Chapter 4
Characterization of Communication Signals and Systems

by Simon Haykin
# Table of Contents

4.1 Representation of Band-Pass Signals and Systems
   4.1.1 Representation of Band-Pass Signals
   4.1.2 Representation of Linear Band-Pass Systems
   4.1.3 Response of a Band-Pass System to a Band-Pass Signal
   4.1.4 Representation of Band-Pass Stationary Stochastic Processes

4.2 Signal Space Representations
   4.2.1 Vector Space Concepts
   4.2.2 Signal Space Concepts
   4.2.3 Orthogonal Expansions of Signals
Table of Contents

4.3 Representation of Digitally Modulated Signals
   4.3.1 Memoryless Modulation Methods
   4.3.2 Linear Modulation with Memory
   4.3.3 Non-linear Modulation Methods with Memory – CPFSK and CPM

4.4 Spectral Characteristics of Digitally Modulated Signals
   4.4.1 Power Spectra of Linearly Modulated Signals
   4.4.2 Power Spectra of CPFSK and CPM Signals (*)
   4.4.3 Power Spectra of Modulated Signals with Memory (*)
4.1 Representation of Band-Pass Signals and Systems

- The channel over which the signal is transmitted is limited in bandwidth to an interval of frequencies centered about the carrier.

- Signals and channels (systems) that satisfy the condition that their bandwidth is much smaller than the carrier frequency are termed narrowband band-pass signals and channels (systems).

- With no loss of generality and for mathematical convenience, it is desirable to reduce all band-pass signals and channels to equivalent low-pass signals and channels.
4.1.1 Representation of Band-Pass Signals

Suppose that a real-valued signal \( s(t) \) has a frequency content concentrated in a narrow band of frequencies in the vicinity of a frequency \( f_c \), as shown in the following figure:

Our object is to develop a mathematical representation of such signals.
A signal that contains only the positive frequencies in $s(t)$ may be expressed as:

$$S_+(f) = 2u(f)S(f)$$

$$s_+(t) = \int_{-\infty}^{\infty} S_+(f) \cdot e^{j2\pi ft} df$$

$$= F^{-1}\left[2u(f)\right] \ast F^{-1}\left[S(f)\right]$$

where $S(f)$ is the Fourier transform of $s(t)$ and $u(f)$ is the unit step function, and the signal $s_+(t)$ is called the analytic signal or the pre-envelope of $s(t)$.

$$F^{-1}\left[2u(f)\right] = \delta(t) + \frac{j}{\pi t}$$

$$s_+(t) = \left[\delta(t) + \frac{j}{\pi t}\right] \ast s(t) = s(t) + j\frac{1}{\pi t} \ast s(t) \equiv s(t) + \hat{s}(t)$$
4.1.1 Representation of Band-Pass Signals

- Define:
  \[
  \hat{s}(t) = \frac{1}{\pi t} * s(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau
  \]

- A filter, called a **Hilbert transformer**, is defined as:
  \[
  h(t) = \frac{1}{\pi t}, \quad -\infty < t < \infty
  \]

- The signal \( \hat{s}(t) \) may be viewed as the output of the Hilbert transformer when excited by the input signal \( s(t) \).

- The frequency response of this filter is:
  \[
  H(f) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} dt = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{t} e^{-j2\pi ft} dt = \begin{cases} 
  -j & (f > 0) \\
  0 & (f = 0) \\
  j & (f < 0)
  \end{cases}
  \]
4.1.1 Representation of Band-Pass Signals

- We observe that $|H(f)|=1$ and the phase response $\Theta(f)=-\pi/2$ for $f>0$ and $\Theta(f)=\pi/2$ for $f<0$. Thus, this filter is basically a 90 degrees phase shifter for all frequencies in the input signal.

- The analytic signal $s_{+}(t)$ is a band-pass signal. To obtain an equivalent low-pass representation, we define:

$$S_{l}(f) = S_{+}(f + f_c)$$

$$s_{l}(t) = s_{+}(t) e^{-j2\pi f_c t} = \left[s(t) + \hat{s}(t)\right] e^{-j2\pi f_c t}$$

- In general, $s_{l}(t)$ is complex-valued:

$$s_{l}(t) = x(t) + jy(t)$$

$$s(t) = x(t) \cos 2\pi f_c t - y(t) \sin 2\pi f_c t$$

$$\hat{s}(t) = x(t) \sin 2\pi f_c t + y(t) \cos 2\pi f_c t$$
4.1.1 Representation of Band-Pass Signals

- $s(t) = x(t) \cos 2 \pi f_c t - y(t) \sin 2 \pi f_c t$ is the desired form for the representation of a band-pass signal. The low-frequency signal components $x(t)$ and $y(t)$ may be viewed as amplitude modulations impressed on the carrier components $\cos 2 \pi f_c t$ and $\sin 2 \pi f_c t$, respectively.

- $x(t)$ and $y(t)$ are called the *quadrature components* of the band-pass signal $s(t)$.

- $s(t)$ can also be written as:

  $$s(t) = \text{Re}\left\{ x(t) + jy(t) \right\} = \text{Re}\left[ s_l(t)e^{j2\pi f_c t} \right]$$

- The low pass signal $s_l(t)$ is usually called the *complex envelope* of the real signal $s(t)$ and is basically the equivalent low-pass signal.
4.1.1 Representation of Band-Pass Signals

- $s_f(t)$ can be also be written as:

$$s_f(t) = a(t) e^{j\theta(t)}$$

where $a(t) = \sqrt{x^2(t) + y^2(t)}$ and $\theta(t) = \tan^{-1} \frac{y(t)}{x(t)}$

- $s(t)$ can be represented as:

$$s(t) = \text{Re} \left[ s_f(t) e^{j2\pi f_c t} \right] = \text{Re} \left[ a(t) e^{j[2\pi f_c t + \theta(t)]} \right]$$

$$= a(t) \cos \left[ 2\pi f_c t + \theta(t) \right]$$

$a(t)$ is called the *envelope* of $s(t)$, and $\theta(t)$ is called the *phase* of $s(t)$.
4.1.1 Representation of Band-Pass Signals

Three equivalent representations of band-pass signals:

\[ s(t) = x(t) \cos 2\pi f_c t - y(t) \sin 2\pi f_c t \]
\[ = \text{Re}\left[ s_l(t) e^{j2\pi f_c t} \right] \]
\[ = a(t) \cos\left[ 2\pi f_c t + \theta(t) \right] \]

The Fourier transform of \( s(t) \) is:

\[
S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt = \int_{-\infty}^{\infty} \left\{ \text{Re}\left[ s_l(t) e^{j2\pi f_c t} \right] \right\} e^{-j2\pi ft} dt
\]

\[
S(f) = \frac{1}{2} \int_{-\infty}^{\infty} \left[ s_l(t) e^{j2\pi f_c t} + s_l^*(t) e^{-j2\pi f_c t} \right] e^{-j2\pi ft} dt
\]

\[ = \frac{1}{2} \left[ S_l(f - f_c) + S_l^*(-f - f_c) \right] \]
4.1.1 Representation of Band-Pass Signals

The energy in the signal \( s(t) \) is defined as:

\[
\varepsilon = \int_{-\infty}^{\infty} s^2(t) \, dt = \int_{-\infty}^{\infty} \left\{ \text{Re} \left[ s_l(t) e^{j2\pi f_c t} \right] \right\}^2 \, dt
\]

\[
\varepsilon = \frac{1}{4} \int_{-\infty}^{\infty} \left[ s_l^2 e^{j4\pi f_c t} + 2s_ls_l^* + (s_l^*)^2 e^{-j4\pi f_c t} \right] dt
\]

\[
= \frac{1}{2} \int_{-\infty}^{\infty} \left| s_l(t) \right|^2 dt + \frac{1}{4} \int_{-\infty}^{\infty} \left[ a^2(t) e^{j4\pi f_c t + 2\theta(t)} + (a^*(t))^2 e^{-j4\pi f_c t + 2\theta(t)} \right] dt
\]

\[
= \frac{1}{2} \int_{-\infty}^{\infty} \left| s_l(t) \right|^2 dt + \frac{1}{2} \int_{-\infty}^{\infty} \left| s_l(t) \right|^2 \cos \left[ 4\pi f_c t + 2\theta(t) \right] dt
\]

where \( \left| s_l(t) \right|^2 = a^2(t) = (a^*(t))^2 \)
Since the signal $s(t)$ is narrow-band, the real envelope $a(t)=|s_I(t)|$
or, equivalently, $a^2(t)$ varies slowly relative to the rapid
variations exhibited by the cosine function.

The net area contributed by the second integral is very small
relative to the value of the first integral, hence, it can be
neglected.

\[ \varepsilon = \frac{1}{2} \int_{-\infty}^{\infty} |s_I(t)|^2 \, dt \quad (4.1-24) \]
A linear filter or system may be described either by its impulse response $h(t)$ or by its frequency response $H(f)$, which is the Fourier transform of $h(t)$. Since $h(t)$ is real, $H^*(-f)=H(f)$, because:

$$H^*(-f) = \left(\int_{-\infty}^{\infty} h(t) e^{-j2\pi f t} dt\right)^*$$

$$= \int_{-\infty}^{\infty} h^*(t) e^{-j2\pi f t} dt$$

Note that $h(t)$ is real.

$$= \int_{-\infty}^{\infty} h(t) e^{-j2\pi f t} dt = H(f)$$

Define $H_l(f-f_c) = \begin{cases} H(f) & (f > 0) \\ 0 & (f < 0) \end{cases}$

then $H_l^*(-f-f_c) = \begin{cases} 0 & (f > 0) \\ H^*(-f) = H(f) & (f < 0) \end{cases}$
As a result

\[ H(f) = H_l(f - f_c) + H_l^*(-f - f_c) \]

thus

\[ h(t) = h_l(t)e^{j2\pi f_c t} + h_l^* e^{-j2\pi f_c t} \]

\[ = 2 \text{Re} \left[ h_l(t)e^{j2\pi f_c t} \right] \]

where

\[ \int_{-\infty}^{\infty} H_l^*(-f - f_c) e^{j2\pi f t} df \quad \text{using} \quad x = -f - f_c \]

\[ = \int_{-\infty}^{\infty} H_l^*(x) e^{-j2\pi f_c x} dx \cdot e^{-j2\pi f_c t} \]

\[ = \left( \int_{-\infty}^{\infty} H_l(x) e^{j2\pi f_c x} dx \right)^* \cdot e^{-j2\pi f_c t} = h_l^*(t) \cdot e^{-j2\pi f_c t} \]

In general, the impulse response \( h_l(t) \) of the equivalent low-pass system is complex-valued.
We have shown in Sections 4.1.1 and 4.1.2 that narrow band band-pass signals and systems can be represented by equivalent low-pass signals and systems.

We demonstrate in this section that the output of a band-pass system to a band-pass input signal is simply obtained from the equivalent low-pass input signal and the equivalent low-pass impulse response of the system.

The output of the band-pass system is also a band-pass signal, and, therefore, it can be expressed in the form:

\[ r(t) = \int_{-\infty}^{\infty} s(\tau) h(t - \tau) \, d\tau \]

where \( r(t) \) is related to the input signal \( s(t) \) and the impulse response \( h(t) \) by the convolution integral.

\[ r(t) = \text{Re} \left[ r_l(t) e^{j2\pi f_c t} \right] \]
4.1.3 Response of a Band-Pass System to a Band-Pass Signal

The output of the system in the frequency domain is:

\[ R(f) = S(f)H(f) \]


\[ = \frac{1}{2} \left[ S_l(f-f_c) + S_l^*(-f-f_c) \right] \left[ H_l(f-f_c) + H_l^*(-f-f_c) \right] \]

For a narrow band signal, \( S_l(f-f_c) \approx 0 \) for \( f < 0 \) and \( H_l^*(-f-f_c) = 0 \) for \( f > 0 \).

\[ S_l(f-f_c)H_l^*(-f-f_c) = 0 \]

4.1-27

For a narrow band signal, \( S_l^*(-f-f_c) \approx 0 \) for \( f > 0 \) and \( H_l(f-f_c) = 0 \) for \( f < 0 \).

\[ S_l^*(-f-f_c)H_l(f-f_c) = 0 \]

4.1-26

\[ R(f) = \frac{1}{2} \left[ S_l(f-f_c)H_l(f-f_c) + S_l^*(-f-f_c)H_l^*(-f-f_c) \right] \]

\[ = \frac{1}{2} \left[ R_l(f-f_c) + R_l^*(-f-f_c) \right] \]

\[ R_l(f) \equiv S_l(f)H_l(f) \]

\[ r_l(t) = \int_{-\infty}^{\infty} s_l(\tau) h_l(t-\tau) d\tau \]
4.1.4 Representation of Band-Pass Stationary Stochastic Processes

In this section, we extend the representation to sample functions of a *band-pass stationary stochastic process*. In particular, we derive the relations between the correlation functions and power spectra of the band-pass signal and the correlation function and power spectra of the equivalent low-pass signal.

