Chapter 2 Random Process and Optimal Filtering

Contents

Random Process

References

. Ziemer, R.E. and Tranter,W.H. , Principles of Communications, 6th Edition, Wiley , 2008

Couch, II, L.W., Digital and Analog Communication Systems, Seventh Ed. Pearson Prentice Hall, 2007.

.Therrien , C.W., Discrete Random Signals and Statistical Signal Processing. Prentice Hall,1992.

.Manolakis, D.G., Ingle , V.K., and Kogon ,S.M., Statistical and Adaptive Signal Processing , McGraw- Hill , 2000.

.Woods, J.W. and Stark, H., Probability and random Processes with Applications to Signal Processing ,Prentice Hall, 3rd Edition , 2002 .

.Leon-Garcia , A., Probability and Random Processes for Electrical Engineering , 3rd Edition, Prentice Hall .2008.

.Haykin, S.,, Adaptive Filter Theory, 4th ed. , Prentice Hall , 2002.

.

2. 1 Random Variable

- A **random variable** $X(A)$ represents the fundamental relationship between a random event A and a real number . For notational convenience, we usually designate the random variable by X .
- The random variable may be discrete or continuous.

2.1.1 Probability Distribution Function

- The distribution function $F_X(x)$ of the random variable X is given by

  $$F_X(x) = P(X \leqq x) \qquad (2.1)$$

  where $P(X \leqq x)$ is the probability that the value taken by the random variable X is less than or equal to a real number x .
- $F_X(x)$ is also called the **cumulative distribution function (CDF)** .

- The distribution function $F_X(x)$ has the following properties :

  1. $0 \leqq F_X(x) \leqq 1$
  2. $F_X(x_1) \leqq F_X(x_2)$ if $x_1 \leqq x_2$
  3. $F_X(-\infty) = 0$
  4. $F_X(+\infty) = 1$

- The probability distribution function (PDF) of the random variable X is defined as

$$p_X(x) = d\,F_X(x) \,/\, dx \qquad (2.2)$$

The probability of the event $x_1 \leqq X \leqq x_2$ equals

$$\begin{aligned} P(x_1 \leqq X \leqq x_2) &= P(X \leqq x_2) - P(X \leqq x_1) \\ &= F_X(x_2) - F_X(x_1) \\ &= \int p_X(x)dx \qquad (2.3) \end{aligned}$$

- **The probability function has the following properties :**

  1. $p_X(x) \geqq 0$

  2. $\int_{-\infty}^{\infty} p_X(x)dx = 1$

- **In the following , for ease of notation , we often omit the subscript *X* and write the PDF of a continuous random variable *X* simply as $p(x)$.**

  We will use the designation $p(X = x_i)$ for the probability of a discrete random variable *X* , where *X* can take on discrete values only .

- **Example 1.1 : Exponential Random Variable**

$$P(X > x) = e^{-\lambda x} \qquad x \geqq 0$$

The CDF of *X* is

$$F_X(x) = P(X \leqq x) = 1 - P(X > x)$$
$$= 0 \qquad x < 0$$
$$1 - e^{-\lambda x} \qquad x \geqq 0$$

- **Example 1.1.2 : Uniform Random Variable**

$$p(x) = \begin{cases} 1/(b-a) & a \leqq x \leqq b \\ 0 & \text{otherwise} \end{cases} \qquad (2.4)$$

Here we have

$E[X] = (b-a)/2$

$VAR[X] = (b-a)^2 / 12$

- **Example 1.1.3 Gaussian Random Variable**

$$p(x) = 1/\sqrt{(2\pi\sigma^2)} \; exp\{-(x-m)^2/2\sigma^2\} \qquad (2.5)$$

2.1.2 Statistical Averages

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

$$m_X = E[X] = \int_{-\infty}^{\infty} x \, p_X(x) \, dx \qquad (2.6)$$

- **The mean-square value of X is given by**

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \, p_X(x) \, dx \qquad (2.7)$$

- **The variance of X is defined as**

$$\text{var}(X) = \sigma_X^2 = E[(X - m_X)^2]$$
$$= \int_{-\infty}^{\infty} (x - m_X)^2 \, p_X(x) \, dx \qquad (2.8)$$

- **The variance and the mean-square are related by**

$$\sigma_X^2 = E[X^2 - 2 m_X X + m_X^2]$$
$$= E[X^2] - m_X^2 \qquad (2.9)$$

■ Chebyshev's Inequality

The variance $\sigma_X{}^2$ of a random variable X is a measure of the spread of the x values about their mean .

The Chebyshev inequality states that the x-values tend to cluster about their mean in the sense that the probability of a value not occurring in the near vicinity of the mean is small ; and it is the smaller the variance.

$$Pro\,[\,|\,x-m\,|\geqq \Delta\,] \leqq \sigma^2 / \Delta^2 \qquad (2.10)$$

■ Appendix

1. We have

$$Pro\,[\,|x-m| \geqq \Delta\,] \leqq \sigma^2 / \Delta^2$$

For $\Delta = k\,\sigma$, then

$$Pro\,[\,|x-m| \geqq k\,\sigma\,] \leqq 1 / k^2$$

***2.* Schwarz Inequality**

$$|\,(h, g)\,| = \|h\| \; \|g\|$$

3. Chi-squared density function

Foe a gaussian distribution function

$$f_X(x) = 1/\sqrt{(2\pi\sigma^2)} \;\; exp\,\{ - x^2 / 2\sigma^2 \}$$

define Y= X²

The pdf of Y is given by

$$f_Y(y) = 1/\sqrt{(2\pi y\sigma^2)} \;\; exp\,\{ - y^2 / 2\sigma^2 \}.$$

This is a Chi-squared distribution function.

2.1.3 Conditional Probability and Bayes' Rule

- **Consider a combined experiment in which a joint event occurs with joint probability $P(A, B)$.**
- **Suppose that the event B has occurred and we wish to determine the probability of occurrence of the event A.**

  This is called the conditional probability of the event A given the occurrence of the event B and is defined as

$$P(A \mid B) = P(A,B) / P(B) \qquad (2.11)$$

  provided that $P(B) > 0$.

In a similar manner, the probability of the event B conditioned on the occurrence of the event A is defined as

$$P(B \mid A) = P(A,B) / P(A) \qquad (2.12)$$

provided that $P(A) > 0$.

These relations may also expressed as

$$P(A,B) = P(A \mid B)\, P(B) = P(\mathrm{B} \mid A)\, P(A) \qquad (2.13)$$

2.1.4 Conditional Probability Density

- **A pair of two different random variables ,**
  $X = (X_1, X_2)$, may be thought of as a vector-valued random variable. Its statistical description requires knowledge of the joint probability density $p(x_1, x_2)$ . .
- **A quantity that provides a measure for the degree of dependence of the two random variables on each other is the conditional probability density $p(x_1 \mid x_2)$ of x_1 given x_2 , and $p(x_2 \mid x_1)$ of x_2 given x_1 .**
- **Bayes' rule :**

$$p(x_1, x_2) = p(x_1 \mid x_2)\, p(x_2) = p(x_2 \mid x_1)\, p(x_1) \qquad (2.14)$$

- **Two random variables are independent if they do not conditioned each other , that is , if**

$$p(x_2 \mid x_1) = p(x_2) \quad \text{and} \quad p(x_1 \mid x_2) = p(x_1)$$

Then, $p(x_1, x_2) = p(x_1)\, p(x_2)$

2.1.5 Gaussian Random Variable

- **A Gaussian random variable *X* is one whose probability density function can be written in the general form**

$$p(x) = \{ 1 / \sqrt{(2\pi\sigma^2)} \} \; exp [- (x - m)^2 / 2\sigma^2)]$$

where *m* is the mean and σ^2 is the variance.

