Chapter 6

RANDOM SIGNALS AND NOISE

Introduction

In previous chapters we dealt entirely with deterministic signals, where the signals were function of time indicating that the behavior of the signal is known or determined for all time. This chapter deals with random signals whose exact behavior cannot be described in advance. Random signals occur in communication both as unwanted noise and as desired information-bearing waveforms. If a message to be received is specified (i.e., if it is known beforehand), then it contains no uncertainty and conveys no information to the receiver. The noise signals are also unpredictable otherwise they can simply be subtracted at the receiver end.

Due to lack of detailed knowledge of the time variation of a random signal, the mathematical model based in terms of probabilities and statistical properties is used. Thus, this chapter deals with the terms for the description of random signals. The major topics include probabilities, random variables, statistical averages, and important probability models.

6.1 Random Variables

An experiment whose outcome cannot be exactly predicted is called a random experiment (e.g. tossing of a coin, drawing of a card from a deck of playing cards, etc.). The sample space is the set of collective outcomes of a random experiment. A particular outcome of an experiment is called a sample point or sample. Collection of outcomes is called an event. Thus, an event is a subset of sample space.

A random variable is a real valued function defined over a sample space of random experiment. Thus, the random variable is a rule or functional that maps the sample points into real number. It is also known as stochastic variable, or random function or stochastic function. In the following text, the random variables are denoted by upper-case letters such as X, Y, etc. and the value assumed by them is denoted by lower-case letter with subscripts such as x_1, x_2, y_1, y_2, etc. The random variables are of two types

1. **Discrete Random variables**

   A random variable whose outcome takes on a finite number of values is known as a discrete random variable. In other words if the sample space S contains a countable number of sample points, then X will be a discrete random variable. For example, the experiment of two tosses of a fair coin, there are eight (i.e., 2^2=4) possible outcomes of this experiment. These will constitute the sample space which consists of {Head, Head}, {Head, Tail}, {Tail, Head} and {Tail, Tail}. Each {} item is an event in the sample space S. A random variable X is simply the function that takes each {} and maps it into a number as shown in Fig.6.1.

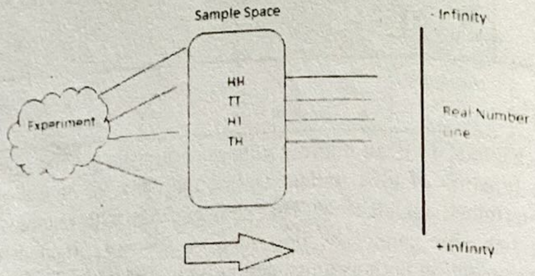

Fig.6.1: Mapping of random variable into a real number

Let the number of heads be the random variable X. The sample space S, and the random variable X will then be as shown below.

S	=	[HH	HT	TH	TT]
X	=	[x_1	x_2	x_3	x_4]
	=	[2	1	1	0]

Many other random variables can be defined over the same sample space, such as, number of tails, square of the number of heads, difference of number of heads and number of tails, etc.

2. **Continuous Random Variable**

A random variable that takes on an infinite number of values is known as continuous random variables. As there are infinite possible values of X, the probability that it takes on any particular value is zero. Hence, the probability function in this case cannot be defined as in the discrete case. In other words, if the sample space S contains a infinite number of sample points, then X will be a continuous random variable. In a continuous case, the probability that lies between two different values is non-zero. For example, if X represents the weight of a person, then the probability that it is exactly 70 kg would be zero, but the probability that it is between 55 kg and 87 kg would be non-zero.

6.1.1 Cumulative Distribution Function (CDF) and Probability Density Function (PDF)

The Cumulative Distribution Function (CDF) of a random variable is defined as the probability that the random variable X takes values less than or equal to x.

$$(CDF): F_X(x) = P(X \le x) \qquad (6.1)$$

Here, x is a dummy variable which is fixed and CDF is denoted by $F_X(x)$.

Important properties of CDF are

1. The CDF is always bounded by 0 and 1.

$$0 \le F_X(x) \le 1$$

2. $F_X(\infty) = 1$

Here, $F_X(\infty) = P(X \le \infty)$, includes all the probability of all possible outcome.

3. $F_X(-\infty) = 0$

The random variable X cannot have any value which is less than or equal to $-\infty$.

4. $F_X(x_1) \le F_X(x_2)$ for $x_1 < x_2$

This property states that $F_X(x)$ is a non-decreasing function.

5. $P(x_1 < X \le x_2) = F_X(x_2) - F_X(x_1)$

Defines the area under the PDF curve over the interval x_1 to x_2.

Probability Density Function (PDF)

PDF is a statistical measure that defines a probability distribution for a random variable and is often denoted as $f_X(x)$. PDF is more convenient way of describing a continuous random variable. The graphical representation of PDF indicates the interval under which the variable will fall. Mathematically, the probability density function is defined as the derivative of the cumulative distribution function. Thus, we have

$$PDF: f_X(x) = \frac{d}{dx} F_X(x) \qquad (6.2)$$

The probability and PDF is related in such a way that the probability of observing X in any interval from x_1 to x_2 is given by the area under the PDF curve over the interval x_1 to x_2.

i.e.,
$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f_X(x)dx \qquad (6.3)$$