Suppose that $n(t)$ is a sample function of a wide-sense stationary stochastic process with zero mean and power spectral density $\Phi_{nn}(f)$. The power spectral density is assumed to be zero outside of an interval of frequencies centered around $f_c$, where $f_c$ is termed the *carrier frequency*. The stochastic process $n(t)$ is said to be a *narrowband band-pass process* if the width of the spectral density is much smaller than $f_c$. 
4.1.4 Representation of Band-Pass Stationary Stochastic Processes

- Under this condition, a sample function of the process \( n(t) \) can be represented by the following equations from Section 4.1.1:

\[
\begin{align*}
n(t) &= a(t) \cos[2\pi f_c t + \theta(t)] \\
&= x(t) \cos 2\pi f_c t - y(t) \sin 2\pi f_c t \\
&= \text{Re}\left[ z(t) e^{j2\pi f_c t} \right]
\end{align*}
\]

- \( a(t) \) is the envelope and \( \theta(t) \) is the phase of the real-valued signal.
- \( x(t) \) and \( y(t) \) are the quadrature components of \( n(t) \).
- \( z(t) \) is called the complex envelope of \( n(t) \).
- If \( n(t) \) is zero mean, then \( x(t) \) and \( y(t) \) must also have zero mean values.
- The stationarity of \( n(t) \) implies that:

\[
\begin{align*}
\phi_{xx}(\tau) &= \phi_{yy}(\tau) \\
\phi_{xy}(\tau) &= -\phi_{yx}(\tau)
\end{align*}
\]

Proved next.
4.1.4 Representation of Band-Pass Stationary Stochastic Processes

Proof of $\phi_{xx}(\tau) = \phi_{yy}(\tau)$ and $\phi_{xy}(\tau) = -\phi_{yx}(\tau)$

Autocorrelation function of $n(t)$ is:

$\phi_{nn}(\tau) = E\left[ n(t)n(t+\tau) \right]$

$= E\left\{ \left[ x(t)\cos 2\pi f_c t - y(t)\sin 2\pi f_c t \right] \times \left[ x(t+\tau)\cos 2\pi f_c (t+\tau) - y(t+\tau)\sin 2\pi f_c (t+\tau) \right] \right\}$

$= \phi_{xx}(\tau)\cos 2\pi f_c t \cos 2\pi f_c (t+\tau) + \phi_{yy}(\tau)\sin 2\pi f_c t \sin 2\pi f_c (t+\tau)$

$- \phi_{xy}(\tau)\sin 2\pi f_c t \cos 2\pi f_c (t+\tau) - \phi_{yx}(\tau)\cos 2\pi f_c t \sin 2\pi f_c (t+\tau)$

by using:

$\cos A \cos B = \frac{1}{2} \left[ \cos (A-B) + \cos (A+B) \right]$  

$\sin A \sin B = \frac{1}{2} \left[ \cos (A-B) - \cos (A+B) \right]$  

$\sin A \cos B = \frac{1}{2} \left[ \sin (A-B) + \sin (A+B) \right]$
4.1.4 Representation of Band-Pass Stationary Stochastic Processes

We can obtain:

\[ \phi_{nn}(\tau) = E\left[n(t)n(t + \tau)\right] \]

\[ = \frac{1}{2}\left[\phi_{xx}(\tau) + \phi_{yy}(\tau)\right]\cos 2\pi f_c \tau + \frac{1}{2}\left[\phi_{xx}(\tau) - \phi_{yy}(\tau)\right]\cos 2\pi f_c (2t + \tau) \]

\[ - \frac{1}{2}\left[\phi_{yx}(\tau) - \phi_{xy}(\tau)\right]\sin 2\pi f_c \tau - \frac{1}{2}\left[\phi_{yx}(\tau) + \phi_{xy}(\tau)\right]\sin 2\pi f_c (2t + \tau) \]

Since \( n(t) \) is stationary, the right-hand side must be independent of \( t \).

As a result, \( \phi_{xx}(\tau) = \phi_{yy}(\tau) \) and \( \phi_{xy}(\tau) = -\phi_{yx}(\tau) \) [Q.E.D.]

Therefore, \( \phi_{nn}(\tau) = \phi_{xx}(\tau)\cos 2\pi f_c \tau - \phi_{yx}(\tau)\sin 2\pi f_c \tau \)

Note that this equation is identical in form to:

\[ n(t) = x(t)\cos 2\pi f_c t - y(t)\sin 2\pi f_c t \]
The autocorrelation function of the equivalent low-pass process $z(t) = x(t) + jy(t)$ is defined as:

$$\phi_{zz}(\tau) = \frac{1}{2} E\left[ z^*(t) z(t + \tau) \right]$$

$$\phi_{zz}(\tau) = \frac{1}{2} \left[ \phi_{xx}(\tau) + \phi_{yy}(\tau) - j\phi_{xy}(\tau) + j\phi_{yx}(\tau) \right]$$

Since $\phi_{xx}(\tau) = \phi_{yy}(\tau)$ and $\phi_{xy}(\tau) = -\phi_{yx}(\tau)$ we obtain:

$$\phi_{zz}(\tau) = \phi_{xx}(\tau) + j\phi_{yx}(\tau)$$

This equation relates the autocorrelation function of the complex envelope to the autocorrelation and cross-correlation functions of the quadrature components.
By combining $\phi_{zz}(\tau) = \phi_{xx}(\tau) + j\phi_{yx}(\tau)$
and $\phi_{nn}(\tau) = \phi_{xx}(\tau)\cos 2\pi f_c \tau - \phi_{yx}(\tau)\sin 2\pi f_c \tau$
we can obtain: $\phi_{nn}(\tau) = \text{Re}[\phi_{zz}(\tau)e^{j2\pi f_c \tau}]$
Therefore, the autocorrelation function $\phi_{nn}(\tau)$ of the band-pass stochastic process is uniquely determined from the autocorrelation function $\phi_{zz}(\tau)$ of the equivalent low-pass process $z(t)$ and the carrier frequency $f_c$.
The power density spectrum of the stochastic process $n(t)$ is:

$$
\Phi_{nn}(f) = \int_{-\infty}^{\infty} \left\{ \text{Re}\left[ \phi_{zz}(\tau)e^{j2\pi f_c \tau} \right] \right\} e^{-j2\pi f \tau} \, d\tau
= \frac{1}{2} \left[ \Phi_{zz}(f - f_c) + \Phi_{zz}(-f - f_c) \right]
$$
Properties of the quadrature components

\[ \phi_{yx}(\tau) = \phi_{xy}(-\tau) \quad (2.2-10) \]
\[ \phi_{xy}(\tau) = -\phi_{yx}(\tau) \quad (4.1-41) \]

\[ \Rightarrow \quad \phi_{xy}(\tau) = -\phi_{xy}(-\tau) \]
\[ \Rightarrow \quad \phi_{xy}(\tau) \text{ is an odd function of } \tau \text{ and } \phi_{xy}(0) = 0. \]

\[ \phi_{zz}(\tau) = \frac{1}{2} [\phi_{xx}(\tau) + \phi_{yy}(\tau) - j\phi_{xy}(\tau) + j\phi_{yx}(\tau)] \]

- If \( \phi_{xy}(\tau) = 0 \) for all \( \tau \), then \( \phi_{zz}(\tau) \) is real (from 4.1-48) and the power spectral density satisfies \( \Phi_{zz}(f) = \Phi_{zz}(-f) \) (i.e. \( \Phi_{zz}(f) \) is symmetric about \( f = 0 \)).

- If \( n(t) \) is Gaussian, \( x(t) \) and \( y(t + \tau) \) are jointly Gaussian. For \( \tau = 0 \), they are statistically independent, and the joint PDF is:

\[ p(x, y) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2 + y^2\right)/2\sigma^2}, \quad \sigma^2 = \phi_{xx}(0) = \phi_{yy}(0) = \phi_{nn}(0) \]
Representation of white noise

White noise is a stochastic process that is defined to have a flat (constant) power spectral density over the entire frequency range. This type of noise can’t be expressed in terms of quadrature components, as a result of its wideband character.

In the demodulation of narrowband signals in noise, it is mathematically convenient to model the additive noise process as white and to represent the noise in terms of quadrature components. This can be accomplished by postulating that the signals and noise at the receiving terminal have passed through an ideal band-pass filter.
Representation of white noise (cont.)

- The noise resulting from passing the white noise process through a spectrally band-pass filter is termed **band-pass white noise** and has the power spectral density:

\[
\Phi_{nn}(f) = \frac{1}{2} N_0 
\]

- The band-pass white noise can be represented by:

\[
\begin{align*}
n(t) &= a(t) \cos \left[ 2\pi f_c t + \theta(t) \right] \\
&= x(t) \cos 2\pi f_c t - y(t) \sin 2\pi f_c t \\
&= \text{Re} \left[ z(t) e^{j2\pi f_c t} \right]
\end{align*}
\]
4.1.4 Representation of Band-Pass Stationary Stochastic Processes

**Representation of white noise (cont.)**

- The equivalent low-pass noise $z(t)$ has a power spectral density:

$$\Phi_{zz}(f) = \begin{cases} N_0 \left( |f| \leq \frac{1}{2} B \right) \\
0 \quad \left( |f| > \frac{1}{2} B \right) \end{cases}$$

$$\phi_{zz}(\tau) = N_0 \frac{\sin \pi B \tau}{\pi \tau}$$

and

$$\phi_{zz}(\tau) = \lim_{B \to \infty} N_0 \delta(\tau)$$

- The power spectral density for white noise and band-pass white noise is symmetric about $f=0$, so $\phi_{yx}(\tau) = 0$ for all $\tau$.

$$\phi_{zz}(\tau) = \phi_{xx}(\tau) = \phi_{yy}(\tau) \quad \text{(from 4.1-48)}$$
4.2 Signal Space Representations

We will demonstrate that signals have characteristics that are similar to vectors and develop a vector representation for signal waveforms.

4.2.1 Vector Space Concepts

A vector $\mathbf{v}$ in an $n$-dimensional space is characterized by its $n$ components $[v_1 \ v_2 \ \cdots \ v_n]$ and may also be represented as a linear combination of unit vectors $\mathbf{e}_i$, $1 \leq i \leq n$,

$$\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i$$

The inner product of two $n$-dimensional vectors is defined as:

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \sum_{i=1}^{n} v_{1i} v_{2i}$$
A set of \( m \) vectors \( v_k, 1 \leq k \leq m \) are orthogonal if:
\[
v_i \cdot v_j = 0 \quad \text{for all} \quad 1 \leq i, j \leq m, \quad \text{and} \quad i \neq j.
\]

The \textit{norm} of a vector \( v \) is denoted by \( ||v|| \) and is defined as:
\[
||v|| = (v \cdot v)^{1/2} = \sqrt{\sum_{i=1}^{n} v_i^2}
\]

A set of \( m \) vectors is said to be \textit{orthonormal} if the vectors are orthogonal and each vector has a unit norm.

A set of \( m \) vectors is said to be \textit{linearly independent} if no one vector can be represented as a linear combination of the remaining vectors.

Two \( n \)-dimensional vectors \( v_1 \) and \( v_2 \) satisfy the triangle inequality:
\[
||v_1 + v_2|| \leq ||v_1|| + ||v_2||
\]
4.2.1 Vector Space Concepts

- **Cauchy-Schwarz inequality:**
  \[ | \mathbf{v}_1 \cdot \mathbf{v}_2 | \leq \| \mathbf{v}_1 \| + \| \mathbf{v}_2 \| \]

- The norm square of the sum of two vectors may be expressed as:
  \[ \| \mathbf{v}_1 + \mathbf{v}_2 \|^2 = \| \mathbf{v}_1 \|^2 + \| \mathbf{v}_2 \|^2 + 2 \mathbf{v}_1 \cdot \mathbf{v}_2 \]

- **Linear transformation** in an \( n \)-dimensional vector space:
  \[ \mathbf{v}' = A \mathbf{v} \]

- In the special case where \( \mathbf{v}' = \lambda \mathbf{v} \), \( A \mathbf{v} = \lambda \mathbf{v} \)
  the vector \( \mathbf{v} \) is called an *eigenvector* and \( \lambda \) is the corresponding *eigenvalue*. 
4.2.1 Vector Space Concepts

**Gram-Schmidt procedure** for constructing a set of orthonormal vectors.

- Arbitrarily selecting a vector $v_1$ and normalizing its length:
  $$u_1 = \frac{v_1}{\|v_1\|}$$

- Select $v_2$ and subtract the projection of $v_2$ onto $u_1$.
  $$u'_2 = v_2 - (v_2 \cdot u_1)u_1$$

- Normalize the vector $u'_2$ to unit length.
  $$u_2 = \frac{u'_2}{\|u'_2\|}$$

- Selecting $v_3$:
  $$u'_3 = v_3 - (v_3 \cdot u_1)u_1 - (v_3 \cdot u_2)u_2$$
  $$u_3 = \frac{u'_3}{\|u'_3\|}$$

- By continuing this procedure, we construct a set of orthonormal vectors.
4.2.2 Signal Space Concepts

- The *inner product* of two generally complex-valued signals \( x_1(t) \) and \( x_2(t) \) is denote by \( < x_1(t), x_2(t) > \) and defined as:

\[
<x_1(t), x_2(t)> = \int_{a}^{b} x_1(t)x_2^*(t) \, dt
\]

- The signals are *orthogonal* if their inner product is zero.

- The *norm* of a signal is defined as:

\[
\| x(t) \| = \left( \int_{a}^{b} |x(t)|^2 \, dt \right)^{1/2}
\]

- A set of \( m \) signals are *orthonormal* if they are orthogonal and their norms are all unity.