- **CDF of a Gaussian random variable**

$$F_X(x) = \int_{-\infty}^{x} p(y)\, dy$$

$$= \int_{-\infty}^{(x-m)/\sigma} 1 / \sqrt{(2\pi)} \} \; exp [- t^2 / 2) \; dt$$

$$= 1 - Q[(x-m)/\sigma] \qquad (2.15)$$

where $Q(x) = 1 / \sqrt{(2\pi)} \} \int_{-\infty}^{x} exp (- t^2 / 2) \; dt$ **(2.16)**

Gaussian function with mean = 0 , variance = 1

$$g(x) = \{ 1/\sqrt{(2\pi)} \}\, exp\,(- x^2 / 2) \qquad (2.17)$$

Error function

$$erf\,(x) = \{ 2 / \sqrt{(\pi)} \} \int_0^x exp\,(- x^2)\, dt \qquad (2.18)$$

Complementary error function

$$erfc\,(x) = 1 - erf\,(x)$$

$$= 2 / \sqrt{(\pi)} \} \int_x^\infty exp\,(- x^2)\, dt \qquad (2.19)$$

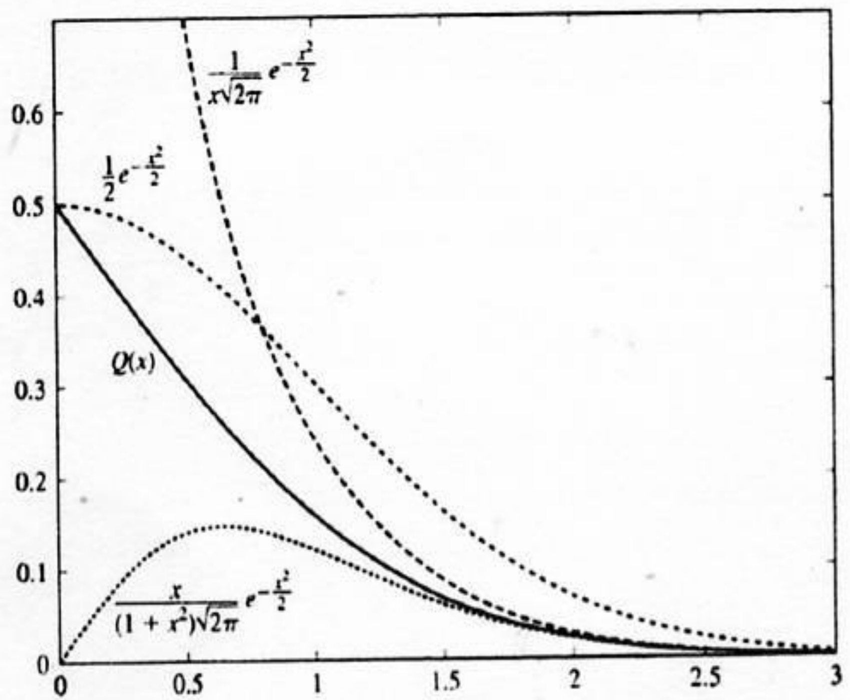
***Q*-function**

$$Q(x) = \{ 1/\sqrt{(2\pi)} \} \int_x^\infty exp\,(- x^2 / 2)\, dt \qquad (2.20)$$

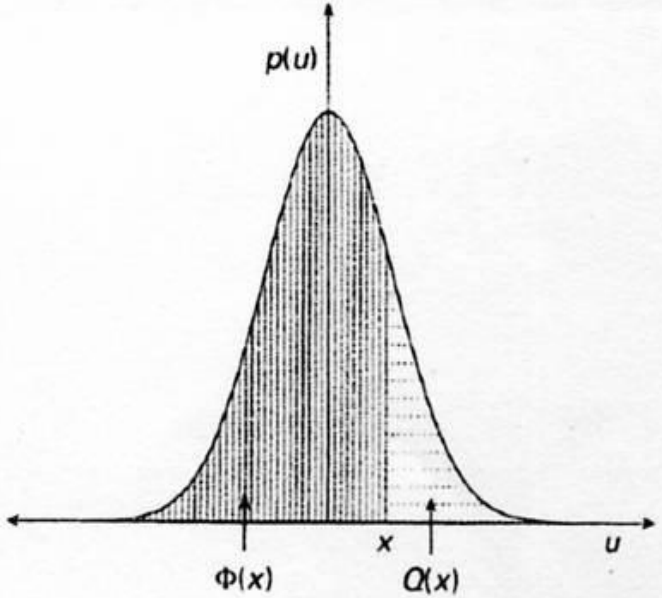
Relation between *erfc* (*x*) and *Q*- function :

$$Q\,(x) = \tfrac{1}{2}\; erfc\,(x / \sqrt{2})$$

$$erfc\,(x) = 2\; Q(x\sqrt{2})$$



$\frac{1}{x\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
$\frac{1}{2}e^{-\frac{x^2}{2}}$
$Q(x)$
$\frac{x}{(1+x^2)\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3



$p(u)$
x
u
$\Phi(x)$
$Q(x)$

2.1.6 Multiple Random Variables

- **Joint CDF**

The joint CDF of multiple random variable X_i , $i = 1, 2, \ldots, n$, is defined as

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \leqq x_1, X_2 \leqq x_2, \ldots, X_n \leqq x_n)$$

δ

$$= \int_{-\infty}^{x1} \int_{-\infty}^{x2} \ldots \int_{-\infty}^{xn} p(x_1, x_2, \ldots, x_n)\, d x_1\, d x_2 \ldots dx_n \qquad (2.21)$$

and $$p(x_1, x_2, \ldots, x_n) = \frac{\delta^n}{\delta x_1\, \delta x_2, \ldots \delta x_n} F(x_1, x_2, \ldots, x_n) \qquad (2.22)$$

- **The correlation between two random variables is defined as**

$$R_{XY} = E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, p_{xy}(x,y)\, dx\, dy$$

$$= m_{XY} \qquad (2.23)$$

Two random variables , X and Y , are said to be uncorrelated if $m_{XY} = m_X\, m_Y$

- **The covariance between two random variables , X and Y , is defined as**

$$\sigma_{xy} = \mathrm{Cov}(X,Y) = E[(X - m_x)(Y - m_y)]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - m_x)(y - m_y)\, p_{XY}(x,y)\, dxdy \qquad (2.24)$$

The correlation coefficient is defined as

$$\rho_{XY} = \sigma_{xy} / \sigma_x \sigma_y \quad , \quad -1 < \rho_{XY} < 1 \qquad (2.25)$$

- **If X and Y are statistically independent , then $\rho_{XY} = 0$**

  Note that random variables , X and Y , are independent if

  $$P(X = x_i,\ Y = y_i) = P(X = x_i)\, P(Y = y_i)$$

- **Gaussian Bivariate Distribution**

The bivariate Gaussian PDF of two random variables, X and Y , is expressed as

$$p_{XY}(x,y) = \{ 1 / [(2\pi\, \sigma_x \sigma_y)\sqrt{(1-\rho^2)}] \} \, exp \{ -(1/2) / (1-\rho^2) [(x-m_x)^2/\sigma_x^2 + (y-m_y)^2/\sigma_y^2 - 2\rho(x-m_x)(y-m_y)/\sigma_x\sigma_y] \} \qquad (2.26)$$

If they are independent , then

$$p_{XY}(x,y) = \{ 1 / (2\pi\, \sigma_x \sigma_y) \} \, exp \{ -(1/2) [(x-m_x)^2/\sigma_x^2 + (y-m_y)^2/\sigma_y^2] \} \qquad (2.27)$$

2.1.7 Sum of iid Random Variables and the Central Limit Theorem

- **Suppose that X_i, $i = 1, 2, \ldots n$, are statistically independent and identically distributed (iid) random variables, each having a finite mean m_x and a finite variance σ^2.**

- **Let Y_n be defined as the normalized sum, called the sample mean :**

$$Y_n = (1/n) \Sigma_{i=1}^{n} X_i \qquad (2.30)$$

The mean of Y_n is

$$E[Y_n] = m_y = (1/n) \Sigma_{i=1}^{n} E[X_i]$$
$$= m_x \qquad (2.31)$$

The variance of Y_n is

$$\sigma_y^{\,2} = E\,(Y_n^{\,2}) - m_y^{\,2}$$

$$= (\,1/\,n^2\ )\,\Sigma_{i=1}^{n}\,\Sigma_{i=1}^{n}\,\mathrm{E}(\,X_i\,X_j\,)$$

$$= (1/n\,)\sigma_x^{\,2} \qquad (2.32)$$

- ## Central Limit Theorem

Define the normalized random variable

$$Z_n = (Y_n - m_n) / \sigma_y = \Sigma_{i=1}^{n} (X_i - m)/ (\sigma_x \sqrt{n}) \quad (2.33)$$

Then the random variable Z_n has a distribution that is asymptotically unit normal.
That is, as n becomes large , the distribution of Z_n approaches that of a zero-mean Gaussian random variable with unit variance.

