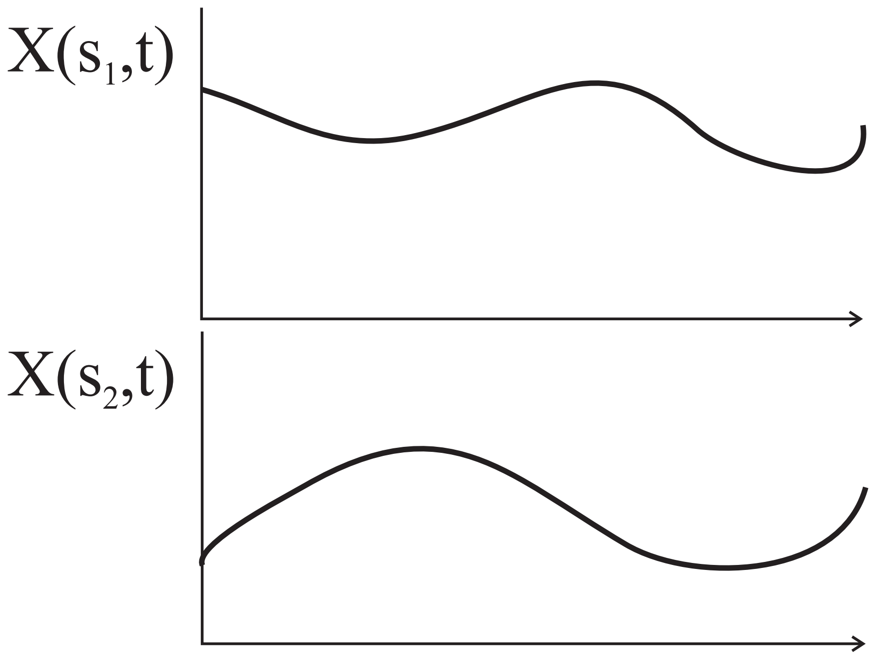
ELEC 400: Random Signals
Lecture Notes Set 14

1 Random Processes

A useful extension of the idea of random variables is the random process. While the random variable X is defined as a univariate function $\mathrm{X}(s)$ where s is the outcome of a random experiment, the random process is a bivariate function $\mathrm{X}(s,t)$ where s is the outcome of a random experiment and t is an index variable such as time. Examples of random processes are the voltages in a circuit over time, light intensity over location. The random process, for two outcomes s_1 and s_2 can be plotted as



Just as for random variables, usually the s is not explicitly written in $\mathrm{X}(s,t)$ so the random process is denoted as $X(t)$.

1.1 Classification of Random Processes

Random processes are classified according to the type of the index variable and classification of the random variables obtained from samples of the random process. The major classification are given below:

Name	Domain of t	Classification of $\mathrm{X}(t,s)$ for a fixed t
Continuous Random Process	all $t \in [-\infty, \infty]$	Continuous random variable
Discrete Random Process	all $t \in [-\infty, \infty]$	Discrete random variable
Continuous Random Sequence	countable set: $\{t_1, t_2, ...\}$	Continuous Random Variable
Discrete Random Sequence	countable set: $\{t_1, t_2, ...\}$	Discrete Random Variable

Examples of each type of random process are:

Continuous Random Process: Voltage in a circuit, temperature at a given location over time, temperature at different positions in a room.

Discrete Random Process: Quantized voltage in a circuit over time.

Continuous Random Sequence: Sampled voltage in a circuit over time.

Discrete Random Sequence: Sampled and quantized voltage from a circuit over time.

1.2 Deterministic and Non-deterministic Random Processes

A random process is called deterministic if future values of a random process can be perfectly predicted from past values. If a process does not have this property it is called non-deterministic.
Example: A random process over time is defined as

$$\mathrm{X}(t) = A\cos(\omega_0 t + \Theta)$$

where A and ω_0 are known constants and Θ is a random variable. Since from a few samples of $\mathrm{X}(t)$ taken at different known times it is possible to calculate θ and thus determine the sample function of $\mathrm{X}(t)$ for all future values of t, this process is deterministic.

