overline{P}_B$, averaged through all the "ups and downs" of the fading experience in a slow-fading channel can be computed as follows:

$$\overline{P}_B = \int_0^\infty P_B(x)\, p(x)\, dx \tag{15.38}$$

where $P_B(x)$ is the bit-error probability for a given modulation scheme at a specific value of SNR = x, where $x = \alpha^2 E_b/N_0$, and $p(x)$ is the pdf of x due to the fading conditions. With E_b and N_0 constant, α is used to represent the amplitude variations due to fading (see Section 15.2.2).

For Rayleigh fading, α has a Rayleigh distribution so α^2, and consequently x, have a chi-squared distribution. Thus, following the form of Equation (15.15),

$$p(x) = \frac{1}{\Gamma} \exp\left(-\frac{x}{\Gamma}\right) \qquad x \geq 0 \tag{15.39}$$

where $\Gamma = \overline{\alpha^2}\, E_b/N_0$ is the SNR averaged through the "ups and downs" of fading. If each diversity (signal) branch, $i = 1, \ldots, M$, has an instantaneous SNR = γ_i, and we assume that each branch has the same average SNR given by Γ, then

$$p(\gamma_i) = \frac{1}{\Gamma} \exp\left(-\frac{\gamma_i}{\Gamma}\right) \qquad \gamma_i \geq 0 \tag{15.40}$$

The probability that a single branch has SNR less than some threshold γ, is

$$P(\gamma_i \leq \gamma) = \int_0^\gamma p(\gamma_i) d\gamma_i = \int_0^\gamma \frac{1}{\Gamma} \exp\left(-\frac{\gamma_i}{\Gamma}\right) d\gamma_i \tag{15.41}$$

$$= 1 - \exp\left(-\frac{\gamma}{\Gamma}\right)$$

The probability that all M independent signal diversity branches are received simultaneously with an SNR less than some threshold value γ, is

$$P(\gamma_1, \ldots, \gamma_M \leq \gamma) = \left[1 - \exp\left(-\frac{\gamma}{\Gamma}\right)\right]^M \tag{15.42}$$

The probability that any single branch achieves SNR > γ is

$$P(\gamma_i > \gamma) = 1 - \left[1 - \exp\left(-\frac{\gamma}{\Gamma}\right)\right]^M \tag{15.43}$$

This is the probability of exceeding a threshold when selection diversity is used.

Example 15.1 Benefits of Diversity

Assume that 4-branch diversity is used and that each branch receives an independently Rayleigh-fading signal. If the average SNR is Γ = 20 dB, determine the probability that all 4 branches are received simultaneously with an SNR less than 10 dB (and also, the probability that this threshold will be exceeded). Compare the results with the case in which no diversity is used.

Solution

Using Equation (15.42) with $\gamma = 10$ dB, and $\gamma/\Gamma = 10$ dB $- 20$ dB $= -10$ dB $= 0.1$, we solve for the probability that the SNR will drop below 10 dB, as follows:

$$P(\gamma_1, \gamma_2, \gamma_3, \gamma_4 \leq 10 \text{ dB}) = [1 - \exp(-0.1)]^4 = 8.2 \times 10^{-5}$$

Or, using selection diversity, we can say that

$$P(\gamma_i > 10 \text{ dB}) = 1 - 8.2 \times 10^{-5} = 0.9999$$

Without diversity,

$$P(\gamma_1 \leq 10 \text{ dB}) = [1 - \exp(-0.1)]^1 = 0.095$$

$$P(\gamma_1 > 10 \text{ dB}) = 1 - 0.095 = 0.905$$

15.5.5.1 Diversity Combining Techniques

The most common techniques for combining diversity signals are *selection, feedback, maximal ratio, and equal gain.* For systems using spatial diversity, selection involves the sampling of M antenna signals and sending the largest one to the demodulator. Selection diversity combining is relatively easy to implement, however, it is not optimal because it does not make use of all the received signals simultaneously.

With feedback or scanning diversity, instead of using the largest of M signals, the M signals are scanned in a fixed sequence until one that exceeds a given threshold is found. This one becomes the chosen signal until it falls below the established threshold and the scanning process starts again. The error-performance of this technique is somewhat inferior to the other methods, but feedback diversity is quite simple to implement.

In the case of maximal-ratio combining, the signals from all of the M branches are weighted according to their individual SNRs and then summed. The individual signals must be cophased before being summed. The control algorithms for setting gains and delays are similar to those used in equalizers and in Rake receivers. Maximal-ratio combining produces an average SNR $\overline{\gamma_M}$ equal to the sum of the individual average SNRs, as shown below [30].

$$\overline{\gamma_M} = \sum_{i=1}^{M} \overline{\gamma_i} = \sum_{i=1}^{M} \Gamma = M\Gamma \tag{15.44}$$

where we assume that each branch has the same average SNR given by $\overline{\gamma_i} = \Gamma$. Thus, maximal ratio combining can produce an acceptable average SNR, even when none of the individual $\overline{\gamma_i}$ is acceptable. It uses each of the M branches in a cophased and weighted manner such that the largest possible SNR is available at the receiver. Equal-gain combining is similar to maximal-ratio combining, except that the weights are all set to unity. The possibility of achieving an acceptable output SNR from a number of unacceptable inputs is still retained. The performance is marginally inferior to maximal ratio combining. (See reference [52] for a detailed treatment of diversity combining.)

15.5.5 Modulation Types for Fading Channels

It should be apparent that an amplitude-based signaling scheme, such as amplitude shift keying (ASK) or quadrature amplitude modulation (QAM), is inherently vulnerable to performance degradation in a fading environment. Thus, for fading channels, the preferred choice for a signaling scheme is a frequency or phase-based modulation type.

In considering orthogonal FSK modulation for fading channels, the use of MFSK (with $M = 8$ or larger) is useful because its error performance is better than binary signaling. In slow Rayleigh fading channels, binary DPSK and 8-FSK perform within 0.1 dB of each other [19]. At first glance, one might argue that a higher order orthogonal alphabet expands the transmission bandwidth, which at some point may cause the coherence bandwidth of the channel to be exceeded, leading to frequency-selective fading. However, for MFSK, the transmission bandwidth that must be available is much larger than the bandwidth of the propagating signal. For example, consider the case of 8-FSK and a symbol rate of 10,000 symbols/s. The transmission bandwidth is $MR_s = 80{,}000$ Hertz. This is the bandwidth that must be available for the system's use. However, each time that a symbol is transmitted, only one single-sideband tone (having a spectral occupancy of 10,000 hertz) is sent, not the whole alphabet. In considering PSK modulation for fading channels, higher order modulation alphabets perform poorly. MPSK with $M = 8$ or larger should be avoided [19]. Example 15.2 examines a mobile communication system to substantiate such avoidance.

Example 15.2 Variations in a Mobile Communication System

The Doppler spread $f_d = V/\lambda$ shows that the fading rate is a direct function of velocity. Table 15.2 shows the Doppler spread versus vehicle speed at carrier frequencies of 900 MHz and 1800 MHz. Calculate the phase variation per symbol for the case of signaling with QPSK modulation at the rate of 24.3 kilosymbols/s. Assume that the carrier frequency is 1800 MHz and that the velocity of the vehicle is 50 miles/hr (80 km/hr). Repeat for a vehicle speed of 100 miles/hr.

Solution

$$\Delta\theta/\text{symbol} = \frac{f_d \text{ Hz}}{R_s \text{ symbols/s}} \times 360^\circ$$

$$= \frac{132 \text{ Hz}}{24.3 \times 10^3 \text{ symbols/s}} \times 360^\circ$$

$$= 2^\circ/\text{symbol}$$

At a velocity of 100 miles/hr: $\Delta\theta/\text{symbol} = 4^\circ/\text{symbol}$

Thus, it should be clear why MPSK with a value of $M > 4$ is not generally used to transmit information in a multipath environment.

TABLE 15.2 Doppler Spread versus Vehicle Speed

Velocity		Doppler (Hz)	Doppler (Hz)
miles/hr	km/hr	900 MHz ($\lambda = 33$ cm)	1800 MHz ($\lambda = 16.6$ cm)
3	5	4	8
20	32	27	54
50	60	66	132
80	108	106	212
120	192	160	320

15.5.6 The Role of an Interleaver

In Section 8.2, the various attributes of an interleaver are described. For transmission in a multipath environment, the primary benefit of an interleaver is to provide time diversity (when used along with error-correction coding). The larger the time span over which the channel symbols are separated, the greater chance that contiguous bits (after deinterleaving) will have been subjected to uncorrelated fading manifestations, and thus the greater chance of achieving effective diversity. Figure 15.19 illustrates the benefits of providing an interleaver time span T_{IL}, which is large compared with the channel coherence time T_0, for the case of DBPSK modulation with soft-decision decoding of a rate $\frac{1}{2}$, $K = 7$ convolutional code, over a slow Rayleigh-fading channel. It should be apparent that an interleaver having the largest ratio of T_{IL}/T_0 is the best performing (large demodulated BER leading to small decoded BER). This leads to the conclusion that T_{IL}/T_0 should be some large number—say, 1000 or 10,000. However, in a real-time communication system, this is not possible because the inherent time delay associated with an interleaver would be excessive. As described in Section 8.2.1 for the case of a block interleaver, before the first row of an array can be transmitted, virtually the entire array must be loaded. Similarly at the receiver, before the array can be deinterleaved, virtually the entire array must be stored. This leads to a delay of one block of data each at the transmitter and receiver. In Example 15.3, it is shown that for a cellular telephone system with a carrier frequency of 900 MHz, a T_{IL}/T_0 ratio of 10 is about as large as one can implement without suffering excessive delay.

It is interesting to note that the interleaver provides no benefit against multipath unless there is motion between the transmitter and receiver (or motion of objects within the signal-propagating paths). As the motion increases in velocity, so does the benefit of a given interleaver to the error performance of the system. (Don't use this as an excuse for exceeding a highway speed limit.) This is shown in Figure 15.20, where part (a) of the figure shows a terrain that is mapped out with attenuation factors $\{\alpha_i\}$, for a particular mobile communications link over a particular terrain. In the region between the points d_0 and d_1, the attenuation factor is α_1. Between the points d_1 and d_2, the attenuation factor is α_2, and so forth. Assume that the points d_i are equally separated by a distance Δd. Part (b) of the figure shows an automobile that is traveling at a slow speed; as the vehicle traverses a distance Δd, nine symbols are emitted from its transmitter. Assume that the inter-

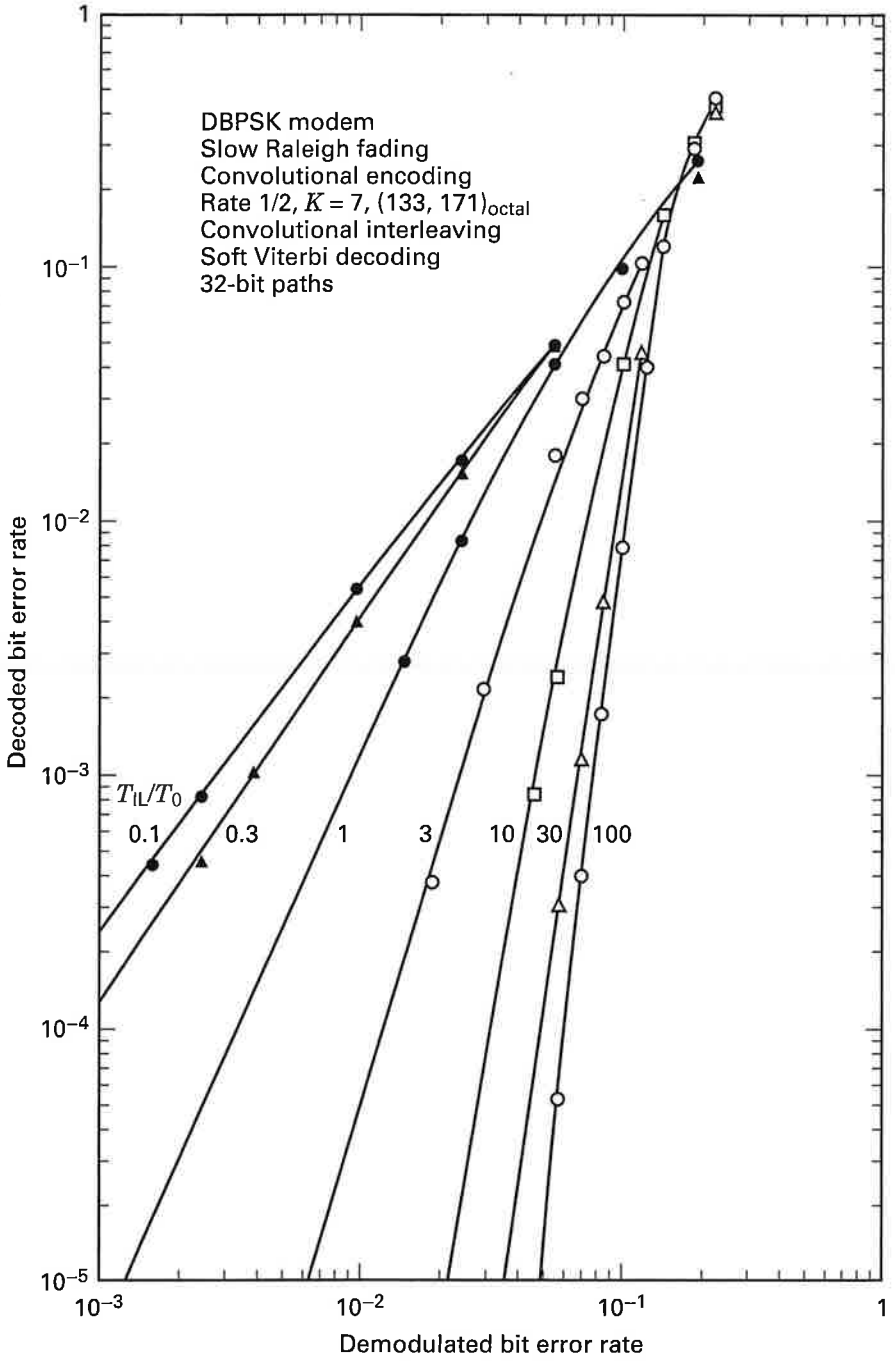
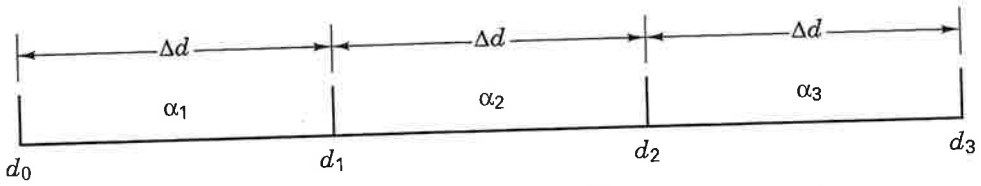
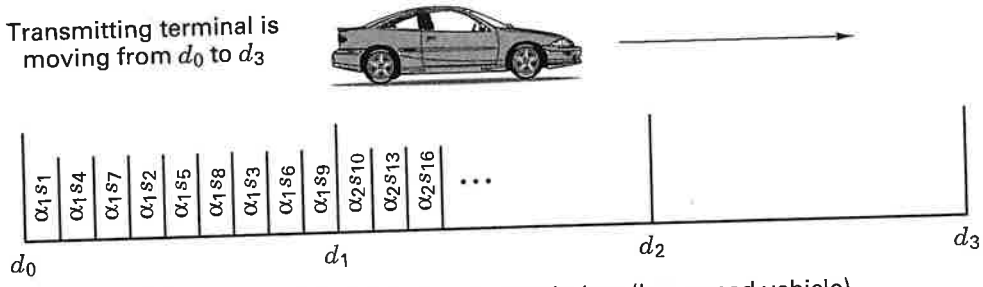


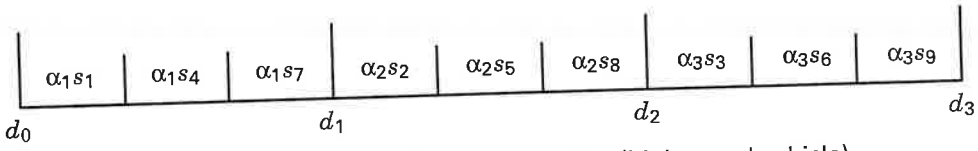
Figure 15.19 Error performance for various ratios of interleaver span to coherence time.



(a) Attenuation α_i versus distance



(b) Transmission of interleaved symbols s_i (low speed vehicle)



(c) Transmission of interleaved symbols s_i (high speed vehicle)

Figure 15.20 The benefits of interleaving improve with increased vehicle speed.

leaver has a span of three symbol intervals, so that symbols s_1 through s_9 appear in the permuted order shown in part (b) of the figure. Notice that all nine of the symbols experience the same attenuation α_1, so that after deinterleaving there is no benefit obtained by using an interleaver with this small a span. Now consider part (c) of the figure, where the vehicle is moving 3 times faster than it is in part (b); thus, as the vehicle traverses a distance Δd, only three symbols are emitted from its transmitter. As before, the symbols are affected by the regional attenuation yielding the nine symbol sequence shown in part (c) of the figure. After deinterleaving of the sequence shown in part (c), we have the following attenuation factor-symbol pairs result: $\alpha_1 s_1$, $\alpha_2 s_2$, $\alpha_3 s_3$, $\alpha_1 s_4$, $\alpha_2 s_5$, $\alpha_3 s_6$, $\alpha_1 s_7$, $\alpha_2 s_8$, $\alpha_3 s_9$. It can be seen that adjacent symbols are affected by different attenuation factors. Thus, the interleaver with too small of a span to yield any benefit at low speeds does provide benefit at faster speeds.

Figure 15.21 also provides evidence that, although communications degrade with increased speed of the mobile unit (the fading rate increases), the benefit of an interleaver is enhanced with increased speed. Figure 15.21 shows the results of field testing performed on a CDMA system meeting the Interim Specification 95

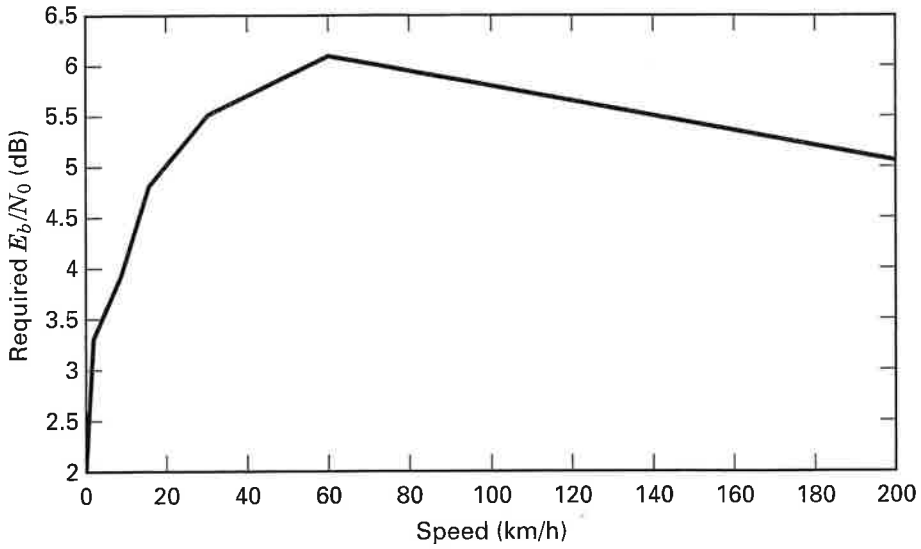


Figure 15.21 Typical E_b/N_0 performance versus vehicle speed for 850 MHz links to achieve a frame-error rate of 1 percent over a Rayleigh channel with two independent paths.

(IS-95) over a link comprising a moving vehicle and a base station [53]. The figure shows a plot of required E_b/N_0 versus vehicle speed to maintain a frame (20 ms of data) error rate of 1%. The best performance (smallest E_b/N_0 requirement) is achieved at low speeds from 0 to 20 km/hr. This slow-speed region is where the system power-control methods can most effectively compensate for the effects of slow fading; at these slow speeds the interleaver cannot provide any benefit, and the plot shows a steep degradation as a function of speed. For velocity in the range of 20–60 km/hr, the steepness of this degradation is reduced. This is the range in which the dynamics of the system power control cannot quite keep up with the increase in fading rapidity; and at the same time, the interleaver does not yet provide sufficient benefit. The speed of 60 km/hr represents the worst error-performance case for this system. As the vehicle goes faster than 60 km/hr, the power control no longer provides any benefits against fading, but the interleaver provides a steadily increasing (with speed) performance improvement. The interleaver's task of transforming the effects of a deep fade (time-correlated degradation events) into random events becomes easier with increased speed. In summary, the system error performance over a fading channel typically degrades with increased speed because of the increase in Doppler spread or fading rapidity. However, the action of an interleaver in the system provides mitigation, which becomes more effective at higher speeds. This trend toward improved error performance doesn't continue indefinitely. Eventually, the irreducible error floor seen in Figure 15.15 dominates. Therefore, if the type of measurements shown in Figure 15.21 were made at speeds beyond 200 km/hr, there

would be a point at which the curve would turn around and steadily show the degradation effects due to the increased Doppler.

15.6 SUMMARY OF THE KEY PARAMETERS CHARACTERIZING FADING CHANNELS

We summarize the conditions that must be met so the channel does not introduce frequency-selective distortion and fast-fading distortion. Combining the expressions given by Equations (15.22), (15.32) and (15.33), we obtain

$$f_0 > W > f_d \tag{15.45}$$

or

$$T_m < T_s < T_0 \tag{15.46}$$

In other words, it is desired that the channel coherence bandwidth exceed the signaling rate, which in turn should exceed the fading rate of the channel. Recall that without distortion mitigation, f_0 sets an upper limit and f_d sets a lower limit on the signaling rate.

15.6.1 Fast-Fading Distortion: Case 1

If the conditions of Equations (15.45) and (15.46) are not met, distortion will result unless appropriate mitigation is provided. Consider the fast-fading case in which the signaling rate is less than the channel fading rate. That is,

$$f_0 > W < f_d \tag{15.47}$$

Mitigation consists of using one or more of the following methods (see Figure 15.18):

- Choose a modulation/demodulation technique that is most robust under fast-fading conditions. That means, for example, avoiding schemes that require PLLs for carrier recovery, since the fast fading could keep a PLL from achieving lock conditions.
- Incorporate sufficient redundancy so that the transmission symbol rate exceeds the channel fading rate but at the same time does not exceed the coherence bandwidth. The channel can then be classified as flat fading. However, as pointed out in Section 15.3.3, even flat-fading channels will experience frequency-selective distortion whenever a channel transfer function exhibits a spectral null near the signal band center. Since this happens only occasionally, mitigation can be accomplished by adequate error-correction coding and interleaving.
- The previously mentioned two mitigation approaches should result in the demodulator operating at the Rayleigh limit [19]. (See Figure 15.17.) However, the probability of error versus E_b/N_0 curve may exhibit flattening (as seen in

Figure 15.15) due to the FM noise that results from the random Doppler spreading. The use of an in-band pilot tone and a frequency-control loop can decrease the level at which the performance curve exhibits the flattening effect.

- To avoid the error floor due to random Doppler spreading, the signaling rate should be increased to about 100–200 times the fading rate [27]. This is one motivation for designing mobile communication systems to operate in a time-division multiple access (TDMA) mode.
- Incorporate error-correction coding and interleaving to further improve system performance.

15.6.2 Frequency-Selective Fading Distortion: Case 2

Consider the frequency-selective case in which the coherence bandwidth is less than the symbol rate, while the symbol rate is greater than the Doppler spread. That is,

$$f_0 < W > f_d \tag{15.48}$$

Since the transmission symbol rate exceeds the channel fading rate, there is no fast-fading distortion. However, mitigation of frequency-selective effects is necessary. One or more of the following techniques may be considered (see Figure 15.18):

- Adaptive equalization, spread spectrum (DS or FH), OFDM, pilot signal. The European GMS system uses a midamble training sequence in each transmission time slot, so that the receiver can estimate the impulse response of the channel. A Viterbi equalizer (explained later) is implemented for mitigating the frequency-selective distortion.
- Once the distortion effects have been reduced, diversity techniques (as well as error-correction coding and interleaving) should be introduced in order to approach AWGN performance. For direct-sequence spread-spectrum (DS/SS) signaling, the use of a Rake receiver (explained later) can be used for providing diversity by coherently combining multipath components that would otherwise be lost.

14.6.3 Fast-Fading and Frequency-Selective Fading Distortion: Case 3

Consider the case in which the channel coherence bandwidth is less than the signaling rate, which in turn is less than the fading rate. This condition is mathematically described by

$$f_0 < W < f_d \tag{15.49}$$

or

$$f_0 < f_d \tag{15.50}$$

Clearly, the channel exhibits both fast fading and frequency-selective fading. Recall from Equations (15.45) and (15.46) that f_0 sets an upper limit and f_d sets a lower limit on the signaling rate. Thus, the condition described by Equation (15.50) presents a difficult design problem, because, unless distortion mitigation is provided, the *maximum* allowable signaling rate is, strictly speaking, *less* than the *minimum* allowable signaling rate. Mitigation in this case is similar to the initial approach outlined in case 1:

- Choose a modulation/demodulation technique that is most robust under fast fading conditions.
- Use transmission redundancy in order to increase the transmitted symbol rate.
- Provide some form of frequency-selective fading mitigation in a manner similar to that outlined in case 2.
- Once the distortion effects have been reduced, introduce some form of diversity (as well as error-correction coding and interleaving), in order to approach AWGN performance.

Example 15.3 Equalizers and Interleavers for Mobile Communications

Consider a cellular telephone that is located in a vehicle travelling at 60 miles per hour (96 km/hr). The carrier frequency is 900 MHz. Use the GSM equalizer test profile shown in Figure 15.22 to determine the following: (a) the rms delay spread, σ_τ, (b) the maximum allowable signal bandwidth $W \approx 1/T_s$ that does not require the use of an equalizer, (c) when operating over a channel with the delay spread found in part (a), which of the following systems requires an equalizer—the United States Digital Cellular Standard (USDC) known as IS-54 (updated to IS-136), the Global System for Mobile (GSM) Communications, CDMA systems designed to meet IS-95? The bandwidths and symbol rates of these systems are USDC: $W = 30$ kHz, $1/T_s = 24.3$ kilosymbols/s; GSM: $W = 200$ kHz, $1/T_s = 271$ kilosymbols/s; and IS-95: $W = 1.25$ MHz, $1/T_s = 9.6$ kilosymbols/s. (d) The total (transmitter plus receiver) time delay caused by the interleaver, when the ratio of interleaver span to coherence time T_{IL}/T_0 is equal to 10. If the total tolerable time delay (transmitter plus receiver) for speech is 100 ms, can such an interleaver be implemented for speech? (e) Repeat parts (a) though (d) for a carrier frequency of 1900 MHz.

Solution

(a) In Figure 15.22, the GSM test profile shows an idealized multipath component (finger) located at each of six delay times $\{\tau_k\}$ from 0 to 16 μs. Each finger can be designated by $S(\tau_k)$, its average relative power, which in this profile is unity (0 dB). The profile represents a *fictitious* multipath environment used for equalization testing [15]. With the finger locations shown on the figure, the mean delay spread is computed as follows:

$$\bar{\tau} = \frac{\sum_k S(\tau_k)\tau_k}{\sum_k S(\tau_k)} = \frac{0 + 3.2 + 6.4 + 9.6 + 12.8 + 16.0}{6} = 8\ \mu s$$

The second moment of delay spread, $\overline{\tau^2}$, and the rms delay spread, σ_τ, are respectively computed as

Figure
profile.

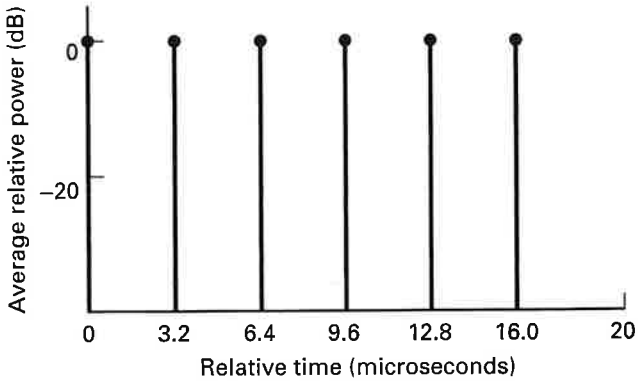


Figure 15.22 GSM equalizer test profile.

$$\overline{\tau^2} = \frac{\sum_k S(\tau_k)\,\tau_k^2}{\sum_k S(\tau_k)} = \frac{0 + 3.2^2 + 6.4^2 + 9.6^2 + 12.8^2 + 16.0^2}{6} = 93.87\ \mu s^2$$

and using Equation (15.17), we compute

$$\sigma_\tau = \sqrt{\overline{\tau^2} - (\bar{\tau})^2} = \sqrt{93.87 - 8^2} = 5.5\ \mu s$$

(b) Using Equation (15.21), the channel coherence bandwidth is determined to be

$$f_0 \approx \frac{1}{5\sigma_\tau} = \frac{1}{5 \times 5.5\ \mu s} = 36.4\ \text{kHz}$$

Thus, the maximum allowable signal bandwidth that will not require the implementation of an equalizer is $W = 36.4$ kHz.

(c) For the various system bandwidths given in this example, it is apparent that the need for an equalizer in USDC is marginal, while in GSM it is definitely required. With regard to systems that are designed to meet IS-95, since the signaling rate or transmission bandwidth W of 1.25 MHz is much larger than the coherence bandwidth f_0 of 36.4 kHz, the system exhibits frequency-selective fading. However, in such direct-sequence spread-spectrum (DS/SS) systems, W is purposely spread with the intent of exceeding f_0 and thus mitigating the effects of frequency-selective fading. An equalizer is only required if ISI poses a problem, and ISI is not a problem if the symbol rate is smaller than the coherence bandwidth (or the symbol duration is larger than the multipath spread). Hence, in the IS-95 example, since the symbol rate of 9.6 ksymbols/s is considerably smaller than the coherence bandwidth, an equalizer is not needed. A Rake receiver, which is described in Section 15.7.2, is used for exploiting path diversity; at the chip level its implementation resembles that of an equalizer.

(d) To determine the interleaver delay, we compute the Doppler spread and coherence time using Equations (15.25) and (15.29), as follows:

$$f_d = \frac{V}{\lambda} = \frac{\dfrac{96{,}000\ \text{m/hr}}{3600\ \text{s/hr}}}{\dfrac{3 \times 10^8\ \text{m/s}}{9 \times 10^8\ \text{Hz}}} = 80\ \text{Hz}, \qquad \text{Thus, } T_0 \approx \frac{0.5}{f_d} = 6.3\ \text{ms}$$

Based on the requirement that $T_{IL}/T_0 = 10$, the interleaver span is $T_{IL} = 63$ ms, making the total transmitter plus receiver delay time equal to 126 ms. For speech, this may be in the marginally acceptable range. Mobile systems often use interleavers with shorter spans that produce one-way delays in the range of 20–40 ms.

(e) Repeating for a carrier frequency of 1900 MHz, the coherence bandwidth calculations are unaffected by the change in carrier frequency, but the Doppler spread, coherence time, and interleaver delay must be computed again. The results are

$$f_d = \frac{V}{\lambda} = 169 \text{ Hz}, \qquad \text{Thus, } T_0 \approx \frac{0.5}{f_d} = 3 \text{ ms}$$

Thus, the interleaver span is $T_{IL} = 30$ ms, making the total transmitter plus receiver delay equal to 60 ms, which is acceptable for speech signals.

15.7 APPLICATIONS: MITIGATING THE EFFECTS OF FREQUENCY-SELECTIVE FADING

15.7.1 The Viterbi Equalizer as Applied to GSM

Figure 15.23 shows the GSM time-division multiple access (TDMA) frame, having a duration of 4.615 ms and comprising 8 slots, one assigned to each active mobile user. A normal transmission burst occupying one time slot contains 57 message bits on each side of a 26-bit midamble, called a *training* or *sounding sequence.* The slot-time duration is 0.577 ms (or the slot rate is 1733 slots/s). The purpose of the midamble is to assist the receiver in estimating the impulse response of the channel adaptively (during the time duration of each 0.577 ms slot). In order for the technique to be effective, the fading characteristics of the channel must not change appreciably during the time interval of one slot. In other words, there cannot be any fast-fading degradation during a time slot when the receiver analyzes the midamble distortion; otherwise, efforts to compensate for the channel's fading characteristics will not be effective. Consider for example a GSM receiver used aboard a high-speed train, traveling at a constant velocity of 200 km/hr (55.56 m/s). Assume the carrier frequency to be 900 MHz, (the wavelength is $\lambda = 0.33$ m). From Equation (15.29), the distance corresponding to a half-wavelength is traversed in

$$T_0 \approx \frac{\lambda/2}{V} \approx 3 \text{ ms} \tag{15.51}$$

which, as indicated in Equation (15.51), corresponds approximately to the coherence time. Therefore, the channel coherence time is more than 5 times greater than the slot time of 0.577 ms. The time needed for a significant change in channel fading characteristics is relatively long compared with the time duration of one slot. Note that the choices made for GSM in the design of its TDMA slot time and midamble were undoubtedly influenced by the need to preclude fast-fading effects, which could cause the equalizer to be ineffective. The GSM symbol rate (or bit rate, since the modulation is binary) is 271 kilosymbols/s and the bandwidth W is 200 kHz. Since the typical rms delay spread σ_τ in an urban environment is on the order of

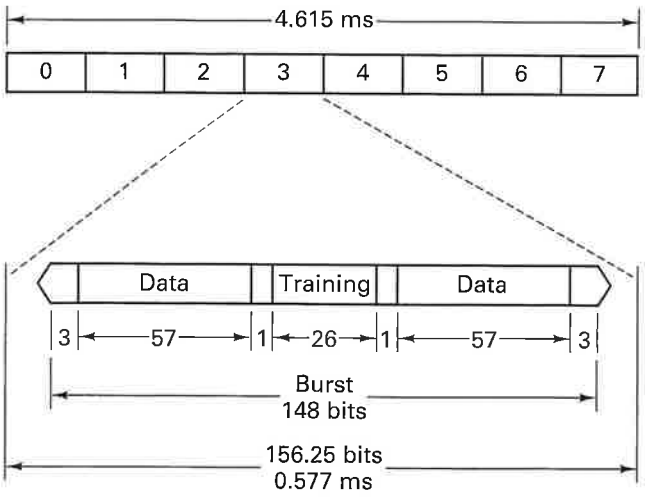


Figure 15.23 The GSM TDMA frame and time slot containing a normal burst.

2 μs, then using Equation (15.21), it can be seen that the resulting coherence bandwidth f_0 is approximately 100 kHz. It should therefore be apparent that since $f_0 < W$, the GSM receiver must utilize some form of mitigation to combat frequency-selective distortion. To accomplish this goal, the Viterbi equalizer is typically implemented.