2.1.8 Transformation of Random Variables

- **Let X_1 and X_2 be continuous random variables with a joint PDF $f_{X1\,,\,X2}(x_1\,,\,x_2)$, and consider the transformation defined by**

$$y_1 = h_1(x_1\,,\,x_2) \quad \text{and} \quad y_2 = h_2(x_1\,,\,x_2)$$

Which are assumed to be one-to-one and continuously differentiable.

The Jacobian of this transformation is defined by the matrix determinant

$$J(x_1\,,\,y_1\;;\,x_2\;\;y_2) = \det \begin{pmatrix} \delta x_1 / \delta y_1 & \delta x_1 / \delta y_2 \\ \delta x_2 / \delta y_1 & \delta x_2 / \delta y_2 \end{pmatrix}$$

$$\neq 0 \qquad (2.34)$$

The joint pdf of $Y1$ and $Y2$ is given by

$$f_{Y1\ Y2}\ (\mathrm{y}_1, \mathrm{y}_2) = f_{X1\ X2}(x_1, x_2)\ |\ J(\qquad)\ | \qquad (2.35)$$

- **Example 1.1: Rayleigh distribution**

Let $Y = \sqrt{(X_1^2 + X_2^2)}$
where X_1 and X_2 are independent zero- mean Gaussian random variables.

$$f_{X,\ X}(x_1, x_2) = (1/2\pi\sigma^2)\ exp\ [-\ (x_1^2 + x_2^2) / 2\ \sigma^2\ ]$$

we will transform the two Gaussian random variables from Cartesian to polar coordinates.
Denote that $R = \sqrt{(X_1^2 + X_2^2)}$
and $\Theta = tan^{-1}(X_2 / X_1)$
The inverse transformation is

$$X_1 = R\ cos\Theta\ ,\ X_2 = R\ sin\Theta\ ,$$

The Jacobian is then calculate as

$$J = \det \qquad = \det$$

$$= r$$

The joint pdf of R and Θ is then

$$f_{R\,\Theta}(r, \theta) = f_{X1,\,X2}(x_1, x_2)\; J$$

$$= (r/2\pi\sigma^2)\, exp\, [-(x_1^2 + x_2^2)/2\sigma^2]$$

$$= (r/2\pi\sigma^2)\, exp\, [-r^2/2\sigma^2]$$

$$r \geqq 0, \quad 0 \leqq \theta \leqq 2\pi \qquad (2.36)$$

The marginal distribution of R and Θ are then given by

$$f_R(r) = (r/\sigma^2)\, exp(-r^2/2\sigma^2) \qquad r \geqq 0 \qquad (2.37)$$

and $$f_\Theta(\theta) = 1/(2\pi)\,, \quad 0 \leqq \theta \leqq 2\pi \qquad (2.38)$$

Note that $f_R(r)$ is the PDF of a Rayleigh random variable .
It is also denoted as Rayleigh distribution.

- **Example 1.2 : Rician distribution**

If X_1 and X_2 are two independent Gaussian random variables with means m_1 and m_2 , respectively, and a common variance σ^2 , then the new random variable

$$R = \sqrt{(X_1^2 + X_2^2)}$$

has a Rician distribution.

The pdf of R is given by

$$f_R(r) = (r/\sigma^2)\, exp\{-(r^2 + s^2)/2\sigma^2)\, I_0(r s/\sigma^2)\} \qquad r \geqq 0 \qquad (2.39)$$

where $I_0(.)$ is the modified Bessel function of zero order , and $s = \sqrt{(m_1^2 + m_2^2)}$, $K = s^2/2\sigma^2$ is denoted as Rician factor

-

Appendix: Complex Random Variables

- $R_{XX} = E[\, XX^H]$

  where H denotes the Hermitian transpose operation.

- $R_{XY} = E[\, XY^H]$

- $C_{XY} = E[(X - m_X)(Y - m_Y)^H]$

2.2 Random Signal and Random Process

2.2.1 Random Vector

- **For a vector of random variables $\mathbf{X} = [\, X_1 \; X_2 \ldots \; X_N \,]^T$,**

  we can construct a corresponding mean vector that is a column vector of the same dimension and whose components are the means of the elements of $\mathbf{X}$.

  That is, $m_x = E[X] = \{\, E[\, X_1 \,] \;\; E[\, X_2 \,] \ldots \;\; E[X_N \,] \,\}^T$ (2.44)

  The correlation matrix is defined as

  $$R_{XX} = E\,[\, X\, X^T \,] \qquad (2.45)$$

  Similarly , the covariance matrix is defined as

  $$C_{XX} = E\,[(\, x - m_x\,)\,(x - m_x\,)^T\,] \qquad (2.46)$$

- **Theorem :**

  Correlation matrices and covariance matrices are symmetric and positive definite.

- **Correlation between two random vectors X and Y is given by**

  $R_{XY} = E[\, XY^T \,]$ **(2.47)**

Joint Expectations for Two Random vectors

(a) For two random vectors, X and Y,

the **cross-correlation matrix** is defined as

$$R_{XY} = E[XY^T] \qquad (2.48)$$

and the **cross-covariance matrix** is define as

$$C_{XY} = E[(X - m_x)(Y - m_y)^T] \qquad (2.49)$$

where m_x and m_y are mean vectors of X and Y, respectively.

It can be shown that $R_{XY} = C_{XY} + m_X m_Y^T \qquad (2.50)$

(b) The random vectors X and Y are said to be **uncorrelated** if $R_{XY} = m_X m_Y^T$; or , equivalently, the cross-covariance $C_{XY} = 0$

(c) Two vectors are said to be **orthogonal** if $R_{XY} = 0$, or the cross-correlation is zero

