Chapter 2
Deterministic and Random Signal Analysis

Wireless Information Transmission System Lab.
Institute of Communications Engineering
National Sun Yat-sen University
2.1. Bandpass and Lowpass Signal Representation
2.2. Signal Space Representation of Waveforms
2.3. Some Useful Random Variables
2.4. Bounds on Tail Probabilities
2.5. Limit Theorems for Sums of Random Variables
2.6. Complex Random Variables
2.7. Random Processes
2.8. Series Expansion of Random Processes
2.9. Bandpass and Lowpass Random Processes
Signal Definitions

\[
\Pi(t) = \begin{cases} 
1 & |t| < \frac{1}{2} \\
\frac{1}{2} & t = \pm \frac{1}{2} \\
0 & \text{otherwise}
\end{cases}
\]

\[
sinc(t) = \begin{cases} 
\frac{\sin(\pi t)}{\pi t} & t \neq 0 \\
1 & t = 0
\end{cases}
\]

and

\[
\text{sgn}(t) = \begin{cases} 
1 & t > 0 \\
-1 & t < 0 \\
0 & t = 0
\end{cases}
\]

\[
\Lambda(t) = \Pi(t) \ast \Pi(t) = \begin{cases} 
t + 1 & -1 \leq t < 0 \\
-t + 1 & 0 \leq t < 1 \\
0 & \text{otherwise}
\end{cases}
\]

The unit step signal \( u_{-1}(t) \) is defined as

\[
u_{-1}(t) = \begin{cases} 
1 & t > 0 \\
\frac{1}{2} & t = 0 \\
0 & t < 0
\end{cases}
\]
Chapter 2.1 : Bandpass and Lowpass Signal Representations
In many cases the information signal is a low frequency (baseband) signal, and the available spectrum of the communication channel is at higher frequencies.

At the transmitter the information signal is translated to a higher frequency signal that matches the properties of the communication channel.

This is the modulation process in which the baseband information signal is turned into a bandpass modulated signal.

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).
We will show that any real, narrowband, and high frequency signal – called a bandpass signal – can be represented in terms of a complex low frequency signal, called the lowpass equivalent of the original bandpass signal.

This result makes it possible to work with the lowpass equivalents of bandpass signals instead of directly working with them, thus greatly simplifying the handling of bandpass signals.

That is so because applying signal processing algorithms to lowpass signals is much easier due to lower required sampling rates which in turn result in lower rates of the sampled data.
The Fourier transform of a signal provides the frequency content or spectrum of a signal: \( X(f) = \mathcal{F}\{x(t)\} \)

The Fourier transform of a real signal:

- Hermitian symmetry: \( X(-f) = X^*(f) \)
  \[ \rightarrow |X(-f)| = |X(f)|: \text{even function} \]
  \[ \rightarrow \angle X(-f) = -\angle X(f): \text{odd function} \]

- All information about the signal is in the positive (or negative) frequencies, and \( x(t) \) can be perfectly reconstructed by specifying \( X(f) \) for \( f \geq 0 \).

- **Bandwidth**: the smallest range of positive frequencies such that \( X(f) = 0 \) when \( |f| \) is outside this range.

- The bandwidth of a real signal is one-half of its frequency support set.

- The *frequency support*, the range of frequencies for which \( X(f) \neq 0 \), is \([-W, +W]\).
Lowpass Signals

- **Lowpass signal (baseband signal):**
  - A signal whose spectrum is located around zero frequency
  - Examples: speech, music, video signals
  - Lowpass signal → Usually Low frequency → Slowly varying in time

- The bandwidth of a real low-pass signal is the minimum positive $W$ such that $X(f)=0$ outside $[-W,+W]$. 

![Diagram of lowpass signal spectrum](image)
Chapter 2.1-1 Bandpass and Lowpass Signals

Positive & Negative Spectrum

Positive spectrum and negative spectrum

\[
X_+(f) = \begin{cases} 
X(f) & f > 0 \\
\frac{1}{2} X(f) & f = 0 \\
0 & f < 0 
\end{cases} \quad X_-(f) = \begin{cases} 
X(f) & f < 0 \\
\frac{1}{2} X(f) & f = 0 \\
0 & f > 0 
\end{cases}
\]

For a real signal \(x(t)\)

\(X_-(f) = X_+^*(-f)\) \quad \(X(-f) = X^*(f)\)

For a complex signal, \(X(f)\) is not symmetric

- Signal cannot be reconstructed from information in positive frequency
- Bandwidth is defined as one-half of the entire range of frequencies over which the spectrum is nonzero, i.e., one-half of the frequency support of the signal.
Bandpass Signals

- The spectral characteristics of the signal and the communication channel do not always match.
- The signal should be modulated to match the spectral characteristics of the channel, where the spectrum of the lowpass signal is translated to a higher frequency. The resulting modulated signal is a bandpass signal.
- A bandpass signal is a *real signal* whose frequency content, or spectrum, is located around some frequency $\pm f_0$, which is far from zero.
  - $f_0$: *central frequency*
  - High frequency $\rightarrow$ rapid variation in time
  - $X_+(f)$ is nonzero only in $[f_0-W/2, f_0+W/2]$
Bandpass Signals

- The spectrum of a real-valued bandpass signal.

- The magnitude spectrum (solid line) is even.

- The phase spectrum (dashed line) is odd.

- The central frequency is not necessarily the midband frequency of the bandpass signal.

- $X_+(f)$ is sufficient to reconstruct $X(f)$:

$$X(f) = X_+(f) + X_-(f) = X_+(f) + X^*_-(f)$$
Pre-envelop

◊ Pre-envelop of a signal $x(t)$ (or analytic signal)

$$x_+(t) = F^{-1} \{ X_+(f) \}$$

$$= F^{-1} \{ X(f)u_-(f) \}$$

$$= x(t) \ast \left( \frac{1}{2} \delta(t) + j \frac{1}{2\pi t} \right)$$

$$= \frac{1}{2} x(t) + \frac{j}{2} \hat{x}(t)$$

◊ $\hat{x}(t) = x(t) \ast (1/\pi t)$ is the Hilbert transform of $x(t)$ with spectrum $F[\hat{x}(t)] = -j \text{sgn}(f)X(f)$

◊ Phase shift of $-\pi/2$ in positive spectrum

◊ Phase shift of $+\pi/2$ in negative spectrum
A filter, called a *Hilbert transformer*, is defined as:

\[ h(t) = \frac{1}{\pi t}, \quad -\infty < t < \infty \]

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

The frequency response of this filter is:

\[
H(f) = \int_{-\infty}^{\infty} h(t) e^{-j 2\pi ft} dt = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{t} e^{-j 2\pi ft} dt = \begin{cases} -j & (f > 0) \\
0 & (f = 0) \\
j & (f < 0) \end{cases}
\]

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.
**Chapter 2.1-2  Lowpass Equivalent of Bandpass Signals**

**Lowpass Equivalent of Signals**

- **Lowpass equivalent** (or complex envelop) of \( x(t) \)
  \[
  x_l(t) = F^{-1}\{X_l(f)\}
  \]

- where \( X_l(f) = 2X_+(f + f_0) \)

- Spectrum of \( x_l(t) \) is located around zero frequency
  \[
  x_l(t) = F^{-1}\{2X_+(f + f_0)\} \\
  = 2x_+(t)e^{-j2\pi f_0 t} \\
  = (x(t) + j\hat{x}(t))e^{-j2\pi f_0 t} \quad (2.1-6)
  \]

\[
= [x(t)\cos 2\pi f_0 t + \hat{x}(t)\sin 2\pi f_0 t] \\
+ j[\hat{x}(t)\cos 2\pi f_0 t - x(t)\sin 2\pi f_0 t] \\
(2.1-7)
\]
Chapter 2.1-2 Lowpass Equivalent of Bandpass Signals

From 2.1-6

Since

\[ x(t) = \text{Re}\left\{ x_i(t)e^{j2\pi f_0 t} \right\} = \frac{1}{2} \left( x_i(t)e^{j2\pi f_0 t} + x_i^*(t)e^{-j2\pi f_0 t} \right) \] (2.1-8)

we have

\[ X(f) = \frac{1}{2} \left( X_i(f - f_0) + X_i^*(f + f_0) \right) \]

The low pass equivalent \( x_i(t) \) is complex,

let \( x_i(t) = x_i(t) + j x_q(t) \) \hspace{1cm} (2.1-10)

- **In-phase**: \( x_i(t) = x(t)\cos 2\pi f_0 t + \hat{x}(t)\sin 2\pi f_0 t \) \hspace{1cm} (2.1-11)
- **Quadrature-phase**: \( x_q(t) = \hat{x}(t)\cos 2\pi f_0 t - x(t)\sin 2\pi f_0 t \)

\[ x(t) = x_i(t)\cos 2\pi f_0 t - x_q(t)\sin 2\pi f_0 t \] \hspace{1cm} (2.1-12)

\[ \hat{x}(t) = x_q(t)\cos 2\pi f_0 t + x_i(t)\sin 2\pi f_0 t \]

Bandpass signal \( x(t) \) can be expressed in terms of two lowpass signals.
Chapter 2.1-2 Lowpass Equivalent of Bandpass Signals

- \( x_l(t) \) can be also be written as:

\[
x_l(t) = r_x(t) e^{j \theta_x(t)}
\]

where \( r_x(t) = \sqrt{x_i^2(t) + x_q^2(t)} \) and \( \theta_x(t) = \tan^{-1} \frac{x_q(t)}{x_i(t)} \)

- \( x(t) \) can be represented as:

From 2.1-8

\[
x(t) = \text{Re} \left[ x_l(t) e^{j 2\pi f_0 t} \right] = \text{Re} \left[ r_x(t) e^{j [2\pi f_0 t + \theta_x(t)]} \right]
\]

\[
= r_x(t) \cos \left[ 2\pi f_0 t + \theta_x(t) \right]
\]

\( r_x(t) \) is called the \textit{envelope} of \( x(t) \), and \( \theta_x(t) \) is called the \textit{phase} of \( x(t) \).
Three equivalent representations of band-pass signals:

\[ x(t) = x_i(t) \cos 2\pi f_0 t - x_q(t) \sin 2\pi f_0 t \]

\[ = \text{Re} \left[ x_l(t) e^{j2\pi f_0 t} \right] \]

\[ = r_x(t) \cos \left[ 2\pi f_0 t + \theta_x(t) \right] \]

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

\[ X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt = \int_{-\infty}^{\infty} \left\{ \text{Re} \left[ x_l(t) e^{j2\pi f_0 t} \right] \right\} e^{-j2\pi ft} dt \]

\[ X(f) = \frac{1}{2} \int_{-\infty}^{\infty} \left[ x_l(t) e^{j2\pi f_0 t} + x_l^*(t) e^{-j2\pi f_0 t} \right] e^{-j2\pi ft} dt \]

\[ = \frac{1}{2} \left[ X_l(f - f_0) + X_l^*(-f - f_0) \right] \]
The energy in the signal $x(t)$ is defined as:

$$E_x = \int_{-\infty}^{\infty} x^2(t) \, dt = \int_{-\infty}^{\infty} \left\{ \text{Re} \left[ x_l(t) e^{j2\pi f_0 t} \right] \right\}^2 \, dt$$

$$E_x = \frac{1}{4} \int_{-\infty}^{\infty} \left[ x_l^2 e^{j4\pi f_0 t} + 2x_l x_l^* + (x_l^*)^2 e^{-j4\pi f_0 t} \right] \, dt$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} |x_l(t)|^2 \, dt + \frac{1}{4} \int_{-\infty}^{\infty} \left[ r_x^2(t) e^{j(4\pi f_0 t + 2\theta_x(t))} + (r_x^*(t))^2 e^{-j(4\pi f_0 t + 2\theta_x(t))} \right] \, dt$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} |x_l(t)|^2 \, dt + \frac{1}{2} \int_{-\infty}^{\infty} |x_l(t)|^2 \cos \left[ 4\pi f_0 t + 2\theta_x(t) \right] \, dt$$

where $|x_l(t)|^2 = r_x^2(t) = (r_x^*(t))^2$
Modulation from two lowpass signals to a bandpass signal

\[ x(t) = \text{Re} \left\{ x_i(t)e^{j2\pi f_0 t} \right\} \]

From 2.1-8

\[ x(t) = x_i(t) \cos 2\pi f_0 t - x_q(t) \sin 2\pi f_0 t \]

From 2.1-12
Demodulation: extract two lowpass signals from a bandpass signal

(1) \( x_i(t) = (x(t) + j\hat{x}(t))e^{-j2\pi f_0 t} \)  \( \text{From 2.1-6} \)

(2) \( x_i(t) = x(t)\cos 2\pi f_0 t + \hat{x}(t)\sin 2\pi f_0 t \)  
\( x_q(t) = \hat{x}(t)\cos 2\pi f_0 t - x(t)\sin 2\pi f_0 t \)  \( \text{From 2.1-11} \)
Chapter 2.1-3  Energy Considerations

◊ Energy of a signal $x(t)$ (Rayleigh’s relation, Table 2.0-1)

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |X(f)|^2 \, df$$

◊ $X(f) = X_+(f) + X_-(f)$, and $X_+(f)$ and $X_-(f)$ are non-overlapped

$$E_x = \int_{-\infty}^{\infty} |X_+(f) + X_-(f)|^2 \, df$$

$$= \int_{-\infty}^{\infty} |X_+(f)|^2 \, df + \int_{-\infty}^{\infty} |X_-(f)|^2 \, df$$

$$= 2 \int_{-\infty}^{\infty} |X_+(f)|^2 \, df = \boxed{2E_{x_+}}$$

$$= 2 \int_{-\infty}^{\infty} \left| \frac{X_l(f)}{2} \right|^2 \, df = \boxed{\frac{1}{2}E_{x_l}}$$
Chapter 2.1-3  Energy Considerations

- **Inner product** of two signals (Parseval’s relation)
  \[ \langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t)y^*(t)dt = \int_{-\infty}^{\infty} X(f)Y^*(f)dt \]

- \[ E_x = \langle x(t), x(t) \rangle \]
- \[ \langle x(t), y(t) \rangle = \frac{1}{2} \text{Re}\{\langle x_l(t), y_l(t) \rangle\} \] (Problem 2.2)

- **Cross-correlation coefficient** of two signals
  \[ \rho_{x,y} = \frac{\langle x(t), y(t) \rangle}{\sqrt{E_x E_y}} \]

- Two signals are **orthogonal** if \( \rho_{x,y} = 0 \)
  \[ \rho_{x,y} = \text{Re}(\rho_{x_l,y_l}) \]

- Orthogonal in baseband \( \rightarrow \) Orthogonal in pass band
Chapter 2.1-4 Lowpass Equivalent of a Bandpass System

- **Bandpass System**: transfer function $H(f)$ is located around $\pm f_0$
  
  $h(t) = \text{Re}\left\{ h_l(t)e^{j2\pi f_0 t} \right\}$

- $h_l(t)$ is the lowpass equivalent of $h(t)$

- If a bandpass signal $x(t)$ passes thru a bandpass system $h(t)$, the output $y(t)$ have spectrum

  $$Y(f) = X(f)H(f)$$

- The spectrum of lowpass equivalent of $y(t)$ is

  $$Y_l(f) = 2Y(f + f_0)u_{-1}(f + f_0)$$

  $$= 2X(f + f_0)H(f + f_0)u_{-1}(f + f_0)$$

  (using eq. 2.1-5)

  $$= \frac{1}{2}\left[ 2X(f + f_0)u_{-1}(f + f_0) \right]\left[ 2H(f + f_0)u_{-1}(f + f_0) \right]$$

  $$= \frac{1}{2} X_l(f)H_l(f)$$

  $y_l(t) = \frac{1}{2} x_l(t) * h_l(t)$
Chapter 2.2: Signal Space Representation of Waveforms
Let \( \mathbf{v}_1 = [v_{11}, v_{12}, \ldots, v_{1n}]' \) and \( \mathbf{v}_2 = [v_{21}, v_{22}, \ldots, v_{2n}]' \) be \( n \times 1 \) complex vector.