Figure 15.24 illustrates the basic functional blocks used in a GSM receiver for estimating the channel impulse response. This estimate is used to provide the detector with channel-corrected reference waveforms [54], as explained below. In the final step, the Viterbi algorithm is used to compute the MLSE of the message bits. A received signal can be described in terms of the transmitted signal convolved with the impulse response of the channel. Let $s_{tr}(t)$ denote the transmitted midamble training sequence, and $r_{tr}(t)$ denote the corresponding received midamble training sequence. Thus,

$$r_{tr}(t) = s_{tr}(t) * h_c(t) \tag{15.52}$$

where $*$ denotes convolution and noise has been neglected. At the receiver, since $r_{tr}(t)$ is part of the received normal burst, it is extracted and sent to a filter having impulse response $h_{mf}(t)$ that is matched to $s_{tr}(t)$. This matched filter yields at its output an estimate of $h_c(t)$, denoted $h_e(t)$, and developed from Equation (15.52) as

$$\begin{aligned} h_e(t) &= r_{tr}(t) * h_{mf}(t) \\ &= s_{tr}(t) * h_c(t) * h_{mf}(t) \\ &= R_s(t) * h_c(t) \end{aligned} \tag{15.53}$$

where $R_s(t) = s_{tr}(t) * h_{mf}(t)$ is the autocorrelation function of $s_{tr}(t)$. If $s_{tr}(t)$ is designed to have a highly peaked (impulse-like) autocorrelation function $R_s(t)$, then $h_e(t) \approx h_c(t)$. Next, using a windowing function $w(t)$, we truncate $h_e(t)$ to form a computationally affordable function $h_w(t)$. The time duration of $w(t)$, denoted L_o,

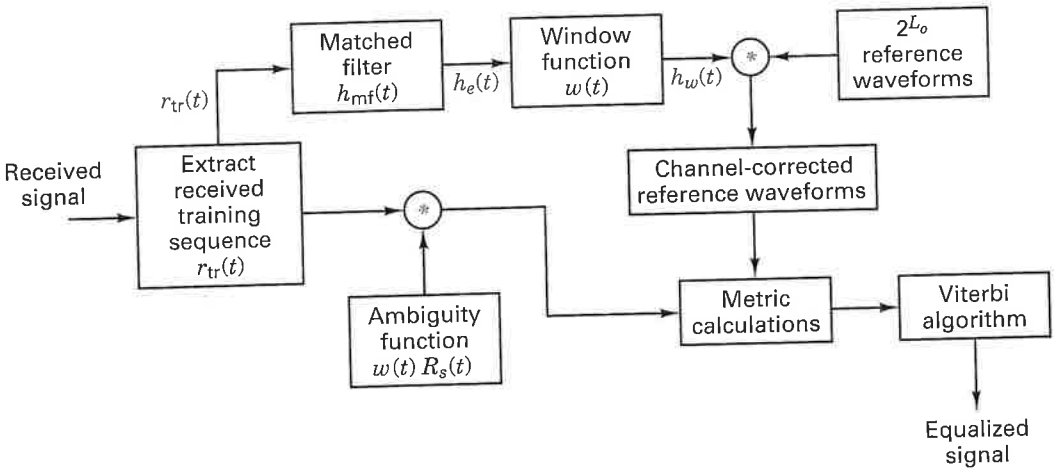


Figure 15.24 The Viterbi equalizer as applied to GSM.

must be large enough to compensate for the effect of typical channel-induced ISI. The term L_o consists of the sum of two contributions—namely, L_{CISI}, corresponding to the controlled ISI caused by Gaussian filtering of the baseband waveform (which then modulates the carrier using MSK), and L_C, corresponding to the channel-induced ISI caused by multipath propagation. Therefore, L_o can be written as

$$L_o = L_{\text{CISI}} + L_C \tag{15.54}$$

The GSM system is required to provide distortion mitigation caused by signal dispersion having delay spreads of approximately 15–20 μs. Since, in GSM, the bit duration is 3.69 μs, we can express L_o in units of bit intervals. Thus, the Viterbi equalizer used in GSM has a memory of 4 to 6 bit intervals. For each L_o-bit interval in the message, the function of the Viterbi equalizer is to find the most likely L_o-bit sequence out of the 2^{L_o} possible sequences that might have been transmitted. Determining the most likely transmitted L_o-bit sequence requires that 2^{L_o} meaningful reference waveforms be created by modifying (or distorting) the 2^{L_o} ideal waveforms (generated at the receiver) in the same way that the channel has distorted the transmitted slot. Therefore, the 2^{L_o} reference waveforms are convolved with the windowed estimate of the channel impulse response $h_w(t)$, in order to generate the distorted or so-called channel-corrected reference waveforms. Next, the channel-corrected reference waveforms are compared with the received data waveforms to yield metric calculations. However, before the comparison takes place, the received data waveforms are convolved with the known windowed autocorrelation function $w(t)R_s(t)$, transforming them in a manner comparable to the transformation applied to the reference waveforms. This filtered message signal is compared with all possible 2^{L_o} channel-corrected reference signals, and metrics are computed in a manner similar to that used in the Viterbi decoding algorithm (VDA).

The VDA yields the maximum likelihood estimate of the transmitted data sequence [34].

Note that most equalizing techniques use filters in order to compensate for the non-optimum properties of $h_c(t)$; that is, equalizing filters attempt to modify the distorted pulse shapes. However, the operation of a Viterbi equalizer is quite different. It entails making measurements of $h_c(t)$, and then providing a means of adjusting the receiver to the channel environment. The goal of such adjustments is to enable the detector to make good estimates from the distorted pulse sequence. With a Viterbi equalizer, the distorted samples are not reshaped or directly compensated in any way; instead the mitigating technique is for the receiver to adjust itself in such a way that it can better deal with the distorted samples.

15.7.2 The Rake Receiver Applied to Direct-Sequence Spread-Spectrum (DS/SS) Systems

Interim Specification 95 (IS-95) describes a DS/SS cellular system that uses a Rake receiver [35–37] to provide path diversity. The Rake receiver searches through the different multipath delays for code correlation and thus recovers delayed signals, which are then optimally combined with the output of other independent correlators. In Figure 15.25, the power profiles associated with the five chip transmissions of the code sequence 1 0 1 1 1 are shown, where the observation times are labeled t_{-4} for the earliest transmission and t_0 for the latest. Each abscissa shows three "fingers" or components arriving with delays τ_1, τ_2, and τ_3. Assume that the intervals between the transmission times t_i and the intervals between the delay times τ_i are each one chip in duration. From this, one can conclude that the finger arriving at

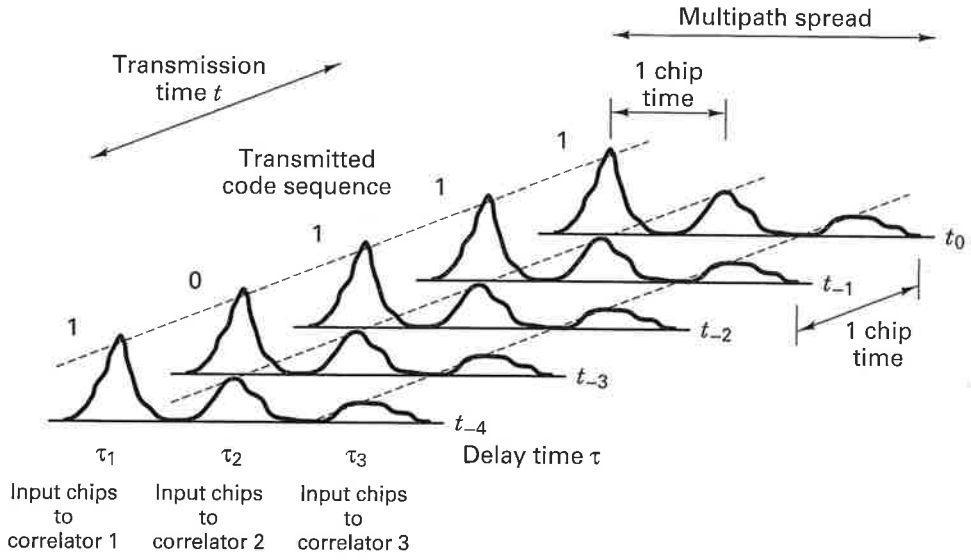


Figure 15.25 Example of received chips seen by a 3-finger rake receiver.

the receiver at time t_{-4}, with delay τ_3, is time coincident with two other fingers, namely the fingers arriving at times t_{-3} and t_{-2}, with delays τ_2 and τ_1, respectively. Since, in this example, the delayed components are separated by at least one chip time, they can be resolved. At the receiver, there must be a sounding device that is dedicated to estimating the τ_i delay times. Note that, for a terrestrial mobile radio system, the fading rate is relatively slow (in the order of milliseconds) or the channel coherence time is relatively large, compared with the chip time duration ($T_0 > T_{ch}$). Hence, the changes in τ_i occur slowly enough that the receiver can readily adapt to them.

Once the τ_i delays are estimated, a separate correlator is dedicated to recovering each resolvable multipath finger. In this example, there would be three such dedicated correlators, each one processing a delayed version of the same chip sequence 1 0 1 1 1. In Figure 15.25, each correlator receives chips with power profiles represented by the sequence of fingers shown along a diagonal line. For simplicity, the chips are all shown as positive signaling elements. In reality, these chips form a PN sequence, which of course, contains both positive and negative pulses. Each correlator attempts to correlate these arriving chips with the same appropriately synchronized PN code. At the end of a symbol interval (typically, there are hundreds or even thousands of chips per symbol), the outputs of the correlators are coherently combined, and a symbol detection is made. Figure 15.26 illustrates the

15.8

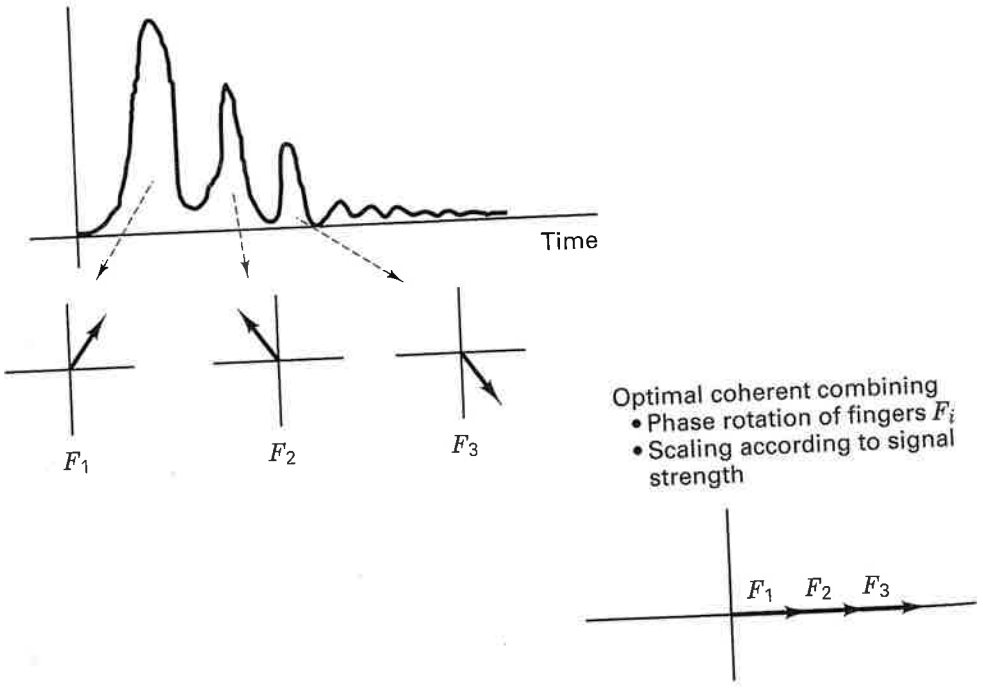


Figure 15.26 Coherent combining of multipath returns in the rake receiver.

phase rotation of fingers (F_i) provided by the Rake receiver, in order to facilitate coherent combining. At the chip level, the Rake receiver resembles an equalizer, but its real function is to exploit the path diversity.

The interference-suppression capability of DS/SS systems stems from the fact that a code sequence arriving at the receiver time-shifted by merely one chip will have very low correlation to the particular PN code with which the sequence is correlated. Therefore, any code chips that are delayed by one or more chip times will be suppressed by the correlator. The delayed chips only contribute to raising the interference level (correlation sidelobes). The mitigation provided by the Rake receiver can be termed *path diversity*, since it allows the energy of chips that arrive via multiple paths to be combined coherently. Without the Rake receiver, this energy would be transparent and therefore lost to the DS/SS receiver. In Figure 15.25, looking vertically above point τ_3 it is clear that there is interchip interference due to different fingers arriving simultaneously. The spread-spectrum processing gain allows the system to endure such interference at the chip level. No other equalization is deemed necessary in IS-95.

15.8 CONCLUSION

In this chapter, the major elements that contribute to fading in certain communication channels have been characterized. Figure 15.1 was presented as a guide for characterizing the fading phenomena. Two types of fading, large scale and small scale, were described. Two manifestations of small-scale fading (signal dispersion and fading rapidity) were examined. Each examination involved two views, one in time and the other in frequency. Two degradation categories were defined for dispersion: frequency-selective fading and flat fading. Two degradation categories were defined for fading rapidity: fast and slow fading. The small-scale fading degradation categories were summarized in Figure 15.7. A mathematical model using correlation and power density functions was presented in Figure 15.8. This model yields a useful symmetry to help us visualize the Fourier transform and duality relationships that describe the fading phenomena. Further, mitigation techniques for ameliorating the effects of each degradation category were presented; these were summarized in Figure 15.18. Finally, mitigation methods that have been applied in two different systems, GSM and CDMA systems meeting IS-95, were described.

REFERENCES

1. Rappaport, T. S., *Wireless Communications,* Chapters 3 and 4, Prentice Hall, Upper Saddle River, New Jersey, 1996.
2. Greenwood, D. and Hanzo, L., "Characterization of Mobile Radio Channels," *Mobile Radio Communications,* edited by R. Steele, Chapter 2, Pentech Press, London, 1994.
3. Lee, W. C. Y., "Elements of Cellular Mobile Radio Systems," *IEEE Trans. on Vehicular Technology,* vol. V–35, no. 2, May 1986, pp. 48–56.

4. Okumura, Y., et. al., "Field Strength and its Variability in VHF and UHF Land Mobile Radio Service," *Review of the Elec. Comm. Lab,* vol. 16, nos. 9 & 10, 1968, pp. 825–73.
5. Hata, M., "Empirical Formulae for Propagation Loss in Land Mobile Radio Services," *IEEE Trans. on Vehicular Technology,* vol. VT–29, no. 3, 1980, pp. 317–25.
6. Seidel, S. Y., et. al., "Path Loss, Scattering and Multipath Delay Statistics in Four European Cities for Digital Cellular and Microcellular Radiotelephone," *IEEE Transactions on Vehicular Technology,* Vol. 40, No. 4, November 1991, pp. 721–30.
7. Cox, D. C., Murray, R., and Norris, A., "800 MHz Attenuation Measured in and around Suburban Houses," *AT&T Bell Laboratory Technical Journal,* Vol. 673, No. 6, July–August 1984, pp. 921–54.
8. Schilling, D. L., et al., "Broadband CDMA for Personal Communications Systems," *IEEE Communications Magazine,* vol. 29, no. 11, November 1991, pp. 86–93.
9. Andersen, J. B., Rappaport, T. S., Yoshida, S., "Propagation Measurements and Models for Wireless Communications Channels, *IEEE Communications Magazine,* vol. 33, no. 1, Jan. 1995, pp. 42–49.
10. Proakis, J. G., *Digital Communications,* Chapter 7, McGraw–Hill Book Company, New York, 1983.
11. Schwartz, M., *Information, Transmission, Modulation, and Noise,* Second Edition, McGraw Hill, New York, 1970.
12. Amoroso, F., "Investigation of Signal Variance, Bit Error Rates and Pulse Dispersion for DSPN Signalling in a Mobile Dense Scatterer Ray Tracing Model," *Int'l Journal of Satellite Communications,* vol. 12, 1994, pp. 579–88.
13. Bello, P. A., "Characterization of Randomly Time-Variant Linear Channels," *IEEE Trans. on Commun. Syst.,* Dec. 1963, pp. 360–93.
14. Green, P. E., Jr., "Radar Astronomy Measurement Techniques," *MIT Lincoln Laboratory,* Lexington, Mass., Tech. Report No. 282, December 1962.
15. Pahlavan, K., and Levesque, A. H., *Wireless Information Networks,* Chapters 3 and 4, John Wiley and Sons, New York, 1995.
16. Lee, W. Y. C., *Mobile Cellular Communications,* McGraw-Hill Book Co., New York, 1989.
17. Amoroso, F. "Use of DS/SS Signaling to Mitigate Rayleigh Fading in a Dense Scatterer Environment," *IEEE Personal Communications,* vol. 3, no. 2, April 1996, pp. 52–61.
18. Clarke, R. H., "A Statistical Theory of Mobile Radio Reception," *Bell System Technical J.,* vol. 47, no. 6, July–August 1968, pp. 957–1000.
19. Bogusch, R. L., *Digital Communications in Fading Channels: Modulation and Coding,* Mission Research Corp., Santa Barbara, California, Report no. MRC-R-1043, March 11, 1987.
20. Amoroso, F., "The Bandwidth of Digital Data Signals," *IEEE Communications Magazine,* vol. 18, no. 6, November 1980, pp. 13–24.
21. Bogusch, R. L., et al., "Frequency Selective Propagation Effects on Spread-Spectrum Receiver Tracking," *Proceedings of the IEEE,* vol. 69, no. 7, July 1981, pp. 787–96.
22. Jakes, W. C. (Ed.), *Microwave Mobile Communications,* John Wiley & Sons, New York, 1974.
23. *Joint Technical Committee of Committee T1 R1P1.4 and TIA TR46.33/TR45.4.4 on Wireless Access,* "Draft Final Report on RF Channel Characterization," Paper No. JTC(AIR)/94.01.17–238R4, January 17, 1994.
24. Bello, P. A. and Nelin, B. D., "The Influence of Fading Spectrum on the Binary Error Probabilities of Incoherent and Differentially Coherent Matched Filter Receivers," *IRE Transactions on Commun. Syst.,* vol. CS-10, June 1962, pp. 160–68.

25. Amoroso, F., "Instantaneous Frequently Effects in a Doppler Scattering Environment," *IEEE International Conference on Communications,* June 7–10, 1987, pp. 1458–66.
26. Fung, V., Rappaport, T. S., and Thoma, B., "Bit-Error Simulation for π/4 DQPSK Mobile Radio Communication Using Two-Ray and Measurement-Based Impulse Response Models," *IEEE J. Sel. Areas Commun.,* vol. 11, no. 3, April 1993, pp. 393–94.
27. Bateman, A. J. and McGeehan, J. P., "Data Transmission over UHF Fading Mobile Radio Channels," *IEEE Proceedings,* vol. 131, Pt. F, No. 4, July 1984, pp. 364–74.
28. Feher, K. *Wireless Digital Communications,* Prentice Hall, Upper Saddle River, New Jersey, 1995.
29. Davarian, F., Simon, M., And Sumida, J., "DMSK: A Practical 2400–bps Receiver for the Mobile Satellite Service," *Jet Propulsion Laboratory* Publication 85–51 (MSAT–X Report No. 111), June 15, 1985.
30. Rappaport, T. S., *Wireless Communications,* Chapter 6, Prentice Hall, Upper Saddle River, New Jersey, 1996.
31. Bogusch, R. L., Guigliano, F. W., and Knepp, D. L., "Frequency-Selective Scintillation Effects and Decision Feedback Equalization in High Data-Rate Satellite Links," *Proceedings of the IEEE,* vol. 71, no. 6, June 1983, pp. 754–67.
32. Qureshi, S. U. H., "Adaptive Equalization," *Proceedings of the IEEE,* vol. 73, no. 9, September 1985, pp. 1340–87.
33. Forney, G. D., "The Viterbi Algorithm," *Proceedings of the IEEE,* vol. 61, no. 3, March 1978, pp. 268–78.
34. Viterbi, A. J. and Omura, J. K., *Principles of Digital Communication and Coding,* McGraw-Hill, New York, 1979.
35. Price, R. and Green P. E. Jr., "A Communication Technique for Multipath Channels," *Proceedings of the IRE,* March 1958, pp. 555–70.
36. Turin, G. L., "Introduction to Spread-Spectrum Antimultipath Techniques and their Application to Urban Digital Radio," *Proceedings of the IEEE,* vol. 68, no. 3, March 1980, pp. 328–53.
37. Simon, M. K., Omura, J. K., Scholtz, R. A., and Levitt, B. K., *Spread Spectrum Communications Handbook,* McGraw Hill Book Co., 1994.
38. Birchler, M. A. and Jasper, S. C., "A 64 kbps Digital Land Mobile Radio System Employing M-16QAM," *Proceedings of the 1992 IEEE Int'l. Conference on Selected Topics in Wireless Communications,* Vancouver, British Columbia, June 25–26, 1992, pp. 158–62.
39. Sari, H., Karam, G., and Jeanclaude, I., "Transmission Techniques for Digital Terrestrial TV Broadcasting," *IEEE Communications Magazine,* vol. 33, no. 2, February 1995, pp. 100–109.
40. Cavers, J. K., The Performance of Phase Locked Transparent Tone-in-Band with Symmetric Phase Detection, *IEEE Trans. on Commun.,* vol. 39, no. 9, Sept. 1991, pp. 1389–99.
41. Moher, M. L. and Lodge, J. H., "TCMP—A Modulation and Coding Strategy for Rician Fading Channel," *IEEE Journal on Selected Areas in Communications,* vol. 7, no. 9, December 1989, pp. 1347–55.
42. Harris, F., "On the Relationship Between Multirate Polyphase FIR Filters and Windowed, Overlapped FFT Processing," *Proceedings of the Twenty Third Annual Asilomar Conference on Signals, Systems, and Computers,* Pacific Grove, California, Oct. 30 to Nov. 1, 1989, pp. 485–88.
43. Lowdermilk, R. W., and Harris, F., "Design and Performance of Fading Insensitive Orthogonal Frequency Division Multiplexing (OFDM) using Polyphase Filtering

Techniques," *Proceedings of the Thirtieth Annual Asilomar Conference on Signals, Systems, and Computers,* Pacific Grove, California, November 3–6, 1996.

44. Kavehrad, M. and Bodeep, G. E., "Design and Experimental Results for a Direct-Sequence Spread-Spectrum Radio Using Differential Phase-Shift Keying Modulation for Indoor Wireless Communications," *IEEE JSAC* vol. SAC-5, no. 5, June 1987, pp. 815–23.
45. Hess, G. C., *Land-Mobile Radio System Engineering,* Artech House, Boston, 1993.
46. Hagenauer, J., and Lutz, E., "Forward Error Correction Coding for Fading Compensation in Mobile Satellite Channels," *IEEE JSAC,* vol. SAC–5, no. 2, February 1987, pp. 215–25.
47. McLane, P. I., et al., "PSK and DPSK Trellis Codes for Fast Fading, Shadowed Mobile Satellite Communication Channels," *IEEE Trans. on Comm.,* vol. 36, no. 11, November 1988, pp. 1242–46.
48. Schlegel, C., and Costello, D. J., Jr., "Bandwidth Efficient Coding for Fading Channels: Code Construction and Performance Analysis," *IEEE JSAC,* vol. 7, no. 9, December 1989, pp. 1356–68.
49. Edbauer, F., "Performance of Interleaved Trellis-Coded Differential 8-PSK Modulation over Fading Channels," *IEEE J. on Selected Areas in Comm.,* vol. 7, no. 9, December 1989, pp. 1340–46.
50. Soliman, S., and Mokrani, K., "Performance of Coded Systems over Fading Dispersive Channels," *IEEE Trans. on Communications,* vol. 40, no. 1, January 1992, pp. 51–59.
51. Divsalar, D. and Pollara, F. "Turbo Codes for PCS Applications," *Proc. ICC '95,* Seattle, Washington, June 18–22, 1995, pp. 54–59.
52. Simon, M and Alouini, M-S., *Digital Communications over Fading Channels: A Unified Approach to Performance Analysis,* John Wiley, New York, 2000.
53. Padovani, R., "Reverse Link Performance of IS-95 Based Cellular Systems," *IEEE Personal Communications,* Third Quarter 1994, pp. 28–34.
54. Hanzo, L. and Stefanov, J., "The Pan-European Digital Cellular Mobile Radio System—Known as GSM," *Mobile Radio Communications,* edited by R. Steele, Chapter 8, Pentech Press, London, 1992.

PROBLEMS

15.1. The probability density function for a Rayleigh continuous random variable, is given in Equation (15.15).

(a) Find an expression for the *distribution function,* as described in Section 1.5.1.

(b) For a signal transmitted over a mobile wireless channel, experiencing Rayleigh fading, use the distribution function to estimate the percentage of time that the signal level is 15 dB below the rms value.

(c) Repeat part b) for a signal level that is 5 dB below the rms value.

15.2. A signal within a mobile wireless system undergoes time spreading. The symbol rate $R_s = 20$ ksymbols/sec. Channel measurements indicate that the mean excess delay is 10 μs, while the second moment of the excess delay is 1.8×10^{-10} s^2.

(a) Calculate the coherence bandwidth, f_0, if it is defined as the frequency interval over which the channel's complex frequency transfer function has a correlation of at least 0.9.

(b) Repeat part a) if f_0 is defined for a bandwidth interval having a correlation of at least 0.5.

(c) Determine whether the signaling will undergo frequency-selective fading.

15.3. Consider a channel whose power-density profile is made up of three impulse functions with relative power and time-delay locations as follows: –20-dB at 0 μs, 0-dB at 2 μs, and –10-dB at 3 μs.

(a) Calculate the mean excess delay.

(b) Calculate the second moment of the excess delay.

(c) Calculate the rms delay spread.

(d) Estimate the coherence bandwidth (corresponding to a correlation ≥ 0.9).

(e) If the receiver is located in an airplane traveling at 800 km/hr, and the time required to traverse a half wavelength is 100 μs, calculate an approximate value for the transmission frequency.

15.4. Given a mobile wireless system with a carrier frequency of f_c = 900 MHz, and a Doppler frequency of f_d = 50 Hz. Assuming that the dense-scatterer model applies,

(a) Plot the *Doppler power spectral density S(v)* in the range of $f_c \pm f_d$ (use about 10 points).

(b) Explain why $S(v)$ has the response it does at the boundaries.

(c) Calculate the coherence time T_0, assuming that the channel's response to sinusoids yields a correlation of at least 0.5.

15.5. For each of the fading-effect categories below, name an application that generally fits that category. Provide numerical justification.

(a) frequency-selective, fast-fading

(b) frequency-selective, slow-fading

(c) flat-fading, fast-fading

(d) flat-fading, slow-fading

15.6. **(a)** What is the relationship between the signal power-density profile, as measured by the rms delay σ_τ, and the Doppler power-spectral density, as measured by the fading bandwidth f_d?

(b) What is the relationship between the spaced-frequency correlation function, as measured by the coherence bandwidth f_0, and the spaced-time correlation function, as measured by the coherence time T_0?

15.7. Consider an indoor narrowband mobile communication system that is characterized by a power density profile made up of four impulse functions with relative power and time-delay locations as follows: 0-dB at 0 ns, –3-dB at 100 ns, –3-dB at 200 ns, and –6-dB at 300 ns. What is the maximum symbol rate that such a system can support without using an equalizer? To find coherence bandwidth, use the criterion that spaced frequency tones are correlated to at least 0.5.

15.8. Consider a mobile wireless system, using QPSK modulation at 24.3 ksymbols/s, and a carrier frequency of 1900 MHz. What is the fastest speed in km/hr that is permissible for a vehicle using such a system, if it is required that the change in phase, $\Delta\theta$, due to spectral broadening (Doppler spread) does not exceed 5°/symbol?

15.9. For an interleaver to provide meaningful time diversity, a rule of thumb requires that the span of the interleaver, T_{IL}, should be at least ten times greater than the channel coherence time, T_0. Show a graph of T_{IL} versus frequency. (Plot three frequency points: 300 MHz, 3 GHz, and 30 GHz) for the following mobile telephone users.

(a) A pedestrian walking at a speed of 1 m/s.

(b) A high-speed train with a speed of 50 m/s.

(c) If the telephone handles real-time speech, which of the six plotted points represent cases where meaningful time diversity can be achieved by using an interleaver span of ten times T_0?

(d) What general conclusions can you draw?

15.10. A transmitted signal having a bandwidth of 5 kHz, and propagating over a channel having a coherence bandwidth of 50 kHz, is clearly an example of a flat-fading channel. Explain how such a channel can occasionally experience frequency-selective fading.

15.11. Consider a TDMA mobile wireless system, having a carrier frequency of 1900 MHz, that operates on a high-speed train at a speed of 180 km/hr. For learning the channel's impulse response in order to provide equalization, each user's transmission consists of training bits in addition to data bits. It is required that the training sequence should consist of 20 bits, should not occupy more than 20% of the total bits, and that these training bits should be embedded in the data, at least every $T_0/4$ s. Assuming binary modulation, what is the slowest transmission rate that would support these requirements without fast fading?

15.12. (a) In the late 1990's in Japan, a digital portable telephone system, called the Portable Handyphone System (PHS) was specified. The specification for PHS includes a carrier spacing of 300 kHz. Is the standard susceptible to frequency-selective fading in environments for which the channel has an rms delay spread on the order of 300 ns?

(b) A telephone standard called Digital Enhanced Cordless Telecommunications (DECT) was designed for high traffic density and short range (indoor) communications. The specification for DECT includes a carrier spacing of 1.728 MHz. Consider that the rms delay spread equals 150 ns. Calculate whether an equalizer needs to be included within the DECT receiver.

15.13. An interleaver span should be at least 10 times greater than the channel coherence time in order to provide significant time diversity for a mobile wireless system. Consider using such an interleaver in the design of a 1 GHz mobile system for pedestrians walking at a speed of 0.5 m/s. How large would the interleaver span have to be? Is this feasible for a real-time voice system?

15.14. If we desire to keep the transmitter-plus-receiver interleaver delay below 100 ms, what is the largest ratio of interleaver span to channel coherence time, T_{IL}/T_0, that can be used, for the following cases?

(a) The channel has a fading rate of 100 Hz.

(b) The channel has a fading rate of 1000 Hz.

15.15. A mobile wireless system is configured to support a data rate of 200 kbits/s using QPSK modulation and a carrier frequency of 1900 MHz. It is intended for use in a vehicle that typically travels at a speed of 96 km/hr.

(a) What change in phase angle value, $\Delta\theta$ per symbol, can be expected?

(b) What is the value of $\Delta\theta$ per symbol, if the data rate is decreased to 100 kbits/s?

(c) Repeat part (b) for a speed of 48 km/hr.

(d) Draw some general conclusions.

15.16. A channel exhibiting multipath fading has an rms delay spread of $\sigma_\tau = 10$ μs, and a Doppler spread of $f_d = 1$ Hz. The baseband pulse duration is chosen to be $T_s = 1$ μs.

(a) What is the value of the channel coherence bandwidth?

(b) What is the value of the channel coherence time?

(c) How would you classify the channel with regard to frequency selectivity and fading rapidity.

(d) How would you change the pulse duration (or data rate) to help mitigate the effects of fading?

15.17. For a mobile wireless application, a phase-based modulation scheme is particularly susceptible to phase distortion. Such distortion can be avoided if the rate of signaling is at least 100 times greater than the fading rate [27]. Consider a radio system operating at a carrier frequency of 1900 MHz, and moving at a velocity of 96 km/hr. In order to avoid degradation due to fast fading, what is the slowest symbol rate that should be sent over such a system?

15.18. Consider a mobile communication system which has a framing and time-slot arrangement as shown below in Figure P15.1.

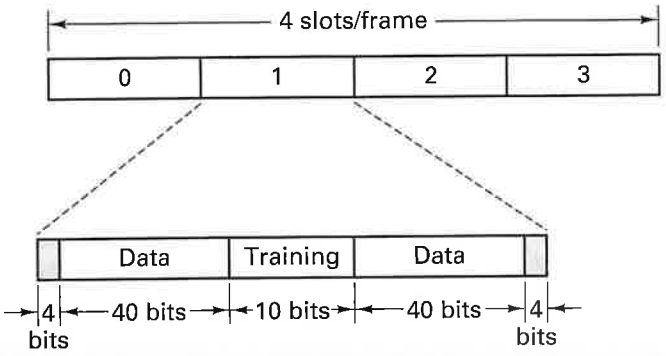


Figure P15.1 TDMA frame format.

There are a total of 4 time slots per frame; each user is assigned a single time slot per frame. Each time slot consists of 98 bits as shown in Figure P15.1. The signaling consists of QPSK modulation with a carrier frequency of 700 MHz. The symbol rate is 33.6 ksymbols/s, and the transmission bandwidth is 47 kHz. The system must perform normally at velocities up to 100 km/hr. Measurements performed on the physical channel conclude that the typical rms delay spread is in the order of 4 μs.

(a) Assuming that the training sequence yields an estimate of the channel impulse response during each slot time, will this design suffer the degradation effects of fast fading?

(b) Will this design suffer the degradation effects of frequency-selective fading?

15.19. The total permissible delay for data transmitted over a particular mobile wireless channel is limited to 340 ms. The data rate is equal to 19.2 ksymbols/s, and the data is interleaved to provide time diversity. The system exhibits delays as shown in Table P15.1.

TABLE P15.1 Delay values in ms

Delay T	Value (in ms)
Encoder	2
Modulator	10
Channel	0.3
Demodulator	25
Decoder	$2 \times 10^8 / f_{clk}$

The delay in milliseconds for the decoder, is given by $2 \times 10^8 / f_{clk}$, where f_{clk} is the decoder clock speed. Calculate the minimum decoder clock speed that would be required for the following interleaver sizes.

(a) 100 bits
(b) 1000 bits
(c) 2850 bits
(d) What can you conclude about the decoder speed as the interleaver size increases?

15.20. Consider a mobile wireless orthogonal FDM (OFDM) system that is intended to operate in a vehicle with speeds up to 80 km/hr in an urban environment having a coherence bandwidth of 100 kHz. The carrier frequency is 3 GHz, and it is required that data be transmitted at a symbol rate of 1024 ksymbols/s. Choose an appropriate subcarrier plan, with the following goals: 1) avoid the use of an equalizer, and 2) minimize any effects due to fast fading. The plan should address how many subcarriers should be used, how far apart in frequency they should be separated, and what value of symbol rate per subcarrier should be used.

15.21. A mobile wireless system uses direct-sequence spread-spectrum (DS/SS) signaling to mitigate the effects of a received signal having two components—a direct path and a reflected path. The reflected path is 120 m longer than the direct path.What should the chip rate be, in order for such a system to provide mitigation against the effects of multipath?

15.22. It is well known that direct-sequence spread-spectrum (DS/SS) signaling can be used as a mitigation technique to protect against the channel-induced ISI of a frequency-selective channel. Yet, when you examine Figure 15.25 at some instant of time, say time τ_3, there exists interchip interference. Doesn't there need to be further equalization techniques used, in order to overcome this chip-level interference? Explain.