2.2.2 Random Process and Statistical Average

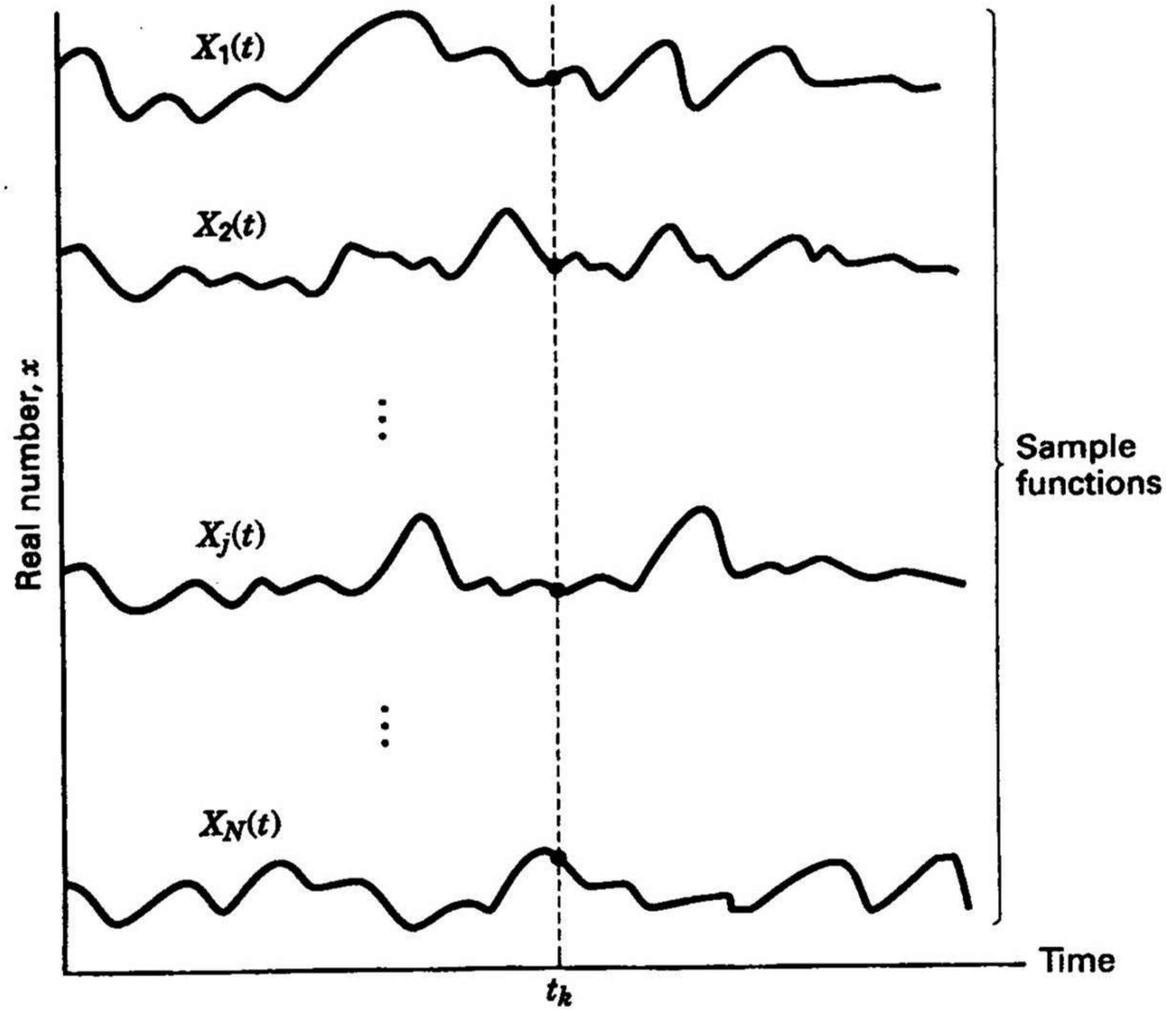
2.2.2.1 Random Process

- **A random process (or stochastic process) $X(A,t)$ can be viewed as a function of two variables : an event A and time t .**
- **Fig.1 illustrates a random process. In the figure, there are N sample functions of time, $\{ X_j(t) \}$.**

  Each of the sample functions can be regarded as the output of a different noise generator. For a specific event A_j , there is a a single time function $X(A_j ,t) = X_j(t)$, i.e. a sample function.

  The totality of all sample functions is called an ensemble.

  For a specific time t_k, $X(A_j ,t) = X_j(t_k)$ is simply a number.
- **For notational simplicity , we shall designate the random process by $X(t)$.**



Random process.

2.2.2.2 Statistical Averages

- **A random process whose distribution functions are continuous can be described statistically with a probability density function (PDF). In general , the form of the PDF of a random process will be different for different times.**
- **The mean of a random process**

$$E\{X(t_k)\} = \int_{-\infty}^{\infty} x \, p_X(x) dx \equiv m_X \qquad (2.78)$$

where $X(t_k)$ is the random variable obtained by observing the random process at time t_k . The pdf of $X(t_k)$ is designated as $p_X(x)$.

- **The autocorrelation function of a random process is defined as**

$$R_X(t_1, t_2) = E[X(t_1) X(t_2)] \qquad (2.79)$$

The autocorrelation function is a measure of the degree to which two time samples of the same random process are related.

- **In general, the *n*th moment is defined as**

$$E\{X^n\} = \int_{-\infty}^{\infty} x^n \, p_X(x) \, dx \qquad (1.80)$$

- **$E\{(X - m_X)^n\}$ is called the *n*th central moment , and when n=2 , the central moment is called the variance of the random process , denoted by σ_X^2.**

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - m_X)^2 \, p_X(x) \, dx \qquad (2.81)$$

- **Stationary Random Process (in the strict sense) :**
The statistics of a stationary random process are invariant to any translation of the time axis. That is,

$$p(x(t_1), x(t_2), \ldots, x(t_n)) = p(x(t_1+\tau), x(t_2+\tau), \ldots, x(t_n+\tau))$$

(2.82)

- **Wide-Sense Stationary (WSS) Random Process**

A random process is said to be wide-sense stationary (WSS) if two of its statistics , mean and autocorrelation, are invariant to a time shift. That is,

$$E\{ X(t) \} = m_X = \text{ a constant}$$

and $$R_X(t_1, t_2) = R_X(t_1 - t_2) = R_X(\tau) \qquad (2.83)$$

where $\tau = (t_1 - t_2)$ **(2.84)**

- **Properties of Autocorrelation of a Real-Valued WSS Random Process**

1. $R_X(\tau) = R_X(-\tau)$
2. $\left| R_X(\tau) \right| \leqq R_X(0)$ **for all** τ
3. $R_X(\tau) \longleftarrow\longrightarrow S_X(f)$ = **power spectral density**
4. $R_X(0) = E[X^2(t)]$ = **average power of the signal**

Stationary and Ergodicity

A stationary random process is said to be ergodic if time averages of a sample function are equal the corresponding ensemble average (or expectation) at a particular point in time.

That is,

$$m_X = \lim_{T \to \infty} (1/T) \int_{-T/2}^{T/2} X(t)\, dt \qquad (2.85)$$

and

$$R_X = \lim_{T \to \infty} (1/T) \int_{-T/2}^{T/2} X(t)\, X(t+\tau)\, dt \qquad (2.86)$$

Example : Gaussian Random Process

- A random process $x(t)$ is said to be Gaussian if the random variables

  $x_1 = x(t_1)$, $x_2 = x(t_2)$, ... , $x_n = x(t_N)$

  have an N-dimensional Gaussian PDF for any N and x_1 , x_2 ,..., x_N
- The N-dimensional Gaussian PDF is

$$f_X(x) = \{1/(2\pi)^{N/2} \mid \mathrm{Det}\ C \mid^{1/2}\}$$
$$exp\{-(1/2)\,[(x-m)^T C^{-1}(x-m)]\}$$

  where m is the mean vector , C is the covariance matrix of x .
- For a wide-sense stationary process , $m_i = E[x(t_i)] = m_j = E[x(t_j)]$, and the element of the covariance matrix become

$$c_{ij} = E[(x_i - m_i)(x_j - m_j)] = E[(x_i - m_i)]\ E[(x_i - m_j)]$$

- If , in addition, the x_i happen to be uncorrelated (e.g., white noise) ,

  $E[(x_i - m_i)] = E[(x_i - m_j)]$ for $i \neq j$

  then $c_{ii} = \sigma^2$, $c_{ij} = 0$ for $i \neq j$

2.2.2.3 Power Spectral Density

- **Definition of power spectral density (PSD) :**

For a random process $X(t)$, define a truncated version of the random process as

$$X_a(t) = \begin{cases} X(t) & |t| \leqq a \\ 0 & |t| > a \end{cases} \qquad (2.87)$$

The energy of this random process is

$$E_{Xa} = \int_{-a}^{a} X^2(t)\, dt = \int_{-\infty}^{\infty} X_a^{\ 2}(t)\, dt \qquad (2.88)$$

Hence, the time-average power is

$$P_{Xa} = (1/2a) \int_{-\infty}^{\infty} X_a^{\ 2}(t)\, dt = (1/2a) \int_{-\infty}^{\infty} X_a^{\ 2}(f)\, df \qquad (2.89)$$

- **The last quantity is obtained using Parseval's theorem. The quantity X_a(f) is the Fourier transform of X_a(t) .**
- **Note that P_{Xa} is a random variable and so to get the ensemble average power , we must take an expectation ,**

$$P_{Xa} = E[P_{Xa}] = (1/2a) \int_{-\infty}^{\infty} E[\,|X_a(f)|^2]\, df \qquad (2.90)$$

The power in the (untruncated) random process $X(t)$ is then found by passing to the limit a $\rightarrow \infty$,

$$P_{Xa} = \lim_{a \to \infty} (1/2a) \int_{-\infty}^{\infty} E[\,|X_a(f)|^2]\, df$$

$$= \int_{-\infty}^{\infty} \lim_{a \to \infty} (1/2a)\, E[\,|X_a(f)|^2]\, df \qquad (2.91)$$