2 Stationarity and Independence

A random process is called stationary if its statistical properties do not change over time. For example, ideally, a lottery machine is stationary in that the properties of its random number generator are not a function of when the machine is activated. The temperature random process for a given outdoor location over time is not stationary when considered over the period of a whole year as the mean temperature at different times of the year will vary. If a random process is not stationary it is called non-stationary. This is a good natural language definition but we require a more rigorous mathematical definition. This motivates us to come up with a good method of describing random processes in a mathematical way.

2.1 Distribution and Density Functions of Random Processes

We will designate the cumulative distribution function (CDF) of random process $\mathrm{X}(t)$ at time t_1 as

$$\mathrm{F}_X(x_1; t_1) = \mathrm{P}\left[X(t_1) \le x_1\right] \tag{1}$$

We can extend this to second order or higher distribution functions easily. The second order distribution function for random process $\mathrm{X}(t)$ is given by

$$\mathrm{F}_X(x_1, x_2; t_1, t_2) = \mathrm{P}\left\{X(t_1) \le x_1 \text{ and } X(t_2) \le x_2\right\} \tag{2}$$

The N^{th} order distribution function for random process $\mathrm{X}(t)$ is

$$\mathrm{F}_X\left(x_1, ..., x_N; t_1, ..., t_N\right) = \mathrm{P}\left\{X\left(t_1\right) \le x_1, ..., X\left(t_N\right) \le x_N\right\} \tag{3}$$

We can extend our idea of probability density functions (PDFs) from random variables to density functions for random processes. The first order density function for random process $\mathrm{X}(t)$ is then

$$\mathrm{f}_X(x_1;t_1) = \frac{\partial}{\partial x_1}\,\mathrm{F}_X\left(x_1;t_1\right) \tag{4}$$

The second order density function is then

$$\mathrm{f}_X(x_1,x_2;t_1,t_2) = \frac{\partial^2}{\partial x_1 \partial x_2}\,\mathrm{F}_X\left(x_1,x_2;t_1,t_2\right) \tag{5}$$

By extension, the N^{th} order density function is then

$$f_X\left(x_1,...,x_N;t_1,...,t_N\right) = \frac{\partial^N}{\partial x_1 ... \partial x_N}\,\mathrm{F}_X\left(x_1,...,x_N;t_1,...,t_N\right) \tag{6}$$

2.1.1 Independent Processes

Two random process $\mathrm{X}(t)$ and $\mathrm{Y}(t)$ are called independent if all possible random variables generated by sampling from $\mathrm{X}(t)$ are independent of all possible random variables generated by sampling from $\mathrm{Y}(t)$. This definition can be rewritten in terms of distribution and density functions so we can apply it mathematical problem solving.
If we define the N^{th} order joint CDF of $\mathrm{X}(t)$ and $\mathrm{Y}(t)$ as

$$\begin{aligned}&\mathrm{F}_{X,Y}\left(x_1,...,x_N,y_1,...,y_N;t_1,...t_N,t'_1,...,t'_N\right)\\ &\qquad = \mathrm{P}\left\{\mathrm{X}\left(t_1\right)\le x_1,...,\mathrm{X}\left(t_N\right)\le x_N,\mathrm{Y}\left(t'_1\right)\le y_1,...,\mathrm{Y}\left(t'_N\right)\le y_N\right\}\end{aligned} \tag{7}$$

We can define the joint PDF of $\mathrm{X}(t)$ and $\mathrm{Y}(t)$ as

$$\begin{aligned}&\mathrm{f}_{X,Y}\left(x_1,...,x_N,y_1,...,y_N;t_1,...,t_N,t'_1,...,t'_N\right)\\ &\qquad = \frac{\partial^{2N}}{\partial x_1...\partial x_N \partial y_1 ... \partial y_N}\,\mathrm{F}_{X,Y}\left(x_1,...,x_N,y_1,...,y_N;t_1,...,t_N,t'_1,...,t'_N\right)\end{aligned} \tag{8}$$

Independence of $\mathrm{X}(t)$ and $\mathrm{Y}(t)$ can then be redefined as allowing (8) to be factored as

$$\begin{aligned}&\mathrm{f}_{X,Y}\left(x_1,...,x_N,y_1,...,y_N;t_1,...,t_N,t'_1,...,t'_N\right)\\ &\qquad = \mathrm{f}_X\left(x_1,...,x_N;t_1,...t_N\right)\mathrm{f}_Y\left(y_1,...,y_N;t'_1,...t'_N\right)\end{aligned} \tag{9}$$

for all selections of x_1,...,x_N,y_1,...,y_N,t_1,...,t_N, t'_1,...,t'_N, and N.