15.23. CDMA and TDMA are unique in the sense that each of these multiple-access schemes is associated with a property for mitigating one of the fading degradation types. Against which type of degradation does each of these schemes provide a "natural" protection.

15.24. Consider a diversity scheme consisting of four branches as shown in Figure P15.2. Each branch is responsible for processing an independent Rayleigh faded signal, $r(t)$. At a particular instant in time, the received signal can be expressed as a four branch vector $\mathbf{r} = [r_1\ r_2\ r_3\ r_4]$, where r_i represents the voltage signal at branch i. Also, the gain

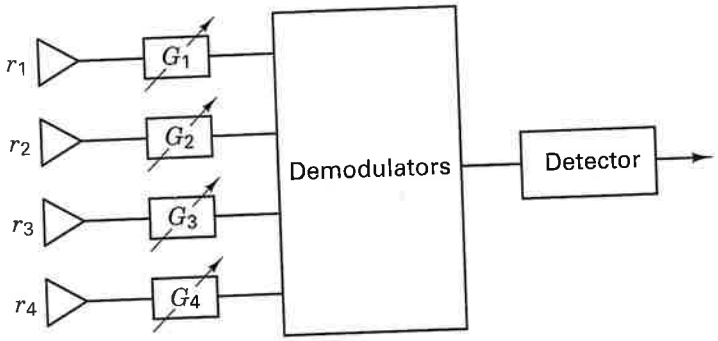


Figure P15.2 Four branch diversity receiver.

in the branches can be expressed with a four-branch vector $\mathbf{G} = [G_1\ G_2\ G_3\ G_4]$, where G_i represents the voltage gain at branch i. Consider a moment in time when the measured values of $\mathbf{r}$ equal [0.87, 1.21, 0.66, 1.90], and the associated gains of $\mathbf{G}$ equal [0.5, 0.8, 1.0, 0.8]. The average noise power of each branch N is equal to 0.25.

(a) Calculate the SNR applied to the detector?

(b) It can be shown [1] that the SNR is maximized when the value of each G_i is equal to r_i^2/N. Based on this condition, determine the maximum achievable SNR.

15.25. A system makes use of branch diversity in order to improve the receiver SNR. It is assumed that each branch receives an independently Rayleigh-faded signal. The receiver must meet a requirement that the probability of all branches being received with an SNR less than some threshold is equal to 10^{-4}, where the threshold is chosen to be 5 dB, and the average SNR is equal to 15 dB.

(a) Calculate the number of diversity branches M required in the receiver to meet this requirement?

(b) Based on the result from part a), calculate the probability that any single branch achieves an SNR > 5 dB?

15.26. A diversity scheme used in a receiver consists of two single branches. The following is received from each branch:

$$\begin{bmatrix}\text{Branch 1}\\ \text{Branch 2}\end{bmatrix} = \begin{bmatrix}1.85 & 1.91 & -1.31 & -1.58 & 1.21 & 1.93 & 1.11 & -1.67 & 2.13 & -2.25\\ 1.67 & 1.69 & -2.13 & -1.26 & 1.74 & 1.76 & 1.29 & -1.93 & 2.31 & -1.08\end{bmatrix}$$

The first row shows the voltage samples received from the first branch and the second row shows the voltage samples received from the second branch. Each column corresponds to a particular instant in time. Consider that the average noise power in each branch is equal to 0.25 W and assume that the above values have been cophased for use in both maximum ratio and equal gain combining. The instantaneous voltage gains provided by an attenuator for branch 1 and branch 2 are $G_1 = 1.2$ and $G_2 = 1.4$ respectively. Also with feedback diversity assume that the threshold has been set to an SNR of 5 dB. Calculate which branch is output to the detector if the following diversity-combining techniques are used:

(a) Selection.

(b) Feedback.

Calculate the value of SNR that is outputted to the detector when the following diversity combining schemes are used.

(c) Maximum Ratio.

(d) Equal Gain.

15.27. The response of a channel to ideal positive and negative pulses are spread over three pulse-time durations, as shown in Figure P15.3. Thus, for a sequence of transmitted pulses, the received waveform consists of the superposition of $L = 3$ events (segments from three pulses)—the current pulse, plus the memory of two past pulses. Use an *encoding trellis diagram* to represent this channel-induced ISI, and annotate each trellis branch with a voltage value resulting from that transition. Imagine that the system is initially flushed into the 00 state by transmitting two negative-polarity pulses. Then, consider the transmission of the sequence 11011 using the ideal pulses in Figure P15.3. Determine the amplitude values of the received distorted signal, and show its path along the trellis diagram. Hint: This binary finite-state machine has 2^{L-1} states. Use quadrille paper to perform the superposition needed to represent the distorted waveforms characterizing the channel. The trellis diagram is formed as

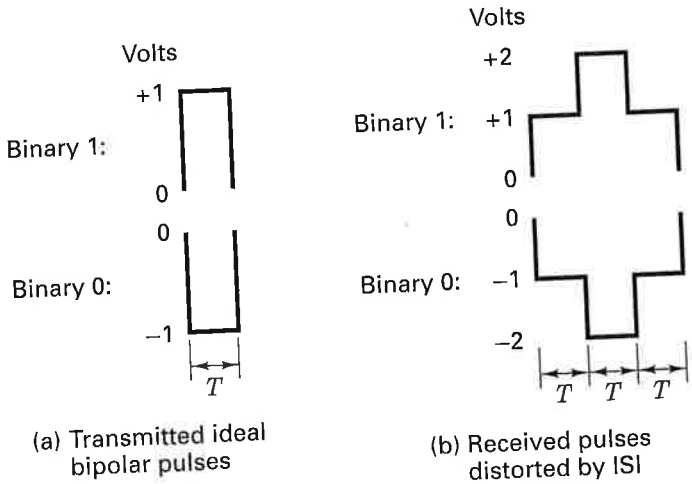


Figure P15.3

in Section 7.2.3, with the exception that voltage levels are used here instead of code bits.

15.28. Use the channel characteristics and training sequence outlined in Problem 15.27, and add noise voltages of {+1 −1 +1 −1 +1} to the received distorted waveform. Use a *decoding trellis diagram* to illustrate how the Viterbi decoding algorithm is used in this equalization process, and show the computations that yield the first message bit. Hint: The process resembles the decoding of convolutionally encoded bits, with the exception that voltage levels are used here instead of code bits.

15.29. A mobile communication system uses a Viterbi equalizer in order to combat the effects of fading. The symbol rate is equal to 160 ksymbols/s, and BPSK modulation is used. The signal dispersion resulting from channel-induced ISI is equal to 25 μs.

(a) Calculate the approximate amount of memory, L_o, in units of bit intervals required to be included in the Viterbi equalizer.

(b) How is the memory size affected if the symbol rate is doubled?

QUESTIONS

15.1. What are the two fading mechanisms that characterize small-scale fading? Explain the *Fourier transform* and *duality* relationships among the time and frequency descriptions of these mechanisms. (See Sections 15.2 through 15.4.)

15.2. What are the differences between *Rician* fading and *Rayleigh* fading? (See Section 15.2.2.)

15.3. Define the following parameters: rms delay spread, coherence bandwidth, coherence time, Doppler spread. How are they related? (See Sections 15.3 and 15.4.)

15.4. What are the two *degradation categories* that characterize signal time-spreading, and what are the two degradation categories that characterize the time-variant nature of the channel? (See Sections 15.3 and 15.4.)

15.5. Why are the two basic fading mechanisms that characterize small-scale fading considered to be *independent* of one another? (See Section 15.4.1.1.)

15.6. Why does *signal distortion* caused by fading cause more serious degradation effects compared with a *loss* in SNR? (See Section 15.5.)

15.7. What techniques can be used for mitigating the effects of frequency-selective fading? What techniques can be used for mitigating the effects of fast fading? (See Section 15.5.)

15.8. What are the various ways of introducing signal *diversity*? (See Section 15.5.3.)

15.9. If there is *no motion* between a transmitter and receiver, how large would the interleaver span need to be in order to provide any time-diversity benefits? (See Section 15.5.6.)

EXERCISES

Using the Companion CD, run the exercises associated with Chapter 15.

APPENDIX A

A Review of Fourier Techniques

A.1 SIGNALS, SPECTRA, AND LINEAR SYSTEMS

Electrical communication signals consist of time-varying voltage or current waveforms, typically described in the time domain. It is also convenient to describe such signals in the frequency domain. A signal's frequency-domain description is called its *spectrum.* Spectral concepts are important in communication analysis and design; they can describe a signal by its average power or energy content at various frequencies, and they illustrate how much of the electromagnetic spectrum (bandwidth) the signal occupies. Broadcast stations are required by the Federal Communications Commission (FCC) to operate at their assigned frequency with very tight tolerances on the occupied bandwidth; for example, amplitude-modulated (AM) radio channels are spaced 10 kHz apart, and television channels are spaced 6 MHz apart. Our interest in spectra and Fourier techniques has to do with the real-world constraints of ensuring that our communication signals are confined to specified spectral boundaries.

Frequency spectral characteristics can be ascribed to both signal waveforms and to circuits. When we say that a particular spectrum describes a *signal,* we mean that one way of characterizing the signal waveform is to specify its amplitude and phase as a function of frequency. However, when we talk about the spectral attributes of a *circuit* we are referring to the output versus input frequency-domain transfer function of the circuit; in other words, we are characterizing the circuit by how much of a specific input signal spectrum is allowed to pass through it.

A.2 FOURIER TECHNIQUES FOR LINEAR SYSTEM ANALYSIS

Fourier techniques are often used for analyzing linear circuits or systems in the following ways: (1) by predicting the system response, (2) by determining the system dynamic specification (transfer function), and (3) by evaluating or interpreting test

results. Item 1, predicting system response, is illustrated schematically in Figure A.1. Let the input be an arbitrary periodic waveform with period equal to T_0 seconds. Fourier techniques allow us to describe such an input as a sum of sinusoidal waveforms, as shown in the figure. The lowest-frequency sinusoid, or the *fundamental* frequency of the input periodic, has frequency $1/T_0$ hertz; the balance of the sinusoids have frequencies that are integral *harmonics* ($2/T_0$, $3/T_0$, . . .) of this fundamental frequency. An important attribute of a linear system is that *superposition* applies, which means that the response to the sum of excitations is the sum of the responses to the individually applied excitations. In fact, this is used as a definition of linearity. Specifically, if

$$y_1(t) = \text{system response to } x_1(t)$$

$$y_2(t) = \text{system response to } x_2(t)$$

and

$$ay_1(t) + by_2(t) = \text{system response to } ax_1(t) + bx_2(t)$$

for all a, b, $x_1(t)$, and $x_2(t)$, then the system is linear. A consequence of this definition is that the output response of a *linear* system with sinusoidal input waveforms must be made up of sinusoidal waveforms having the *same frequencies* as the input waveforms; such a system is typically specified by an output versus input *frequency transfer function* (magnitude and phase versus frequency) as shown in Figure A.2. Figure A.2a illustrates a typical example of signal magnitude versus

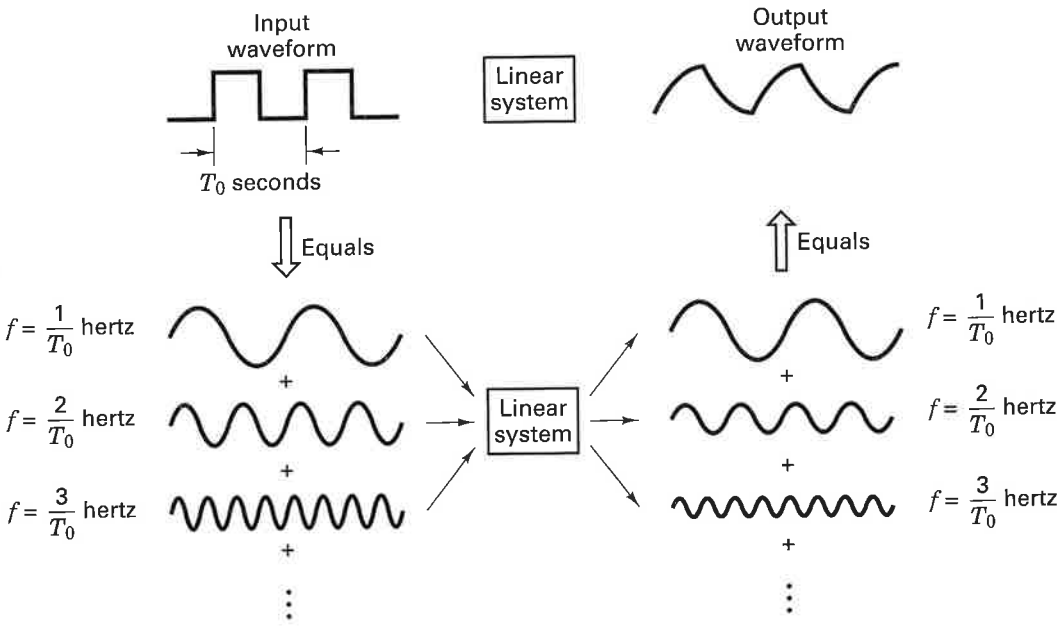


Figure A.1 Predicting system response.

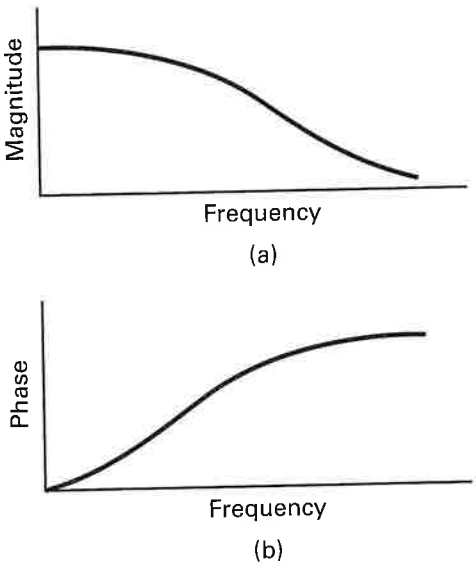


Figure A.2 System transfer function. (a) Magnitude response. (b) Phase response.

frequency; similarly, Figure A.2b illustrates a typical example of signal phase versus frequency.

The system transfer function serves as a performance specification; it describes the system response to each of the component sinusoids. Therefore, with the system transfer function in hand, one can predict each of the resulting output components. Using the principle of superposition, the final step of the analysis is to sum the individual output responses, thus forming the resulting overall response to the input periodic. (See Figure A.1.) In a similar manner, one can determine a system's transfer function, or evaluate a system's test results from knowledge of the input and output waveforms.

The development of Fourier methods had a major impact on the analysis of linear systems; it provided the translation between transient phenomena and sinusoidal techniques, and it simplified the analysis of linear systems under the excitation of any arbitrary input waveform. Just as logarithms allow the operation of multiplication to be treated as addition, so Fourier techniques allow the replacement of complex waveforms with sinusoidal components and sinusoidal methods.

A.2.1 Fourier Series Transform

Signals that are periodic with finite energy within each period can be represented by the *Fourier series.* The following equation describes such an arbitrary periodic waveform $x(\lambda)$ in terms of an infinite number of increasing harmonic sine and cosine components:

$$x(\lambda) = \tfrac{1}{2}a_0 + a_1 \cos\lambda + a_2 \cos 2\lambda + a_3 \cos 3\lambda + \cdots + b_1 \sin\lambda + b_2 \sin 2\lambda + b_3 \sin 3\lambda + \cdots \quad \text{(A.1)}$$

The terms cos λ and sin λ are called the *fundamental terms*; the terms cos $n\lambda$ and sin $n\lambda$, for $n > 1$, are called *harmonic terms,* where n is an integer. The terms a_n and b_n represent the coefficients of the fundamental and harmonics, and $\frac{1}{2}a_0$ is the constant or dc term.

The function $x(\lambda)$ must have a period of 2π, or a submultiple thereof, and it must be single valued. The Fourier series can be thought of as a "recipe" for synthesizing any *arbitrary periodic* waveform using sinusoidal components. To be useful, the series must converge; that is, the sum of the series, as more and more of the higher harmonics are added, must approach a limit.

The process of synthesizing an arbitrary periodic waveform, from the coefficient values describing the mix of harmonics, is termed *synthesis.* The inverse process of calculating the coefficient values is termed *analysis.* Calculation of the coefficients is facilitated by the fact that the average of the sine and cosine cross-products is zero, as well as the average of any sinusoid. The following equations illustrate the basic averaging properties of the sine, cosine, and their products and cross-products:

$$\left.\begin{aligned} \int_{-\pi}^{\pi} \sin m\lambda \, d\lambda &= 0 \\ \int_{-\pi}^{\pi} \cos m\lambda \, d\lambda &= 0 \\ \int_{-\pi}^{\pi} \sin m\lambda \cos n\lambda \, d\lambda &= 0 \end{aligned}\right\} \text{where } m \text{ and } n \text{ are any integers} \tag{A.2}$$

$$\left.\begin{aligned} \int_{-\pi}^{\pi} \sin m\lambda \sin n\lambda \, d\lambda &= 0 \\ \int_{-\pi}^{\pi} \cos m\lambda \cos n\lambda \, d\lambda &= 0 \end{aligned}\right\} \text{for } m \neq n \tag{A.3}$$

$$\left.\begin{aligned} \int_{-\pi}^{\pi} (\sin m\lambda)^2 \, d\lambda &= \pi \\ \int_{-\pi}^{\pi} (\cos m\lambda)^2 \, d\lambda &= \pi \end{aligned}\right\} \text{for } m = n \tag{A.4}$$

Consider how one could go about finding the value of the coefficient, a_n or b_n, in Equation (A.1). To find the coefficient a_3, for example, we can multiply both sides of Equation (A.1) by cos $3\lambda \, d\lambda$ and integrate, as follows:

$$\int_{-\pi}^{\pi} x(\lambda)\cos 3\lambda\, d\lambda = \int_{-\pi}^{\pi} \tfrac{1}{2}a_0 \cancelto{0}{\cos 3\lambda\, d\lambda} + \int_{-\pi}^{\pi} a_1 \cancelto{0}{\cos\lambda \cos 3\lambda\, d\lambda}$$
$$+ \int_{-\pi}^{\pi} a_2 \cancelto{0}{\cos 2\lambda \cos 3\lambda\, d\lambda} + \int_{-\pi}^{\pi} a_3 (\cos 3\lambda)^2\, d\lambda + \cdots$$
$$+ \int_{-\pi}^{\pi} b_1 \cancelto{0}{\sin\lambda \cos 3\lambda\, d\lambda} + \int_{-\pi}^{\pi} b_2 \cancelto{0}{\sin 2\lambda \cos 3\lambda\, d\lambda}$$
$$+ \int_{-\pi}^{\pi} b_3 \cancelto{0}{\sin 3\lambda \cos 3\lambda\, d\lambda} + \cdots$$

$$\int_{-\pi}^{\pi} x(\lambda)\cos 3\lambda\, d\lambda = \int_{-\pi}^{\pi} a_3 (\cos 3\lambda)^2\, d\lambda = a_3\pi$$

$$a_3 = \frac{1}{\pi}\int_{-\pi}^{\pi} x(\lambda)\cos 3\lambda\, d\lambda$$

We can generalize the preceding analysis to get

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(\lambda)\cos n\lambda\, d\lambda \tag{A.5}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(\lambda)\sin n\lambda\, d\lambda \tag{A.6}$$

a_0 is found by solving Equation (A.5) with $n = 0$. This results in

$$\tfrac{1}{2}a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(\lambda)\, d\lambda \tag{A.7}$$

which represents the zero-frequency term, or the average value of the periodic waveform. The synthesis process of Equation (A.1) can be expressed in more compact form as follows:

$$x(\lambda) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos n\lambda + b_n \sin n\lambda) \tag{A.8}$$

There are several ways to express the *transform pair* (analysis and synthesis) of the Fourier series. The most common form makes use of the following identities to express the sine and cosine in exponential form:

$$\cos\lambda = \frac{e^{j\lambda} + e^{-j\lambda}}{2} \tag{A.9}$$

$$\sin\lambda = \frac{e^{j\lambda} - e^{-j\lambda}}{2j} \tag{A.10}$$

A periodic function with period T_0 seconds has frequency components of f_0, $2f_0$, $3f_0$, . . . , where $f_0 = 1/T_0$ is called the *fundamental frequency.* We also refer to the frequency components as ω_0, $2\omega_0$, $3\omega_0$, . . . , where $\omega_0 = 2\pi/T_0$ is called the fundamental *radian* frequency. The terms f and ω are each used to denote frequency. When f is used, frequency in hertz is intended; when ω is used, frequency is radians/second is intended. Let us replace the $n\lambda$ terms of Equations (A.5) to (A.8) with $2\pi nf_0t = 2\pi nt/T_0$ as the general argument of the sinusoidal components, where n is an integer. For $n = 1$, nf_0 represents the fundamental frequency; for $n > 1$, nf_0 represents harmonics of the fundamental frequency. Using Equations (A.8) to (A.10), we can express $x(t)$ in exponential form as follows:

$$x(t) = \frac{a_0}{2} + \frac{1}{2}\sum_{n=1}^{\infty} [(a_n - jb_n)e^{j2\pi nf_0t} + (a_n + jb_n)e^{-j2\pi nf_0t}] \tag{A.11}$$

Let c_n denote the complex coefficients, or spectral components of $x(t)$, related to a_n and b_n by

$$c_n = \begin{cases} \frac{1}{2}(a_n - jb_n) & \text{for } n > 0 \\ \dfrac{a_0}{2} & \text{for } n = 0 \\ \frac{1}{2}(a_n + jb_n) & \text{for } n < 0 \end{cases} \tag{A.12}$$

Then we can simplify Equation (A.11), writing

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{j2\pi nf_0t} \tag{A.13}$$

where the coefficients of the exponential harmonics are

$$c_n = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)e^{-j2\pi nf_0t}\,dt \tag{A.14}$$

To verify Equation (A.14), we multiply both sides of Equation (A.13) by $e^{-j2\pi mf_0t}\,dt/T_0$, integrate over the interval $(-T_0/2, T_0/2)$, and use the relationship

$$\frac{1}{T_0}\int_{-T_0/2}^{T_0/2} e^{j(n-m)2\pi f_0t}\,dt = \delta_{nm} = \begin{cases} 1 & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases} \tag{A.15}$$

where δ_{nm} is known as the *Kronecker delta.* By multiplying and integrating in this way we obtain, for all integers m,

$$\frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)e^{-j2\pi mf_0t}\,dt = \sum_{n=-\infty}^{\infty} c_n\delta_{nm} = c_m \tag{A.16}$$

In general, the coefficient c_n is a complex number; it can be expressed in the form

$$c_n = |c_n|\,e^{j\theta_n} \tag{A.17}$$

$$c_{-n} = |c_n|e^{-j\theta_n} \tag{A.18}$$

where

$$|c_n| = \tfrac{1}{2}\sqrt{a_n^2 + b_n^2} \tag{A.19}$$

$$\theta_n = \tan^{-1} -\frac{b_n}{a_n} \tag{A.20}$$

$$b_0 = 0 \quad \text{and} \quad c_0 = \frac{a_0}{2}$$

The value of $|c_n|$ defines the magnitude of the nth harmonic component of the periodic waveform, so that a plot of $|c_n|$ versus frequency, called the *magnitude spectrum,* yields the magnitude of each of the n discrete harmonics in the signal. Similarly, a plot of θ_n versus frequency, called the *phase spectrum,* yields the phase of each harmonic component in the signal.

The Fourier coefficients of a real-valued periodic time function exhibit the relationship

$$c_{-n} = c_n^* \tag{A.21}$$

where c_n^* is the complex conjugate of c_n. We therefore have

$$|c_{-n}| = |c_n| \tag{A.22}$$

and the magnitude spectrum is an even function of frequency. Similarly, the phase spectrum θ_n is an odd function of frequency, because from Equation (A.20),

$$\theta_{-n} = -\theta_n \tag{A.23}$$

The Fourier series is particularly useful in characterizing arbitrary periodic waveforms, with finite energy in each period, as presented above. The Fourier series can also be used to characterize nonperiodic signals having finite energy over a finite interval. However, a more convenient frequency-domain representation for such signals uses the Fourier integral transform. (See Section A.2.3.)

A.2.2 Spectrum of a Pulse Train

A signal of great interest in digital communications is an ideal periodic sequence of rectangular pulses, called a *pulse train,* illustrated in Figure A.3. For the pulse train, $x_p(t)$, with pulse amplitude A, pulse width T, and period T_0, the reader can verify, using Equations (A.14) and (A.10), the following expression for the Fourier series coefficients:

$$c_n = \frac{AT}{T_0}\frac{\sin(\pi nT/T_0)}{\pi nT/T_0} = \frac{AT}{T_0}\operatorname{sinc}\frac{nT}{T_0} \tag{A.24}$$

In this expression,

$$\operatorname{sinc} y = \frac{\sin \pi y}{\pi y}$$

The sinc function, as shown in Figure A.4, has a maximum value of unity at $y = 0$ and approaches zero as y approaches infinity, oscillating through positive and

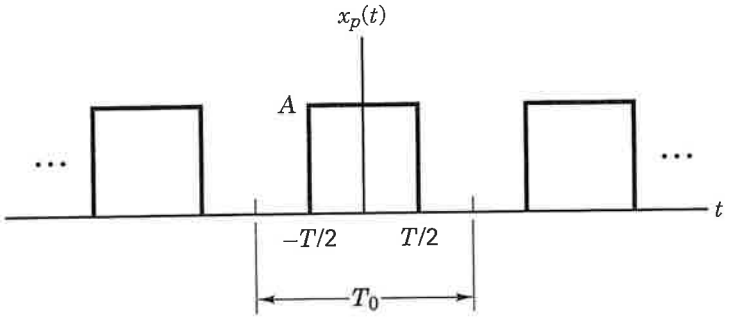


Figure A.3 Pulse train.

negative values. It goes through zero at $y = \pm 1, \pm 2, \ldots$ The pulse train magnitude spectrum, $|c_n|$ as a function of n/T_0, is plotted in Figure A.5a, and the phase spectrum, θ_n, is plotted in Figure A.5b. The positive and negative frequencies of the two-sided spectrum represent a useful way of expressing the spectrum mathematically; of course, only the positive frequencies can be reproduced in a laboratory.

Synthesis is performed by substituting the coefficients of Equation (A.24) into Equation (A.13). The resulting series yields the original ideal pulse train, $x_p(t)$, synthesized from its component parts:

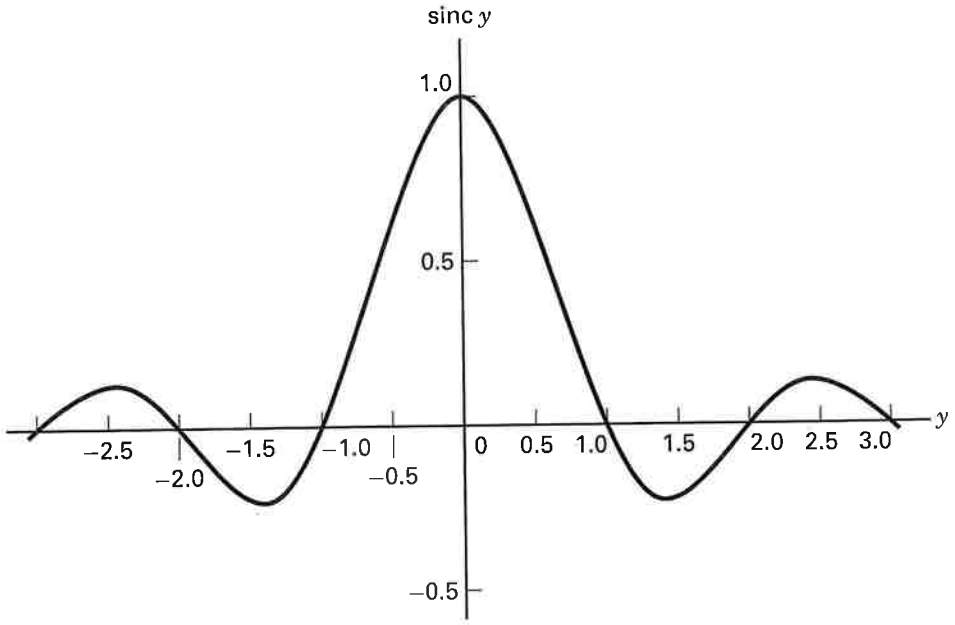


Figure A.4 Sinc function.

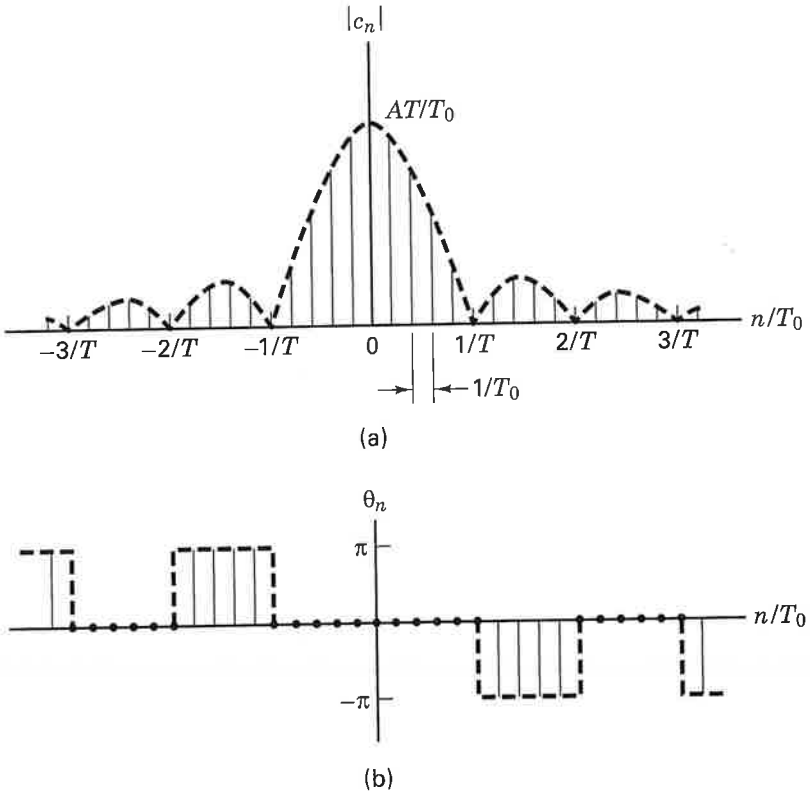


Figure A.5 Spectrum of a pulse train. (a) Magnitude spectrum. (b) Phase spectrum.

$$x_p(t) = \frac{AT}{T_0} \sum_{n=-\infty}^{\infty} \text{sinc}\frac{nT}{T_0} e^{j2\pi n f_0 t} \tag{A.25}$$

The ideal periodic pulse train contains frequency components at all integer multiples of the fundamental. In communication systems the significant portion of a baseband signal's power or energy is often assumed to be contained within the frequencies from zero to the first null of the magnitude spectrum (see Figure A.5a). Therefore, $1/T$ is often used as a measure of signal *bandwidth,* in hertz, for a pulse train with pulse width T. Note that bandwidth is inversely proportional to pulse width; the narrower are the pulses, the wider is the bandwidth associated with these pulses. Also, notice that the spacing between spectral lines $\Delta f = 1/T_0$ is inversely proportional to the pulse period; as the period increases, the lines move closer together.

A.2.3 Fourier Integral Transform

In communication systems we often encounter nonperiodic signals having finite energy in a finite interval, and having zero energy outside this interval. Such signals can be conveniently characterized using the Fourier integral transform, or simply

the *Fourier transform.* We can describe the nonperiodic signal as a periodic one, in the limiting sense. For example, consider the pulse train shown in Figure A.3. As $T_0 \to \infty$ and the pulse train approaches a single pulse, $x(t)$, the number of spectral lines approaches infinity and the spectral plot approaches a smooth frequency spectrum $X(f)$. For this limiting case, we can define the Fourier integral transform pair.

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\,dt \tag{A.26}$$

and

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,dt \tag{A.27}$$

where f is frequency measured in hertz. This pair can be used to describe the time–frequency relationship for nonperiodic signals.

Henceforth, the Fourier integral transform operation will be designated by the notation $\mathcal{F}\{\cdot\}$, and the inverse Fourier integral transform will be designated by $\mathcal{F}^{-1}\{\cdot\}$. The relationship between the time and frequency domains will be indicated by using the double arrow as follows:

$$x(t) \leftrightarrow X(f)$$

This notation indicates that $X(f)$ is the Fourier transform of $x(t)$ and that $x(t)$ is the inverse Fourier transform of $X(f)$. In the typical communications context, $x(t)$ is a real-valued function and $X(f)$ is a complex function, having real and imaginary components; in polar form, the spectrum, $X(f)$, can be specified by a magnitude characteristic and a phase characteristic:

$$X(f) = |X(f)|\,e^{j\theta(f)} \tag{A.28}$$

The properties of $X(f)$, the spectrum of a nonperiodic waveform, are similar to those of the spectrum for a periodic waveform, presented in Equations (A.17) to (A.23); that is, when $x(t)$ is real valued,

$$X(-f) = X^*(f) \tag{A.29}$$

$$= |X(f)|\,e^{-j\theta(f)} \tag{A.30}$$

where X^* is the complex conjugate of X. The magnitude spectrum $|X(f)|$ is an even function of f and the phase spectrum is an odd function of f. In many cases $X(f)$ is either purely real or purely imaginary, and only one plot suffices to describe it.

A.3 FOURIER TRANSFORM PROPERTIES

There are many excellent references dealing with the details of Fourier transforms and their properties [1–4]. In this appendix we will emphasize the properties that are fundamental to communication systems. Some of the key features affecting sig-

nal transmission in communication systems are time delay, phase shift, multiplication by other signals, frequency translation, waveform convolution, and spectral convolution. We shall focus on the Fourier properties (shifting and convolution) needed to describe these key communication features.

A.3.1 Time Shifting Property

If $x(t) \leftrightarrow X(f)$, then

$$\mathcal{F}\{x(t - t_0)\} = \int_{-\infty}^{\infty} x(t - t_0)e^{-j2\pi ft}\,dt \tag{A.31}$$

Let $\mu = t - t_0$; then

$$\mathcal{F}\{x(t - t_0)\} = \int_{-\infty}^{\infty} x(\mu)e^{-j2\pi f(\mu + t_0)}d\mu$$

$$= X(f)e^{-j2\pi ft_0}$$

As a signal is delayed in time, the magnitude of its frequency spectrum remains unchanged, but its phase spectrum experiences a phase shift. A time shift of t_0 in the time domain is equivalent to multiplication by $e^{-j2\pi ft_0}$ (a phase shift of $-2\pi ft_0$) in the frequency domain.

