Chapter 5
Random Variables and Processes

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

Table of Contents

- 5.1 Introduction
- 5.2 Probability
- 5.3 Random Variables
- 5.4 Statistical Averages
- 5.5 Random Processes
- 5.6 Mean, Correlation and Covariance Functions
- 5.7 Transmission of a Random Process through a Linear Filter
- 5.8 Power Spectral Density
- 5.9 Gaussian Process
- 5.10 Noise
- 5.11 Narrowband Noise

5.1 Introduction

- Fourier transform is a mathematical tool for the representation of *deterministic signals*.
- *Deterministic signals*: the class of signals that may be modeled as completely specified functions of time.
- A signal is "random" if it is not possible to predict its precise value in advance.
- A *random process* consists of an ensemble (family) of *sample functions*, each of which varies randomly with time.
- A *random variable* is obtained by observing a random process at a fixed instant of time.

5.2 Probability

- Probability theory is rooted in phenomena that, explicitly or implicitly, can be modeled by an experiment with an outcome that is subject to chance.
  - Example: Experiment may be the observation of the result of tossing a fair coin. In this experiment, the possible outcomes of a trial are "heads" or "tails".
- If an experiment has K possible outcomes, then for the kth possible outcome we have a point called the *sample point,* which we denote by s_k. With this basic framework, we make the following definitions:
  - The set of all possible outcomes of the experiment is called the *sample space*, which we denote by S.
  - An *event* corresponds to either a single sample point or a set of sample points in the space S.

5.2 Probability

- A single sample point is called an *elementary event*.
- The entire sample space S is called the *sure event*; and the *null set* ϕ is called the *null* or *impossible event*.
- Two events are *mutually exclusive* if the occurrence of one event precludes the occurrence of the other event.

- A *probability measure* **P** is a function that assigns a non-negative number to an event A in the sample space S and satisfies the following three properties (axioms):

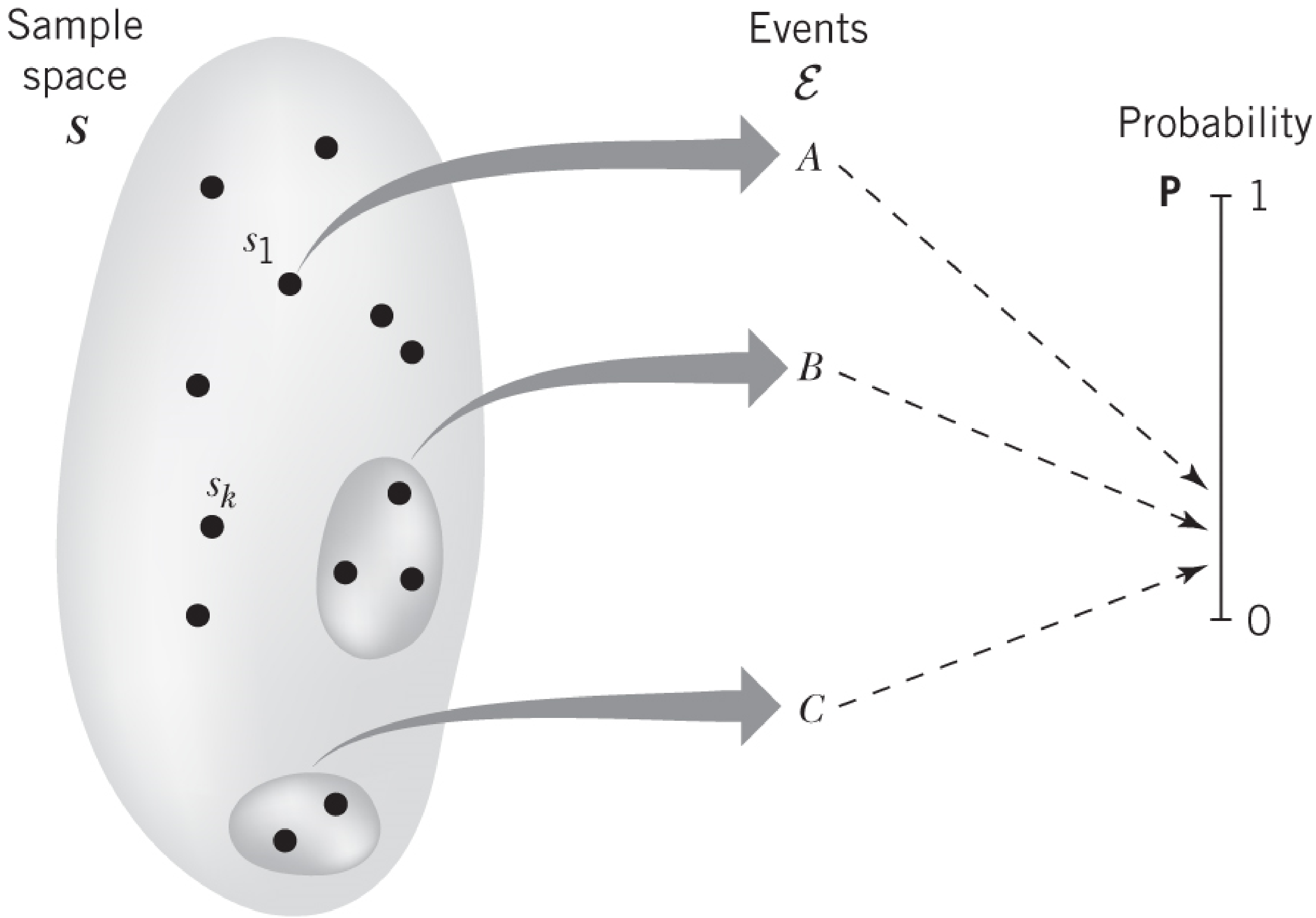
1. $0 \le \mathbf{P}[A] \le 1$ (5.1)

2. $\mathbf{P}[S] = 1$ (5.2)

3. If A and B are two mutually exclusive events, then

$$\mathbf{P}[A \cup B] = \mathbf{P}[A] + \mathbf{P}[B] \quad (5.3)$$

5.2 Probability



5.2 Probability

- The following properties of probability measure **P** may be derived from the above axioms:

1. $\mathbf{P}\left[\overline{A}\right] = 1 - \mathbf{P}\left[A\right]$ (5.4)

2. When events A and B are not mutually exclusive:

$$\mathbf{P}\left[A \cup B\right] = \mathbf{P}\left[A\right] + \mathbf{P}\left[B\right] - \mathbf{P}\left[A \cap B\right] \quad (5.5)$$

3. If $A_1, A_2, \ldots, A_m$ are mutually exclusive events that include all possible outcomes of the random experiment, then

$$\mathbf{P}\left[A_1\right] + \mathbf{P}\left[A_2\right] + \ldots + \mathbf{P}\left[A_m\right] = 1 \quad (5.6)$$

5.2 Probability

- Let $\mathbf{P}[B|A]$ denote the probability of event B, given that event A has occurred. The probability $\mathbf{P}[B|A]$ is called the *conditional probability* of B given A.
- $\mathbf{P}[B|A]$ is defined by

$$\mathbf{P}[B|A] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[A]} \qquad (5.7)$$

- Bayes' rule
  - We may write Eq.(5.7) as $\mathbf{P}[A \cap B] = \mathbf{P}[B|A]\mathbf{P}[A]$ (5.8)
  - It is apparent that we may also write $\mathbf{P}[A \cap B] = \mathbf{P}[A|B]\mathbf{P}[B]$ (5.9)
  - From Eqs.(5.8) and (5.9), provided $\mathbf{P}[A] \neq 0$, we may determine $\mathbf{P}[B|A]$ by using the relation

$$\mathbf{P}[B|A] = \frac{\mathbf{P}[A|B]\mathbf{P}[B]}{\mathbf{P}[A]} \qquad (5.10)$$

5.2 Conditional Probability

- Suppose that the condition probability $\mathbf{P}[B|A]$ is simply equal to the elementary probability of occurrence of event B, that is

$$\mathbf{P}[B|A] = \mathbf{P}[B] \Longrightarrow \mathbf{P}[A \cap B] = \mathbf{P}[A]\mathbf{P}[B]$$

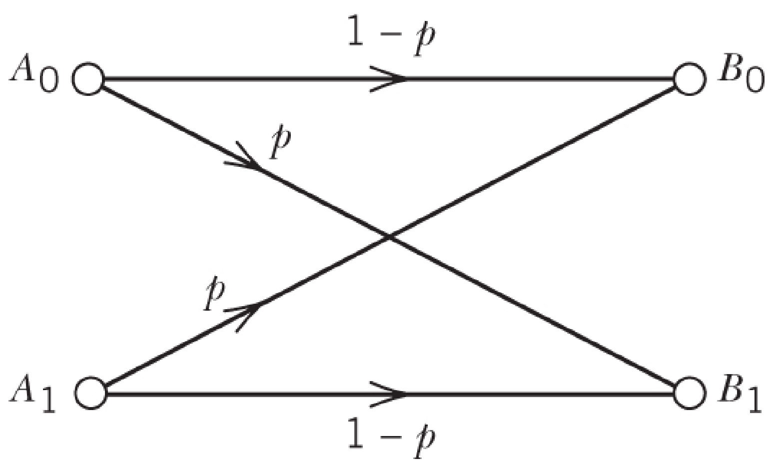
so that

$$\mathbf{P}[A|B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]} = \frac{\mathbf{P}[A]\mathbf{P}[B]}{\mathbf{P}[B]} = \mathbf{P}[A] \qquad (5.13)$$

  - Events A and B that satisfy this condition are said to be *statistically independent*.

