The Characterization of Fading Channels

by

Bernard Sklar

Introduction

When the mechanisms that cause fading in communication channels were first modeled in the 1950s and 1960s, the principles developed were primarily applied to over-the-horizon communications covering a wide range of frequency bands. The 3-30 MHz high-frequency (HF) band used for ionospheric propagation, as well as the 300 MHz-3 GHz ultra-high-frequency (UHF) and the 3-30 GHz super-high-frequency (SHF) bands used for tropospheric scatter, are examples of channels that are affected by fading phenomena. Although the fading effects in mobile radio channels are somewhat different from those encountered in ionospheric and tropospheric channels, the early models are still quite useful in helping to characterize the fading effects in mobile digital communication systems. This article emphasizes so-called *Rayleigh fading*, primarily in the UHF band, which affects mobile systems such as cellular and personal communication systems (PCS). The primary goal is to characterize the fading channel and in so doing to describe the fundamental fading manifestations and types of degradation.

The Challenge of Communicating Over Fading Channels

In the analysis of communication system performance, the classical (ideal) additive white Gaussian noise (AWGN) channel, with statistically independent Gaussian noise samples corrupting data samples free of intersymbol interference (ISI), is the usual starting point for developing basic performance results. The primary source of performance degradation is thermal noise generated in the receiver. Often, external interference received by the antenna is more significant than the thermal noise. This external interference can sometimes be characterized as having a broadband spectrum and quantified by a parameter called *antenna temperature*. The thermal noise usually has a flat power spectral density over the signal band and a zero-mean Gaussian voltage probability density function (pdf). When modeling practical systems, the next step is the introduction of bandlimiting filters. Filtering in the transmitter usually serves to satisfy some regulatory requirement on spectral containment. Filtering in the receiver is often the result of implementing a matched filter [1]. Due to the bandlimiting and phase-distortion properties of

filters, special signal design and equalization techniques may be required to mitigate the filter-induced ISI.

If a radio channel's propagating characteristics are not specified, one usually infers that the signal attenuation versus distance behaves as if propagation takes place over ideal free space. The model of free space treats the region between the transmit and receive antennas as being free of all objects that might absorb or reflect radio frequency (RF) energy. It also assumes that, within this region, the atmosphere behaves as a perfectly uniform and nonabsorbing medium. Furthermore, the earth is treated as being infinitely far away from the propagating signal (or, equivalently, as having a reflection coefficient that is negligible). Basically, in this idealized free-space model, the attenuation of RF energy between the transmitter and receiver behaves according to an inverse-square law. The received power expressed in terms of transmitted power is attenuated by a factor $L_s(d)$. This factor, expressed below, is called *path loss* or *free space loss*, and is predicated on the receiving antenna being isotropic [1].

$$L_s(d) = \left(\frac{4\pi d}{\lambda} \right)^2 \tag{1}$$

In Equation (1), d is the distance between the transmitter and the receiver, and λ is the wavelength of the propagating signal. For this case of idealized propagation, received signal power is very predictable. For most practical channels, where signal propagation takes place in the atmosphere and near the ground, the free-space propagation model is inadequate to describe the channel behavior and predict system performance. In a wireless mobile communication system, a signal can travel from transmitter to receiver over multiple reflective paths. This phenomenon, referred to as *multipath* propagation, can cause fluctuations in the received signal's amplitude, phase, and angle of arrival, giving rise to the terminology *multipath fading.* Another name, *scintillation,* having originated in radio astronomy, is used to describe the fading caused by physical changes in the propagating medium, such as variations in the electron density of the ionospheric layers that reflect high-frequency (HF) radio signals. Both names, fading and scintillation, refer to a signal's random fluctuations; the main difference is that scintillation involves mechanisms (such as electrons) that are much smaller than a wavelength. The end-to-end modeling and design of systems that incorporate techniques to mitigate the effects of fading are usually more challenging than those whose sole source of performance degradation is AWGN.

Characterizing Mobile-Radio Propagation

Figure 1 represents an overview of fading-channel manifestations. It starts with two types of fading effects that characterize mobile communications: large-scale fading and small-scale fading. *Large-scale fading* represents the average signal-power attenuation or the path loss due to motion over large areas. In Figure 1, the large-scale fading manifestation is shown in blocks 1, 2, and 3. This phenomenon is affected by prominent terrain contours (hills, forests, billboards, clumps of buildings, and so on) between the transmitter and the receiver. The receiver is often said to be "shadowed" by such prominences. The statistics of large-scale fading provide a way of computing an estimate of path loss as a function of distance. This is described in terms of a mean-path loss (*n*th-power law) and a log-normally distributed variation about the mean. *Small-scale fading* refers to the dramatic changes in signal amplitude and phase that can be experienced as a result of small changes (as small as a half wavelength) in the spatial positioning between a receiver and a transmitter. As indicated in Figure 1 blocks 4, 5, and 6, small-scale fading manifests itself in two mechanisms: time-spreading of the signal (or *signal dispersion*) and time-variant behavior of the channel. For mobile-radio applications, the channel is time-variant because motion between the transmitter and the receiver results in propagation path changes. The rate of change of these propagation conditions accounts for the fading rapidity (rate of change of the fading impairments). Small-scale fading is called *Rayleigh fading* if there are multiple reflective paths that are large in number and there is no line-of-sight signal component; the envelope of such a received signal is statistically described by a Rayleigh pdf. When a dominant nonfading signal component is present, such as a line-of-sight propagation path, the small-scale fading envelope is described by a *Rician* pdf [2]. In other words, the small-scale fading statistics are said to be Rayleigh whenever the line-of-sight path is blocked, and Rician otherwise. A mobile radio roaming over a large area must process signals that experience both types of fading: small-scale fading superimposed on large-scale fading.

Large-scale fading (attenuation or path loss) can be considered to be a spatial average over the small-scale fluctuations of the signal. It is generally evaluated by averaging the received signal over 10-30 wavelengths, in order to decouple the small-scale (mostly Rayleigh) fluctuations from the large-scale shadowing effects (typically log-normal). There are three basic mechanisms that impact signal propagation in a mobile communication system: reflection, diffraction, and scattering.

- *Reflection* occurs when a propagating electromagnetic wave impinges upon a smooth surface with very large dimensions compared to the RF signal wavelength (λ).

- *Diffraction* occurs when the propagation path between the transmitter and receiver is obstructed by a dense body with dimensions that are large when compared to λ, causing secondary waves to be formed behind the obstructing body. Diffraction is a phenomenon that accounts for RF energy travelling from transmitter to receiver without a line-of-sight path between the two. It is often termed *shadowing* because the diffracted field can reach the receiver even when shadowed by an impenetrable obstruction.

- *Scattering* occurs when a radio wave impinges on either a large rough surface or any surface whose dimensions are on the order of λ or less, causing the energy to be spread out (scattered) or reflected in all directions. In an urban environment, typical signal obstructions that yield scattering include lampposts, street signs, and foliage. The name *scatterer* applies to any obstruction in the propagation path that causes a signal to be reflected or scattered.

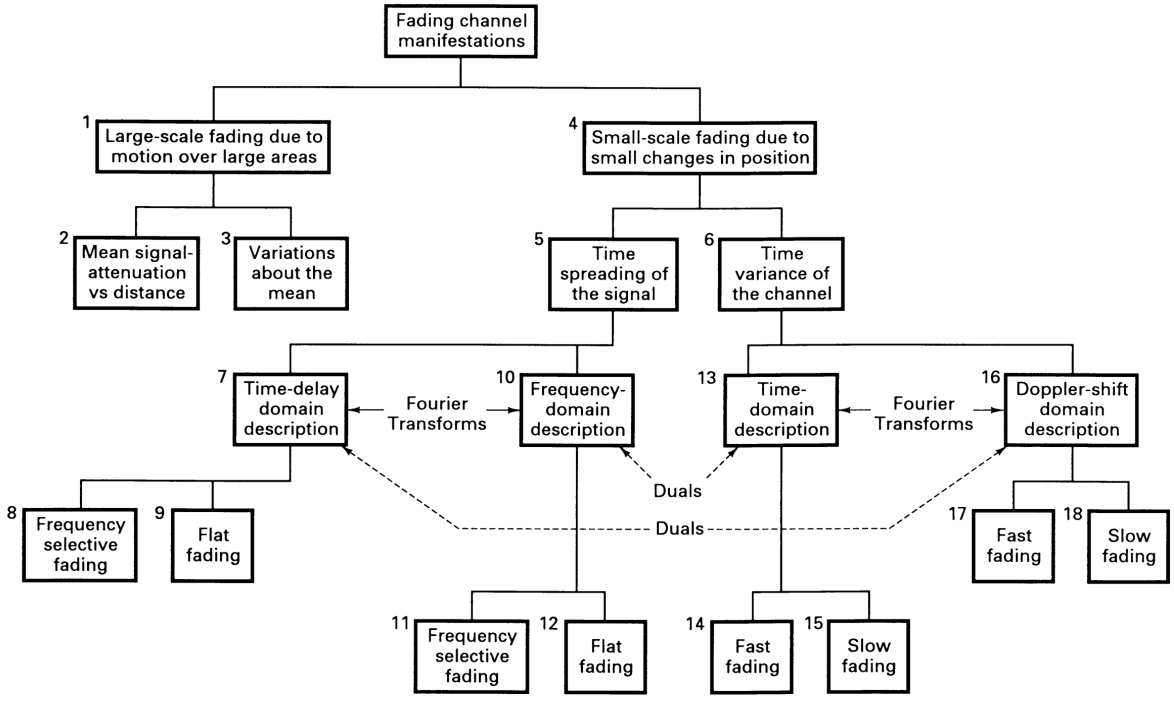


Figure 1

Fading channel manifestations.