A.3.2 Frequency Shifting Property

If $x(t) \leftrightarrow X(f)$, then

$$\mathcal{F}\{x(t)e^{j2\pi f_0 t}\} = \int_{-\infty}^{\infty} x(t)e^{j2\pi f_0 t}\,e^{-j2\pi ft}\,dt$$

$$= \int_{-\infty}^{\infty} x(t)e^{-j2\pi(f - f_0)t}\,dt \tag{A.32}$$

$$= X(f - f_0)$$

This is the basic *frequency translating* property that describes the shifted spectrum resulting from multiplying a signal by $e^{j2\pi f_0 t}$. Equation (A.32) can be used in conjunction with Equation (A.9) to yield the Fourier transform of a waveform multiplied by a cosine wave, as follows:

$$x(t)\cos 2\pi f_0 t = \tfrac{1}{2}[x(t)e^{j2\pi f_0 t} + x(t)e^{-j2\pi f_0 t}] \tag{A.33}$$

$$x(t)\cos 2\pi f_0 t \leftrightarrow \tfrac{1}{2}[X(f - f_0) + X(f + f_0)]$$

This property is also called the *mixing* or *modulation* theorem. Multiplication of an arbitrary signal by a sinusoid of frequency f_0 translates the original signal spectrum by f_0, and also by $-f_0$.

A.4 USEFUL FUNCTIONS

A.4.1 Unit Impulse Function

A useful function in communication theory is the unit impulse or *Dirac delta* function, $\delta(t)$. The impulse function can be developed from any of several fundamental functions (e.g., a rectangular pulse or a triangular pulse). In each development, the impulse function is defined in the limiting sense (the pulse amplitude approaches infinity, the pulse width approaches zero, but the area under the pulse is constrained to be unity) [5]. The unit impulse function has the following important properties:

$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1 \tag{A.34}$$

$$\delta(t) = 0 \quad \text{for } t \neq 0 \tag{A.35}$$

$$\delta(t) \text{ is unbounded at } t = 0 \tag{A.36}$$

$$\mathcal{F}\{\delta(t)\} = \mathcal{F}^{-1}\{\delta(f)\} = 1 \tag{A.37}$$

$$\int_{-\infty}^{\infty} x(t)\delta(t - t_0)\, dt = x(t_0) \tag{A.38}$$

Equation (A.38) is known as the *sifting* or *sampling property;* the unit impulse multiplier selects a sample of the function $x(t)$ evaluated at $t = t_0$.

In some problems, it is useful to use the following equivalent integrals for an impulse function, defined in the time domain or the frequency domain [3].

$$\delta(t) = \int_{-\infty}^{\infty} e^{j2\pi ft}\, df \tag{A.39}$$

$$\delta(f) = \int_{-\infty}^{\infty} e^{-j2\pi ft}\, dt \tag{A.40}$$

A.4.2 Spectrum of a Sinusoid

For the purpose of representing a sinusoidal waveform by a Fourier transform, the waveform may be assumed to exist only in the interval $(-T_0/2 < t < T_0/2)$. Under these conditions the function has a Fourier transform as long as T_0 is finite. In the limit, T_0 is made very large, but finite. The spectrum of the waveform $x(t) = A \cos 2\pi f_0 t$ can be found by using Equations (A.9) and (A.26):

$$X(f) = \int_{-\infty}^{\infty} \frac{A}{2} (e^{j\,2\pi f_0 t} + e^{-j\,2\pi f_0 t}) e^{-j\,2\pi f t}\, dt$$

$$= \frac{A}{2} \int_{-\infty}^{\infty} e^{-j\,2\pi(f - f_0)t} + e^{-j\,2\pi(f + f_0)t} dt$$

As described in Equation (A.40), the foregoing integral expression can be equated to unit impulse functions located at frequencies $\pm f_0$ as follows:

$$X(f) = \frac{A}{2}[\delta(f - f_0) + \delta(f + f_0)] \tag{A.41}$$

Similarly, the spectrum of a sine waveform $y(t) = A \sin 2\pi f_0 t$ can be shown to be equal to

$$Y(f) = \frac{A}{2j}[\delta(f - f_0) - \delta(f + f_0)] \tag{A.42}$$

The cosine waveform spectrum is shown in Figure A.6, and the sine waveform spectrum is shown in Figure A.7. Each of the impulse functions shown on these spectral plots is depicted as a spike with a weight of $A/2$ or $-A/2$.

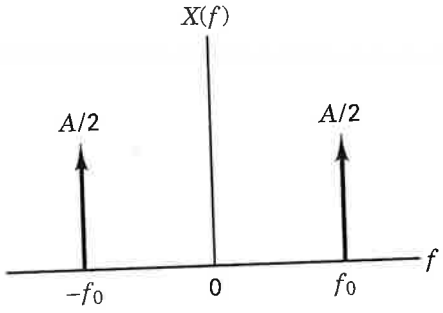


Figure A.6 Spectrum for $x(t) = A \cos 2\pi f_0 t$.

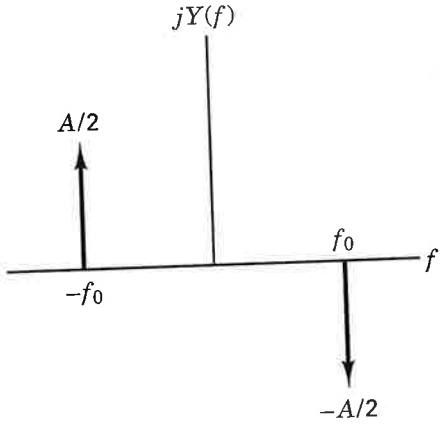


Figure A.7 Spectrum for $y(t) = A \sin 2\pi f_0 t$.

A.5 CONVOLUTION

Convolution was used by Oliver Heaviside in the late nineteenth century to calculate electrical circuit output current when the input voltage waveform was more complicated than a simple battery source. The use of the methods of Heaviside predates the use of the analytical methods developed by Fourier and Laplace (even though publications by Fourier and Laplace came earlier).

The response of a circuit to an impulse voltage $v(t) = \delta(t)$ is called the *impulse response* and is denoted by $h(t)$, as shown in Figure A.8; it is simply the output voltage that would result if the input were a delta function. Heaviside approximated an arbitrary voltage waveform, like the one shown in Figure A.9a, by a set of equally spaced pulses. Such pulses of finite height and nonzero duration are shown in Figure A.9b. In the limit as the pulse width $\Delta\tau$ approaches zero, each pulse approaches an impulse function with weight equal to the area under that pulse. In the following discussion we shall refer to these equally spaced pulses as *impulses* even though they are impulses *only in the limit.*

Care needs to be taken with the notation of time, since we are interested in the times at which impulses are applied and also the times at which their output responses are observed. We need to identify these two different time sequences; we shall use the following notation:

1. Time of the input application will be termed τ, so that the input voltage impulses are designated $v(\tau_1), v(\tau_2), \ldots, v(\tau_N)$.
2. Time of the output response will be termed t, so that the output currents are designated $i(t_1), i(t_2), \ldots, i(t_N)$.

Heaviside found the response or current produced by each input impulse independently; then he added the individual responses to get the total current. The weight of the impulse produced by the rectangular voltage at time τ_1 is the product $v(\tau_1)\ \Delta\tau$. The series of impulses can approximate the arbitrary input voltage as closely as desired by allowing $\Delta\tau$ to approach zero. Note again that the instant at which an impulse is applied is called τ_i, and the instance at which the system response is determined is called t_i, where τ is the input time variable, t is the output time variable, and $i = 1, \ldots, N$.

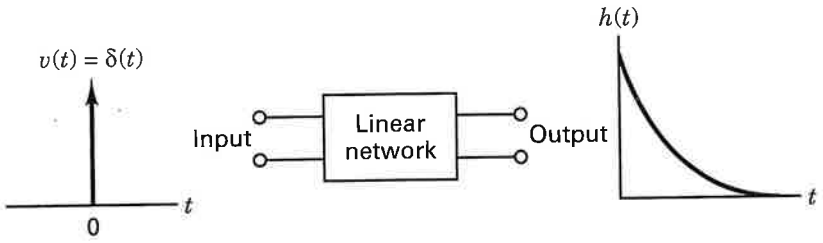


Figure A.8 Impulse response of a linear system.

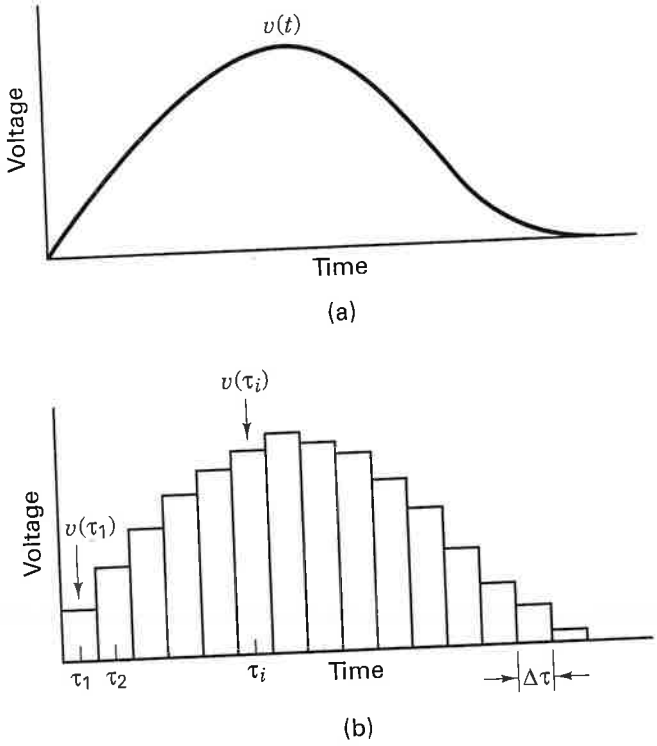


Figure A.9 (a) Input voltage waveform. (b) Approximate input voltage waveform.

Figure A.10 illustrates the output response $i(t) = A_1 h(t - \tau_1)$ to an impulse with height $v(\tau_1)$. Since the input impulse at τ_1 is *not* a *unit* impulse, we weight it with its strength or area, $A_1 = v(\tau_1)\,\Delta\tau$. At some time t_1, where $t_1 > \tau_1$, the output response to the impulse $v(\tau_1)$ is expressed as

$$i(t_1) = A_1 h(t_1 - \tau_1) \qquad \text{for } t_1 > \tau_1$$

as shown in Figure A.10. When there are several input impulses, the total output response for a linear system is simply the sum of the individual responses. Figure A.11 illustrates the response of the network to two input impulses. For N impulses, the output current measured at time t_1 can be expressed as

$$i(t_1) = A_1 h(t_1 - \tau_1) + A_2 h(t_1 - \tau_2) + \cdots + A_N(t_1 - \tau_N)$$

where the impulses are applied at $\tau_1, \tau_2, \ldots, \tau_N$ and where $t_1 > \tau_N$.

Any impulses applied at times greater than t_1 are disregarded, for they contribute nothing to $i(t_1)$. This corresponds to the *causality* requirement for physically realizable systems, which states that the system response must be zero prior to the application of the excitation. By generalizing, we get the output current at any time t, namely,

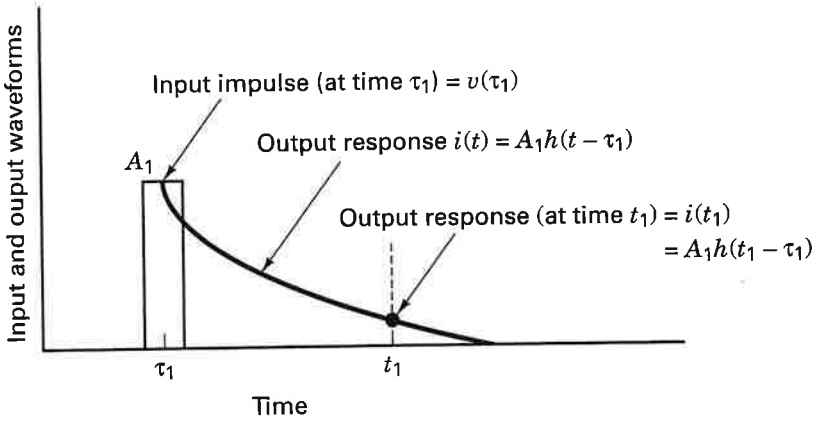


Figure A.10 Output response to an impulse at time τ_1.

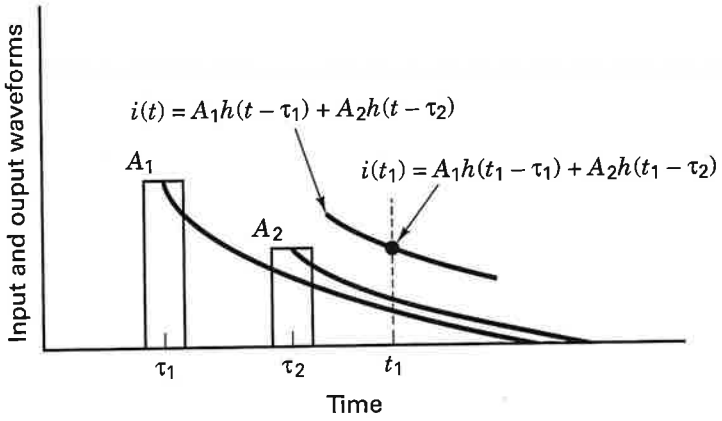


Figure A.11 Output response to two impulses.

$$i(t) = A_1h(t - \tau_1) + A_2h(t - \tau_2) + \cdots + A_Nh(t - \tau_N)$$

or

$$i(t) = \sum_{j=1}^{N} v(\tau_j)\,\Delta\tau h(t - \tau_j) \tag{A.43}$$

since the height of the impulse at τ_j is equal to $v(\tau_j)$. As $\Delta\tau$ approaches zero, the sum of the input impulses approaches the actual applied voltage $v(\tau)$; we can replace $\Delta\tau$ with $d\tau$, and the summation becomes the *convolution integral:*

$$i(t) = \int_{-\infty}^{\infty} v(\tau)h(t - \tau)\,d\tau \tag{A.44a}$$

or

$$i(t) = \int_{-\infty}^{\infty} v(t - \tau)\, h(\tau)\, d\tau \tag{A.44b}$$

In shorthand notation, this is expressed as

$$i(t) = v(t) * h(t) \tag{A.45}$$

In summary, $i(t)$ is the sum of the individual impulse responses as a function of output time t. Each impulse response is due to an impulse applied at some input time τ and is weighted by the strength of that impulse.

A.5.1 Graphical Illustration of Convolution

Consider that an input square pulse $v(t)$ is applied to a linear network whose impulse response is labeled $h(t)$ as shown in Figure A.12a. The output response is characterized by the convolution integral expressed in Equation (A.44).

The independent variable in the convolution integral is τ. The functions $v(\tau)$ and $h(-\tau)$ are shown in Figure A.12b. Note that $h(-\tau)$ is obtained by folding $h(\tau)$ about $\tau = 0$. The term $h(t-\tau)$ represents the function $h(-\tau)$ shifted by t seconds along the positive τ axis. Figure A.12c shows the function $h(t_1 - \tau)$. The value of the convolution integral at $t = t_1$ is given by Equation (A.44) evaluated at $t = t_1$. This is simply the area under the product curve of $v(\tau)$ and $h(t_1 - \tau)$, shown shaded in Figure A.12d. Similarly, the convolution integral evaluated at $t = t_2$ is equal to the shaded area in Figure A.12e. Figure A.12f is a plot of the output response as a result of the square pulse input to the circuit with impulse response shown in Figure A.12a. Each evaluation of the convolution integral, at some time t_i, yields one point, $i(t_i)$, on the plot of Figure A.12f.

A.5.2 Time Convolution Property

If $x_1(t) \leftrightarrow X_1(f)$, and $x_2(t) \leftrightarrow X_2(f)$, then

$$x_1(t) * x_2(t) = \int_{-\infty}^{\infty} x_1(\tau)\, x_2(t - \tau)\, d\tau$$

$$\mathcal{F}\{x_1(t) * x_2(t)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1(\tau)\, x_2(t - \tau)\, d\tau\, e^{-j2\pi ft}\, dt$$

For linear systems, we may exchange the order of integration as follows:

$$\mathcal{F}\{x_1(t) * x_2(t)\} = \int_{-\infty}^{\infty} x_1(\tau)\, d\tau \int_{-\infty}^{\infty} x_2(t - \tau) e^{-j2\pi ft}\, dt \tag{A.46}$$

By the Fourier *time shifting property,* the second integral expression of the right-hand side is equal to $X_2(f)e^{-j2\pi f\tau}$:

$$\mathcal{F}\{x_1(t) * x_2(t)\} = X_2(f) \int_{-\infty}^{\infty} x_1(\tau) e^{-j2\pi f\tau}\, d\tau \tag{A.47}$$

$$= X_1(f) X_2(f)$$

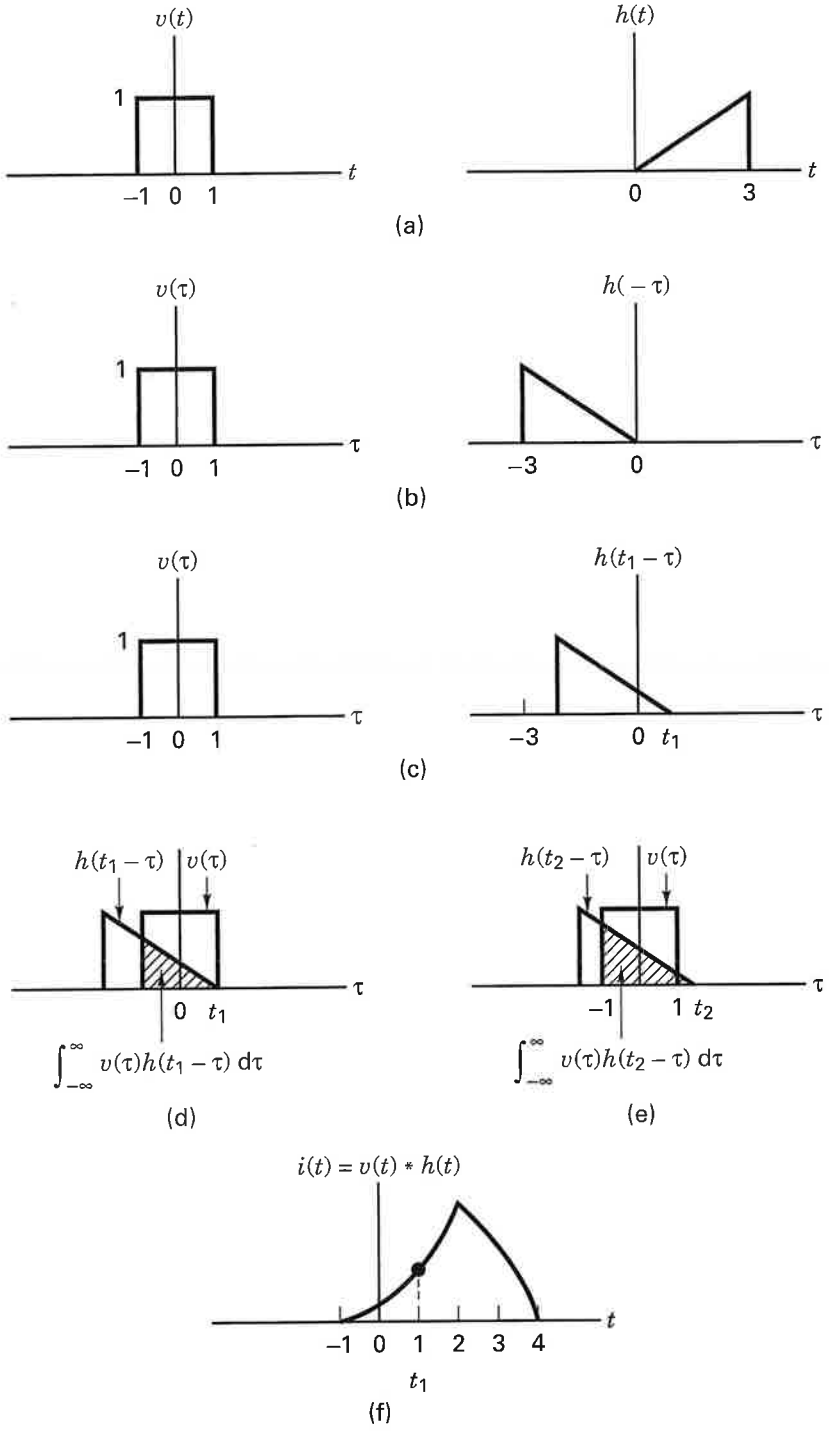


Figure A.12 Graphical example of convolution.

Therefore, the operation of *convolution* in the time domain can be replaced by *multiplication* in the frequency domain.

A.5.3 Frequency Convolution Property

Because of the symmetry of the Fourier transform pair in Equations (A.26) and (A.27), it can be shown that multiplication in the time domain transforms to convolution in the frequency domain:

$$x_1(t)x_2(t) \leftrightarrow X_1(f) * X_2(f) \tag{A.48}$$

The properties that transform multiplication in one domain to convolution in the other domain are particularly useful, since one operation is often easier to perform than the other. For example, we discussed earlier that Heaviside used convolution to solve for the output current of a linear system when the input was excited by an arbitrary voltage waveform. Such methods involve the (sometimes tedious) convolution of an input waveform with the impulse response of a system. Since convolution in the time domain is transformed into multiplication in the frequency domain, as shown in Equation (A.47), for a linear system we can simply multiply the input waveform spectrum by the system transfer function. The output waveform is then found by taking the inverse Fourier transform of the product:

$$i(t) = \mathcal{F}^{-1}\{V(f)H(f)\} \tag{A.49}$$

Solutions of the form shown in Equation (A.49) are often much easier to perform than those described by Equation (A.45). However, under certain circumstances, the operation of convolution is so simple that it can be performed graphically, by inspection. For example, suppose that we wished to multiply an arbitrary waveform by some fixed frequency cosine wave, such as a carrier wave, in the case of modulation. By applying Equation (A.48), we can convolve the spectrum of the arbitrary waveform with the spectrum of the cosine wave. This is easily accomplished, as is shown in the next section.

A.5.4 Convolution of a Function with a Unit Impulse

By the property shown in Equation (A.47), it should be clear that if

$$x(t) \leftrightarrow X(f)$$

and since

$$\delta(t) \leftrightarrow 1$$

then

$$x(t) * \delta(t) \leftrightarrow X(f) \tag{A.50}$$

It should also be evident that

$$x(t) * \delta(t) = x(t) \tag{A.51}$$

and

$$X(f) * \delta(f) = X(f) \tag{A.52}$$

We therefore conclude that convolution of a function with a unit impulse function reproduces the original function. A simple extension of Equation (A.52) yields

$$X(f) * \delta(f - f_0) = X(f - f_0) \tag{A.53}$$

Figure A.13 illustrates the ease of convolving the spectrum of an arbitrary waveform with the spectrum of a cosine wave. Figure A.13a shows an arbitrary baseband spectrum $X(f)$. Figure A.13b shows a spectrum, $Y(f) = \delta(f - f_0) + \delta(f + f_0) = \mathcal{F}\{2 \cos 2\pi f_0 t\}$. The output, $Z(f) = X(f) * Y(f)$, in Figure A.13c is obtained by convolving the waveform spectrum with the impulse functions of $Y(f)$ according to Equation (A.53), where the impulses act as sampling functions. Hence, for this simple example, convolution can be performed graphically by sweeping the sampling impulses past the waveform spectrum. Multiplication by the impulse functions at each step in the sweep yields replications of the waveform spectrum. The result, shown in Figure A.13c, is a shifted version of the original spectrum $X(f)$ to the locations of the impulse functions in Figure A.13b.

A.5.5 Demodulation Application of Convolution

In Section A.5.4 we examined a waveform multiplied by $2 \cos 2\pi f_0 t$. We illustrated the frequency-domain view of convolving the waveform spectrum with a cosine-wave spectrum. In this section we look at the reverse process. A waveform that has been multiplied by $2 \cos 2\pi f_0 t$ is to be demodulated (the waveform is to be restored to its baseband frequency range).

Figure A.14a represents the spectrum, $Z(f)$, of the waveform that has been upshifted in frequency. We can demodulate this upshifted waveform and recover the baseband waveform, by multiplying it by $2 \cos 2\pi f_0 t$. Instead, we shall illustrate the detection process in the frequency domain by convolving $Z(f)$ with the spectrum of the carrier, $Y(f) = \delta(f - f_0) + \delta(f + f_0)$, shown in Figure a.14b.

A simple extension of Equations (A.52) and (A.53) yields

$$X(f - f_0) * \delta(f - f_1) = X(f - f_0 - f_1) \tag{A.54}$$

Therefore, the result of demodulation, $X(f) = Z(f) * Y(f)$ is obtained by applying Equation (A.54). The resulting signal spectrum appears at baseband (detected) and also at frequencies $\pm 2f_0$, as shown in Figure A.14c. As in the previous section, the convolution can be performed graphically. The resulting Figure A.14c contains the following terms:

$$\begin{aligned} &[Z(f - f_0) + Z(f + f_0)] * [\delta(f - f_0) + \delta(f + f_0)] \\ &\quad = Z(f - f_0) * \delta(f - f_0) + Z(f - f_0) * \delta(f + f_0) \\ &\qquad + Z(f + f_0) * \delta(f - f_0) + Z(f + f_0) * \delta(f + f_0) \\ &\quad = 2Z(f) + Z(f - 2f_0) + Z(f + 2f_0) \end{aligned} \tag{A.55}$$

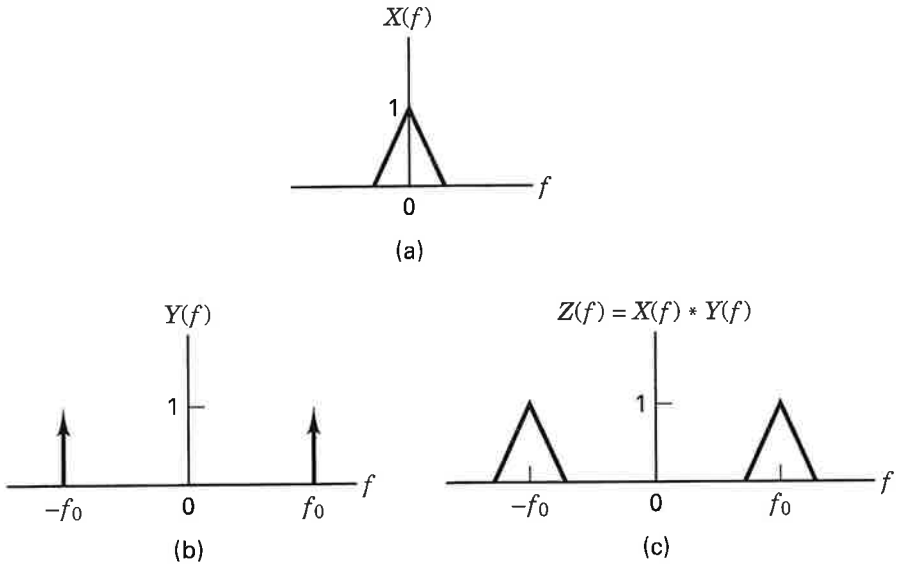


Figure A.13 Convolving a signal spectrum with a cosine-wave spectrum.

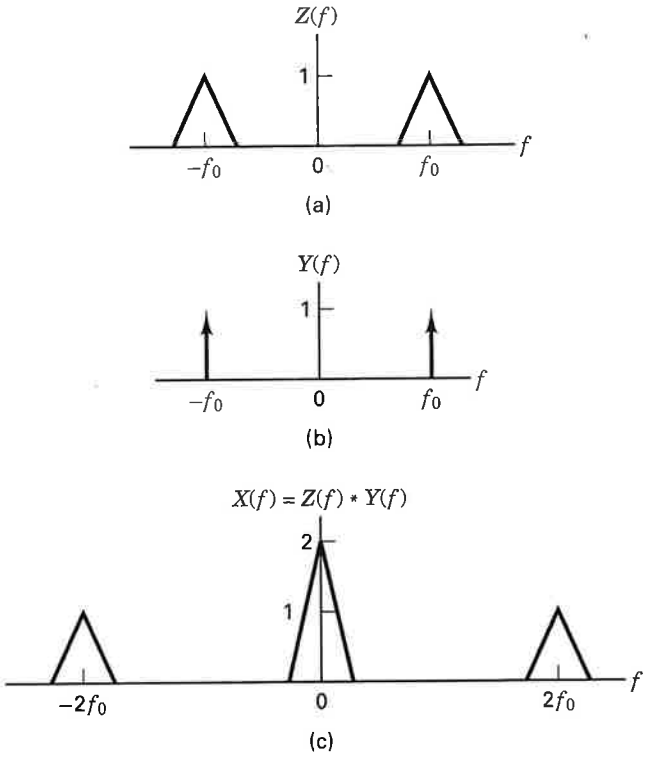


Figure A.14 Demodulation application.

Notice that the resulting terms consist of the baseband spectrum plus terms associated with higher-frequency components. The result is typical of the detection process; the higher-frequency terms are filtered and discarded, leaving the demodulated baseband spectrum.

A.6 TABLES OF FOURIER TRANSFORMS AND OPERATIONS

Commonly used Fourier transforms are listed in Table A.1 and Fourier operations in Table A.2.

TABLE A.1 Fourier Transforms

$x(t)$	$X(f)$
1. $\delta(t)$	1
2. 1	$\delta(f)$
3. $\cos 2\pi f_0 t$	$\frac{1}{2}[\delta(f-f_0)+\delta(f+f_0)]$
4. $\sin 2\pi f_0 t$	$\frac{1}{2j}[\delta(f-f_0)-\delta(f+f_0)]$
5. $\delta(t-t_0)$	$\exp(-j2\pi f t_0)$
6. $\exp(j2\pi f_0 t)$	$\delta(f-f_0)$
7. $\exp(-a\lvert t\rvert),\quad a>0$	$\frac{2a}{a^2+(2\pi f)^2}$
8. $\exp\left[-\pi\left(\frac{t}{T}\right)^2\right]$	$T\exp[-\pi(fT)^2]$
9. $u(t)=\begin{cases}1 & \text{for } t>0\\ 0 & \text{for } t<0\end{cases}$	$\frac{1}{2}\delta(f)+\frac{1}{j2\pi f}$
10. $\exp(-at)\,u(t),\quad a>0$	$\frac{1}{a+j2\pi f}$
11. $t\exp(-at)\,u(t),\quad a>0$	$\frac{1}{(a+j2\pi f)^2}$
12. $\text{rect}\left(\frac{t}{T}\right)$	$T\,\text{sinc}\,fT$
13. $\cos 2\pi f_0 t\left[\text{rect}\left(\frac{t}{T}\right)\right]$	$\frac{T}{2}[\text{sinc}(f-f_0)T+\text{sinc}(f+f_0)T]$
14. $W\,\text{sinc}\,Wt$	$\text{rect}\left(\frac{f}{W}\right)$
15. $\begin{cases}1-\frac{\lvert t\rvert}{T} & \text{for } \lvert t\rvert\le T\\ 0 & \text{for } \lvert t\rvert>T\end{cases}$	$T\,\text{sinc}^2 fT$
16. $\sum_{m=-\infty}^{\infty}\delta(t-mT_0)$	$\frac{1}{T_0}\sum_{n=-\infty}^{\infty}\delta\left(f-\frac{n}{T_0}\right)$

Note: rect $(f/2W) = 1$ for $-W < f < W$, 0 for $|f| > W$, and sinc $x = (\sin \pi x)/\pi x$.

TABLE A.2 Fourier Operations

Operation	$x(t)$	$X(f)$
1. Scaling	$x(at)$	$\frac{1}{\lvert a\rvert} X\left(\frac{f}{a}\right)$
2. Time shifting	$x(t-t_0)$	$X(f)\exp(-j2\pi f t_0)$
3. Frequency shifting	$x(t)\exp(j2\pi f_0 t)$	$X(f-f_0)$
4. Time differentiation	$\frac{d^n x}{dt^n}$	$(j2\pi f)^n X(f)$
5. Frequency differentiation	$(-j2\pi t)^n(x(t))$	$\frac{d^n X}{df^n}$
6. Time integration	$\int_{-\infty}^{t} x(\tau)d\tau$	$\frac{1}{j2\pi f}X(f)+\frac{1}{2}X(0)\delta(f)$
7. Time convolution	$x_1(t)*x_2(t)$	$X_1(f)X_2(f)$
8. Frequency convolution	$x_1(t)x_2(t)$	$X_1(f)*X_2(f)$

REFERENCES

1. Papoulis, A., *Signal Analysis,* McGraw-Hill Book Company, New York, 1977.
2. Panter, P. F., *Modulation, Noise, and Spectral Analysis,* McGraw-Hill Book Company, New York, 1965.
3. Bracewell, R., *The Fourier Transform and Its Applications,* McGraw-Hill Book Company, New York, 1978.
4. Haykin, S., *Communication Systems,* John Wiley & Sons, Inc., New York, 1983.
5. Schwartz, M., *Information, Transmission, Modulation, and Noise,* McGraw-Hill Book Company, New York, 1980.

B.1 BA

APPENDIX B

Fundamentals of Statistical Decision Theory

The basic elements of a statistical decision problem are (1) a set of hypotheses that characterize the possible true states of nature, (2) a test in which data are obtained from which we wish to infer the truth, (3) a decision rule that operates on the data to decide in an optimal fashion which hypothesis best describes the true state of nature, and (4) a criterion of optimality. These fundamental steps are treated in the material that follows. The *optimality criterion* we will choose for the decision rule is to minimize the probability of making an erroneous decision, although other criteria are possible [1].

The subject of statistical decision theory and hypothesis testing builds on the mathematical discipline of probability theory and random variables. It is assumed that the reader has a familiarity with these subjects; if not, Reference [2] is a suggested resource.

B.1 BAYES' THEOREM

The mathematical foundations of hypothesis testing rest on Bayes' theorem, which is derived from the definition of the relationship between the conditional and joint probability of the random variables A and B:

$$P(A|B)P(B) = P(B|A)P(A) = P(A, B) \tag{B.1}$$

A statement of the theorem is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{B.2}$$

Bayes' theorem allows us to infer the conditional probability $P(A|B)$, from the conditional probability $P(B|A)$.

B.1.1 Discrete Form of Bayes' Theorem

Bayes' theorem can be expressed in discrete form as

$$P(s_i|z_j) = \frac{P(z_j|s_i)P(s_i)}{P(z_i)} \qquad \begin{matrix} i = 1, \ldots, M \\ j = 1, \ldots \end{matrix} \tag{B.3}$$

where

$$P(z_j) = \sum_{i=1}^{M} P(z_j|s_i)P(s_i)$$

In a communications application, s_i is the ith signal class, from a set of M classes, and z_j is the jth sample of a received signal. Equation (B.3) can be thought of as the description of an experiment involving a received sample and some statistical knowledge of the signal classes to which the received sample may belong. The probability of occurrence of the ith signal class, $P(s_i)$, before the experiment, is called the *a priori probability*. As a result of examining a particular received sample, z_j, we can find a statistical measure of the *likelihood* that z_j belongs to class s_i from the conditional probability density function (pdf) $P(z_j|s_i)$. *After* the experiment, we can compute the *a posteriori probability*, $P(s_i|z_j)$, which can be thought of as a "refinement" of our prior knowledge. Thus we enter into the experiment with some a priori knowledge concerning the probability of the state of nature, and after examining a sample signal, we are provided with an "after-the-fact" a posteriori probability. The parameter $P(z_j)$ is the probability of the received sample, z_j, over the entire space of signal classes. This term, $P(z_j)$, can be thought of as a scaling factor, since its value is the same for *each* signal class.