Define $S_X(f) = \lim\ (1/2a)\ E[\ |X_a(f)|^2\ ]$ **(2.92)**

Then, the average power in the process can be expressed

as $P_X = \int_{-\infty}^{\infty} S_X(f)\, df$ **(2.93)**

$S_X(f)$ **is denoted as power spectral density of the random process** $X(t)$ **.**

- **Note : Parseval's energy theorem**

The energy of a non-periodic signal $g(t)$ **is equal to the total area under the curve of the energy density spectrum** $S_g(f)$ **, where**

$$E_g = \int_{-\infty}^{\infty} |g(t)|^2\, dt$$

$$= \int_{-\infty}^{\infty} |G(f)|^2\, df \qquad (2.94)$$

and $g(t) \leftrightarrow G(f)$ **(2.95)**

- **Wiener-Khinchine Relation :**

For a wide- sense stationary random process $X(t)$ whose autocorrelation function is given by $R_X(\tau)$, the power spectral (PSD) of the process is

$$S_X(f) = F\{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau)\, e^{-j2\pi f\tau}\, d\tau \tag{2.96}$$

In other words, the autocorrelation function and power spectral density for a Fourier transform pair.

2.4.4 Cross Correlation

- **Definition : The cross correlation between two random processes $X(t)$ and $Y(t)$ is defined as**

$$R_{XY}(t_1, t_2) = \mathrm{E}[X(t_1)\, Y(t_2)] \tag{2.97}$$

- **Two random processes $X(t)$ and $Y(t)$ are jointly stationary if both $X(t)$ and $Y(t)$ are individually stationary , and the cross correlation $R_{XY}(t_1, t_2)$ depends only on $\tau = (t_1 - t_2)$.**

It follows that

$$R_{XY}(t_1, t_2) = R_{XY}(t_1 - t_2) = R_{XY}(\tau)$$

- **Example :**

If two random processes $X(t)$ and $Y(t)$ are jointly stationary and

$Z(t) = X(t) + Y(t)$ then the autocorrelation of Z(t) is

$$R_Z(t+\tau, t) = R_X(\tau) + R_Y(\tau) + R_{XY}(\tau) + R_{XY}(-\tau)$$

2.2.3 Response of a Linear Time-Invariant System to Random Signals

- **Consider a linear time-invariant (LTI) system characterized by its impulse response *h*(*t*), *or*, equivalently, by its frequency response *H* (*f*) , where *h*(*t*) and *H(f)* are a Fourier transform pair. That is,**

$$H(f) = \int_{-\infty}^{\infty} h(t)\, e^{-j2\pi f t}\, dt \qquad (2.101)$$

$$h(t) = \int_{-\infty}^{\infty} H(f)\, e^{j2\pi f t}\, df \qquad (2.102)$$

- **Let *x*(t) be the input signal to the system and let *y*(t) denote the output signal. Then *y* (t) can be expressed in terms the convolution integral**

$$y(t) = \int_{-\infty}^{\infty} h(\tau)\, x(t-\tau)\, d\tau \qquad (2.103)$$

- **Now , suppose that x (t) is a sample function of a stationary stochastic process X (t). Then, the output y (t) is a sample function of a stochastic process Y (t) . The statistical averages are given as follows.**

The mean value of Y (t) is

$$m_Y(t) = E[Y(t)] = \int_{-\infty}^{\infty} h(\tau)\, E[X(t-\tau)]\, d\tau$$

$$= m_x \int_{-\infty}^{\infty} h(\tau)\, d\tau$$

$$= m_x H(0) \qquad (2.104)$$

where $H(0)$ is the frequency response of the linear system at $f = 0$.

The autocorrelation function of the output is

$$\Psi_{yy}(t_1, t_2) = (1/2)\, E[Y_{t1}, Y_{t2}^*]$$

$$= (1/2) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\beta)\, h^*(\alpha)\, E[X(t-\beta) X^*(t-\alpha)]\, d\alpha\, d\beta$$

$$(2.105)$$

After some mathematical manipulations, we finally obtain

$$\psi_{yy}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\beta)\, h^*(\alpha)\, \psi_{xx}(\tau+\alpha-\beta)\, d\alpha\, d\beta \qquad (2.106)$$

- By evaluating the Fourier transform of both sides of the above equation , we obtain the power spectral density of the output process in the form

$$\Phi_{yy}(f) = \Phi_{xx}(f)\, |H(f)|^2 \qquad (2.107)$$

When the autocorrelation function $\psi_{yy}(\tau)$ is desired, it can be evaluated by

$$\psi_{yy}(\tau) = \int_{-\infty}^{\infty} \Phi_{yy}(f)\, e^{j2\pi f\tau}\, df \qquad (2.108)$$

and $\psi_{yy}(0) = \int_{-\infty}^{\infty} \Phi_{xx}(f)\, |H(f)|^2\, df$

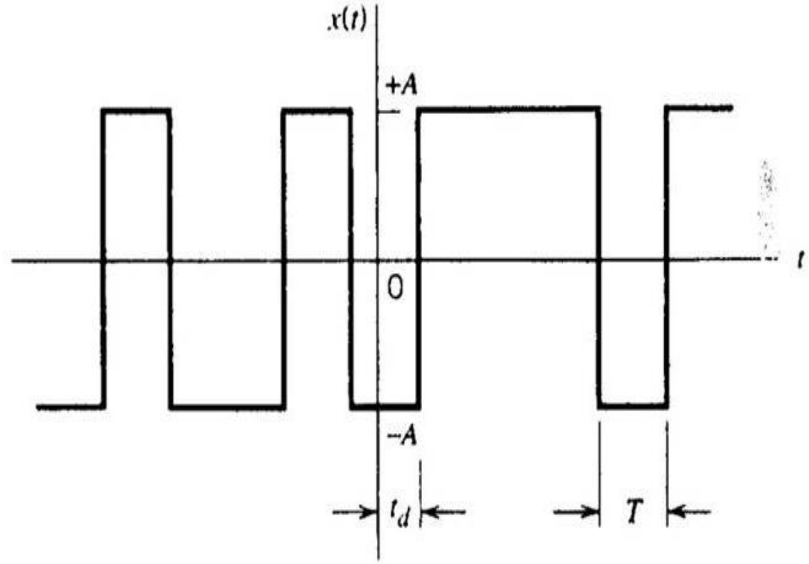
Example : Random Binary signal

The figure shows the sample function $x(t)$ of a random process $X(t)$ consisting of a random binary sequence of binary symbols , 1 and 0 . The following assumptions are made :

1. The symbols 1 and 0 are represented by rectangular pulses of amplitudes +A and –A , respectively.
2. The pulses are not synchronized , so the staring time of the first complete pulse for positive time is equaly to lie between 0 and T. Thus τ_d is a random variable uniformly distributed between *0* and *T* .
3. The amplitude level –A and +A occur with equal probability.

   Thus $E[X(t)] = 0$ for all *t* .

Consider the first case when $| t_k - t_j | > T$, *the random variables* $X(t_k)$ and $X(t_j)$ occur in different pulse intervals and are ,therefore , independent . Thus we have $E[X(t_k)\, X(t_j)] = E[X(t_k)]\, E[X(t_j)] = 0$



Sample function of random binary wave.

Consider next the case when $|t_k - t_i| < T$, with $t_k = 0$, $t_i < t_k$, *or* $t_i > t_k$.