- A set of \( m \) signals is *linearly independent*, if no signal can be represented as a linear combination of the remaining signals.
4.2.2 Signal Space Concepts

The *triangle inequality* for two signals is:

\[
\|x_1(t) + x_2(t)\| \leq \|x_1(t)\| + \|x_2(t)\|
\]

The *Cauchy-Schwarz inequality* is:

\[
\left| \int_a^b x_1(t) x_2^*(t) \, dt \right| \leq \left( \int_a^b |x_1(t)|^2 \, dt \right)^{1/2} \left( \int_a^b |x_2(t)|^2 \, dt \right)^{1/2}
\]

with equality when \(x_2(t) = ax_1(t)\), where \(a\) is any complex number.
Suppose that \( s(t) \) is a deterministic, real-valued signal with finite energy:

\[
\mathcal{E}_s = \int_{-\infty}^{\infty} \left[ s(t) \right]^2 dt
\]

Suppose that there exists a set of functions \( \{f_n(t), n=1,2,\ldots,K\} \) that are orthonormal in the sense that:

\[
\int_{-\infty}^{\infty} f_n(t) f_m(t) dt = \begin{cases} 
0 & (m \neq n) \\
1 & (m = n)
\end{cases}
\]

We may approximate the signal \( s(t) \) by a weighted linear combination of these functions, i.e.,

\[
s(t) = \sum_{k=1}^{K} s_k f_k(t)
\]
The approximation error incurred is:

\[ e(t) = s(t) - \hat{s}(t) \]

The energy of the approximation error:

\[ \varepsilon_e = \int_{-\infty}^{\infty} \left[ s(t) - \hat{s}(t) \right]^2 dt = \int_{-\infty}^{\infty} \left[ s(t) - \sum_{k=1}^{K} s_k f_k(t) \right]^2 dt \quad (4.2-24) \]

To minimize the energy of the approximation error, the optimum coefficients in the series expansion of \( s(t) \) may be found by:

- Differentiating Equation 4.2-24 with respect to each of the coefficients \( \{s_k\} \) and setting the first derivatives to zero.
- Use a well-known result from estimation theory based on the mean-square-error criterion, which is that the minimum of \( \varepsilon_e \) with respect to the \( \{s_k\} \) is obtained when the error is orthogonal to each of the functions in the series expansion.
Using the second approach, we have:

\[
\int_{-\infty}^{\infty} \left[ s(t) - \sum_{k=1}^{K} s_k f_k(t) \right] f_n(t) \, dt = 0, \quad n = 1, 2, \ldots, K
\]

Since the functions \( \{f_n(t)\} \) are orthonormal, we have:

\[
s_n = \int_{-\infty}^{\infty} s(t) f_n(t) \, dt, \quad n = 1, 2, \ldots, K
\]

Thus, the coefficients are obtained by projecting the signals \( s(t) \) onto each of the functions.

The minimum mean square approximation error is:

\[
\mathcal{E}_{\text{min}} = \int_{-\infty}^{\infty} e(t) s(t) \, dt = \int_{-\infty}^{\infty} \left[ s(t) \right]^2 \, dt - \int_{-\infty}^{\infty} \sum_{k=1}^{K} s_k f_k(t) s(t) \, dt
\]

\[
= \mathcal{E}_s - \sum_{k=1}^{K} s_k^2
\]
4.2.3 Orthogonal Expansions of Signals

- When the minimum mean square approximation error $\varepsilon_{\text{min}}=0$,

$$
\varepsilon_s = \sum_{k=1}^{K} s_k^2 = \int_{-\infty}^{\infty} [s(t)]^2 \, dt
$$

- Under such condition, we may express $s(t)$ as:

$$
s(t) = \sum_{k=1}^{K} s_k f_k(t)
$$

- When every finite energy signal can be represented by a series expansion of the form for which $\varepsilon_{\text{min}}=0$, the set of orthonormal functions $\{f_n(t)\}$ is said to be complete.
**Example 4.2-1 Trigonometric Fourier Series:**

Consider a finite energy signal \( s(t) \) that is zero everywhere except in the range \( 0 \leq t \leq T \) and has a finite number of discontinuities in this interval. Its periodic extension can be represented in a Fourier series as:

\[
s(t) = \sum_{k=0}^{\infty} \left( a_k \cos \frac{2\pi k t}{T} + b_k \sin \frac{2\pi k t}{T} \right)
\]

where the coefficients \( \{a_k, b_k\} \) that minimize the mean square error are given by:

\[
a_k = \frac{1}{\sqrt{T}} \int_0^T s(t) \cos \frac{2\pi k t}{T} \, dt, \quad b_k = \frac{1}{\sqrt{T}} \int_0^T s(t) \sin \frac{2\pi k t}{T} \, dt
\]

The set of trigonometric functions \( \left\{ \sqrt{2/T} \cos 2\pi k t / T, \sqrt{2/T} \sin 2\pi k t / T \right\} \) is complete, and the series expansion results in zero mean square error.
4.2.3 Orthogonal Expansions of Signals

**Gram-Schmidt procedure**

- Constructing a set of orthonormal waveforms from a set of finite energy signal waveforms \( \{s_i(t), i=1,2,\cdots,M\} \).
- Begin with the first waveform \( s_1(t) \) which has energy \( \varepsilon_1 \).

The first orthonormal waveform is:

\[
f_1(t) = \frac{s_1(t)}{\sqrt{\varepsilon_1}}
\]

- The 2nd waveform is constructed from \( s_2(t) \) by first computing the projection of \( f_1(t) \) onto \( s_2(t) \):

\[
c_{12} = \int_{-\infty}^{\infty} s_2(t) f_1(t) \, dt
\]

- Then \( c_{12} f_1(t) \) is subtracted from \( s_2(t) \):

\[
f_2'(t) = s_2(t) - c_{12} f_1(t)
\]
Gram-Schmidt procedure (cont.)

- If \( \varepsilon_2 \) denotes the energy of \( f'_2(t) \), the normalized waveform that is orthogonal to \( f_1(t) \) is:

\[
f_2(t) = \frac{f'_2(t)}{\sqrt{\varepsilon_2}}
\]

- In general, the orthogonalization of the \( k \)th function leads to

\[
f_k(t) = \frac{f'_k(t)}{\sqrt{\varepsilon_k}} \quad \text{where} \quad f'_k(t) = s_k(T) - \sum_{i=1}^{k-1} c_{ik} f_i(t)
\]

\[
c_{ik} = \int_{-\infty}^{\infty} s_k(t) f_i(t) \, dt, \quad i = 1, 2, \ldots, k-1
\]

- The orthogonalization process is continued until all the \( M \) signal waveforms have been exhausted and \( N \leq M \) orthonormal waveforms have been constructed.
Gram-Schmidt procedure (cont.)

Once we have constructed the set of orthonormal waveforms \( \{ f_n(t) \} \), we can express the \( M \) signals \( \{ s_n(t) \} \) as linear combinations of the \( \{ f_n(t) \} \):

\[
s_k(t) = \sum_{n=1}^{N} s_{kn} f_n(t), \quad k = 1, 2, \ldots, M
\]

\[
\varepsilon_k = \int_{-\infty}^{\infty} \left[ s_k(t) \right]^2 dt = \sum_{n=1}^{N} s_{kn}^2 = \| s_k \|^2
\]

\[
s_k = \begin{bmatrix} s_{k1} & s_{k2} & \ldots & s_{kN} \end{bmatrix}
\]

Each signal may be represented as a point in the \( N \)-dimensional signal space with coordinates \( \{ s_{ki}, i=1,2,\ldots,N \} \).
4.2.3 Orthogonal Expansions of Signals

- **Gram-Schmidt procedure (cont.)**
  - The energy in the $k$th signal is simply the square of the length of the vector or, equivalently, the square of the Euclidean distance from the origin to the point in the $N$-dimensional space.
  - Any signal can be represented geometrically as a point in the signal space spanned by the $\{f_n(t)\}$.
  - The functions $\{f_n(t)\}$ obtained from the Gram-Schmidt procedure are not unique.
  - If we alter the order in which the orthogonalization of the signals $\{s_n(t)\}$ is performed, the orthonormal waveforms will be different.
  - Nevertheless, the vectors $\{s_n(t)\}$ will retain their geometrical configuration and their lengths will be invariant to the choice of orthonormal functions $\{f_n(t)\}$. 
Consider the case in which the signal waveforms are band-pass and represented as:

\[ s_m(t) = \text{Re}\left[ s_{lm}(t) e^{j2\pi f_c t} \right], \quad m = 1, 2, \ldots, M \]

\[ \varepsilon_m = \int_{-\infty}^{\infty} s_m^2(t) \, dt = \frac{1}{2} \int_{-\infty}^{\infty} |s_{lm}(t)|^2 \, dt \quad \text{(from 4.1-24)} \]

Similarity between any pair of signal waveforms is measured by the normalized cross correlation:

\[ \frac{1}{\sqrt{\varepsilon_m \varepsilon_k}} \int_{-\infty}^{\infty} s_m(t) s_k(t) \, dt = \text{Re}\left\{ \frac{1}{2\sqrt{\varepsilon_m \varepsilon_k}} \int_{-\infty}^{\infty} s_{lm}(t) s_{lk}^*(t) \, dt \right\} \quad (4.2-44) \]

Complex-valued cross-correlation coefficient \( \rho_{km} \) is defined as:

\[ \rho_{km} = \frac{1}{2\sqrt{\varepsilon_m \varepsilon_k}} \int_{-\infty}^{\infty} s_{lm}^*(t) s_{lk}(t) \, dt \quad (4.2-45) \]
4.2.3 Orthogonal Expansions of Signals

The Euclidean distance between a pair of signals is defined as:

$$d_{km}^{(e)} = \| s_m - s_k \| = \left( \int_{-\infty}^{\infty} \left[ s_m(t) - s_k(t) \right]^2 dt \right)^{1/2}$$

$$= \left\{ \varepsilon_m + \varepsilon_k - 2\sqrt{\varepsilon_m \varepsilon_k} \operatorname{Re}(\rho_{km}) \right\}^{1/2}$$

When $\varepsilon_m = \varepsilon_k = \varepsilon$ for all $m$ and $k$, this expression simplifies to:

$$d_{km}^{(e)} = \left\{ 2\varepsilon \left[ 1 - \operatorname{Re}(\rho_{km}) \right] \right\}^{1/2}$$

From 4.2-44 & 4.2-45.
4.3 Representation of Digitally Modulated Signals

Digitally modulated signals, which are classified as linear, are conveniently expanded in terms of two orthonormal basis functions of the form:

\[ f_1(t) = \sqrt{\frac{2}{T}} \cos 2\pi f_c t \quad f_2(t) = -\sqrt{\frac{2}{T}} \sin 2\pi f_c t \]

If \( s_{lm}(t) \) is expressed as \( s_{lm}(t) = x_l(t) + jy_l(t) \), \( s_m(t) \) may be expressed as:

\[ s_m(t) = x_l(t) f_1(t) + y_l(t) f_2(t) \]  

From 4.2-42.

In the transmission of digital information over a communication channel, the modulator is the interface device that maps the digital information into analog waveforms that match the characteristics of the channel.
The mapping is generally performed by taking blocks of $k = \log_2 M$ binary digits at a time from the information sequence $\{a_n\}$ and selecting one of $M = 2^k$ deterministic, finite energy waveforms $\{s_m(t), m=1,2,\ldots,M\}$ for transmission over the channel.

When the mapping is performed under the constraint that a waveform transmitted in any time interval depends on one or more previously transmitted waveforms, the modulator is said to have memory. Otherwise, the modulator is called memoryless.

Functional model of passband data transmission system
The digital data transmits over a band-pass channel that can be \textit{linear} or \textit{nonlinear}.

This mode of data transmission relies on the use of a sinusoidal carrier wave modulated by the data stream.

In digital passband transmission, the incoming data stream is modulated onto a carrier (usually sinusoidal) with fixed frequency limits imposed by a band-pass channel of interest.

The modulation process making the transmission possible involves switching (keying) the amplitude, frequency, or phase of a sinusoidal carrier in some fashion in accordance with the incoming data.

There are three basic signaling schemes: amplitude-shift keying (ASK), frequency-shift Keying (FSK), and phase-shift keying (PSK).
4.3 Representation of Digitally Modulated Signals

- Illustrative waveforms for the three basic forms of signaling binary information. (a) ASK (b) PSK (c) FSK.
Unlike ASK signals, both PSK and FSK signals have a constant envelope. This property makes PSK and FSK signals impervious to amplitude nonlinearities.

In practice, we find that PSK and FSK signals are preferred to ASK signals for passband data transmission over nonlinear channels.

Digital modulation techniques may be classified into coherent and noncoherent techniques, depending on whether the receiver is equipped with a phase-recovery circuit or not.

The phase-recovery circuit ensures that the oscillator supplying the locally generated carrier wave in the receiver is synchronized (in both frequency and phase) to the oscillator supplying the carrier wave used to originally modulated the incoming data stream in the transmitter.
4.3 Representation of Digitally Modulated Signals

- **M-ary signaling scheme:**
  - For almost all applications, the number of possible signals $M=2^n$.
  - Symbol duration $T=nT_b$, where $T_b$ is the bit duration.
  - We have $M$-ary ASK, $M$-ary PSK, and $M$-ary FSK.
  - We can also combine different methods of modulation into a hybrid form. For example, $M$-ary amplitude-phase keying (APK) and $M$-ary quadrature-amplitude modulation (QAM).
  - $M$-ary PSK and $M$-ary QAM are examples of linear modulation.
  - An $M$-ary PSK signal has a constant envelope, whereas an $M$-ary QAM signal involves changes in the carrier amplitude.
4.3 Representation of Digitally Modulated Signals

M-ary signaling scheme (cont.):

- $M$-ary PSK can be used to transmit digital data over a **nonlinear band-pass channel**, whereas $M$-ary QAM requires the use of a **linear channel**.
- $M$-ary PSK, and $M$-ary QAM are commonly used in **coherent systems**.
- ASK and FSK lend themselves naturally to use in **non-coherent systems** whenever it is impractical to maintain carrier phase synchronization.
- We can’t have noncoherent PSK.
4.3.1 Memoryless Modulation Methods

Pulse-amplitude-modulated (PAM) signals

*Double-sideband (DSB)* signal waveform may be represented as:

\[ s_m(t) = \text{Re} \left[ A_m g(t) e^{j2\pi f_c t} \right] \]

\[ = A_m g(t) \cos 2\pi f_c t, \quad m = 1, 2, ..., M, \quad 0 \leq t \leq T \]

where \( A_m \) denote the set of \( M \) possible amplitudes corresponding to \( M=2^k \) possible \( k \)-bit blocks of symbols.