Some of the important properties of PDF are

1. The CDF can be derived by integrating PDF.

$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

2. PDF is a non-negative function for all values of x

$$f_X(x) \ge 0 \qquad for\,all\,x$$

3. The area under PDF curve is always equal to unity.

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

The distribution of a discrete random variable can be characterized through its **Probability Mass Function (PMF)**. Mathematically, it is defined as

$$f_X(x_i) = P(X = x_i) \quad for\,all\,i$$

Thus,

$$f_X(x) \geq 0 \;\; and \;\; \sum_i f_X(x_i) = 1$$

6.1.2 Joint Cumulative Distribution and Probability Density Function

Previously chapter focused the situations involving only one random variable. However, many times, the outcome of an experiment requires more than one random variables to describe the experiment. Consider the situations involving two random variables X and Y. Now the (two dimensional) Joint CDF for two random variables is defined as

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) \tag{6.4}$$

$F_{XY}(x,y)$ is the probability associated with the set of all those sample points such that under X, their transformed values will be less than or equal to x and at the same time, under Y, the transformed values will be less than or equal to y. In other words, $F_{XY}(x_1, y_1)$ is the probability associated with the set of all sample points whose transformation does not fall outside the shaded region in the two dimensional space shown in Fig. 6.2.

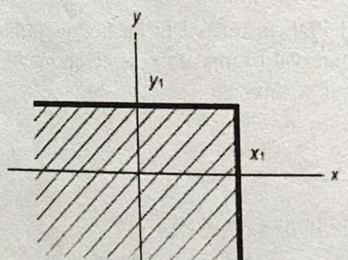

Fig.6.2. *Space of (X,Y) corresponding to $F_{XY}(x_1,y_1)$*

Looking at the sample space S, let A be the set of all those sample points $s \in S$ such that $X \leq x_1$. Similarly, if B is comprised of all those sample points $s \in S$ such that $Y \leq y_1$; then $F_{XY}(x_1, y_1)$ is the probability associated with the event AB.

Properties of the two dimensional CDF are:

1. $F_{XY}(x, y) \geq 0, \quad -\infty < x < \infty, \; -\infty < y < \infty$
2. $F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = 0$
3. $F_{XY}(\infty, \infty) = 1$
4. $F_{XY}(\infty, y) = F_Y(y)$
5. $F_{XY}(x, \infty) = F_X(x)$
6. If $x_2 > x_1$ and $y_2 > y_1$, then

$$F_{XY}(x_2, y_2) \geq F_{XY}(x_2, y_1) \geq F_{XY}(x_1, y_1)$$

We define the two dimensional joint PDF as

$$f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) \tag{6.5}$$

Properties of the two dimensional PDF are:

1. $f_{XY}(x, y) \geq 0$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dxdy = 1$
3. $P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dxdy$

6.1.3 Common Probability Model

There are various types of random variables with their own probability density function which are different from each other. Practically, it is very difficult to study all these probability distribution functions. Hence, the PDF's of all the random variables are approximated to some standard probability density functions. Some commonly used standard PDF's are.

Uniform Distribution

If a continuous random variable X is equally likely to be observed in a finite range and is likely to have a zero value outside this finite range, then the random variable is said to have a uniform distribution. The plot of PDF versus the dummy variable x is shown in Fig.6.3.

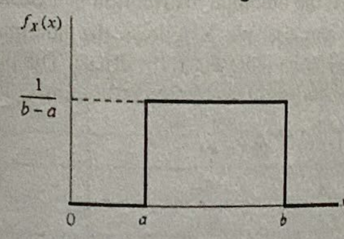

Fig.6.3. *Uniform Distribution*

The PDF of a random variable having uniform distribution is given by

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & for \;\; a \le x \le b \\ 0 & elsewhere \end{cases} \qquad (6.6)$$

The Bernoulli distribution

Consider a probabilistic experiment involving the discrete random variables X that takes one of two possible values:

- The value 1 with probability p;
- The value 0 with probability 1-p

Such a random variable is called the Bernoulli random variable, the probability distribution of which is defined by

$$f_X(x) = \begin{cases} 1-p; \; x=0 \\ p; \; x=1 \\ 0 ; otherwise \end{cases} \qquad (6.7)$$

This probability distribution is illustrated in Fig.6.4 the two delta functions weighted 1/2, shown in Fig.6.4 represents the probability at each sample points x=0 and x=1

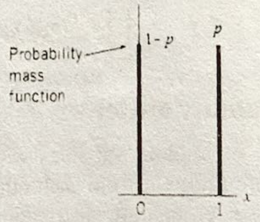

***Fig.6.4** Bernoulli distribution*

Binomial distribution

The binomial distribution is the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p. Therefore, the probability of an event is defined by its binomial distribution. A success/failure experiment is called a Bernoulli experiment or Bernoulli trial as previously explained. When n = 1, the binomial distribution is a Bernoulli distribution. In general, if the random variable X follows the binomial distribution with parameters n and p, X can be represented by B(n,p). The probability of getting exactly k successes in n trials is given by the density function in Eq.(6.8) below.

$$f_X(x) = \binom{n}{k} p^k (1-p)^{n-k} \delta(x-k) \qquad (6.8)$$

Where, $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$

Poisson Distribution

As the number n increases, the binomial distribution becomes difficult to handle. If 'n' is very large, probability 'p' is very small, then the binomial distribution can be approximated by the Poisson distribution. The probability of random variable having Poisson distribution is given by Eq.(6.9)

$$P(X=k) = \frac{(np)^k \cdot e^{-np}}{k!} = \frac{\lambda^k \cdot e^{-\lambda}}{k!} \qquad (6.9)$$

Where, λ is the parameter of the distribution. We say X follows a Poisson distribution with parameter λ

Gaussian (Normal) distribution

Gaussian distribution is one of the most commonly approximated distribution in the statistical analysis of communications systems. The majority of noise processes observed in practice are Gaussian and many naturally occurring experiments are characterized by continuous random variables with Gaussian distribution. The PDF of continuous random variable having Gaussian distribution is given by Eq.(6.10)

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad (6.10)$$

Where, μ is the mean and σ^2 is the variance of random variable.