5.2 Conditional Probability

- Example 5.1 Binary Symmetric Channel
  - This channel is said to be *discrete* in that it is designed to handle discrete messages.
  - The channel is *memoryless* in the sense that the channel output at any time depends only on the channel input at that time.
  - The channel is *symmetric*, which means that the probability of receiving symbol 1 when 0 is sent is the same as the probability of receiving symbol 0 when symbol 1 is sent.



5.2 Conditional Probability

- Example 5.1 Binary Symmetric Channel (continued)
  - The *a priori probabilities* of sending binary symbols 0 and 1:

$$\mathbf{P}\left[A_0\right] = p_0 \qquad \mathbf{P}\left[A_1\right] = p_1$$

  - The *conditional probabilities of error*:

$$\mathbf{P}\left[B_1 \middle| A_0\right] = \mathbf{P}\left[B_0 \middle| A_1\right] = p$$

  - The probability of receiving symbol 0 is given by:

$$\mathbf{P}\left[B_0\right] = \mathbf{P}\left[B_0 \middle| A_0\right]\mathbf{P}\left[A_0\right] + \mathbf{P}\left[B_0 \middle| A_1\right]\mathbf{P}\left[A_1\right] = (1-p)\,p_0 + pp_1$$

  - The probability of receiving symbol 1 is given by:

$$\mathbf{P}\left[B_1\right] = \mathbf{P}\left[B_1 \middle| A_0\right]\mathbf{P}\left[A_0\right] + \mathbf{P}\left[B_1 \middle| A_1\right]\mathbf{P}\left[A_1\right] = pp_0 + (1-p)\,p_1$$

5.2 Conditional Probability

- Example 5.1 Binary Symmetric Channel (continued)
  - The *a posteriori probabilities* $\mathbf{P}[A_0|B_0]$ and $\mathbf{P}[A_1|B_1]$:

$$\mathbf{P}\left[A_0 \middle| B_0\right] = \frac{\mathbf{P}\left[B_0 \middle| A_0\right]\mathbf{P}\left[A_0\right]}{\mathbf{P}\left[B_0\right]} = \frac{(1-p)\,p_0}{(1-p)\,p_0 + pp_1}$$

$$\mathbf{P}\left[A_1 \middle| B_1\right] = \frac{\mathbf{P}\left[B_1 \middle| A_1\right]\mathbf{P}\left[A_1\right]}{\mathbf{P}\left[B_1\right]} = \frac{(1-p)\,p_1}{pp_0 + (1-p)\,p_1}$$

5.3 Random Variables

- We denote the *random variable* as $X(s)$ or just X.
- X is a function, s is the outcome of the experiment.
- Random variable may be discrete or continuous.
- Consider the random variable X and the probability of the event $X \leq x$. We denote this probability by $P[X \leq x]$.
- To simplify our notation, we write

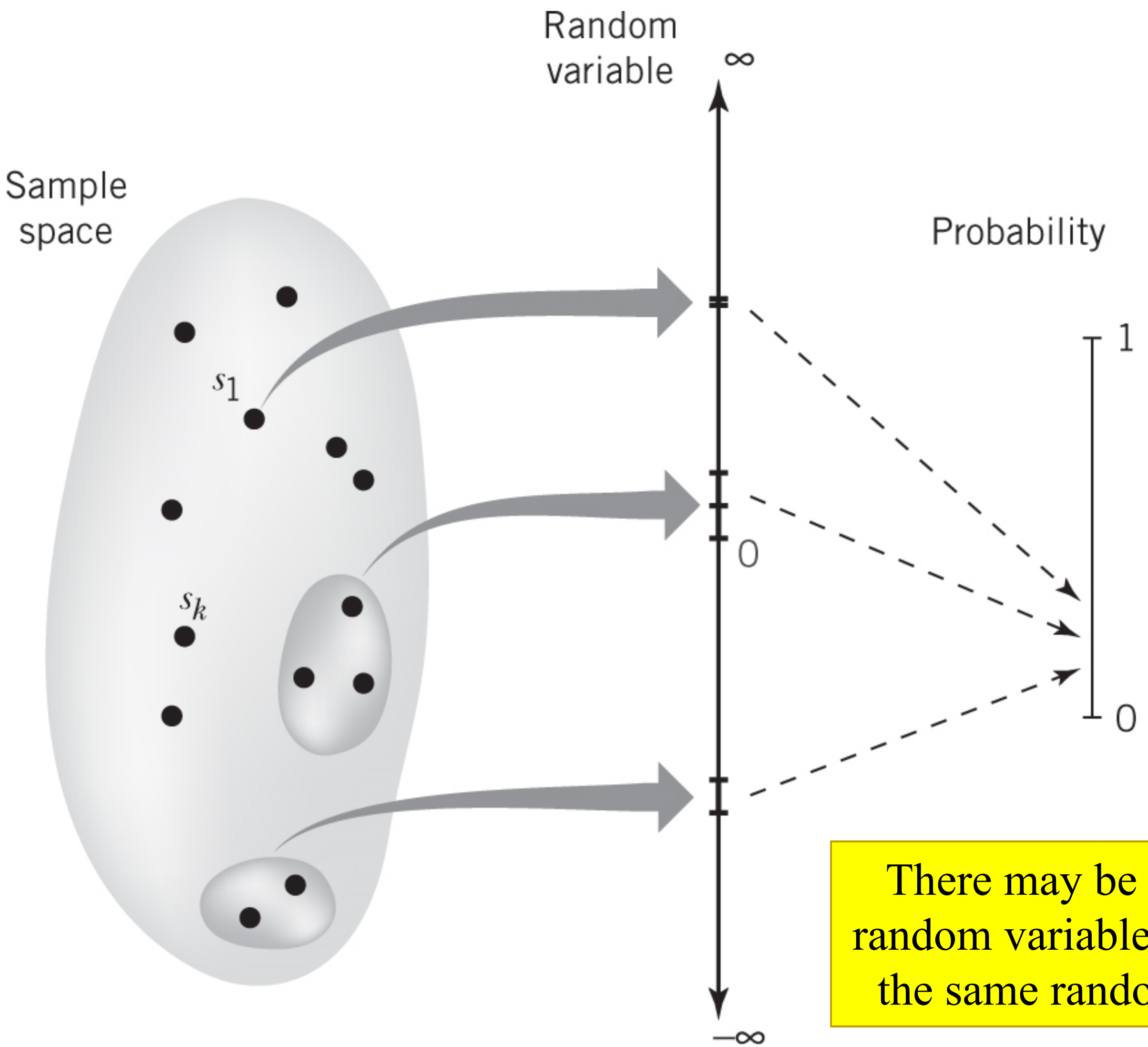
$$F_X(x) = \mathbf{P}[X \leq x] \qquad (5.15)$$

  - The function $F_X(x)$ is called the *cumulative distribution function* (cdf) or simply the *distribution function* of the random variable X.
- The distribution function $F_X(x)$ has the following properties:

1. $0 \leq F_x(x) \leq 1$

2. $F_x(x_1) \leq F_x(x_2) \quad \text{if } x_1 < x_2$

5.3 Random Variables



There may be more than one random variable associated with the same random experiment.

5.3 Random Variables

- If the distribution function is continuously differentiable, then

$$f_X(x) = \frac{d}{dx} F_X(x) \qquad (5.17)$$

  - $f_X(x)$ is called the *probability density function* (pdf) of the random variable X.