- The signal amplitudes \( A_m \) take the discrete values:
  \[ A_m = (2m - 1 - M)d, \quad m = 1, 2, ..., M, \quad -(M-1)d ... (M-1)d \]

- \( 2d \) is the distance between adjacent signal amplitudes.

- \( g(t) \) is a real-valued signal pulse whose shape influences the spectrum of the transmitted signal.

- The *symbol rate* is \( R/k, T_b = 1/R \) is the *bit interval*, and \( T=k/R=kT_b \) is the *symbol interval*. 
4.3.1 Memoryless Modulation Methods

- Pulse-amplitude-modulated (PAM) signals (cont.)
  - The $M$ PAM signals have energies:
    \[ \varepsilon_m = \int_0^T s_m^2(t) \, dt = \frac{1}{2} A_m^2 \int_0^T g^2(t) \, dt = \frac{1}{2} A_m^2 \varepsilon_g \]
  - These signals are one-dimensional and are represented by:
    \[ s_m(t) = s_m f(t) \]
  - $f(t)$ is defined as the unit-energy signal waveform given as:
    \[ f(t) = \sqrt{\frac{2}{\varepsilon_g}} g(t) \cos 2\pi f_c t \]
  - $s_m = A_m \sqrt{\frac{1}{2} \varepsilon_g}$, $m = 1, 2, \ldots, M$
  - Digital PAM is also called amplitude-shift keying (ASK).
4.3.1 Memoryless Modulation Methods

- Pulse-amplitude-modulated (PAM) signals (cont.)

  - Signal space diagram for digital PAM signals:

    (a) \( M = 2 \)
    
    - 0
    - 1
    
    (b) \( M = 4 \)
    
    - 00
    - 01
    - 11
    - 10
    
    (c) \( M = 8 \)
    
    - 000
    - 001
    - 011
    - 010
    - 110
    - 111
    - 101
    - 100
4.3.1 Memoryless Modulation Methods

Pulse-amplitude-modulated (PAM) signals (cont.)

- **Gray encoding**: The mapping of \( k \) information bits to the \( M=2^k \) possible signal amplitudes may be done in a number of ways. The preferred assignment is one in which the adjacent signal amplitudes differ by one binary digit.

- The **Euclidean distance** between any pair of signal points is:
  \[
  d_{mn}^{(e)} = \sqrt{(s_m - s_n)^2} = \sqrt{\frac{1}{2} \varepsilon_g |A_m - A_n|} = d \sqrt{2\varepsilon_g} |m - n|
  \]

- The **minimum Euclidean distance** between any pair of signals is:
  \[
  d_{\text{min}}^{(e)} = d \sqrt{2\varepsilon_g}
  \]
4.3.1 Memoryless Modulation Methods

- Pulse-amplitude-modulated (PAM) signals (cont.)
  - **Single Sideband (SSB)** PAM is represented by:
    \[
    s_m(t) = \Re \left\{ A_m \left[ g(t) \pm j \hat{g}(t) \right] e^{j2\pi f_c t} \right\}, \quad m = 1, 2, \ldots, M
    \]
  
  where \( \hat{g}(t) \) is the *Hilbert transform* of \( g(t) \).

- The digital PAM signal is also appropriate for transmission over a channel that does not require carrier modulation and is called *baseband signal*:
  \[
  s_m(t) = A_m g(t), \quad m = 1, 2, \ldots, M
  \]

- If \( M=2 \), the signals are called *antipodal* and have the special property that:
  \[
  s_1(t) = -s_2(t)
  \]
4.3.1 Memoryless Modulation Methods

- Pulse-amplitude-modulated (PAM) signals (cont.)
  - Four-amplitude level baseband and band-pass PAM signals
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals (Binary Phase-Shift Keying)
  - In a coherent binary PSK system, the pair of signals $s_1(t)$ and $s_2(t)$ used to represent binary symbols 1 and 0, respectively is defined by
    
    $$s_1(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t)$$
    
    $$s_2(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t + \pi) = -\sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t)$$

  where $0 \leq t \leq T_b$ and $E_b$ is the transmitted signal energy per bit.

  - A pair of sinusoidal waves that differ only in a relative phase-shift of 180 degrees are referred to as antipodal signals.
Phase-modulated signals (Binary Phase-Shift Keying)

To ensure that each transmitted bit contains an integral number of cycles of the carrier wave, the carrier frequency $f_c$ is chosen equal to $n_c/T_b$ for some fixed integer $n_c$.

In the case of binary PSK, there is only one basis function of unit energy:

$$\phi_1(t) = \sqrt{2} \cos(2\pi f_c t), \quad 0 \leq t < T_b$$

$$s_1(t) = \sqrt{E_b} \phi_1(t), \quad s_2(t) = -\sqrt{E_b} \phi_1(t), \quad 0 \leq t < T_b$$

The coordinates of the message points are:

$$s_{11} = \int_0^{T_b} s_1(t)\phi_1(t) dt = +\sqrt{E_b} \quad s_{21} = \int_0^{T_b} s_2(t)\phi_1(t) dt = -\sqrt{E_b}$$
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals (Binary Phase-Shift Keying)
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals (Quadriphase-Shift Keying)
  - Quadriphase-Shift Keying (QPSK)

\[
s_i(t) = \begin{cases} 
\sqrt{\frac{2E}{T}} \cos \left[ 2\pi f_c t + (2i - 1) \frac{\pi}{4} \right], & 0 \leq t \leq T \\
0, & \text{elsewhere}
\end{cases}
\]

where \( i = 1, 2, 3, 4; \) \( E \) is the transmitted signal energy per symbol, and \( T \) is the symbol duration.

- Equivalently

\[
s_i(t) = \sqrt{\frac{2E}{T}} \cos \left( 2i - 1 \frac{\pi}{4} \right) \cos (2\pi f_c t) - \sqrt{\frac{2E}{T}} \sin \left( 2i - 1 \frac{\pi}{4} \right) \sin (2\pi f_c t)
\]
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals (Quadriphase-Shift Keying)
  - Defined a pair of quadrature carriers:
    \[ \phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T \]
    \[ \phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad 0 \leq t \leq T \]
  - There are four message points, and the associated signal vectors are defined by
    \[
    s_i(t) = \begin{bmatrix}
    \sqrt{E} \cos\left((2i - 1)\frac{\pi}{4}\right) \\
    \sqrt{E} \sin\left((2i - 1)\frac{\pi}{4}\right)
    \end{bmatrix}, \quad i = 1,2,3,4
    \]
Phase-modulated signals (Quadriphase-Shift Keying)

- Each possible value of the phase corresponds to a unique dibit.
- For example, we may choose the Gray coding.

<table>
<thead>
<tr>
<th>Gray-encoded Input Dibit</th>
<th>Phase of QPSK Signal (radians)</th>
<th>Coordinates of Message Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>$\pi/4$</td>
<td>$+\sqrt{E/2}$ $-\sqrt{E/2}$</td>
</tr>
<tr>
<td>00</td>
<td>$3\pi/4$</td>
<td>$-\sqrt{E/2}$ $-\sqrt{E/2}$</td>
</tr>
<tr>
<td>01</td>
<td>$5\pi/4$</td>
<td>$-\sqrt{E/2}$ $+\sqrt{E/2}$</td>
</tr>
<tr>
<td>11</td>
<td>$7\pi/4$</td>
<td>$+\sqrt{E/2}$ $+\sqrt{E/2}$</td>
</tr>
</tbody>
</table>
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals (Quadriphase-Shift Keying)
- Signal space diagram of coherent QPSK system
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals (Quadriphase-Shift Keying)

Input binary sequence

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Dibit</td>
<td>01</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

(a)

Odd-numbered sequence 0
Polarity of coefficient $s_1$

- $s_1 \phi_1(t)$

(b)

Even-numbered sequence 1
Polarity of coefficient $s_2$

- $s_2 \phi_2(t)$

(c)

$s(t)$

(d)
Phase-modulated signals (\(M\)-ary PSK)

The \(M\) signal waveforms are represented as:

\[
s_m(t) = \text{Re} \left[ g(t) e^{j2\pi(m-1)/M} e^{j2\pi f_c t} \right], \quad m = 1, 2, \ldots, M, \quad 0 \leq t \leq T
\]

\[
= g(t) \cos \left[ 2\pi f_c t + \frac{2\pi}{M} (m-1) \right]
\]

\[
= g(t) \cos \frac{2\pi}{M} (m-1) \cos 2\pi f_c t - g(t) \sin \frac{2\pi}{M} (m-1) \sin 2\pi f_c t
\]

Digital phase modulation is usually called phase-shift keying (PSK).
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals ($M$-ary PSK)
  - Signal space diagram for octaphase shift keying (i.e., $M=8$)
4.3.1 Memoryless Modulation Methods

◼ Phase-modulated signals (\(M\)-ary PSK)

◼ The signal waveforms have equal energy:

\[
\varepsilon = \int_0^T s_m^2(t) \, dt = \frac{1}{2} \int_0^T g^2(t) \, dt = \frac{1}{2} \varepsilon_g
\]

◼ The signal waveforms may be represented as a linear combination of two orthonormal signal waveforms:

\[
s_m(t) = s_{m1} f_1(t) + s_{m2} f_2(t)
\]

\[
f_1(t) = \sqrt{\frac{2}{\varepsilon_g}} g(t) \cos 2\pi f_c t \quad \text{and} \quad f_2(t) = -\sqrt{\frac{2}{\varepsilon_g}} g(t) \sin 2\pi f_c t
\]

◼ The two-dimensional vectors \(s_m = [s_{m1} \ s_{m2}]\) are given by:

\[
s_m = \begin{bmatrix} \sqrt{\frac{2}{\varepsilon_g}} \cos \frac{2\pi}{M} (m - 1) \\ \sqrt{\frac{2}{\varepsilon_g}} \sin \frac{2\pi}{M} (m - 1) \end{bmatrix}, \quad m = 1, 2, \ldots, M
\]
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals ($M$-ary PSK)
  - The Euclidean distance between signal points is:
    \[ d_{mn}^{(e)} = \| s_m - s_n \| \]
    \[ = \sqrt{(s_{m1} - s_{n1})^2 + (s_{m2} - s_{n2})^2} = \left\lfloor \varepsilon_g \left[ 1 - \cos \frac{2\pi}{M} (m - n) \right] \right\rfloor^{1/2} \]
  - The minimum Euclidean distance corresponds to the case in which $|m-n|=1$, i.e., adjacent signal phases.
    \[ d_{\text{min}}^{(e)} = \sqrt{\varepsilon_g \left( 1 - \cos \frac{2\pi}{M} \right)} \]
  - Average probability of symbol error for coherent $M$-ary PSK:
    \[ P_e \approx \text{erfc} \left( \sqrt{\frac{E}{N_0}} \sin \left( \frac{\pi}{M} \right) \right) \]
Phase-modulated signals (Offset QPSK)

- The carrier phase changes by ±180 degrees whenever both the in-phase and quadrature components of the QPSK signal changes sign.
- This can result in problems for power amplifiers.
- The problem may be reduced by using offset QPSK.
- In offset QPSK, the bit stream responsible for generating the quadrature component is delayed (i.e. offset) by half a symbol interval with respect to the bit stream responsible for generating the in-phase component.

\[ \phi_i(t) = \sqrt{\frac{2}{T}} \cos\left(2\pi f_c t\right), \quad 0 \leq t \leq T \]
Phase-modulated signals (Offset QPSK)

- The two basis functions of offset QPSK are defined by
  \[
  \phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T
  \]
  \[
  \phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad \frac{T}{2} \leq t \leq \frac{3T}{2}
  \]

- The phase transitions likely to occur in offset QPSK are confined to ±90 degrees.
- However, ±90 degrees phase transitions in offset QPSK occur twice as frequently.
Phase-modulated signals (Offset QPSK)

- Amplitude fluctuations in offset QPSK due to filtering have a smaller amplitude than in the case of QPSK.
- The offset QPSK has exactly the same probability of symbol error in an AWGN channel as QPSK.
- The reason for the equivalence is that the statistical independence of the in-phase and quadrature components applies to both QPSK and offset QPSK.
4.3.1 Memoryless Modulation Methods

Phase-modulated signals (Offset QPSK)
4.3.1 Memoryless Modulation Methods

- Phase-modulated signals (π/4-Shifted QPSK)
  - The carrier phase used for the transmission of successive symbols is alternately picked from one of the two QPSK constellations in the following figure and then the other.
4.3.1 Memoryless Modulation Methods

Phase-modulated signals (\(\pi/4\)-Shifted QPSK)

- It follows that a \(\pi/4\)-shifted QPSK signal may reside in any one of eight possible phase states:

<table>
<thead>
<tr>
<th>Gray-Encoded Input Dibit</th>
<th>Phase Change, (\Delta \theta) (radians)</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>(\pi/4)</td>
</tr>
<tr>
<td>01</td>
<td>(3\pi/4)</td>
</tr>
<tr>
<td>11</td>
<td>(-3\pi/4)</td>
</tr>
<tr>
<td>10</td>
<td>(-\pi/4)</td>
</tr>
</tbody>
</table>
Phase-modulated signals (π/4-Shifted QPSK)

- Attractive features of the π/4-shifted QPSK scheme
  - The phase transitions from one symbol to the next are restricted to ±π/4 and ±3π/4.
  - Envelope variations due to filtering are significantly reduced.
  - π/4-shifted QPSK signals can be noncoherently detected, thereby considerably simplifying the receiver design.
  - Like QPSK signals, π/4-shifted QPSK can be differently encoded, in which case we should really speak of π/4-shifted DQPSK.
4.3.1 Memoryless Modulation Methods

- **Dual-Carrier Modulation (DCM)**
  - Adopted in Multi-band OFDM (Ultra Wideband)
  - The coded and interleaved binary serial input data, \( b[i] \) where \( i = 0, 1, 2, \cdots \), shall be divided into groups of 200 bits and converted into 100 complex numbers using a technique called dual-carrier modulation.
  - The conversion shall be performed as follows:
    1. The 200 coded bits are grouped into 50 groups of 4 bits. Each group is represented as \( (b[g(k)], b[g(k)+1], b[g(k)+50], b[g(k)+51]) \), where \( k \in [0, 49] \) and
      \[
      g(k) = \begin{cases} 
      2k & k \in [0, 24] \\
      2k + 50 & k \in [25, 49] 
      \end{cases}
      \]
4.3.1 Memoryless Modulation Methods

- **Dual-Carrier Modulation (DCM)**

2. Each group of 4 bits \(b[g(k)], b[g(k)+1], b[g(k) + 50], b[g(k) + 51]\) shall be mapped onto a four-dimensional constellation, and converted into two complex numbers \((d[k], d[k + 50])\).