Figure 1 may serve as a table of contents for the sections that follow. The two manifestations of small-scale fading, signal time-spreading (signal dispersion) and the time-variant nature of the channel are examined in two domains: time and frequency, as indicated in Figure 1 blocks 7, 10, 13, and 16. For signal dispersion, the fading degradation types are categorized as being frequency-selective or frequency-nonselective (flat), as listed in blocks 8, 9, 11, and 12.

For the time-variant manifestation, the fading degradation types are categorized as fast-fading or slow-fading, as listed in blocks 14, 15, 17, and 18. (The labels indicating Fourier transforms and duals are explained later.)

Figure 2 is a convenient pictorial (not a precise graphical representation) showing the various contributions that must be considered when estimating path loss for link budget analysis in a mobile radio application [3]. These contributions are: (1) mean path loss as a function of distance, due to large-scale fading; (2) near-worst-case variations about the mean path loss or large-scale fading margin (typically 6-10 dB); and (3) near-worst-case Rayleigh or small-scale fading margin (typically 20-30 dB). In Figure 2, the annotations "≈1-2%" indicate a suggested area (probability) under the tail of each pdf as a design goal. Hence, the amount of margin indicated is intended to provide adequate received signal power for approximately 98-99% of each type of fading variation (large- and small-scale).

Using complex notation, a transmitted signal is written as follows:

$$s(t) = \mathrm{Re}\left\{g(t)e^{j2\pi f_c t}\right\} \tag{2}$$

where $\mathrm{Re}\{\bullet\}$ denotes the real part of $\{\bullet\}$, and f_c is the carrier frequency. The baseband waveform $g(t)$ is called the *complex envelope* of $s(t)$, and can be expressed as follows [1]:

$$g(t) = \left|g(t)\right| e^{j\phi(t)} = R(t)e^{j\phi(t)} \tag{3}$$

where $R(t) = |\, g(t)|$ is the envelope magnitude, and $\phi(t)$ is its phase. For a purely phase- or frequency-modulated signal, $R(t)$ will be constant, and in general will vary slowly compared to $t = 1/f_c$.

In a fading environment, $g(t)$ will be modified by a complex dimensionless multiplicative factor $\alpha(t)e^{-j\theta(t)}$. (We show this derivation later.) The *modified* baseband waveform can be written as $\alpha(t)e^{-j\theta(t)}g(t)$, but for now let's examine the magnitude, $\alpha(t)R(t)$, of this envelope, which can be expressed in terms of three positive terms, as follows [4]:

$$\alpha(t)R(t) = m(t) \times r_0(t) \times R(t) \tag{4}$$

where $m(t)$ is called the *large-scale-fading component* of the envelope, and $r_0(t)$ is called the *small-scale-fading component*. Sometimes, $m(t)$ is referred to as the *local mean* or *log-normal fading* because generally its measured values can be statistically described by a log-normal pdf, or equivalently, when measured in decibels, $m(t)$ has a Gaussian pdf. Furthermore, $r_0(t)$ is sometimes referred to as *multipath* or *Rayleigh fading.* For the case of a mobile radio, Figure 3 illustrates the relationship between $\alpha(t)$ and $m(t)$. In this figure, we consider that an *unmodulated* carrier wave is being transmitted, which in the context of Equation (4) means that for all time, $R(t) = 1$. Figure 3a is a representative plot of signal power received versus antenna displacement (typically in units of wavelength). The signal power received is of course a function of the multiplicative factor $\alpha(t)$. Small-scale fading superimposed on large-scale fading can be readily identified. The typical antenna displacement between adjacent signal-strength nulls, due to small-scale fading, is approximately a half wavelength. In Figure 3b, the large-scale fading or local mean, $m(t)$, has been removed in order to view the small-scale fading, $r_0(t)$, referred to some average constant power. Recall that $m(t)$ can generally be evaluated by averaging the received envelope over 10-30 wavelengths. The log-normal fading is a relatively slow-varying function of position, while the Rayleigh fading is a relatively fast-varying function of position. Note that for an application involving motion, such as a radio in a moving vehicle, a function of position is tantamount to a function of time. In the sections that follow, some of the details regarding the statistics and mechanisms of large-scale and small-scale fading are enumerated.

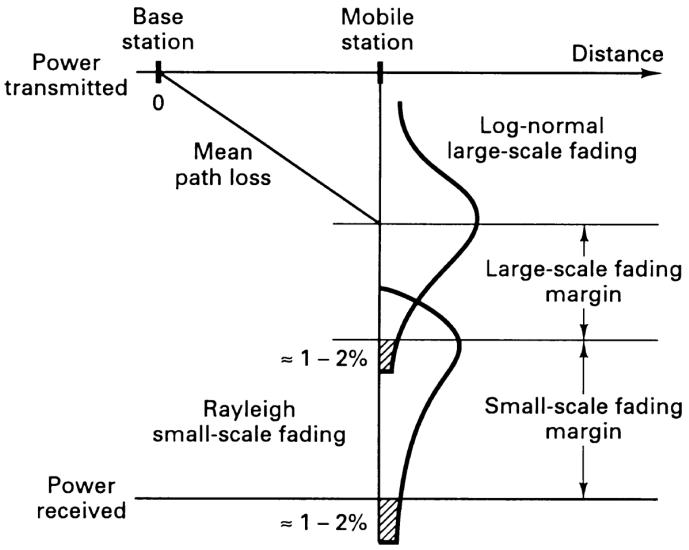
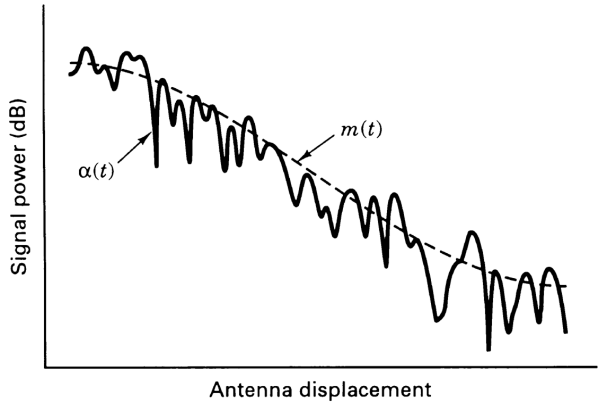
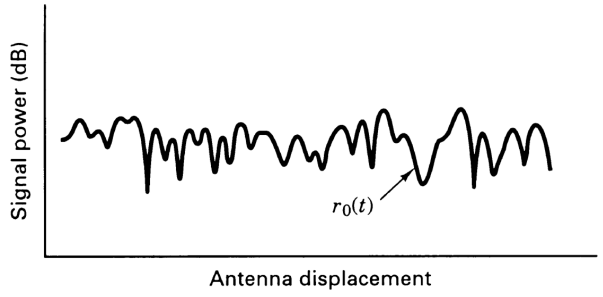


Figure 2

Link budget considerations for a fading channel.



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



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

Figure 3

Large-scale fading and small-scale fading.

Large-Scale Fading

For mobile radio applications, Okumura [5] made some of the earlier comprehensive path-loss measurements for a wide range of antenna heights and coverage distances. Hata [6] 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 transmitter and receiver is proportional to an nth-power of d relative to a reference distance d_0 [2].

$$\overline{L_p}(d) \propto \left(\frac{d}{d_0}\right)^n \tag{5}$$

$\overline{L_p}(d)$ is often stated in decibels, as shown below.

$$\overline{L_p}(d)\ (\text{dB}) = L_s(d_0)\ (\text{dB}) + 10\,n\,\log\left(\frac{d}{d_0}\right) \tag{6}$$

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 (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 10n dB/decade. The value of the exponent n depends on the frequency, antenna heights, and propagation environment. In free space, where signal propagation follows an inverse-square law, n is equal to 2, as shown in Equation (1). In the presence of a very strong guided-wave phenomenon (such as urban streets), n can be lower than 2. When obstructions are present, n is larger. Figure 4 shows a scatter plot of path loss versus distance for measurements made at several sites in Germany [7]. 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.