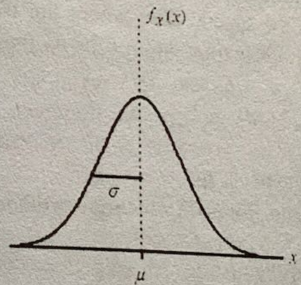

***Fig.6.5** Gaussian (Normal) PDF*

The mean and variance of common probability models is given in the Table.6.1

Table.6.1: Mean and variance of various distributions

Random variable (X)	Mean E[X]	Variance (σ_x^2)
Uniform	$(a+b)/2$	$(b-a)^2/12$
Bernoulli	p	$p(1-p)$
Binomial	np	$np(1-p)$
Poisson	λ	λ
Gaussian (Normal)	m_x	σ_x^2

6.1.4 Statistical Average

For many of the applications, the PDF provides more information about the random variable than actually needed. Therefore, it is always simpler and more convenient to describe a random variable by a few characteristics numbers only. These numbers are the various statistical averages or mean values. Some of these averages provide a simple and fairly adequate (though incomplete) description of the random variable. We now define a few of these averages and explore their significance.

Mean Value

The mean value (also called the expected value or average value) of random variable X is defined by

$$m_x = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{6.11}$$

Where E denotes the expectation operator. Note that m_x is a constant. If the function g(x) transforms X into some other random variable, then the expected value is given by,

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx \tag{6.12}$$

n^{th} moment of random variable X

The n^{th} moment of random variable is defined as the mean value of X^n. This can also be viewed as the special case where the transformation function g(x) is equal to X^n. Thus, the expression for the n_{th} moment of X is given by,

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)dx \tag{6.13}$$

The most widely used moments are

- The first moment (i.e., for n=1) which gives the mean value

$$m_x = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{6.14}$$

- The second moment (i.e., for n=2) which gives the mean square value

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)dx \tag{6.15}$$

Central moment

The central moment is the expected value of the difference between the random variable X and its mean value m_x. Thus, n_{th} central moment is defined as

$$E[(X - m_x)^n] = \int_{-\infty}^{\infty} (x - m_x)^n f_X(x)dx \tag{6.16}$$

i. **Variance**

From the Eq.(6.16) we see that the first central moment i.e., for n=1 is always zero because,

$$E[(X - m_x)] = \int_{-\infty}^{\infty} (x - m_x) f_X(x)dx = m_x - m_x = 0 \tag{6.17}$$

The second central moment i.e., for n=2, of a random variable X is called variance. The expression for variance is

$$Var = E[(X - m_x)^2] = \int_{-\infty}^{\infty} (x - m_x)^2 f_X(x)dx \tag{6.18}$$

Simplifying,

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

From linear property of expectation,

$$E[X^2 - 2m_x X + m_x^2] = E[X^2] - 2m_x E[X] + m_x^2$$

$$= E[X^2] - 2m_x^2 + m_x^2$$

$$E[X^2 - 2m_x X + m_x^2] = E[X^2] - m_x^2 \tag{6.19}$$

The symbol σ_x^2 is generally used to denote the variance.

The standard deviation σ_x of a random variable is defined as the square root of the variance σ_x^2. Thus standard deviation is expressed as

$$\sigma_x = \sqrt{E[X^2] - m_x^2} \tag{6.20}$$

The standard deviation σ_x of a random variable X is the measure of the width of its PDF. The larger the value of σ_x, the wider is the PDF. The value of standard deviation indicates the deviation in the values of the random variable from its mean value.

ii. **Covariance**

Covariance of random variables X and Y gives us some important information about the dependence between X and Y. Covariance is mathematically represented by the following Eq.(6.21).

$$\text{cov}[X,Y] = E[(X - m_x)(Y - m_y)] \tag{6.21}$$

Expanding the expression of covariance, we get

$$\text{cov}[X,Y] = E[(X - m_x)(Y - m_y)]$$

$$\text{cov}[X,Y] = E[XY] - E[Xm_y] - E[m_x Y] + E[m_x m_y]$$

$$\text{cov}[X,Y] = E[XY] - m_y E[X] - m_x E[Y] + m_x m_y$$

$$\text{cov}[X,Y] = E[XY] - m_x m_y - m_y m_x + m_x m_y$$

$$\text{cov}[X,Y] = E[XY] - m_x m_y \tag{6.22}$$

Thus, when $E[XY] = m_x m_y$, than, cov[X,Y]=0 and the random variables X and Y are said to be uncorrelated.

iii. Correlation coefficient

The covariance cov[X,Y], normalized with respect to the product $\sigma_x\sigma_y$ is termed as the correlation coefficient and is denoted by ρ, that is

$$\rho = \frac{E[XY]-m_x m_y}{\sigma_x \sigma_y} = \frac{\text{cov}[X,Y]}{\sigma_x \sigma_y} \tag{6.23}$$

Also we note that,

$$-1 < \rho < 1$$

where σ_x and σ_y are the standard deviation of random variable X and Y respectively.

Following are the basic statements related to the correlation between X and Y.

- X and Y are uncorrelated, if and only if cov[X,Y]=0.
- X and Y are orthogonal, if and only if E[X,Y]=0.

6.2 Random Process

The random process is a natural extension of the random variable (RV). Consider, for example, the temperature X of a certain city at noon. The temperature X is an RV and takes on different values every day. To get the complete statistics of X, we need to record values of X at noon over many days (a large number of trials). From this data, we can determine $f_X(x)$, the PDF of the RV X (the temperature at noon).