3. The complex numbers shall be normalized using a normalization factor \(K_{MOD}\).

- The normalization factor \(K_{MOD} = 10^{-1/2}\) is used for the dual-carrier modulation.

- An approximate value of the normalization factor may be used, as long as the device conforms to the modulation accuracy requirements.
4.3.1 Memoryless Modulation Methods

- Dual-Carrier Modulation (DCM)
## 4.3.1 Memoryless Modulation Methods

### Dual-Carrier Modulation (DCM) Encoding Table

<table>
<thead>
<tr>
<th>Input Bit ((b[g(k)], b[g(k)+1], b[g(k) + 50], b[g(k) + 51]))</th>
<th>(d[k]) (l)-out</th>
<th>(d[k]) (Q)-out</th>
<th>(d[k + 50]) (l)-out</th>
<th>(d[k + 50]) (Q)-out</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>-3</td>
<td>-3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0001</td>
<td>-3</td>
<td>1</td>
<td>1</td>
<td>-3</td>
</tr>
<tr>
<td>0010</td>
<td>-3</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0011</td>
<td>-3</td>
<td>3</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>0100</td>
<td>-1</td>
<td>-3</td>
<td>-3</td>
<td>1</td>
</tr>
<tr>
<td>0101</td>
<td>-1</td>
<td>-1</td>
<td>-3</td>
<td>-3</td>
</tr>
<tr>
<td>0110</td>
<td>-1</td>
<td>1</td>
<td>-3</td>
<td>3</td>
</tr>
<tr>
<td>0111</td>
<td>-1</td>
<td>3</td>
<td>-3</td>
<td>-1</td>
</tr>
<tr>
<td>1000</td>
<td>1</td>
<td>-3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>1001</td>
<td>1</td>
<td>-1</td>
<td>3</td>
<td>-3</td>
</tr>
<tr>
<td>1010</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1011</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>-1</td>
</tr>
<tr>
<td>1100</td>
<td>3</td>
<td>-3</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>1101</td>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>-3</td>
</tr>
<tr>
<td>1110</td>
<td>3</td>
<td>1</td>
<td>-1</td>
<td>3</td>
</tr>
<tr>
<td>1111</td>
<td>3</td>
<td>3</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>
4.3.1 Memoryless Modulation Methods

- **Quadrature amplitude modulation (QAM)**
  - **Quadrature PAM** or **QAM**: The bandwidth efficiency of PAM/SSB can also be obtained by simultaneously impressing two separate \( k \)-bit symbols from the information sequence \( \{a_n\} \) on two quadrature carriers \( \cos 2\pi f_c t \) and \( \sin 2\pi f_c t \).
  - The signal waveforms may be expressed as:

\[
s_m(t) = \text{Re} \left[ (A_{mc} + jA_{ms}) g(t) e^{j2\pi f_c t} \right], \quad m = 1, 2, ..., M, \quad 0 \leq t \leq T
\]

\[
= A_{mc} g(t) \cos 2\pi f_c t - A_{ms} g(t) \sin 2\pi f_c t
\]

\[
s_m(t) = \text{Re} \left[ V_m e^{j\theta_m} g(t) e^{j2\pi f_c t} \right] = V_m g(t) \cos (2\pi f_c t + \theta_m)
\]

\[
V_m = \sqrt{A_{mc}^2 + A_{ms}^2} \quad \text{and} \quad \theta_m = \tan^{-1} \left( \frac{A_{ms}}{A_{mc}} \right)
\]
4.3.1 Memoryless Modulation Methods

- Quadrature amplitude modulation (QAM) (cont.)
  - We may select a combination of $M_1$-level PAM and $M_2$-phase PSK to construct an $M = M_1 M_2$ combined PAM-PSK signal constellation.
4.3.1 Memoryless Modulation Methods

Quadrature amplitude modulation (QAM) (cont.)

As in the case of PSK signals, the QAM signal waveforms may be represented as a linear combination of two orthonormal signal waveforms \(f_1(t)\) and \(f_2(t)\):

\[
s_m(t) = s_{m1}f_1(t) + s_{m2}f_2(t)
\]

\[
f_1(t) = \sqrt{\frac{2}{\varepsilon_g}} g(t) \cos 2\pi f_c t \quad \text{and} \quad f_2(t) = -\sqrt{\frac{2}{\varepsilon_g}} g(t) \sin 2\pi f_c t
\]

\[
s_m = [s_{m1} \quad s_{m2}] = \begin{bmatrix} A_{mc} \sqrt{\frac{1}{2}} \varepsilon_g & A_{ms} \sqrt{\frac{1}{2}} \varepsilon_g \end{bmatrix}
\]

The Euclidean distance between any pair of signal vectors is:

\[
d_{mn}^{(e)} = \|s_m - s_n\| = \sqrt{\frac{1}{2} \varepsilon_g \left[ (A_{mc} - A_{nc})^2 + (A_{ms} - A_{ns})^2 \right]}
\]
4.3.1 Memoryless Modulation Methods

- Quadrature amplitude modulation (QAM) (cont.)
  - Several signal space diagrams for *rectangular QAM*:

\[
d^{(e)}_{\text{min}} = d \sqrt{2\varepsilon_g}
\]
4.3.1 Memoryless Modulation Methods

- Multidimensional signals:
  - We may use either the time domain or the frequency domain or both in order to increase the number of dimensions.
  - Subdivision of time and frequency axes into distinct slots:
4.3.1 Memoryless Modulation Methods

Orthogonal multidimensional signals

Consider the construction of \( M \) equal-energy orthogonal signal waveforms that differ in frequency:

\[
s_m(t) = \Re\left[ s_{lm}(t) e^{j2\pi f_c t} \right], \quad m = 1, 2, \ldots, M, \quad 0 \leq t \leq T
\]

\[
= \sqrt{\frac{2\varepsilon}{T}} \cos\left[ 2\pi f_c t + 2\pi m\Delta f t \right]
\]

\[
s_{lm}(t) = \sqrt{\frac{2\varepsilon}{T}} e^{j2\pi m\Delta f t}, \quad m = 1, 2, \ldots, M, \quad 0 \leq t \leq T
\]

This type of frequency modulation is called frequency-shift keying (FSK).
Orthogonal multidimensional signals (cont.)

These waveforms have equal cross-correlation coefficients:

\[
\rho_{km} = \frac{2\varepsilon}{T} \int_0^T e^{j2\pi(m-k)\Delta f} dt = \frac{\sin \pi T (m-k) \Delta f}{\pi T (m-k) \Delta f} e^{j\pi T (m-k) \Delta f}
\]

\[
\rho_r \equiv \text{Re}(\rho_{km}) = \frac{\sin[\pi T (m-k) \Delta f]}{\pi T (m-k) \Delta f} \cos[\pi T (m-k) \Delta f]
\]

\[
= \frac{\sin[2\pi T (m-k) \Delta f]}{2\pi T (m-k) \Delta f}
\]

Note that \(\text{Re}(\rho_{km})=0\) when \(\Delta f=1/2T\) and \(m \neq k\).

\(|\rho_{km}|=0\) for multiple of \(1/T\).
4.3.1 Memoryless Modulation Methods

- Orthogonal multidimensional signals (cont.)
  - Cross-correlation coefficient as a function of frequency separation for FSK signals:
4.3.1 Memoryless Modulation Methods

Orthogonal multidimensional signals (cont.)

For $\Delta f = 1/2T$, the $M$-FSK signals are equivalent to the $N$-dimensional vectors:

$$
\begin{align*}
    s_1 &= \begin{bmatrix} \sqrt{\varepsilon} & 0 & 0 & \ldots & 0 & 0 \end{bmatrix} \\
    s_2 &= \begin{bmatrix} 0 & \sqrt{\varepsilon} & 0 & \ldots & 0 & 0 \end{bmatrix} \\
    \vdots & \vdots \\
    s_N &= \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & \sqrt{\varepsilon} \end{bmatrix}
\end{align*}
$$

The distance between pairs of signals is:

$$
    d_{km}^{(e)} = \sqrt{2\varepsilon}, \quad \text{for all } m, k
$$
4.3.1 Memoryless Modulation Methods

- Orthogonal multidimensional signals (cont.)
  - Orthogonal signals for $M=N=3$ and $M=N=2$. 

![Diagram showing orthogonal signals for $M=N=3$ and $M=N=2$.](image)
Binary Frequency-Shift Keying

In a binary FSK system, symbols 1 and 0 are distinguished from each other by transmitting one of two sinusoidal waves that differ in frequency by a fixed amount.

\[ s_i(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_i t), & 0 \leq t \leq T_b \\ 0, & \text{elsewhere} \end{cases} \]

The transmitted frequency is

\[ f_i = \frac{n_c + i}{T_b} \quad \text{for some fixed integer } n_c \text{ and } i = 1, 2 \]
**4.3.1 Memoryless Modulation Methods**

- **Binary Frequency-Shift Keying**
  - The FSK signal is a continuous phase signal in the sense that phase continuity is always maintained, including the inter-bit switching times.
  - This form of digital modulation is an example of *continuous-phase frequency-shift keying* (CPFSK).
  - The signal $s_1(t)$ and $s_2(t)$ are orthogonal

$$
\phi_i(t) = \begin{cases} 
\sqrt{\frac{2}{T_b}} \cos(2\pi f_i t), & 0 \leq t \leq T_b \\
0, & \text{elsewhere}
\end{cases}
$$
4.3.1 Memoryless Modulation Methods

Binary Frequency-Shift Keying

\[ s_{ij} = \int_{0}^{T_b} s_i(t) \phi_j(t) \, dt \]

\[ = \int_{0}^{T_b} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_i t) \sqrt{\frac{2}{T_b}} \cos(2\pi f_i t) \, dt \]

\[ = \begin{cases} \sqrt{E_b}, & i = j \\ 0, & i \neq j \end{cases} \]

\[ s_1 = \begin{bmatrix} \sqrt{E_b} \\ 0 \end{bmatrix} \quad s_2 = \begin{bmatrix} 0 \\ \sqrt{E_b} \end{bmatrix} \]
4.3.1 Memoryless Modulation Methods

Biorthogonal signals

- A set of $M$ biorthogonal signals can be constructed from $M/2$ orthogonal signals by simply including the negatives of the orthogonal signals.
- The correlation between any pair of waveforms is either $\rho_r = -1$ or 0.

Signal space diagrams for $M=4$ and $M=6$ biorthogonal signals.
Simplex signals

For a set of $M$ orthogonal waveforms $\{s_m(t)\}$ or their vector representation $\{s_m\}$ with mean of:

$$s = \frac{1}{M} \sum_{m=1}^{M} s_m$$

*Simplex signals* are obtained by translating the origin of the $m$ orthogonal signals to the point $\bar{s}$.

$$s'_m = s_m - \bar{s}, \quad m = 1, 2, \ldots, M$$

The energy per waveform is:

$$\|s'_m\|^2 = \|s_m - \bar{s}\|^2 = \varepsilon - \frac{2}{M} \varepsilon + \frac{1}{M} \varepsilon = \varepsilon \left(1 - \frac{1}{M}\right)$$
4.3.1 Memoryless Modulation Methods

**Simplex signals (cont.)**

- The cross correlation of any pair of signals is (4.2-27):

\[
\text{Re}(\rho_{mn}) = \frac{s'_m \cdot s'_n}{\|s'_m\| \cdot \|s'_n\|} = \frac{-1/M}{1 - 1/M}
\]

\[
= -\frac{1}{M - 1}
\]

- The set of simplex waveforms is *equally correlated* and requires less energy, by the factor \(1 - 1/M\), than the set of orthogonal waveforms.

Signal space diagrams for \(M\)-ary simplex signals.
Signal waveforms from binary codes

A set of $M$ signaling waveforms can be generated from a set of $M$ binary code words of the form:

$$C_m = [c_{m1}, c_{m2}, \ldots, c_{mN}], \quad m = 1, 2, \ldots, M, \quad c_{mj} = 0 \text{ or } 1.$$

Each component of a code word is mapped into an elementary binary PSK waveform:

$$c_{mj} = 1 \Rightarrow s_{mj}(t) = \sqrt{\frac{2\varepsilon_c}{T_c}} \cos 2\pi f_c t, \quad 0 \leq t \leq T_c$$

$$c_{mj} = 0 \Rightarrow s_{mj}(t) = -\sqrt{\frac{2\varepsilon_c}{T_c}} \cos 2\pi f_c t, \quad 0 \leq t \leq T_c$$

where $T_c = T/N$ and $\varepsilon_c = \varepsilon /N$. 