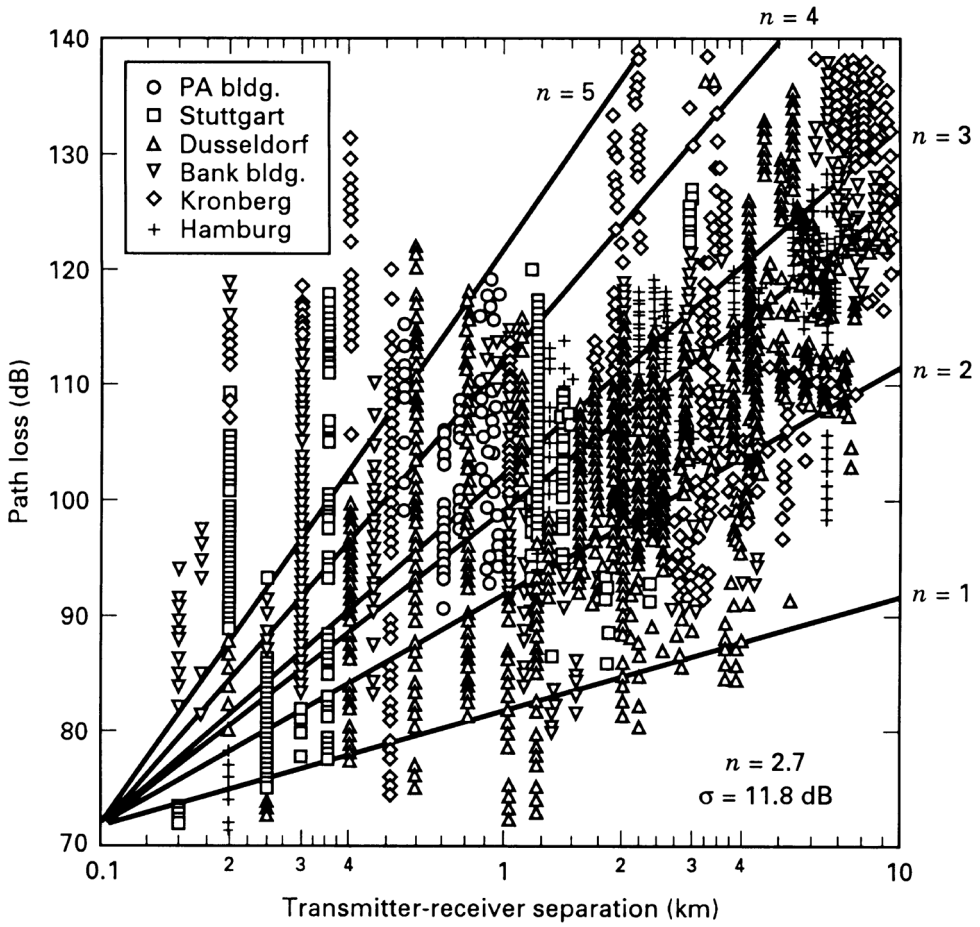


Figure 4

Path loss versus distance, measured in several German cities.

The path loss versus distance expressed in Equation (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 since the environment of different sites may be quite different for similar transmitter-receiver (T-R) separations. Figure 4 illustrates that path-loss variations can be quite large. Measurements have shown that for any value of d, the path loss $L_p(d)$ is a random variable having a log-normal distribution about the mean distant-dependent value $\overline{L_p}(d)$ [8]. Thus, path loss $L_p(d)$ can be expressed in terms of $\overline{L_p}(d)$ as expressed in Equation (6), plus a random variable X_σ, as follows [2]:

$$L_p(d)\ \text{(dB)} = L_S(d_0)\ \text{(dB)} + 10n \log_{10}(d/d_0) + X_\sigma\ \text{(dB)} \tag{7}$$

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 (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 6-10 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 σ of X_σ. There are several good references dealing with the measurement and estimation of propagation path loss for many different applications and configurations [2, 5-9].

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, $r(t)$, can be written as

$$r(t) = \sum_n \alpha_n(t)\, s[t - \tau_n(t)] \tag{8}$$

Substituting Equation (2) into Equation (8), we write the received bandpass signal as follows:

$$\begin{aligned} r(t) &= \mathrm{Re}\left(\left\{ \sum_n \alpha_n(t)\, g[t - \tau_n(t)] \right\} e^{j2\pi f_c [t - \tau_n(t)]} \right) \\ &= \mathrm{Re}\left(\left\{ \sum_n \alpha_n(t)\, e^{-j2\pi f_c \tau_n(t)} g[t - \tau_n(t)] \right\} e^{j2\pi f_c t} \right) \end{aligned} \tag{9}$$

From Equation (9), it follows that the equivalent received *baseband* signal is

$$z(t) = \sum_n \alpha_n(t)\, e^{-j2\pi f_c \tau_n(t)} g[t - \tau_n(t)] \tag{10}$$

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 (10), reduces to

$$z(t) = \sum_n \alpha_n(t)\, e^{-j2\pi f_c \tau_n(t)} = \sum_n \alpha_n(t)\, e^{-j\theta_n(t)} \tag{11}$$

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 (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, namely fading of $z(t)$. Equation (11) can be expressed more compactly as the net received envelope, which is the summation over all the scatterers, as follows:

$$z(t) = \alpha(t) e^{-j\theta(t)} \tag{12}$$

where $\alpha(t)$ is the resultant magnitude, and $\theta(t)$ is the resultant phase. The right side of Equation (12) represents the same complex multiplicative factor that was described earlier. Equation (12) is an important result because it tells us that, even though a *bandpass* signal $s(t)$ as expressed in Equation (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 5 illustrates the primary mechanism that causes fading in multipath channels, as described by Equations (11) and (12). In the figure, a reflected signal has a phase delay (a function of additional path length) with respect to a 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 $x_n(t) + j\, y_n(t) = \alpha_n(t) e^{-j\theta_n(t)}$. 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 (4). For the case of an unmodulated carrier wave as shown in Equation (12), $r_0(t)$ is the magnitude of $z(t)$, as follows:

$$r_0(t) = \sqrt{x_r^2(t) + y_r^2(t)} \tag{13}$$

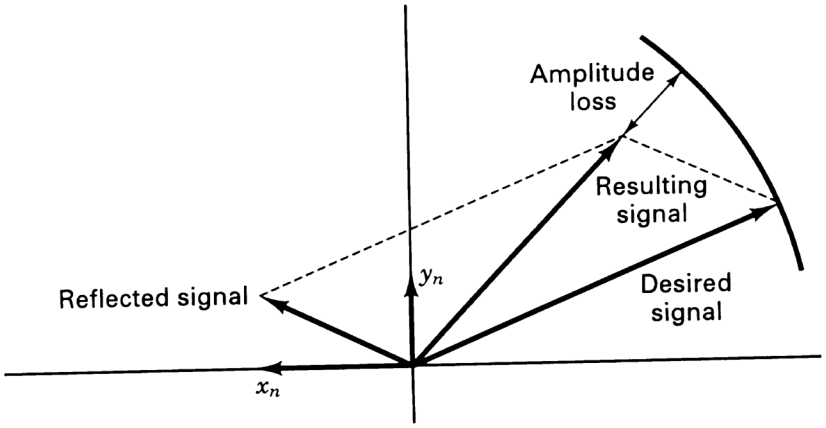


Figure 5

Effect of a multipath reflected signal on a desired signal.

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, 3].

$$p(r_0) = \begin{cases} \dfrac{r_0}{\sigma^2} \exp\left[-\dfrac{\left(r_0^{\,2} + A^2\right)}{2\sigma^2} \right] I_0\left(\dfrac{r_0 A}{\sigma^2} \right) & \text{for } r_0 \geq 0,\ A \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

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 $I_0(\bullet)$ is the modified Bessel function of the first kind and zero order [12]. The Rician distribution is often described in terms of a parameter K that 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 follows:

$$p(r_0) = \begin{cases} \frac{r_0}{\sigma^2} \exp\left[-\frac{r_0^2}{2\sigma^2} \right] & \text{for } r_0 \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

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 article, it will be assumed unless stated otherwise 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.