The *inner product* between \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \) is:

\[
\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \sum_{i=1}^{n} v_{1i} v_{2i}^* = \mathbf{v}_2^H \mathbf{v}_1
\]

- \( \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle^* = \sum_{i=1}^{n} (v_{2i} v_{1i}^*)^* = \sum_{i=1}^{n} v_{1i} v_{2i}^* \)
- \( \langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = 2 \text{Re}\{ \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \} \)

A vector can be represented as a linear combination of an *orthonormal basis* \( \{\mathbf{e}_i\} \),

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

where \( v_i = \langle \mathbf{v}, \mathbf{e}_i \rangle \) is the projection of \( \mathbf{v} \) onto \( \mathbf{e}_i \).

The *norm* of a \( n \)-dim vector \( \mathbf{v} \) is defined as:

\[
\|\mathbf{v}\| = (\langle \mathbf{v}, \mathbf{v} \rangle)^{1/2} = \sqrt{\sum_{i=1}^{n} |v_i|^2}
\]
Two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are *orthogonal* if
\[\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0\]

A set of $m$ vectors $\{\mathbf{v}_k\}$ are *orthogonal* if
\[\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0, \forall i \neq j\]

A set of $m$ vectors $\{\mathbf{v}_k\}$ are *orthonormal* if vectors are *orthogonal* and each has a *unit norm*.

A set of $m$ vectors $\{\mathbf{v}_k\}$ are *linearly independent* if no one vector is the linear combination of the remaining vectors.
Important Inequalities:

- **Triangle inequality:** \( \| \mathbf{v}_1 + \mathbf{v}_2 \| \leq \| \mathbf{v}_1 \| + \| \mathbf{v}_2 \| \)
  with equality if \( \mathbf{v}_1 = a \mathbf{v}_2, a > 0 \)

- **Cauchy-Schwarz inequality:** \( \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \leq \| \mathbf{v}_1 \| \cdot \| \mathbf{v}_2 \| \)
  with equality if \( \mathbf{v}_1 = a \mathbf{v}_2, \forall a \in \mathbb{C} \)

- \( \| \mathbf{v}_1 + \mathbf{v}_2 \|^2 = \| \mathbf{v}_1 \|^2 + \| \mathbf{v}_2 \|^2 + 2 \text{Re} \{ \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \} \)
  when \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \) are *orthogonal* \( \rightarrow \) Pythagorean relation

\[ \| \mathbf{v}_1 + \mathbf{v}_2 \|^2 = \| \mathbf{v}_1 \|^2 + \| \mathbf{v}_2 \|^2 \]

- The vector \( \mathbf{v} \) is an *eigen-vector* of a matrix \( \mathbf{A} \) if \( \mathbf{A} \mathbf{v} = \lambda \mathbf{v} \)
  where \( \lambda \) is the corresponding *eigen-value*
Gram-Schmidt Procedure

To construct a set of orthonormal vectors from a set of $n$-dimensional vectors $v_i$, $1 \leq i \leq m$

(1) Normalizing the length of first vector by

$$u_1 = v_1 / \|v_1\|$$

(2) Subtract the projection of $v_2$ on $v_1$,

$$u_2' = v_2 - (\langle v_2, u_1 \rangle)u_1$$

then normalize the length of $u_2'$ by $u_2 = u_2' / \|u_2'\|$

(3) Subtract the projection of $v_3$ on $v_1$ and $v_2$,

$$u_3' = v_3 - (\langle v_3, u_1 \rangle)u_1 - (\langle v_3, u_2 \rangle)u_2$$

then normalize the length of $u_2'$ by $u_3 = u_3' / \|u_3'\|$
Chapter 2.2-2 Signal Space Concepts

- The *inner product* of two complex signals is
  \[ \langle x_1(t), x_2(t) \rangle = \int_{-\infty}^{\infty} x_1(t)x_2^*(t)dt \]

- Two signals are *orthogonal* if \( \langle x_1(t), x_2(t) \rangle = 0 \)

- The *norm* of a signal is
  \[ \| x(t) \| = \left( \int_{-\infty}^{\infty} |x_1(t)|^2 dt \right)^{1/2} = \sqrt{E_x} \]

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

- A set of \( m \) signals is *linearly independent* if
  no signal is a linear combination of the remaining signals
Chapter 2.2-2 Signal Space Concepts

◊ **Triangle inequality** : \[ \|x_1(t) + x_2(t)\| \leq \|x_1(t)\| + \|x_2(t)\| \]

◊ **Cauchy-Schwarz inequality** : \[ \left| \langle x_1(t), x_2(t) \rangle \right| \leq \|x_1(t)\| \cdot \|x_2(t)\| = \sqrt{E_{x_1} E_{x_2}} \]

or \[ \left| \int_{-\infty}^{\infty} x_1(t)x_2^*(t)dt \right| \leq \left[ \int_{-\infty}^{\infty} |x_1(t)|^2 dt \right]^{1/2} \left[ \int_{-\infty}^{\infty} |x_2(t)|^2 dt \right]^{1/2} \]

with equality if \( x_2(t) = ax_1(t) \), \( a \) is any complex constant
Suppose that $s(t)$ is deterministic with finite energy

$$E_s = \int_{-\infty}^{\infty} |s(t)|^2 \, dt < \infty$$

Suppose existing a set of orthonormal functions $\phi_1(t), \ldots, \phi_K(t)$

$$\langle \phi_n(t), \phi_m(t) \rangle = \int_{-\infty}^{\infty} \phi_n(t)\phi_m^*(t) \, dt = \begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases}$$

The signal $s(t)$ may be approximated as a weighted linear combination of these functions:

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

Approximation error:

$$e(t) = s(t) - \hat{s}(t)$$

Select $\{s_k\}$ to minimize energy of approximation error

$$E_e = \int_{-\infty}^{\infty} |s(t) - \hat{s}(t)|^2 \, dt = \int_{-\infty}^{\infty} \left| s(t) - \sum_{k=1}^{K} s_k \phi_k(t) \right|^2 \, dt$$
From estimation theory, the minimum of $E_e$ is achieved when error is orthogonal to the set of functions, i.e.,

$$
\int_{-\infty}^{\infty} \left[ s(t) - \sum_{k=1}^{K} s_k \phi_k(t) \right] \phi_n^*(t) dt = 0 \quad n=1,2,..,K
$$

Since $\{\phi_n(t)\}$ are orthonormal, the above equation reduces to:

$$
\phi_n^*(t) \delta(t-t_k) = \int_{-\infty}^{\infty} s(t) \phi_n^*(t) dt \quad n=1,2,..,K
$$

Because $\hat{s}(t)$ is a linear combination of $\{\phi_n(t)\}$

The minimum mean square approximation error is

$$
E_{e,\text{min}} = \int_{-\infty}^{\infty} e(t) s^*(t) dt
$$

$$
= \int_{-\infty}^{\infty} |s(t)|^2 dt - \int_{-\infty}^{\infty} \sum_{k=1}^{K} s_k \phi_k(t) s^*(t) dt
$$

$$
= E_S - \sum_{k=1}^{K} |s_k|^2
$$

$E_{e,\text{min}} = 0$ if $E_S = \sum_{k=1}^{K} |s_k|^2$ 

$\{\phi_n(t)\}$ is complete
Chapter 2.2-3 Orthogonal Expansions of Signals

[Example]

Consider a finite energy real signal \( s(t) \), and \( s(t) = 0 \) when \( t \not\in [0, T] \). Its periodic expansion can be represented in a Fourier series as:

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

where \( \{a_k, b_k\} \) that minimize MSE are given by:

\[
a_k = \frac{2}{T} \int_0^T s(t) \cos \frac{2\pi kt}{T} \, dt
\]

\[
b_k = \frac{2}{T} \int_0^T s(t) \sin \frac{2\pi kt}{T} \, dt
\]

Since \( \left[ \sqrt{\frac{1}{T}}, \sqrt{\frac{2}{T}} \cos \frac{2\pi kt}{T}, \sqrt{\frac{2}{T}} \sin \frac{2\pi kt}{T} \right] \) is a complete set for the expansion of periodic signals on \([0,T]\),

The series expansion has zero mean-square-error.
Chapter 2.2-4 Gram-Schmidt Procedures

- \{s_m(t), m=1,2,\ldots,M\}: set of finite energy signal waveforms

- To construct a set of orthonormal waveforms from \{s_m(t)\}

- The first orthonormal waveform is \( \phi_1(t) = \frac{s_1(t)}{\sqrt{E_1}} \)

- The projection of \( s_2(t) \) on \( f_1(t) \) is \( c_{21} = \langle s_2(t), \phi_1(t) \rangle = \int_{-\infty}^{\infty} s_1(t)\phi_1^*(t)dt \)

- Subtract the projection of \( s_2(t) \) on \( f_1(t) \): \( \gamma_2(t) = s_2(t) - c_{21}\phi_1(t) \)

- The second orthonormal waveform is \( \phi_2(t) = \frac{\gamma_2(t)}{\sqrt{E_2}} \)
  where \( E_2 = \int_{-\infty}^{\infty} |\gamma_2(t)|^2 dt \)
In general, the orthogonalization of the $k$-th function is

$$\phi_k(t) = \gamma_k(t) / \sqrt{E_k}$$

where:

$$\gamma_k(t) = s_k(t) - \sum_{i=1}^{k-1} c_{ki} \phi_i(t)$$

$$c_{ki} = \langle s_k(t), \phi_i(t) \rangle = \int_{-\infty}^{\infty} s_k(t) \phi_i^*(t) dt, \quad i = 1, 2, \ldots, k - 1$$

$$E_k = \int_{-\infty}^{\infty} |\gamma_k(t)|^2 dt$$

The orthogonalization process continues until $\{s_m(t)\}$ are exhausted, and $N \leq M$ orthonormal waveforms are obtained.

N=M if $\{s_m(t)\}$ are linearly independent.
Probability and Random Variables

Wireless Information Transmission System Lab.
Institute of Communications Engineering
National Sun Yat-sen University
Sample space or certain event of a die experiment:

\[ S = \{1, 2, 3, 4, 5, 6\} \]

The six outcomes are the sample points of the experiment.

An event is a subset of \( S \), and may consist of any number of sample points. For example:

\[ A = \{2, 4\} \]

The complement of the event \( A \), denoted by \( \overline{A} \), consists of all the sample points in \( S \) that are not in \( A \):

\[ \overline{A} = \{1, 3, 5, 6\} \]
Two events are said to be *mutually exclusive* if they have no sample points in common – that is, if the occurrence of one event excludes the occurrence of the other. For example:

\[ A = \{2, 4\}; \quad B = \{1, 3, 6\} \]

\( A \) and \( \overline{A} \) are mutually exclusive events.

The *union* (sum) of two events in an event that consists of all the sample points in the two events. For example:

\[ C = \{1, 2, 3\} \]

\[ D = B \cup C = \{1, 2, 3, 6\} \]

\[ A \cup \overline{A} = S \]
The *intersection* of two events is an event that consists of the points that are common to the two events. For example:

\[ E = B \cap C = \{1,3\} \]

When the events are mutually exclusive, the intersection is the *null event*, denoted as \( \phi \). For example:

\[ A \cap \overline{A} = \phi \]
Associated with each event $A$ contained in $S$ is its probability $P(A)$.

Three postulations:
- $P(A) \geq 0$.
- The probability of the sample space is $P(S) = 1$.
- Suppose that $A_i$, $i = 1, 2, \ldots$, are a (possibly infinite) number of events in the sample space $S$ such that
  $$A_i \cap A_j = \emptyset; \quad i \neq j = 1, 2, \ldots$$
  Then the probability of the union of these mutually exclusive events satisfies the condition:
  $$P\left( \bigcup_i A_i \right) = \sum_i P(A_i)$$


"Joint events" and "joint probabilities" (two experiments)

- If one experiment has the possible outcomes $A_i$, $i = 1,2,\ldots,n$, and the second experiment has the possible outcomes $B_j$, $j = 1,2,\ldots,m$, then the combined experiment has the possible joint outcomes $(A_i, B_j)$, $i = 1,2,\ldots,n$, $j = 1,2,\ldots,m$.

- Associated with each joint outcome $(A_i, B_j)$ is the joint probability $P(A_i, B_j)$ which satisfies the condition:

$$0 \leq P(A_i, B_j) \leq 1$$

- Assuming that the outcomes $B_j$, $j = 1,2,\ldots,m$, are mutually exclusive, it follows that:

$$\sum_{j=1}^{m} P(A_i, B_j) = P(A_i)$$

- If all the outcomes of the two experiments are mutually exclusive, then:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P(A_i, B_j) = \sum_{i=1}^{n} P(A_i) = 1$$
Conditional probabilities

The conditional probability of the event \( A \) given the occurrence of the event \( B \) is defined as:

\[
P(A \mid B) = \frac{P(A, B)}{P(B)}
\]

provided \( P(B) > 0 \).

- \( P(A, B) = P(A \mid B)P(B) = P(B \mid A)P(A) \)
- \( P(A, B) \) is interpreted as the probability of \( A \cap B \). That is, \( P(A, B) \) denotes the simultaneous occurrence of \( A \) and \( B \).
- If two events \( A \) and \( B \) are mutually exclusive, \( A \cap B = \phi \), then \( P(A \mid B) = 0 \).
- If \( B \) is a subset of \( A \), we have \( A \cap B = B \) and \( P(A \mid B) = 1 \).
Probability

◊ **Bayes’ theorem:**

◊ If \( A_i, \ i = 1, 2, \ldots, n, \) are mutually exclusive events such that

\[
\bigcup_{i=1}^{n} A_i = S
\]

and \( B \) is an arbitrary event with nonzero probability, then

\[
P(A_i \mid B) = \frac{P(A_i, B)}{P(B)} = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)P(A_j)}
\]

\( P(B) = \sum_{j=1}^{n} P(B, A_j) = \sum_{j=1}^{n} P(B \mid A_j)P(A_j) \)

◊ \( P(A_i) \) represents their *a priori probabilities* and \( P(A_i \mid B) \) is the *a posteriori probability* of \( A_i \) conditioned on having observed the received signal \( B \).
Statistical independence

- If the occurrence of $A$ does not depend on the occurrence of $B$, then $P(A \mid B) = P(A)$.
- $P(A, B) = P(A \mid B)P(B) = P(A)P(B)$
- When the events $A$ and $B$ satisfy the relation $P(A, B) = P(A)P(B)$, they are said to be statistically independent.

Three statistically independent events $A_1, A_2,$ and $A_3$ must satisfy the following conditions:

\[
P(A_1, A_2) = P(A_1)P(A_2)
\]
\[
P(A_1, A_3) = P(A_1)P(A_3)
\]
\[
P(A_2, A_3) = P(A_2)P(A_3)
\]
\[
P(A_1, A_2, A_3) = P(A_1)P(A_2)P(A_3)
\]
Random Variables, Probability Distributions, and Probability Densities

- Given an experiment having a sample space \( S \) and elements \( s \in S \), we define a function \( X(s) \) whose domain is \( S \) and whose range is a set of numbers on the real line.