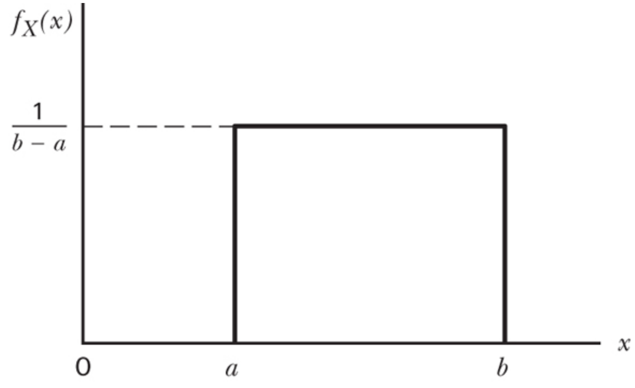
- Probability of the event $x_1 < X \le x_2$ equals

$$\begin{aligned}\mathbf{P}[x_1 < X \le x_2] &= \mathbf{P}[X \le x_2] - \mathbf{P}[X \le x_1] \\ &= F_X(x_2) - F_X(x_1) \xrightarrow{x_1 = -\infty} F_X(x) = \int_{-\infty}^{x} f_X(\xi)\, d\xi \qquad (5.19) \\ &= \int_{x_1}^{x_2} f_X(x)\, dx\end{aligned}$$

- Probability density function must always be a nonnegative function, and with a total area of one.

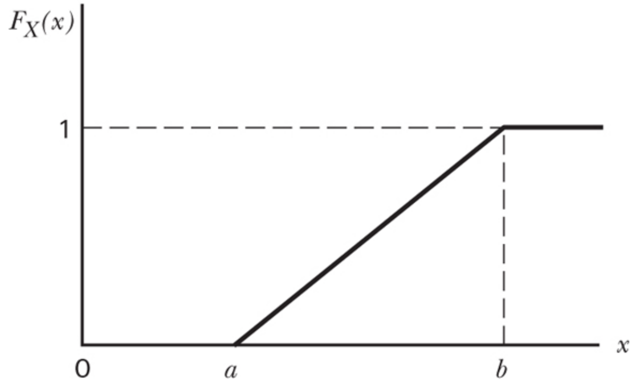
- Example 5.2 Uniform Distribution

$$f_X(x) = \begin{cases} 0, & x \le a \\ \dfrac{1}{b-a}, & a < x \le b \\ 0, & x > b \end{cases}$$



(a)

$$F_X(x) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x \le b \\ 0, & x > b \end{cases}$$



(b)

5.3 Random Variables

- Several Random Variables
  - Consider two random variables X and Y. We define the *joint distribution function* $F_{X,Y}(x,y)$ as the probability that the random variable X is less than or equal to a specified value x and that the random variable Y is less than or equal to a specified value y.

$$F_{X,Y}(x,y) = \mathbf{P}[X \le x, Y \le y] \qquad (5.23)$$

  - Suppose that joint distribution function $F_{X,Y}(x,y)$ is continuous everywhere, and that the partial derivative

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} \qquad (5.24)$$

exists and is continuous everywhere. We call the function $f_{X,Y}(x,y)$ the *joint probability density function* of the random variables X and Y.

5.3 Random Variables

- Several Random Variables
  - The joint distribution function $F_{X,Y}(x,y)$ is a monotone-nondecreasing function of both x and y.
  - $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(\xi,\eta)\,d\xi\,d\eta = 1$
  - *Marginal density* $f_X(x)$

$$F_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X,Y}(\xi,\eta)\,d\xi\,d\eta \longrightarrow f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,\eta)\,d\eta \quad (5.27)$$

  - Suppose that X and Y are two continuous random variables with joint probability density function $f_{X,Y}(x,y)$. The *conditional probability density function of Y given that $X = x$* is defined by

$$f_Y(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad (5.28)$$

5.3 Random Variables

- Several Random Variables
  - If the random variable X and Y are *statistically independent*, then knowledge of the outcome of X can in no way affect the distribution of Y.

$$f_Y(y|x) = f_Y(y) \xrightarrow{\text{by}(5.28)} f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad (5.32)$$

$$\Downarrow$$

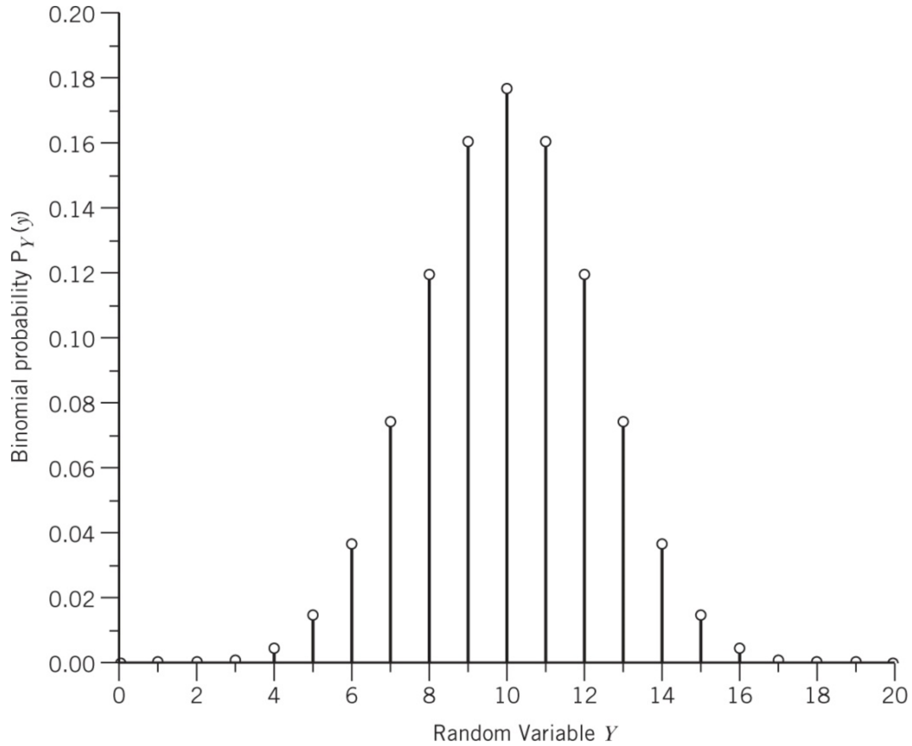
$$\mathbf{P}[X \in A,\ Y \in B] = \mathbf{P}[X \in A]\mathbf{P}[Y \in B] \quad (5.33)$$

- Example 5.3 Binomial Random Variable
  - Consider a sequence of coin-tossing experiments where the probability of a head is p and let X_n be the Bernoulli random variable representing the outcome of the nth toss.
  - Let Y be the number of heads that occur on N tosses of the coins:

$$Y = \sum_{n=1}^{N} X_n$$

$$\mathbf{P}[Y = y] = \binom{N}{y} p^y (1-p)^{N-y}$$

$$\binom{N}{y} = \frac{N!}{y!(N-y)!}$$



5.4 Statistical Averages

- The *expected value* or *mean* of a random variable X is defined by

$$\mu_x = \mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx \qquad (5.36)$$

- Function of a Random Variable
  - Let X denote a random variable, and let $g(X)$ denote a real-valued function defined on the real line. We denote as

$$Y = g(X) \qquad (5.37)$$

  - To find the expected value of the random variable Y.

$$\mathbf{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y)\,dy \longrightarrow \mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_x(x)\,dx \qquad (5.38)$$

5.4 Statistical Averages

- Example 5.4 Cosinusoidal Random Variable
  - Let $Y=g(X)=\cos(X)$
  - X is a random variable uniformly distributed in the interval $(-\pi, \pi)$

$$f_X(x) = \begin{cases} \dfrac{1}{2\pi}, & -\pi < x < \pi \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \mathbf{E}[Y] &= \int_{-\pi}^{\pi} (\cos x)\left(\frac{1}{2\pi}\right) dx \\ &= \frac{1}{2\pi} \sin x \Big|_{x=-\pi}^{\pi} \\ &= 0 \end{aligned}$$

5.4 Statistical Averages

- Moments
  - For the special case of $g(X) = X^n$, we obtain the *nth moment* of the probability distribution of the random variable X; that is

$$\mathbf{E}\left[X^n\right] = \int_{-\infty}^{\infty} x^n f_X(x)\,dx \qquad (5.39)$$

  - *Mean-square value* of X:

$$\mathbf{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx \qquad (5.40)$$

  - The *nth central moment* is

$$\mathbf{E}\left[\left(X - \mu_X\right)^n\right] = \int_{-\infty}^{\infty} \left(x - \mu_X\right)^n f_X(x)\,dx \qquad (5.41)$$

5.4 Statistical Averages

- For $n = 2$ the second central moment is referred to as the *variance* of the random variable X, written as

$$\mathrm{Var}[X] = \mathbf{E}\left[(X-\mu_X)^2\right] = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)\,dx \quad (5.42)$$

  - The variance of a random variable X is commonly denoted as σ_X^2.
  - The square root of the variance is called the *standard deviation* of the random variable X.