In such a situation, we can see that ,from the figure , that the random variables $X(t_k)$ and $X(t_i)$ occur in the same pulse interval If and ony if the delay τ_d satisfies the condition

$$\tau_d < T - |t_k - t_i|$$

Thus we obtain the conditional expectation

$$E[X(t_k)\,X(t_i) \mid \tau_d] = \begin{cases} A^2 & \tau_d < T - |t_k - t_i| \\ 0 & \text{elsewhere} \end{cases}$$

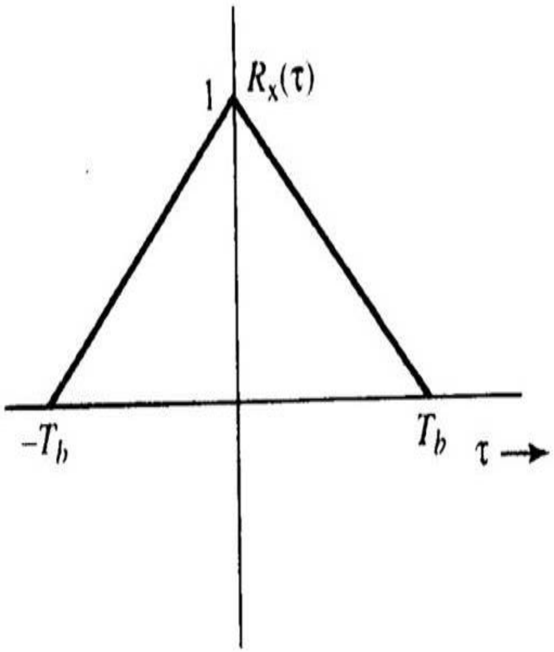
Averaging this result over all possible values of τ_d , we get

$$E[X(t_k)\,X(t_i)] = \int_0^{T - |t_k - t_i|} (A^2/T)\, d\tau_d$$

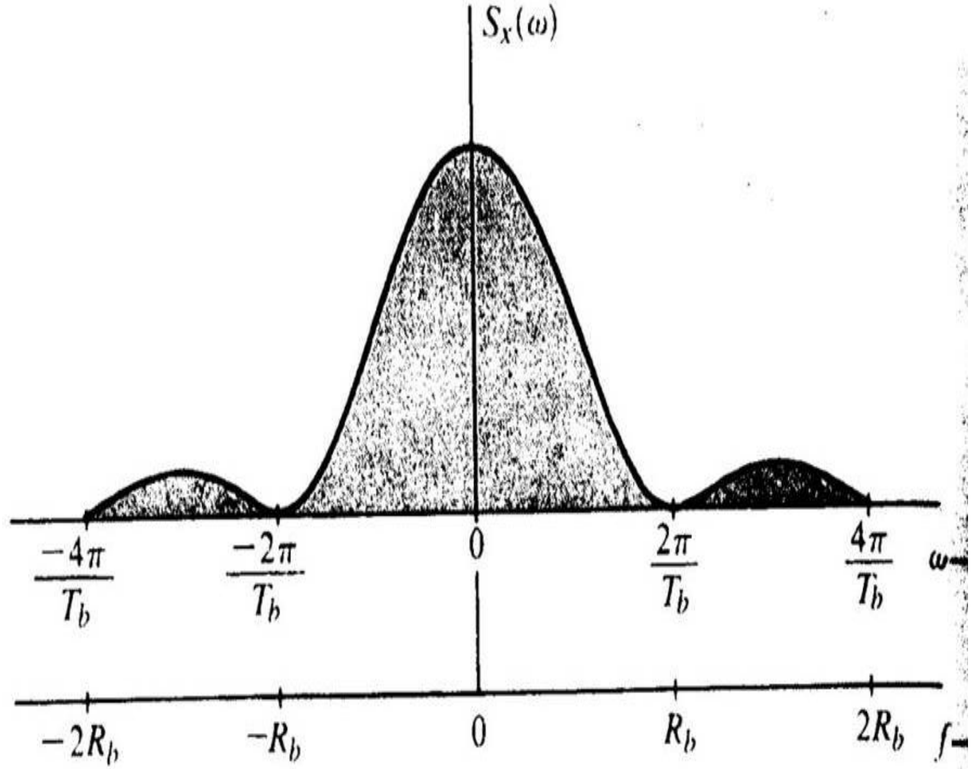
$$= A^2 (1 - |t_k - t_i| / T) \quad , \quad |t_k - t_i| < T$$

By same reasoning for any other values of t_k , we conclude that the autocorrelation function of a random binary wave can be expressed as

$$R_X(\tau) = \begin{cases} A^2 (1 - |\tau| / T) & |\tau| < T \\ 0 & |\tau| \geqq T \end{cases}$$

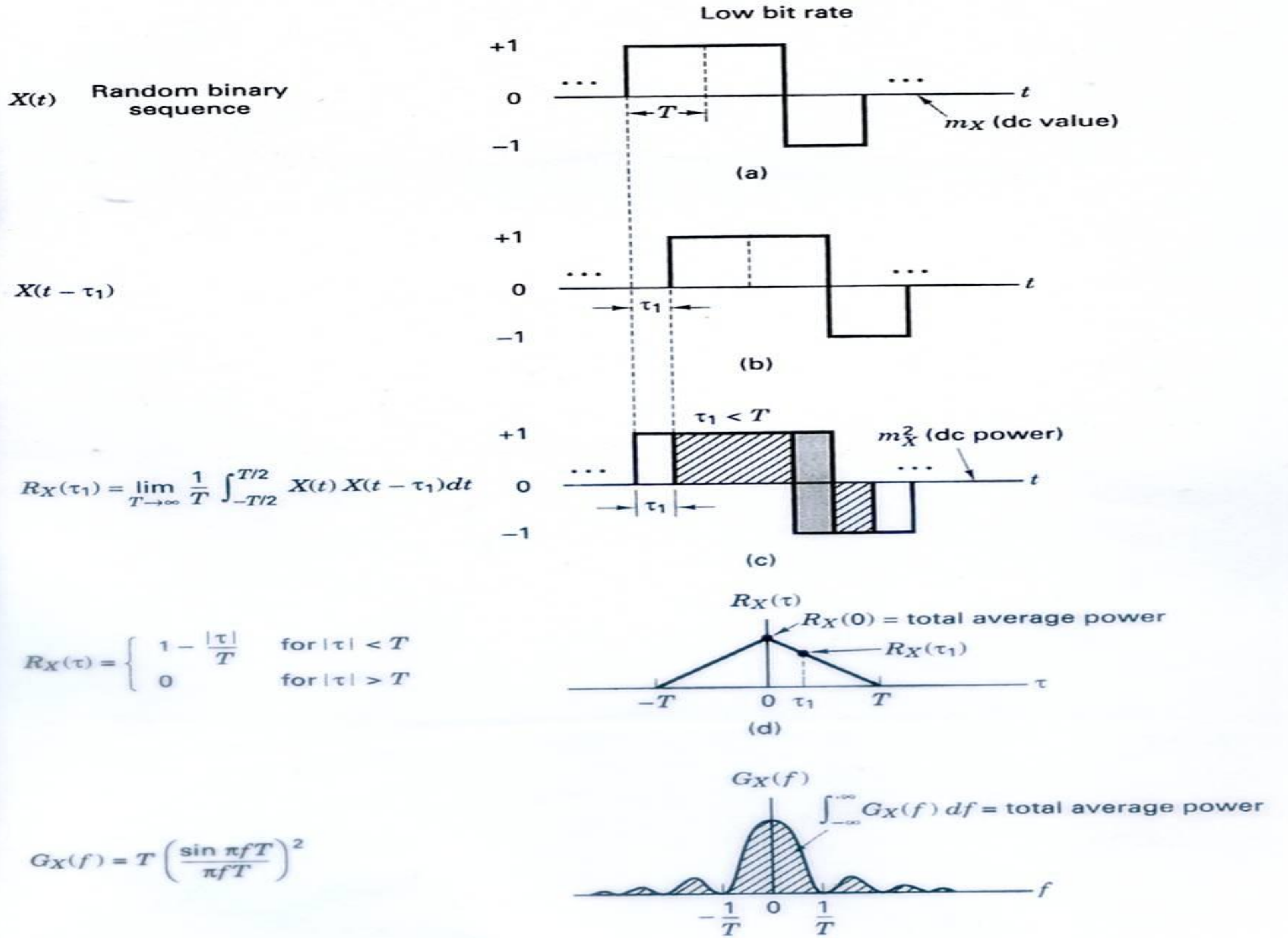

1
$R_x(\tau)$
$-T_b$
T_b
$\tau \rightarrow$

(c)


$S_x(\omega)$
$\frac{-4\pi}{T_b}$
$\frac{-2\pi}{T_b}$
0
$\frac{2\pi}{T_b}$
$\frac{4\pi}{T_b}$
$\omega \rightarrow$
$-2R_b$
$-R_b$
0
R_b
$2R_b$
$f \rightarrow$

(d)

- **Random Polar Binary Signal**



Autocorrelation and power spectral density.