- The function \( X(s) \) is called a random variable.

- Example 1: If we flip a coin, the possible outcomes are head (H) and tail (T), so \( S \) contains two points labeled H and T. Suppose we define a function \( X(s) \) such that:
  \[
  X(s) = \begin{cases} 
  +1 & (s = H) \\ 
  -1 & (s = T)
  \end{cases}
  \]
  Thus we have mapped the two possible outcomes of the coin-flipping experiment into the two points (+1,-1) on the real line.

- Example 2: Tossing a die with possible outcomes \( S = \{1,2,3,4,5,6\} \). A random variable defined on this sample space may be \( X(s) = s \), in which case the outcomes of the experiment are mapped into the integers 1,…,6, or, perhaps, \( X(s) = s^2 \), in which case the possible outcomes are mapped into the integers \{1,4,9,16,25,36\}.
Give a random variable $X$, let us consider the event \( \{ X \leq x \} \) where 

\( x \) is any real number in the interval \((-\infty, \infty)\). We write the 

probability of this event as \( P(X \leq x) \) and denote it simply by \( F(x) \), i.e.,

\[
F(x) = P(X \leq x), \quad -\infty < x < \infty
\]

The function \( F(x) \) is called the \textit{probability distribution function} of the random variable \( X \).

It is also called the \textit{cumulative distribution function (CDF)}.

\[
0 \leq F(x) \leq 1
\]

\[
F(-\infty) = 0 \text{ and } F(\infty) = 1.
\]
The derivative of the CDF $F(x)$, denoted as $p(x)$, is called the *probability density function (PDF)* of the random variable $X$.

$$p(x) = \frac{dF(x)}{dx}, \quad -\infty < x < \infty$$

$$F(x) = \int_{-\infty}^{x} p(u)du, \quad -\infty < x < \infty$$

When the random variable is discrete, the derivative of the CDF is called the *probability mass function (PMF)*.
Determining the probability that a random variable $X$ falls in an interval $(x_1, x_2)$, where $x_2 > x_1$.

\[ P(X \leq x_2) = P(X \leq x_1) + P(x_1 < X \leq x_2) \]

\[ F(x_2) = F(x_1) + P(x_1 < X \leq x_2) \]

\[ \Rightarrow P(x_1 < X \leq x_2) = F(x_2) - F(x_1) \]

\[ = \int_{x_1}^{x_2} p(x)\,dx \]

The probability of the event $\{x_1 < X \leq x_2\}$ is simply the area under the PDF in the range $x_1 < X \leq x_2$. 

48
Random Variables, Probability Distributions, and Probability Densities

- Multiple random variables, joint probability distributions, and joint probability densities: (two random variables)

Joint CDF: \( F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} p(u_1, u_2) du_1 du_2 \)

Joint PDF: \( p(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2) \)

\[ \int_{-\infty}^{\infty} p(x_1, x_2) dx_1 = p(x_2) \quad \int_{-\infty}^{\infty} p(x_1, x_2) dx_2 = p(x_1) \]

The PDFs \( p(x_1) \) and \( p(x_2) \) obtained from integrating over one of the variables are called *marginal* PDFs.

\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, x_2) dx_1 dx_2 = F(\infty, \infty) = 1 \]

Note that: \( F(-\infty, -\infty) = F(-\infty, x_2) = F(x_1, -\infty) = 0. \)
Random Variables, Probability Distributions, and Probability Densities

◊ Statistically independent random variables:
  ◊ If the experiments result in mutually exclusive outcomes, the probability of an outcome is independent of an outcome in any other experiment.
  ◊ The joint probability of the outcomes factors into a product of the probabilities corresponding to each outcome.
  ◊ Consequently, the random variables corresponding to the outcomes in these experiments are independent in the sense that their joint PDF factors into a product of marginal PDFs.
  ◊ The multidimensional random variables are statistically independent if and only if:
    \[ F(x_1, x_2, \ldots, x_n) = F(x_1)F(x_2)\ldots F(x_n) \]
    \[ p(x_1, x_2, \ldots, x_n) = p(x_1)p(x_2)\ldots p(x_n) \]
Problem: given a random variable $X$, which is characterized by its PDF $p(x)$, determine the PDF of the random variable $Y=g(X)$, where $g(X)$ is some given function of $X$.

- When the mapping $g$ from $X$ to $Y$ is one-to-one, the determination of $p(y)$ is relatively straightforward.

- However, when the mapping is not one-to-one, as in the case, for example, when $Y=X^2$, we must be very careful in our derivation of $p(y)$.
Functions of Random Variables

- Transformation of a random variable
Example 2.1-1:

- $Y = aX + b$; where $a$ and $b$ are constant and assume that $a > 0$.
- $F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P(X \leq \frac{y - b}{a})$

\[
F_Y(y) = \int_{-\infty}^{(y-b)/a} p_X(x) \, dx = F_X\left(\frac{y - b}{a}\right)
\]

$p_Y(y) = \frac{1}{a} p_X\left(\frac{y - b}{a}\right)$

($F_Y(y)$ 对 $y$ 作微分)
Example 2.1-1 (cont.):

\[ Y = aX + b; \text{ where } a \text{ and } b \text{ are constant and assume that } a > 0. \]
Example 2.1-2:

- \( Y = aX^3 + b; \ a > 0. \)
- The mapping between \( X \) and \( Y \) is one-to-one.
- \( F_Y(y) = P(Y \leq y) = P(aX^3 + b \leq y) \)
  \[
  = P \left[ X \leq \left( \frac{y-b}{a} \right)^{1/3} \right] = F_X \left( \left( \frac{y-b}{a} \right)^{1/3} \right)
  \]
- Differentiation with respect to \( y \) yields the desired relationship between the two PDFs
  \[
p_Y(y) = \frac{1}{3a \left[ (y-b)/a \right]^{2/3}} p_X \left( \left( \frac{y-b}{a} \right)^{1/3} \right)
  \]
Example 2.1-3:

- $Y=ax^2+b$; $a>0$.
- The mapping between $X$ and $Y$ is not one-to-one.
Functions of Random Variables

Example 2.1-3 (cont.):

- $Y = aX^2 + b; \ a > 0.$

$$F_Y(y) = P(Y \leq y) = P(aX^2 + b \leq y) = P\left(|X| \leq \frac{y-b}{a}\right)$$

$$F_Y(y) = F_X\left(\sqrt{\frac{y-b}{a}}\right) - F_X\left(-\sqrt{\frac{y-b}{a}}\right)$$

- Differentiating with respect to $y$, we obtain the PDF of $Y$ in terms of the PDF of $X$ in the form:

$$p_Y(y) = \frac{p_X\left[\sqrt{(y-b)/a}\right]}{2a\left[\sqrt{(y-b)/a}\right]} + \frac{p_X\left[-\sqrt{(y-b)/a}\right]}{2a\left[\sqrt{(y-b)/a}\right]}$$
Example 2.1-3 (cont.):

- $Y = g(X) = aX^2 + b; \ a > 0$ has two solutions:

  \[ x_1 = \sqrt{\frac{y - b}{a}}, \quad x_2 = -\sqrt{\frac{y - b}{a}} \]

- $p_Y(y)$ consists of two terms corresponding to these two solutions:

  \[
p_Y(y) = p_x \left[ x_1 = \sqrt{(y - b) / a} \right] \cdot \frac{1}{g'(x_1)} + p_x \left[ x_1 = -\sqrt{(y - b) / a} \right] \cdot \frac{1}{g'(x_1)}
  \]

- In general, suppose that $x_1, x_2, \ldots, x_n$ are the real roots of $g(x) = y$. The PDF of the random variable $Y = g(X)$ may be expressed as

  \[
p_Y(y) = \sum_{i=1}^{n} \frac{p_X(x_i)}{|g'(x_i)|} \quad \text{Why?}
  \]
Function of one random variable (discrete case):

- Suppose $X$ is a discrete random variable that can have one of $n$ values $x_1, x_2, \ldots, x_n$.
- Let $g(X)$ be a scalar-valued function.
- $Y = g(X)$ is a discrete random variable that can have one of $m$, $m \leq n$, values $y_1, y_2, \ldots, y_m$.
- If $g(X)$ is a one-to-one mapping, then $m$ will be equal to $n$.
- If $g(X)$ is many-to-one, then $m$ will be smaller than $n$.
- The CDF of $Y$ can be obtained easily from the CDF function of $X$ as:
  \[ P(Y = y_i) = \sum P(X = x_i) \]
  where the sum is over all values of $x_i$ that map to $y_i$. 
Function of one random variable (continuous case):

- If $X$ is a continuous random variable, then the pdf of $Y=g(X)$ can be obtained from the pdf of $X$.
- Let $y$ be a particular value of $Y$ and let $x^{(1)}$, $x^{(2)}$, ..., $x^{(n)}$ be roots of the equation $y=g(x)$. That is $y=g(x^{(1)})=g(x^{(2)})= ... = y=g(x^{(n)})$.
- $P(y < Y \leq y + \Delta y) = f_Y(y)\Delta y$ as $\Delta y \to 0$
  
  $= P[\{x : y < g(x) \leq y + \Delta y\}]$
Function of one random variable (continuous case):
Functions of Random Variables

◊ Function of one random variable (continuous case):

\[ P(y < Y < y + \Delta y) = P(x^{(1)} < X < x^{(1)} + \Delta x^{(1)}) + P(x^{(2)} - \Delta x^{(2)} < X < x^{(2)}) \]
\[ + P(x^{(3)} < X < x^{(3)} + \Delta x^{(3)}) \]
\[ = f_X(x^{(1)})\Delta x^{(1)} + f_X(x^{(2)})|\Delta x^{(2)}| + f_X(x^{(3)})\Delta x^{(3)} = f_Y(y)\Delta y \]

Since the slope \( g'(x) \) of \( g(x) \) is \( \Delta y / \Delta x \), we have

\[ \Delta x^{(1)} = \frac{\Delta y}{g'(x^{(1)})} \quad \Delta x^{(2)} = \frac{\Delta y}{g'(x^{(2)})} \quad \Delta x^{(3)} = \frac{\Delta y}{g'(x^{(3)})} \]

\[ f_Y(y)\Delta y = \frac{f_X(x^{(1)})}{g'(x^{(1)})}\Delta y + \frac{f_X(x^{(2)})}{|g'(x^{(2)})|}\Delta y + \frac{f_X(x^{(3)})}{g'(x^{(3)})}\Delta y \]

\[ \Rightarrow f_Y(y) = \sum_{i=1}^{k} \frac{f_X(x^{(i)})}{|g'(x^{(i)})|} \]
Functions of Random Variables

◊ Function of several random variables (continuous case):
  ◦ Goal: find the joint distribution of \( n \) random variables \( Y_1, Y_2, \ldots, Y_n \) given the distribution of \( n \) related random variables, \( X_1, X_2, \ldots, X_n \), where

\[
Y_i = g_i(X_1, X_2, \ldots, X_n), \quad i = 1, 2, \ldots, n
\]

◊ Let’s start with a mapping of two random variables onto two other random variables:

\[
Y_1 = g_1(X_1, X_2) \quad Y_2 = g_2(X_1, X_2)
\]

◊ Suppose \( (x_1^{(i)}, x_2^{(i)}), \ i = 1, 2, \ldots, k \) are the \( k \) roots of \( y_1 = g_1(x_1, x_2) \) and \( y_2 = g_2(x_1, x_2) \). We need to find the region in the \( x_1, x_2 \) plane such that

\[
y_1 < g_1(x_1, x_2) < y_1 + \Delta y_1 \text{ and } y_2 < g_2(x_1, x_2) < y_2 + \Delta y_2
\]
Function of several random variables (continuous case):

There are $k$ such regions as shown in the following figure for the case of $k=3$. 
Function of several random variables (continuous case):

- Each region consists of a parallelogram and the area of each parallelogram is equal to \( \frac{\Delta y_1 \Delta y_2}{|J(x_1^{(i)}, x_2^{(i)})|} \) where \( J(x_1, x_2) \) is the Jacobian of the transformation defined as

\[
J(x_1, x_2) = \begin{vmatrix}
\frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\
\frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2}
\end{vmatrix}
\]

By summing the contribution from all regions, we obtain the joint pdf of \( Y_1 \) and \( Y_2 \) as

\[
f_{Y_1, Y_2}(y_1, y_2) = \sum_{i=1}^{k} \frac{f_{X_1, X_2}(x_1^{(i)}, x_2^{(i)})}{|J(x_1^{(i)}, x_2^{(i)})|}
\]
Functions of Random Variables

- Function of several random variables (continuous case):
  - We can generalize this result to the $n$-variate case as
    
    $$f_Y(y_1, y_2, \ldots, y_n) = \sum_{i=1}^{k} f_X\left(x_1^{(i)} = g_1^{-1}, x_2^{(i)} = g_2^{-1}, \ldots, x_n^{(i)} = g_n^{-1}\right) \left|J^{(i)}\right|$$

    where $J^{(i)}$ denotes the Jacobian of the transformation defined by the following determinant when evaluating at the $i$-th solution $x_1^{(i)} = g_1^{-1}, x_2^{(i)} = g_2^{-1}, \ldots, x_n^{(i)} = g_n^{-1}$ of $y = g(x) = [g_1(x), g_2(x), \ldots, g_n(x)]$:

    $$J^{(i)} = \begin{vmatrix}
    \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\
    \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\
    \vdots & \vdots & \ddots & \vdots \\
    \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_n}
    \end{vmatrix}$$

    at $x^{(i)}_1 = g_1^{-1}, x^{(i)}_2 = g_2^{-1}, \ldots, x^{(i)}_n = g_n^{-1}$.
Example 2.1-4

\[ Y_i = \sum_{j=1}^{n} a_{ij}X_j, \quad i = 1, 2, \ldots, n \]

\[ Y = AX \quad \text{where} \ A \text{ is an} \ n \times n \text{ matrix.} \]

\[ X = A^{-1}Y \quad \Rightarrow \quad X_i = \sum_{j=1}^{n} b_{ij}Y_j, \quad i = 1, 2, \ldots, n \]

\[ \{b_{ij}\} \text{are the elements of the inverse matrix} \ A^{-1}. \]

The Jacobian of this transformation is \( J = \det A. \)

\[ p_Y(y_1, y_2, \ldots, y_n) \]

\[ = p_X \left( x_1 = \sum_{j=1}^{n} b_{1j} \cdot y_j, x_2 = \sum_{j=1}^{n} b_{2j} \cdot y_j, \ldots, x_n = \sum_{j=1}^{n} b_{nj} \cdot y_j \right) \frac{1}{|\det A|} \]
Functions of Random Variables

- The \textit{mean} or \textit{expected value} of $X$, which characterized by its PDF $p(x)$, is defined as:
  \[ E(X) \equiv m_x = \int_{-\infty}^{\infty} xp(x)dx \]
  This is the \textit{first moment} of random variable $X$.