-

$$\begin{aligned}\sigma_X^2 &= \mathrm{Var}[X] = \mathbf{E}\left[(X-\mu_X)^2\right] \\ &= \mathbf{E}\left[X^2 - 2\mu_X X + \mu_X^2\right] \\ &= \mathbf{E}\left[X^2\right] - 2\mu_X \mathbf{E}[X] + \mu_X^2 \\ &= \mathbf{E}\left[X^2\right] - \mu_X^2 \qquad (5.44)\end{aligned}$$

5.4 Statistical Averages

- *Characteristic function* $\phi_X(\upsilon)$ is defined as the expectation of the complex exponential function $\exp(j\upsilon X)$, as shown by

$$\psi_X(j\upsilon) = \mathbf{E}\left[\exp(j\upsilon X)\right] = \int_{-\infty}^{\infty} f_X(x)\exp(j\upsilon X)\,dx \qquad (5.45)$$

  - In other words , the characteristic function $\phi_X(\upsilon)$ is the Fourier transform of the probability density function $f_X(x)$.

- Analogous with the inverse Fourier transform:

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi_X(j\upsilon)\exp(-j\upsilon X)\,d\upsilon \qquad (5.46)$$

5.4 Statistical Averages

- Characteristic functions
  - First moment (mean) can be obtained by:

$$E(X) = m_x = -j \left. \frac{d\psi(jv)}{dv} \right|_{v=0}$$

  - Since the differentiation process can be repeated, n-th moment can be calculated by:

$$E(X^n) = (-j)^n \left. \frac{d^n\psi(jv)}{dv^n} \right|_{v=0}$$

5.4 Statistical Averages

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

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

$$=E\left[\prod_{i=1}^{n}\left(e^{jvX_i}\right)\right]=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\prod_{i=1}^{n}e^{jvx_i}\right)f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)dx_1dx_2\ldots dx_n$$

Since the random variables are statistically independent,

$$f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\ldots f_{X_n}(x_n) \quad \Rightarrow \quad \psi_Y(jv)=\prod_{i=1}^{n}\psi_{X_i}(jv)$$

If X_i are iid (independent and identically distributed)

$$\Rightarrow \quad \psi_Y(jv)=\left[\psi_X(jv)\right]^n$$

5.4 Statistical Averages

- Characteristic functions
  - The PDF of Y is determined from the inverse Fourier transform of $\Psi_Y(jv)$.
  - 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.

- Example 5.5 Gaussian Random Variable
  - The probability density function of such a Gaussian random variable is defined by:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2} \right), \quad -\infty < x < \infty$$

  - The characteristic function of a Gaussian random variable with mean m_x and variance σ^2 is (Problem 5.1):

$$\psi(jv) = \int_{-\infty}^{\infty} e^{jvx} \left[\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m_x)^2/2\sigma^2} \right] dx = e^{jvm_x - (1/2)v^2\sigma^2}$$

  - It can be shown that the central moments of a Gaussian random variable are given by:

$$E\left[(X-m_x)^k\right] = \mu_k = \begin{cases} 1 \cdot 3 \cdots (k-1)\sigma^k & (\text{even } k) \\ 0 & (\text{odd } k) \end{cases}$$

5.4 Statistical Averages

- Example 5.5 Gaussian Random Variable (cont.)
  - The sum of n statistically independent Gaussian random variables is also a Gaussian random variable.
  - Proof:

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

$$\psi_Y(jv) = \prod_{i=1}^{n} \psi_{X_i}(jv) = \prod_{i=1}^{n} e^{jvm_i - v^2\sigma_i^2/2} = e^{jvm_y - v^2\sigma_y^2/2}$$

where $m_y = \sum_{i=1}^{n} m_i$ and $\sigma_y^2 = \sum_{i=1}^{n} \sigma_i^2$

Therefore, Y is Gaussian-distributed with mean m_y and variance σ_y^2.

5.4 Statistical Averages

- Joint Moments
  - Consider next a pair of random variables X and Y. A set of statistical averages of importance in this case are the *joint moments*, namely, the expected value of $X^i Y^k$, where i and k may assume any positive integer values. We may thus write

$$\mathbf{E}\left[X^i Y^k\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i y^k f_{X,Y}(x,y)\,dxdy \qquad (5.51)$$

  - A joint moment of particular importance is the *correlation* defined by $\mathbf{E}[XY]$, which corresponds to $i = k = 1$.
  - *Covariance* of X and Y :

$$\mathrm{Cov}[XY] = \mathbf{E}\left[\left(X - \mathbf{E}[X]\right)\left(Y - \mathbf{E}[Y]\right)\right] = \mathbf{E}[XY] - \mu_X \mu_Y \quad (5.53)$$

5.4 Statistical Averages

- *Correlation coefficient* of X and Y :

$$\rho = \frac{\text{Cov}[XY]}{\sigma_X \sigma_Y} \qquad (5.54)$$

  - σ_X and σ_Y denote the variances of X and Y.

- We say X and Y are *uncorrelated* if and only if $\text{Cov}[XY] = 0$.
  - Note that if X and Y are *statistically independent*, then they are *uncorrelated*.
  - The converse of the above statement is not necessarily true.

- We say X and Y are *orthogonal* if and only if $\text{E}[XY] = 0$.

5.4 Statistical Averages

- Example 5.6 Moments of a Bernoulli Random Variable
  - Consider the coin-tossing experiment where the probability of a head is p. Let X be a random variable that takes the value 0 if the result is a tail and 1 if it is a head. We say that X is a *Bernoulli random variable*.

$$\mathbf{P}(X=x)=\begin{cases}1-p & x=0\\ p & x=1\\ 0 & \text{otherwise}\end{cases}$$

$$\mathbf{E}[X]=\sum_{k=0}^{1}k\mathbf{P}(X=k)=0\cdot(1-p)+1\cdot p=p$$

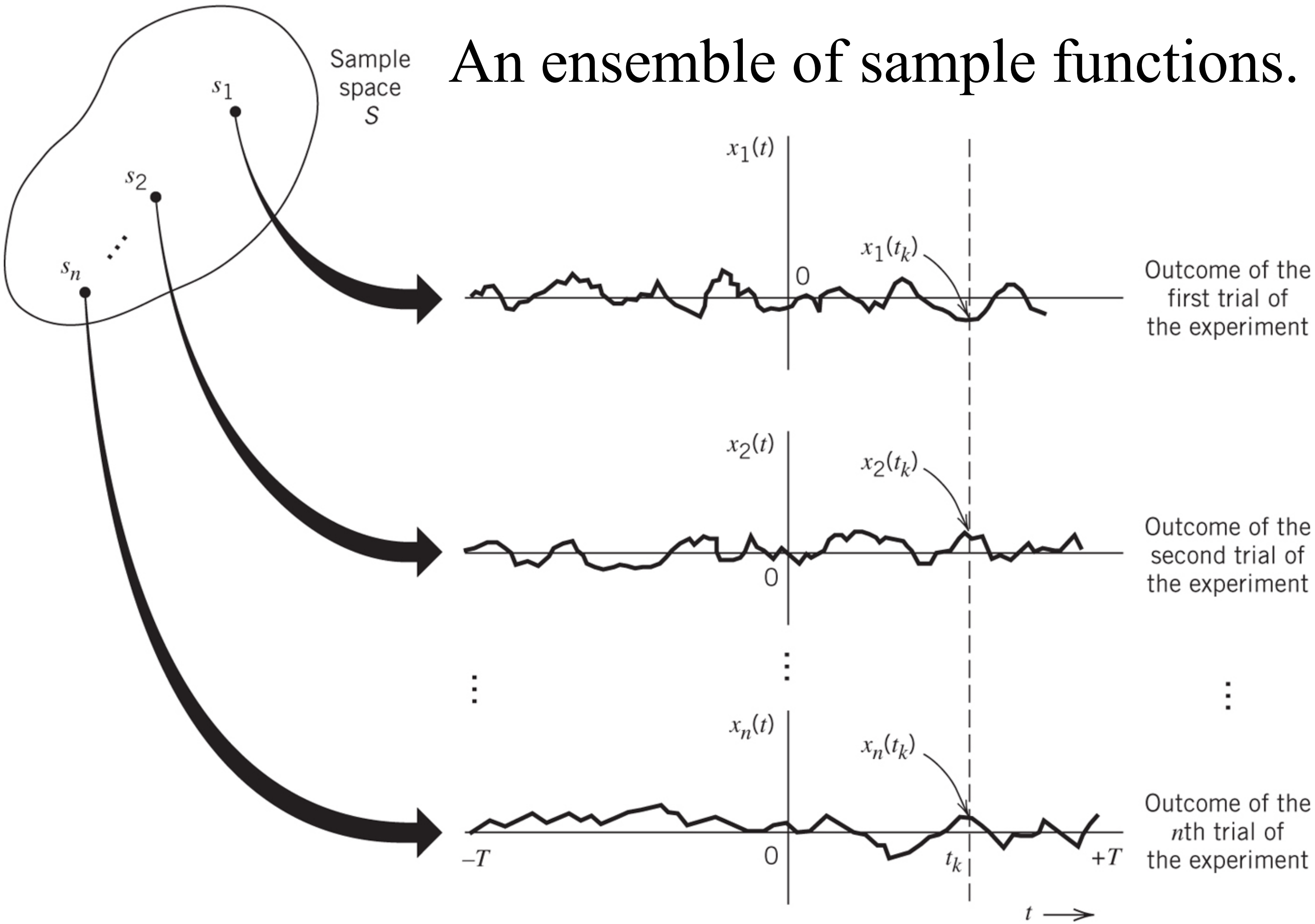
$$\sigma_X^2=\sum_{k=0}^{1}(k-\mu_X)^2\,\mathbf{P}[X=k]$$
$$=(0-p)^2(1-p)+(1-p)^2$$
$$=p(1-p)$$