Example B.1 Use of Bayes' Theorem (Discrete Form)

Given two boxes of parts. Box 1 contains 1000 parts, of which 10% are defective, and box 2 contains 2000 parts, of which 5% are defective. If a box is randomly chosen and then a part is randomly chosen from it, tested, and found to be good, what is the probability that the part came from box 1?

Solution

$$P(\text{box 1}|\text{GP}) = \frac{P(\text{GP}|\text{box 1})P(\text{box 1})}{P(\text{GP})}$$

where GP means "good part."

$$\begin{aligned} P(\text{GP}) &= P(\text{GP}|\text{box 1})\,P(\text{box 1}) + P(\text{GP}|\text{box 2})\,P(\text{box 2}) \\ &= (0.90)(0.5) + (0.95)(0.5) \\ &= 0.450 + 0.475 = 0.925 \end{aligned}$$

$$P(\text{box 1}|\text{GP}) = \frac{0.450}{0.925} = 0.486$$

Before the experiment, the a priori probability of having chosen either box 1 or box 2 was equally likely. After obtaining a good part, the Bayesian computation can be regarded as a way of "fine tuning" our thinking that $P(\text{box 1}) = 0.5$ to yield the a posteri-

ori probability of 0.486. The Bayes' theorem is simply a formalization of common sense. Having selected a good part from one of the two boxes, isn't it intuitively reasonable that there is a higher probability that the part came from the box with the larger concentration of good parts, and a lower probability that it came from the box with the smaller concentration of good parts? The Bayes' theorem has refined the a priori statistic into an a posteriori statistic for the probability of box selection.

(B.3)

Example B.2 Decision Theory Applied to a Betting Game

A box has three coins: a fair coin, a two-headed coin, and a two-tailed coin. You are asked to pick one coin at random, look at one side only, and guess head or tail for the other side. What is the optimum decision strategy for this game?

Solution

We can view this problem as a signal detection problem. A signal is transmitted, but because of the channel noise, the received signal is somewhat obscured. Not being able to look at the other side of the coin is tantamount to receiving a noise perturbed signal. Let H_i represent the hypotheses ($i = F, H, T$), where F, H, and T, stand for fair, head, and tail, respectively:

H_F: H, T (fair coin)

H_H: H, H (two-headed coin)

H_T: T, T (two-tailed coin)

Let z_j represent the received sample ($j = H, T$), where z_H is a head and z_T is a tail. Let the priori probabilities of the hypotheses be equally likely, so that $P(H_F) = P(H_H) = P(H_T) = \frac{1}{3}$. Using Bayes' theorem,

$$P(H_i|z_j) = \frac{P(z_j|H_i)P(H_i)}{\sum_i P(z_j|H_i)P(H_i)}$$

we need to compute the probability for each hypothesis, given each signal class. Thus, we need to examine the results of *six* computations before we can establish an optimum decision strategy. In each case, the value of $P(z_j|H_i)$ can be obtained from the conditional probabilities drawn in Figure B.1. Consider that we choose a coin and view a head (z_H), we compute the following three a posteriori probabilities:

$$P(H_F|z_H) = \frac{(\frac{1}{2})(\frac{1}{3})}{(\frac{1}{2})(\frac{1}{3}) + (1)(\frac{1}{3}) + 0} = \frac{1}{3}$$

$$P(H_H|z_H) = \frac{(1)(\frac{1}{3})}{(\frac{1}{2})(\frac{1}{3}) + (1)(\frac{1}{3}) + 0} = \frac{2}{3}$$

$$P(H_T|z_H) = 0$$

If the received sample is a tail (z_T), we similarly compute

$$P(H_F|z_T) = \frac{1}{3}$$

$$P(H_H|z_T) = 0$$

$$P(H_T|z_T) = \frac{2}{3}$$

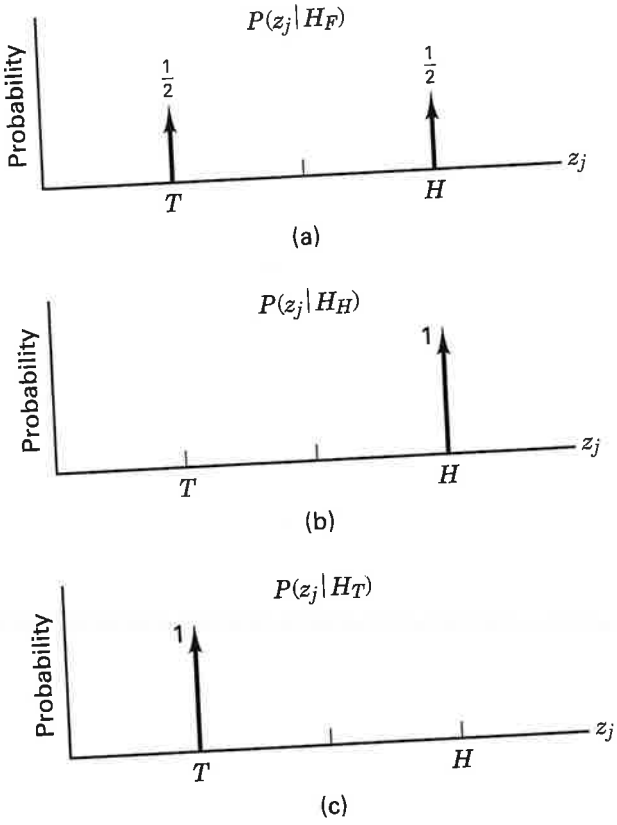


Figure B.1 Conditional probability $P(z_j|H_i)$. (a) Conditioned on the fair-coin hypothesis. (b) Conditioned on the two-headed-coin hypothesis. (c) Conditioned on the two-tailed-coin hypothesis.

The optimum decision strategy then is as follows: If a head, z_H, is received, choose hypothesis H_H (that the other side is also a head). If a tail, z_T, is received, choose hypothesis H_T (that the other side is also a tail).

B.1.2 Mixed Form of Bayes' Theorem

For most communication engineering applications of interest, the possible values of the received samples are *continuous* in range, because of the additive Gaussian noise in the channel. Therefore, the most useful form of Bayes' theorem contains a continuous- instead of a discrete-valued pdf. We shall rewrite Equation (B.3) to emphasize this change:

$$P(s_i|z) = \frac{p(z|s_i)P(s_i)}{p(z)} \qquad i = 1, \ldots, M \tag{B.4}$$

$$p(z) = \sum_{i=1}^{M} p(z|s_i)P(s_i)$$

Here, $p(z|s_i)$ is the conditional pdf of the received continuous-valued sample, z, conditioned on the signal class, s_i.

Example B.3 A Pictorial View of Bayes' Theorem

Consider two signal classes, s_1 and s_2, characterized by the triangular-shaped conditional pdfs, $p(z|s_1)$ and $p(z|s_2)$, illustrated in Figure B.2. A signal is received; it might have any value on the z-axis. If the pdfs did not overlap, we could classify the signal with certainty. For the example shown in Figure B.2, we need a rule to help us classify received signals, since some signals will fall in the region where the two pdfs overlap. Consider a received signal, z_a. Assume that the two signal classes, s_1 and s_2, are equally likely, and calculate the two alternative a posteriori probabilities. Suggest a decision rule that the receiver should use for deciding to which signal class z_a belongs. Repeat this for signal z_b.

Solution

From Figure B.2 we can see that $p(z_a|s_1) = 0.5$ and $p(z_a|s_2) = 0.3$. Thus,

$$P(s_1|z_a) = \frac{p(z_a|s_1)P(s_1)}{p(z_a|s_1)P(s_1) + p(z_a|s_2)P(s_2)}$$

$$= \frac{(0.5)(0.5)}{(0.5)(0.5) + (0.3)(0.5)} = \frac{5}{8}$$

and

$$P(s_2|z_a) = \frac{(0.3)(0.5)}{(0.5)(0.5) + (0.3)(0.5)} = \frac{3}{8}$$

One rule is to decide that the received signal belongs to the class with the maximum a posteriori probability (class s_1). An equivalent rule, for the case of equal a priori probabilities, is to examine the value of the pdf conditioned on each signal class (referred to as the likelihood of the signal class) and choose the class with the maximum. Examine Figure B.2 and notice that this *maximum likelihood rule* parallels our intuition. The likelihood that signal z_a belongs to each class corresponds to an encircled point on each pdf. The maximum likelihood rule is to choose the signal class that yields the largest conditional probability of all the alternatives. We repeat the computations for the received signal z_b, as follows:

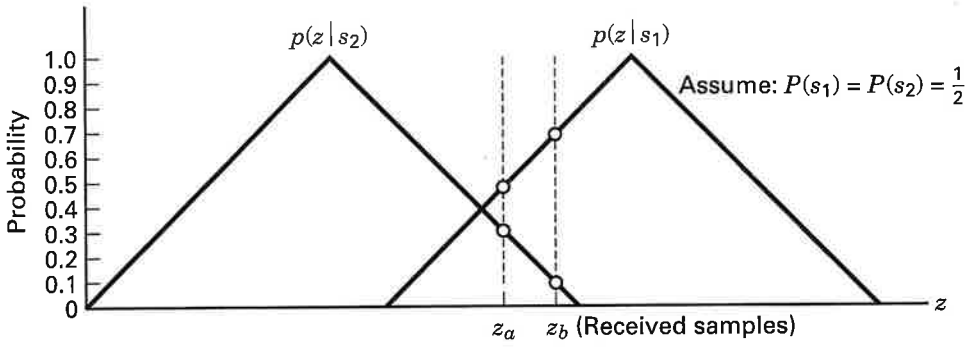


Figure B.2 Pictorial view of Bayes' theorem.

$$P(s_1|z_b) = \frac{(0.7)(0.5)}{(0.7)(0.5) + (0.1)(0.5)} = \frac{7}{8}$$

$$P(s_2|z_b) = \frac{(0.1)(0.5)}{(0.7)(0.5) + (0.1)(0.5)} = \frac{1}{8}$$

As before, the maximum likelihood rule dictates that we choose signal class s_1. Notice that in the case of received sample z_b, we can have greater confidence in the correctness of our choice compared to the case of signal z_a. This is because the ratio of $p(z_b|s_1)$ to $p(z_b|s_2)$ is considerably larger than the ratio of $p(z_a|s_1)$ to $p(z_a|s_2)$.

B.2 DECISION THEORY

B.2.1 Components of the Decision Theory Problem

Having reviewed hypothesis testing based on Bayesian statistics, let us examine more carefully the components of the decision theory problem in the context of a communication system, as shown in Figure B.3. The signal source at the transmitter consists of a set $\{s_i(t)\}$, $i = 1, \ldots, M$, of waveforms (or hypotheses). A signal waveform $r(t) = s_i(t) + n(t)$ is received, where $n(t)$ is an additive white Gaussian noise (AWGN) process introduced in the channel. At the receiver, the waveform is reduced to a single number $z(t = T)$, which may appear anywhere on the z-axis.

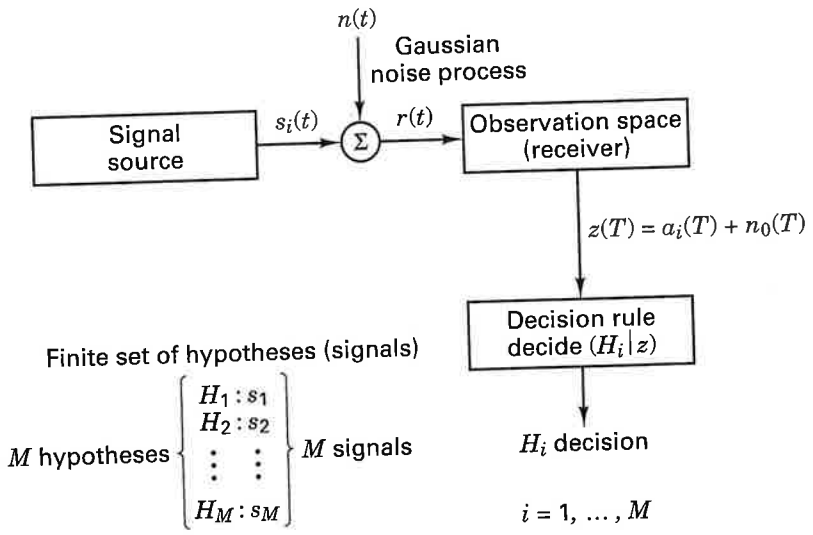


Figure B.3 Components of the decision theory problem in the context of a communication system.

Because the noise is a Gaussian process and the receiver is assumed linear, the output, $z(t)$, is also a Gaussian process [1], and the number, $z(T)$, is a *continuous-valued random variable.*

$$z(T) = a_i(T) + n_0(T) \tag{B.5}$$

The sample $z(T)$ is made up of a signal component, $a_i(T)$, and a noise component, $n_0(T)$. The time T is the symbol duration. At each kT, where k is an integer, the receiver uses a decision rule for deciding which signal class has been received. For ease of notation, Equation (B.5) is sometimes written simply as $z = a_i + n_0$, where the functional dependence on T is implicit.

B.2.2 The Likelihood Ratio Test and the Maximum A Posteriori Criterion

A reasonable starting point for establishing the receiver decision rule for the case of *two* signal classes is

$$P(s_1|z) \underset{H_2}{\overset{H_1}{\gtrless}} P(s_2|z) \tag{B.6}$$

Equation (B.6) states that we should choose hypothesis H_1 if the a posteriori probability $P(s_1|z)$ is greater than the a posteriori probability $P(s_2|z)$. Otherwise, we should choose hypothesis H_2.

We can replace the posteriori probabilities of Equation (B.6) with their equivalent expressions from Bayes' theorem [Equation (B.4)], yielding

$$p(z|s_1)\,P(s_1) \underset{H_2}{\overset{H_1}{\gtrless}} P(z|s_2)\,P(s_2) \tag{B.7}$$

We now have a decision rule in terms of pdfs (likelihoods). If we rearrange Equation (B.7) and put it in the form

$$\frac{p(z|s_1)}{p(z|s_2)} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{P(s_2)}{P(s_1)} \tag{B.8}$$

then the left-hand ratio is known as the *likelihood ratio* and the entire equation is often referred to as the *likelihood ratio test.* Equation (B.8) corresponds to making a decision based on a comparison of a measurement of a received signal to a threshold. Since the test is based on choosing the signal class with maximum a posteriori probability, the decision criterion is called the *maximum a posteriori* (MAP) criterion. It is also called the *minimum error criterion,* since on the average, this criterion yields the minimum number of incorrect decisions. It should be emphasized that this criterion is optimum only when each of the error types are equally harmful or costly. When some of the error types are more costly than others, a criterion that incorporates relative cost of the errors should best be employed [1].

B.2.3 The Maximum Likelihood Criterion

Very often there is no knowledge available about the a priori probabilities of the hypotheses or signal classes. Even when such information is available, its accuracy is sometimes mistrusted. In those instances, decisions are usually made by assuming the most conservative a priori probabilities possible; that is, the values of the a priori probabilities are selected so that the classes are *equally likely*. When this is done, the MAP criterion is known as the *maximum likelihood criterion*, and Equation (B.8) can be written as

$$\frac{p(z|s_1)}{p(z|s_2)} \underset{H_2}{\overset{H_1}{\gtrless}} 1 \tag{B.9}$$

Notice that the maximum likelihood criterion of Equation (B.9) is the same as the maximum likelihood rule that was described in Example B.3.

B.3 SIGNAL DETECTION EXAMPLE

B.3.1 The Maximum Likelihood Binary Decision

The pictorial view of the decision process in Example B.3 dealt with triangular-shaped probability density functions as a convenient example. Figure B.4 illustrates the conditional pdfs for the binary noise-perturbed output signals, $z(T) = a_1 + n_0$ and $z(T) = a_2 + n_0$ from a typical receiver. The signals a_1 and a_2 are mutually independent and are equally likely. The noise n_0 is assumed to be an independent Gaussian random variable with zero mean, variance σ_0^2, and pdf given by

$$p(n_0) = \frac{1}{\sigma_0\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{n_0^2}{\sigma_0^2}\right)\right] \tag{B.10}$$

We can therefore write the likelihood ratio, described in Equation (B.8), as

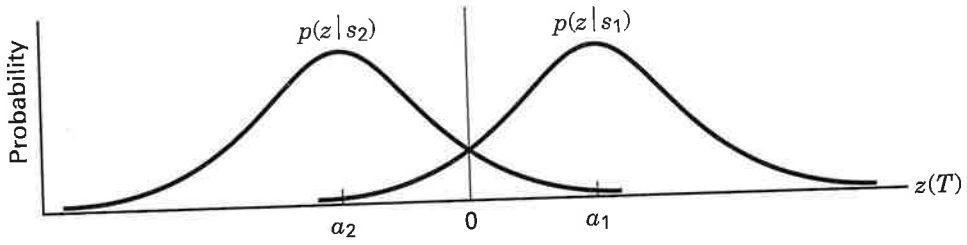


Figure B.4 Conditional pdfs for a typical binary receiver.

$$\Lambda(z) = \frac{p(z|s_1)}{p(z|s_2)}$$

$$= \frac{\dfrac{1}{\sigma_0\sqrt{2\pi}} \exp\left[-\dfrac{1}{2}\left(\dfrac{z-a_1}{\sigma_0}\right)^2\right]}{\dfrac{1}{\sigma_0\sqrt{2\pi}} \exp\left[-\dfrac{1}{2}\left(\dfrac{z-a_2}{\sigma_0}\right)\right]} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{P(s_2)}{P(s_1)} \tag{B.11}$$

$$= \frac{\exp\left(-\dfrac{z^2}{2\sigma_0^2}\right) \exp\left(-\dfrac{a_1^2}{2\sigma_0^2}\right) \exp\left(\dfrac{2za_1}{2\sigma_0^2}\right)}{\exp\left(-\dfrac{z^2}{2\sigma_0^2}\right) \exp\left(-\dfrac{a_2^2}{2\sigma_0^2}\right) \exp\left(\dfrac{2za_2}{2\sigma_0^2}\right)} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{P(s_2)}{P(s_1)}$$

$$= \exp\left[\frac{z(a_1-a_2)}{\sigma_0^2} - \frac{a_1^2-a_2^2}{2\sigma_0^2}\right] \underset{H_2}{\overset{H_1}{\gtrless}} \frac{P(s_2)}{P(s_1)}$$

where a_1 is the receiver output signal component when $s_1(t)$ is sent and a_2 is the output signal component when $s_2(t)$ is sent. The inequality relationship described by Equation (B.11) is preserved for any *monotonically* increasing (or decreasing) transformation. Therefore, to simplify Equation (B.11), we take the natural logarithm of both sides, resulting in the log-likelihood ratio.

$$L(z) = \frac{z(a_1-a_2)}{\sigma_0^2} - \frac{a_1^2-a_2^2}{2\sigma_0^2} \underset{H_2}{\overset{H_1}{\gtrless}} \ln\frac{P(s_2)}{P(s_1)} \tag{B.12}$$

When the classes are equally likely,

$$\ln\frac{P(s_2)}{P(s_1)} = 0$$

so that

$$z \underset{H_2}{\overset{H_1}{\gtrless}} \frac{a_1^2-a_2^2}{2(a_1-a_2)}$$

$$z \underset{H_2}{\overset{H_1}{\gtrless}} \frac{a_1+a_2}{2} = \gamma_0 \tag{B.13}$$

For *antipodal signals*, $s_1(t) = -s_2(t)$ and $a_1 = -a_2$; thus, we can write

$$z \underset{H_2}{\overset{H_1}{\gtrless}} 0 \tag{B.14}$$

Therefore, the maximum likelihood rule for the case of equally likely antipodal signals compares the received sample to a zero threshold, which is tantamount to deciding $s_1(t)$ if the sample is positive, and $s_2(t)$ if the signal is negative.

B.3.2 Probability of Bit Error

For the binary example in Section B.3.1, we want to compute the bit error probability, P_B, using the decision rule in Equation (B.13). The probability of an error is calculated by summing the probabilities of the various ways that an error can be made:

$$P_B = P(H_2|s_1)P(s_1) + P(H_1|s_2)P(s_2) \tag{B.15}$$

That is, given that class $s_1(t)$ was transmitted, an error results if hypothesis H_2 is chosen; or, given that class $s_2(t)$ was transmitted, an error results if hypothesis H_1 is chosen. For the special case of symmetric probability density functions, and for $P(s_1) = P(s_2) = 0.5$, we can write

$$P_B = P(H_2|S_1) = P(H_1|S_2) \tag{B.16}$$

The probability of an error, P_B, is equal to the probability that an incorrect hypothesis, H_1, will be decided when $s_2(t)$ is sent, or that H_2 will be decided when $s_1(t)$ is sent. Thus P_B is numerically equal to the area under the "tail" of either pdf, $p(z|s_1)$ or $p(z|s_2)$, falling on the *incorrect side* of the threshold. We can therefore compute P_B by integrating $p(z|s_1)$ between the limits $-\infty$ and γ_0 or by integrating $p(z|s_2)$ between the limits γ_0 and ∞, as follows:

$$\begin{aligned} P_B &= \int_{\gamma_0=(a_1+a_2)/2}^{\infty} p(z|s_2)\,dz \\ &= \int_{(a_1+a_2)/2}^{\infty} \frac{1}{\sigma_0\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-a_2}{\sigma_0}\right)^2\right] dz \end{aligned} \tag{B.17}$$

Let

$$u = \frac{z - a_2}{\sigma_0}$$

Then $\sigma_0\,du = dz$, and

$$P_B = \int_{u=(a_1-a_2)/2\sigma_0}^{u=\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du = Q\left(\frac{a_1 - a_2}{2\sigma_0}\right) \tag{B.18}$$

where $Q(x)$, called the *complementary error function* or *co-error function,** is tabulated in Table B.1.

Another form of the co-error function which is frequently used is

$$\text{erfc}\,(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp\,(-u^2)\,du \tag{B.19}$$

The two co-error functions, $Q(x)$ and erfc (x), are related as follows:

*Note that the co-error function is defined in several ways; however, all definitions are essentially equivalent.

TABLE

x
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
2.4
2.5
2.6
2.7
2.8
2.9
3.0
3.1
3.2
3.3
3.4

TABLE B.1 Complementary Error Function $Q(x) = \int_x^\infty (1/\sqrt{2\pi}) \exp(-u^2/2)du$

	$Q(x)$									
x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2168	0.2148
0.8	0.2169	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0094	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002

$$\text{erfc}(x) = 2Q(x\sqrt{2}) \tag{B.20}$$

$$Q(x) = \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right) \tag{B.21}$$

REFERENCES

1. Van Trees, H. L., *Detection, Estimation, and Modulation Theory,* Part 1, John Wiley & Sons, Inc., New York, 1968.
2. Papoulis, A., *Probability, Random Variables, and Stochastic Processes,* McGraw-Hill Book Company, New York, 1965.

APPENDIX C

Response of Correlators to White Noise

The inputs to a bank of N correlators represent a white Gaussian noise process, $n(t)$, with zero mean and two-sided power spectral density, $N_0/2$. The output of each correlator at time $t = T$ is a *Gaussian random variable* defined by

$$n_j = \int_0^T n(t)\psi_j(t)\,dt \quad j = 1, \ldots, N \tag{C.1}$$

where $\{\psi_j(t)\}$ forms an orthonormal set. Since n_j is Gaussian, it is characterized completely by its mean and variance. The mean is

$$\overline{n}_j = \mathbf{E}\{n_j\} = \mathbf{E}\left\{\int_0^T n(t)\psi_j(t)\,dt\right\} \tag{C.2}$$

where $\mathbf{E}\{\cdot\}$ is the expected value operator. The variance of n_j is

$$\sigma_j^2 = \mathbf{E}\{n_j^2\} - \overline{n}_j^2 \tag{C.3}$$

$$= \mathbf{E}\left\{\int_0^T n(t)\psi_j(t)\,dt \int_0^T n(s)\psi_j(s)\,ds\right\} - \overline{n}_j^2 \tag{C.4}$$

$$= \int_0^T \int_0^T \mathbf{E}\{n(t)n(s)\psi_j(t)\psi_j(s)\}\,dt\,ds - \overline{n}_j^2 \tag{C.5}$$

Since $n(t)$ is a zero-mean process,

$$\mathbf{E}\{n(t)\} = 0 \tag{C.6}$$

which implies that

$$\bar{n}_j = \mathbf{E}\{n_j\} = 0 \tag{C.7}$$

The autocorrelation function of the process $n(t)$ is

$$R_n(t, s) = \mathbf{E}\{n(t)n(s)\} \tag{C.8}$$

If the noise $n(t)$ is assumed stationary, then $R_n(t, s)$ is a function of only the time difference $\tau = t - s$. From Equation (C.5), we then have

$$\sigma_j^2 = \text{var}\,\{n_j\} = \int_0^T \int_0^T R_n(\tau)\psi_j(t)\psi_j(s)\,dt\,ds \tag{C.9}$$

For a stationary random process, the power spectral density $G_n(f)$ and the autocorrelation function $R_n(\tau)$ form a Fourier transform pair. Thus, we can write

$$R_n(\tau) = \int_{-\infty}^{\infty} G_n(f)e^{j2\pi f\tau}df \tag{C.10}$$

Since $n(t)$ is white noise, its power spectral density $G_n(f)$ is $N_0/2$ for all f, and we can write Equation (C.10) as

$$R_n(\tau) = \int_{-\infty}^{\infty} \frac{N_0}{2} e^{j2\pi f\tau}df = \frac{N_0}{2}\delta(\tau) \tag{C.11}$$

where $\delta(\tau)$ is the unit impulse function defined in Section A.4.1. Substituting Equation (C.11) into Equation (C.9), we get

$$\sigma_j^2 = \frac{N_0}{2}\int_0^T \int_0^T \delta(t - s)\psi_j(t)\psi_j(s)\,dt\,ds \tag{C.12}$$

$$= \frac{N_0}{2}\int_0^T \psi_j^2(t)\,dt = \frac{N_0}{2} \qquad j = 1, \ldots, N \tag{C.13}$$

where we have utilized the *sifting property* of the unit impulse function (see Section A.4.1) and the fact that $\{\psi_j(t)\}$, $j = 1, \ldots, N$, constitutes an orthonormal set. Thus, for white Gaussian noise with two-sided power spectral density $N_0/2$ watts/hertz, the output noise power from each of the N correlators is equal to $N_0/2$ watts.

APPENDIX D

Often-Used Identities

$$\cos x \cos y = \tfrac{1}{2}\cos(x + y) + \tfrac{1}{2}\cos(x - y) \tag{D.1}$$

$$\sin x \sin y = -\tfrac{1}{2}\cos(x + y) + \tfrac{1}{2}\cos(x - y) \tag{D.2}$$

$$\sin x \cos y = \tfrac{1}{2}\sin(x + y) + \tfrac{1}{2}\sin(x - y) \tag{D.3}$$

$$\cos x \sin y = \tfrac{1}{2}\sin(x + y) - \tfrac{1}{2}\sin(x - y) \tag{D.4}$$

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y \tag{D.5}$$

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y \tag{D.6}$$

$$\cos^2 x = \tfrac{1}{2}(1 + \cos 2x) \tag{D.7}$$

$$\sin^2 x = \tfrac{1}{2}(1 - \cos 2x) \tag{D.8}$$

$$\sin x \cos x = \tfrac{1}{2}\sin 2x \tag{D.9}$$

$$\sin x + \sin y = 2\sin\tfrac{1}{2}(x + y)\cos\tfrac{1}{2}(x - y) \tag{D.10}$$

$$\sin x - \sin y = 2\cos\tfrac{1}{2}(x + y)\sin\tfrac{1}{2}(x - y) \tag{D.11}$$

$$\cos x + \cos y = 2\cos\tfrac{1}{2}(x + y)\cos\tfrac{1}{2}(x - y) \tag{D.12}$$

$$\cos x - \cos y = -2\sin\tfrac{1}{2}(x + y)\sin\tfrac{1}{2}(x - y) \tag{D.13}$$

$$\sin x = \frac{e^{jx} - e^{-jx}}{2j} \tag{D.14}$$

$$\cos x = \frac{e^{jx} + e^{-jx}}{2} \tag{D.15}$$

$$P_B = \frac{1}{n}\sum_{j=2}^{n} j\binom{n}{j} p^j(1-p)^{n-j} = p - p(1-p)^{n-1} \tag{D.16}$$

Proof:

$$j\binom{n}{j} = j\frac{n!}{j!\,(n-j)!} = \frac{n!}{(j-1)!\,(n-j)!} = n\frac{(n-1)!}{(j-1)!\,[(n-1)-(j-1)]!}$$

$$= n\binom{n-1}{j-1}$$

$$P_B = \sum_{j=2}^{n}\binom{n-1}{j-1} p^j(1-p)^{n-j} = p\sum_{j=2}^{n}\binom{n-1}{j-1} p^{j-1}(1-p)^{(n-1)-(j-1)}$$

Change of parameter: $i = (j-1)$

Therefore, $(j=2)$ becomes $(i=1)$, and $(j=n)$ becomes $(i=n-1)$.

$$P_B = p\sum_{i=1}^{n-1}\binom{n-1}{i} p^i(1-p)^{(n-1)-i}$$

$$= p\sum_{i=0}^{n-1}\left[\binom{n-1}{i} p^i(1-p)^{(n-1)-i}\right] - p\binom{n-1}{0} p^0(1-p)^{(n-1)-0}$$

$$= p[1-(1-p)^{n-1}]$$

$$= p - p(1-p)^{n-1}$$

APPENDIX E

s-Domain, *z*-Domain and Digital Filtering

Robert W. Stewart
Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow, Scotland, UK

In Appendix A, Equation (A.26) and (A.27) defined the Fourier and inverse Fourier transform, respectively. Although the Fourier transform is useful for a steady-state frequency analysis of a system, it is not particularly useful for transient analysis; and for some functions, the Fourier integral does not in fact exist, whereas the Laplace transform, discussed in this appendix, does. Hence, to allow a more in-depth analysis of linear systems, we often choose to use the Laplace transform. In terms of Laplace and Fourier defining equations, it is straightforward to show that the Laplace transform is an extension of the Fourier transform. If the system under analysis is a discrete time system rather than continuous time system, then we can use the notationally simpler z-transform, which can be directly derived from the Laplace transform. Another reason for using the Laplace transform (for continuous time analysis) and z-transform (for discrete time analysis) is that operations that are "awkward" in the time domain, such as convolution, can be performed more easily in the s-domain or z-domain.

Hence, in this appendix the s-domain, z-domain, and discrete frequency transform are reviewed and thereafter the ubiquitous digital filter is introduced with varied reference to the Laplace- and z-transforms.

E.1 THE LAPLACE TRANSFORM

Recall the Fourier transform in Equation (A.26) of Appendix A, namely,

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\,dt \quad \text{or} \quad X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt \tag{E.1}$$

where $\omega = 2\pi f$.

If we define a new function $v(t)$ that is $x(t)$ multiplied by the real exponential time function $e^{-\sigma t}$, where σ is a real number, then $v(t) = x(t)e^{-\sigma t}$. Hence, the Fourier transform of the function $v(t)$ is

$$V(\omega) = \int_{-\infty}^{\infty} v(t)e^{-j\omega t}\,dt = \int_{-\infty}^{\infty} x(t)e^{-\sigma t}e^{-j\omega t}\,dt = \int_{-\infty}^{\infty} x(t)e^{-(\sigma+j\omega)t}\,dt \quad \text{(E.2)}$$

Therefore, we can rewrite Equation (E.1) as

$$X(\sigma + j\omega) = \int_{-\infty}^{\infty} x(t)e^{-(\sigma+j\omega)t}\,dt \quad \text{(E.3)}$$

Letting s be the complex frequency $s = \sigma + j\omega$, we can now define the *Laplace transform* of a time domain signal $x(t)$ as

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}\,dt \quad \text{(E.4)}$$

where "s" is the Laplace variable. If we rewrite the inverse Fourier transform in Equation (A.27) in terms of angular frequency $\omega = 2\pi f$, then $d\omega/df = 2\pi$ and

$$x(t) = \int_{-\infty}^{\infty} X(\omega)\,e^{j\omega t}\,\frac{d\omega}{2\pi} \quad \text{(E.5)}$$

Since $s = \sigma + j\omega$, it follows that $ds/d\omega = j$, and we can therefore specify the *inverse Laplace transform* as

$$x(t) = \frac{1}{j2\pi}\int_{\sigma-j\infty}^{\sigma+j\infty} X(s)e^{st}\,ds \quad \text{(E.6)}$$

Equations (E.4) and (E.6) are referred to as a *Laplace transform pair* $[x(t) \leftrightarrow X(s)]$, or more precisely, the *two-sided or bilateral Laplace transform pair.* If we (reasonably) assume that the signal $x(t)$ does not exist before $t = 0$ (i.e., it is causal), then the transform can thus be termed *one sided* or *unilateral* and written as

$$X(s) = \int_{0}^{\infty} x(t)e^{-st}\,dt \quad \text{(E.7)}$$

The *inverse unilateral Laplace transform* is the same as Equation (E.6). Hence, it is also appropriate to refer to Equation (E.6) and (E.7) as the *one-sided Laplace transform pair,* or *unilateral Laplace transform pair.*

E.1.1 Standard Laplace Transforms

Table E.1 shows some standard one-sided Laplace transforms. Note that the (two-sided) Laplace transform of Equation (E.4) is identical to the Fourier transform of Equation (A.26) if we set $s = j\omega$, where $\omega = 2\pi f$. In order to develop the Laplace transform, $x(t)$ was multiplied by a "convergence factor" $e^{-\sigma t}$, where σ is any real number. Hence, when actually evaluating the integration, the Laplace transform may exist for many functions that do not have a corresponding Fourier transform.

TABLE E.1 Laplace Transforms

Waveform Type	Time Function	Laplace Transform
Impulse	$\delta(t)$	1
Unit Step function	$u(t)$	$\frac{1}{s}$
Ramp function	$tu(t)$	$\frac{1}{s^2}$
Exponential	$e^{at}u(t)$	$\frac{1}{s-a}$
	$te^{at}u(t)$	$\frac{1}{(s-a)^2}$
Sine wave	$\sin(\omega t)u(t)$	$\frac{\omega}{(s^2+\omega^2)}$
Cosine wave	$\cos(\omega t)u(t)$	$\frac{s}{s^2+\omega^2}$
Damped sine wave	$e^{at}\sin(\omega t)u(t)$	$\frac{\omega}{(s-a)^2+\omega^2}$
Damped cosine wave	$e^{at}\cos(\omega t)u(t)$	$\frac{(s-a)}{(s-a)^2+\omega^2}$

One of the key advantages of the Laplace transform is the capability of transforming functions that are not absolutely integrable.