High bit rate

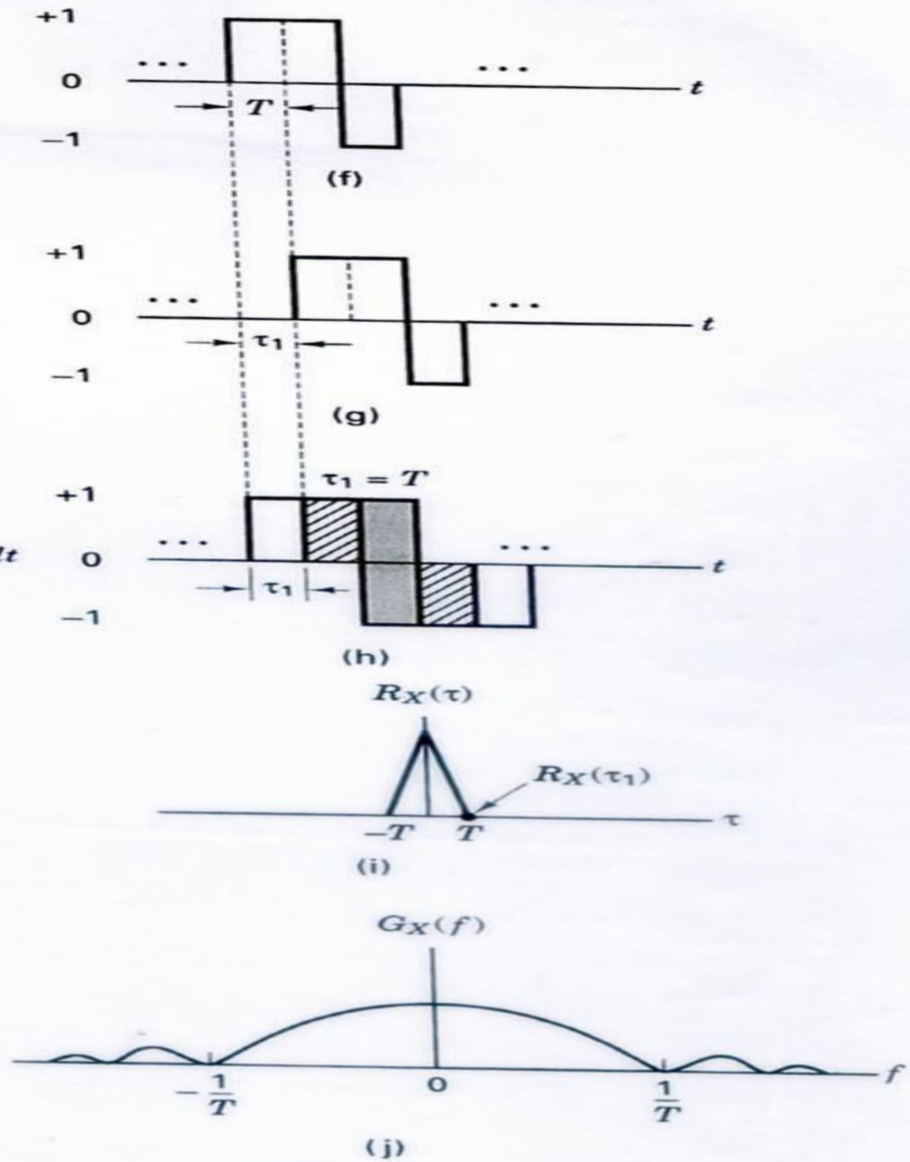
$X(t)$ Random binary sequence

$X(t-\tau_1)$

$$R_X(\tau_1) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)\, X(t-\tau_1)\,dt$$

$$R_X(\tau) = \begin{cases} 1 - \dfrac{|\tau|}{T} & \text{for } |\tau| < T \\ 0 & \text{for } |\tau| > T \end{cases}$$

$$G_X(f) = T\left(\frac{\sin \pi f T}{\pi f T}\right)^2$$



continued

2.2.4 Bandpass Random Process

- **We define a bandpass (or narrowband) process as a real , zero-mean , and WSS random process by**

$$X(t) = \mathrm{Re}\,[\, g(t)\, exp\, (\, j2\pi f_0 t + \theta_c\,)\,]$$
$$= X_i(t)\cos(\,2\pi f_0 t + \theta_c\,) - X_q(t)\sin(\,2\pi f_0 t + \theta_c\,) \qquad (2.109)$$

where $X_i(t)$ and $X_q(t)$ are denoted as the equivalent lowpass in-phase component and quadrature component , respectively, and θ_c is an independent random variable uniformly distributed over (0 , 2π) .

The lowpass equivalent process is given by

$$g(t) = X_i(t) + j\,X_q(t) \qquad (2.110)$$

The constant θ_c is often called the random start-up phase.

- **We can show that [Couch, pp. 446-452]**

1. $g(t)$ is a complex WSS baseband process .

2. $X_i(t)$ and $X_q(t)$ are jointly WSS zero-mean random processes .

3. $X_i(t)$ and $X_q(t)$ have the same power spectral density.

$$S_{Xi}(f) = S_{Xq}(f) = \begin{cases} [S_X(f - f_c) + S_X(f + f_c)] & |f| < B \\ 0 & \text{otherwise} \end{cases}$$

where B is the bandwidth of g(t).

4. Autocorrelation function

$$R_X(\tau) = \tfrac{1}{2} \, \mathrm{Re}\,\{R_g(\tau)\, exp\,(j\, 2\pi f_0 \tau)\}$$

5. Power spectral density

$$S_X(f) = \tfrac{1}{4}\, [S_g(f - f_c) + S_g(-f - f_c)]$$

- **Example : Filtered White Gaussian Noise**

White Gaussian noise with power spectral density of $N_0 / 2$ passes through an ideal bandpass filter with transfer function

$$H(f) = \begin{cases} 1 & |f - f_0| < B \\ 0 & \text{otherwise} \end{cases}$$

where $B < f_0$. .

The output , called filtered Gaussian white noise , is denoted

by $w(t)$.

The power spectral density of the filtered noise will be

$$S_w(f) = (N_0 / 2)\, |H(f)|^2$$

The filtered white Gaussian noise can also expressed as

$$w(t) = w_i(t)\ \cos(2\pi f_0 t) - w_q(t)\ \sin(2\pi f_0 t)$$

where $w_i(t)$ and $w_q(t)$ are the in-phase and quadrature components of $w(t)$, respectively, and are lowpass processes.

The power spectral density of the lowpass- equivalent processes are given by

$$S_{wi}(f) = S_{wq}(f) = \begin{cases} N_0 & |f| < B \\ 0 & \text{otherwise} \end{cases}$$

and $$S_g(f) = \begin{cases} 2N_0 & |f| < B \\ 0 & \text{otherwise} \end{cases}$$

Power of the bandpass Gaussian noise = $2N_0 B$

Example : Power Spectral Density of BPSK signal

The BPSK signal can be expressed by

$$v\,(\mathrm{t}) = x(\mathrm{t}) \cos\,(2\pi f_0 t + \theta_c)$$

where x(t) represents the polarity binary data and θ_c is the random start-up phase.