$$\mathbf{E}[X_jX_k]=\begin{cases}\mathbf{E}[X_j]\mathbf{E}[X_k] & j\neq k\\ \mathbf{E}[X_j^2] & j=k\end{cases}$$
$$=\begin{cases}p^2 & j\neq k\\ p & j=k\end{cases}$$

where the $\mathbf{E}[X_j^2]=\sum_{k=0}^{1}k^2\mathbf{P}[X=k]$.

5.5 Random Processes

An ensemble of sample functions.



For a fixed time instant t_k, $\{x_1(t_k), x_2(t_k), \ldots x_n(t_k)\} = \{X(t_k, s_1), X(t_k, s_2), \ldots, X(t_k, s_n)\}$ constitutes a random variable.

5.5 Random Processes

- At any given time instant, the value of a stochastic process is a random variable indexed by the parameter t. We denote such a process by $X(t)$.
- In general, the parameter t is continuous, whereas X may be either continuous or discrete, depending on the characteristics of the source that generates the stochastic process.
- The noise voltage generated by a single resistor or a single information source represents a single realization of the stochastic process. It is called a *sample function*.

5.5 Random Processes

- The set of all possible sample functions constitutes an *ensemble* of sample functions or, equivalently, the *stochastic process* $X(t)$.
- In general, the number of sample functions in the ensemble is assumed to be extremely large; often it is infinite.
- Having defined a stochastic process $X(t)$ as an ensemble of sample functions, we may consider the values of the process at any set of time instants $t_1>t_2>t_3>\ldots>t_n$, where n is any positive integer.
- In general, the random variables $X_{t_i} \equiv X\left(t_i\right), i=1,2,\ldots,n,$ are characterized statistically by their joint PDF $f_X\left(x_{t_1}, x_{t_2}, \ldots, x_{t_n}\right)$.

5.5 Random Processes

- Stationary stochastic processes
  - Consider another set of n random variables $X_{t_i+t} \equiv X(t_i+t)$, $i=1,2,\ldots,n$, where t is an arbitrary time shift. These random variables are characterized by the joint PDF $f_X\left(x_{t_1+t}, x_{t_2+t}, \ldots, x_{t_n+t}\right)$.
  - The joint PDFs of the random variables X_{t_i} and $X_{t_i+t}, i=1,2,\ldots,n,$ may or may not be identical. When they are identical, i.e., when

$$f_X\left(x_{t_1}, x_{t_2}, \ldots, x_{t_n}\right) = f_X\left(x_{t_1+t}, x_{t_2+t}, \ldots, x_{t_n+t}\right)$$

    for all t and all n, it is said to be stationary in the strict sense(SSS).
  - When the joint PDFs are different, the stochastic process is *non-stationary*.

5.5 Random Processes

- Averages for a stochastic process are called *ensemble averages*.
- The nth moment of the random variable X_{t_i} is defined as:

$$E\left(X_{t_i}^n\right)=\int_{-\infty}^{\infty} x_{t_i}^n f_X\left(x_{t_i}\right) dx_{t_i}$$

- In general, the value of the nth moment will depend on the time instant t_i if the PDF of X_{t_i} depends on t_i.
- When the process is stationary, $f_X\left(x_{t_i+t}\right)=f_X\left(x_{t_i}\right)$ for all t. Therefore, the PDF is independent of time, and, as a consequence, the nth moment is independent of time.

5.5 Random Processes

- Two random variables: $X_{t_i} \equiv X(t_i), i = 1, 2.$
  - The correlation is measured by the *joint moment*:

$$E\left(X_{t_1} X_{t_2}\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_{t_1} x_{t_2} f_X\left(x_{t_1}, x_{t_2}\right) dx_{t_1} dx_{t_2}$$

  - Since this joint moment depends on the time instants t_1 and t_2, it is denoted by $R_X(t_1, t_2)$.
  - $R_X(t_1, t_2)$ is called the *auto-correlation function* of the stochastic process.
  - For a stationary stochastic process, the joint moment is:

$$E(X_{t_1} X_{t_2}) = R_X(t_1, t_2) = R_X(t_1 - t_2) = R_X(\tau)$$

  - $R_X(-\tau) = E(X_{t_1} X_{t_1+\tau}) = E(X_{t_1+\tau} X_{t_1}) = E(X_{t_1'} X_{t_1'-\tau}) = R_X(\tau)$
  - Average power in the process $X(t)$: $R_X(0) = E(X_t^2)$.

5.5 Random Processes

- Wide-sense stationary (WSS)
  - A *wide-sense stationary process* has the property that the mean value of the process is independent of time (a constant) and where the autocorrelation function satisfies the condition that $R_X(t_1,t_2)=R_X(t_1-t_2)$.
  - Wide-sense stationarity is a less stringent condition than strict-sense stationarity.

5.5 Random Processes

- Auto-covariance function
  - The *auto-covariance function* of a stochastic process is defined as:

$$Cov\left(X_{t_1}, X_{t_2}\right) = E\left\{\left[X_{t_1} - m\left(t_1\right)\right]\left[X_{t_2} - m\left(t_2\right)\right]\right\}$$
$$= R_X\left(t_1, t_2\right) - m\left(t_1\right) m\left(t_2\right)$$

  - When the process is stationary, the auto-covariance function simplifies to:

$$Cov(X_{t_1}, X_{t_2}) = C_X(t_1 - t_2) = C_X(\tau) = R_X(\tau) - m^2$$

  - For a Gaussian random process, higher-order moments can be expressed in terms of first and second moments. Consequently, a Gaussian random process is completely characterized by its first two moments.

5.6 Mean, Correlation and Covariance Functions

- Consider a random process $X(t)$. We define the mean of the process $X(t)$ as the expectation of the random variable obtained by observing the process at some time t, as shown by

$$\mu_X(t) = \mathbf{E}[X(t)] = \int_{-\infty}^{\infty} x f_{X(t)}(x)\,dx \qquad (5.57)$$

- A random process is said to be *stationary to first order* if the distribution function (and therefore density function) of $X(t)$ does not vary with time.

$$f_{X(t_1)}(x) = f_{X(t_2)}(x) \text{ for all } t_1 \text{ and } t_2 \;\rightarrow\; \mu_X(t) = \mu_X \text{ for all } t \qquad (5.59)$$

  - The mean of the random process is a constant.
  - The variance of such a process is also constant.

5.6 Mean, Correlation and Covariance Functions

- We define the *autocorrelation function* of the process $X(t)$ as the expectation of the product of two random variables $X(t_1)$ and $X(t_2)$.

$$R_X(t_1, t_2) = \mathbf{E}\left[X(t_1)X(t_2)\right]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1)X(t_2)}(x_1, x_2)\, dx_1 dx_2 \qquad (5.60)$$

- We say a random process $X(t)$ is *stationary to second order* if the joint distribution $f_{X(t_1)X(t_2)}(x_1, x_2)$ depends on the difference between the observation time t_1 and t_2.

$$R_X(t_1, t_2) = R_X(t_2 - t_1) \quad \text{for all } t_1 \text{ and } t_2 \qquad (5.61)$$

- The *autocovariance function* of a stationary random process $X(t)$ is written as

$$C_X(t_1, t_2) = \mathbf{E}\left[\left(X(t_1) - \mu_X\right)\left(X(t_2) - \mu_X\right)\right] = R_X(t_2 - t_1) - \mu_X^2 \quad (5.62)$$

5.6 Mean, Correlation and Covariance Functions

- For convenience of notation, we redefine the autocorrelation function of a stationary process $X(t)$ as

$$R_X(\tau) = \mathbf{E}\left[X(t+\tau)X(t)\right] \text{ for all } t \qquad (5.63)$$

- This autocorrelation function has several important properties:

1. $$R_X(0) = \mathbf{E}\left[X^2(t)\right] \qquad (5.64)$$

2. $$R_X(\tau) = R_X(-\tau) \qquad (5.65)$$

3. $$\left|R_X(\tau)\right| \le R_X(0) \qquad (5.67)$$

- Proof of (5.64) can be obtained from (5.63) by putting $\tau = 0$.