But the temperature is also a function of time. This means that the temperature at 1 p.m., may have an entirely different distribution from that of the temperature at noon. Still, the two temperatures may be related, via a joint probability density function. Thus, this random temperature X is a function of time and can be expressed as X(t). If the random variable is defined for a time interval $t \in [t_a, t_b]$, i.e., suppose from 6 am to 6 pm, then, $X(t)$ is a function of time and is random for every instant $t \in [t_a, t_b]$. An RV that is a function of time is called a random process, or stochastic process. Communication signals as well as noises, typically random and varying with time, are well characterized by random processes.

To specify an RV X, we run multiple trials of the experiment and from the outcomes estimate $f_X(x)$. Similarly, to specify the random process X(t), we do the same thing for each time instant t. To continue with our example of the random process X(t), the temperature of the city we need to record daily temperatures for each value of t (for each time of the day) This can be done by recording temperature at every instant of the day, which gives a waveform $X(t,\lambda_i)$, where λ_i indicates the day for which record was taken. This process is repeated everyday for large number of days. Each waveform thus obtained as shown in Fig.6.6 is a sample function (rather than sample point) of the random process and the collection of all possible waveforms (i.e., sample functions) is known as the ensemble (corresponding to the sample space) of the random process X(t).

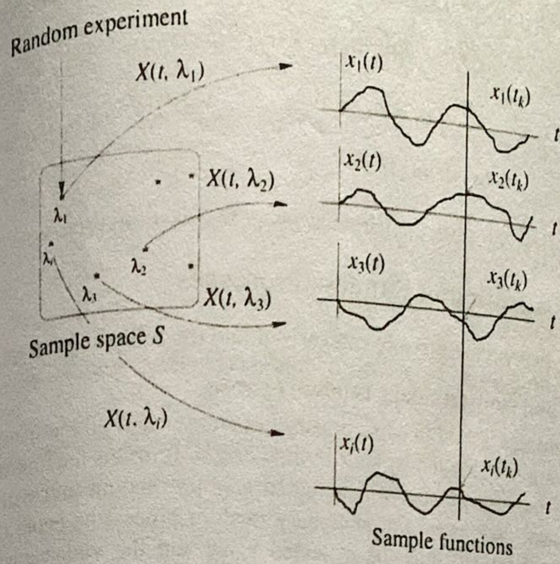

Fig.6.6 Random process

6.2.1 Ensemble Averages

For each time instance of a random process, the average value, variance etc. can be calculated from all sample functions $X(t,\lambda_i)$.

1. **Expected Value (E[X(t)])**

For a random process $X(t)$ with PDF $f_{X(t)}(x)$, the expected value is given by

$$E[X(t)] = \int_{-\infty}^{\infty} x\, f_{X(t)}(x)\,dx = m_X(t) \tag{6.24}$$

2. **Variance $[\sigma_X^2(t)]$**

$$\sigma_X^2(t) = E[|X(t)-m_X(t)|^2] = \int_{-\infty}^{\infty} |x-m_X(t)|^2 f_{X(t)}(x)\,dx \tag{6.25}$$

3. **Autocorrelation and Autocovariance**

We are interested in how the value of a random process X(t) evaluated at t_2 depends on its value at time t_1. At t_2 and t_2 the random process is characterized by random variables X_1 and X_2, respectively.

The relationship between X_1 and X_2 is given by the joint probability density function, $f_{X_1X_2}(x_1,x_2)$

Autocorrelation function

$$R_{XX}(t_1,t_2) = E[X_1X_2] \tag{6.26}$$

$$R_{XX}(t_1,t_2) = E[X(t_1)X_2(t_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\, f_{X_1X_2}(x_1,x_2)\,dx_1\,dx_2 \tag{6.27}$$

Autocovariance Function

$$C_{XX}(t_1,t_2) = E[((X(t_1)-m_X(t_1))((X(t_2)-m_X(t_2))]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-m_X(t_1))(x_2-m_X(t_2))\,f_{X_1X_2}(x_1,x_2)\,dx_1\,dx_2$$

$$= R_{XX}(t_1,t_2)-m_X(t_1)m_X(t_2)$$

$C_{XX}(t_1,t_2)$ describes the variance $\sigma_X^2(t)$ of a random process.

6.2.2 Classification of Random Processes

The random process are broadly categorized into the following categories

Stationary and Nonstationary random process

A random process is said to be stationary in strict sense if all the statistics of the process are not affected by shift in time origin. It means the time shift of a stationary random process will result in another random process having same statistical properties. For stationary random processthe entire PDF is independent of time t, thus the expected value and the variance are also constant over time.

$$f_{X(t)}(x)=f_{X(t+t_0)}(x) \tag{6.28}$$

$$m_X(t)=m_X(t+t_0)=m_X \tag{6.29}$$

$$\sigma_X(t)=\sigma_X(t+t_0)=\sigma_X \tag{6.30}$$

Autocorrelation and Autocovariance Function of a Stationary Random Process

The joint probability density function of a stationary process (with regards to the different realizations) does not change if a constant value t is added to both t_1 and t_2.

$$f_{X(t_1)X(t_2)}(x_1,x_2)=f_{X(t_1+t)X(t_2+t)}(x_1,x_2) \tag{6.31}$$

The autocorrelation function then only depends on the difference τ between t_1 and t_2.