The PSD of v (t) is found by

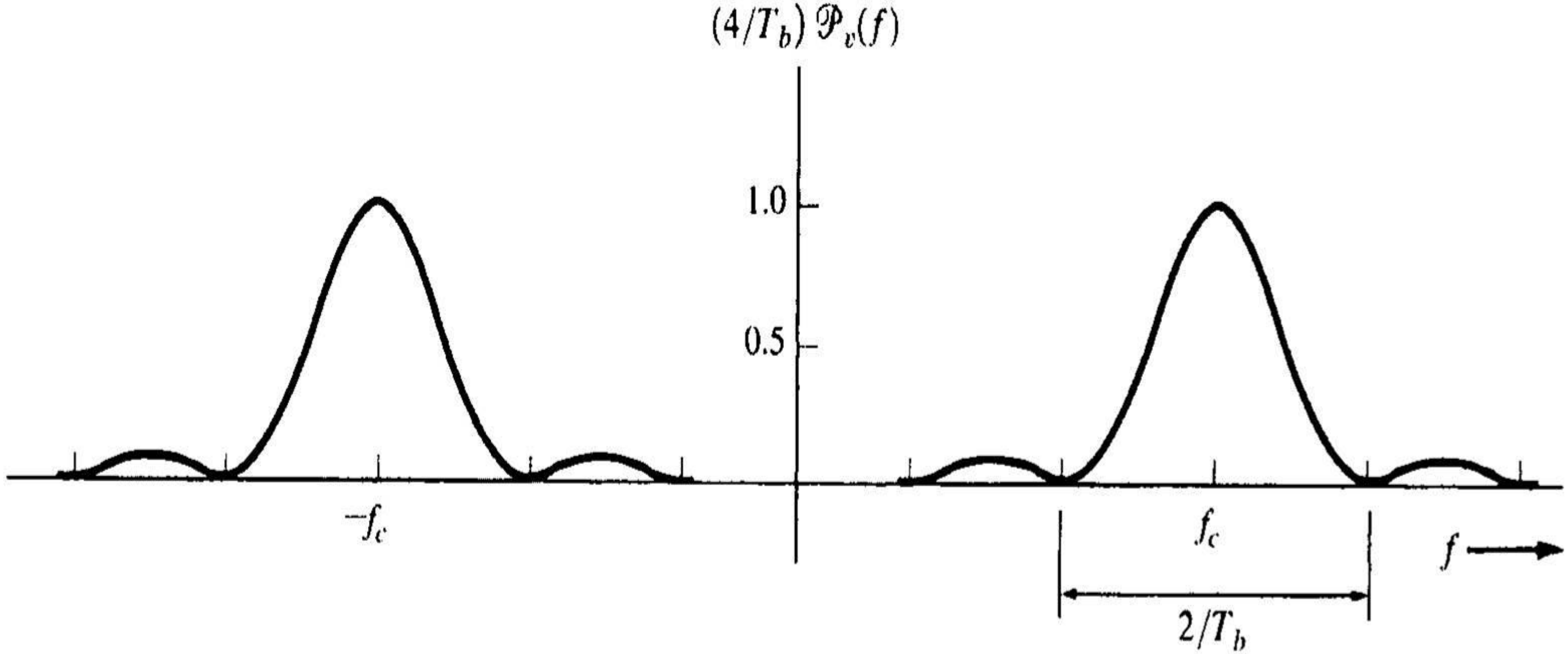
$$S_v(f) = ¼\ [S_x(f - f_c) + S_x(-f - f_c)]$$

The PSD of the polar baseband signal with equally likely binary data is given by

$$S_x(f) = T_b\ (\sin \pi f T_b\ / \pi f T_b)^2$$

We then obtain the PSD for the BPSK signal

$$S_v(f) = (1/4)\, T_b\ \{\ [\sin \pi(f - f_c)T_b\ / \pi(f - f_c)T_b]^2 + [\sin \pi(f + f_c)T_b\ / \pi(f + f_c)T_b]^2\ \}$$



Power spectrum for a BPSK signal.

Exercise #2

The PDF of a Rayleigh –distributed random variable X is given by

$$p(x) = \left(x / \sigma^2 \right) \; exp\left(-x^2 / 2\sigma^2 \right) \qquad x \geqq 0$$

Find the mean and variance of X.

Answer : **mean =** $\sigma / \sqrt{(\pi/2)}$

variance = $\sqrt{(2-\pi/2)}\ \sigma$

2.2.5 Markov Processes

- **Markov Process :**

A random process, X(t), is said to be a Markov process if for any time instants, $t_1 < t_2 < \dots < t_n < t_{n+1}$, the random process satisfies

$$F_X\left(X(t_{n+1}) \leqq x_{n+1} \mid X(t_n) = x_n ,\ X(t_{n-1}) = x_{n-1} , \dots, X(t_1) = x_1 \right)$$

$$= F_X\left(X(t_{n+1}) \leqq x_{n+1} \mid X(t_n) = x_n \right)$$

The Markovian property states that given the present, the future is independent of the past .

In other words, the future of the random process depends only on where it is now and not on how it got there.

Example #1 Sinusoidal wave with random phase

$X(t) = A \cos (2\pi f_c t + \Theta)$

where A is a constant and Θ is a random variable with uniform pdf over the interval $[-\pi , \pi]$, i.e.,

$$f_{\Theta}(\theta) = \begin{cases} 1/2\pi, & -\pi \leqq \theta \leqq \pi \\ 0, & \text{elsewhere} \end{cases}$$

1. Find the autocorrelation function of X(t)

   Ans. $R_X(\tau) = (A^2/2) \cos (2\pi f_c)$

2. Find the power spectral density of X(t)

   Ans. $S_X(f) = (A^2/2) [\delta(f- f_c) + \delta(f- f_c)]$

Example # 2

If $Y(t) = X(t) \cos(2\pi f_c t + \Theta)$

where X(t) is a stationary random process and Θ is a randomvariable with uniform pdf over the interval $[-\pi, \pi]$

Find the autocorrelation function and power spectral density of X (t) .

Ans. $R_Y(\tau) = (1/2)\, R_X(\tau) \cos(2\pi f_c \tau)$

$$S_Y(f) = (1/2)\,[S_X(f - f_c) + S_X(f + f_c)]$$

Example #3

A stationary Gaussian process X(t) with zero-mean and PSD $S_X(f)$ is applied to a linear filter whose impulse response is a rectangular function of time , duration = T , height = 1/ T .

Y(t) is the output at time t .

1. find the mean and variance of Y(t)
2. what is the probability density function of Y.?
3. Find the output power spectral density.

Ans. $H(f) = \exp(-j\pi fT)\, \sin(\pi fT)/(\pi fT)$

$$S_{YY}(f) = [\sin^2(\pi fT)/(\pi fT)^2]\, S_{XX}(f)$$

Appendix

2.A. Gram-Schmidt Orthogonalization Process

- Suppose that the subspace Y is defined by means of a non-orthogonal basis , such as a collection of random variables

  $$Y = \{ y_1 , y_2 , \ldots, y_M \} \tag{1.68}$$

  Which may be mutually correlated. The subspace Y is defined again as the linear span of this basis. The Gram-Schmidt orthogonalization process is a recursive procedure of generating an orthogonal basis $\varepsilon_1 , \varepsilon_2 , \ldots, \varepsilon_M$ from $y_1, y_2, \ldots, y_M$.

  The basic idea of the method is this :

  a. Initiate the procedure by selecting $\varepsilon_1 = y_1$

  b. Consider y_2 and decompose it relative to ε_1 . Then the component of y_2 which is perpendicular to ε_1 is selected as ε_2 ; so that $(\varepsilon_1 , \varepsilon_2) = 0$.

c. **Take y_3 and decompose it relative to the subspace spanned by $\{\varepsilon_1, \varepsilon_2\}$ and take the corresponding perpendicular component to be ε_3, and so on.**

For example, the first three steps of the procedure are

$$\varepsilon_1 = y_1$$

$$\varepsilon_2 = y_2 - E[y_2 \varepsilon_1] E[\varepsilon_1\varepsilon_1]^{-1} \varepsilon_1$$

$$\varepsilon_3 = y_3 - E[y_3\varepsilon_1] E[\varepsilon_1\varepsilon_1]^{-1} \varepsilon_1 - E[y_3\varepsilon_2] E[\varepsilon_2\varepsilon_2]^{-1} \varepsilon_2$$

***d*. At the *n*th iteration step**

$$\varepsilon_n = y_n - \sum_{i=1}^{n-1} E[y_n\varepsilon_i] E[\varepsilon_i\varepsilon_i]^{-1} \varepsilon_i \ , \quad 2 \leqq n \leqq M \qquad (1.69)$$

The basis $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M\}$ generated in this way is orthogonal by construction.