- Proof of (5.65):

$$R_X(\tau) = \mathbf{E}\left[X(t+\tau)X(t)\right] = \mathbf{E}\left[X(t)X(t+\tau)\right] = R_X(-\tau)$$

- Proof of (5.67):

$$\mathbf{E}\left[\left(X(t+\tau) \pm X(t)\right)^2\right] \geq 0$$

$$\rightarrow \mathbf{E}\left[X^2(t+\tau)\right] \pm 2\mathbf{E}\left[X(t+\tau)X(t)\right] + \mathbf{E}\left[X^2(t)\right] \geq 0$$

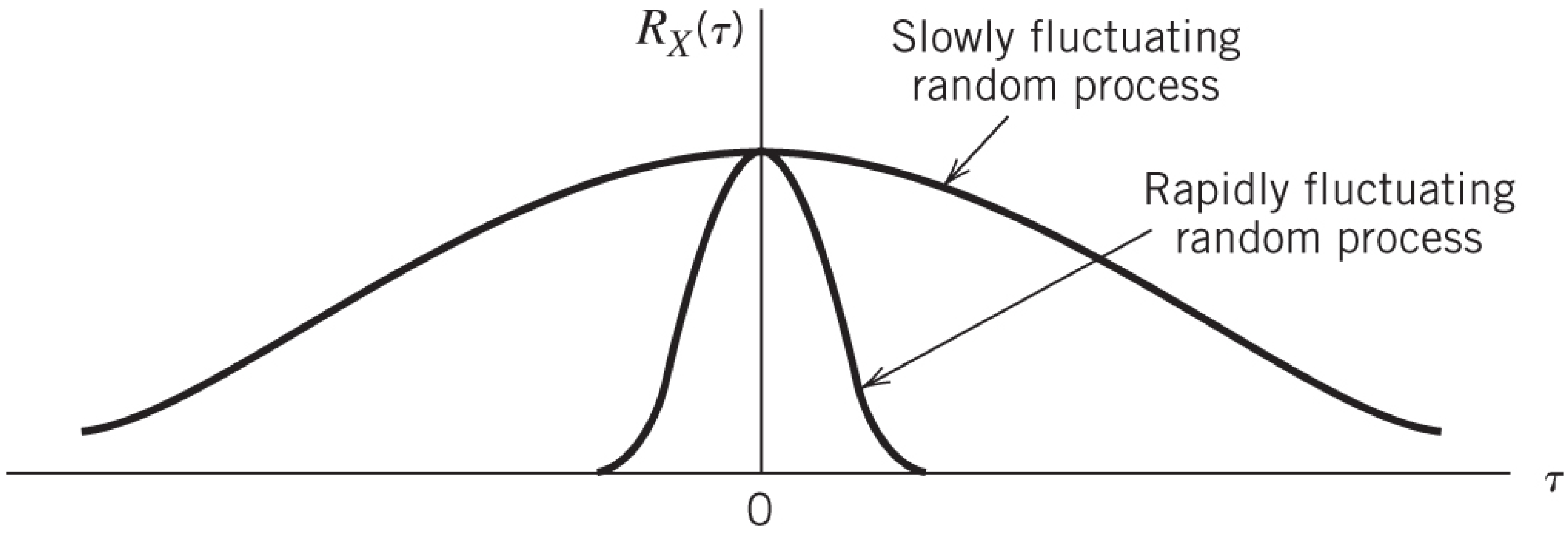
$$\rightarrow 2R_X(0) \pm 2R_X(\tau) \geq 0$$

$$\rightarrow -R_X(0) \leq R_X(\tau) \leq R_X(0)$$

$$\rightarrow \left|R_X(\tau)\right| \leq R_X(0)$$

5.6 Mean, Correlation and Covariance Functions

- The physical significance of the autocorrelation function $R_X(\tau)$ is that it provides a means of describing the "interdependence" of two random variables obtained by observing a random process $X(t)$ at times τ seconds apart.



- Example 5.7 Sinusoidal Signal with Random Phase.
  - Consider a sinusoidal signal with random phase:

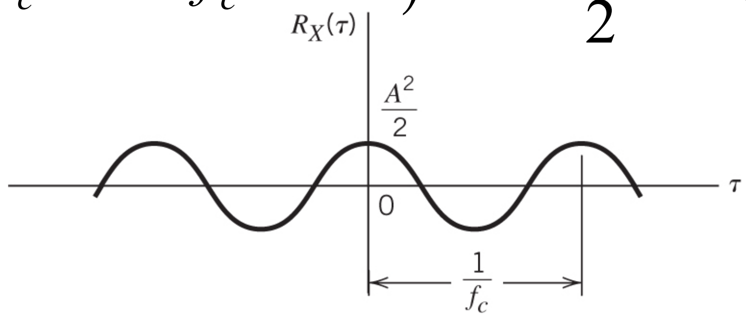
$$X(t) = A\cos(2\pi f_c t + \Theta) \qquad f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi}, & -\pi \le \theta \le \pi \\ 0, & \text{elsewhere} \end{cases}$$

$$R_X(\tau) = \mathbf{E}\left[X(t+\tau)X(t)\right]$$

$$= \mathbf{E}\left[\frac{A^2}{2}\cos(4\pi f_c t + 2\pi f_c \tau + 2\Theta)\right] + \frac{A^2}{2}\mathbf{E}\left[\cos(2\pi f_c \tau)\right]$$

$$= \frac{A^2}{2}\int_{-\pi}^{\pi}\frac{1}{2\pi}\cos(4\pi f_c t + 2\pi f_c \tau + 2\theta)\,d\theta + \frac{A^2}{2}\cos(2\pi f_c \tau)$$

$$= \frac{A^2}{2}\cos(2\pi f_c \tau)$$



- Averages for joint stochastic processes
  - Let $X(t)$ and $Y(t)$ denote two stochastic processes and let $X_{t_i} \equiv X(t_i)$, $i=1,2,\ldots,n$, $Y_{t'_j} \equiv Y(t'_j)$, $j=1,2,\ldots,m$, represent the random variables at times $t_1>t_2>t_3>\ldots>t_n$, and $t'_1>t'_2>t'_3>\ldots>t'_m$, respectively. The two processes are characterized statistically by their joint PDF:

$$f_{XY}\left(x_{t_1}, x_{t_2}, \ldots, x_{t_n}, y_{t'_1}, y_{t'_2}, \ldots, y_{t'_m}\right)$$

  - The *cross-correlation function* of $X(t)$ and $Y(t)$, denoted by $R_{xy}(t_1,t_2)$, is defined as the joint moment:

$$R_{xy}(t_1,t_2) = E(X_{t_1} Y_{t_2}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{t_1} y_{t_2} f_{XY}(x_{t_1}, y_{t_2}) dx_{t_1} dy_{t_2}$$

  - The *cross-covariance* is:

$$Cov(X_{t_1}, Y_{t_2}) = R_{xy}(t_1,t_2) - m_x(t_1) m_y(t_2)$$

5.6 Mean, Correlation and Covariance Functions

- Averages for joint stochastic processes
  - When the process are jointly and individually stationary, we have $R_{xy}(t_1,t_2)=R_{xy}(t_1-t_2)$, and $\mu_{xy}(t_1,t_2)=\mu_{xy}(t_1-t_2)$:

$$R_{xy}(-\tau)=E(X_{t_1}Y_{t_1+\tau})=E(X_{t_1^{'}-\tau}Y_{t_1^{'}})=E(Y_{t_1^{'}}X_{t_1^{'}-\tau})=R_{yx}(\tau)$$

  - The stochastic processes $X(t)$ and $Y(t)$ are said to be *statistically independent* if and only if :

$$f_{XY}(x_{t_1},x_{t_2},\ldots,x_{t_n},y_{t_1^{'}},y_{t_2^{'}},\ldots,y_{t_m^{'}})=f_X(x_{t_1},x_{t_2},\ldots,x_{t_n})f_Y(y_{t_1^{'}},y_{t_2^{'}},\ldots,y_{t_m^{'}})$$

    for all choices of t_i and t'_i and for all positive integers n and m.
  - The processes are said to be *uncorrelated* if

$$R_{xy}(t_1,t_2)=E(X_{t_1})E(Y_{t_2}) \qquad \Rightarrow Cov(X_{t_1},Y_{t_2})=0$$