- The \textit{n-th moment} is defined as:
  \[ E(X^n) = \int_{-\infty}^{\infty} x^n p(x)dx \]

- Define $Y=g(X)$, the expected value of $Y$ is:
  \[ E(Y) = E[g(X)] = \int_{-\infty}^{\infty} g(x) p(x)dx \]
The *n*-th central moment of the random variable $X$ is:

$$E(Y) = E[(X - m_x)^n] = \int_{-\infty}^{\infty} (x - m_x)^n \, p(x) \, dx$$

When $n=2$, the central moment is called the *variance* of the random variable and denoted as $\sigma_x^2$:

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 \, p(x) \, dx$$

$$\sigma_x^2 = E(X^2) - [E(X)]^2 = E(X^2) - m_x^2$$

In the case of two random variables, $X_1$ and $X_2$, with joint PDF $p(x_1, x_2)$, we define the *joint moment* as:

$$E(X_1^k X_2^n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^k x_2^n \, p(x_1, x_2) \, dx_1 \, dx_2$$
The joint central moment is defined as:

\[
E[(X_1 - m_1)^k (X_2 - m_2)^n]
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - m_1)^k (x_2 - m_2)^n p(x_1, x_2) dx_1 dx_2
\]

If \(k=n=1\), the joint moment and joint central moment are called the correlation and the covariance of the random variables \(X_1\) and \(X_2\), respectively.

The correlation between \(X_i\) and \(X_j\) is given by the joint moment:

\[
E(X_i X_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_i x_j p(x_i, x_j) dx_i dx_j
\]
The covariance between $X_i$ and $X_j$ is given by the joint central moment:

$$\mu_{ij} = E[(X_i - m_i)(X_j - m_j)]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - m_i)(x_j - m_j)p(x_i, x_j)dx_i dx_j$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i x_j - x_j m_i - x_i m_j + m_i m_j)p(x_i, x_j)dx_i dx_j$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_i x_j p(x_i, x_j)dx_i dx_j - m_i m_j$$

$$= E(X_i X_j) - m_i m_j$$

The $n \times n$ matrix with elements $\mu_{ij}$ is called the covariance matrix of the random variables, $X_i$, $i=1,2, \ldots, n$. 
Two random variables are said to be *uncorrelated* if
\[ E(X_iX_j) = E(X_i)E(X_j) = m_i m_j. \]
Uncorrelated \(\rightarrow\) Covariance \(\mu_{ij} = 0.\)
If \(X_i\) and \(X_j\) are statistically independent, they are uncorrelated.
If \(X_i\) and \(X_j\) are uncorrelated, they are *not necessary* statistically independently.
Two random variables are said to be *orthogonal* if
\[ E(X_iX_j) = 0. \]
Two random variables are orthogonal if they are uncorrelated and either one or both of them have zero mean.
 Characteristic functions

- The characteristic function of a random variable $X$ is defined as the statistical average:

$$E(e^{jvX}) \equiv \psi(jv) = \int_{-\infty}^{\infty} e^{jvx} p(x) dx$$

- $\psi(jv)$ may be described as the Fourier transform of $p(x)$.

- The inverse Fourier transform is:

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(jv)e^{-jvx} dv$$

- First derivative of the above equation with respect to $v$:

$$\frac{d\psi(jv)}{dv} = j\int_{-\infty}^{\infty} xe^{jvx} p(x) dx$$
◊ Characteristic functions (cont.)

◊ First moment (mean) can be obtained by:

\[ E(X) = m_x = -j \frac{d\psi(jv)}{dv} \bigg|_{v=0} \]

◊ Since the differentiation process can be repeated, \( n \)-th moment can be calculated by:

\[ E(X^n) = (-j)^n \frac{d^n\psi(jv)}{dv^n} \bigg|_{v=0} \]

◊ Suppose the characteristic function can be expanded in a Taylor series about the point \( v=0 \):

\[ \psi(jv) = \sum_{n=0}^{\infty} \left[ \frac{d^n\psi(jv)}{dv^n} \right]_{v=0} \frac{v^n}{n!} \Rightarrow \psi(jv) = \sum_{n=0}^{\infty} E(X^n) \frac{(jv)^n}{n!} \]
Characteristic functions (cont.)

- Determining the PDF of a sum of statistically independent random variables:

\[ Y = \sum_{i=1}^{n} X_i \implies \psi_Y(jv) = E(e^{jvY}) = E\left[\exp\left(jv\sum_{i=1}^{n} X_i\right)\right] \]

\[ = E\left[\prod_{i=1}^{n} (e^{jvX_i})\right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{i=1}^{n} e^{jvX_i}\right) p(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n \]

Since the random variables are statistically independent,

\[ p(x_1, x_2, \ldots, x_n) = p(x_1) p(x_2) \ldots p(x_n) \implies \psi_Y(jv) = \prod_{i=1}^{n} \psi_{X_i}(jv) \]

If \( X_i \) are iid (independent and identically distributed)

\[ \implies \psi_Y(jv) = [\psi_X(jv)]^n \]
Characteristic functions (cont.)

The PDF of $Y$ is determined from the inverse Fourier transform of $\Psi_Y(j\nu)$.

Since the characteristic function of the sum of $n$ statistically independent random variables is equal to the product of the characteristic functions of the individual random variables, it follows that, in the transform domain, the PDF of $Y$ is the $n$-fold convolution of the PDFs of the $X_i$.

Usually, the $n$-fold convolution is more difficult to perform than the characteristic function method in determining the PDF of $Y$. 
Characteristic functions (for \( n \)-dimensional random variables)

If \( X_i, i=1,2,\ldots,n \), are random variables with PDF \( p(x_1,x_2,\ldots,x_n) \), the \( n \)-dimensional characteristic function is defined as:

\[
\psi(j\nu_1,j\nu_2,\ldots,j\nu_n) = E \left[ \exp \left( j \sum_{i=1}^{n} \nu_i X_i \right) \right] \\
= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left( j \sum_{i=1}^{n} \nu_i x_i \right) p(x_1,x_2,\ldots,x_n) dx_1 dx_2 \cdots dx_n
\]

For two dimensional characteristic function:

\[
\psi(j\nu_1,j\nu_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j(\nu_1 x_1 + \nu_2 x_2)} p(x_1,x_2) dx_1 dx_2
\]

\[
E(X_1X_2) = -\left. \frac{\partial^2 \psi(j\nu_1,j\nu_2)}{\partial \nu_1 \partial \nu_2} \right|_{\nu_1=\nu_2=0}
\]
Chapter 2.3: Some Useful Random Variables

Wireless Information Transmission System Lab.
Institute of Communications Engineering
National Sun Yat-sen University
Some Useful Random Variables

◊ In this section we list some of the frequently encountered random variables, their probability density functions (PDFs), their cumulative distribution functions (CDFs), and their moments.

◊ Our main emphasis will be on the Gaussian random variable and many random variables that are derived from the Gaussian random variable.
The Bernoulli Random Variable

- The Bernoulli random variable is a discrete binary-valued random variable taking values 1 and 0 with probabilities $p$ and $1 - p$, respectively.
- The probability mass function (PMF) for this random variable is given by
  \[ P[X = 1] = p \]
  \[ P[X = 0] = 1 - p \]
- The mean and variance are given by
  \[ E[X] = p \]
  \[ \text{var}[X] = p(1 - p) \]
The Binomial Random Variable

- The binomial random variable models the sum of \( n \) independent Bernoulli random variables with common parameter \( p \).
- The PMF of this random variable is given by
  \[
  P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \ldots, n
  \]
- The mean and variance are given by
  \[
  E[X] = np \\
  \text{var}[X] = np(1 - p)
  \]
The uniform random variable is a continuous random variable with PDF

\[ p(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \]

where \( b > a \) and the interval \([a, b]\) is the range of the random variable.

The mean and variance are given by

\[ E[X] = \frac{b + a}{2} \]

\[ \text{var}[X] = \frac{(b - a)^2}{12} \]
Also known as *Normal* random variable.

The Gaussian random variable is described in terms of two parameters $m \in \mathbb{R}$ and $\sigma > 0$ by the PDF

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

The mean and variance are given by

$$E[X] = m$$

$$\text{var}[X] = \sigma^2$$

The Gaussian random variable is denoted by $X \sim \mathcal{N}(m, \sigma^2)$

When $m = 0, \sigma = 1 \rightarrow \mathcal{N}(0, 1)$ is a *standard normal*. 
The Gaussian Random Variable

**Q-function**: 

\[ Q(x) = P[\mathcal{N}(0,1) > x] = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt \quad x \geq 0 \]

The CDF of a Gaussian random variable is given by

\[ F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-m)^2}{2\sigma^2}} dt \]

\[ = 1 - \int_x^\infty \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-m)^2}{2\sigma^2}} dt \]

\[ = 1 - \int_{\frac{x-m}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \]

\[ = 1 - Q \left( \frac{x-m}{\sigma} \right) \]
In general if $X \sim \mathcal{N}(m, \sigma^2)$, then
\[
P[X > \alpha] = Q\left(\frac{\alpha - m}{\sigma}\right)
\]
\[
P[X < \alpha] = Q\left(\frac{m - \alpha}{\sigma}\right)
\]

Some of the important properties of the Q function:
\[
Q(0) = \frac{1}{2} \quad Q(\infty) = 0
\]
\[
Q(-\infty) = 1 \quad Q(-x) = 1 - Q(x)
\]
Some useful bounds for the Q function for $x > 0$ are:

- $Q(x) \leq \frac{1}{2} e^{-\frac{x^2}{2}}$
- $Q(x) < \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- $Q(x) > \frac{x}{(1+x^2)\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

For large $x$ we have

$Q(x) \approx \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

for large $x$ since $1+x^2 \approx x^2$
The complementary error function is defined as
\[ \text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \]

Relation with the Q function as follows:
\[ Q(x) = \frac{1}{2} \text{erfc}\left( \frac{x}{\sqrt{2}} \right) \]
\[ \text{erfc}(x) = 2Q(\sqrt{2}x) \]

The characteristic function of a Gaussian random variable is given by
\[ \Phi_X(\omega) = E\left[ e^{i\omega X} \right] = \int_{-\infty}^{\infty} e^{i\omega x} \left[ \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma^2}} \right] dx = e^{i\omega m - \frac{1}{2}\omega^2 \sigma^2} \]
For an $N(m, \sigma^2)$ random variable, we have

$$E[(X - m)^2] = \begin{cases} 
1 \times 3 \times 5 \times \cdots \times (2k - 1) \sigma^{2k} & \text{for } n = 2k \\
0 & \text{for } n = 2k + 1
\end{cases}$$

from which we can obtain moments of the Gaussian random variable.

The sum of $n$ independent Gaussian random variables is a Gaussian random variable whose mean and variance are the sum of the means and the sum of the variances of the ransom variables, respectively.
The Chi-Square ($\chi^2$) Random Variable

- If $\{X_i, i = 1, \ldots, n\}$ are iid (independent and identically distributed) zero-mean Gaussian random variables with common variance $\sigma^2$ and we define

$$X = \sum_{i=1}^{n} X_i^2$$

then $X$ is a Chi-Square ($\chi^2$) random variable with $n$ degrees of freedom.
The Chi-Square ($\chi^2$) Random Variable

◇ The PDF of Chi-Square random variable is given by

$$p(x) = \begin{cases} 
\frac{1}{2^{n/2} \Gamma\left(\frac{n}{2}\right) \sigma^n} x^{n/2-1} e^{-x/2\sigma^2} & x > 0 \\
0 & \text{otherwise}
\end{cases}$$

where $\Gamma(x)$ is the *Gamma function* defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad t \geq 0$$

◇ The mean and variance of $X$ is given by

$$E[X] = n\sigma^2$$
$$\text{var}[X] = 2n\sigma^4$$
Some properties of Gamma function:

\[ \Gamma(x + 1) = x \Gamma(x) \quad \Gamma(1) = 1 \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \]

\[ \Gamma\left(\frac{n}{2} + 1\right) = \begin{cases} 
\left(\frac{n}{2}\right)! & n \text{ even and positive} \\
\sqrt{\pi} \frac{n(n - 2)(n - 4) \cdots 3 \times 1}{n + 1 \choose 2} & n \text{ odd and positive}
\end{cases} \]

When \( n = 2m \), the CDF of the Chi-Square random variable with \( n \) degrees of freedom has a closed form:

\[ F(x) = \begin{cases} 
1 - e^{-\frac{x}{2\sigma^2}} \sum_{k=0}^{m-1} \frac{1}{k!} \left(\frac{x}{2\sigma^2}\right)^k & x > 0 \\
0 & \text{otherwise}
\end{cases} \]
The Chi-Square ($\chi^2$) Random Variable

- The characteristic function of Chi-Square random variable with $n$ degrees of freedom is given by

$$\Phi(\omega) = \left( \frac{1}{1 - 2j\omega\sigma^2} \right)^{\frac{n}{2}}$$

- When $n = 2$, Chi-Square random variable is an exponential random variable with mean equal to $2\sigma^2$

$$p(x) = \begin{cases} 
\frac{1}{2\sigma^2} e^{-\frac{x}{2\sigma^2}} & x > 0 \\
0 & \text{otherwise}
\end{cases} \quad (2.3-27)$$
The Chi-Square ($\chi^2$) Random Variable

- The Chi-Square random variable is a spatial case of a **Gamma random variable**.
- A Gamma random variable is defined by a PDF of the form:
  
  $$ p(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & x \geq 0, \lambda > 0, \alpha > 0 \\ 0 & \text{otherwise} \end{cases} $$

- A Chi-Square random variable is a gamma random variable with $\lambda = 1/2\sigma^2$ and $\alpha = n/2$. 
The noncentral \( \chi^2 \) random variable with \( n \) degrees of freedom is defined similarly to a \( \chi^2 \) random variable in which \( X_i \)'s are independent Gaussians with common variance \( \sigma^2 \) but with different means denoted by \( m_i \).

The PDF of random variable is given by:

\[
p(x) = \begin{cases} 
\frac{1}{2\sigma^2} \left( \frac{x}{s^2} \right)^{n-2} e^{-\frac{s^2+x}{2\sigma^2}} I_{x/2-1} \left( \frac{s}{\sigma^2} \sqrt{x} \right) & x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

where \( s \) is defined as: \( s = \sqrt{\sum_{i=1}^{n} m_i^2} \)
\( I_\alpha(x) \) is the modified Bessel function of the first kind and order \( \alpha \) given by
\[
I_\alpha(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{\alpha+2k}}{k! \Gamma(\alpha + k + 1)}, \quad x \geq 0
\]

When \( \alpha = 0 \), the function \( I_0(x) \) can be written as
\[
I_0(x) = \sum_{k=0}^{\infty} \left( \frac{x^k}{2^k k!} \right)^2
\]
and for \( x > 1 \) can be approximated by
\[
I_0(x) \approx \frac{e^x}{\sqrt{2\pi x}}
\]
The CDF of this random variable, when $n = 2m$, can be written in the form

$$F(x) = \begin{cases} 
1 - Q_m \left( \frac{s}{\sigma}, \sqrt{x} \right) & x > 0 \\
0 & \text{otherwise}
\end{cases}$$

where $Q_m(a, b)$ is the generalized Marcum Q function and is defined as

$$Q_m(a, b) = \int_b^\infty x \left( \frac{x}{a} \right)^{m-1} e^{-\left(x^2 + a^2\right)/2} I_{m-1}(ax) dx$$

$$= Q_1(a, b) + e^{-\left(a^2 + b^2\right)/2} \sum_{k=1}^{m-1} \left( \frac{b}{a} \right)^k I_k(ab)$$
The Marcum $Q$ function is defined as

$$Q_1(a, b) = \int_b^\infty xe^{-\frac{a^2+x^2}{2}} I_0(ax) \, dx$$

or

$$= e^{-\frac{a^2+b^2}{2}} \sum_{k=0}^{\infty} \left( \frac{a}{b} \right)^k I_k(ab), \quad b \geq a > 0$$

Some properties of this function:

$Q_1(x, 0) = 1$

$Q_1(0, x) = e^{-\frac{x^2}{2}}$

$Q_1(a, b) \approx Q(b-a)$ for $b \gg 1$ and $b \gg b-a$
The Noncentral Chi-Square ($\chi^2$) Random Variable

- The mean and variance is given by
  \[ E[X] = n\sigma^2 + s^2 \]
  \[ \text{var}[X] = 2n\sigma^4 + 4\sigma^2s^2 \]

- The characteristic function is given by
  \[ \Phi(\omega) = \left( \frac{1}{1 - 2j\omega\sigma^2} \right)^n e^{\frac{j\omega s^2}{1 - 2j\omega\sigma^2}} \]
The Rayleigh Random Variable

- If $X_1$ and $X_2$ are iid Gaussian distributed with $\mathcal{N}(0, \sigma^2)$, then $X = \sqrt{X_1^2 + X_2^2}$ is Rayleigh random variable.