As indicated in Figure 1, blocks 4, 5, and 6, small-scale fading manifests itself in two mechanisms:

- Time-spreading of the underlying digital pulses within the signal
- A time-variant behavior of the channel due to motion (for example, a receive antenna on a moving platform)

Figure 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 traveling at a constant velocity). In Figure 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 that results from the fading channel's non-optimum 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 6, either antenna position or transmission time can be used to illustrate this time-variant behavior. Figures 6a-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 [13], where λ is the wavelength of the carrier frequency. For each of the 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 7 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 7, the time-spreading mechanism will be characterized in the time-delay domain as a multipath delay spread, and in the frequency domain 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.

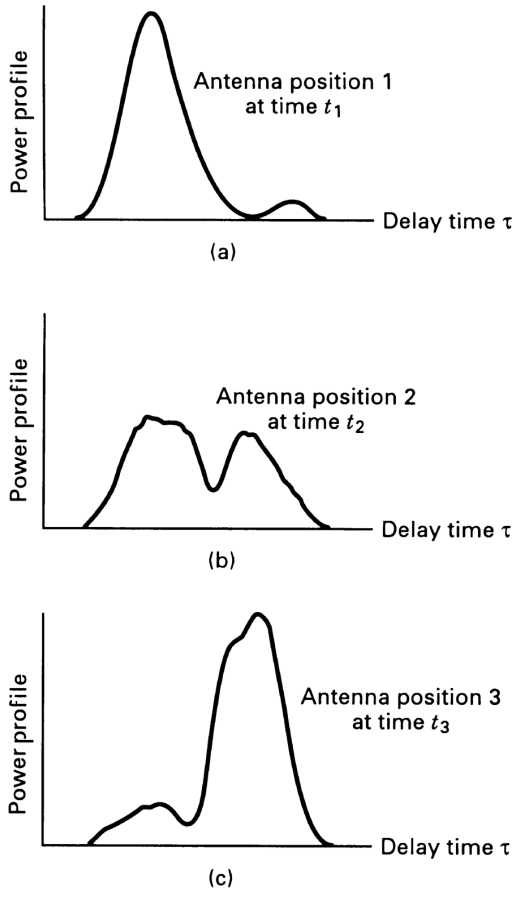


Figure 6

Response of a multipath channel to a narrowband pulse versus delay, as a function of antenna position.

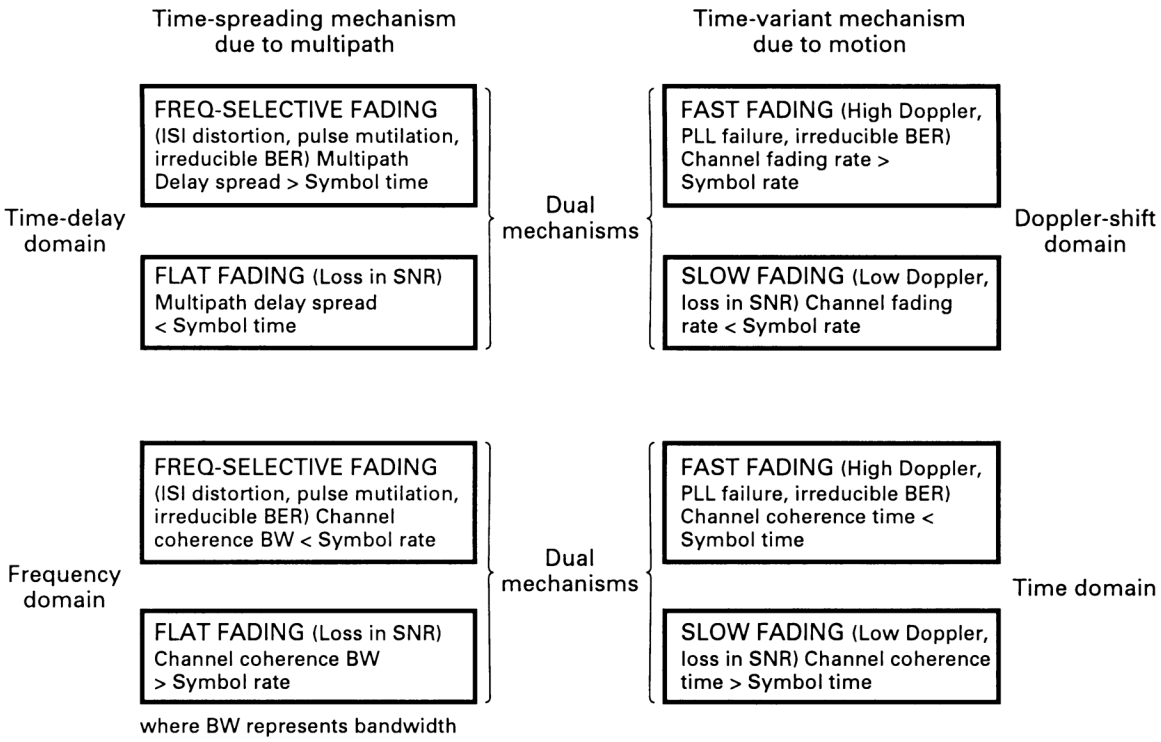


Figure 7

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

Signal Time-Spreading Viewed in the Time-Delay Domain

A simple way to model the fading phenomenon was introduced by Bello [14] 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 [3, 14] 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 8 contains four functions that make up this model [3, 11, 14-16]. These functions are now examined, starting with Figure 8a and proceeding counterclockwise to Figure 8d.

In Figure 8a, a *multipath-intensity profile*, $S(\tau)$, is plotted 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 [11, 16]. 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 [16].

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.

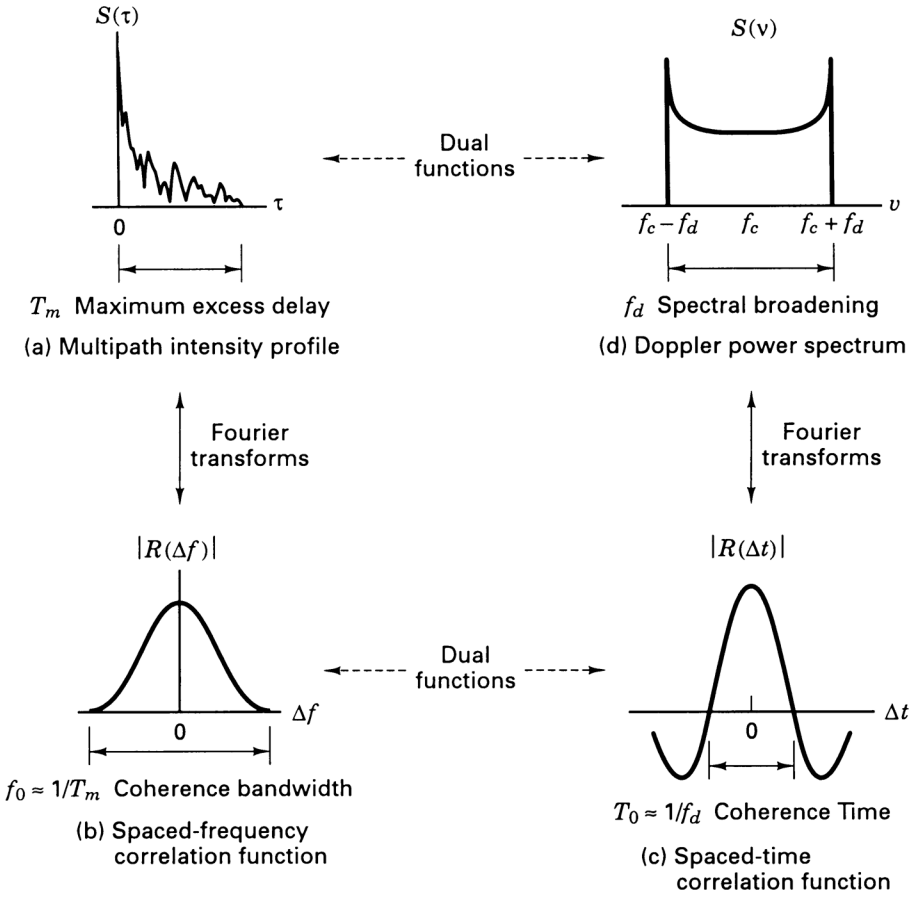


Figure 8

Relationships among the channel correlation functions and power density functions.

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*, as indicated in Figure 1, blocks 8 and 9, and Figure 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 later in this article.

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 when viewing degradation in the frequency domain, where the phenomenon is more easily described.) For loss in SNR due to flat fading, the appropriate mitigation technique is to improve the received SNR (or 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 objective.