$$R_{XX}(t_1,t_2)=E[X(t_1)X(t_2)] \tag{6.32}$$

$$R_{XX}(t_1,t_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1\,x_2\,f_{X(t_1)X(t_2)}(x_1,x_2)\,dx_1\,dx_2 \tag{6.33}$$

$$R_{XX}(t_1,t_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1\,x_2\,f_{X(0)X(t_2-t_1)}(x_1,x_2)\,dx_1\,dx_2 \tag{6.34}$$

$$R_{XX}(t_1,t_2)=R_{XX}(0,t_2-t_1)=R_{XX}(|t_2-t_1|)=R_{XX}(\tau) \tag{6.35}$$

Since the average value is constant, the autocovariance function is given by

$$C_{XX}(t_1,t_2)=E[((X(t_1)-m_X(t_1))((X(t_2)-m_X(t_2))] \tag{6.36}$$

$$C_{XX}(t_1,t_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-m_X(t_1))(x_2-m_X(t_2))\,f_{X_1X_2}(x_1,x_2)\,dx_1\,dx_2 \tag{6.37}$$

$$C_{XX}(t_1,t_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-m_X(t_1))(x_2-m_X(t_2))\,f_{X(0)X(t_2-t_1)}(x_1,x_2)\,dx_1\,dx_2 \tag{6.38}$$

$$C_{XX}(t_1,t_2)=C_{XX}(0,t_2-t_1)=C_{XX}(|t_2-t_1|)=C_{XX}(\tau) \tag{6.39}$$

Properties of the Autocorrelation function of a stationary random process

- Symmetry: $R_{XX}(t)=R_{XX}(-\tau)$
- Mean Square Average: $R_{XX}(0)=E[X(t)^2]\geq 0$
- Maximum: $R_{XX}(0)\geq |R_{XX}(\tau)|$
- Periodicity: if $R_{XX}(0)=R_{XX}(t_o)$. then $R_{XX}(\tau)$is periodic with period t_0.

The random process $X(t)$ representing the humidity of a city is an example of a **nonstationary** random process because the humidity statistics (i.e,. mean, for example) depends on the time of the day.

Wide-sense (or Weakly) stationary random process

A process that is not stationary in the strict sense, but still may have mean and autocorrelation function that are independent of the shift of time origin. A random process $X(t)$ is called WSS if the following three properties are satisfied:

- The average value of the random process is a constant:

  $m_X(t)=m_X$
- The autocorrelation and autocovariance function only depends on the time difference $\tau=t_1-t_2$.

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

  $C_{XX}(t_1,t_2)=C_{XX}(|t_2-t_1|)=C_{XX}(\tau)$
- The variance is a constant and finite:

  $\sigma_X^2=C_{XX}(0)=(R_{XX}(0)-m_X^2)<\infty$

Obviously all the strict sense stationary process (SSSP) is Wide Sense Stationary Process (WSSP) but vice versa is not always true.

6.2.3 Ergodic random process (Ergodicity)

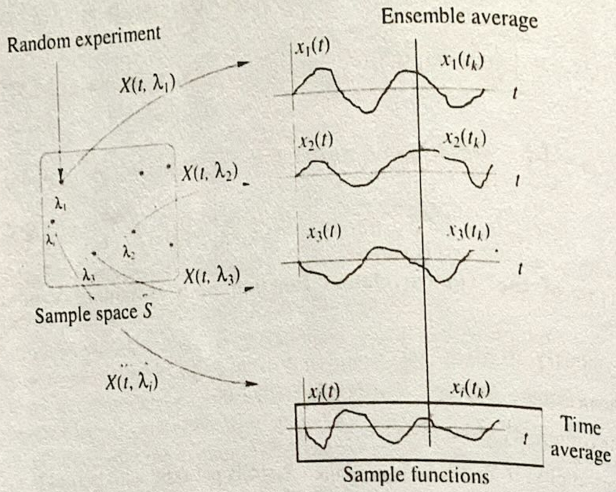

Fig.6.7 Ergodic Process

A wide sense stationary random process $X(t)$ is called ergodic, if the time average of each sample function $X(t, \lambda_i)$ converge towards the corresponding ensemble average.

It is know that the expectations or ensemble averages of a random process $X(t)$ are averages "**across the process.**" For example, the mean of a stochastic process $X(t)$ at some fixed time is the expectation of the random variable $X(t_k)$ that describes all possible values of sample functions of the process $X(t)$ sampled at time $t = t_k$. Naturally, we may also define long-term sample averages or time averages that are averages "**along the process.**" Whereas in ensemble averaging a set of independent realizations of the process $X(t)$ sampled at some fixed time t_k is considered, in time averaging the focus is on a single waveform evolving across time t and representing one waveform realization of the process $X(t)$.

With time averages providing the basis of a practical method for possible estimation of ensemble averages of a stochastic process, we would like to explore the conditions under which this estimation is justifiable. To address this important issue, consider the sample function $X(t, \lambda_i)$ of a weakly stationary process $X(t)$ observed over the interval $-T \le t \le T$. The time-average value of the sample function $X(t, \lambda_i)$ is defined by the definite integral.

$$m_{X(\lambda_i)} = \frac{1}{2T}\int_{-T}^{T} X(t,\lambda_i)\,dt \tag{6.40}$$

Clearly, the time average $mX_{(\lambda_i)}$ is a random variable, as its value depends on the observation interval and which particular sample function of the process $X(t)$. Since the process $X(t)$ is assumed to be weakly stationary, the mean of the time average ($mX_{(\lambda_i)}$) is given by (after interchanging the operations of expectation and integration. which is permissible because both operations are linear)

$$E[m_{X(\lambda_i)}] = \frac{1}{2T}\int_{-T}^{T} E[X(t,\lambda_i)]\,dt \tag{6.41}$$

$$E[m_{X(\lambda_i)}] = \frac{1}{2T}\int_{-T}^{T} m_X\,dt \tag{6.42}$$

$$E[m_{X(\lambda_i)}] = m_X \tag{6.43}$$

where m_X is the mean of the process $X(t)$. Accordingly, the time average $mX_{(\lambda_i)}$ represents an unbiased estimate of the ensemble-averaged mean m_X. Most importantly, we say that the process $X(t)$ is ergodic in the mean if two conditions are satisfied:

- The time average $mX_{(\lambda_i)}$ approaches the ensemble average in the limit as the observation interval approaches infinity; that is,

$$\lim_{T\to\infty} m_{X(\lambda_i)} = m_X$$

- The variance of $mX_{(\lambda_i)}$ (i.e., $\sigma^2_{X(\lambda_i)}$) approaches the ensemble average in the limit as the observation interval approaches infinity; that is,

$$\lim_{T\to\infty} \sigma^2_{X(\lambda_i)} = 0$$

The other time average of particular interest is the autocorrelation function $R_{XX}(t, \lambda_i)$, defined in terms of the sample function $X(t, \lambda_i)$ observed over the interval $-T \le t \le T$. We may define the time-averaged autocorrelation function of $X(t, \lambda_i)$ as

$$R_{XX}(\tau,\lambda_i) = \frac{1}{2T}\int_{-T}^{T} x(t,\lambda_i)\,x(t+\tau,\lambda_i)\,dt = \frac{1}{2T}\int_{-T}^{T} x(t)\,x(t+\tau)\,dt \tag{6.44}$$

This second time average should also be viewed as a random variable with a mean and variance of its own. In a manner similar to ergodicity of the mean, we say that the process $X(t, \lambda_i)$ is ergodic in the autocorrelation function if the following two limiting conditions are satisfied:

$$\lim_{T\to\infty} R_{XX}(\tau,\lambda_i) = R_{XX}(\tau) \tag{6.45}$$

$$\lim_{T\to\infty} \mathrm{var}[R_{XX}(\tau,\lambda_i)] = 0 \tag{6.46}$$

With the property of ergodicity confined only to the mean and autocorrelation functions, it follows that all ergodic processes are weakly stationary. In communications, all the information bearing signals are considered to be WSSP and furthermore the ergodic. With these assumptions, the time averaged values of a single realization of the signal is sufficient to describe the statistical properties of the signal.

6.3 Power Spectral Density Function

The power spectral density function (psdf) of a WSS random process $X(t)$ is defined as the Fourier Transform of the autocorrelation (AC) function $R_{XX}(\tau)$. In other words the psdf and the AC function of a WSSP are Fourier Transform pairs.

$$S_{XX}(w) = F\{R_{XX}(\tau)\} = \int_{-\infty}^{\infty} R_{XX}(\tau)\exp(-jw\tau)\,d\tau \tag{6.47}$$

Inverse transform:

$$R_{XX}(\tau) = F^{-1}\{(S_{XX}(w)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(w)\exp(jw\tau)\,dw \tag{6.48}$$

Properties of Power Spectral Density

- $S_{XX}(w) = S_{XX}(-w)$
- $S_{XX}(w) \geq 0$
- $\mathrm{Im}\{S_{XX}(w)\} = 0$

6.4 Noise in communication

1. **White noise (WN)**

It is an ideal case of representing noise in communication. In communications system noise analysis is generally based on an white noise. The white noise has flat spectrum density, $N_W(f)$ over $-\infty < f < \infty$. The flat spectrum density of WN is expressed as

$$N_W(f) = \frac{N_0}{2}, \; for \; -\infty < f < \infty \tag{6.49}$$

$$N_W(f) = N_0, \; for \; 0 \leq f < \infty \; or \; -\infty < f < 0 \tag{6.50}$$

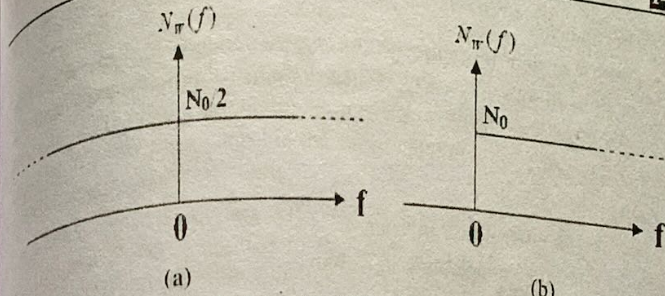

Fig.6.8 (a) Power Spectral Density of White Noise ($N_W(f)$) over $-\infty < f < \infty$.
(b) Power Spectral Density of White Noise ($N_W(f)$) over $0 < f < \infty$

The average or mean value of noise is zero. If the probability of occurrence of a white noise is specified by a Gaussian distribution function, it is called white Gaussian noise.

Autocorrelation function of white noise may be obtained simply by taking the inverse Fourier Transform of psdf.

$$R_{WN}(\tau) = F^{-1}[N_W(f)] \tag{6.51}$$

$$R_{WN}(\tau) = \frac{N_0}{2}\delta(t) \tag{6.52}$$

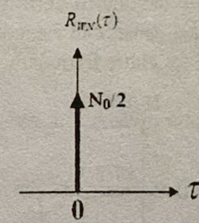

Fig.6.9 Autocorrelation function of white noise

From Eq.(6.52), it is evident that

$$R_{WN}(\tau) = 0 \;\; for \;\; \tau \neq 0$$

It indicates that any two different samples of white noise, no matter how close they are in time shift ($\tau \to 0$), are uncorrelated i.e., the white noise is an example of a perfect random process. White noise has no physical significance, as it does not exist in nature. Because of its convenient mathematical properties, the white noise is used to model the noise.