4.3.1 Memoryless Modulation Methods
4.3.1 Memoryless Modulation Methods

Signal waveforms from binary codes (cont.)

- The $M$ code words $\{C_m\}$ are mapped into a set of $M$ waveforms $\{s_m(t)\}$.
- The waveforms can be represented in vector form as:
  \[
  s_m = \begin{bmatrix} s_{m1} & s_{m2} & \cdots & s_{mN} \end{bmatrix}, \quad m = 1, 2, \ldots, M, \quad \text{where} \quad s_{mj} = \pm \sqrt{\varepsilon / N}.
  \]
- $N$ is called the block length of the code and is also the dimension of the $M$ waveforms.

Signal space diagrams for signals generated from binary codes.
4.3.1 Memoryless Modulation Methods

Signal waveforms from binary codes (cont.)

- Each of the $M$ waveforms has energy $\varepsilon$.
- Any adjacent signal points have a cross-correlation coefficient:

$$\rho_r = \frac{\varepsilon (1 - 2/N)}{\varepsilon} = \frac{N-2}{N}$$

- The corresponding distance is:

$$d^{(e)} = \sqrt{2\varepsilon (1 - \rho_r)} = \sqrt{4\varepsilon / N}$$
4.3.2 Linear Modulation with Memory

- There are some modulation signals with dependence between the signals transmitted in successive symbol intervals.
- This signal dependence is usually introduced for the purpose of shaping the spectrum of the transmitted signal so that it matches the spectral characteristics of the channel.

Examples of baseband signals and the corresponding data sequence:
4.3.2 Linear Modulation with Memory

- **NRZ**: the binary information digit 1 is represented by a rectangular pulse of polarity $A$ and the binary digit 0 is represented by a rectangular pulse of polarity $-A$.
  - The NRZ modulation is memoryless and is equivalent to a binary PAM or a binary PSK signal in a carrier-modulated system.

- **NRZI**: the signal is different from the NRZ signal in that transitions from one amplitude level to another occur only when a 1 is transmitted.
  - This type of signal encoding is called *differential encoding*.
  - The encoding operation is described by the relation:
    \[ b_k = a_k \oplus b_{k-1} \]
    where $\{a_k\}$ is the binary information sequence into the encoder, $\{b_k\}$ is the output sequence of the encoder.
NRZI (cont.)

- When \( b_k = 1 \), the transmitted waveform is a rectangular pulse of amplitude \( A \), and when \( b_k = 0 \), the transmitted waveform is a rectangular pulse of amplitude \(-A\).
- The differential encoding operation introduces memory in the signal.
- The combination of the encoder and the modulator operations may be represented by a state diagram (Markov chain):

Input bit

State \( S_0 \) \( S_0 = 0 \) \( 1/s(t) \) \( 0/s(t) \)

State \( S_1 \) \( S_1 = 1 \) \( 1/s(t) \) \( 0/s(t) \)

0/s(t)

1/s(t)
NRZI (cont.)

The state diagram may be described by two transition matrices corresponding to the two possible input bits \{0,1\}.

When \(a_k=0\), the encoder stays in the same state and the state transition matrix for a zero is:

\[
T_1 = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

where \(t_{ij}=1\) if \(a_k\) results in a transition from state \(i\) to state \(j\), \(i=1,2\), and \(j=1,2\); otherwise \(t_{ij}=0\).

The state transition matrix for \(a_k=1\) is:

\[
T_2 = \begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix}
\]
NRZI (cont.)

The trellis diagram for the NRZI signal:

Delay modulation is equivalent to encoding the data sequence by a run-length-limited code called a Miller code and using NRZI to transmit the encoded data (will be shown in Chapter 9).
Delay modulation (cont.):

- The signal of delay modulation may be described by a state diagram that has four states:

  ![State Diagram](image)

  - There are two elementary waveforms $s_1(t)$ and $s_2(t)$ and their negatives $-s_1(t)$ and $-s_2(t)$, which are used for transmitting the binary information.

---

**4.3.2 Linear Modulation with Memory**

---

---

---
4.3.2 Linear Modulation with Memory

**Delay modulation**

- The state transition matrices can be obtained from the figure in previous page:

  - When $a_k = 0$: 
    $$ T_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} $$
    $$ s_3 \rightarrow s_1 \\ s_4 \rightarrow s_1 \\ s_1 \rightarrow s_4 \\ s_2 \rightarrow s_4 $$

  - When $a_k = 1$: 
    $$ T_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} $$
    $$ s_1 \rightarrow s_2 \\ s_3 \rightarrow s_2 \\ s_2 \rightarrow s_3 \\ s_4 \rightarrow s_3 $$
4.3.2 Linear Modulation with Memory

- Modulation techniques with memory such as NRZI and Miller coding are generally characterized by a $K$-state Markov chain with \textit{stationary state probabilities} \{\(p_i, i = 1,2,\cdots,K\}\) and \textit{transition probabilities} \{\(p_{ij}, i,j = 1,2,\cdots,K\}\). Associated with each transition is a signal waveform \(s_j(t), j = 1,2,\cdots,K\).

- \textit{Transition probability matrix} \((P)\) can be arranged in matrix form as:

\[
P = \begin{bmatrix}
p_{11} & p_{12} & \cdots & p_{1K} \\
p_{21} & p_{22} & \cdots & p_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
p_{K1} & p_{K2} & \cdots & p_{KK}
\end{bmatrix}
\]
4.3.2 Linear Modulation with Memory

- The transition probability matrix can be obtained from the transition matrices \( \{ T_i \} \) and the corresponding probabilities of occurrence of the input bits:

\[
P = \sum_{i=1}^{2} q_i T_i
\]

where \( q_1 = P(a_k=0) \) and \( q_2 = P(a_k=1) \).

- For the NRZI signal with equal state probabilities \( p_1 = p_2 = 0.5 \),

\[
P = \begin{bmatrix}
1 & 1 \\
\frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix} = \frac{1}{2} \cdot \begin{bmatrix}1 & 1 \\
0 & 1
\end{bmatrix} + \frac{1}{2} \cdot \begin{bmatrix}0 & 1 \\
1 & 0
\end{bmatrix}
\]
The transition probability matrix for the Miller-coded signal with equally likely symbols ($q_1=q_2=0.5$ or, equivalently, $p_1=p_2=p_3=p_4=0.25$) is:

\[
P = \begin{bmatrix}
0 & \frac{1}{2} & 0 & \frac{1}{2} \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\end{bmatrix}
\]

The transition probability matrix is useful in the determination of the spectral characteristics of digital modulation techniques with memory.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

Introduction

- In this section, we consider a class of digital modulation methods in which the phase of the signal is constrained to be continuous.
- This constraint results in a phase or frequency modulator that has memory.
- The modulation method is also non-linear.

Continuous-phase FSK (CPFSK)

- A conventional FSK signal is generated by shifting the carrier by an amount \( f_n = \frac{1}{2} \Delta f I_n \), \( I_n = \pm 1, \pm 3, \ldots, \pm (M - 1) \), to reflect the digital information that is being transmitted.
- This type (conventional type) of FSK signal is memoryless.
Continuous-phase FSK (CPFSK) (cont.)

The switching from one frequency to another may be accomplished by having $M=2^k$ separate oscillators tuned to the desired frequencies and selecting one of the $M$ frequencies according to the particular $k$-bit symbol that is to be transmitted in a signal interval of duration $T=k/R$ seconds.

The reasons why we have CPFSK: (or the defects of conventional FSK)

- Such abrupt switching from one oscillator output to another in successive signaling intervals results in relatively large spectral side lobes outside of the main spectral band of the signal.
- Consequently, this method requires a large frequency band for transmission of the signal.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

Continuous-phase FSK (CPFSK) (cont.)

Solution:

- To avoid the use of signals having large spectral side lobes, the information-bearing signal frequency modulates a single carrier whose frequency is changed continuously.
- The resulting frequency-modulated signal is phase-continuous and, hence, it is called continuous-phase FSK (CPFSK).
- This type (continuous-phase type) of FSK signal has memory because the phase of the carrier is constrained to be continuous.
Continuous-phase FSK (CPFSK) (cont.)

- In order to represent a CPFSK signal, we begin with a PAM signal:
  \[ d(t) = \sum_n I_n g(t - nT) \]

- \( d(t) \) is used to frequency-modulate the carrier.

- \( \{I_n\} \) denotes the sequence of amplitudes obtained by mapping \( k \)-bit blocks of binary digits from the information sequence \( \{a_n\} \) into the amplitude levels \( \pm 1, \pm 3, \ldots, \pm (M-1) \).

- \( g(t) \) is a rectangular pulse of amplitude \( 1/2T \) and duration \( T \) seconds.
Continuous-phase FSK (CPFSK) (cont.)

- Equivalent low-pass waveform \( v(t) \) is expressed as

\[
v(t) = \sqrt{\frac{2\varepsilon}{T}} \exp\left\{ j \left[ 4\pi f_d \int_{-\infty}^{t} d(\tau) d\tau + \phi_0 \right] \right\}
\]

- \( f_d \) is the *peak frequency deviation*, \( \phi_0 \) is the initial phase of the carrier.

- The carrier-modulated signal may be expressed as

\[
s(t) = \sqrt{\frac{2\varepsilon}{T}} \cos\left[ 2\pi f_c t + \phi(t; I) + \phi_0 \right]
\]

where \( \phi(t; I) \) represents the time-varying phase of the carrier.
Continuous-phase FSK (CPFSK) (cont.)

\[ \phi(t; I) = 4\pi T f_d \int_{-\infty}^{t} d(\tau) d\tau = 4\pi T f_d \int_{-\infty}^{t} \left[ \sum_{n} I_n g(\tau - nT) \right] d\tau \]

\[ = 2\pi f_d T \sum_{k = -\infty}^{n-1} I_k + 2\pi f_d (t - nT) I_n \quad \text{for} \quad nT \leq t \leq (n+1)T \]

\[ \theta_n = \pi h \sum_{k = -\infty}^{n-1} I_k \]

\[ h = 2 f_d T \]

\[ q(t) = \begin{cases} 
0 & (t < 0) \\
\frac{t}{2T} & (0 \leq t \leq T) \\
\frac{1}{2} & (t > T) 
\end{cases} \]

Note that, although \( d(t) \) contains discontinuities, the integral of \( d(t) \) is continuous. Hence, we have a continuous-phase signal.

\( \theta_n \) represents the accumulation (memory) of all symbols up to time \( nT \).

Parameter \( h \) is called the modulation index.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

Continuous-phase modulation (CPM)

- CPFSK becomes a special case of a general class of continuous-phase modulated (CPM) signals in which the carrier phase is

\[ \phi(t; I) = 2\pi \sum_{k=-\infty}^{n} I_k h_k q(t - kT), \quad nT \leq t \leq (n + 1)T \]

- when \( h_k = h \) for all \( k \), the modulation index is fixed for all symbols.
- when \( h_k \) varies from one symbol to another, the CPM signal is called multi-\( h \). (In such a case, the \( \{h_k\} \) are made to vary in a cyclic manner through a set of indices.)

- The waveform \( q(t) \) may be represented in general as the integral of some pulse \( g(t) \), i.e.,

\[ q(t) = \int_{0}^{t} g(\tau) d\tau \]
Continuous-phase modulation (CPM) (cont.)

- If \( g(t) = 0 \) for \( t > T \), the CPM signal is called **full response CPM**. (Fig a. b.)
- If \( g(t) \neq 0 \) for \( t > T \), the modulated signal is called **partial response CPM**. (Fig c. d.)

![Graphs showing different cases of CPM]

- LREC, \( L=1 \), results in CPFSK
- LREC, \( L=2 \)
- LRC, \( L=1 \)
- LRC, \( L=2 \)
Continuous-phase modulation (CPM) (cont.)

- The CPM signal has memory that is introduced through the phase continuity.
- For \( L > 1 \), additional memory is introduced in the CPM signal by the pulse \( g(t) \).
- Three popular pulse shapes are given in the following table.
  - LREC denotes a rectangular pulse of duration \( LT \).
  - LRC denotes a raised cosine pulse of duration \( LT \).
  - *Gaussian minimum-shift keying* (GMSK) pulse with bandwidth parameter \( B \), which represents the -3 dB bandwidth of the Gaussian pulse.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- Continuous-phase modulation (CPM) (cont.)
  - Some commonly used CPM pulse shapes
    - LREC
      \[ g(t) = \begin{cases} 
      \frac{1}{2LT} & (0 \leq 1 \leq LT) \\
      0 & \text{(otherwise)} 
      \end{cases} \]
    - LRC
      \[ g(t) = \begin{cases} 
      \frac{1}{2LT} \left(1 - \cos \frac{2\pi t}{LT}\right) & (0 \leq 1 \leq LT) \\
      0 & \text{(otherwise)} 
      \end{cases} \]
    - GMSK
      \[ g(t) = \left\{ Q\left[2\pi B\left(t - \frac{T}{2}\right)(\ln 2)^{1/2}\right] - Q\left[2\pi B\left(t + \frac{T}{2}\right)(\ln 2)^{1/2}\right]\right\} \]
      \[ Q(t) = \int_{t}^{-\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dt \]
Minimum-shift keying (MSK).