Signal Time-Spreading Viewed in the Frequency Domain

A completely analogous characterization of signal dispersion can be specified in the frequency domain. Figure 8b shows the function $\left| R(\Delta f) \right|$, designated a *spaced-frequency* correlation function; 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 frequency components have a strong potential for amplitude correlation. That is, a signal's spectral components in that range are affected by the channel in a similar manner, for example, exhibiting fading or no fading. Note that f_0 and T_m are reciprocally related (within a multiplicative constant). As an approximation, it is possible to say that

$$f_0 \approx 1/T_m \tag{16}$$

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. It is most often characterized in terms of its root mean squared (rms) value, called the *rms delay spread*, σ_τ, where

$$\sigma_\tau = \sqrt{\overline{\tau^2} - (\overline{\tau})^2} \tag{17}$$

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

No exact relationship exists between coherence bandwidth and delay spread. An approximation can be derived from signal analysis (usually using Fourier transform techniques) of actual signal dispersion measurements in particular 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 as follows [17]:

$$f_0 \approx \frac{1}{50\sigma_\tau} \tag{18}$$

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 [18, 19] is generally accepted as a useful model for an urban propagation environment. This model is called the *dense-scatterer* channel model. With the use of such a model, coherence bandwidth has similarly been defined [18] for a bandwidth interval over which the channel's complex frequency transfer function has a correlation of at least 0.5, to be

$$f_0 = \frac{0.276}{\sigma_\tau} \tag{19}$$

Studies involving ionospheric effects often employ the following definition [20]:

$$f_0 = \frac{1}{2\pi\sigma_\tau} \tag{20}$$

A more popular approximation of f_0 corresponding to a bandwidth interval having a correlation of at least 0.5 is as follows [2]:

$$f_0 = \frac{1}{5\sigma_\tau} \tag{21}$$

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, and so on). It is important to note that neither of the parameters in Equation (21) depends on signaling speed. A system's signaling speed only influences its transmission bandwidth, W.

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 (QPSK, MSK, spread spectrum, and so on) [21]. 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 to those components contained within the coherence bandwidth. Figure 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 9a is the frequency transfer function of a frequency-selective channel ($f_0 < W$). Figure 9a shows that various spectral components of the transmitted signal will be affected differently.

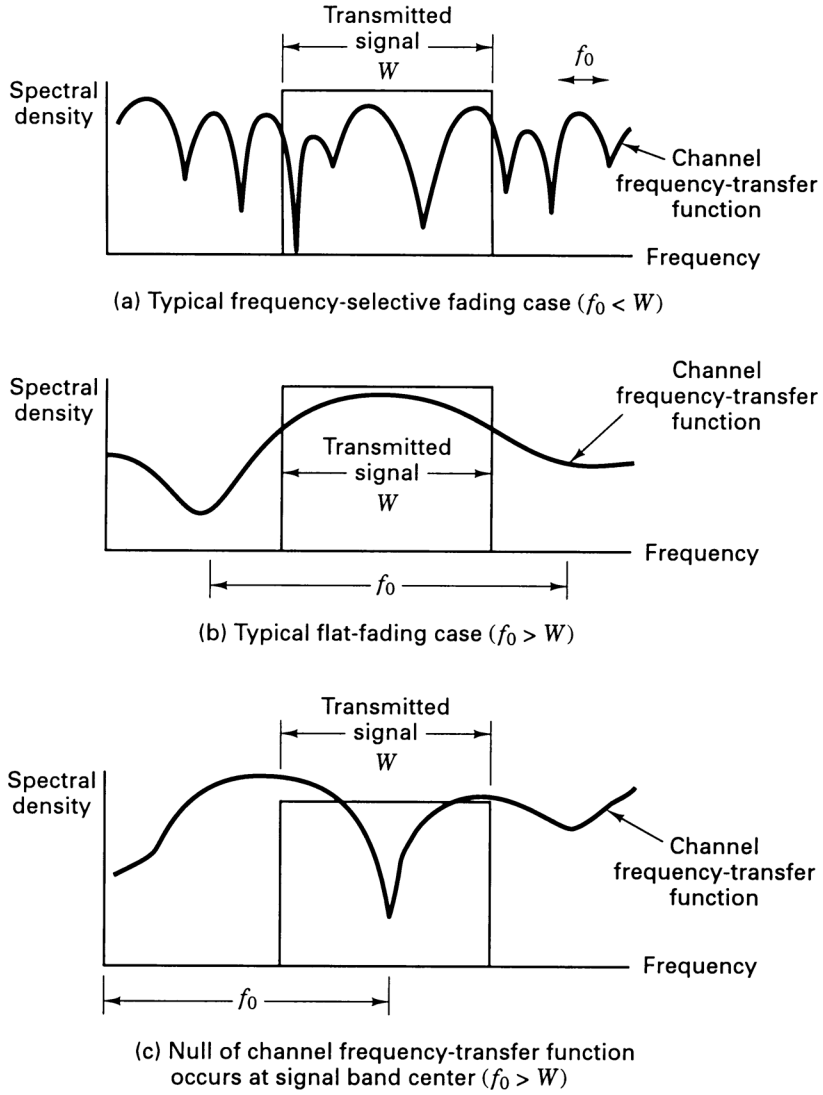


Figure 9

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

Frequency-nonselective or 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 (fading or no fading). This is illustrated in Figure 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 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

$$f_0 > W \approx \frac{1}{T_s} \tag{22}$$

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 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 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 [18]. Thus, even though 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 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 condition shown in Figure 9c. By comparison, it should be clear that in Figure 9a the fading is independent of the position of the signal band, and frequency-selective fading occurs all the time, not just occasionally.

Examples of Flat Fading and Frequency-Selective Fading

Figure 10 shows some examples of flat fading and frequency-selective fading for a direct-sequence spread-spectrum (DS/SS) system [20, 21]. In Figure 10, there are three plots of the output of a pseudo-noise (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 shown in Figure 8a. The key difference is that the

amplitudes shown in Figure 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 10 is drawn from a satellite-to-ground communications link exhibiting scintillation because of atmospheric disturbances. However, Figure 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 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 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 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 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% of the spread-spectrum bandwidth. The coherence bandwidths (relative to the spread-spectrum signaling speed) shown in 10b and 10c depict channels that can be categorized as *moderately* and *highly* frequency-selective, respectively. 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 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 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.