E.1.2 Laplace Transform Properties

If we know the Laplace transform pair $y(t) \leftrightarrow Y(s)$, then it can be shown that for a delayed version of the signal denoted as $y(t-t_0)$,

$$y(t - t_0) \leftrightarrow e^{-st_0} Y(s) \tag{E.8}$$

This is the simple time-shift property of Laplace transforms. Table (E.2) shows other properties of Laplace transforms. All of these properties can be verified by direct evaluation of the Laplace transform property of interest. Note that the relationship of $s = j\omega$ between the Laplace and Fourier transform means that there is a simple equivalence between like transforms of Table E.1 with those in Table A.1, as well as like operations in Table E.2 with those in Table A.2.

TABLE E.2 Laplace Operations

Property	Time Function	Laplace Transform
General or arbitrary function	$x(t)$	$X(s)$
General or arbitrary function	$y(t)$	$Y(s)$
Linearity	$ax(t) + by(t)$	$aX(s) + bY(s)$
Time shift ($\tau > 0$)	$x(t-\tau)$	$e^{-s\tau}X(s)$
Time scaling	$x(at)$	$\frac{1}{a}X\left(\frac{s}{a}\right)$
Modulation	$e^{-at}x(t)$	$X(s-a)$
Differentiation	$\frac{dx(t)}{dt}$	$sX(s) - x(0)$
Integration	$\int_{-\infty}^{t} x(\tau)d\tau$	$\frac{X(s)}{s}$
Convolution	$x(t)*y(t)$	$X(s)Y(s)$

E.1.3 Using the Laplace Transform

Laplace transforms are useful where (time) differential equations have to be solved or convolution operations have to be performed. For example, to find the current $i(t)$ in the simple RC circuit in Figure E.1, note that the voltage across the capacitor (integrating component) and resistor is the input voltage

$$v_{\text{in}}(t) = i(t)R + \frac{q}{C} = i(t)R + \frac{1}{C}\int_0^t i(t)\,dt \tag{E.9}$$

If the input voltage is a unit-step power source, $v_{\text{in}}(t) = u(t)$, and q is the charge across the capacitor (in Coulombs); taking Laplace transforms of Equation (E.9), using Tables E.1 and E.2, we obtain

$$V_{\text{in}}(s) = RI(s) + \frac{I(s)}{sC} \text{ and } \Rightarrow I(s) = \frac{V_{\text{in}}(s)}{R + 1/(sC)} = \frac{1/R}{s + 1/(RC)} \tag{E.10}$$

(Note that for the unit step, $V_{\text{in}}(s) = 1/s$.) Next, transforming back to the time domain (based again on the Laplace transform tables) gives

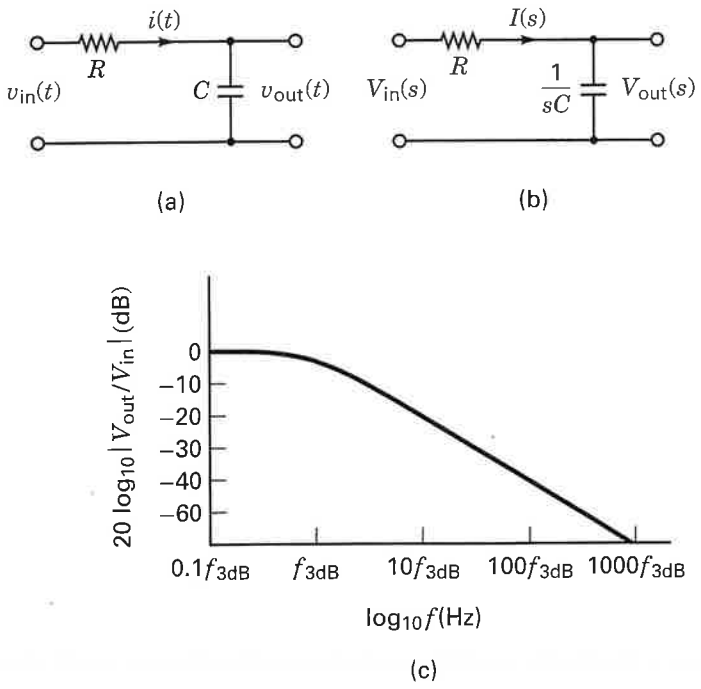


Figure E.1 (a) Resistor capacitor circuit, (b) Laplace transform representation, (c) Magnitude response.

$$i(t) = \frac{1}{R} e^{-t/(RC)} \tag{E.11}$$

E.1.4 Transfer Function

Using the Laplace transform, we can specify the (s-variable) transfer function of a linear system. From Equation (E.10), if the resistance $R = 0$, then the impedance of the capacitor can be calculated as

$$Z_c = \frac{V_{in}(s)}{I(s)} = \frac{1}{sC} \tag{E.12}$$

The input and output Laplace voltages can be specified as

$$V_{in}(s) = I(s)\,R + \frac{I(s)}{sC} \qquad V_{out}(s) = \frac{I(s)}{sC} \tag{E.13}$$

and, therefore, the (Laplace) transfer function can be specified as

$$H(s) = \frac{V_{out}(s)}{V_{in}(s)} = \frac{\frac{I(s)}{sC}}{I(s)R + \frac{I(s)}{sC}} = \frac{1}{sRC + 1} \tag{E.14}$$

E.1.5 RC Circuit Low Pass Filtering

Let the input to the RC circuit be a complex sinusoid $v_{in}(t) = e^{j\omega t}$. From Appendix A, we note that we can equate this to the Fourier transform by setting $s = j\omega$, where $\omega = 2\pi f$. Thus, the frequency response can be realized from the transfer function

$$\frac{V_{out}(f)}{V_{in}(f)} = \frac{1}{j\omega RC + 1} = \frac{1}{j2\pi fRC + 1} = \frac{1}{\sqrt{(2\pi fRC)^2 + 1}} e^{-j[\arctan(2\pi fRC)]} \tag{E.15}$$

For small values of f, then $|H(f)| \approx 1$; and for large values of f, then $|H(f)| \approx 0$. When $f = f_0 = 1/(2\pi RC)$, then $|H(f)| \approx 1/\sqrt{2}$. Noting that $20 \log_{10}(1/\sqrt{2}) = -3$ dB, then f_0 is the "–3-dB" frequency when the output power is half of the input power. Therefore, Equation (E.15) specifies a low-pass filter that is the same as was presented in Equation (1.63). Low frequencies are passed and high frequencies are attenuated; this is shown in the magnitude frequency response of Figure E.1c.

E.1.6 Poles and Zeroes

Linear systems and hence (linear) analog filters can be represented by time-domain differential equations. For example, consider the second-order example

$$y(t) = A\frac{d^2x(t)}{dt^2} + B\frac{dx(t)}{dt} + Cx(t) + D\frac{d^2y(t)}{dt^2} + E\frac{dy(t)}{dt} \tag{E.16}$$

The various orders of differentiation and/or integration are implemented in the real world by using capacitive and inductive components in combination with feedback amplifiers of given orders [2]. Taking the Laplace transform of Equation (E.16) yields the more mathematically (and notationally) convenient Laplace equation

$$Y(s) = As^2X(s) + BsX(s) + CX(s) + Ds^2Y(s) + EsY(s) \tag{E.17}$$

and the transfer function

$$H(s) = \frac{Y(s)}{X(s)} = \frac{As^2 + Bs + C}{-Ds^2 - Es + 1} = \frac{A(s - a_0)(s - a_1)}{-D(s - b_0)(s - b_1)} \tag{E.18}$$

The roots $\{a_0, a_1\}$ of the numerator s-polynomial are termed the *zeroes,* and the roots $\{b_0, b_1\}$ of the denominator s-polynomial are termed the *poles.* Note that if A, B, and C are real values, then the zeroes $\{a_0, a_1\}$ will be complex conjugates.

E.1.7 Linear System Stability

Consider briefly the single-pole equation corresponding to an arbitrary linear system:

$$H(s) = \frac{1}{s - \sigma} \tag{E.19}$$

The impulse response of this circuit can be found (by using Table E.1) to be the inverse of the transfer function in Equation (E.19); and if $\sigma = \rho + j\zeta$, then the impulse response is

$$h(t) = e^{\sigma t} = e^{\rho t} e^{j\zeta t} \tag{E.20}$$

Noting that $\mathrm{Re}[\sigma] = \rho$, then the impulse response is diverging for increasing t (time) if $\rho > 0$. However, for $\rho < 0$, the impulse response is converging for increasing t. The term $e^{j\zeta t}$ is a complex (oscillating) sinusoid (as discussed in Section A.2.1). In slightly different language, we could state that *a system is stable if all s-domain poles have a real component less than zero.*

Thus, if the poles were plotted on the complex s-plane, then all poles must be on the left-hand side of the complex plane. Figure E.2 illustrates the complex s-plane regions of stability and an example of a *stable* 3rd-order transfer function, where the poles are in the left-hand side of the plane, i.e. the poles have a negative real component. Note that the zeroes may be in either the left- or right-hand side of the s-plane and have no bearing on stability.

For a circuit with more than one pole, one can think of the transfer function as a cascade of single-pole circuits, as follows:

$$H(s) = \frac{(s-a_0)(s-a_1)(s-a_2)}{(s-b_0)(s-b_1)(s-b_2)} = (s-a_0)(s-a_1)(s-a_2)\left[\frac{1}{s-b_0}\right]\left[\frac{1}{s-b_1}\right]\left[\frac{1}{s-b_2}\right] \tag{E.21}$$

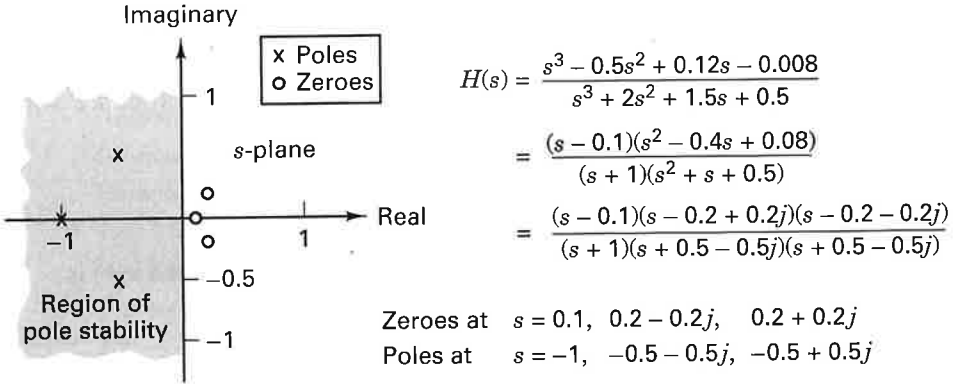


Figure E.2 Plotting the poles and zeroes of an s-domain transfer function.

For stability, all poles must lie on the left-hand side of the complex plane. Note that for real circuits with real Laplace coefficients (i.e. in Equation (E.16) A, B, C, D, E would be all real valued), the poles and zeroes will either be real or occur as complex conjugate pairs as illustrated in Figure (E.2).

For our previous example of an RC circuit, the transfer function in Equation (E.14) is unconditionally stable since $2\pi RC$ is always a positive value—which is of course the expected result. Instability will only be seen in linear systems that have feedback (recursion), such as filters implemented with inverting or noninverting operational amplifier circuits.

E.2 THE z-TRANSFORM

The z-transform is essentially the discrete equivalent of the Laplace transform. It allows for conventional mathematical analysis (transient and steady state) and manipulation of signals and systems. Probably the most frequent modern use of the z-transform is for specifying discrete systems and analyzing their stability.

The z-transform allows the convolution of an input signal and a discrete linear system to be calculated in a mathematically tractable form. Also, the "poles" and "zeroes" of a system can be observed, leading to information regarding the dynamic behavior and stability of a discrete system. It is important to note that the poles and zeroes of the z-transform are different from the poles and zeroes of the Laplace transform.

E.2.1 Calculating the z-Transform

We can derive the z-transform from the Laplace transform of Equation (E.4) by considering a signal $x(t)$ that is sampled every T seconds, thus yielding a discrete time signal of samples $x(0), x(T), x(2T), \ldots = \{x(kT)\}$. The sampled data represent a set of weighted and translated delta functions whose Laplace transform (by the time shifting property) can be written as

$$X(s) = \sum_{k=0}^{\infty} x(kT)e^{-skT} \tag{E.22}$$

If we employ the parameter $z = e^{sT}$ and replace discrete time kT with a sample number k, we get

$$X(z) = \sum_{k=0}^{\infty} x(k)z^{-k} \tag{E.23}$$

As an example, the z-transform of a simple unit step function $u(k)$ is

$$U(z) = \sum_{k=0}^{\infty} u(k)z^{-k} = 1 + z^{-1} + z^{-2} + z^{-3} + \ldots = \frac{1}{1 - z^{-1}} \tag{E.24}$$

Note the geometric series form and the assumption that $|z| < 1$ (the region of convergence). As in the case of the Laplace transform, the z-transform can be

TABLE E.3 z-Transforms of Some Simple Functions

Waveform Type	Time Function	z-Transform
Impulse	$\delta(k)$	1
Delayed impulse	$\delta(k-m)$	z^{-m}
Unit step function	$u(k)$	$\frac{z}{z-1}$
Ramp function	$ku(k)$	$\frac{z}{(z-1)^2}$
Exponential	$e^{ak}u(k)$	$\frac{z}{z-e^a}$
Sine wave	$\sin(\omega k)u(k)$	$\frac{z\sin(\omega)}{z^2-2z\cos(\omega)+1}$
Cosine wave	$\cos(\omega k)u(k)$	$\frac{z[z-\cos(\omega)]}{z^2-2z\cos(\omega)+1}$

tabulated in tables, such as Table E.3, and operations of the z-transform can be tabulated in tables, such as Table E.4.

E.2.2 The Inverse *z*-Transform

We can transform back from the z-domain to the time domain using the inverse z-transform [2], as follows:

$$x(k) = z^{-1}\{X(z)\} = \frac{1}{j2\pi}\oint_C X(z)\, z^{k-1} dz \tag{E.25}$$

TABLE E.4 z-Transform Operations

Property	Time Function	Z Transform
General or arbitrary function	$x(t)$	$X(z)$
General or arbitrary function	$y(t)$	$Y(z)$
Linearity	$ax(t)+by(t)$	$aX(z)+bY(z)$
Time shift	$x(k-m)$	$z^{-m}X(z)$
Modulation	$e^{-j\omega k}x(k)$	$X(e^{j\omega}z)$
Exponential scaling	$a^k x(k)$	$X(z/a)$
Ramp scaling	$kx(k)$	$-z\frac{d}{dz}X(z)$
Convolution	$x(k)*h(k)$	$X(z)H(z)$

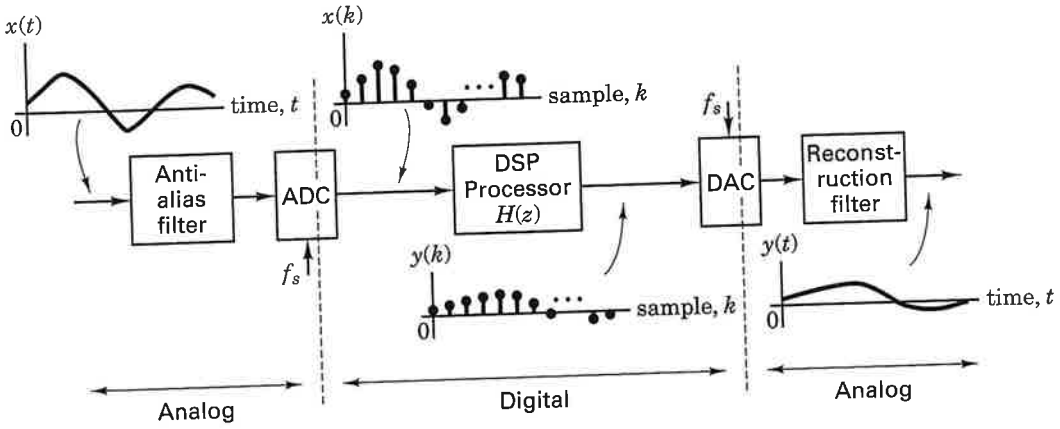


Figure E.3 Digital filter equations are implemented on the DSP Processor which processes the time sampled data signal to produce a time sampled output data signal.

Here, the complex integration $\oint$ is over any simple contour in the region of convergence of $X(z)$ that circles the point $z = 0$. The evaluation of the inverse z-transform would appear to be somewhat more complicated than for the z-transform. Typical operations required include partial fraction evaluation, polynomial division, the residue theorem, and difference equation synthesis. Hence, most z-transformations, and inverse z-transformations are calculated using tables of standard integrals and properties, and explicit evaluation of Equation (E.25) can usually be avoided. In modern-day DSP analysis, software packages such as SystemView [1] are used to manipulate discrete signals and systems, and the z-transform is largely used simply as a convenient analytical notation for the specification and stability of discrete signals and systems.

E.3 DIGITAL FILTERING

Using appropriate analog and digital components, a digital filter can be set up to perform a desired frequency discrimination or phase-modifying function. Figure E.3 shows the components required for a digital filter that produces the filtered sequence $y(k)$ for the input sequence $x(k)$ [2]. A general digital filter output $y(k)$ is produced from a weighted sum of past inputs $x(k)$ and past outputs $y(k - n)$, with $n > 0$. A signal flow graph (composed of only adders, multipliers and sample delays), for a four-feedforward weight and three-weight feedback digital filter, is shown in Figure E.4. (Note that a single sample time delay is indicated by the symbol Δ. Often, one sees such time-domain signal flow graphs drawn with a mixture of time-domain and z-domain notation, using z^{-1} to represent a delay; although it is a common practice, it is not precise.

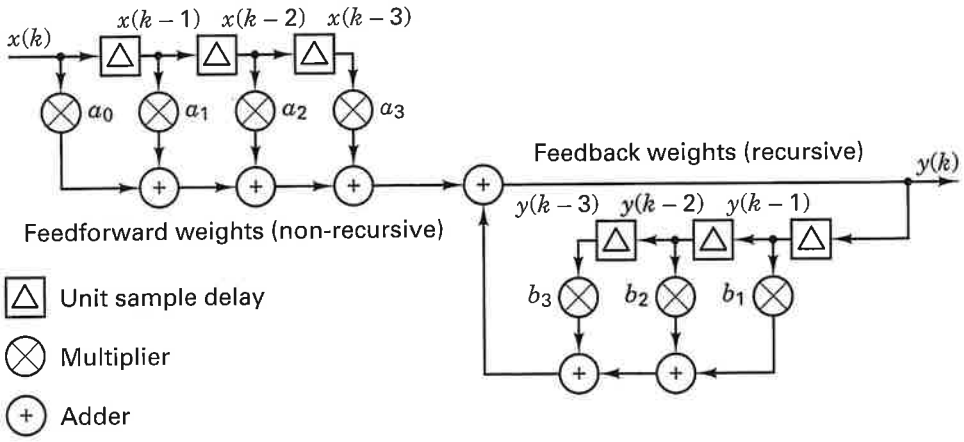


Figure E.4 A generic digital filter.

The output of this filter is given as

$$\begin{aligned} y(k) &= a_0x(k) + a_1x(k-1) + a_2x(k-2) + a_3x(k-3) + \\ &\quad + b_1y(k-1) + b_2y(k-2) + b_3y(k-3) \\ &= \sum_{n=0}^{3} a_nx(k-n) + \sum_{m=1}^{3} b_my(k-m) \end{aligned} \tag{E.26}$$

Taking the z-transform of Equation (E.26) gives

$$\begin{aligned} Y(z) &= a_0X(z) + a_1X(z)z^{-1} + a_2X(z)z^{-2} + a_3X(z)z^{-3} \\ &\quad + b_1\,Y(z)z^{-1} + b_2Y(z)z^{-2} + b_3\,Y(z)z^{-3} \end{aligned} \tag{E.27}$$

E.3.1 Digital Filter Transfer Function

The transfer function of the digital filter in Figure E.4 is produced by rearranging Equation (E.27) to yield

$$\begin{aligned} H(z) &= \frac{Y(z)}{X(z)} = \frac{a_0 + a_1z^{-1} + a_2z^{-2} + a_3z^{-3}}{1 - b_1z^{-1} - b_2z^{-2} - b_3z^{-3}} \\ &= \frac{a_0(1 - \alpha_1z^{-1})(1 - \alpha_2z^{-1})(1 - \alpha_3z^{-1})}{(1 - \beta_1z^{-1})(1 - \beta_2z^{-1})(1 - \beta_3z^{-1})} \\ &= \frac{a_0(z - \alpha_1)(z - \alpha_2)(z - \alpha_3)}{(z - \beta_1)(z - \beta_2)(z - \beta_3)} = \frac{A(z)}{B(z)} \end{aligned} \tag{E.28}$$

where the α values are the z-domain *zeroes* of the filter and the β values are the z-domain *poles* of the filter, found by finding the roots of the numerator polyno-

mial $A(z) = 0$ and the denominator polynomial $B(z) = 0$, respectively. For a digital filter similar to the signal flow graph in Figure E.4 but with N feedforward weights and $M - 1$ feedback weights, the numerator and denominator polynomials will be of order N and M, respectively, in the transfer function analogous to Equation (E.28).

E.3.2 Single Pole Filter Stability

Due to the presence of numerical feedback in the signal flow graph, the digital filter may be (numerically) unstable. For example, consider the single-feedback weight filter of Figure E.5:

$$y(k) = x(k) + by(k - 1) \tag{E.29}$$

The impulse response of this filter (i.e., applying a discrete unit impulse $\delta(k)$, following the principles of convolution described in Section A.5) is

$$h(k) = b^k \tag{E.30}$$

If $|b| < 1$, then the filter impulse response is converging (stable); and if $|b| > 1$, the filter impulse response is diverging (unstable). Figure E.5 illustrates a converging impulse response where $|b| < 1$, and, more precisely, $-1 < b < 1$. Taking the z-transform of Equation (E.29) yields

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - bz^{-1}} = \frac{z}{z - b} \tag{E.31}$$

From Equation (E.31), we show the z-domain signal flow graph in Figure E.5b which corresponds to the time domain signal flow graph of Figure E.5a. The delay elements in Figure E.5b (represented by Δ in Figure E.5a) are now represented by z^{-1}, and the inputs and outputs are specified as the z-transforms $X(z)$ and $Y(z)$. Note, however, that the general topology of the two signal flow graphs is the same. (This perhaps indicates why digital filter signal flow graphs are often drawn (imprecisely) with a mixed time-domain and z-domain notation). We can state the

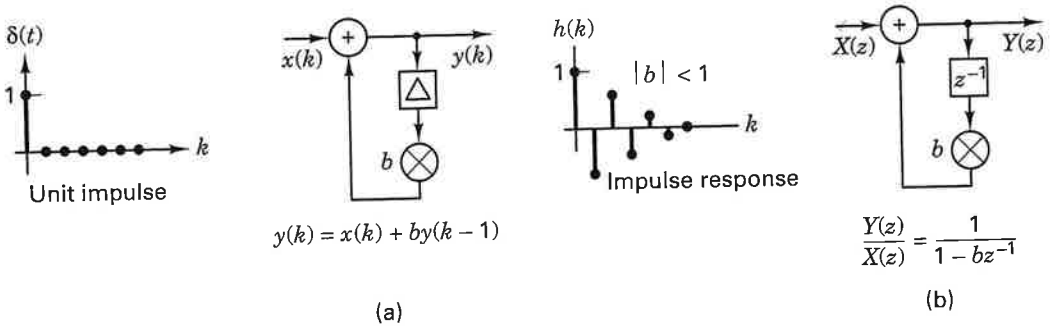


Figure E.5 (a) A single feedback weight filter time domain signal flow graph, and (b) the z-domain signal flow graph equivalent.

stability criteria of $|b| < 1$ as follows: *for stability, the magnitude of the poles (or roots of the denominator polynomial) of the digital filter transfer function must be less than one.*

E.3.3 General Digital Filter Stability

From the factorized transfer function in Equation (E.28), we can redraw the (time domain) signal flow graph of Figure E.4 and the z-domain signal flow graph of Figure E.6. This signal flow graph essentially corresponds to rewriting Equation (E.28) in the form

$$H(z) = a_0(1 - \alpha_1 z^{-1}) \cdot (1 - \alpha_2 z^{-1}) \cdot (1 - \alpha_3 z^{-1}) \cdot \left[\frac{1}{1 - \beta_1 z^{-1}}\right] \cdot \left[\frac{1}{1 - \beta_2 z^{-1}}\right] \cdot \left[\frac{1}{1 - \beta_3 z^{-1}}\right] \tag{E.32}$$

This explicitly shows the first-order sections for each zero and pole of the filter. In order for the filter to be stable, then all of the pole values $\{\beta_1, \beta_2, \beta_3\}$ in the cascade must have a magnitude of less than 1. If any one first-order section is unstable (or diverging), then so is the entire cascade. As was noted for the Laplace transform, the z-domain poles (and zeroes) may be complex, which is the reason that magnitude rather than amplitude is the stability criterion. (Note that the signal flow graph implementation of Figure E.6 is only for analysis purposes, and a digital filter would never be practically implemented in this factored form because some of the multiplier values may be complex and thus computational requirements would be unnecessarily increased for the implementation of a real coefficient filter.)

E.3.4 z-Plane Pole-Zero Diagram and the Unit Circle

If the complex poles and zeroes of a filter or linear system are plotted on the real and imaginary axes, the resulting plane can be referred to as the z-plane. We can verify the stability of a system by observing that all poles lie within the *unit circle.* Figure E.7 shows the z-plane for a filter with transfer function

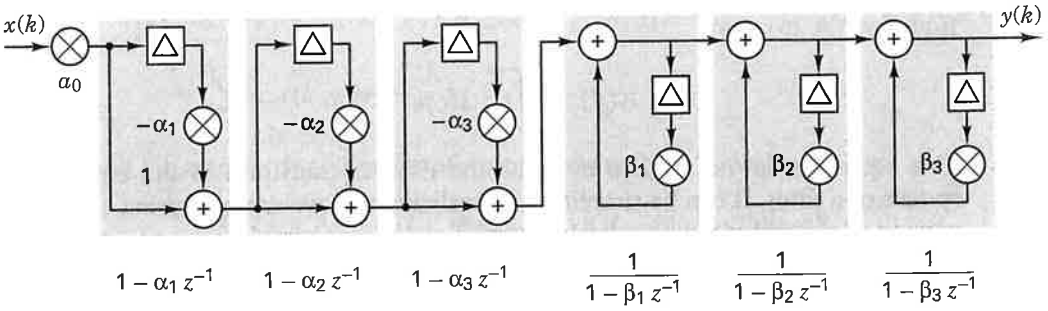


Figure E.6 Digital filter as a cascade of feedforward and feedback first order sections.

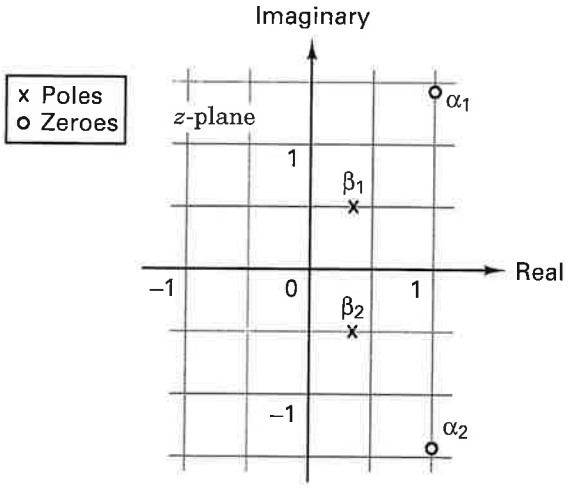


Figure E.7 Poles and zeroes plotted on the z-plane.

$$H(z) = \frac{1 - 2z^{-1} + 3z^{-2}}{1 - \frac{2}{3}z^{-1} + \frac{1}{3}z^{-2}}$$

$$= \frac{(1 - (1 + j\sqrt{2})z^{-1})(1 - (1 - j\sqrt{2})z^{-1})}{(1 - (1/3 + j\sqrt{2}/3)z^{-1})(1 - (1/3 - j\sqrt{2}/3)z^{-1})} \tag{E.33}$$

$$= \frac{(1 - \alpha_1 z^{-1})(1 - \alpha_2 z^{-1})}{(1 - \beta_1 z^{-1})(1 - \beta_2 z^{-1})}$$

with zeroes at $z = 1 + j\sqrt{2}$ and $z = 1 - j\sqrt{2}$ and poles at $z = 1/3 + j\sqrt{2}/3$ and $z = 1/3 - j\sqrt{2}/3$. This filter is clearly stable, as all poles lie within the unit circle.

E.3.5 Discrete Fourier Transform of Digital Filter Impulse Response

The frequency response of a digital filter is calculated from the discrete Fourier transform (DFT) of the filter impulse response. Recall, from the Fourier transform Equation (A.26), that

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\,dt \tag{E.34}$$

This equation can be used to evaluate the Fourier transform of the impulse response of a filter. It can be simplified by realizing that we are now using a sampled version of a signal $x(t)$ with samples every $T_s = 1/f_s$ seconds:

$$X(f) = \int_{-\infty}^{\infty} x(kT_s)e^{-j2\pi fkT_s}d(kT_s) = \sum_{k=-\infty}^{\infty} x(kT_s)e^{-j2\pi fkT_s} = \sum_{k=-\infty}^{\infty} x(kT_s)e^{-(j2\pi fk)/f_s} \tag{E.35}$$

Of course, the digital filter impulse response is causal and the first sample of the impulse response is at $k = 0$, and the last sample is at $k = N - 1$, giving a total of N samples in the transform. Hence, for this finite number of samples, rewriting Equation (E.35) with respect to the sample number k rather than explicit time kT_s gives

$$X(f) = \sum_{k=0}^{N-1} x(k)e^{-(j\,2\pi fk)/f_s} \tag{E.36}$$

Note that Equation (E.36) is actually calculated for a continuous frequency variable f. In reality we need only evaluate this equation at certain discrete frequencies, which are the zero frequency (dc) and harmonics of the "fundamental" frequency, for a total of N discrete frequencies $0, f_0, 2f_0$, up to f_s, where $f_0 = 1/NT_s$:

$$X\left(\frac{nf_s}{N}\right) = \sum_{k=0}^{N-1} x(k)e^{-(j\,2\pi kf_s n)/Nf_s} \quad \text{for } n = 0 \text{ to } N-1 \tag{E.37}$$

Simplifying to use only the time index k and the frequency index n gives the *discrete Fourier transform (DFT)*

$$X(n) = \sum_{k=0}^{N-1} x(k)e^{-(j\,2\pi kn)/N} \quad \text{for } n = 0 \text{ to } N-1 \tag{E.38}$$

Given that the discrete signal $x(k)$ was sampled at the rate of f_s samples/s, then the signal has image (or alias) components above $f_s/2$. Hence, when evaluating Equation (E.38), it is only necessary to evaluate up to $f_s/2$. Note that Equation (E.38) is the same as the z-transform of Equation (E.23) if we set $z = e^{(j2\pi n)/N}$ for a sequence of length N samples.

E.4 FINITE IMPULSE RESPONSE FILTER DESIGN

By far the most common type of digital filter is the finite impulse response (FIR) filter, which as the name suggests has an impulse response of finite duration. This filter has no feedback weights (recall Figure E.4) and therefore we can conclude that it is unconditionally stable. The output of the FIR filter in Figure E.8 is

$$\begin{aligned} y(k) &= a_0 x(k) + a_1 x(k-1) + a_2 x(k-2) + a_3 x(k-3) + \cdots + a_{N-1} x(k-N+1) \\ &= \sum_{n=0}^{N-1} a_n x(k-n) \end{aligned} \tag{E.39}$$

and the transfer function therefore contains only zeros and no poles:

$$\begin{aligned} H(z) &= a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + \cdots + a_{N-1} z^{-N+1} \\ &= a_0(1 - \alpha_1 z^{-1})(1 - \alpha_2 z^{-1})(1 - \alpha_3 z^{-1}) \ldots (1 - \alpha_N z^{-1}) \end{aligned} \tag{E.40}$$

The FIR filter is essentially a moving average calculation, whereby the output is a weighted average of the last N input samples. Hence, this type of filter is often

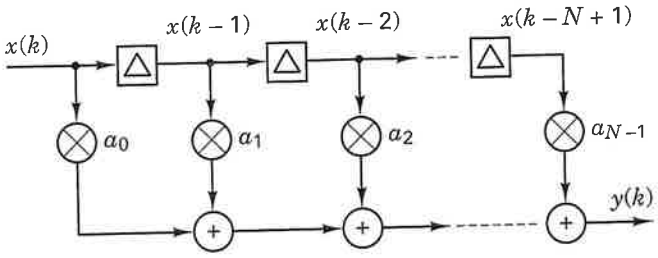


Figure E.8 Finite impulse response digital filter.

termed a *moving average filter.* Other names include *tapped delay line,* and *transversal filter.*

E.4.1 FIR Filter Design

Using modern DSP analysis software, such as SystemView [1], FIR digital filters are designed based on a magnitude frequency plot with specified tolerances and user requirements, as illustrated in Figure E.9 for low-pass filter. Classical filter design techniques, such as Parks-McLellan, Remez Exchange, Kaiser window and so on [4], are then used to realize a suitable frequency response that satisfies the user requirements with a minimum number of weights. Unless otherwise specified in the design process, most FIR filters are designed to have linear phase or a constant group delay (corresponding to a symmetrical impulse response).

Figure E.10 shows the impulse response and frequency response of a digital filter design with parameters of cut-off frequency 1000 Hz, stopband attenuation of 20 dB, passband ripple of 3 dB and transition band of 500 Hz; the sampling frequency, f_s is 10,000 Hz. If a filter with a more stringent frequency response is required (such as more stopband attenuation) then it is likely that the filter design procedure will produce an FIR with more filter weights [4].

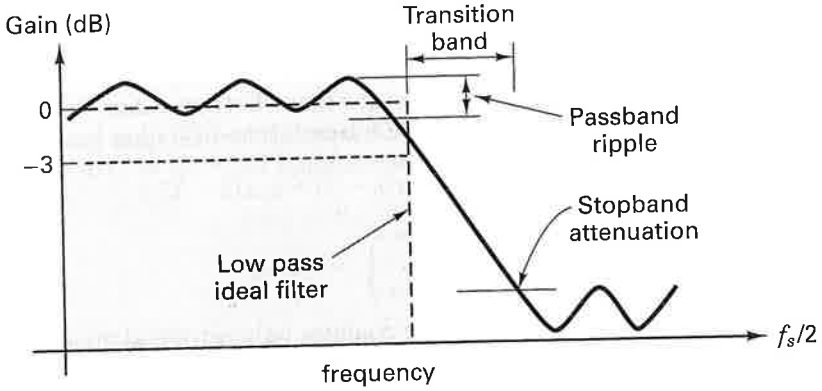


Figure E.9 Generic low pass filter magnitude response. The more stringent the filter requirements of stopband attenuation, transition bandwidth and to a lesser extent passband ripple, the more weights that are required.