- Rayleigh random variable is square root of
  - A Chi-Square random variable with two degrees of freedom
  - An exponential random variable (please see 2.3-27)

- The PDF of Rayleigh random variable is given by

\[
p(x) = \begin{cases} 
  \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & x > 0 \\
  0 & \text{otherwise}
\end{cases}
\]
The mean and variance is given by

\[ E[X] = \sigma \sqrt{\frac{\pi}{2}} \]

\[ \text{var}[x] = \left(2 - \frac{\pi}{2}\right)\sigma^2 \]

In general, the \( n \)th moment of a Rayleigh random variable is given by

\[ E[X^k] = \left(2\sigma^2\right)^{k/2} \Gamma\left(\frac{k}{2} + 1\right) \]

The characteristic function is given by

\[ \Phi_X(\omega) = 1_F\left(1, \frac{1}{2}; -\frac{1}{2} \omega^2 \sigma^2\right) + j \sqrt{\frac{\pi}{2}} \omega \sigma^2 e^{-\frac{\omega^2 \sigma^2}{2}} \]
The Rayleigh Random Variable

\[ \hypergeom{1}{1}{a, b; x} \text{ is the confluent hypergeometric function defined by} \]
\[ \hypergeom{1}{1}{a, b; x} = \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b)x^k}{\Gamma(a)\Gamma(b+k)k!}, \quad b \neq 0, -1, -2, \ldots \]
\[ = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_{0}^{1} e^{xt}t^{a-1}(1-t)^{b-a-1} \, dt \]

\[ \text{In Beaulieu (1990), it is shown that} \]
\[ \hypergeom{1}{1}{1, \frac{1}{2}; -x} = -e^{-x} \sum_{k=0}^{\infty} \frac{x^k}{(2k-1)k!} \]
The Rayleigh Random Variable

- The CDF of a Rayleigh random variable is given by

\[
F(x) = \begin{cases} 
1 - e^{-\frac{x^2}{2\sigma^2}} & x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

- For generalized Rayleigh random variable, we have

\[
X = \sqrt{\sum_{i=1}^{n} X_i^2}, \quad X_i \sim \mathcal{N}(0, \sigma^2)
\]

- The PDF of general case is given by

\[
p(x) = \begin{cases} 
\frac{x^{n-1}}{2^{\frac{n-2}{2}} \sigma^n \Gamma\left(\frac{n}{2}\right)} e^{-\frac{x^2}{2\sigma^2}} & x \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]
The Rayleigh Random Variable

- If $n = 2m$, the CDF for the generalized Rayleigh is given by

$$F(x) = \begin{cases} 
1 - e^{-x^2/(2\sigma^2)} \sum_{k=0}^{m-1} \frac{1}{k!} \left( \frac{x^2}{2\sigma^2} \right)^k & x \geq 0 \\
0 & \text{otherwise}
\end{cases}$$

- The $k$th moment of generalized Rayleigh random variable for any integer value of $n$ (even or odd) is given by

$$E(X^k) = \left( 2\sigma^2 \right)^{k/2} \frac{\Gamma\left( (n+k)/2 \right)}{\Gamma\left( n/2 \right)}, \quad k \geq 0$$
\begin{itemize}
  \item Rayleigh distribution
    \begin{itemize}
    \item Mean: \( r_{\text{mean}} = E[R] = \int_0^\infty r p(r)dr = \sigma \sqrt{\frac{\pi}{2}} = 1.2533\sigma \)
    \item Variance: \( \sigma_r^2 = E[R^2] - E^2[R] = \int_0^\infty r^2 p(r)dr - \frac{\sigma^2 \pi}{2} \)
      \[= \sigma^2 \left(2 - \frac{\pi}{2}\right) = 0.4292 \sigma^2\]
    \item Median value of \( r \) is found by solving: \( \frac{1}{2} = \int_0^{r_{\text{median}}} p(r)dr \)
      \( r_{\text{median}} = 1.177 \sigma \)
    \item Moments of \( R \) are: \( E[R^k] = \left(2\sigma^2\right)^{k/2} \Gamma\left(1+\frac{k}{2}\right)\)
    \item Most likely value: \( \max \{ p_R(r) \} = \sigma \)
    \end{itemize}
\end{itemize}
Rayleigh distribution

Received signal envelope voltage $r$ (volts)
◊ Rayleigh distribution

◊ Probability That Received Signal Doesn’t Exceed A Certain Level (R)

\[
F_R(r) = \int_0^r p(u)du
\]

\[
= \int_0^r \frac{u}{\sigma^2} \exp\left(-\frac{u^2}{2\sigma^2}\right)du
\]

\[
= -\exp\left(-\frac{u^2}{2\sigma^2}\right)|_0^r
\]

\[
= 1 - \exp\left(-\frac{r^2}{2\sigma^2}\right)
\]
◊ Rayleigh distribution

◊ Mean value:

\[ r_{\text{mean}} = E[R] = \int r p(r) dr \]

\[
= \int_0^\infty \frac{r^2}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr
= \int_0^\infty -rd \exp\left(-\frac{r^2}{2\sigma^2}\right) \]

\[
= -r \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right) \bigg|_0^\infty + \int_0^\infty \exp\left(-\frac{r^2}{2\sigma^2}\right) dr
= \sqrt{2\pi} \sigma \cdot \int_0^\infty \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr
= \sqrt{2\pi} \sigma \cdot \frac{1}{2} = \sigma \sqrt{\frac{\pi}{2}} = 1.2533\sigma
\]
Rayleigh distribution:

Mean square value:

\[ E[R^2] = \int_0^\infty r^2 p(r)dr \]

\[ = \int_0^\infty \frac{r^3}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right)dr = \int_0^\infty -r^2 \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right)dr \]

\[ = -r^2 \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right)\bigg|_0^\infty + \int_0^\infty 2r \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right)dr \]

\[ = -2\sigma^2 \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right)\bigg|_0^\infty = 2\sigma^2 \]

0-0=0
Rayleigh distribution

Variance:

\[
\sigma_r^2 = E[R^2] - E^2[R]
\]

\[
= (2\sigma^2) - \left(\sigma \cdot \sqrt{\frac{\pi}{2}}\right)^2
\]

\[
= \sigma^2 \cdot \left(2 - \frac{\pi}{2}\right) = 0.4292\sigma^2
\]
◊ Rayleigh distribution
  ◊ Most likely value
    ◊ Most Likely Value happens when: \( dp(r) / dr = 0 \)

\[
\frac{dp(r)}{dr} = \frac{1}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) - \frac{2r^2}{2\cdot\sigma^4} \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right) = 0
\]

\[\Rightarrow \quad r = \sigma\]

\[\Rightarrow \quad p(r)\big|_{r=\sigma} = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right)\big|_{r=\sigma} = \frac{\exp\left(-\frac{1}{2}\right)}{\sigma} = \frac{0.6065}{\sigma}\]
◊ Rayleigh distribution

◊ Characteristic function

\[ \psi_R(jv) = \int_0^\infty \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} e^{jvr} \, dr \]

\[ = \int_0^\infty \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \cos(vr) \, dr + j\int_0^\infty \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \sin(vr) \, dr \]

\[ = _1F_1 \left( 1, \frac{1}{2}; -\frac{1}{2}v^2\sigma^2 \right) + j\sqrt{\frac{1}{2}} \pi v\sigma^2 e^{-v^2\sigma^2/2} \]

where \(_1F_1\left(1, \frac{1}{2}; -a\right)\) is the confluent hypergeometric function:

\[ _1F_1(\alpha; \beta; x) = \sum_{k=0}^{\infty} \frac{\Gamma(\alpha + k)\Gamma(\beta)x^k}{\Gamma(\alpha)\Gamma(\beta + k)k!}, \quad \beta \neq 0, -1, -2, \ldots \]
Rayleigh distribution

Rayleigh distribution is frequently used to model the statistics of signals transmitted through radio channels such as cellular radio.

Consider a carrier signal $s$ at a frequency $\omega_0$ and with an amplitude $a$:

$$s = a \cdot \exp(j\omega_0 t)$$

The received signal $s_r$ is the sum of $n$ waves:

$$s_r = \sum_{i=1}^{n} a_i \exp[j(\omega_0 t + \theta_i)] \equiv r \exp[j(\omega_0 t + \theta)]$$

where $r \exp(j\theta) = \sum_{i=1}^{n} a_i \exp(j\theta_i)$
Rayleigh distribution

- Define: \( r \exp(j\theta) = \sum_{i=1}^{n} a_i \cos \theta_i + j \sum_{i=1}^{n} a_i \sin \theta_i \equiv x + jy \)

We have: \( x \equiv \sum_{i=1}^{n} a_i \cos \theta_i \) and \( y \equiv \sum_{i=1}^{n} a_i \sin \theta_i \)

where: \( r^2 = x^2 + y^2 \quad x = r \cos \theta \quad y = r \sin \theta \)

- Because (1) \( n \) is usually very large, (2) the individual amplitudes \( a_i \) are random, and (3) the phases \( \theta_i \) have a uniform distribution, it can be assumed that (from the central limit theorem) \( x \) and \( y \) are both Gaussian variables with means equal to zero and variance:

\[
\sigma_x^2 = \sigma_y^2 \equiv \sigma^2
\]
◊ Rayleigh distribution

◊ Because $x$ and $y$ are independent random variables, the joint distribution $p(x, y)$ is

$$p(x, y) = p(x)p(y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

◊ The distribution $p(r, \theta)$ can be written as a function of $p(x, y)$:

$$p(r, \theta) = |J|p(x, y)$$

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r$$

$$p(r, \theta) = \frac{r}{2\pi\sigma} \exp\left(-\frac{r^2}{2\sigma^2}\right)$$
Rayleigh distribution

Thus, the Rayleigh distribution has a PDF given by:

\[
p_R(r) = \frac{2\pi}{\sigma^2} e^{-r^2/2\sigma^2} \begin{cases} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} & r \geq 0 \\ 0 & \text{otherwise} \end{cases}
\]

The probability that the envelope of the received signal does not exceed a specified value \( R \) is given by the corresponding cumulative distribution function (CDF):

\[
F_R(r) = \int_0^r \frac{u}{\sigma^2} e^{-u^2/2\sigma^2} du = 1 - \exp^{-r^2/2\sigma^2}, \quad r \geq 0
\]
If \( X_1 \) and \( X_2 \) are independent Gaussian distributed with \( N(m_1, \sigma^2) \) and \( N(m_2, \sigma^2) \), then \( X = \sqrt{X_1^2 + X_2^2} \) is Ricean random variable.

Ricean random variable is square root of a noncentral Chi-square random variable.

The PDF of Ricean random variable is given by

\[
p(x) = \frac{x}{\sigma^2} \exp \left( -\frac{x^2 + s^2}{2\sigma^2} \right) I_0 \left( \frac{sx}{\sigma^2} \right), \quad x \geq 0
\]

\( s = \sqrt{m_1^2 + m_2^2} \)

For \( s = 0 \), Ricean = Rayleigh

For large \( s \), Ricean \( \approx \) Gaussian
The Ricean Random Variable

CDF of Ricean random variable is given by

\[
F(x) = \begin{cases} 
\frac{x}{\sigma^2} I_0 \left( \frac{sx}{\sigma^2} \right) e^{-\frac{x^2+s^2}{2\sigma^2}} & x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

The first two moments are given by

\[
E[X] = \sigma \sqrt{\frac{\pi}{2}} \, \Gamma \left( \frac{1}{2} \right) _1 \, _1F_1 \left( -\frac{1}{2}, 1, -\frac{s^2}{2\sigma^2} \right)
\]

\[
E[X^2] = 2\sigma^2 + s^2
\]

In general, the \( k \)th moment is given by

\[
E[X^k] = (2\sigma^2)^{k/2} e^{-s^2/2\sigma^2} \Gamma \left( 1 + \frac{k}{2} \right) _1 \, _1F_1 \left( -\frac{k}{2}, 1; -\frac{s^2}{2\sigma^2} \right); \quad k \geq 0
\]
The Ricean Random Variable

- Let $K = s^2/(2\sigma^2)$ (Rice factor)
  \[ A = s^2 + 2\sigma^2 \]

  then we have: $E[X^2] = A$

- The PDF of Ricean random variable written in $K$ and $A$:

\[
p(x) = \begin{cases} 
  \frac{2(K+1)}{A} xe^{-\frac{K+1}{A}x^2 + \frac{AK}{K+1}} I_0 \left(2x \sqrt{\frac{K(K+1)}{A}}\right) & \text{for } x \geq 0 \\
  0 & \text{otherwise}
\end{cases}
\]
The Ricean Random Variable

- If \( \{X_i, i=1,2,\ldots,n\} \) be iid Gaussian random variables with mean \( m_i \) and variance \( \sigma^2 \), then
  \[
  X = \sqrt{\sum_{i=1}^{n} X_i^2}
  \]
  where \( X \) is a generalized Ricean random variable.

- PDF: \( p(x) = \frac{x^{n/2}}{\sigma^2 s^{(n-2)/2}} e^{-(x^2+s^2)/2\sigma^2} I_{n/2-1} \left( \frac{xs}{\sigma^2} \right), \quad x \geq 0 \)

- CDF with \( n = 2m \): \( F(x) = 1 - Q_m \left( \frac{s}{\sigma}, \frac{x}{\sigma} \right), \quad x \geq 0 \)

- The \( k \)th moment in general case is given by
  \[
  E\left[ X^k \right] = \left(2\sigma^2 \right)^{k/2} e^{-s^2/2\sigma^2} \frac{\Gamma\left( (n+k)/2 \right)}{\Gamma\left( n/2 \right)} \text{$_1$F$_1$} \left( \frac{n+k}{2}, \frac{n}{2}; \frac{s^2}{2\sigma^2} \right)
  \]
Rice distribution

- When there is a dominant stationary (non-fading) signal component present, such as a line-of-sight (LOS) propagation path, the small-scale fading envelope distribution is Rice.

\[ s_r = r' \exp[j(\omega_0 t + \theta)] + A \exp(j\omega_0 t) \]

\[ \equiv [(x + A) + jy] \exp(j\omega_0 t) \equiv r \exp[j(\omega_0 t + \theta)] \]

\[ r^2 = (x + A)^2 + y^2 \]

\[ x + A = r \cos \theta \]

\[ y = r \sin \theta \]
Rice Distribution in Wireless Communications

◊ Rice distribution

◊ By following similar steps described in Rayleigh distribution, we obtain:

\[
p(r) = \begin{cases} 
\frac{r}{\sigma_r^2} \exp\left( -\frac{r^2 + A^2}{2\sigma_r^2} \right) I_0 \left( \frac{Ar}{\sigma_r^2} \right) & \text{for } (A \geq 0, r \geq 0) \\
0 & \text{for } (r < 0)
\end{cases}
\]

where

\[
I_0 \left( \frac{Ar}{\sigma_r^2} \right) = \frac{1}{2\pi} \int_0^{2\pi} \exp \left( \frac{Ar \cos \theta}{\sigma_r^2} \right) d\theta
\]

is the modified zeroth-order Bessel function.

\[
I_0(x) = \sum_{i=0}^{\infty} \left( \frac{x^i}{i! \cdot 2^i} \right)
\]
Rice distribution

- The Rice distribution is often described in terms of a parameter $K$ which is defined as the ratio between the deterministic signal power and the variance of the multi-path. It is given by $K = A^2/(2\sigma^2)$ or in terms of dB:

$$k(dB) = 10 \cdot \log \frac{A^2}{2\sigma^2} \quad [dB]$$

- The parameter $K$ is known as the Rice factor and completely specifies the Rice distribution.