- MSK is a special form of binary CPFSK (and, therefore, CPM) in which the modulation index $h=1/2$.
- The phase of the carrier in the interval $nT \leq t \leq (n+1)T$ is
  \[ \phi(t; I) = \frac{1}{2} \pi \sum_{k=-\infty}^{n-1} I_k + \pi I_n q(t-nT) \]
  
  \[ = \theta_n + \frac{1}{2} \pi I_n \left( \frac{t-nT}{T} \right) , \quad nT \leq t \leq (n+1)T \]

- The modulated carrier signal is
  \[ s(t) = A \cos \left[ 2\pi f_c t + \theta_n + \frac{1}{2} \pi I_n \left( \frac{t-nT}{T} \right) \right] \]

  \[ = A \cos \left[ 2\pi \left( f_c + \frac{1}{4T} I_n \right) t - \frac{1}{2} n\pi I_n + \theta_n \right] , \quad nT \leq t \leq (n+1)T \]
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

Minimum-shift keying (MSK) (cont.)

- The expression indicates that the binary CPFSK signal can be expressed as a sinusoid having one of two possible frequencies in the interval \( nT \leq t \leq (n+1)T \). If we define these frequencies as

\[
    f_1 = f_c + \frac{1}{4T} \\
    f_2 = f_c - \frac{1}{4T}
\]

- Then the binary CPFSK signal may be written in the form

\[
    s_i(t) = A \cos \left[ 2\pi f_i t + \theta_n + \frac{1}{2} n\pi (-1)^{i-1} \right], \quad i = 1, 2
\]
Minimum-shift keying (MSK) (cont.)

Why binary CPFSK with \( h = 1/2 \) is called minimum-shift keying (MSK)?

Because the frequency separation \( \Delta f = f_2 - f_1 = 1/2 T \), and \( \Delta f = 1/2 T \) is the minimum frequency separation that is necessary to ensure the orthogonality of the signals \( s_1(t) \) and \( s_2(t) \) over a signaling interval of length \( T \).

The phase in the \( n \)th signaling interval is the phase state of the signal that results in phase continuity between adjacent interval.
Minimum-shift keying (MSK) (Haykin)

Consider a continuous-phase frequency-shift keying (CPFSK) signal, which is defined for the interval $0 \leq t \leq T_b$ as follows:

$$s(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_1 t + \theta(0)] & \text{for symbol 1} \\ \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_2 t + \theta(0)] & \text{for symbol 0} \end{cases}$$

- $E_b$ is the transmitted signal energy per bit.
- $T_b$ is the bit duration.
- The phase $\theta(0)$, denoting the value of the phase at time $t = 0$, sums up the past history of the modulation process up to time $t = 0$. 

4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM
Minimum-shift keying (MSK) (Haykin)

Another useful way of representing the CPFSK signal $s(t)$ is to express it in the conventional form of an angle-modulated signal as follows:

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_c t + \theta(t)]$$  (*)

θ (t) is the phase of $s(t)$.

When the phase θ (t) is a continuous function of time, we find that the modulated signal $s(t)$ itself is also continuous at all times, including the inter-bit switching times.
Minimum-shift keying (MSK) (Haykin)

- The phase $\theta (t)$ of a CPFSK signal increases or decreases linearly with time during each bit duration of $T_b$ seconds

$$\theta (t) = \theta (0) \pm \frac{\pi h}{T_b} t, \quad 0 \leq t \leq T_b$$

The plus (minus) sign corresponds to sending symbol 1 (0).

- We can find that

$$f_c + \frac{h}{2T_b} = f_1 \quad f_c - \frac{h}{2T_b} = f_2$$

- We thus get

$$f_c = \frac{1}{2}(f_1 + f_2) \quad h = T_b(f_1 + f_2)$$
Minimum-shift keying (MSK) (Haykin)

- $h$ is referred to as the *deviation ratio*.
- With $h=1/2$, the frequency deviation equals half the bit rate. This is the minimum frequency spacing that allows the two FSK signals representing symbols 1 and 0 to be coherently orthogonal.
- At time $t = T_b$:
  \[
  \theta(T_b) - \theta(0) = \begin{cases} 
  \pi h & \text{for symbol 1} \\
  -\pi h & \text{for symbol 0}
  \end{cases}
  \]

- This is to say, the sending of symbol 1 increases the phase of a CPFSK signal $s(t)$ by $\pi h$ radians, whereas the sending of symbol 0 reduces it by an equal amount.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- Minimum-shift keying (MSK) (Haykin)
- Phase Trellis
Minimum-shift keying (MSK) (Haykin)

Phase trellis of the sequence 1101000 ($h = 1/2$)

From (*), we may express the CPFSK signal $s(t)$:

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos \left[ \theta(t) \right] \cos (2\pi f_c t) - \sqrt{\frac{2E_b}{T_b}} \sin \left[ \theta(t) \right] \sin (2\pi f_c t)$$

$$\equiv S_l \cos (2\pi f_c t) - S_Q \sin (2\pi f_c t) \quad (**)$$
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- Minimum-shift keying (MSK) (Haykin)
  - Consider first the in-phase component $S_I$
  - With the deviation ratio $h=1/2$,
    \[ \theta(t) = \theta(0) \pm \frac{\pi}{2T_b} t, \quad 0 \leq t \leq T_b \]
    where the plus (minus) sign corresponds to symbol 1 (0).
  - A similar result holds for $\theta(t)$ in the interval $-T_b \leq t \leq 0$, except that the algebraic sign is not necessarily the same in both intervals.
  - Since the phase $\theta(0)$ is 0 or $\pi$, depending on the past history of the modulation process, we find that, in the interval $-T_b \leq t \leq T_b$, the polarity of $\cos[\theta(t)]$ depends only on $\theta(0)$, regardless of the sequence of 1s and 0s transmitted before or after $t = 0$. 
Minimum-shift keying (MSK) (Haykin)

Thus, for $-T_b \leq t \leq T_b$, the in-phase component $s_I(t)$ consists of a half-cycle cosine pulse defined as follows:

$$s_I(t) = \sqrt{\frac{2E_b}{T_b}} \cos \left[ \theta(t) \right] = \sqrt{\frac{2E_b}{T_b}} \cos \left[ \theta(0) \right] \cos \left( \frac{\pi}{2T_b} t \right)$$

$$= \pm \sqrt{\frac{2E_b}{T_b}} \cos \left( \frac{\pi}{2T_b} t \right), \quad -T_b \leq t \leq T_b$$

where the plus (minus) sign corresponds to $\theta(0) = 0$ ($\theta(0) = \pi$).
Minimum-shift keying (MSK) (Haykin)

Similarly, for $0 \leq t \leq 2T_b$, the quadrature component $s_Q(t)$ consists of a half-cycle sine pulse, whose polarity depends only on $\theta (T_b)$, defined as follows:

$$s_Q(t) = \sqrt{\frac{2E_b}{T_b}} \sin \left[ \theta (t) \right] = \sqrt{\frac{2E_b}{T_b}} \sin \left[ \theta (T_b) \right] \sin \left( \frac{\pi}{2T_b} t \right)$$

$$= \pm \sqrt{\frac{2E_b}{T_b}} \sin \left( \frac{\pi}{2T_b} t \right), \quad 0 \leq t \leq 2T_b$$

where the plus (minus) sign corresponds to $\theta (T_b) = \pi / 2$ ($\theta (T_b) = -\pi / 2$).
Minimum-shift keying (MSK) (Haykin)

MSK signal may assume any one of four possible forms:

- The phase $\theta(0)=0$ and $\theta(T_b)=\pi/2$, corresponding to the transmission of symbol 1.
- The phase $\theta(0)=\pi$ and $\theta(T_b)=\pi/2$, corresponding to the transmission of symbol 0.
- The phase $\theta(0)=\pi$ and $\theta(T_b)=-\pi/2$, corresponding to the transmission of symbol 1.
- The phase $\theta(0)=0$ and $\theta(T_b)=-\pi/2$, corresponding to the transmission of symbol 0.
Minimum-shift keying (MSK) (Haykin)

Define the orthonormal basis functions:

\[
\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos\left(\frac{\pi}{2T_b} t\right) \cos(2\pi f_c t), \quad 0 \leq t \leq T_b
\]

\[
\phi_2(t) = \sqrt{\frac{2}{T_b}} \sin\left(\frac{\pi}{2T_b} t\right) \sin(2\pi f_c t), \quad 0 \leq t \leq T_b
\]

From (**) , we have:

\[
s(t) = s_1 \phi_1(t) + s_2 \phi_2(t), \quad 0 \leq t \leq T_b
\]

where the coefficients \( s_1 \) and \( s_2 \) are related to the phase states \( \theta(0) \) and \( \theta(T_b) \), respectively.
Minimum-shift keying (MSK) (Haykin)

\[ s_1 = \int_{-T_b}^{T_b} s(t) \phi_1(t) \, dt = \sqrt{E_b} \cos[\theta(0)], \quad -T_b \leq t \leq T_b \]

\[ s_2 = \int_{0}^{2T_b} s(t) \phi_2(t) \, dt = -\sqrt{E_b} \sin[\theta(T_b)], \quad 0 \leq t \leq 2T_b \]

- Both integrals are evaluated for a time interval equal to twice the bit duration.
- Both the lower and upper limits of the product integration used to evaluate the coefficient \( s_1 \) are shifted by the bit duration \( T_b \) with respect to those used to evaluate the coefficient \( s_2 \).
- The time interval \( 0 \leq t \leq T_b \), for which the phase states \( \theta(0) \) and \( \theta(T_b) \) are defined, is common to both integrals.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- Minimum-shift keying (MSK) (Haykin)
- Signal space diagram for MSK system


### Minimum-shift keying (MSK) (Haykin)

In QPSK the transmitted symbol is represented by any one of the four message points, whereas in MSK one of two message points is used to represent the transmitted symbol at any one time.

<table>
<thead>
<tr>
<th>Transmitted Binary Symbol, $0 \leq t \leq T_b$</th>
<th>Phase States (radians)</th>
<th>Coordinates of Message Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0$</td>
<td>$\theta(0)$ $\theta(T_b)$</td>
<td>$s_1$ $s_2$</td>
</tr>
<tr>
<td>$1$</td>
<td>$\pi$ $\pi/2$</td>
<td>$-\sqrt{E_b}$ $\sqrt{E_b}$</td>
</tr>
<tr>
<td>$0$</td>
<td>$\pi$ $+\pi/2$</td>
<td>$-\sqrt{E_b}$ $-\sqrt{E_b}$</td>
</tr>
<tr>
<td>$1$</td>
<td>$0$ $+\pi/2$</td>
<td>$\sqrt{E_b}$ $-\sqrt{E_b}$</td>
</tr>
</tbody>
</table>
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

**Figure 6.30** (a) Input binary sequence. (b) Waveform of scaled time function $s_1 \phi_1(t)$. (c) Waveform of scaled time function $s_2 \phi_2(t)$. (d) Waveform of the MSK signal $s(t)$ obtained by adding $s_1 \phi_1(t)$ and $s_2 \phi_2(t)$ on a bit-by-bit basis.
Minimum-shift keying (MSK) (Haykin)

- Coherent detection of the MSK signal
  - In the case of an AWGN channel, the received signal is given by
    \[ x(t) = s(t) + w(t) \]
    where \( s(t) \) is the transmitted MSK signal, and \( w(t) \) is the sample function of a white Gaussian noise process of zero mean and power spectral density \( N_0/2 \).
  - For the optimal decision of \( \theta(0) \),
    \[ x_1 = \int_{-T_b}^{T_b} x(t) \phi_1(t) \, dt = s_1 + w_1, \quad -T_b \leq t \leq T_b \]
    If \( x_1 > 0 \) \( \Rightarrow \) \( \hat{\theta}(0) = 0 \)
    If \( x_1 < 0 \) \( \Rightarrow \) \( \hat{\theta}(0) = \pi \)
Minimum-shift keying (MSK) (Haykin)

- Coherent detection of the MSK signal (cont.)
  - For the optimum detection of $\theta (T_b)$,
    \[
    x_2 = \int_0^{2T_b} x(t) \phi_2(t) \, dt = s_2 + w_2, \quad 0 \leq t \leq 2T_b
    \]
  
  If $x_2 > 0 \implies \hat{\theta}(T_b) = -\pi / 2$

  If $x_2 < 0 \implies \hat{\theta}(T_b) = \pi / 2$

- Error probability of MSK
  \[
  P_e = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{E_b}{N_0}} \right)
  \]

  which is exactly the same as that for binary PSK and QPSK.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- Minimum-shift keying (MSK) (Haykin)
- Block diagrams for MSK transmitter
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- Minimum-shift keying (MSK) (Haykin)
- Block diagrams for coherent MSK receiver
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- **Minimum-shift keying (MSK) (Haykin)**

- **Power spectra of MSK signals**

\[
S_B(f) = 2 \left[ \frac{\psi_g(f)}{2T_b} \right] = \frac{32E_b}{\pi^2} \left[ \frac{\cos(2\pi T_b f)}{16T_b^2 f^2 - 1} \right]^2
\]

MSK does not produce as much interference outside the signal band of interest as QPSK.
Minimum-shift keying (MSK) (cont.)

- MSK may also be represented as a form of four-phase PSK.
- The equivalent low-pass digitally modulated signal in the form

\[ v(t) = \sum_{n=-\infty}^{\infty} \left[ I_{2n} g(t-2nT) - jI_{2n+1} g(t-2nT-T) \right] \]

- \( g(t) \) is a sinusoidal pulse

\[
g(t) = \begin{cases} 
\sin \frac{\pi t}{2T} & (0 \leq t \leq 2T) \\
0 & \text{(otherwise)}
\end{cases}
\]
Minimum-shift keying (MSK) (cont.)