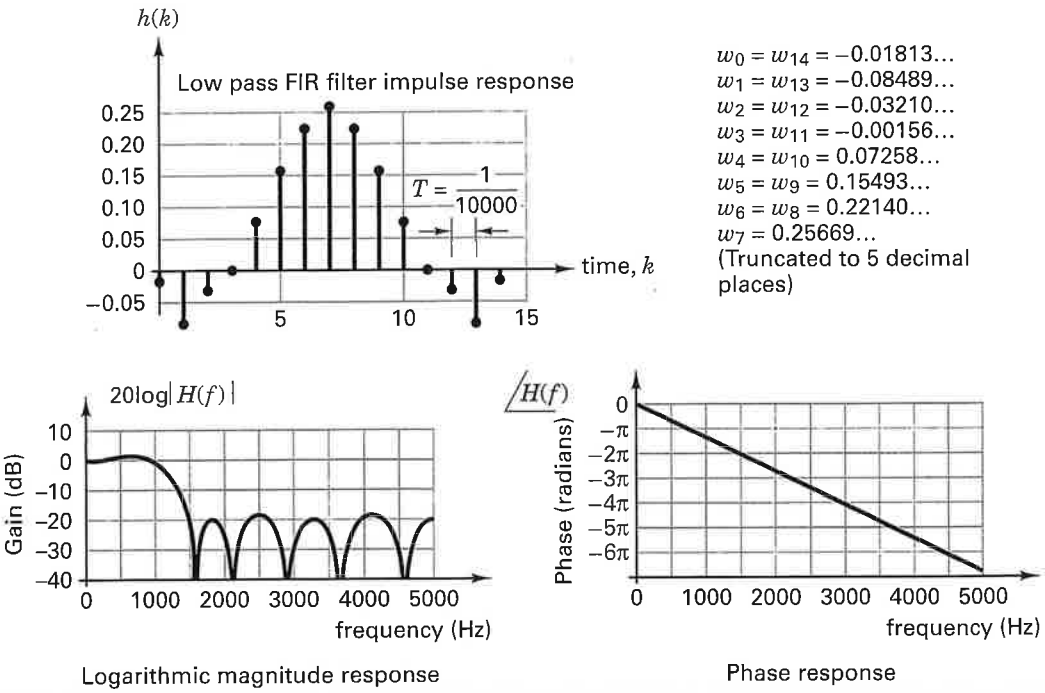


Figure E.10 The impulse response $h(n) = w_n$ and the frequency response $H(f)$ of a low-pass filter with 15 weights, a sampling rate of 10,000 Hz, designed from specified cut-off frequency designed at around 1000Hz.

E.4.2 The FIR Differentiator

Consider the simple digital filtering *differentiator* shown in Figure E.11. We can intuitively reason that this filter is a high-pass filter by observing the output for both low- and high-frequency input sinusoids. The output sequence of this filter is given by

$$y(k) = [x(k) - x(k-1)] \tag{E.41}$$

Taking the z-transform of Equation (E.41) gives

$$Y(z) = [X(z) - X(z)z^{-1}] \tag{E.42}$$

Therefore, the transfer function is given by

$$\frac{Y(z)}{X(z)} = (1 - z^{-1}) \tag{E.43}$$

Figure E.12 illustrates why this filter acts to produce a high-pass function. Essentially the output of the filter is the difference between the last two samples. If the

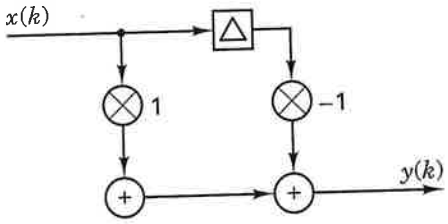


Figure E.11 Differentiator/high pass filter.

difference between successive samples was small (as for a low frequency), then the output is small. If the difference is large (as for a high frequency), then the output is large. If a dc signal was input, then the output amplitude would be zero, i.e. infinite attenuation. The frequency response can also be calculated from the Fourier transform of the impulse response.

If the filter weights are changed to $\{1/T, -1/T\}$ from $\{1, -1\}$ where the sampling frequency is $f_s = 1/T$, then, for low-frequency inputs, $y(k)$ is approximately the differential of the input:

$$y(k) \approx \frac{x(k) - x(k-1)}{T} \approx \frac{dx(t)}{dt} \quad \text{and} \quad \frac{Y(z)}{X(z)} = \frac{1}{T}(1 - z^{-1}) \qquad \text{(E.44)}$$

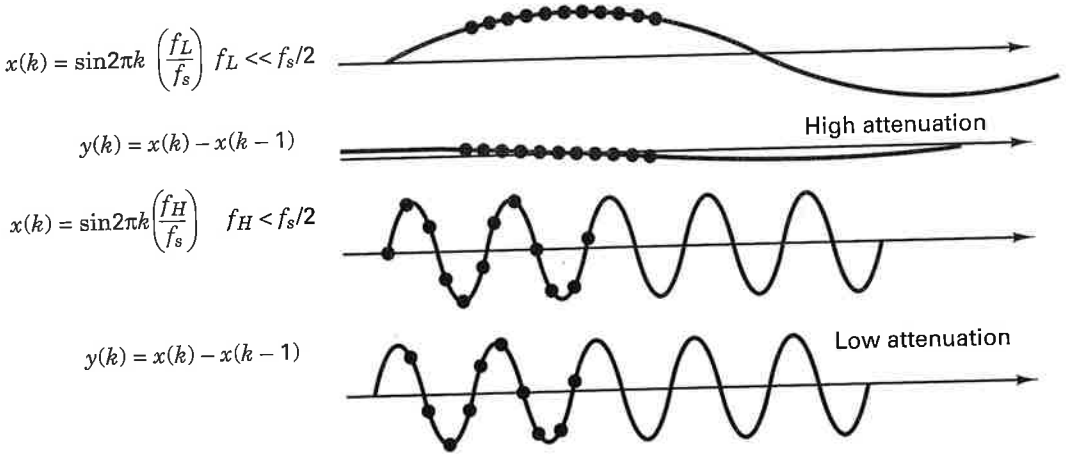


Figure E.12 Digital filter differentiator acting as a high pass filter.

E.5 INFINITE IMPULSE RESPONSE FILTER DESIGN

Infinite impulse response (IIR) filters are usually designed from analog prototypes using a mapping from the s-plane to the z-domain. As their name would suggest, the impulse response (assuming infinite precision arithmetic) can be infinite in duration. IIR filters have both feedforward and feedback weights, as was shown in Figure E.4. IIR filters can have a very long impulse response for just a few weights, due to the recursive nature of the signal flow graph. Hence, it may be possible to design an IIR filter with fewer weights than an FIR filter, for the same functional magnitude response. In general, an IIR digital filter will not have linear phase.

E.5.1 Backward Difference Operator

Equation (E.44) represents a means of relating the (continuous time) Laplace transform variable s to the (discrete time) z-transform variable z. Given that the Laplace domain representation of time differentiation (d/dt) is the variable s, for example,

$$y(t) = \frac{dx(t)}{dt} \quad \Rightarrow Y(s) = sX(s) \tag{E.45}$$

then, given, for instance, a general low-pass Butterworth characteristic

$$H(s) = \frac{1}{s^2 + \sqrt{2}\,s + 1} \tag{E.46}$$

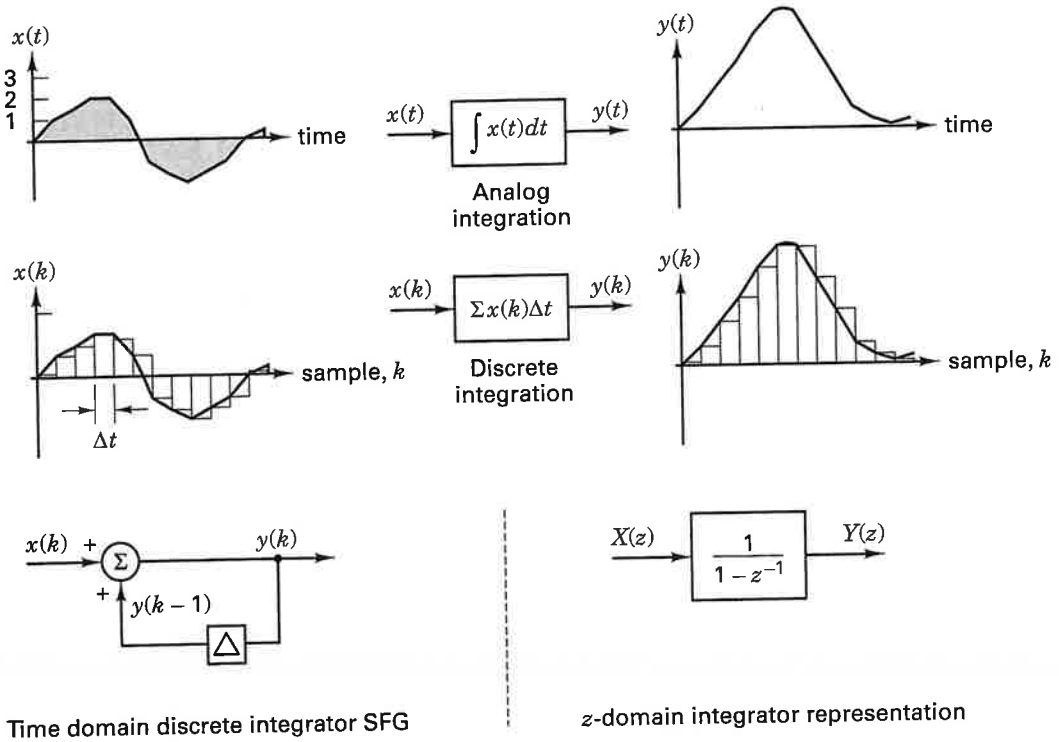
we could produce a discrete approximation to this analog circuit by substituting the approximation

$$s \approx \frac{1}{T}(1 - z^{-1}) \tag{E.47}$$

into Equation (E.46) to produce a z-domain equation

$$\begin{aligned} H(z) = H(s)\Big|_{s=\frac{1}{T}(1-z^{-1})} &= \frac{1}{\frac{1}{T^2}(1 - z^{-1})^2 + \sqrt{2}\frac{1}{T}(1 - z^{-1}) + 1} \\ &= \frac{T^2}{(1 - 2z^{-1} + z^{-2}) + \sqrt{2}\,T(1 - z^{-1}) + T^2} \\ &= \frac{T^2}{z^{-2} - (\sqrt{2}T + 2)z^{-1} + (1 + \sqrt{2}T + T^2)} \end{aligned} \tag{E.48}$$

At low frequencies, where the approximation of Equation (E.47) is "good," then it may be that this transform produces a "reasonable" digital filter equivalent of the analog low-pass Butterworth. (Equation (E.47) is sometimes called the "backward



Time domain discrete integrator SFG

z-domain integrator representation

Figure E.13 Single pole filter acting as an integrator. Note that a feedback weight of just less than 1 is often included in the feedback loop to introduce some "forgetting" into the integrator.

difference operator"). Unfortunately, this mapping is very poor at high frequencies, and consequently, it cannot be used to produce high-pass filters. Hence, it is rarely used.

E.5.2 IIR Filter Design using the Bilinear Transform

The bilinear transform is obtained by replacing s with the approximation

$$s \approx \frac{2}{T}\frac{(1 - z^{-1})}{(1 + z^{-1})} \tag{E.49}$$

This substitution provides a mapping that preserves the stability of the analog prototype and will produce filters that are greatly improved over the backward difference operator of Equation (E.47) [2]. SystemView [1] uses the bilinear transform to produce digital filters from standard analog prototypes, such as Butterworth, Elliptic, and Chebychev. Note that the bilinear transform will always yield a filter that has both poles and zeroes, and it therefore represents an infinite impulse response (IIR) design.

E.5.3 The IIR Integrator

A digital *integrator* is essentially a one-weight IIR filter:

$$y(k) = x(k) + y(k-1) = \sum_{i=0}^{k} x(i) \tag{E.50}$$

In the z-domain, the transfer function of a discrete integrator is obtained from

$$Y(z) = X(z) + z^{-1} Y(z) \tag{E.51}$$

yielding

$$\frac{Y(z)}{X(z)} = \frac{1}{1 - z^{-1}} \tag{E.52}$$

Figure E.13 shows the implementation of the simple digital integrator and a graphical representation of the relationship with continuous time integration (i.e., area under the graph).

If a weight value of less than one is included in the feedback loop (say 0.99), then the integrator is often called a *leaky integrator.* When viewed in the frequency domain, a (leaky) integrator has the characteristics of a simple low-pass filter.

REFERENCES

1. SystemView DSP Communications Software. Elanix, Westlake Village, CA, 2000.
2. B. Porat. A Course in Digital Signal Processing. John Wiley and Sons, 1997.
3. T. K. Moon, W. C. Stirling. Mathematical Methods and Algorithms for Digital Signal Processing. Prentice Hall, 2000.
4. R. W. Stewart. The DSPedia: A Multimedia Resource for DSP, BlueBox Multimedia, UK, 2000.

APPENDIX F

List of Symbols

a_{ij}	Coefficient of jth basis function
a_j	Signal component output of jth correlator
A	Peak amplitude of a waveform
A_e	Effective area (of an antenna)
B_L	Single-sided loop bandwidth
c	Speed of light $\approx 3 \times 10^8$ m/s
C	Channel capacity
C	Electrical capacitance
$C/_kT°$	Ratio of average carrier power to noise power spectral density
d	Distance
d_0	Reference distance
d_f	Free distance
$d_{\min}$	Minimum distance
D	Delay time (of message)
D	Redundancy of a language
D	Decryption transformation
e	The natural number 2.7183
$\mathbf{e}$	Error pattern vector
$e(t)$	Error signal
$\mathbf{e}(X)$	Error pattern polynomial
E	Encryption transformation
E_x	Energy of waveform $x(t)$
$\mathbf{E}[X]$	Expected value of the random variable X
EIRP	Effective radiated power with reference to an isotropic source
E_b/J_0	Ratio of bit energy to jammer power spectral density
E_b/N_0	Ratio of bit energy to noise power spectral density
E_c/N_0	Ratio of channel symbol energy to noise power spectral density
f	Frequency (hertz)
f_c	Carrier-wave frequency
f_m	Maximum frequency
f_0	Coherence bandwidth
f_d	Doppler frequency spread
f_s	Sampling frequency
f_l	Lower cutoff filter frequency
f_u	Upper cutoff filter frequency
F	Noise figure
$\mathcal{F}\{x\}$	Fourier transform of the function $x(t)$
$\mathcal{F}^{-1}\{X\}$	Inverse Fourier transform of the function $X(f)$
F	Field
F^*	Finite field
$g(t)$	Pseudorandom code function
$g(t)$	Signal waveform (baseband)
$\mathbf{g}(X)$	Generator polynomial (for a cyclic code)
G	Antenna gain
G	Coding gain
$\mathbf{G}$	Generator matrix (for a linear block code)
G	Normalized total message traffic
G_p	Processing gain
$G_x(f)$	Power spectral density of waveform $x(t)$
$h(t)$	Impulse response of a network
$h_c(t)$	Channel impulse response
$\mathbf{H}$	Parity-check matrix for a code
H_i	The ith hypothesis
$\mathbf{H}_k$	Hadamard matrix
$H(f)$	Frequency transfer function of a network
$H_0(f)$	Optimum frequency transfer function
$H(X)$	Entropy of information source X
$H(X\|Y)$	conditional entropy (entropy of X, given Y)
$i(t)$	Electrical current waveform
I	Electrical current
$I_0(x)$	Zero-order modified Bessel function of the first kind
$I(X)$	Self-information of information source X
J	Received average jammer power
J_0	Jammer power spectral density
J/S	Ratio of received average jammer power to average signal power
k	Number of bits per M-ary signal set
k/n	Code rate (ratio of number of data bits to total bits in codeword)
K	Constraint length of a convolutional encoder
K	Key, dictating a specific encryption or decryption transformation
ℓ	Number of quantization bits
$\ell(d_k)$	Likelihood of data bit d_k
L	Look-ahead length for convolutional feedback decoding
L	Number of branch words in sequence
L	Number of quantization levels
$L(d_k)$	Log-likelihood ratio of data bit d_k
L_e	Extrinsic log-likelihood ratio
L_s	Free space path loss
L_o	Other losses
L_p	Path loss

L_o	Observation time
L_c	Observation time for channel-induced ISI
L_{CISI}	Observation time for controlled ISI
L_c	Channel log-likelihood ratio
$\mathbf{m}$	Message vector
$\mathbf{m}(X)$	Message polynomial
m_i	Data bit
M	Margin
M	Waveform- or signal-set size
(n, k)	Code designation by number of total bits (n) and data bits (k) in codeword
$\bar{n}$	Average number of bits per character
n_0	Noise random variable output of correlator at symbol time $t = T$
$n(t)$	Gaussian noise process
N	Noise power
N	Unicity distance
N_0	Level of single-sided power spectral density of white noise
NSR	Ratio of average noise power to average signal power
p_c	Probability of channel symbol error
p_i	parity bit
$p(t)$	Instantaneous power
$p(x)$	Probability density function of a continuous random variable
$p(x\|y)$	Probability density function of x conditioned on y
$\mathbf{P}$	Parity array
P_B	Probability of bit error
P_E	Probability of symbol error
P_{FA}	Probability of false alarm
P_m	Probability of miss
P_M	Probability of message or block error
P_{nd}	Probability of undetected error
P_r/N_0	Ratio of received average signal power to noise power spectral density
$\mathbf{p}(X)$	Remainder polynomial
$P(X)$	Probability of a discrete random variable
P_x	Average power in waveform $x(t)$
q	Quantization step size (quantile interval)
$\mathbf{q}(X)$	Quotient polynomial
$Q(x)$	Complementary error function (integral of the tail beyond x of the Gaussian density function)
r	Filter roll-off factor
r	True rate of a language
r'	Absolute rate of a language
$r(t)$	Received signal waveform
R	Data rate (bits/second)
$R(\Delta f)$	Spaced-frequency correlation function
$R(\Delta t)$	Spaced-time correlation function
R_c	Code-bit rate or channel-bit rate (code bits/second)
R_{ch}	Chip rate (chips/second)
R_s	Symbol rate (symbols/second)
$R_x(\tau)$	Autocorrelation function of waveform $x(t)$
$\mathcal{R}$	Electrical resistance
$s(t)$	Signal waveform
$\hat{s}(t)$	Estimate of signal waveform
$S(v)$	Doppler power spectral density
$S(\tau)$	Multipath intensity profile
$\mathbf{s}$	Signal vector
sgn x	Sign function of x
S_k	State at time k
S	Signal power
$\mathbf{S}$	Syndrome vector
SJR	Ratio of average signal power to average jammer power
SNR	Ratio of average signal power to average noise power
S/N	Ratio of signal power to noise power
$S(f)$	Fourier transform of the waveform s(t)
$\mathbf{S}(X)$	Syndrome polynomial
t	Number of errors correctable in an error-correcting code
t	Independent time variable
t_0	Time delay
t_{ij}	Amount of message traffic from i to j
T	Pulse width
T	Symbol interval
$T(D)$	Transfer function or generating function of convolutional code
T_{ch}	Duration of a chip
T_{hop}	Duration of a hop
T_s	Sampling interval
T°	Temperature
T°_A	Antenna temperature
T°_L	Effective line temperature
T_m	Multipath delay time (maximum)
T_0	Coherence time
T_{IL}	Interleaver span
T°_R	Effective receiver temperature
T°_S	System temperature
T_{acq}	Time to acquire
u_i	Code symbol
$u(t)$	Unit step function
$\mathbf{U}$	Codeword vector
$\mathbf{U}(X)$	Codeword polynomial
v	Relative velocity
$v(t)$	Electrical voltage waveform
var (X)	Variance of random variable X
V	Velocity
$w(t)$	Jammer waveform
W	Bandwidth
W_f	Filter bandwidth
W_{DSB}	Double-sideband bandwidth
W_N	Noise equivalent bandwidth
W_{ss}	Spread-spectrum bandwidth
$z(t)$	Output of matched filter or correlator
α_k	Forward state metric at time k
β_k	Reverse state metric at time k
Γ	SNR averaged through the "ups and downs" of fading
Γa	State metric for state a
γ	threshold (decision) level
γ_0	Optimum threshold level
$\gamma_{U^{(m)}}$	Likelihood of codeword $\mathbf{U}^{(m)}$
δ	Fractional frequency drift per day
δ_k	Branch metric at time k
δ_{mn}	Kronecker delta function
$\delta(t)$	Impulse (Dirac delta) function
ϵ	Error
ζ	Loop damping characteristic (second-order loop)
η	Antenna efficiency
$\theta(t)$	Time-varying phase
$\Theta(\omega)$	Fourier transform of $\theta(t)$
κ	Boltzmann's constant, 1.38×10^{-23}J/K
$\Lambda(d_k)$	Likelihood ratio of data bit d_k
λ	Joint probability
λ	Wavelength
λ	Packet arrival rate
π	Pi, 3.14159
ρ	Fraction of the frequency band being jammed
ρ	Fraction of the time the jammer is "on"
ρ	Normalized loop signal-to-noise ratio
ρ	Normalized message throughput
ρ	Number of erasures correctable in an error-correcting code
ρ	Time-correlation coefficient
ρ_0	Value of ρ that maximizes bit error probability (worst-case jamming)
σ_τ	rms delay spread
σ_X	Standard deviation of random variable X
σ^2_X	Variance of random variable X
τ	Pulse width
τ	Time shift (independent variable of the autocorrelation function)
$\psi_j(t)$	Basis function
$\Psi_x(f)$	Energy spectral density of waveform $x(t)$
ω	Radian frequency (radians per second)

Index

Absolute bandwidth, 49, 54
Access algorithms *See* Multiple access
ACI (adjacent channel interference), 249
ACK (acknowledgment), 316, 682–684
Acquisition:
  carrier wave, 603–625
  for spread spectrum systems, 746–751
Adaptive equalizing filter. *See* Equalization
Adaptive prediction. *See* Source coding, DPCM
ADC (analog-to-digital converter), 811
Adjacent channel interference (ACI), 249
AGC (automatic gain control), 810
A-law compression, 833–835
Aliasing, 65, 69–73
ALOHA system. *See* Multiple access, access algorithms
AMI (alternate mark inversion) PCM, 88
Amplitude phase keying (APK). *See* Modulation/demodulation, bandpass
Amplitude shift keying (ASK). *See* Modulation/demodulation, bandpass
AM/PM (amplitude-to-phase) conversion, 247, 296
Analog-to-digital converter (ADC), 811
Antenna, 247–248, 251–254, 256–257, 278
  beamwidth, 252, 254
  effective area, 251–254
  efficiency, 247–248, 252
  gain, 252–253
  reciprocity theorem, 253–254
  temperature, 278
Anti-jam (AJ) margin, 754–760
Antipodal signals, 129–130, 132–134, 173–174
APK (amplitude phase keying). *See* Modulation/demodulation
ARQ (automatic repeat request), 316
ASCII code, 58–59
ASK (amplitude shift keying). *See* Modulation/demodulation
Atmospheric effects, 248–249, 282–283
Authentication, for encryption systems, 891
Autocorrelation function, 19–20, 24–30, 32
Automatic gain control (AGC), 810
Automatic repeat request (ARQ), 316
Auto-zero, 809
Averages. *See* Random signals
AWGN channel, 30–33
Backoff, in TWTAs, 295–296, 690
Bandlimited channels, modulation and coding for, 566–590
Bandwidth, 42–49, 54, 71–72
  absolute, 49, 54
  bounded power spectral density, 49, 54
  characteristics of filters, 42–45
  coherence bandwidth. *See* Fading channels
  dilemma, 47
  fractional power containment, 49, 54
  half-power, 48, 54
  noise equivalent, 48–49, 54
  null-to-null, 49, 54
  transition, 71, 72
Bandwidth efficiency plane, 534–537
  bandwidth-limited systems, 540, 543, 547, 690–692
  power-limited systems, 541, 545, 547
Bandwidth-efficient signaling:
  coded modulation. *See* Trellis-coded modulation
  commercial telephone modems, 567
  duobinary, 94–100
  MPSK (multiple phase shift keying). *See* Modulation/demodulation, bandpass, *M*-ary
  MSK (minimum shift keying). *See* Modulation/demodulation, bandpass
  OQPSK (offset QPSK). *See* Modulation/demodulation, bandpass
  QAM (quadrature amplitude modulation). *See* Modulation/demodulation, bandpass
  QPSK (quaternary PSK). *See* Modulation/demodulation, bandpass
Barker sequences, 626
Baseband signals:
  bandwidth, 45–49, 89–94, 137–138
  definition, 6
  transmission. *See* Modulation/demodulation. *Also, see* Equalization. *Also, see* Pulse shaping
Baseband system, 56–58
Basis functions, 110–117, 134–135
Bayes' theorem, 1035–1040
  discrete form, 1036–1038
  mixed form, 1038–1040
BCH codes, 370–374
Bessel function, 215
Binary symmetric channel (BSC), 318–319, 396–394, 399–400
Binary tree search, in multiple access, 686
Biorthogonal codes, 311–312
Bi-phase PCM, 87–90
Bipolar signaling, 132–134
Bit:
  binary digit, 12
  information content, 529–531
Bit error probability. *See* Error performance, channel coding. *Also, see* Error performance, signaling
BLADES system, 768–769
Block coding. *See* Coding, channel. *Also, see* Source coding
Boltzmann's constant, 258
Bounded distance decoders, 345
BPSK (binary phase shift keying). *See* Modulation/demodulation, bandpass
BSC (binary symmetric channel), 318–319, 396–394, 399–400
Butterworth filter. *See* Filters
Capacity of a channel, 525–525
Carrier sense multiple access (CSMA) networks, 708–713
Carrier wave, 45–46, 168–169
Catastrophic convolutional codes, 414–415
CCI (co-channel interference), 249–250
CDMA (code division multiple access), 672–674, 769–771, 776–795
Central limit theorem, 31
Channel:
  binary symmetric, 318–319, 396–394, 399–400
  capacity of, 525–526
  definition, 244
  discrete memoryless (DMC), 318
  distortionless, 36–37
  Gaussian, 319–320, 400–401
  multipath, 945–958
Character coding, 58–62
Cipher feedback, 919–920
Ciphers, classic. *See* Encryption
CIRC (cross-interleave Reed–Solomon code), 469–475
Co-channel interference (CCI), 249
Code distance. *See* Distance
Code division multiple access (CDMA), 672–674, 769–771, 776–795
Code rate, 320, 382, 385, 388
Codes, source. *See* Source coding
Code weight, 342–343, 346
Coding (and decoding), channel:
  block codes, 328–374
    bounds:
      Hamming, 349–350
      Plotkin, 351–352
      cosets, 336–337
    cyclic codes, 356–374
    generator matrix, 331–334
    generator polynomial, 331–334
    parity codes, 321–323
    parity check matrix, 334–335
    rectangular (product) codes, 323
    Reed-Solomon (R–S), 437–460
      compact disc, 469–475
      decoding, 454–460
      encoding, 450–454
      error probability, 438–444
      finite fields, 445–450
      performance against burst noise, 441
      shortening, 471–472
    standard array, 336–338, 349–356
    syndrome testing, 335–341
    systematic codes, 333–334
    vector spaces, 329–331
  channel models:
    BSC (binary symmetric channel), 318–319, 396–394, 399–400

DMC (discrete memoryless channel), 318
Gaussian, 319–320, 400–401
hard versus soft decision, 396–400
concatenated codes, 468–469
convolutional codes, 382–429
add-compare-select, 405–408
catastrophic, 414–415
connection representation, 385–387
feedback decoding, 427–429
free distance, 409–410
hard-decision decoding, 396–400
impulse response, 386–388
maximum likelihood decoding, 395–396, 401–405
path memory, 408–416
performance bounds, 415–416
polynomial representation, 388–389
sequential decoding, 422–427
soft-decision decoding, 396–398, 400, 420–421
state representation, 389–391
synchronization, 408
transfer function, 410–412
tree diagram, 391–393
trellis diagram, 393–394
Viterbi decoding, 401–405
interleaving (and deinterleaving), 461–468
block, 461–465
convolutional, 466–468
properties:
capability, 326–327, 342–349, 366–374, 412–413, 415–418
closure, 329, 352
distance, 342–343, 408–410
erasure correction, 348–349
error correction, 336–345, 412–413, 415–418
error detection, 322–323, 336–345
gain, 325, 416–418
rate, 320, 382, 385, 388
redundancy, 320
weight, 342–343, 346
structured sequences, 317–328
trade-offs, 324–326
turbo codes, 475–511
branch metric, 503–504
error performance, 496–497
feedback decoder, 493–496
forward state metric, 499–501
log-likelihood ratio, 479–480, 510–511
MAP (maximum a posteriori) algorithm, 498–509
MAP decoding example, 504–509
product code example, 482–488
recursive systematic codes, 488-493
reverse state metric, 499-500, 502–503
types of error control, 315–317
automatic repeat request (ARQ), 316–317
terminal connectivity, 315–316
waveform coding, 305–314
antipodal, 307–308
biorthogonal, 311–312
correlation detection of, 313–314
M-ary signaling, 219–234, 308
orthogonal, 310–311
transorthogonal (simplex), 312–313
Coherence bandwidth. See Fading channels, characterization, small-scale fading
Coherence time. *See* Fading channels, characterization, small-scale fading
Coherent demodulation. *See* Modulation/demodulation
Compact disc (CD) digital audio system, 469–475
Companding, 84–85, 827–834
Complementary error function, Q(x), 122, 210–213, 1043–1045
Complex envelope, 204–206
Compression, nonuniform quantization:
A-law, 84–85, 833-834
logarithmic, 827–829
μ-law, 84, 829–833
Compression buffer, TDMA, 698–704
Concatenated coding systems, 468–469. *See* turbo codes
Confusion, in encryption systems, 683– 685
Constraint length, 382–384
Continuous phase modulation (CPM), 559–563, 631–635
Convolutional coding. *See* Coding, channel
Convolution integral, 34, 124, 1027
Correlation, 124–126
Correlation receiver, 178–183
Correlative coding, 94–100
Cosets, linear block codes, 336–337
Cosine filter, 97–99
Cross-correlation coefficient, 307–309, 312
Cross-interleave Reed–Solomon code (CIRC), 469–474
Cryptographic attacks. *See* Encryption
CSMA (carrier sense multiple access) networks, 708–713
Cycle slips, 612–614
Cyclic codes, 356–366, 370–373

DAMA (demand assignment multiple access), 678–686
Data encryption standard (DES), 909–915
Decibels (dB), 40–41, 83, 263
Decision theory, 1035–1045
likelihood ratio test, 119–120, 1041–1042
maximum a posteriori criterion (MAP), 1042
maximum likelihood criterion, 120, 1042
minimum error criterion, 120, 1041
Decision threshold, 119–120, 177–178, 182–183, 1042
Decoding. *See* Coding, channel
Decryption. *See* Encryption
Deinterleaving. *See* Coding, channel, interleaving
Delay-locked loop (DLL), 751–753
Delta function
Dirac, 16, 1023
Kronecker, 110, 1017
Delta modulation (DM). *See* Source coding, DPCM
Demand assignment multiple access (DAMA), 678–686
Demodulation. *See* Modulation/demodulation
DES (data encryption standard), 909–915
Detection. *See* Modulation/demodulation
Differential coding, 87, 96–97, 194–197, 206–209
Differential phase shift keying (DPSK). *See* Modulation/demodulation, bandpass
Differential pulse code modulation (DPCM). *See* Source coding
Diffusion, in encryption systems, 905
Digital communications:
advantages, 3–4
cost, 4
nomenclature, 11–12
performance, versus analog, 13
waveforms,
bandpass. *See* Modulation/demodulation
baseband. *See* Modulation/demodulation
classification, 13–16
Dirac delta function, 16, 1023
Direct sequencing (DS) spread spectrum systems, 732–735, 769–771, 776–779
Discrete memoryless channel (DMC), 318
Discrete sources, 14, 804–809
Distance
Euclidean, 400–401, 420–421, 585–589
free, 408–409
Hamming, 342, 396–400
minimum, 342
Distortionless transmission, 36–39
Dithering, 823–826
DM (delta modulation). *See* Source coding, DPCM
DMC (discrete memoryless channel), 318
Doppler shift, 601–602, 607–609, 645–646
DPCM (differential pulse code modulation). *See* Source coding
DPSK (differential phase shift keying). *See* Modulation/demodulation
Duobinary signaling. *See* Bandwidth efficient signaling
Duplex terminal connectivity, 315

Early Bird (INTELSAT I), 689
Early/late-gate synchronization, 628–630, 752–753
Eb/N0 (normalized SNR). *See* Modulation/demodulation
EBCDIC code, 59
EIRP (effective isotropic radiated power), 252–254
Encryption (and decryption):
classic ciphers, 894–897
Caesar cipher, 894–895
Polybius square, 894–895
Trithemius progressive key, 895–896
Vigenere key, 895–897
cryptographic attacks, 893–894
chosen plaintext, 894
cyphertext only, 893
known plaintext, 893–894
data encryption standard (DES), 909–915
goals, 893
models, 891–893
perfect secrecy, 897–900
entropy, 900–902
equivocation, 900–902
ideal secrecy, 902–904
rate of a language, 902
redundancy, 902
unicity distance, 902–904
PGP (pretty good privacy), 929–940
authentication and signature, 937–939

Encryption (and decryption)
PGP (pretty good privacy) (*cont.*)
CAST, 931–932
Diffie-Hellman (Elgamal variant), 935–936
digital signature standard, 939Triple-DES, 931–932
IDEA, 932–935
MD5, 939
RSA, 936, 939
SHA-1, 939
square-and-multiply algorithm, 937
practical security, 905–909
confusion, 905–907
diffusion, 905
permutation, 907–908
product cipher, 908–909
substitution, 905–907
pretty good privacy (PGP). *See* PGP
public key cryptosystems, 920–928
knapsack problem, 925–928
Merkle–Hellman scheme, 928
Rivest–Shamir–Adelman scheme, 923–925
signature authentication, 921–922
trapdoor one-way function, 922
stream encryption, 915–919
self-synchronous, 919–920
synchronous, 919–920
Energy density reduction. *See* Spread spectrum systems
Energy signal, definition, 14–16
Energy spectral density (ESD), 17–18
Ensemble averages, 21–23
Entropy, 529–534, 900–902
Equalization, 149–161
channel characterization, 149–151
eye diagram, 151–152
filter types:
transversal, 152–157
zero-forcing solution, 154–156
minimum mean square error, 156–157
decision feedback, 157–158
preset, 158
adaptive, 158–160
filter update rate, 160–161
Equivocation, 532–534, 900–902
Erasure correction, 348–349
Ergodicity, 25–26
Error correction capability:
block codes, 326–327, 342–349, 366–374
convolutional codes, 412–413, 415–418
Error detection, 316–317, 321–323,342
Error performance, channel coding:
P_B (bit error probability), 311–312, 367, 370, 415–416, 439–443
P_E (symbol or codeword error probability), 311–312, 439
P_M (message error probability), 323, 327
P_{nd} (undetected error probability), 322, 346
Error performance, signaling:
P_B (bit error probability):
bipolar signaling, 121–135
coherent BPSK, 209–21, 219
coherent FSK, 213, 219
coherently detected differentially encoded PSK, 211–212
coherent M-ary orthogonal, 225–229
coherent multiple phase, 220–224
differentially detected PSK (DPSK), 216–219
noncoherent FSK, 213–216, 219
OQPSK and MSK, 562
QAM, 563–566
unipolar signaling, 131–135
P_B versus P_E:
multiple phase signaling, 234–235
orthogonal signaling, 232–234
P_E (symbol error probability):
multiple phase signaling, 229–230
orthogonal signaling, 230–232
Error probability plane, 522–523
Euclidean distance, 332–334, 420–423
Euler's totient function, 923
Expansion buffer, TDMA, 702–704
Extension codes, 808–809, 860, 864–865