- Example 5.9 Quadrature-Modulated Processes
  - Consider a pair of quadrature-modulated processes $X_1(t)$ and $X_2(t)$:

$$X_1(t) = X(t)\cos(2\pi f_c t + \Theta)$$
$$X_2(t) = X(t)\sin(2\pi f_c t + \Theta)$$

$$\begin{aligned}
R_{12}(\tau) &= \mathbf{E}\left[X_1(t)X_2(t-\tau)\right] \\
&= \mathbf{E}\left[X(t)X(t-\tau)\cos(2\pi f_c t + \Theta)\sin(2\pi f_c t - 2\pi f_c \tau + \Theta)\right] \\
&= \mathbf{E}\left[X(t)X(t-\tau)\right]\mathbf{E}\left[\cos(2\pi f_c t + \Theta)\sin(2\pi f_c t - 2\pi f_c \tau + \Theta)\right] \\
&= \frac{1}{2}R_X(\tau)\mathbf{E}\left[\sin(4\pi f_c t - 2\pi f_c \tau + 2\Theta) - \sin(2\pi f_c \tau)\right] \\
&= -\frac{1}{2}R_X(\tau)\sin(2\pi f_c \tau)
\end{aligned}$$

$$R_{12}(0) = \mathbf{E}\left[X_1(t)X_2(t)\right] = 0$$

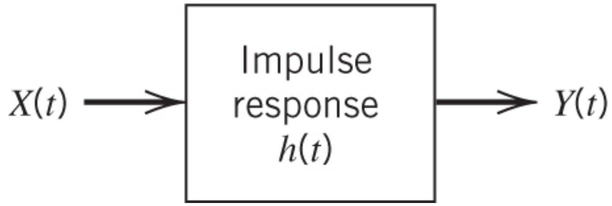
- Ergodic Processes
- In many instances, it is difficult or impossible to observe all sample functions of a random process at a given time.
- It is often more convenient to observe a single sample function for a long period of time.
- For a sample function $x(t)$, the *time average* of the mean value over an observation period $2T$ is

$$\mu_{x,T} = \frac{1}{2T}\int_{-T}^{T} x(t)\,dt \qquad (5.84)$$

- For many stochastic processes of interest in communications, the time averages and ensemble averages are equal, a property known as *ergodicity*.
- This property implies that whenever an ensemble average is required, we may estimate it by using a time average.

5.7 Transmission of a Random Process Through a Linear Filter

- Suppose that a random process $X(t)$ is applied as input to linear time-invariant filter of impulse response $h(t)$, producing a new random process $Y(t)$ at the filter output.



- Assume that $X(t)$ is a wide-sense stationary random process.
- The mean of the output random process $Y(t)$ is given by

$$\begin{aligned}\mu_Y(t) &= \mathbf{E}\left[Y(t)\right] = \mathbf{E}\left[\int_{-\infty}^{\infty} h(\tau_1) X(t-\tau_1)\, d\tau_1\right] \\ &= \int_{-\infty}^{\infty} h(\tau_1) \mathbf{E}\left[X(t-\tau_1)\right] d\tau_1 \\ &= \int_{-\infty}^{\infty} h(\tau_1) \mu_X(t-\tau_1)\, d\tau_1 \qquad (5.86)\end{aligned}$$

5.7 Transmission of a Random Process Through a Linear Filter

- When the input random process $X(t)$ is wide-sense stationary, the mean $\mu_X(t)$ is a constant μ_X, then mean $\mu_Y(t)$ is also a constant μ_Y .

$$\mu_Y(t) = \mu_X \int_{-\infty}^{\infty} h(\tau_1) d\tau_1 = \mu_X H(0) \qquad (5.87)$$

where $H(0)$ is the zero-frequency (dc) response of the system.

- The autocorrelation function of the output random process $Y(t)$ is given by:

$$R_Y(t,u) = \mathbf{E}\left[Y(t)Y(u)\right] = \mathbf{E}\left[\int_{-\infty}^{\infty} h(\tau_1) X(t-\tau_1) d\tau_1 \int_{-\infty}^{\infty} h(\tau_2) X(u-\tau_2) d\tau_2\right]$$

$$= \int_{-\infty}^{\infty} d\tau_1 h(\tau_1) \int_{-\infty}^{\infty} d\tau_2 h(\tau_2) \mathbf{E}\left[X(t-\tau_1) X(u-\tau_2)\right]$$

$$= \int_{-\infty}^{\infty} d\tau_1 h(\tau_1) \int_{-\infty}^{\infty} d\tau_2 h(\tau_2) R_X(t-\tau_1, u-\tau_2)$$

- When the input $X(t)$ is a wide-sense stationary random process, the autocorrelation function of $X(t)$ is only a function of the difference between the observation times:

$$R_Y(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau - \tau_1 + \tau_2)d\tau_1 d\tau_2 \qquad (5.90)$$

- If the input to a stable linear time-invariant filter is a wide-sense stationary random process, then the output of the filter is also a wide-sense stationary random process.

5.8 Power Spectral Density

- The Fourier transform of the autocorrelation function $R_X(\tau)$ is called the *power spectral density* $S_X(f)$ of the random process $X(t)$.

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau)\exp(-j2\pi f\tau)\,d\tau \qquad (5.91)$$

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f)\exp(j2\pi f\tau)\,df \qquad (5.92)$$

- Equations (5.91) and (5.92) are basic relations in the theory of spectral analysis of random processes, and together they constitute what are usually called the *Einstein-Wiener-Khintchine relations*.

5.8 Power Spectral Density

- Properties of the Power Spectral Density
  - **Property 1:** $S_X(0)=\int_{-\infty}^{\infty} R_X(\tau)\,d\tau$ (5.93)
    - Proof: Let $f=0$ in Eq. (5.91)
  - **Property 2:** $\mathbf{E}\left[X^2(t)\right]=\int_{-\infty}^{\infty} S_X(f)\,df$ (5.94)
    - Proof: Let $\tau=0$ in Eq. (5.92) and note that $R_X(0)=\mathbf{E}[X^2(t)]$.
  - **Property 3:** $S_X(f)\geq 0$ for all f (5.95)
  - **Property 4:** $S_X(-f)=S_X(f)$ (5.96)
    - Proof: From (5.91)

$$S_X(-f)=\int_{-\infty}^{\infty} R_X(\tau)\exp(j2\pi f\tau)\,d\tau \underset{R_X(\tau)=R_X(-\tau)}{\overset{\tau\to-\tau}{=}} \int_{-\infty}^{\infty} R_X(\tau)\exp(-j2\pi f\tau)\,d\tau = S_X(f)$$

Proof of Eq. (5.95)

- It can be shown that (see eq. 5.106) $S_Y(f) = S_X(f)\left|H(f)\right|^2$

$$R_Y(\tau) = \int_{-\infty}^{\infty} S_Y(f)\exp(j2\pi f\tau)\,df = \int_{-\infty}^{\infty} S_X(f)\left|H(f)\right|^2 \exp(j2\pi f\tau)\,df$$

$$R_Y(0) \underset{(5.64)}{=} E\left[Y^2(t)\right] = \int_{-\infty}^{\infty} S_X(f)\left|H(f)\right|^2 df \geq 0 \quad \text{for any } \left|H(f)\right|$$

- Suppose we let $|H(f)|^2=1$ for any arbitrarily small interval $f_1 \leq f \leq f_2$, and $H(f)=0$ outside this interval. Then, we have:

$$\int_{f_1}^{f_2} S_X(f)\,df \geq 0$$

  This is possible if an only if $S_X(f)\geq 0$ for all f.

- Conclusion: $S_X(f)\geq 0$ for all f.