- As $A \to 0$, $K \to -\infty$ dB, and as the dominant path decreases in amplitude, the Rice distribution degenerates to a Rayleigh distribution.
Rice Distribution in Wireless Communications

◊ Rice distribution

![Graph showing Rice distribution for different signal envelope levels](image-url)
Rice distribution

\[ s_r = \sum_{i=1}^{n} a_i \exp[j(\omega_0 t + \theta_i)] + A \exp(j\omega_0 t) \]

\[ = \left[ \sum_{i=1}^{n} a_i \exp(j\theta_i) \right] j(\omega_0 t) + A \exp(j\omega_0 t) \]

\[ \equiv r' \exp(j\theta) \exp(j\omega_0 t) + A \exp(j\omega_0 t) \]

\[ = r' \exp[j(\omega_0 t + \theta)] + A \exp(j\omega_0 t) \]

\[ \equiv [(x + A) + jy] \exp(j\omega_0 t) \equiv r \exp[j(\omega_0 t + \theta)] \]

where \[ r' \exp(j\theta) = \sum_{i=1}^{n} a_i \exp(j\theta_i) \]
Rice distribution

Define: \( r' \exp(j \theta) = \sum_{i=1}^{n} a_i \cos \theta_i + j \sum_{i=1}^{n} a_i \sin \theta_i \equiv x + jy \)

We have: \( x \equiv \sum_{i=1}^{n} a_i \cos \theta_i \) and \( y \equiv \sum_{i=1}^{n} a_i \sin \theta_i \)

and \( r^2 = (x + A)^2 + y^2 \quad x + A = r \cos \theta \quad y = r \sin \theta \)

Because (1) \( n \) is usually very large, (2) the individual amplitudes \( a_i \) are random, and (3) the phases \( \theta_i \) have a uniform distribution, it can be assumed that (from the central limit theorem) \( x \) and \( y \) are both Gaussian variables with means equal to zero and variance:

\[ \sigma_x^2 = \sigma_y^2 \equiv \sigma^2 \]
Rice distribution

Because $x$ and $y$ are independent random variables, the joint distribution $p(x, y)$ is

$$p(x, y) = p(x)p(y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

The distribution $p(r, \theta)$ can be written as a function of $p(x, y)$:

$$p(r, \theta) = |J| p(x, y)$$

$$J \equiv \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r$$
Rice distribution

\[ p(r, \theta) = \frac{r}{2\pi \sigma} \exp\left( -\frac{x^2 + y^2}{2\sigma^2} \right) \]

\[ = \frac{r}{2\pi \sigma} \exp\left( -\frac{(r \cos \theta - A)^2 + (r \sin \theta)^2}{2\sigma^2} \right) \]

\[ = \frac{r}{2\pi \sigma} \exp\left( -\frac{r^2 + A^2 - 2Ar \cos \theta}{2\sigma^2} \right) \]

\[ = \frac{r}{\sigma} \exp\left( -\frac{r^2 + A^2}{2\sigma^2} \right) \frac{1}{2\pi} \exp\left( \frac{Ar \cos \theta}{\sigma^2} \right) \]
◊ **Rice distribution**

The Rice distribution has a probability density function (pdf) given by:

\[
p(r) = \int_0^{2\pi} p(r, \theta) d\theta
\]

\[
= \begin{cases} 
\frac{r}{\sigma^2} \exp \left( - \frac{r^2 + A^2}{2\sigma^2} \right) \frac{1}{2\pi} \int_0^{2\pi} \exp \left( \frac{Ar \cos \theta}{\sigma^2} \right) d\theta & r \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]
The Nakagami Random Variable

- Frequently used to characterize the statistics of signals transmitted through multi-path fading channels.
- PDF of *Nakagami*-m random variable
  \[ p(x) = \frac{2}{\Gamma(m)} \left( \frac{m}{\Omega} \right)^m x^{2m-1} e^{-mx^2/\Omega} \]
  - \( \Omega = E\left(X^2\right) \)
  - *Fading Figure* (m) defined as the ratio of moments
    \[ m = \frac{\Omega^2}{E[(R^2 - \Omega)^2]} , \quad m \geq \frac{1}{2} \]
- The nth moment is given by
  \[ E\left[X^n\right] = \frac{\Gamma(m + n/2)}{\Gamma(m)} \left( \frac{\Omega}{m} \right)^{n/2} \]
**The Nakagami Random Variable**

- **Mean and variance**
  
  \[
  E[X] = \frac{\Gamma(m + 1/2)}{\Gamma(m)} \left( \frac{\Omega}{m} \right)^{1/2}
  \]
  
  \[
  \text{var}[X] = \Omega \left( 1 - \frac{1}{m} \left( \frac{\Gamma(m + 1/2)}{\Gamma(m)} \right)^2 \right)
  \]
  
- **$m=1$** → Rayleigh Fading
- **$\frac{1}{2} < m < 1$**, PDF has larger tail
- **$m > 1$**, the tail of PDF decays faster
The Lognormal Random Variable

- Let $Y \sim \mathcal{N}(m, \sigma^2)$, and $Y = \ln X \rightarrow X = e^Y$ is lognormal distributed, then the PDF of $X$ is
  
  $$p(x) = \frac{1}{\sqrt{2\pi \sigma x}} e^{-\frac{(\ln x - m)^2}{2\sigma^2}}, \quad x \geq 0$$

- Mean and variance
  
  $$E[X] = e^{m + \sigma^2/2}$$
  $$\text{var}[X] = e^{2m + \sigma^2} \left(e^{\sigma^2} - 1\right)$$

- Modeling the effect of *shadowing* of the signal due to large obstructions in mobile radio communications
Jointly Gaussian Random Variable

- Let \( X_i, i=1,2,\cdots,n \), are \textit{Gaussian} random variables with means \( m_i \), variances \( \sigma_i^2 \), and covariances \( \mu_{ij}, \ i,j=1,2,\cdots,n \).

- \( X = [X_1, X_2, \ldots, X_n]^t \) is a \textit{Gaussian random vector} and \( \{X_i\} \) are \textit{joint Gaussian random variables} or \textit{multivariate Gaussian random variables}.

- PDF of \( X \) is

\[
p(x) = \frac{1}{(2\pi)^{n/2} (\det C)^{1/2}} \exp \left( -\frac{1}{2} (x-m)^t C^{-1} (x-m) \right)
\]

where \( m = E[X] = [m_1, m_2, \ldots, m_n]^t \)

\( C = E[(x-m) (x-m)^t] \), where \( C_{ij} = \text{Cov}(X_i, X_j) \)
Jointly Gaussian Random Variable

- For \( n=2 \), \( \mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} \), \( \mathbf{C} = \begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix} \)

- Where \( \rho = \frac{\text{Cov}(X_1, X_2)}{\sigma_1 \sigma_2} \) is correlation coefficient

- PDF is given by

\[
p(x_1, x_2) = \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1-\rho^2}} \exp \left\{ -\frac{\left( \frac{x_1-m_1}{\sigma_1} \right)^2 + \left( \frac{x_2-m_2}{\sigma_2} \right)^2}{2(1-\rho^2)} - 2\rho \left( \frac{x_1-m_1}{\sigma_1} \right) \left( \frac{x_2-m_2}{\sigma_2} \right) \right\}
\]

- When \( x_1 \) and \( x_2 \) are uncorrelated, i.e. \( \rho = 0 \),

\[
p(x_1, x_2) = \frac{1}{2\pi \sigma_1 \sigma_2} \exp \left\{ -\frac{(x_1-m_1)^2}{2\sigma_1^2} - \frac{(x_2-m_2)^2}{2\sigma_2^2} \right\} = N(m_1, \sigma_1^2) \times N(m_2, \sigma_2^2)
\]

\( x_1 \) and \( x_2 \) are independent
Jointly Gaussian Random Variable

- $\{X_i\}$ are uncorrelated $\rightarrow$ $\{X_i\}$ are independent.

- Linear combinations of Gaussian random variables are Gaussian random variables.

- Set of Gaussian random variables is jointly Gaussian.

- If $X$ is a Gaussian vector, then $Y = AX$, where $A$ is nonsingular, is also a Gaussian vector.
Chapter 2.4: Bounds on Tail Probabilities
In many cases, the error probability of a communication system is expressed in terms of the probability that a random variable exceeds a certain value, i.e., in the form of $P[X > a]$.

Unfortunately, in many cases these probabilities cannot be expressed in closed form.

In such cases we are interested in finding upper bounds on these tail probabilities.
The Markov Inequality

- The Markov inequality gives an upper bound on the tail probability of nonnegative random variables.
- Assume that $X$ is a non-negative random variable, i.e., $p(x) = 0$ for all $x < 0$.
- For any $\alpha > 0$,
  \[ P[X \geq \alpha] \leq \frac{E[X]}{\alpha} \]

**Proof:**

\[
E[X] = \int_{0}^{\infty} xp(x)dx \geq \int_{\alpha}^{\infty} xp(x)dx \geq \alpha \int_{\alpha}^{\infty} p(x)dx = \alpha P[X \geq \alpha]
\]
Chernov Bound

◊ Chernov bound is very tight and can be applied to all random variable.

◊ Let $\delta$ and $\nu \neq 0$ are arbitrary real numbers

$$P[e^{\nu X} \geq e^{\nu \delta}] \leq \frac{E[e^{\nu X}]}{e^{\nu \delta}} = E[e^{\nu(X - \delta)}]$$

Thus, $P[X \geq \delta] \leq E[e^{\nu(X - \delta)}]$, $\forall \nu > 0$

$P[X \leq \delta] \leq E[e^{\nu(X - \delta)}]$, $\forall \nu < 0$

◊ Since $\nu$ is arbitrary, we want to find the value of $\nu^*$ such that

◊ $E[e^{\nu^*(X - \delta)}]$ is a **tightest upper bound** of $P[X \geq \delta]$ or $P[X \leq \delta]$

◊ $E[e^{\nu^*(X - \delta)}]$ is smallest
Chernov Bound

- Let $g(v) = E[e^{v(X-\delta)}]$
  
  $g'(v) = E[(X-\delta)e^{v(X-\delta)}]$
  
  $g''(v) = E[(X-\delta)^2 e^{v(X-\delta)}]$

- $\because g''(v) > 0$, the smallest $g(v^*)$ occurs when $g'(v) = 0$
  
  $\Leftrightarrow v^*$ is the root of $E[Xe^{vX}] = \delta E[e^{vX}]$ (2.4-7)

- To know $v^*$ is positive or negative:
  
  - If $g'(0) < 0 \rightarrow v^* > 0$, and vice versa
  
  - When $v = 0$, $g'(0) = E[X] - \delta$
  
  - Thus, $P[X \geq \delta] \leq e^{-v^*\delta}E[e^{v^*X}]$, $\forall \delta > E[X]$
  
  $P[X \leq \delta] \leq e^{-v^*\delta}E[e^{v^*X}]$, $\forall \delta < E[X]$
Chernov Bound for Sums of Random Variables

Let $X_i$, $i=1,2,\ldots,n$, be iid random variables, and $Y$ be the sample mean:

$$Y = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Applying Chernov bound,

$$P[Y \geq \delta] = P\left[ \sum_{i=1}^{n} X_i \geq n\delta \right] \leq E\left[ \exp\left( \nu \left( \sum_{i=1}^{n} X_i - n\delta \right) \right) \right] = \left[ E\left[ e^{\nu(X-\delta)} \right] \right]^n, \quad \nu > 0$$

The optimal $\nu^*$ occurs when $E[e^{\nu(X-\delta)}]$ is smallest

$\quad \rightarrow \nu^*$ is the root of $E[Xe^{\nu X}] = \delta E[e^{\nu X}]$  Exactly the same as (2.4-7)

Thus, $P[Y \geq \delta] \leq \left[ E\left[ e^{\nu^*(X-\delta)} \right] \right]^n = e^{-\nu^*\delta} \left[ E\left[ e^{\nu^* X} \right] \right]^n$
Chapter 2.5 : Limit Theorems for Sums of Random Variables

Wireless Information Transmission System Lab.
Institute of Communications Engineering
National Sun Yat-sen University
Limit Theorems for Sums of r.v.s

- Let $X_i, i = 1, 2, \ldots, n$, be i.i.d. r.v., and $E[X_i] = m < \infty$, $Var[X_i] = \sigma^2 < \infty$

- **Strong Law of Large Number (LLN):**
  \[
  Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i \rightarrow m
  \]
  - $Y_n$ converges to $m$ almost everywhere, or almost surely
  \[
  \Pr[Y_n = m] \rightarrow 1, n \rightarrow \infty
  \]

- **Central Limit Theorem (CLT)**
  \[
  \frac{1}{n} \sum_{i=1}^{n} X_i - m \rightarrow N(0,1)
  \]
  - Converges in *distribution*, i.e.
  \[
  \Pr\left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{X_i - m}{\sigma} \leq t \right\} \rightarrow 1 - Q(t)
  \]
Chapter 2.6 : Complex Random Variables

Wireless Information Transmission System Lab.
Institute of Communications Engineering
National Sun Yat-sen University
A complex random variable $Z = X + jY$ can be considered as a pair of real random variables $X$ and $Y$.

If $X$ and $Y$ are jointly Gaussian random variables, $Z$ is a complex Gaussian random variable.

(ex) Consider a complex r.v. $Z = X + jY$, when $X$ and $Y$ are i.i.d. Gaussian $\sim N(0, \sigma^2)$, the PDF of $Z$ is

$$p(z) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2} e^{\frac{|z|^2}{2\sigma^2}}$$

The mean and variance of a complex r.v. are defined by:

- $\mathbb{E}[Z] = \mathbb{E}[X] + j\mathbb{E}[Y]$
- $\text{Var}[Z] = \mathbb{E}[|Z|^2] - |\mathbb{E}[Z]|^2 = \text{Var}[X] + \text{Var}[Y]$
A complex random vector is defined as $Z = X + jY$, where $X$ and $Y$ are real-valued random vectors of size $n$.

For $Z = X + jY$, define the following real-valued matrices.