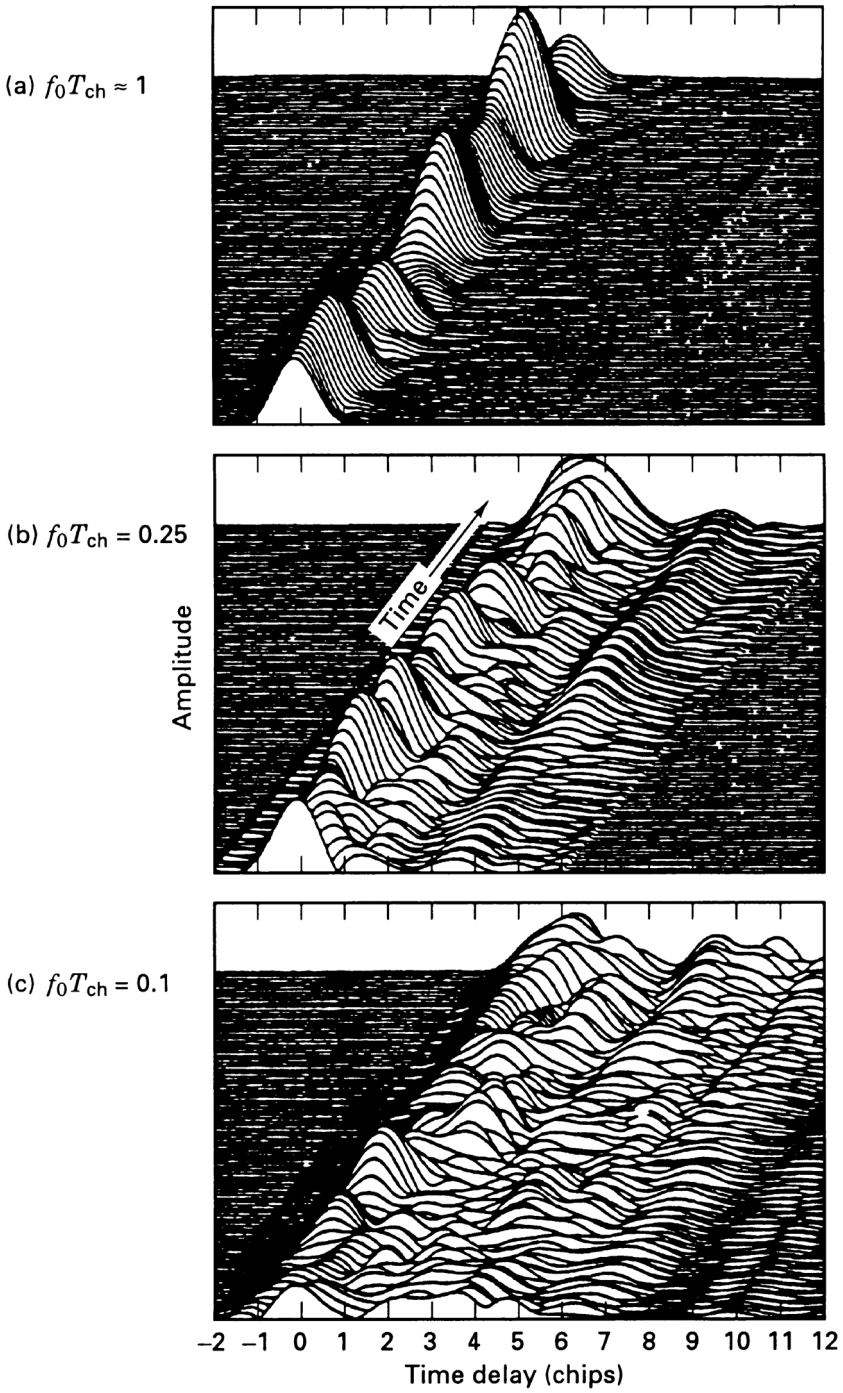


Figure 10

DS/SS matched-filter output time-history examples for three levels of channel conditions, where T_{ch} is the time duration of a chip [20].

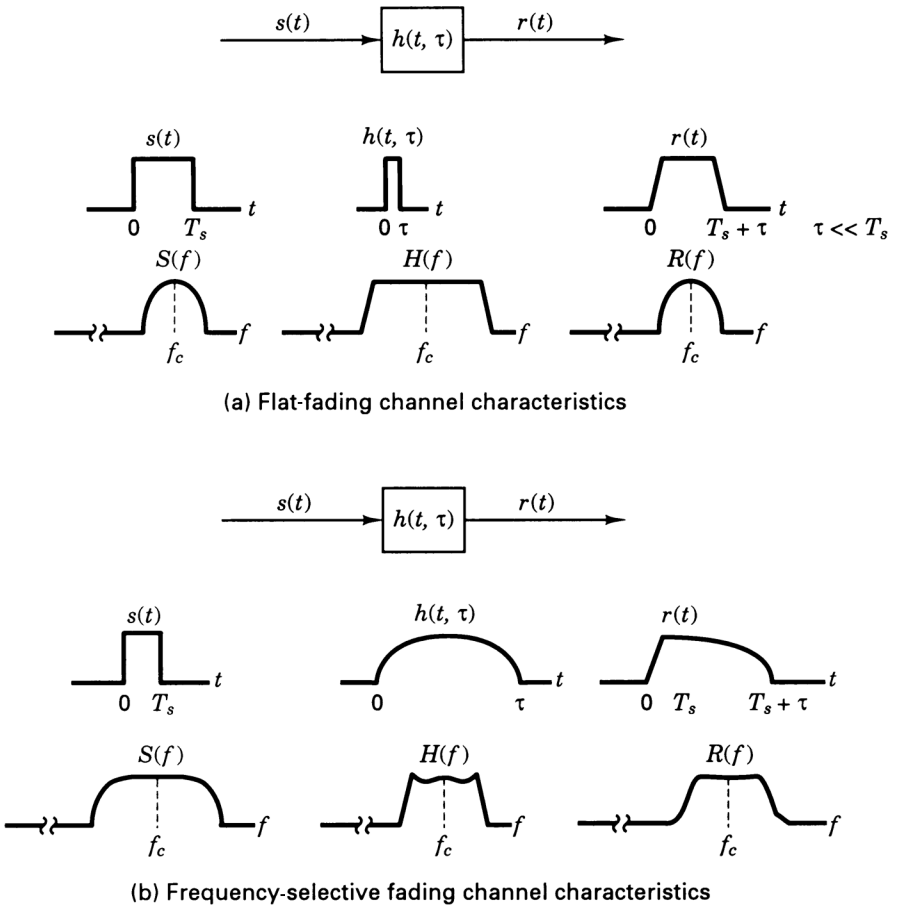


Figure 11

Flat-fading and frequency-selective fading characteristics [2].

Time Variance of the Channel Caused by Motion (Viewed in the Time Domain)

Signal dispersion and coherence bandwidth, described above, 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 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 [16]. 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 (that is, 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 follows [19]:

$$R(\Delta t) = J_0(kV\Delta t) \tag{23}$$

where $J_0(\bullet)$ is the zero-order Bessel function of the first kind [12], $V\Delta t$ 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 [18]. 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).

The Basic Fading Manifestations Are Independent of One Another

For a moving antenna, the fading of a transmitted 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 [18, 19]. 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 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, and so forth). Consider Figure 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.

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 article, it is interesting to note duality when examining time-domain versus frequency-domain relationships.

In Figure 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 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 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

1 as the duality between blocks 10 and 13, and in Figure 7 as the duality between the time-spreading mechanism in the frequency domain and the time-variant mechanism in the time domain.

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 7: *fast fading* and *slow fading.* The term *fast fading* is used to describe 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 to 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 to the time duration of a transmission symbol. Thus, one can expect the channel state to remain virtually unchanged during the time in which a symbol is transmitted. The propagating symbols likely will not suffer from the pulse distortion described above. The primary degradation in a slow-fading channel, as with flat fading, is loss in SNR.

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 7d shows a *Doppler power spectral density* (or Doppler spectrum)*,* $S(v)$, plotted as a function of Doppler-frequency shift, v. 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 as follows [19]:

$$S(\nu) = \frac{1}{\pi f_d \sqrt{1 - \left(\frac{\nu - f_c}{f_d} \right)^2}} \tag{24}$$

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 (24) is classically bowl-shaped, as shown in Figure 8d. Note that the spectral shape is a result of the dense-scatterer channel model. Equation (24) has been shown to match experimental data gathered for mobile radio channels [23]; 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 [24].

In Figure 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 traveling 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

$$f_d = \frac{V}{\lambda} \tag{25}$$

where V is relative velocity and λ is the signal wavelength. f_d is positive when the transmitter and receiver move toward each other, and negative when moving away from each other. 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 [2, 19].

$S(\nu)$ 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 [16]. 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 8, $S(v)$ 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 1 as the duality between blocks 7 and 16, and in Figure 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 $S(v)$ allows estimating 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 (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 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

$$T_0 \approx \frac{1}{f_d} \tag{26}$$

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 the following [3]:

$$T_0 \approx \frac{9}{16\pi f_d} \tag{27}$$

A popular rule defines T_0 as the geometric mean of Equations (26) and (27). This yields the following:

$$T_0 = \sqrt{\frac{9}{16\pi f_d^{\,2}}} = \frac{0.423}{f_d} \tag{28}$$

For the case of a 900 MHz mobile radio, Figure 12 illustrates the typical effect of Rayleigh fading on a signal's envelope amplitude versus time [2]. 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 12 and Equation (25), the time required to traverse a distance $\lambda/2$ (approximately the coherence time) when traveling at a constant velocity V, is as follows:

$$T_0 \approx \frac{\lambda/2}{V} = \frac{0.5}{f_d} \tag{29}$$

Thus, when the interval between fades is approximately $\lambda/2$, as in Figure 12, the resulting expression for T_0 in Equation (29) is quite close to the geometric mean shown in Equation (28). From Equation (29), and using the parameters shown in Figure 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 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.

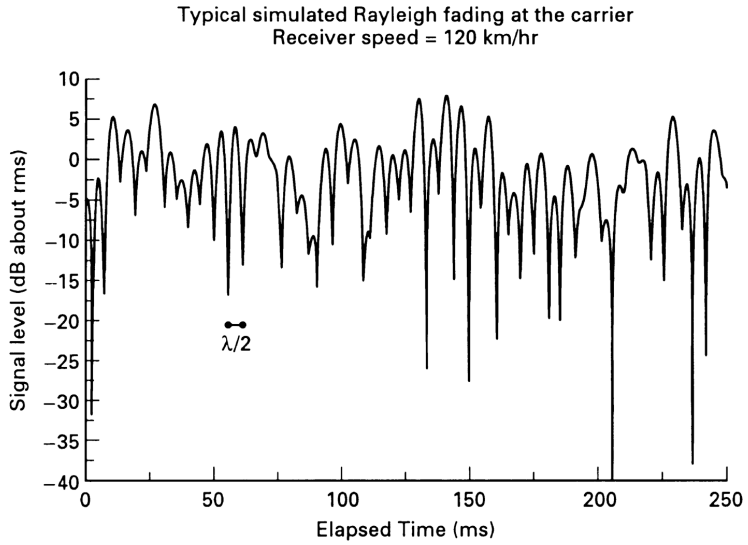


Figure 12

A typical Rayleigh fading envelope at 900 MHz [2].