2.2 First Order Stationarity

Now that we have a notation for describing the density of random processes, we can address the definitions of stationarity. A random process $\mathrm{X}(t)$ is called stationary to order one if its first order density function does not change with a shift in time, or in terms of our density notation:

$$\mathrm{f}_X\left(x_1;t_1\right) = \mathrm{f}_X\left(x_1;t_1+\Delta\right) \tag{10}$$

for all x_1, t_1 and Δ. If $\mathrm{X}(t)$ is stationary to order random variables $X_1 = \mathrm{X}(t_1)$ and $X_2 = \mathrm{X}(t_2)$ will have the same PDF for any selection of t_1 and t_2. This means that the expectation of any function of $\mathrm{X}(t)$ will be a constant over t. That is,

$$\mathrm{E}\left\{\mathrm{g}\left[\mathrm{X}\left(t_1\right)\right]\right\}\mathrm{E}\left\{\mathrm{g}\left[\mathrm{X}\left(t_2\right)\right]\right\} \tag{11}$$

for any function $\mathrm{g}\left(\cdot\right)$, t_1 and t_2.

2.3 Second Order Stationarity

A common mistake when first working with random processes is to mistake first order stationarity for general stationarity. It can be easily shown that first order stationarity is not enough to ensure stationarity of all statistical properties.
Example: A random process is given by

$$\mathrm{X}(t) = N$$

where N is a Gaussian random variable with mean 0 and variance 1:

$$\mathrm{f}_N(n) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{n^2}{2}\right)$$

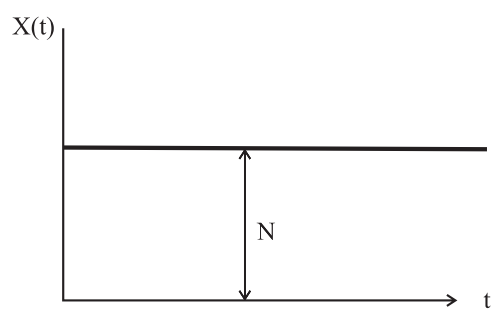
This process is constant over all time but different instances of the random process have different values. The first order density function of $\mathrm{X}(t)$ is then

$$\mathrm{f}_X(x_1; t_1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_1^2}{2}\right)$$

and the second order density function is

$$\begin{aligned} \mathrm{f}_X(x_1, x_2; t_1, t_2) &= \mathrm{f}_X(x_2; t_2 | \mathrm{X}(t_1) = x_1) \, \mathrm{f}_X(x_1; t_1) \\ &= \delta(x_2 - x_1) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_1^2}{2}\right) \end{aligned}$$

where the second line is a result of $\mathrm{X}(t_1) = \mathrm{X}(t_2)$ for all t_1 and t_2 from the definition of the random process. One sample function of $\mathrm{X}(t)$ is shown below:



One sample function of $\mathrm{X}(t)$

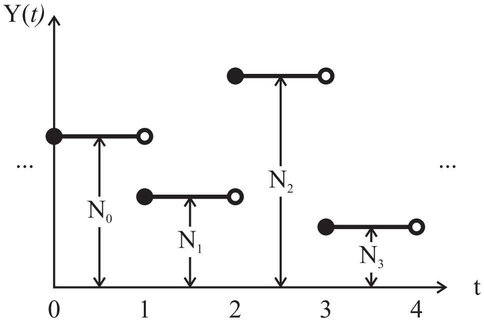
Example 2: A second random process, $\mathrm{Y}(t)$, is defined as

$$\mathrm{Y}(t) = N_i \text{ for } i \le t < i + 1$$

where N_i for $i = ..., -2, -1, 0, 1, 2...$ are independent and identically distributed Gaussian random variables with PDFs of

$$\mathrm{f}_{N_i}(n_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{n_i^2}{2}\right)$$

We get a random process which is constant for period of length 1 but changes value at every integer time index, with its value being independent in each interval. One sample function of $\mathrm{Y}(t)$ is shown below:



One sample function of $\mathrm{Y}(t)$

The first order density function of $\mathrm{Y}(t)$ is given by

$$\mathrm{f}_Y\left(y_1;t_1\right)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{{y_1}^2}{2}\right)$$