- $C_X = E[(X - E[X])(X - E[X])^t]$ : Covariance matrix of $X$
- $C_Y = E[(Y - E[Y])(Y - E[Y])^t]$ : Covariance matrix of $Y$
- $C_{XY} = E[(X - E[X])(Y - E[Y])^t]$
- $C_{YX} = E[(Y - E[Y])(X - E[X])^t]$

- $C_X$ and $C_Y$ are symmetric and non-negative definite
- $C_{XY} = C_{YX}^t$

Let $\tilde{Z} = \begin{bmatrix} X \\ Y \end{bmatrix}$, then $\text{Cov}(\tilde{Z}) = C_{\tilde{Z}} = \begin{bmatrix} C_X & C_{XY} \\ C_{YX} & C_Y \end{bmatrix}$
Chapter 2.6-1 Complex Random Vectors

- $C_Z = E[(Z - E[Z])(Z - E[Z])^H]$: covariance matrix of $Z$
  - Hermitian and non-negative definite
- $\tilde{C}_Z = E[(Z - E[Z])(Z - E[Z])']$: pseudo-covariance matrix of $Z$
  - Skew-Hermitian

- $C_Z = C_X + C_Y + j(C_{YX} - C_{XY})$
- $\tilde{C}_Z = C_X - C_Y + j(C_{XY} + C_{YX})$

Hermitian $A = A^H$
Skew-Hermitian $A^H = -A$
Proper Complex Random Vectors

- \( Z = X + jY \) is *proper* if \[ \tilde{C}_z = C_x - C_y + j(C_{xy} + C_{yx}) = 0 \] (2.6-14)
  - \( C_x = C_y \)
  - \( C_{xy} = -C_{yx} = -C'_{xy} \)
  - Thus, \[ C_z = 2C_x + 2jC_{yx} \]
  - \( \frac{1}{2} \text{Re}[C_z] \)
  - \( \frac{1}{2} \text{Im}[C_z] \)

- (ex 1) When \( n=1 \), \( Z = X + jY \) is proper, which implies
  - \( \text{Var}[X] = \text{Var}[Y] \)
  - \( \text{Cov}[X,Y] = -\text{Cov}[Y,X] \)

\( Z \) is proper if \( X \) and \( Y \) have *equal variances* and are *uncorrelated*.

Jointly Gaussian r.v.: uncorrelated=independent.

Uncorrelated: \( E[xy] = m_x m_y \)
Proper Complex Random Vectors

◊ If $Z = X + jY$ is complex Gaussian, i.e. $X$ and $Y$ are jointly Gaussian, the PDF of $Z$ is

$$p(z) = p(\tilde{z}) = \frac{1}{(2\pi)^n (\det C_{\tilde{z}})^2} e^{-\frac{1}{2}(\tilde{z} - \tilde{m})' C_{\tilde{z}}^{-1} (\tilde{z} - \tilde{m})}$$

◊ where $\tilde{m} = E[\tilde{Z}] = \begin{bmatrix} m_x \\ m_y \end{bmatrix}$, $C_{\tilde{z}} = \begin{bmatrix} C_X & C_{XY} \\ C_{YX} & C_Y \end{bmatrix}$

◊ When $Z$ is proper, it can be shown that the PDF of $Z$ is

$$p(z) = \frac{1}{\pi^n \det C_Z} e^{-\frac{1}{2}(z - m)^\dagger C_Z^{-1} (z - m)}$$
Circular Complex Random Vectors

◊ **Z** is *circular* or *circularly symmetric* if rotating **Z** by any angle does not change the PDF

◊ **Z** is *circularly symmetric* \( \iff \) **Z** and \( e^{j\theta}Z \) have the same PDF

\[
E[Z] = e^{j\theta} E[Z], \forall \theta \iff E[Z] = 0
\]

\[
\tilde{C}_Z = E[ZZ^\prime] = \tilde{C}_{e^{j\theta}Z} = e^{j2\theta} E[ZZ^\prime], \forall \theta \iff \tilde{C}_Z = 0
\]

◊ **Z** is *circularly symmetric* \( \Rightarrow \) **Z** is *zero-mean and proper*

◊ **Z** is *zero-mean, proper* and *Gaussian* \( \iff \) **Z** is *circularly symmetric*

◊ **Z** is *complex proper* \( \iff \) \( AZ + b \) is *complex proper*
Chapter 2.7 : Random Processes
Random Processes

- At any time instant, the value of a random process $X(t)$ is a random variable.
- The r.vs $X_{t_i} \equiv X(t_i), i = 1, 2, ..., n,$ are characterized statistically by their joint PDF $p(x_{t_1}, x_{t_2}, ..., x_{t_n}).$
- *Mean* function:
  $$m_X(t) = E[X(t)]$$
- *Autocorrelation* function:
  $$R_X(t_1, t_2) = E[X(t_1)X^*(t_2)]$$
  - $R_X(t_1, t_2) = R_X(t_2, t_1)$
- *Cross-correlation* function
  $$R_{XY}(t_1, t_2) = E[X(t_1)Y^*(t_2)]$$
  - $R_{XY}(t_1, t_2) = R_{YX}(t_2, t_1)$
Chapter 2.7-1 WSS Random Processes

◊ $X(t)$ is Wise-Sense-Stationary (WSS) if
  
  (1) $m_X(t) = m_X$, $\forall t$
  
  (2) $R_X(t_1, t_2) = R_X(\tau)$, $\tau = t_1 - t_2$

◊ $R_X(-\tau) = R_X^*(\tau)$

◊ $X(t)$ and $Y(t)$ are Jointly WSS if
  
  (1) $X(t)$ and $Y(t)$ are WSS
  
  (2) $R_{XY}(t_1, t_2) = R_{XY}(\tau)$, $\tau = t_1 - t_2$

◊ $R_{XY}(\tau) = R_{YX}^*(-\tau)$

◊ A complex random process $Z(t) = X(t) + jY(t)$ is WSS if $X(t)$ and $Y(t)$ are jointly WSS.
The power spectral density (PSD) function or power spectrum of a WSS r.p. is the distribution of power in terms of frequency.

Wiener-Khinchin theorem states that PSD of a WSS r.p. $X(t)$ is

$$S_X(f) = F[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau)e^{-j2\pi f \tau} d\tau \quad (W/Hz)$$

- $R_X(-\tau) = R_X^*(\tau) \Rightarrow S_X(f) \geq 0$ (Proves are omitted.)
- If WSS $X(t)$ is real, $S_X(f)$ is a real, nonnegative, even function of $f$.

Cross spectral density (CSD) of two jointly WSS r.p. is

$$S_{XY}(f) = F[R_{XY}(\tau)]$$

- $R_{XY}(\tau) = R_{YX}^*(-\tau) \Rightarrow S_{XY}(f) = S_{YX}^*(f)$

The power of a WSS r.p. $X(t)$

$$P_X = E\left[|X(t)|^2\right] = R_X(0) = \int_{-\infty}^{\infty} S_X(f) df$$
If $X(t)$ and $Y(t)$ are jointly WSS, $Z(t) = aX(t) + bY(t)$ is WSS

$$R_Z(\tau) = |a|^2 R_X(\tau) + |b|^2 R_Y(\tau) + ab^* R_{XY}(\tau) + ba^* R_{YX}(\tau)$$

$$S_Z(f) = |a|^2 S_X(f) + |b|^2 S_Y(f) + 2 \text{Re}[ab^* S_{XY}(f)]$$

When $a=b=1$, $Z(t) = X(t) + Y(t)$

$$R_Z(\tau) = R_X(\tau) + R_Y(\tau) + R_{XY}(\tau) + R_{YX}(\tau)$$

$$S_Z(f) = S_X(f) + S_Y(f) + 2 \text{Re}[S_{XY}(f)]$$

When $a=1, b=j$, $Z(t) = X(t) + jY(t)$

$$R_Z(\tau) = R_X(\tau) + R_Y(\tau) + j(R_{YX}(\tau) - R_{XY}(\tau))$$

$$S_Z(f) = S_X(f) + S_Y(f) + 2 \text{Im}[S_{XY}(f)]$$
When a WSS r.p. $X(t)$ passes thru an LTI system $h(t)$

$$Y(t) = X(t) * h(t) = \int_{-\infty}^{\infty} h(\tau)X(t - \tau) d\tau$$

$m_Y = m_X \int_{-\infty}^{\infty} h(t) dt$

$R_{XY}(\tau) = E[X(t + \tau)Y^*(t)] = R_X(\tau) * h^*(-\tau)$

$R_Y(\tau) = E[Y(t + \tau)Y^*(t)] = R_X(\tau) * h(\tau) * h^*(-\tau)$

$Y(t)$ is also WSS, $X(t)$ and $Y(t)$ are jointly WSS

Let $\mathcal{F}\{h(t)\} = H(f)$

$m_Y = m_X H(0)$

$S_{XY}(f) = S_X(f)H^*(f)$

$S_Y(f) = S_X(f)|H(f)|^2$
Gaussian Random Process

- A real r.p. $X(t)$ is Gaussian if $\left[ X(t_1), X(t_2), \cdots, X(t_n) \right]'$ is a Gaussian random vector for any $n$ and any $(t_1, t_2, \ldots, t_n)$.
  That is, random variables $\{X(t_i)\}_{i=1}^n$ are jointly Gaussian r.v.s.
  - *Linear filtering of a Gaussian r.p. is also Gaussian* even if the filtering is time varying.
- Two real r.p. $X(t)$ and $Y(t)$ are jointly Gaussian if
  $\left[ X(t_1), X(t_2), \cdots, X(t_n), Y(t'_1), Y(t'_2), \cdots, Y(t'_m) \right]'$ is a Gaussian random vector for any $n, m$ and for any $(t_1, t_2, \ldots, t_n)$ and $(t'_1, t'_2, \ldots, t'_m)$.
  - Two jointly Gaussian r.p. are uncorrelated $\Rightarrow$ independent.
- $Z(t) = X(t) + jY(t)$ is complex Gaussian if $X(t)$ and $Y(t)$ are jointly Gaussian.
Chapter 2.7-1 WSS Random Processes

White Process

- **White process:** the PSD is a constant for all $f$, usually denoted by $N_0/2$

$$S_X(f) = \frac{N_0}{2} \iff R_X(\tau) = \frac{N_0}{2} \delta(t)$$

- Infinite Energy $\Rightarrow$ Not a physically existing process
- Closely modeling some physical phenomenon, e.g., thermal noise

- **Properties of thermal noise $N(t)$**
  1. stationary
  2. zero-mean
  3. Gaussian
  4. $N(t)$ is white with PSD

$$S_N(f) = \frac{N_0}{2} = \frac{kT}{2}$$

$T$ is ambient temperature ($K$) and $k=1.38 \times 10^{-23}$ is Boltzmann’s const.
Discrete Time Random Process

- Discrete-time process $X(n)$ consisting of an ensemble of sample sequences $\{x(n)\}$ are usually obtained by uniformly sampling a continuous-time random process.
- The PSD of a WSS discrete-time random process is
  \[ S_X(f) = \sum_{m=-\infty}^{\infty} R_X(m)e^{-j2\pi fm} \]
- The autocorrelation is the inverse Fourier transform of PSD
  \[ R_X(m) = \int_{-1/2}^{1/2} S_X(f)e^{j2\pi fm}df \]
- The power of discrete-time random process is
  \[ P = \mathbb{E}\left[|X(n)|^2\right] = R_X(0) = \int_{-1/2}^{1/2} S_X(f)df \]
Cyclostationary Random Processes

- \(X(t)\) is cyclostationary if mean and autocorrelation functions are periodic with period \(T_0\), i.e.,
  \[m_x(t + T_0) = m_x(t)\]
  \[R_x(t_1 + T_0, t_2 + T_0) = R_x(t_1, t_2)\]

- In communication systems, many modulated processes are modeled as cyclostationary.

- The average autocorrelation function is
  \[\overline{R_x(\tau)} = \frac{1}{T_0} \int_0^{T_0} R_x(t + \tau, t) dt\]

- The average power spectral density function is
  \[S_x(f) = F[\overline{R_x(\tau)}]\]
[ex] Let \( \{a_n\} \) be a discrete time WSS r.p. with \( m_a(n)=m_a \), and \( R_a(m)=E[a_{n+m}a_n^*] \). For any deterministic function \( g(t) \), let

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

show that \( X(t) \) is cyclostationary.
Chapter 2.7-3 Proper and Circular Random Process

- For a complex r.p. $Z(t)=X(t)+jY(t)$, let $\bar{Z}(t) = Z(t) - E[Z(t)]$

  1. $C_Z(t+\tau,t) = E[\bar{Z}(t+\tau)\bar{Z}^*(t)]$ : covariance function
  2. $\tilde{C}_Z(t+\tau,t) = E[\bar{Z}(t+\tau)\bar{Z}(t)]$ : pseudo-covariance function

- $C_Z(t+\tau,t) = C_X(t+\tau,t) + C_Y(t+\tau,t) + j(C_{YX}(t+\tau,t) - C_{XY}(t+\tau,t))$
- $\tilde{C}_Z(t+\tau,t) = C_X(t+\tau,t) - C_Y(t+\tau,t) + j(C_{YX}(t+\tau,t) + C_{XY}(t+\tau,t))$

- A complex r.p. $Z(t)$ is proper if $\tilde{C}_Z(t+\tau,t) = 0$
  - $C_X(t+\tau,t) = C_Y(t+\tau,t)$
  - $C_{YX}(t+\tau,t) = -C_{XY}(t+\tau,t)$
  - $C_Z(t+\tau,t) = 2C_X(t+\tau,t) + j2C_{YX}(t+\tau,t)$
Chapter 2.7-3 Proper and Circular Random Process

- If $Z(t)$ is a proper Gaussian r.p., then for all $n$ and $(t_1, t_2, \ldots, t_n)$, 
  \[ [Z(t_1), Z(t_2), \ldots, Z(t_n)]' \]
  is a proper Gaussian vector

- A complex r.p. $Z(t)$ is circular if, for all $\theta$ and $Z(t)$, $e^{i\theta} Z(t)$ have the same statistical properties
- $Z(t)$ is circular $\Rightarrow$ $Z(t)$ is zero-mean and proper
- If $Z(t)$ is Gaussian:
  
  circular $\iff$ zero-mean and proper
- Linearly filtering of a circular Gaussian r.p. is circular Gaussian
Chapter 2.7-4 Markov Chain

- **Markov Chain:**
  - Discrete-time and discrete-valued
  - Current value depends on the entire past values only thru the most recent values.
  - In a $j$th-order Markov chain, the current value depends on the past values only through the most recent $j$ values, i.e.
    \[
    P(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots) = P(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_{n-j} = x_{n-j})
    \]

- Let the set of the *most recent $j$ values* be the *state* of the Markov chain, and the current state is denoted as
  \[
  S_n = (X_n, X_{n-1}, \ldots, X_{n-j+1})
  \]
  - $S_n$ depends only on $S_{n-1}$, *i.e.*
    \[
    P(S_n = s_n | S_{n-1} = s_{n-1}, S_{n-2} = s_{n-2}, \ldots) = P(S_n = s_n | S_{n-1} = s_{n-1})
    \]

  - \{S_n\} is first order Markov chain
  - $X_n$ is a deterministic function of state $S_n$
In general,
\[ P[X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_{n-j} = x_{n-j}] \]
\[ = P[X_n = x_n \mid S_n = s_n]P[S_n = s_n \mid S_{n-1} = s_{n-1}] \]

Finite state machine: at time \( n \), state \( S_n \) takes value in the set \( \{1, 2, \ldots, S\} \) and
\[ P[S_n = s_n \mid S_{n-1} = s_{n-1}, S_{n-2} = s_{n-2}, \ldots] = P[S_n = s_n \mid S_{n-1} = s_{n-1}] \]

The internal development of \( \{S_n\} \) depends on \( \{1, 2, \ldots, S\} \) and \( P[S_n \mid S_{n-1}] \)

A Markov chain is homogeneous if \( P[S_n \mid S_{n-1}] \) is indep. of \( n \)

\( \Rightarrow \) Transition probability form state \( i \) to state \( j \) is indep. of \( n \)
Chapter 2.7-4 Markov Chain

- In homogeneous Markov chain, **State transition matrix**
  (one-step transition matrix) $P$ with the $(i,j)$-th element
  
  $$[P]_{ij} = P_{ij} = P[S_n = j|S_{n-1} = i], \quad 1 \leq i, j \leq S$$
  
  - $P_{ij} \geq 0$ is the transition probability from state $i$ to state $j$
  - $P_{i1} + P_{i2} + \ldots + P_{iS} = 1$

- **State probability vector**:
  
  $$p(n) = [p_1(n) \quad p_2(n) \quad \ldots \quad p_S(n)]$$
  
  - $p_i(n)$: probability of being in state $i$ in time $n$
  - $p_1(n) + p_2(n) + \ldots + p_S(n) = 1$

- $p(n) = p(n-1)P$
- $p(n) = p(0)P^n$
Chapter 2.7-4 Markov Chain

◊ If \( \lim_{n \to \infty} P^n \) exists and
\[
\lim_{n \to \infty} P^n = \begin{bmatrix} p \\ p \\ \vdots \\ p \end{bmatrix},
\]
for any \( p(0) \)
\[
\lim_{n \to \infty} p(n) = \lim_{n \to \infty} p(0)P^n = p(0) \begin{bmatrix} p \\ p \\ \vdots \\ p \end{bmatrix} = p
\]

◊ For any \( p(0) \), the Markov chain stabilizes at \( p \)

◊ \( p : Steady-state \) of Markov chain, which can be obtained by
\[
p \cdot P = p
\]

◊ Does \textit{steady state} exist and is it unique?