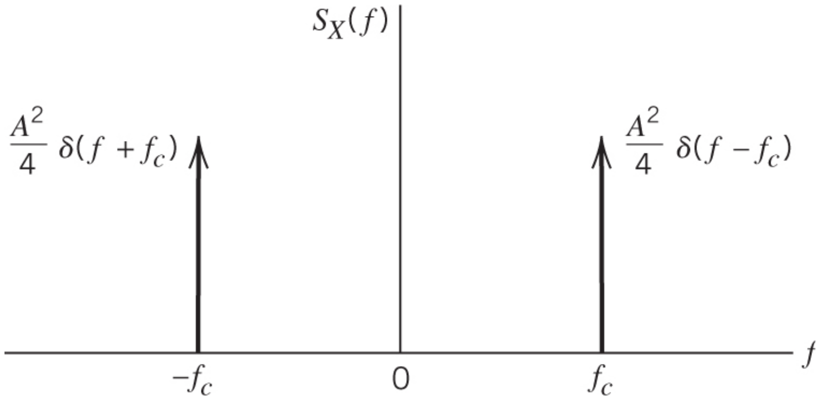
5.8 Power Spectral Density

- Example 5.10 Sinusoidal Signal with Random Phase
  - Consider the random process $X(t)=A\cos(2\pi f_c t+\Theta)$, where Θ is a uniformly distributed random variable over the interval $(-\pi,\pi)$.
  - The autocorrelation function of this random process is given in Example 5.7:

$$R_X(\tau)=\frac{A^2}{2}\cos(2\pi f_c\tau) \quad (5.74)$$

  - Taking the Fourier transform of both sides of this relation:

$$S_X(f)=\frac{A^2}{4}\left[\delta(f-f_c)+\delta(f+f_c)\right] \quad (5.97)$$



- Example 5.12 Mixing of a Random Process with a Sinusoidal Process
  - A situation that often arises in practice is that of mixing (i.e., multiplication) of a WSS random process $X(t)$ with a sinusoidal signal $\cos(2\pi f_c t+\Theta)$, where the phase Θ is a random variable that is uniformly distributed over the interval $(0,2\pi)$.
  - Determining the power spectral density of the random process $Y(t)$ defined by:

$$Y(f) = X(t)\cos(2\pi f_c t + \Theta) \qquad (5.101)$$

  - We note that random variable Θ is independent of $X(t)$.

5.8 Power Spectral Density

- Example 5.12 Mixing of a Random Process with a Sinusoidal Process (continued)
  - The autocorrelation function of $Y(t)$ is given by:

$$
\begin{aligned}
R_Y(\tau) &= \mathbf{E}\left[Y(t+\tau)Y(t)\right] \\
&= \mathbf{E}\left[X(t+\tau)\cos(2\pi f_c t + 2\pi f_c \tau + \Theta)X(t)\cos(2\pi f_c t + \Theta)\right] \\
&= \mathbf{E}\left[X(t+\tau)X(t)\right]\mathbf{E}\left[\cos(2\pi f_c t + 2\pi f_c \tau + \Theta)\cos(2\pi f_c t + \Theta)\right] \\
&= \frac{1}{2}R_X(\tau)\mathbf{E}\left[\cos(2\pi f_c \tau) + \cos(4\pi f_c t + 2\pi f_c \tau + 2\Theta)\right] \\
&= \frac{1}{2}R_X(\tau)\cos(2\pi f_c \tau)
\end{aligned}
$$

↓ Fourier transform

$$
S_Y(f) = \frac{1}{4}\left[S_X(f - f_c) + S_X(f + f_c)\right] \qquad (5.103)
$$

5.8 Power Spectral Density

- Relation among the Power Spectral Densities of the Input and Output Random Processes
  - Let $S_Y(f)$ denote the power spectral density of the output random process $Y(t)$ obtained by passing the random process through a linear filter of transfer function $H(f)$.

$$R_Y(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau-\tau_1+\tau_2)d\tau_1 d\tau_2 \quad (5.90)$$

$$S_Y(f) = \int_{-\infty}^{\infty} R_Y(\tau)e^{-j2\pi f\tau}d\tau$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau-\tau_1+\tau_2)e^{-j2\pi f\tau}d\tau_1 d\tau_2 d\tau$$

Let $\tau-\tau_1+\tau_2=\tau_0$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau_0)e^{-j2\pi f(\tau_0+\tau_1-\tau_2)}d\tau_1 d\tau_2 d\tau_0$$

$$= \int_{-\infty}^{\infty} h(\tau_1)e^{-j2\pi f\tau_1}d\tau_1 \int_{-\infty}^{\infty} h(\tau_2)e^{j2\pi f\tau_2}d\tau_2 \int_{-\infty}^{\infty} R_X(\tau_0)e^{-j2\pi f\tau_0}d\tau_0$$

$$= H(f)H^*(f)S_X(f) = |H(f)|^2 S_X(f) \quad (5.106)$$

5.10 Noise

- The sources of noise may be external to the system (e.g., atmospheric noise, galactic noise, man-made noise), or internal to the system.

- The second category includes an important type of noise that arises from spontaneous fluctuations of current or voltage in electrical circuits. This type of noise represents a basic limitation on the transmission or detection of signals in communication systems involving the use of electronic devices.

- The two most common examples of spontaneous fluctuations in electrical circuits are *shot noise* and *thermal noise*.

5.10 Noise

- Thermal Noise
  - *Thermal noise* is the name given to the electrical noise arising from the random motion of electrons in a conductor.
  - The mean-square value of the thermal noise voltage V_{TN}, appearing across the terminals of a resistor, measured in a bandwidth of Δf Hertz, is given by:

$$\mathbf{E}\left[V_{TN}^2\right] = 4kTR\Delta f \text{ volts}^2$$

k : *Boltzmann's constant*=1.38×10^{-23} joules per degree Kelvin.

T : *Absolute temperature* in degrees Kelvin.

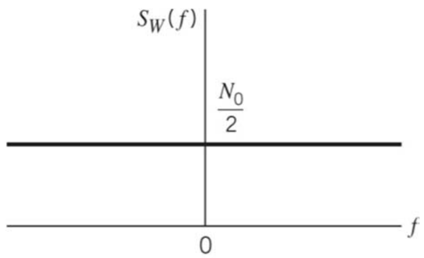
R: The resistance in ohms.

5.10 Noise

- White Noise
  - The noise analysis is customarily based on an idealized form of noise called *white noise*, the power spectral density of which is independent of the operating frequency.
  - *White* is used in the sense that white light contains equal amount of all frequencies within the visible band of electromagnetic radiation.
  - We express the power spectral density of white noise, with a sample function denoted by $w(t)$, as

$$S_W(f) = \frac{N_0}{2}$$

$$N_0 = kT_e$$

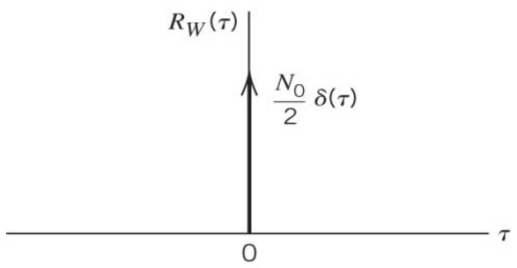


The dimensions of N_0 are in watts per Hertz, k is Boltzmann's constant and T_e is the equivalent noise temperature of the receiver.

5.10 Noise

- White Noise
  - The *equivalent noise temperature* of a system is defined as the temperature at which a noisy resistor has to be maintained such that, by connecting the resistor to the input of a noiseless version of the system, it produces the same available noise power at the output of the system as that produced by all the sources of noise in the actual system.
  - The autocorrelation function is the inverse Fourier transform of the power spectral density:

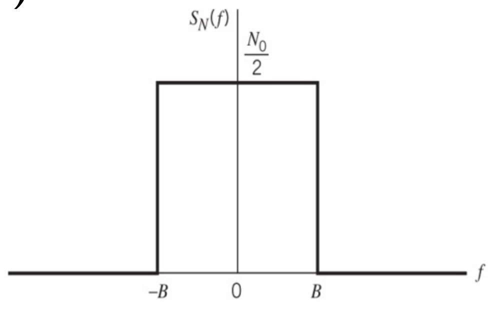
$$R_W(\tau) = \frac{N_0}{2}\delta(\tau)$$



  - Any two different samples of white noise, no matter how closely together in time they are taken, are uncorrelated.
  - If the white noise $w(t)$ is also Gaussian, then the two samples are statistically independent.

- Example 5.14 Ideal Low-Pass Filtered White Noise
  - Suppose that a white Gaussian noise $w(t)$ of zero mean and power spectral density $N_0/2$ is applied to an ideal low-pass filter of bandwidth B and passband amplitude response of one.
  - The power spectral density of the noise $n(t)$ is

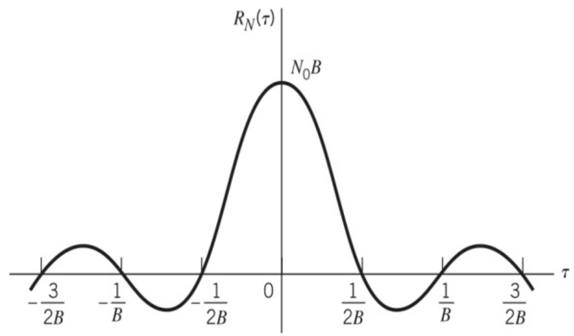
$$S_N(f) = \begin{cases} \dfrac{N_0}{2}, & -B < f < B \\ 0, & |f| > B \end{cases}$$



(a)

  - The autocorrelation function of $n(t)$ is

$$R_N(\tau) = \int_{-B}^{B} \frac{N_0}{2} \exp(j2\pi f\tau)\, df$$
$$= N_0 B \ \text{sinc}(2B\tau)$$



(b)