The second order density function of $\mathrm{Y}(t)$ given in terms of conditional PDFs is

$$\mathrm{f}_Y\left(y_1,y_2;t_1;t_2\right)=\mathrm{f}_Y\left(y_2;t_2\middle|\,\mathrm{Y}(t_1)=y_1\right)\mathrm{f}_Y(y_1;t_1) \tag{12}$$

We note that $\mathrm{Y}(t_1)=\mathrm{Y}(t_2)$ if t_1 and t_2 are located between the same two integers: $i\le t_1,t_2<i+1$ for some i. If this case is not true, then $\mathrm{Y}(t_1)$ and $\mathrm{Y}(t_2)$ are independent. The second order density function of $\mathrm{Y}(t)$ is then

$$\mathrm{f}_Y\left(y_1,y_2;t_1;t_2\right)=\begin{cases}\delta\left(y_1-y_2\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{{y_1}^2}{2}\right) & \text{if } i\le t_1,t_2<i+1 \text{ for some } i\\ \frac{1}{2\pi}\exp\left(-\frac{{y_1}^2+{y_2}^2}{2}\right) & \text{otherwise}\end{cases}$$

Note from the previous examples, that two functions can have the same first order density function but different second order density functions. This motivates the definition for different orders of stationarity.

A random process $\mathrm{X}(t)$ is called second order stationary or stationary to order two if

$$\mathrm{f}_X\left(x_1,x_2;t_1,t_2\right)=\mathrm{f}_X\left(x_1,x_2;t_1+\Delta,t_2+\Delta\right) \tag{13}$$

for all possible selections of x_1, x_2, t_1, t_2 and Δ. It can be easily seen that second order stationarity implies first order stationarity but the reverse is not true.

To study second order stationarity some useful functions have been developed. The first of these is the autocorrelation function of a random process which is defined for a random process $\mathrm{X}(t)$ as

$$\mathrm{R}_{XX}\left(t_1,t_2\right)=\mathrm{E}\left[\mathrm{X}\left(t_1\right)\mathrm{X}\left(t_2\right)\right] \tag{14}$$

If $\mathrm{X}(t)$ is stationary to order two then it can be seen that

$$\begin{aligned}\mathrm{R}_{XX}\left(t_1,t_2\right) &= \mathrm{E}\left[\mathrm{X}\left(t_1\right)\mathrm{X}\left(t_2\right)\right]\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2\,\mathrm{f}_X\left(x_1,x_2;t_1,t_2\right)\mathrm{d}x_1\mathrm{d}x_2\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_1x_2\,\mathrm{f}_X\left(x_1,x_2;t_1,t_1+\tau\right)\mathrm{d}x_1\mathrm{d}x_2 \text{ where } \tau=t_2-t_1\\ &= \mathrm{E}\left[\mathrm{X}\left(t_1\right)\mathrm{X}\left(t_1+\tau\right)\right]\ \text{ for any } t_1\\ &= \mathrm{R}_{XX}\left(\tau\right) \end{aligned} \tag{15}$$

2.4 Wide Sense Stationarity

It should be noted that all processes that are stationary to order two have the property that $\mathrm{R}_{XX}(t_1, t_2) = \mathrm{R}_{XX}(t_2 - t_1)$ but the converse is not true. This property is useful so processes that have this property are given a special name, Wide Sense Stationary. A random process is called Wide Sense Stationary if

$$\begin{aligned} \mathrm{E}[\mathrm{X}(t)] &= \overline{X}, \text{ a constant over all } t, \text{ and} \qquad (16) \\ \mathrm{R}_{XX}(t_1, t_2) &= \mathrm{R}_{XX}(\tau) \text{ where } \tau = t_2 - t_1 \qquad (17) \end{aligned}$$

Example: A random process $\mathrm{X}(t)$ is defined as

$$\mathrm{X}(t) = A\cos(\omega t + \phi)$$

where A and ω are constants and ϕ is a random variable that is uniformly distributed from 0 to 2π. The expected value of $\mathrm{X}(t)$ is

$$\mathrm{E}[\mathrm{X}(t)] = \int_0^{2\pi} \frac{1}{2\pi} A\cos(\omega t + \phi)\,\mathrm{d}\phi = 0$$