◊ Does \( \lim_{n \to \infty} P^n \) exist and have equal rows
Markov chain is *irreducible* if it is possible to move from any state to any other state in finite steps, i.e.,
\[
\forall i \neq j, \exists n < \infty, \text{ s.t. } P[S_{m+n} = j | S_m = i] > 0
\]

Period of state \(i\) is
\[
gcd\{n : P[S_n = i | S_0 = i] > 0\}
\]

State \(i\) is *aperiodic* if its period is one.

Markov chain is *ergodic* if it’s *irreducible* & all states are *aperiodic*

- \(\lim_{n \to \infty} P^n\) exists and *have equal rows*
- The steady state probability *exists* and is *unique*
(ex) A four-state Markov chain has transition probability matrix

\[
P = \begin{bmatrix}
\frac{1}{2} & 1 & 0 & 1 \\
\frac{1}{3} & 0 & 1 & 0 \\
\frac{1}{2} & 0 & 2 & 0 \\
0 & \frac{1}{4} & 0 & \frac{3}{4} \\
\frac{5}{6} & 0 & \frac{1}{6} & 0
\end{bmatrix}
\]
Chapter 2.8 : Series Expansion of Random Processes
Chapter 2.8-1 Sampling Theorem for Band-limited Random Process

- **Sampling**: straightforward method to represent a process with a set of variables
- A deterministic real signal is **bandlimited** if
  \[ X(f) = 0, \text{ for } |f| > W; \text{ } W: \text{highest frequency in } x(t) \]
- **Nyquist rate**: the minimum sampling rate is \( f_N = 2W \)
  - For complex signal whose frequency support is between \([W_1, W_2]\)
    \[ f_N = 2W = W_2 - W_1 \]
  - Signals can be **perfectly reconstructed** if sampling rate is at least 2W
  - Require \(2W \text{ real}\) numbers per second to describe a **real** signal
  - Require \(2W \text{ complex}\) numbers (\(4W \text{ real}\) numbers) per second to describe a **complex** signal
Chapter 2.8-1 Sampling Theorem for Band-limited Random Process

- Sampling below Nyquist rate results in **aliasing**
- A *bandlimited* signal sampling at Nyquist rate can be *reconstructed* from samples using interpolation formula
  \[ x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2W}\right) \text{sinc}\left[2W\left(t - \frac{n}{2W}\right)\right] \]
- Equivalent to pass thru a *lowpass filter*: \( h(t) = \text{sinc}(2Wt) \)
- \( s_n(t) = \text{sinc}\left[2W\left(t - \frac{n}{2W}\right)\right] \) are orthogonal: \( \langle s_n(t), s_m(t) \rangle = \frac{1}{2W} \delta(m - n) \)
A stationary random process $X(t)$ is bandlimited if $S_X(f) = 0$, for $|f| > W$.

$S_X(f) = \mathcal{F}[R_X(\tau)]$, with samples $\{R_X(n/2W)\}$,

$$R_X(\tau) = \sum_{n=-\infty}^{\infty} R_X\left(\frac{n}{2W}\right) \sin c\left[2W\left(\tau - \frac{n}{2W}\right)\right]$$

A bandlimited stationary r.p. $X(t)$ can be represented with

$$X(t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2W}\right) \sin c\left[2W\left(t - \frac{n}{2W}\right)\right]$$

$\{X(n/2W)\}$ are random variables described by joint PDF.

If $X(t)$ is WSS, $\{X(n/2W)\}$ is WSS discrete r.p. with autocorrelation

$$E\left[X\left(\frac{n}{2W}\right)X^*\left(\frac{m}{2W}\right)\right] = R_X\left(\frac{n-m}{2W}\right) = \int_{-W}^{W} S_X(f) e^{j2\pi f \frac{n-m}{2W}} df$$
◊ The *mean square error* of the sampling representation is zero

\[
E \left[ \left| X(t) - \sum_{n=-\infty}^{\infty} X \left( \frac{n}{2W} \right) \text{sinc} \left[ 2W \left( t - \frac{n}{2W} \right) \right] \right|^2 \right] = 0
\]

◊ The sampling representation equals to \( X(t) \) in the *mean-square sense*

◊ If the process \( X(t) \) is filtered white Gaussian noise then it is zero-mean and its power spectrum is flat in the \([-W,W]\) interval.

◊ In this case the samples are uncorrelated, and since they are Gaussian, they are independent as well.
2.8-2 Karhunen-Loeve Expansion

- **K-L Expansion**: an orthonormal expansion applies to random process and results in uncorrelated variables as coefficients.
- There are many ways in which a random process can be expanded in terms of a sequence of random variables and an orthonormal basis.
- However, if we require the additional condition that the random variables be mutually uncorrelated, then the orthonormal bases have to be the solutions of an eigen function problem given by an integral equation whose kernel is the aut-corvariance function of the random process.
- Solving this integral equation results in the orthonormal basis and projecting the random process on this basis results in the sequence of uncorrelated random variables.
2.8-2 Karhunen-Loeve Expansion

- Autocovariance function of a random process $X(t)$
  \[ C_X(t_1, t_2) = R_X(t_1, t_2) - m_X(t_1)m_X(t_2) \]

- In K-L expansion, $X(t)$ can be expanded over an interval of interest $[a, b]$ in terms of an orthonormal basis $\{\phi_n(t)\}$
  \[ \hat{X}(t) = \sum_{n=1}^{\infty} X_n \phi_n(t), \quad a < t < b \]

- $\{X_n\}$ is set of *uncorrelated* coefficients
- $\{\phi_n(t)\}$ are eigenfunctions of $C_X(t_1, t_2)$ with unit norm
  \[ \int_a^b C_X(t_1, t_2) \phi_n(t_2) dt_2 = \lambda_n \phi_n(t_1), \quad a < t_1 < b \]
  \[ \int_a^b |\phi_n(t)|^2 dt = 1 \]
2.8-2 Karhunen-Loeve Expansion

1) \( X_n \) is projection of \( X(t) \) on the bases function \( \phi_n(t) \)
   \[
   X_n = \langle X(t), \phi_n(t) \rangle = \int_a^b X(t)\phi_n^*(t)dt
   \]

2) \( \{X_n\} \) are mutually uncorrelated and
   \[
   \text{COV}[X_n, X_m] = \begin{cases} 
   \lambda_n & n = m \\
   0 & n \neq m 
   \end{cases}
   \]

3) \( E[\hat{X}(t)] = E[X(t)] = m_X(t) \), \( a < t < b \)

4) \( \hat{X}(t) \) equals to \( X(t) \) in the mean-square sense:
   \[
   E\left[\left|X(t) - \hat{X}(t)\right|^2\right] = 0, \quad a < t < b
   \]

5) **Mercer’s theorem**:
   \[
   C_X(t_1, t_2) = \sum_{n=1}^{\infty} \lambda_n \phi_n(t_1)\phi_n(t_2), \quad a < t_1, \ t_2 < b
   \]
6) \( \{\phi_n(t)\} \) form a complete basis for expansion of all signals with finite energy in the interval \([a,b]\); i.e., if \( g(t) \) is

\[
\int_a^b |g(t)|^2 dt < \infty
\]

it can be expanded by

\[
g(t) = \sum_{n=1}^\infty g_n \phi_n(t), \ a < t < b
\]

where

\[
g_n = \langle g(t), \phi_n(t) \rangle = \int_a^b g(t)\phi_n^*(t)dt
\]

- K-L expansion applies to WSS or non-stationary processes
- If \( X(t) \) is Gaussian process, \( \{X_n\} \) are indep. Gaussian r.v.s
Chapter 2.9 : Bandpass and Lowpass Random Process

Wireless Information Transmission System Lab.
Institute of Communications Engineering
National Sun Yat-sen University
Bandpass and Lowpass Random Process

- A WSS random process $X(t)$ is **lowpass** if $S_X(f)$ is located around zero frequency.
- A WSS random process $X(t)$ is **bandpass** if $X(t)$ is real, zero-mean and $S_X(f)$ is located around $f = \pm f_0$.
- The *in-phase* and *quadrature* components of a **bandpass** $X(t)$
  
  \[
  X_i(t) = X(t) \cos 2\pi f_0 t + \hat{X}(t) \sin 2\pi f_0 t \\
  X_q(t) = \hat{X}(t) \cos 2\pi f_0 t - X(t) \sin 2\pi f_0 t
  \]

- $X_i(t)$ and $X_q(t)$ are *jointly WSS zero-mean* processes.
- $X_i(t)$ and $X_q(t)$ have the *same power spectral density*.
- $X_i(t)$ and $X_q(t)$ are both **lowpass** processes, i.e. their power spectral density is located around $f = 0$. 

179
Define the lowpass equivalent process $X_l(t)$ as

$$X_l(t) = X_i(t) + jX_q(t)$$

- $X_l(t)$ is a proper random process.

- Since $X(t)$ by assumption is zero-mean, so is its Hilbert transform.

  - This is obvious since the Hilbert transform is just a filtering operation.
  - It is clear that $X_i(t)$ and $X_q(t)$ are both zero-mean processes.
Bandpass and Lowpass Random Process

◊ Autocorrelation of $X_i(t) = X(t)\cos 2\pi f_0 t + \hat{X}(t)\sin 2\pi f_0 t$

$$R_{X_i}(t+\tau,t) = E[X_i(t+\tau)X_i(t)]$$

$$= E[(X(t+\tau)\cos 2\pi f_0 (t+\tau) + \hat{X}(t+\tau)\sin 2\pi f_0 (t+\tau))$$

$$\times (X(t)\cos 2\pi f_0 t + \hat{X}(t)\sin 2\pi f_0 t)]$$

$$= R_X(\tau)\cos 2\pi f_0 (t+\tau)\cos 2\pi f_0 t$$

$$+ R_{XX}(t+\tau,t)\cos 2\pi f_0 (t+\tau)\sin 2\pi f_0 t$$

$$+ R_{XX}(t+\tau,t)\sin 2\pi f_0 (t+\tau)\cos 2\pi f_0 t$$

$$+ R_{XX}(t+\tau,t)\sin 2\pi f_0 (t+\tau)\sin 2\pi f_0 t$$
Bandpass and Lowpass Random Process

- Since $\hat{X}(t) = X(t) \ast (1/\pi t)$, $\hat{X}(t)$ and $X(t)$ are joint WSS
  \[ R_{\hat{X}\hat{X}}(\tau) = -\hat{R}_X(\tau) \]
  \[ R_{\hat{X}X}(\tau) = \hat{R}_X(\tau) \quad \text{All are functions of } \tau. \]
  \[ R_{\hat{X}\hat{X}}(\tau) = R_X(\tau) \quad \text{See Problem 2.56.} \]

- Thus,
  \[ R_{X_i}(\tau) = R_X(\tau)\cos(2\pi f_0 \tau) + \hat{R}_X(\tau)\sin(2\pi f_0 \tau) \]

- Similarly,
  \[ R_{X_q}(\tau) = R_{X_i}(\tau) = R_X(\tau)\cos(2\pi f_0 \tau) + \hat{R}_X(\tau)\sin(2\pi f_0 \tau) \]
  \[ R_{X_iX_q}(\tau) = -R_{X_qX_i}(\tau) = R_X(\tau)\sin(2\pi f_0 \tau) - \hat{R}_X(\tau)\cos(2\pi f_0 \tau) \]

- $X_i(t)$ and $X_q(t)$ are jointly WSS

- $X_i(t)$ and $X_q(t)$ have equal autocorrelation functions \( \Rightarrow \) equal PSD
Because $F\left[ \hat{R}_X(\tau) \right] = -j \text{sgn}(f)S_X(f)$ and

\[
R_{X_q}(\tau) = R_{X_i}(\tau) = R_X(\tau) \cos(2\pi f_0 \tau) + \hat{R}_X(\tau) \sin(2\pi f_0 \tau)
\]

\[
R_{X_iX_q}(\tau) = -R_{X_qX_i}(\tau) = R_X(\tau) \sin(2\pi f_0 \tau) - \hat{R}_X(\tau) \cos(2\pi f_0 \tau)
\]

- $S_{X_i}(f) = S_{X_q}(f) = \begin{cases} S_X(f + f_0) + S_X(f - f_0) & |f| < f_0 \\ 0 & \text{otherwise} \end{cases}$

- $S_{X_iX_q}(f) = -S_{X_qX_i}(f) = \begin{cases} j[S_X(f + f_0) - S_X(f - f_0)] & |f| < f_0 \\ 0 & \text{otherwise} \end{cases}$

- $X_i(t)$ and $X_q(t)$ are both lowpass processes
- If $S_X(f)$ is symmetric around $\pm f_0$, $S_X(f + f_0) = S_X(f - f_0), |f| < f_0$
  \[ S_{X_iX_q}(f) = 0, R_{X_iX_q}(\tau) = 0 \]
  \[ X_i(t) \text{ and } X_q(t) \text{ are uncorrelated} \]
Bandpass and Lowpass Random Process

- **Lowpass equivalent** of $X(t)$ is defined by
  \[ X_l(t) = X_i(t) + jX_q(t) \]
- $X_l(t)$ is lowpass
- \[ R_{X_l}(\tau) = 2R_{X_i}(\tau) + 2jR_{X_qX_i}(\tau) = 2[R_X(\tau) + j\hat{R}_X(\tau)]e^{-j2\pi f_0\tau} \]
  \[ \Rightarrow R_{X_l}(\tau) = 2 \cdot (\text{lowpass equivalent of } R_X(\tau)) \]
- \[ S_{X_l}(f) = \begin{cases} 4S_X(f + f_0) & |f| < f_0 \\ 0 & \text{otherwise} \end{cases} \]
- $X(t)$ is Gaussian $\Rightarrow X_l(t), X_i(t)$ and $X_q(t)$ are jointly Gaussian
- $X_l(t)$ is Gaussian, zero-mean and proper
  \[ S_X(f + f_0) = S_X(f - f_0), \quad \text{for } |f| < f_0 \]
  \[ \Rightarrow X_l(t) \text{ and } X_q(t) \text{ are independent} \]