Analogy for Spectral Broadening in Fading Channels

Let's 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 13 shows the keying of a digital signal (such as amplitude-shift keying or frequency-shift keying) 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 (that is, 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 13a by an ideal rectangular on-off (switching) function in Figure 13b. The frequency-domain description of this switching function is of the form sinc fT [1].

In Figure 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 13 with the sinc fT function of part (b), yielding the broadened spectrum depicted in part (c). Further, 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 the 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.

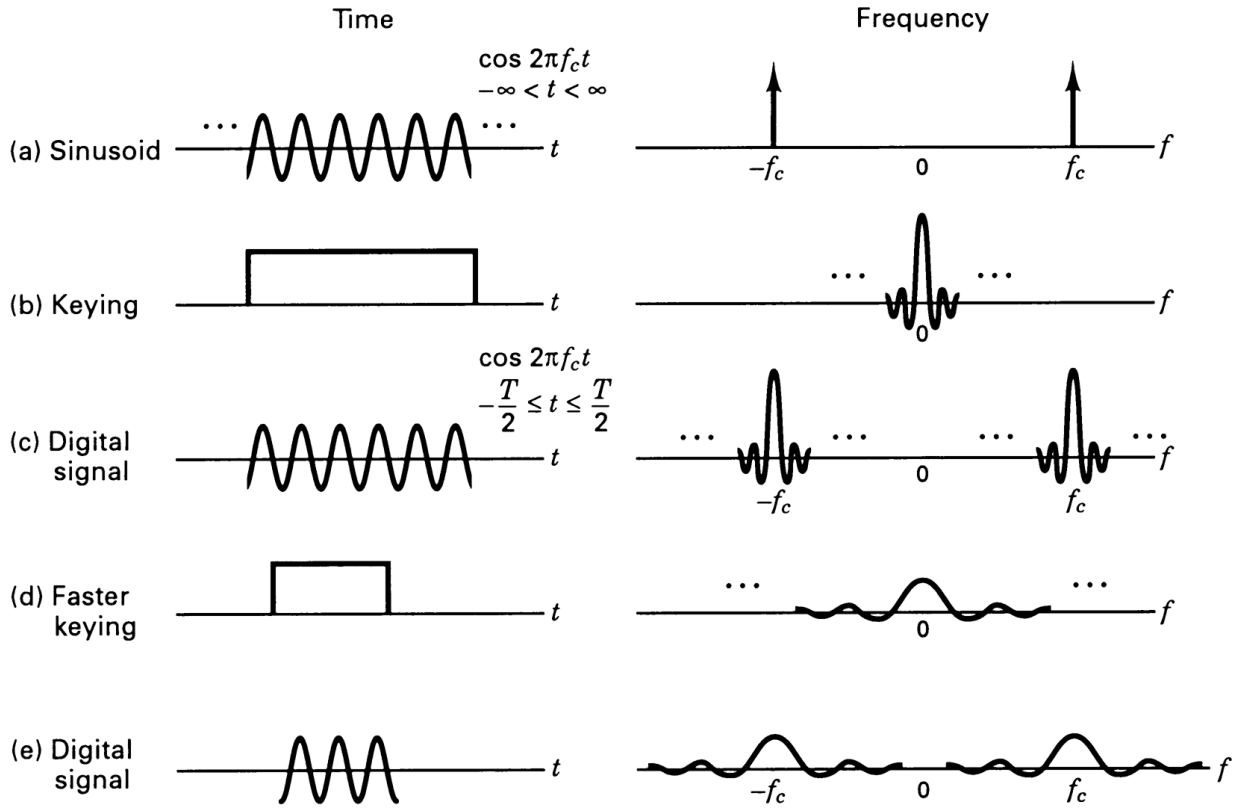


Figure 13

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

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

$$W < f_d \tag{30}$$

or

$$T_s > T_0 \tag{31}$$

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 ensuring that the signaling rate exceeds the channel fading rate. That is,

$$W > f_d \tag{32}$$

or

$$T_s < T_0 \tag{33}$$

Equation (22) showed 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 (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 (32) and (33) don't 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 $W \gg fd$ (or $T_s \ll T_0$). 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 [25]. 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 that can cause frequency swings much larger than $\pm V/\lambda$ when the information is recovered by an instantaneous frequency detector (a nonlinear device) [26]. Figure 14 illustrates how this can happen. At time t_1, owing to vehicle motion, 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 four times that of the specular Doppler alone. Therefore, the instantaneous frequency shift, $d\varphi/dt$, would be about four 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 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_0 = 850 MHz for various simulated mobile speeds [27]. 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-error 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 [28]. The exact factor depends on the signal modulation, receiver design, and required error rate [2, 26, 28-30]. Davarian [30] 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 keyed (DMSK) modulation.

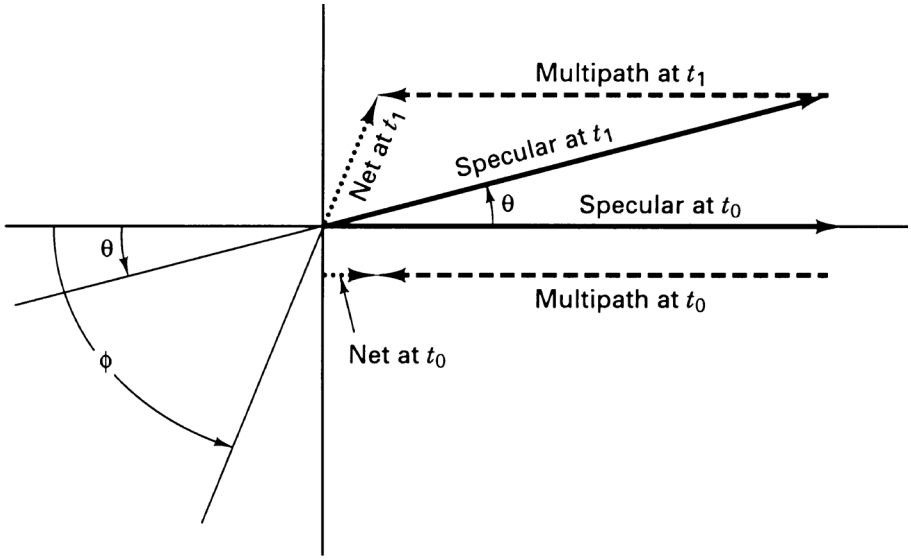


Figure 14

A combination of specular and multipath components can register much larger frequency swings than $\pm V/\lambda$ [26].

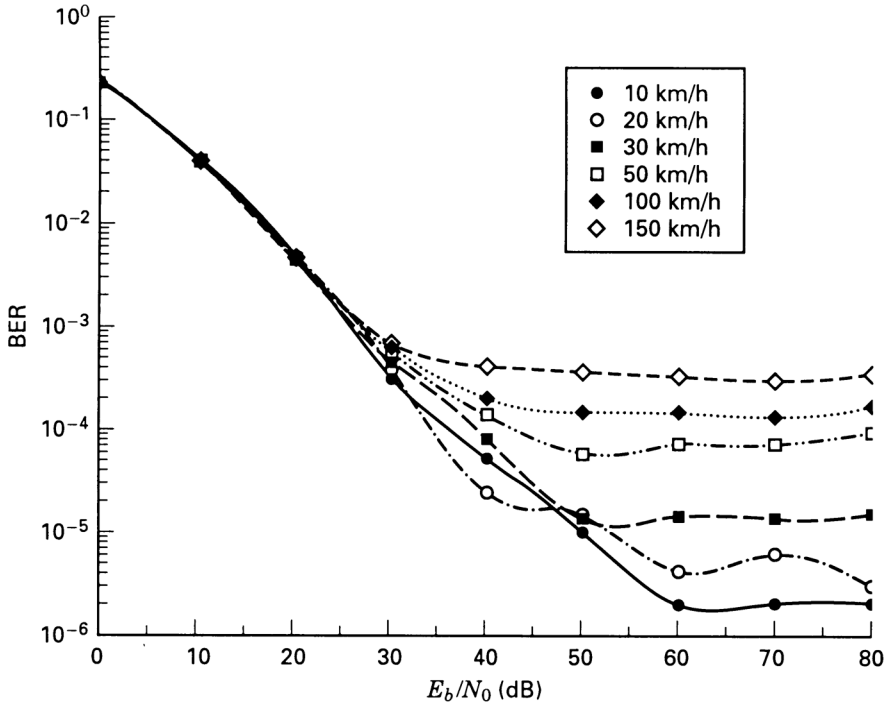


Figure 15

Error performance versus E_b/N_0 for π/4 DQPSK for various mobile speeds: f_c = 850 MHz, R_s = 24 ksymbol/s [27].