Facsimile, 866–868
Fading channels, 945–1011
Characterization, 945–978
large-scale fading, 951–952
small-scale fading, 953–958
coherence bandwidth, 959–965, 992–995
coherence time, 967, 969–971, 973, 992–995
degradation types:
Doppler spreading (fading rate), 969–971, 973–976
fast fading, 969, 982–983
flat fading, 962–966
frequency-selective fading, 962–966
slow fading, 969
mitigation, 978–1001
decision feedback equalizer, 980–981
diversity, 983–987
interleaving, 988–991
orthogonal frequency-division multiplexing (OFDM), 982
pilot signal, 982
rake receiver, 999–1001
spread spectrum, 981–982
Viterbi equalizer, 996–998
False-lock in a phase-locked loop, 615
FDMA (frequency division multiple access), 657–665
FDMA/TDMA combined, 666–667
Feedback decoding, 427–429
Feedback shift registers. *See* Shift registers
FH/MFSK (fast hopping MFSK). *See* Spread spectrum systems
Filters:
Butterworth, 41–42
cosine, 97–98
equalizing. *See* Equalization
ideal, 37–39
raised cosine, 139–149
realizable, 40–42
root-raised cosine, 146–149
transversal. *See* Equalization
Formatting:
character coding, 58–61
pulse code modulation. *See* PCM
sampling. *See* Sampling
quantizing. *See* Source coding
Fourier integral, 1020
Fourier operations, 1034
Fourier series, 1014
Fourier transform, generalized, 113–114
Fourier transform properties, 1021
frequency shifting, 1021
time shifting, 1021
Fourier transform table, 1033
Fractional power containment bandwidth, 49, 54
Free distance, 409–410
Free space, 244–245
Free space loss, 254–257
Frequency division multiple access (FDMA). *See* Multiple access
Frequency hopping (FH). *See* Spread spectrum systems
Frequency reuse, 658–659, 674–675
Frequency transfer function, 35–41, 1013–1014
Frequency translation (heterodyning or mixing), 41–43, 479–483
Friis, H. T., 272–273, 297
FSK (frequency shift keying). *See* Modulation/demodulation

Gardner, F. M., 615, 629, 650
Gaussian channel, 319–320, 400–401
Gaussian density function, 30–31
Generator matrix, 331–334
Generator polynomial, 358–365
Global positioning system, 608
Global System for Mobile (GSM) Communication, 996–998
Golay code, extended, 369–370
Granular quantization noise, 814
G/T° (receiver figure-of-merit), 261, 289
Guard band, 661
Guard time, 665

Half-power bandwidth, 48, 54
Hamming codes, 366–369
Hamming distance, 342–343, 347, 400–402, 408
Hamming weight, 342–343
Hard-decision decoding, 318–319, 397–400
Heterodyning (frequency translation or mixing), 45–47, 661–663
Hofman, L., 643, 651
Huffman code, 862–866
Hypothesis testing, 1041–1043

I (in-phase) channel, 190–191, 198–199, 205–209
IDEA,
Ideal filter, 37–39
Ideal secrecy, 902–904
Identities, often used, 1048–1049
IM (intermodulation) products, 247, 250
Implementation loss, 250, 288
Impulse function, 16, 1023
Impulse response, 34, 386
Information content, 529–533
Information source, 11, 804–810
INTELSAT. *See* Satellite systems
Interference:
ACI (adjacent channel), 249
CCI (co-channel), 249–250
ISI (intersymbol interference), 77–78, 136–142, 235–236
filter roll-off, 139–142
Nyquist filter, 138
Nyquist pulse, 138, 146–149
pulse shaping, 138–142
pulse support time, 146
raised-cosine filter, 139–142
root-raised cosine filter, 146–149
square-root Nyquist pulse, 146–149
Interference rejection, 720–721, 726–727
Interleaving. *See* Coding, channel
Intermodulation (IM) products, 247, 250
ISI (intersymbol interference). *See* Interference
Isotropic radiator, 251–253

Jamming. *See* Spread spectrum systems
JPEG. *See* Source coding, compression algorithms

Knapsack problem. *See* Encryption, public key cryptosystems
Kronecker delta, 110, 1017

LAN (local area network). *See* Multiple access, LAN
Lempel-Ziv codes, 868-869
Lender, A., 94, 100
Likelihood function, 119–121, 181–182, 395–396, 477–481, 1041–1042
Likelihood ratio test, 119–120, 479–480, 510–511, 1041–1042
Limiters, 247
Line loss, 250, 274–276
Linear block coding. *See* Coding, channel, block
Linear feedback shift register (LFSR). *See* Shift registers
Linear independence, 331
Linear predictive coding (LPC), 838–841, 852–853, 872–873
Linear systems, 33–42, 1012–1014
Link availability, 266–270
Link budget analysis, 259–270, 286–295
Link margin. *See* Margin
Local area network (LAN). *See* Multiple access, LAN
Logarithmic compression, 827–829
Log-likelihood function, 1042
Loop, phase-locked. *See* Phase-locked loops (PLL)
LPC (linear predictive coding), 838–841, 852–853, 872–873
LPI (low probability of intercept), 722
LPPF (low probability of position fix), 722
LPSE (low probability of signal exploitation), 722
Luecke, E. J., 627, 651

Manchester coding, 132–134, 708–710
Margin:
  AJ, 757–759
  noise, 262–265
M-ary signaling:
  bandpass, 220–235
  baseband, 91–94
  waveform coding, 305–315
Matched filter, 122-126, 184–188
Maury–Styles sequences, 643
Maximum a posteriori (MAP) criterion, 1041
Maximum likelihood:
  criterion, 1041
  decoding, 395–396
  detection, 119–122
MCPC (multiple channels per carrier), 690–693
Memoryless source, 806
Merkle–Hellman scheme, 927–928
Mid-riser, 812
Mid-tread, 812
Minimum distance, 343, 408–410
Minimum error criterion, 120, 1041
Minimum shift keying (MSK), 559–563
Mixing (heterodyning), 45–47, 661–663
m-law quantization. *See* Compression, nonuniform quantization
Modems, 4–5, 567
Modulation/demodulation, 85–235
  bandpass, 169–235
    APK (amplitude phase keying), 176
    ASK (amplitude shift keying), 175
    correlation receiver, 178–183
    decision regions, 177–178
    DPSK (differential phase shift keying), 194–197, 216–218
    FSK (frequency shift keying), 175, 191–193, 198–204, 213–216, 220–221, 225–229
    M-ary:
      MFSK orthogonal, 225–229, 230–234
      MPSK (multiple phase), 188–191, 207–209, 222-224, 229–230, 234–235
    MSK (minimum shift keying), 559–563
    OQPSK (offset QPSK), 555–563
    PSK (phase shift keying), 173–174, 183–188, 209–211, 218–219, 614–616, 621–623
    QAM (quadrature amplitude modulation), 563–566
    QPSK (quaternary PSK), 188–191, 222–224, 555
    sampled matched filter, 184–188
  baseband, 85–149
    bipolar, 85–90, 132–136
    E_b/N_0, 117–119
    line codes. *See* PCM waveforms
    M-ary, 91–94
    matched filter, 122–131
    maximum likelihood receiver, 119–122
    multilevel, 86–90, 94–100
    PCM waveforms, 85–90
    shaped pulses, 145–149
    unipolar, 85–90, 131–132, 134–136
    vector view of signals and noise, 110–117
Modulation loss, 247
Modulo arithmetic, 96, 329
MPSK (multiple phase shift keying). *See* Modulation/demodulation, bandpass
MSK (minimum shift keying), 559–563
Multilevel signaling, 86–90, 94–100
Multipath channels, 771–772, 947–978
Multiple access:
  access algorithms:
    DAMA (demand assignment), 677–686
    polling (binary tree search), 686–689
    pure ALOHA, 678–682
    R-ALOHA (reservation ALOHA), 683–686
    random access, 678–686
    S-ALOHA (slotted ALOHA), 682–686
  architecture, 677–678
  CDMA (code division), 672–674, 769–771, 776–794
  FDMA (frequency division), 660–672, 690–698
    MCPC (multiple channel per carrier), 690–693
    multiplex telephony, 660–665
    single channel per carrier (SPADE), 693–698
  FDMA/TDMA combined, 666
  frequency reuse, 657–658, 674–676
  information flow, between controller and earth station, 677–678
  for INTELSAT. *See* Satellite systems
  LAN (local area network), 708–713
    CSMA (carrier sense), 708–713
    token ring, 710–713
  PDMA (polarization division), 674–676
  SDMA (space division), 674–676
  TDMA (time division), 665–672
    SS/TDMA (satellite-switched TDMA), 704–708
Multiple phase shift keying (MPSK). *See* Modulation/demodulation, bandpass, M-ary
Multiplexing, 656–666

NAK, 316, 682–684
Newman, F., 643, 651
Noise:
  atmospheric, 248, 282–283
  AWGN, 30–33, 116–117
  galactic (cosmic), 250, 282–285
  IM (intermodulation), 247, 250
  phase, 247
  sky, 282–285
  thermal, 30–31, 258–259
  white, 32–33, 116–117, 258
Noise equivalent bandwidth, 48–49, 54
Noise figure, 270–273, 276–277, 279–281
Noise sources, 245–250
Noise temperature, 273, 276–282
Noise wheel, Rogoff, 727–728
Nomenclature, digital systems, 9–11
Noncoherent demodulation. *See* Modulation/demodulation
Nonregenerative repeaters. *See* Repeater
Nonreturn-to-zero (NRZ) PCM, 86–88
Nonuniform quantization. *See* Source coding, quantization
N-tap linear prediction (LPC), 839–841
Null-to-null bandwidth, 49, 54
Nyquist
  filtering, 137–149
  minimum bandwidth, 137–142, 524–525
  sampling criterion, 63–75

Odenwalder, J. P., 419, 430
One-tap linear prediction (LPC), 838–839
One-time pad, 893
OOK (on–off keying), 175, 216
OQPSK (offset QPSK), 555–563
Orthogonal codes, 307–315
Orthogonal signals, 110–117. *See* Modulation/demodulation, bandpass, FSK. *Also, see* Coding, types of error control
Orthogonal spaces, 110–117
Overload error, 814, 820–823

PAM (pulse amplitude modulation), 91–94
Parity-check codes, 321–323
Parity-check matrix, 334–340
Partial-band noise jamming, 760–762
Partial response signaling, 94–100
Path loss, 254–257
PB (bit error probability). *See* Error performance
PCM (pulse code modulation):
  definition, 79–81
  DPCM (differential). *See* Source coding, DPCM
  multiplex frame structure, 700–703
    European standard, 701–702
    T-Carrier, 702–703
  spectral attributes, 89–90
  waveforms, 85–89
    multilevel, 91–100
    nonreturn to zero, 86–88
    phase encoded, 88
    return to zero, 88
  word size, 91
PE (symbol error probability). *See* Error performance
Perfect codes, 347–348
Permutation, in encryption systems, 907–908
Phase-locked loops (PLL), 603–625
  analysis, nonlinear, 611-614
  Costas loop, 616–617
  cycle slips, 612–614
  errors, steady state, 606–609
  false lock, 615
  linearized loop equation, 604–606
  loop bandwidth, 609–611, 621
  loop filter characteristics, 606–609

Phase-locked loops (PLL) (*cont.*)
loop hang-up, 620
performance, 609–611, 621–623
response to a frequency ramp, 608–609
response to a frequency step, 607–608
response to a phase step, 606–607
signal-to-noise ratio, 615–616, 618, 621–623
squaring loss, 615–616, 618, 621–623
suppressed carrier loops, 614–618
Phase shift keying (PSK). *See* Modulation/demodulation, bandpass
Phasor representation of signals, 171–173
Pietrobon, S., 498, 412
PM (message error probability). *See* Error performance, channel coding
PN sequences. *See* Pseudonoise sequences
Pointing loss, 248
Poisson process, 680–681
Polarization division multiple access (PDMA), 674–676
Polarization loss, 248
Polling, multiple access algorithm, 686–689
Polybinary signaling, 99–100
Polynomial division, 377–363
Power:
average, 15, 18–19, 51–52
decibels, 40–41
definition, 14–15
spectral density (PSD), 17–18, 26–30
Power-limited systems, 541, 545, 547
Predictive coding, 838–841, 852–853, 872–873
Prefix-free codes, 861
Pretty good privacy (PGP), 929–940
Privacy, for encryption systems, 669
Probability of error. *See* Error performance
Processing gain, 735–738
Product cipher, 908–909
Product (rectangular) code, 323
Product integrator (correlator), 124–126, 178–183
Pseudonoise (PN) sequences. *See* Spread spectrum systems
PSK (phase shift keying). *See* Modulation/demodulation, bandpass
Public key cryptosystems. *See* Encryption
Pulse amplitude modulation (PAM), 91–94
Pulse code modulation (PCM). *See* PCM
Pulse jamming, 763–766
Pulse shaping, 136–161. Also, see Equalization
Nyquist pulse and square-root Nyquist pulse, 146–149
Raised-cosine filter, 139–144
Root-raised cosine filter, 146–149

Q (quadrature), channel, 198–199, 205–209, 223–224
QAM (quadrature amplitude modulation), 563–566
QPSK (quaternary PSK), 188–191, 222–224, 555
Quantization. *See* Source coding, amplitude quantizing

Radio maps of the sky, 283–285
Radiometer, 722–723
Radome, 248
Rain, 248, 266–270, 282–283
R-ALOHA (reservation-ALOHA) system, 683–686
Random access, 678–686
Random processes:
autocorrelation function, 24–30
definition, 22
ensemble averages, 21–22
ergodicity, 25–26
Gaussian, 30–31
Markov, 389, 807
Poisson, 680–681
power spectral density, 26–30
sample function, 25
stationarity, 24–25
time averages, 25–26
Random variables:
central limit theorem, 31
Gaussian, 30–31
moments, 22
probability density function (pdf), 21–23, 30–31
Rayleigh, 214
Rician, 214–215
standard deviation, 26
variance, 22, 26
Range equation, 250–254
Rate:
of block codes, 320
of convolutional codes, 382, 385, 388
of a language, 902
Received isotropic power (RIP), 289–290
Receiver figure-of-merit (G/T_∞), 261, 289
Rectangular (product) code, 323
Rect function, 1033
Redundancy of a language, 902
Redundancy-reducing coding. *See* Source coding
Reed–Solomon (R-S) codes, 437–461
decoding, 454–461
encoding, 450–454
error probability, 438–441
finite fields, 445–450
performance against burst noise, 441
Repeaters:
nonlinear repeater amplifiers, 295–296
nonregenerative, 290–295
regenerative, 3–4, 290
Return-to-zero (RZ) PCM, 85–90
Ring network, 710–713
RIP (received isotropic power), 289–290
Rivest–Shamir–Adelman (RSA) scheme, 923–925
Robustness, 741
Roll-off factor, 139–142
Root-raised cosine filter, 146–149
RSA public key cryptosystem, 923–925, 935–936
Run-length codes, 862–870

S-ALOHA (slotted ALOHA) system, 682–686
Sampling, 63–75
aliasing, 69–75
channel noise, 77
flat-top, 69
impulse, 63–66
natural, 66–69
sample and hold, 69, 75–76
signal-to-noise ratio, 78–79
theorem, 63
timing jitter, 77
Why oversample? 72–75
Satellite repeaters. *See* Repeaters
Satellite-switched TDMA (SS/TDMA), 704–708
Satellite systems:
COMSTAR, 674–676
Early Bird, 689
INTELSAT, 689–708
FDM/FM/FDMA (MCPC), 690
link budget for, 263–270
SPADE, 693–698
SS/TDMA, 704–708
TDMA, 665–672
Saturation quantization error, 76–77, 814, 820–823
Schwarz's inequality, 123
Secrecy, perfect, 897–900
Secrecy, practical, 905–909
Self information, 805
Sequential decoding, 422–427
Shannon, C E , 525–529, 590
capacity theorem, 525–529
limit, 528–529
Shift registers
decoding of cyclic codes, 365–366
encoding of cyclic codes, 363–365
linear feedback (LFSR), 729–730, 916–917
vulnerabilities of, 917–919
Sidebands, 45–47, 660–665
Sidelobes, 252
Signal classification, 14–16
analog and discrete, 14
deterministic and random, 14
energy and power, 14–16
periodic and nonperiodic, 14
unit impulse, 16
Signal constellation design, 569–573
Signal-to-noise ratio (SNR) degradation, 245–250
Signal-to-quantization-noise (S/N)q ratio, 78–79
Simplex (transorthogonal) coding, 312–313
Simplex terminal connectivity, 315
Sinc function, 30, 1018–1020
Sky noise, 282–285
Slotted ALOHA (S-ALOHA) system, 682–686
Soft-decision decoding, 319, 396–401, 420–422
Source coding:
adaptive prediction, 850–853
forward adaptation, 851–852
synthesis/analysis coding, 852–853
linear predictive (LPC), 838–841, 852–853, 872–873
block coding, 853–856
codebook coders, 854–855
code population, 855
tree coders, 854–855
trellis coder, 854–855
vector quantizing, 854–855
differential pulse code modulation (DPCM), 835–850
delta modulation (DM), 841–842
N-tap linear prediction (LPC), 839–841
one-tap linear prediction (LPC), 838–839
sigma-delta, 842–850
A-to-D converter (ADC), 847
D-to-A converter (DAC), 848
examples, 870–884
audio compression, 870–875
adaptive differential pulse code modulation (ADPCM), 870
codebook excited linear predictive (CELP), 872–873
standard CCITT G.722, 871–872
standard CCITT G.726, 870
subband partitioned ADPCM, 871–872
image compression, 875–884
Joint Photographic Experts Group (JPEG), 876–880

Motion Picture Experts Group (MPEG), 880–884
LPC (linear predictive coding), 838–841, 852–853, 872–873
source coding for digital data, 860–869
Huffman, 862–868
Lempel-Ziv, 868–869
properties, 860–861
prefix-free, 861
uniquely decodable, 861
run length, 866–868
sources, 804–811
discrete, 804–809
entropy, 805–809
extension codes, 808–809
waveform, 809–811
transform coding, 856–859
quantization for, 857
subband coding, 857–859, 871–872
quantization, 78–85, 811–835
dithering, 823–826
memoryless (instantaneous), 835
noise, 76, 78–79, 813–816
nonuniform, 83–85, 826–835
saturation, 76–77, 820–823
uniform (linear), 81–83, 816–820
Space division multiple access (SDMA), 674–676
Space loss, 248–249, 254
SPADE, 693–698
Spectral characteristics of filters, 42–45
Spectral density, 16–20, 26–30
ESD (energy), 17
PSD (power), 17–20, 26–30
Spilker, J. I., Jr., 643, 645, 651
Spread spectrum systems,
attributes:
energy density reduction, 722–723
fine time resolution, 723–724
interference suppression, 720–721, 726–727
multiple access, 724. *See* CDMA
processing gain, 735–738
robustness, 741–742
CDMA applications, 672–674, 769–771,776–794
cellular systems, 775–794
commercial applications, 769–794
direct sequencing (DS), 732–735, 769–771
FCC Part 15 Rules, 772–773
frequency hopping (FH), 740–745
fast versus slow, 742–745
historical background, 727–728
noise wheel, 727–728
stored reference (SR) system, 727
transmitted reference (TR) system, 727
IS-95 cellular systems, 782–794
jamming:
AJ (anti-Jam) margin, 757–759
BLADES system, 768–769
broadband noise, 759–760
jammer waveforms, 754–756
J/S ratio, 756–757
multiple tone, 763–765
partial-band noise, 760–762
pulse jammer, 763–766
repeat-back jammer, 765–768
multipath channels, 771–772
pseudonoise (PN) sequences, 728–732
PN autocorrelation function, 730–732
randomness properties, 729
shift register generation of, 729–730
vulnerabilities of linear feedback shift registers, 917–919
synchronization, 745–754
acquisition, 746–751
correlator structures, 746–749
serial search, 749–750
sequential estimation, 750–751
tracking, 751–754
Squaring loss, 615–616, 618, 621–623
SS/TDMA (satellite-switched TDMA), 704–708
Standard array, 336–338, 349–356
coset, 336–340
syndrome of a coset, 336–340
State diagram, 389–391, 807
Stationary random process, 24–25
Stored reference spread spectrum system, 727
Subband coding, 857–859, 871–872
Substitution, in encryption systems, 905–907
Successive approximation, analog-to-digital conversion, 835
Superposition, 1012–1013
Surviving path, Viterbi decoding, 402–405
Symbols, 61–62
Symbols used in this book, 1071–1076
Synchronization:
acquisition, carrier wave, 448–451
for convolutional decoding, 408
receiver:
coherent, PLL. *See* Phase-locked loops
continuous phase modulation, 631–639
data aided, 635–637
nondata aided, 637–639
error performance, 621–623, 630–631
frame, 639–643
spectrum analysis techniques, 623–625
symbol, 625–631
word, 640–643
network, 643–649
closed loop, 647–649
frequency reference fluctuations, 645–646
open loop, 644–647
for spread spectrum. *See* Spread spectrum systems
Synchronous altitude, 244, 663
Syndrome testing, 335–341
Synthesis/analysis coding. *See* Source coding
Systematic codes:
block, 333–334, 359–360, 363–365
convolutional, 413–414

Tau-dither loop, 751–754
TDMA (time division multiple access), 665–672
Telephone modems, 567–568
Telephony, 660–665
Temperature:
antenna, 278
effective noise, 258–259, 273–285
sky noise, 282–285
system, 277–282
Terminal connectivity, 315–316
Thermal noise, 27–30, 202–204
Throughput, of a communications resource, 657
Time averaging, 25–26
Time division multiple access (TDMA). *See* Multiple access
Time hopping (TH), 724–726
Token ring network, 710–713
Tone jamming, 763–765
Tracking:
delay lock loop, 751–753
tau-dither loop, 752–754
Trade-offs:
convolutional code rate, 382, 385, 388
link, 296–297
modulation and coding, 522–524, 534–538
Traffic, communication, 704–706
Transfer function, convolutional codes, 409–413
Transfer function, frequency, 35–41, 1013
Transform coding, 856–859
Transition bandwidth, 71
Transmitted reference (TR) spread spectrum system, 727
Transorthogonal (simplex) codes, 312–313
Transversal filter, 152–157
Trapdoor one-way function, 922
Traveling-wave tube amplifier (TWTA), 247, 295–296
Tree diagram, 324–326
Trellis-coded modulation, 573–590
coding gain, 581–582
decoding, 580–583
encoding, 576–580
example, 585–589
idea behind TCM, 574–576
mapping waveforms to trellis transitions, 577–579
multidimensional, 589–590
Trellis diagram, 393–395, 401–408, 579–580, 583, 585
Turbo codes:
branch metric, 499–504
concepts, 477–481
feedback decoder, 493–498
forward state metric, 499–504
likelihood and log-likelihood functions, 477–480
log-likelihood algebra, 481–482, 510–511
MAP decoding algorithm, 498–504
MAP decoding example, 498–504
product-code example, 504–509
recursive systematic codes, 488–493
reverse state metric, 409–504

Ungerboeck, G., 576, 592
Unicity distance, 902–904
Uniform (linear) quantization, 81–83, 816–820
Unipolar signaling, 131–132
Unit impulse function, 16, 1023

Variance, 22, 26
VCO (voltage-controlled oscillator), 603–606, 610–611, 614–617, 620
Vector quantizing, 854–856
Vector representation of signals, 110–117, 222–223, 225–226
Vector spaces and subspaces, 110–117
Velocity error, 608
Viterbi, A. J., 401–405, 430–431, 611–614, 650–652
Viterbi decoding, 401–405

Waveform coding. *See* Coding, channel, waveform
Waveform sources, 809–811
White noise, 32–33, 116–117, 258
response of a correlator to, 117, 1046–1048
power spectral density of, 32, 809–810
Wide-sense stationary, 24–25
Willard sequences, 642–643
Wintz, P. A., 627, 651

LICENSE AGREEMENT AND LIMITED WARRANTY

READ THE FOLLOWING TERMS AND CONDITIONS CAREFULLY BEFORE OPENING THIS SOFTWARE MEDIA PACKAGE. THIS LEGAL DOCUMENT IS AN AGREEMENT BETWEEN YOU AND PRENTICE-HALL, INC. (THE "COMPANY"). BY OPENING THIS SEALED SOFTWARE MEDIA PACKAGE, YOU ARE AGREEING TO BE BOUND BY THESE TERMS AND CONDITIONS. IF YOU DO NOT AGREE WITH THESE TERMS AND CONDITIONS, DO NOT OPEN THE SOFTWARE MEDIA PACKAGE. PROMPTLY RETURN THE UNOPENED SOFTWARE MEDIA PACKAGE AND ALL ACCOMPANYING ITEMS TO THE PLACE YOU OBTAINED THEM FOR A FULL REFUND OF ANY SUMS YOU HAVE PAID.

1. **GRANT OF LICENSE:** In consideration of your payment of the license fee, which is part of the price you paid for this product, and your agreement to abide by the terms and conditions of this Agreement, the Company grants to you a nonexclusive right to use and display the copy of the enclosed software program (hereinafter the "SOFTWARE") on a single computer (i.e., with a single CPU) at a single location so long as you comply with the terms of this Agreement. The Company reserves all rights not expressly granted to you under this Agreement.

2. **OWNERSHIP OF SOFTWARE:** You own only the magnetic or physical media (the enclosed SOFTWARE) on which the SOFTWARE is recorded or fixed, but the Company retains all the rights, title, and ownership to the SOFTWARE recorded on the original SOFTWARE copy(ies) and all subsequent copies of the SOFTWARE, regardless of the form or media on which the original or other copies may exist. This license is not a sale of the original SOFTWARE or any copy to you.

3. **COPY RESTRICTIONS:** This SOFTWARE and the accompanying printed materials and user manual (the "Documentation") are the subject of copyright. You may not copy the Documentation or the SOFTWARE, except that you may make a single copy of the SOFTWARE for backup or archival purposes only. You may be held legally responsible for any copying or copyright infringement which is caused or encouraged by your failure to abide by the terms of this restriction.

4. **USE RESTRICTIONS:** You may not network the SOFTWARE or otherwise use it on more than one computer or computer terminal at the same time. You may physically transfer the SOFTWARE from one computer to another provided that the SOFTWARE is used on only one computer at a time. You may not distribute copies of the SOFTWARE or Documentation to others. You may not reverse engineer, disassemble, decompile, modify, adapt, translate, or create derivative works based on the SOFTWARE or the Documentation without the prior written consent of the Company.

5. **TRANSFER RESTRICTIONS:** The enclosed SOFTWARE is licensed only to you and may not be transferred to any one else without the prior written consent of the Company. Any unauthorized transfer of the SOFTWARE shall result in the immediate termination of this Agreement.

6. **TERMINATION:** This license is effective until terminated. This license will terminate automatically without notice from the Company and become null and void if you fail to comply with any provisions or limitations of this license. Upon termination, you shall destroy the Documentation and all copies of the SOFTWARE. All provisions of this Agreement as to warranties, limitation of liability, remedies or damages, and our ownership rights shall survive termination.

7. **MISCELLANEOUS:** This Agreement shall be construed in accordance with the laws of the United States of America and the State of New York and shall benefit the Company, its affiliates, and assignees.

8. **LIMITED WARRANTY AND DISCLAIMER OF WARRANTY:** The Company warrants that the SOFTWARE, when properly used in accordance with the Documentation, will operate in substantial conformity with the description of the SOFTWARE set forth in the Documentation. The Company does not warrant that the SOFTWARE will meet your requirements or that the operation of the SOFTWARE will be uninterrupted or error-free. The Company warrants that the media on which the SOFTWARE is delivered shall be free from defects in materials and workmanship under normal use for a period of thirty (30) days from the date of your purchase. Your only remedy and the Company's only obligation under these limited warranties is, at the Company's option, return of the warranted item for a refund of any amounts paid by you or replacement of the item. Any replacement of SOFTWARE or media under the warranties shall not extend the original warranty period. The limited warranty set forth above shall not apply to any SOFTWARE which the Company determines in good faith has been subject to misuse, neglect, improper installation, repair, alteration, or damage by you. EXCEPT FOR THE EXPRESSED WARRANTIES SET FORTH ABOVE, THE COMPANY DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. EXCEPT FOR THE EXPRESS WARRANTY SET FORTH ABOVE, THE COMPANY DOES NOT WARRANT, GUARANTEE, OR MAKE ANY REPRESENTATION REGARDING THE USE OR THE RESULTS OF THE USE OF THE SOFTWARE IN TERMS OF ITS CORRECTNESS, ACCURACY, RELIABILITY, CURRENTNESS, OR OTHERWISE.

IN NO EVENT, SHALL THE COMPANY OR ITS EMPLOYEES, AGENTS, SUPPLIERS, OR CONTRACTORS BE LIABLE FOR ANY INCIDENTAL, INDIRECT, SPECIAL, OR CONSEQUENTIAL DAMAGES ARISING OUT OF OR IN CONNECTION WITH THE LICENSE GRANTED UNDER THIS AGREEMENT, OR FOR LOSS OF USE, LOSS OF DATA, LOSS OF INCOME OR PROFIT, OR OTHER LOSSES, SUSTAINED AS A RESULT OF INJURY TO ANY PERSON, OR LOSS OF OR DAMAGE TO PROPERTY, OR CLAIMS OF THIRD PARTIES, EVEN IF THE COMPANY OR AN AUTHORIZED REPRESENTATIVE OF THE COMPANY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN NO EVENT SHALL LIABILITY OF THE COMPANY FOR DAMAGES WITH RESPECT TO THE SOFTWARE EXCEED THE AMOUNTS ACTUALLY PAID BY YOU, IF ANY, FOR THE SOFTWARE.

SOME JURISDICTIONS DO NOT ALLOW THE LIMITATION OF IMPLIED WARRANTIES OR LIABILITY FOR INCIDENTAL, INDIRECT, SPECIAL, OR CONSEQUENTIAL DAMAGES, SO THE ABOVE LIMITATIONS MAY NOT ALWAYS APPLY. THE WARRANTIES IN THIS AGREEMENT GIVE YOU SPECIFIC LEGAL RIGHTS AND YOU MAY ALSO HAVE OTHER RIGHTS WHICH VARY IN ACCORDANCE WITH LOCAL LAW.

ACKNOWLEDGMENT

YOU ACKNOWLEDGE THAT YOU HAVE READ THIS AGREEMENT, UNDERSTAND IT, AND AGREE TO BE BOUND BY ITS TERMS AND CONDITIONS. YOU ALSO AGREE THAT THIS AGREEMENT IS THE COMPLETE AND EXCLUSIVE STATEMENT OF THE AGREEMENT BETWEEN YOU AND THE COMPANY AND SUPERSEDES ALL PROPOSALS OR PRIOR AGREEMENTS, ORAL, OR WRITTEN, AND ANY OTHER COMMUNICATIONS BETWEEN YOU AND THE COMPANY OR ANY REPRESENTATIVE OF THE COMPANY RELATING TO THE SUBJECT MATTER OF THIS AGREEMENT.

Should you have any questions concerning this Agreement or if you wish to contact the Company for any reason, please contact in writing at the address below.

Robin Short
Prentice Hall PTR
One Lake Street
Upper Saddle River, New Jersey 07458

Communications

Digital Communications

SECOND EDITION

Fundamentals and Applications

"This is a remarkably comprehensive treatment of the field, covering in considerable detail modulation, coding (both source and channel), encryption, multiple access and spread spectrum. It can serve both as an excellent introduction for the graduate student with some background in probability theory or as a valuable reference for the practicing communication system engineer. For both communities, the treatment is clear and well presented."

—Andrew Viterbi,
The Viterbi Group

CD-ROM INCLUDED

The CD contains the complete educational version of SystemView by ELANIX, the user friendly software for communications and DSP design. The CD also contains detailed notes for getting started, a comprehensive DSP tutorial, and over 200 additional communication system exercises.

Master the key aspects of digital communications: technology, concepts, and techniques

Digital Communications: Fundamentals and Applications, Second Edition is a thoroughly revised and updated edition of the field's classic, best-selling introduction to the subject. With remarkable clarity, Dr. Bernard Sklar introduces this important discipline, which forms the nucleus of today's wireless and Internet revolutions. He provides a unified structure and context for understanding digital communications—all without sacrificing mathematical precision.

Sklar begins by introducing the fundamentals of signals, spectra, formatting, and baseband transmission. Next, he presents practical coverage of virtually every contemporary modulation, coding, and signal-processing technique, with numeric examples and step-by-step guidance through multiple applications. Coverage includes:

- Signals and processing steps: from the information source through the transmitter, channel, receiver, and up to the information sink
- Key tradeoffs: signal-to-noise ratio, probability of error, and bandwidth
- Trellis-coded modulation and Reed-Solomon codes: what's *behind* the math
- Synchronization and spread spectrum solutions
- Fading channels: causes, effects, and techniques for mitigation
- A complete guide to understanding *turbo codes*: squeezing maximum performance out of digital connections
- Implementing PGP encryption, the de facto industry standard for electronic mail
- Essential concepts for successful system design

Whether you're building wireless systems, xDSL, fiber or coax-based services, satellite networks, or Internet infrastructure, Sklar presents the theory *together with* practical application details. With nearly 500 illustrations and 300 problems and exercises, there's never been a faster way to master fundamentals as well as advanced topics in digital communications.

About the Author

DR. BERNARD SKLAR has over 40 years of experience in technical design and management positions at Republic Aviation, Hughes Aircraft, Litton Industries, and at The Aerospace Corporation, where he helped develop the MILSTAR satellite system. He is now head of advanced systems at Communications Engineering Services, a consulting company he founded in 1984. He has taught engineering courses at several universities, including UCLA and USC, and has trained professional engineers worldwide.

PRENTICE HALL
Upper Saddle River, NJ 07458

www.phptr.com

ISBN 0-13-084788-7