- This type of signal is viewed as a four-phase PSK signal in which the pulse shape is one-half cycle of a sinusoid (0 ~ $\pi$).
- The even-numbered binary-valued ($\pm 1$) symbols $\{I_{2n}\}$ of the information sequence $\{I_n\}$ are transmitted via the cosine of the carrier, while the odd-numbered symbols $\{I_{2n+1}\}$ are transmitted via the sine of the carrier.
- The transmission rate on the two orthogonal carrier components is $1/2T$ bits/s so that the combined transmission rate is $1/T$ bits/s.
- Note that the bit transitions on the sine and cosine carrier components are staggered or offset in time by $T$ seconds.
Minimum-shift keying (MSK) (cont.)

For this reason, the signal

\[ s(t) = A \left\{ \sum_{n=-\infty}^{\infty} I_{2n}g(t - 2nT) \cos 2\pi f_c t \right\} + \left\{ \sum_{n=-\infty}^{\infty} I_{2n+1}g(t - 2nT - T) \sin 2\pi f_c t \right\} \]

is called offset quadrature PSK (OQPSK) or staggered quadrature PSK (SQPSK).

Figure in next page illustrates the representation of an MSK signal as two staggered quadrature-modulated binary PSK signals. The corresponding sum of the two quadrature signals is a constant amplitude, frequency-modulated signal.
4.3.3 Non-linear Modulation Methods with Memory---CPFSK and CPM

- Minimum-shift keying (MSK) (cont.)

(a) In-phase signal component

(b) Quadrature signal component

(c) MSK signal [sum of (a) and (b)]
Minimum-shift keying (MSK) (cont.)

- Compare the waveforms for MSK with OQPSK (pulse $g(t)$ is rectangular for $0 \leq t \leq 2T$) and with conventional QPSK (pulse $g(t)$ is rectangular for $0 \leq t \leq 2T$).
  - All three of the modulation methods result in identical data rates.
  - The MSK signal has continuous phase.
  - The OQPSK signal with a rectangular pulse is basically two binary PSK signals for which the phase transitions are staggered in time by $T$ seconds. Thus, the signal contains phase jumps of $\pm 90^\circ$.
  - The conventional four-phase PSK (QPSK) signal with constant amplitude will contain phase jumps of $\pm 180^\circ$ or $\pm 90^\circ$ every $2T$ seconds.
Minimum-shift keying (MSK) (cont.)

- Compare the waveforms for MSK with OQPSK and QPSK (cont.)
Minimum-shift keying (MSK) (cont.)
4.4 Spectral Characteristics of Digitally Modulated Signals

- In most digital communication systems, the available channel bandwidth is limited.

- The system designer must consider the constraints imposed by the channel bandwidth limitation in the selection of the modulation technique used to transmit the information.

- From the power density spectrum, we can determine the channel bandwidth required to transmit the information-bearing signal.
Beginning with the form

\[ s(t) = \text{Re} \left[ \nu(t) e^{j2\pi f_c t} \right] \]

where \( \nu(t) \) is the equivalent low-pass signal.

Autocorrelation function (from 4.1-50)

\[ \phi_{ss}(\tau) = \text{Re} \left[ \phi_{\nu\nu}(\tau) e^{j2\pi f_c \tau} \right] \]

Power density spectrum (from 4.1-51)

\[ \Phi_{ss}(f) = \frac{1}{2} \left[ \Phi_{\nu\nu}(f - f_c) + \Phi_{\nu\nu}(-f - f_c) \right] \]

First we consider the general form

\[ \nu(t) = \sum_{n=-\infty}^{\infty} I_n g(t - nT) \]

where the transmission rate is \( 1/T = R/k \) symbols/s and \( \{I_n\} \) represents the sequence of symbols.
4.4.1 Power Spectra of Linearly Modulation Signals

Autocorrelation function

\[ \phi_{vv}(t + \tau; t) = \frac{1}{2} E[\nu^*(t)\nu(t + \tau)] \]

\[ = \frac{1}{2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} E[I_n^*I_m]g^*(t - nT)g(t + \tau - mT) \]

We assume the \{I_n\} is WSS with mean \( \mu_i \) and the autocorrelation function

\[ \phi_{ii}(m) = \frac{1}{2} E[I_n^*I_{n+m}] \]

\[ \phi_{vv}(t + \tau; t) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi_{ii}(m - n)g^*(t - nT)g(t + \tau - mT) \quad \text{let } m' = m - n \]

\[ = \sum_{m'=-\infty}^{\infty} \phi_{ii}(m') \sum_{n=-\infty}^{\infty} g^*(t - nT)g\left(t + \tau - (m' + n)T\right) \quad \text{let } m = m' \]

\[ = \sum_{m=-\infty}^{\infty} \phi_{ii}(m) \sum_{n=-\infty}^{\infty} g^*(t - nT)g(t + \tau - nT - mT) \]
The second summation
\[ \sum_{n=-\infty}^{\infty} g^*(t-nT)g(t+\tau-nT-mT) \]

is periodic in the \( t \) variable with period \( T \).

Consequently, \( \phi_{\nu\nu}(t+\tau;t) \) is also periodic in the \( t \) variable with period \( T \). That is
\[ \phi_{\nu\nu}(t+T+\tau;t+T) = \phi_{\nu\nu}(t+\tau;t) \]

In addition, the mean value of \( \nu(t) \), which is
\[ E[\nu(t)] = E\left[ \sum_{n=-\infty}^{\infty} I_n g(t-nT) \right] = \mu_i \sum_{n=-\infty}^{\infty} g(t-nT) \]

is periodic with period \( T \).
Therefore \( \nu(t) \) is a stochastic process having a periodic mean and autocorrelation function. Such a process is called a cyclostationary process or a periodically stationary process in the wide sense.

In order to compute the power density spectrum of a cyclostationary process, the dependence of \( \phi_{\nu \nu}(t+\tau;t) \) on the \( t \) variable must be eliminated. Thus,

\[
\bar{\phi}_{\nu \nu}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} \phi_{\nu \nu}(t+\tau;t)dt
\]

\[
= \sum_{m=-\infty}^{\infty} \phi_{ii}(m) \sum_{n=-\infty}^{\infty} \frac{1}{T} \int_{-T/2}^{T/2} g^*(t-nT)g(t+\tau-nT-mT)dt
\]

\[
= \sum_{m=-\infty}^{\infty} \phi_{ii}(m) \sum_{n=-\infty}^{\infty} \frac{1}{T} \int_{-T/2-nT}^{T/2-nT} g^*(t')g(t'+\tau-mT)dt' \quad (t' = t - nT)
\]
We interpret the integral as the *time-autocorrelation function* of \( g(t) \) and define it as:

\[
\phi_{gg}(\tau) = \int_{-\infty}^{\infty} g^*(t)g(t + \tau)\,dt
\]

Consequently,

\[
\bar{\phi}_{vv}(\tau) = \frac{1}{T} \sum_{m=-\infty}^{\infty} \phi_{ii}(m)\phi_{gg}(\tau - mT)
\]

The (average) power density spectrum of \( v(t) \) is in the form

\[
\Phi_{vv}(f) = \frac{1}{T} \left| G(f) \right|^2 \Phi_{ii}(f)
\]

where \( G(f) \) is the Fourier transform of \( g(t) \), and \( \Phi_{ii}(f) \) denotes the power density spectrum of the information sequence

\[
\Phi_{ii}(f) = \sum_{m=-\infty}^{\infty} \phi_{ii}(m)e^{-j2\pi fmT}
\]
The result illustrates the dependence of the power density spectrum of $v(t)$ on the spectral characteristics of the pulse $g(t)$ and the information sequence $\{I_n\}$.

That is, the spectral characteristics of $v(t)$ can be controlled by (1) design of the pulse shape $g(t)$ and by (2) design of the correlation characteristics of the information sequence.

Whereas the dependence of $\Phi_{nn}(f)$ on $G(f)$ is easily understood upon observation of equation, the effect of the correlation properties of the information sequence is more subtle.

First of all, we note that for an arbitrary autocorrelation $\phi_{ii}(m)$ the corresponding power density spectrum $\Phi_{ii}(f)$ is periodic in frequency with period $1/T$. (see next page)
In fact, the expression relating the spectrum $\Phi_{ii}(f)$ to the autocorrelation $\phi_{ii}(m)$ is in the form of an exponential Fourier series with the $\{\phi_{ii}(m)\}$ as the Fourier coefficients.

$$\phi_{ii}(m) = T \int_{-1/2T}^{1/2T} \Phi_{ii}(f) e^{j2\pi fmT} df$$

Second, let us consider the case in which the information symbols in the sequence are real and mutually uncorrelated. In this case, the autocorrelation function $\phi_{ii}(m)$ can be expressed as (applying 2.2-5 and 2.2-6)

$$\phi_{ii}(m) = \begin{cases} \sigma_i^2 + \mu_i^2 & (m = 0) \\ \mu_i^2 & (m \neq 0) \end{cases}$$

where $\sigma_i^2$ denotes the variance of an information symbol.
4.4.1 Power Spectra of Linearly Modulation Signals

- Substitute for $\phi_{ii}(m)$ in equation, we obtain

$$\Phi_{ii}(f) = \sum_{m=-\infty}^{\infty} \phi_{ii}(m)e^{-j2\pi fmT}$$

It may be viewed as the exponential Fourier series of a periodic train of impulses with each impulse having an area $1/T$.

$$\sum_{n=-\infty}^{\infty} \delta(t-nT_s) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} e^{jn\omega_s t} = \sigma_i^2 + \mu_i^2 \sum_{m=-\infty}^{\infty} e^{-j2\pi fmT} = \sigma_i^2 + \frac{\mu_i^2}{T} \sum_{m=-\infty}^{\infty} \delta(f - \frac{m}{T})$$

where $\omega_s = \frac{2\pi}{T_s}$

- The desired result for the power density spectrum of $v(t)$ when the sequence of information symbols is uncorrelated.

$$\Phi_{vv}(f) = \frac{\sigma_i^2}{T} |G(f)|^2 + \frac{\mu_i^2}{T^2} \sum_{m=-\infty}^{\infty} \left| G\left(\frac{m}{T}\right) \right|^2 \delta(f - \frac{m}{T})$$
4.4.1 Power Spectra of Linearly Modulation Signals

- The expression for the power density spectrum is purposely separated into two terms to emphasize the two different types of spectral components.
- The first term is the continuous spectrum, and its shape depends only on the spectral characteristic of the signal pulse $g(t)$.
- The second term consists of discrete frequency components spaced $1/T$ apart in frequency. Each spectral line has a power that is proportional to $|G(f)|^2$ evaluated at $f = m/T$.
- Note that the discrete frequency components vanish when the information symbols have zero mean, i.e., $\mu_i = 0$. This condition is usually desirable for the digital modulation techniques under consideration, and it is satisfied when the information symbols are equally likely and symmetrically positioned in the complex plane.
Example 4.4-1 To illustrate the spectral shaping resulting from \( g(t) \), consider the rectangular pulse shown in figure. The Fourier transform of \( g(t) \) is

\[
G(f) = A T \frac{\sin \pi f T}{\pi f T} e^{-i \pi f T}
\]

Hence

\[
|G(f)|^2 = (AT)^2 \left( \frac{\sin \pi f T}{\pi f T} \right)^2
\]

Thus

\[
\Phi_{uv}(f) = \sigma_i^2 A^2 T \left( \frac{\sin \pi f T}{\pi f T} \right)^2 + A^2 \mu_i^2 \delta(f)
\]

Decays inversely as the square of the frequency
Example 4.4-2  As a second illustration of the spectral shaping resulting from \( g(t) \), we consider the raised cosine pulse

\[
g(t) = \frac{A}{2} \left[ 1 + \cos \frac{2\pi}{T} \left( t - \frac{T}{2} \right) \right], \quad 0 \leq t \leq T
\]

its Fourier transform is:

\[
G(f) = \frac{A}{2} \frac{\sin \frac{\pi f T}{2}}{\pi f T (1 - f^2 T^2)} e^{-j \pi f T}
\]

Decays inversely as the \( f^6 \)
Example 4.4-3  To illustrate that spectral shaping can also be accomplished by operations performed on the input information sequence, we consider a binary sequence \( \{b_n\} \) from which we form the symbols \( I_n = b_n + b_{n-1} \)

- The \( \{b_n\} \) are assumed to be uncorrelated random variables, each having zero mean and unit variance. Then the autocorrelation function of the sequence \( \{I_n\} \) is

\[
\phi_{ii}(m) = E(I_n I_{n+m}) = E[(b_n + b_{n-1})(b_{n+m} + b_{n+m-1})]
\]

\[
= \begin{cases} 
E\left[b_n^2 + 2b_nb_{n-1} + b_{n-1}^2\right] & (m = 0) \\
E\left[b_n^2 + b_nb_{n+1} + b_{n-1}b_{n+1} + b_nb_{n-1}\right] & (m = +1) \\
E\left[b_nb_{n-1} + b_nb_{n-2} + b_{n-1}^2 + b_{n-1}b_{n-2}\right] & (m = -1) \\
E\left[b_nb_{n+m} + b_nb_{n+m-1} + b_{n-1}b_{n+m} + b_{n-1}b_{n+m-1}\right] & (\text{otherwise}) 
\end{cases}
\]

\( E[X_i X_j] = E[X_i] E[X_j] \)

\( E[(X_i - 0)^2] = 1 \)
Example 4.4-3 (cont.)

Hence, the power density spectrum of the input sequence is (from 4.4-13)

\[ \Phi_{ii}(f) = \sum_{m=-\infty}^{\infty} \phi_{ii}(m) e^{-j2\pi fmT} = 2(1 + \cos 2\pi fT) = 4\cos^2 \pi fT \]

and the corresponding power density spectrum for the (low-pass) modulated signal is (from 4.4-12)

\[ \Phi_{vv}(f) = \frac{4}{T} |G(f)|^2 \cos^2 \pi fT \]