The average power (P_{WN}) of white noise is infinity because the bandwidth is infinity.

$$P_{WN} = \int_{-\infty}^{\infty} \frac{N_0}{2}\,df = \infty \tag{6.53}$$

2. **Thermal Noise**

Thermal or Johnson noise occurs due to random movement of free electrons in any conducting material due to thermal energy and the noise power is equal to

$$P_n = kTB \tag{6.54}$$

Where

P_n the noise power in Watt.

T absolute temperature in Kelvin.

B system bandwidth in Hz.

K Boltzmann's constant (1.38×10^{-23} Joule/Kelvin).

The Power spectral density function (psdf) of thermal noise can be approximated as

$$N_T(f) = \frac{hf}{2(e^{hf/kT}-1)} \quad watt/Hz \tag{6.55}$$

h = plank's constant

And at the origin

$$N_T(f) = \frac{kT}{2} \quad at \; f=0 \tag{6.56}$$

is maximum and gradually decline to zero as frequency increases to infinity. Therefore, thermal noise can be considered as white noise with flat spectrum of $kT/2$ for entire frequency range of interest.

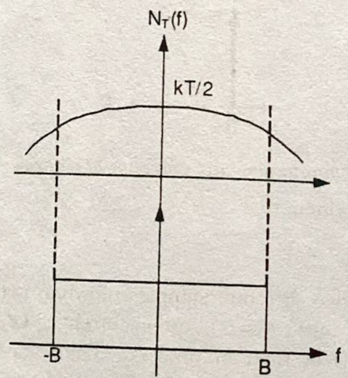

Fig.6.10 psd of Thermal Noise

As the thermal noise is universally present in all communication system, the thermal noise characteristic- additive, white, and Gaussian are most often used to model the noise in communication system, and called as Additive White Gaussian Noise (AWGN)

6.5 Passage of random signal and noise through LTI system

In communications engineering, the information bearing signals with given statistical properties are passed/proceesed, generally, through a linear time invariant (*LTI*) system or combination of the systems.It is therefore becomes essential to know the changes in the statistical properties of the signals when passed through various systems.

Let a Wide Sense Stationary *RP* (WSSP) $X(t)$ with mean value m_x and the AC function $R_{xx}(\tau)$ is applied to the input of a LTI system with impulse response $h(t)$ and transfer function $H(f)$. Now let us find the statistical properties of the signal at the output of the system $Y(t)$.

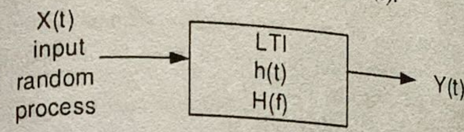

Fig.6.11 Determination of response of a LTI system to a random input

By definition, output of LTI system is

$$Y(t) = X(t) \otimes h(t) \tag{6.57}$$

$$Y(t) = \int_{-\infty}^{\infty} X(t-\alpha)h(\alpha)d\alpha \tag{6.58}$$

$$Y(t) = \int_{-\infty}^{\infty} X(\alpha)h(t-\alpha)d\alpha \tag{6.59}$$

We now determine the Mean and Autocorrelation function of a RP at the output of the system.

Mean value is given by the expectation operation

$$m_y = E[Y(t)] = E[\int_{-\infty}^{\infty} X(t-\alpha)h(\alpha)d\alpha] \tag{6.60}$$

As expectation operator E[.] is linear operation.

$$m_y = \int_{-\infty}^{\infty} E[X(t-\alpha)]h(\alpha)d\alpha \tag{6.61}$$

$$m_y = m_x \int_{-\infty}^{\infty} h(\alpha)d\alpha \tag{6.62}$$

Because as input RP X(t) is WSSP, the mean m_x is a constant and independent of time shift.

$$[E[X(t-\alpha)] = E[X(t)] = m_x]$$

We know

$$H(f) = \int_{-\infty}^{\infty} h(\alpha) e^{-j2\pi f\alpha} d\alpha \tag{6.63}$$

For f=0

$$H(0) = \int_{-\infty}^{\infty} h(\alpha) d\alpha \tag{6.64}$$

So from Eq.(6.67) and Eq.(6.69) we get

$$m_y = m_x H(0) \tag{6.65}$$

Where H(0) is zero frequency response of the system. It is obvious from the Eq. (6.65) that the mean value of output is also independent of the shift in time.

To estimate theautocorrelation function of Y(t). let us first determine the cross correlation between X(t) and Y(t)

$$R_{xy}(t_1,t_2) = R_{xy}(\tau) = E[X(t_1)Y(t_2)] \tag{6.66}$$

$$R_{xy}(t_1,t_2) = E[X(t_1)\int_{-\infty}^{\infty} X(s)h(t_2 - s)ds] \tag{6.67}$$

$$R_{xy}(t_1,t_2) = \int_{-\infty}^{\infty} E[X(t_1)X(s)]h(t_2 - s)ds \tag{6.68}$$

$$R_{xy}(t_1,t_2) = \int_{-\infty}^{\infty} R_{xx}(t_1 - s)h(t_2 - s)ds \tag{6.69}$$

Let. us introduce new variable 'u' such that $s-t_2=u$. then

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} R_{xx}(\tau - u)h(-u)du \tag{6.70}$$