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 (3), a demodulated signal (neglecting noise) is described by Equation (10), which is rewritten below.

$$z(t) = \sum_n \alpha_n(t)\, e^{-j2\pi f_c \tau_n(t)}\, R[t - \tau_n(t)]\, e^{j\phi(t-\tau_n)} \tag{34a}$$

where $R(t) = |g(t)|$ is the envelope magnitude, and $\varphi(t)$ 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 (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 above, 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 test statistic $z(T)$ out of the demodulator in each signaling interval, including the noise $n_0(T)$, as follows:

$$z(T) = \alpha(T)\, R(T) e^{-j\left[\theta(T)-\phi(T)\right]} + n_0(T) \tag{34b}$$

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 versus E_b/N_0 for the basic coherent and noncoherent PSK and orthogonal FSK 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 16 depicts the performance curves. When $(E_b/N_0)E(\alpha^2) \gg 1$, where $E(\bullet)$ represents statistical expectation, the bit-error probability expressions for the basic binary signaling schemes shown in Figure 16 are given in Table 1 [11]. Each of the signaling schemes that manifests a waterfall-shaped performance plot under AWGN interference now exhibits performance that takes the form of inverse linear functions, as a result of the Rayleigh fading. The essence of slow, flat fading is seen in Figure 16, where the near-linear P_B versus E_b/N_0 curves, often referred to as the *Rayleigh limit*, illustrate that for an error probability of 10^{-4}, the average performance degradation is approximately 25 dB.

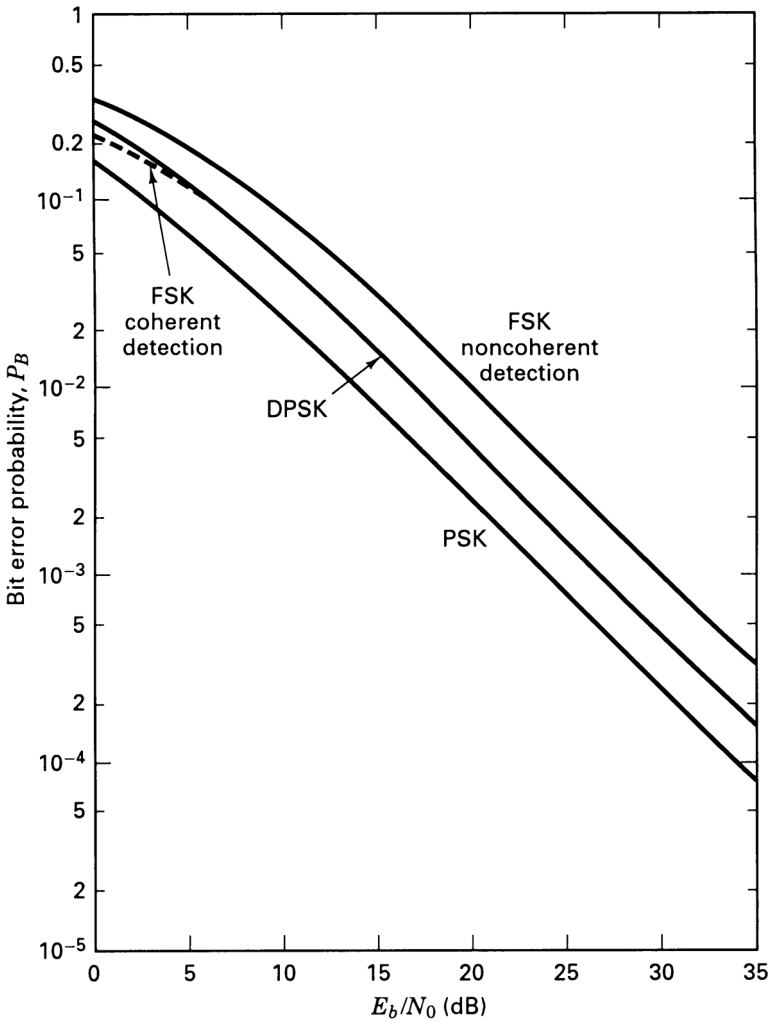


Figure 16

Performance of binary signaling over a slow Rayleigh fading channel [11].

Table 1

Rayleigh-Limit Bit-Error Performance

Where $(E_b/N_0)\mathbf{E}(\alpha^2) \gg 1$

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

Conclusion

This article has characterized the major elements that contribute to fading in certain communication channels. Figure 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 7. A mathematical model using correlation and power density functions was presented in Figure 8. This model yields a useful symmetry to help visualize the Fourier transform and duality relationships that describe the fading phenomena.

References

[1] Sklar, B., *Digital Communications: Fundamentals and Applications, Second Edition* (Upper Saddle River, NJ: Prentice-Hall, 2001).

[2] Rappaport, T.S., *Wireless Communications* (Upper Saddle River, NJ: Prentice-Hall, 1996).

[3] Greenwood, D. and Hanzo, L., "Characterisation of Mobile Radio Channels," *Mobile Radio Communications*, edited by R. Steele (London: Pentech Press, 1994).

[4] 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.

[5] 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.

[6] 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.

[7] Seidel, S.Y., et al., "Path Loss, Scattering and Multipath Delay Statistics in Four European Cities for Digital Cellular and Microcellular Radiotelephone," *IEEE Trans. on Vehicular Technology*, vol. 40, no. 4, November 1991, pp. 721-30.

[8] 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.

[9] Schilling, D.L., et al., "Broadband CDMA for Personal Communications Systems," *IEEE Communications Magazine*, vol. 29, no. 11, November 1991, pp. 86-93.

[10] Andersen, J.B., Rappaport, T.S., and Yoshida, S., "Propagation Measurements and Models for Wireless Communications Channels," *IEEE Communications Magazine*, vol. 33, no. 1, Jan. 1995, pp. 42-49.

[11] Proakis, J.G., *Digital Communications* (New York: McGraw-Hill, 1983).

[12] Schwartz, M., *Information, Transmission, Modulation, and Noise*, Second Edition (New York: McGraw Hill, 1970).

[13] 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.

[14] Bello, P.A., "Characterization of Randomly Time-Variant Linear Channels," *IEEE Trans. on Commun. Syst.*, Dec. 1963, pp. 360-93.

[15] Green, P.E., Jr., "Radar Astronomy Measurement Techniques," *MIT Lincoln Laboratory* (Lexington, Massachusetts) Tech. Report no. 282, December 1962.

[16] Pahlavan, K., and Levesque, A.H., *Wireless Information Networks*, (New York: John Wiley and Sons, 1995).

[17] Lee, W.Y.C., *Mobile Cellular Communications* (New York: McGraw-Hill, 1989).

[18] 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.

[19] Clarke, R.H., "A Statistical Theory of Mobile Radio Reception," *Bell System Technical Journal*, vol. 47, no. 6, July-August 1968, pp. 957-1000.

[20] 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.

[21] Amoroso, F., "The Bandwidth of Digital Data Signals," *IEEE Communications Magazine*, vol. 18, no. 6, November 1980, pp. 13-24.

[22] 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.

[23] Jakes, W.C. (Ed.), *Microwave Mobile Communications* (New York: John Wiley & Sons, 1974).

[24] *Joint Technical Committee of Committee T1 R1P1.4 and TIA TR46.3.3/TR45.4.4 on Wireless Access*, "Draft Final Report on RF Channel Characterization," Paper no. JTC(AIR)/94.01.17-238R4, January 17, 1994.

[25] 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.

[26] Amoroso, F., "Instantaneous Frequency Effects in a Doppler Scattering Environment," *IEEE International Conference on Communications*, June 7-10, 1987, pp. 1458-66.

[27] 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 Journal on Selected Areas in Communication*, vol. 11, no. 3, April 1993, pp. 393-405.

[28] Bateman, A.J. and McGeehan, J.P., "Data Transmission Over UHF Fading Mobile Radio Channels," *IEE Proceedings*, vol. 131, Pt. F, no. 4, July 1984, pp. 364-74.

[29] Feher, K. *Wireless Digital Communications* (Upper Saddle River, NJ: Prentice-Hall, 1995).

[30] 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.

About the Author

Bernard Sklar is the author of *Digital Communications: Fundamentals and Applications, Second Edition* (Prentice-Hall, 2001, ISBN 0-13-084788-7).