Thus the cross correlation between $X(t)$ and $Y(t)$

$$R_{xy}(\tau) = R_{xx}(\tau) \otimes h(-\tau) \tag{6.71}$$

Now. the autocorrelation function of $Y(t)$ is

$$R_{yy}(t_1,t_2) = R_{yy}(\tau) = E[Y(t_1)Y(t_2)] \tag{6.72}$$

$$R_{yy}(t_1,t_2) = E[Y(t_2)\int_{-\infty}^{\infty} X(s)h(t_1 - s)ds] \tag{6.73}$$

$$R_{yy}(t_1,t_2) = \int_{-\infty}^{\infty} E[Y(t_2)X(s)]h(t_1 - s)ds \tag{6.74}$$

$$R_{yy}(t_1,t_2) = \int_{-\infty}^{\infty} R_{xy}(s - t_2)h(t_1 - s)ds \tag{6.75}$$

Again, let $u=s-t_2$, than

$$R_{yy}(t_1,t_2) = \int_{-\infty}^{\infty} R_{xy}(u)h(t_1 - t_2 - u)du \tag{6.76}$$

$$R_{yy}(t_1,t_2) = \int_{-\infty}^{\infty} R_{xy}(u)h(\tau - u)du \tag{6.77}$$

$$R_{yy}(t_1,t_2) = R_{xy}(\tau) \otimes h(\tau) \tag{6.78}$$

So,

$$R_{yy}(t_1,t_2) = R_{yy}(\tau) = R_{xx}(\tau) \otimes h(-\tau) \otimes h(\tau) \tag{6.79}$$

As seen from above equation. the autocorrelation function of output Y(t) is independent of shift in time (only function of time lag $\tau = t_1 - t_2$)

Thus from Eq.(6.70) and Eq.(6.84) it can be concluded that the output of a LTI to a WSSP excitation is also WSSP.

Relation between Power Spectral Density, Autocorrelation function and system Transfer Function

If we have to find the spectrum of output $Y(t)$, we can compute the power spectrum from autocorrelation function as they form a Fourier transform pair.

$$S_y(f) = FT[R_{yy}(\tau)] \tag{6.80}$$

$$S_y(f) = FT[R_{xx}(\tau) \otimes h(-\tau) \otimes h(\tau)] \tag{6.81}$$

$$S_y(f) = FT[R_{xx}(\tau)] \times FT[h(-\tau)] \times FT[h(\tau)] \tag{6.82}$$

We know. $FT[h(\tau)] = H(f)$, then $FT[h(-\tau)] = H^*(f)$. (i.e., complex conjugate of H(f)). Now. Eq.(6.82) can be written as

$$\therefore S_y(f) = S_x(f) \times H(f) \times H^*(f) \tag{6.83}$$

$$S_y(f) = S_x(f) \times |H(f)|^2 \tag{6.84}$$

6.6 Ideal low pass filtering of white noise

As with the signal, the noise is also passed through the system/s and it is equally essential to analyze how the system/s respond to the noise at its input.

Let us apply White Noise with *psdf* $N_W(f)=N_0/2$ to an ideal low-pass flter with bandwidth B.

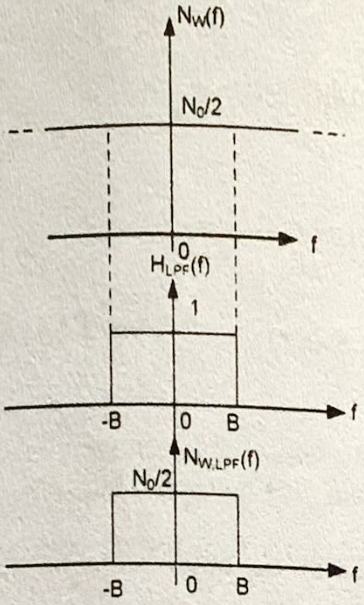

Fig.6.12 White Noise after passing through LPF

From Fig.6.12 it can be seen that the psdf at the output of the filter is

$$N_{LFP}(f) = \begin{cases} \frac{N_0}{2}, & -B < f < B \\ 0, & otherwise \end{cases} \tag{6.85}$$

Now the autocorrelation function is

$$R_{W.LPF}(\tau) = F^{-1}\{(S_{yy}(f)\} = F^{-1}\{N_{LFP}(f)\} \tag{6.86}$$

$$R_{W.LPF}(\tau) = \int_{-B}^{B} N_{LFP}(f)\exp(j2\pi f\tau)\,df \tag{6.87}$$

$$R_{W.LPF}(\tau) = \int_{-B}^{B} \frac{N_0}{2}\exp(j2\pi f\tau)\,df \tag{6.88}$$

$$R_{W.LPF}(\tau) = N_0 B \sin c(2B\tau) \tag{6.89}$$

The autocorrelation function is plotted in Fig.6.13.

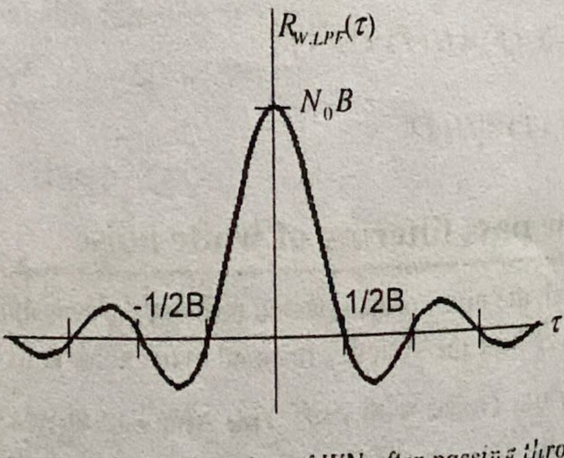

Fig.6.13. Autocorrelation function of WN after passing through LPF.