The autocorrelation function is given by

$$\begin{aligned} \mathrm{R}_X X(t, t+\tau) &= \mathrm{E}[\mathrm{X}(t)\,\mathrm{X}(t+\tau)] \\ &= \mathrm{E}\{A\cos(\omega t + \phi)\,A\cos[\omega(t+\tau)+\phi]\} \\ &= \mathrm{E}\left\{\frac{A^2}{2}[\cos(2\omega t + \omega\tau + \phi) + \cos(-\omega\tau)]\right\} \\ &= \frac{A^2}{2}\mathrm{E}[\cos(2\omega t + \omega\tau + \phi)] + \frac{A^2}{2}\mathrm{E}[\cos(\omega\tau)] \\ &= \frac{A^2}{2}(0) + \frac{A^2}{2}\cos(\omega\tau) \\ &= \frac{A^2}{2}\cos(\omega\tau) \end{aligned}$$

The mean is a constant and the autocorrelation function is only a function of τ so this process is Wide Sense Stationary. It is easily seen that if ϕ is uniformly distributed in $\left[0, \frac{\pi}{4}\right]$ the process is not Wide Sense Stationary as the autocorrelation in this case is a function of both t_1 and t_2.

Second Order Stationarity is a sufficient but not necessary condition for Wide Sense Stationarity. That is, there exist processes which are Wide Sense Stationary but not second order stationary but the reverse is not true.

Two random processes $\mathrm{X}(t)$ and $\mathrm{Y}(t)$ are called jointly Wide Sense Stationary if they are individually Wide Sense Stationary and

$$\mathrm{R}_{XY}(t, t+\tau) = \mathrm{E}[\mathrm{X}(t)\,\mathrm{Y}(t+\tau)] = \mathrm{R}_{XY}(\tau)$$

2.5 Nth Order Stationarity

A random process is called stationary to order N or N^{th} order stationary if

$$\mathrm{f}_X(x_1, ..., x_N; t_1, ..., t_N) = \mathrm{f}_X(x_1, ..., x_N; t_1 + \Delta, ..., t_N + \Delta)$$

for all possible x_1, ..., x_N, t_1, ..., t_N, and Δ. It is easy to see if a random process is stationary to order N it is also stationary to all orders less than N. If a random process is stationary to orders 1,2 up to infinity it is called strictly stationary.

3 Time Averages and Ergodicity

Define the time average operator as

$$\mathrm{A}\left[\cdot\right] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \left[\cdot\right] \mathrm{d}t \tag{18}$$

We denote a single sample function of random process $\mathrm{X}(t)$ as $\mathrm{x}(t)$. The time average of a sample function $\mathrm{x}(t)$ is denoted as

$$\overline{\mathrm{x}} = \mathrm{A}\left[\mathrm{x}(t)\right] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{x}(t)\mathrm{d}t \tag{19}$$

We will also define a time average autocorrelation function as

$$\mathcal{R}_{XX}\left(\tau\right) = \mathrm{A}\left[\mathrm{x}(t)\,\mathrm{x}(t+\tau)\right] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{x}(t)\,\mathrm{x}(t+\tau)\mathrm{d}t \tag{20}$$

It can be seen that in general $\overline{x}$ and $\mathcal{R}_{XX}\left(\tau\right)$ are random variables. It can be easily shown that

$$\mathrm{E}\left[\overline{x}\right] = \overline{X} \tag{21}$$

$$\mathrm{E}\left[\mathcal{R}_{XX}\left(\tau\right)\right] = \mathrm{R}_{XX}\left(\tau\right) \tag{22}$$

Assume that there was some theorem or set of properties of $\mathrm{X}(t)$ that make $\overline{x}$ and $\mathcal{R}_{XX}\left(\tau\right)$ constants for all sample functions $\mathrm{x}(t)$ of $\mathrm{X}(t)$, so that $\overline{x} = \overline{X}$ and $\mathcal{R}_{XX}\left(\tau\right) = \mathrm{R}_{XX}\left(\tau\right)$. We call processes that have these properties ergodic. In natural language, ergodic processes have their time averages equal to their statistical averages. Ergodicity is a restrictive form of stationarity. It is very difficult to prove mathematically and impossible to prove experimentally. It is often, however, assumed to be true for a given observed process to make useful kinds of statistical manipulations possible. For example, any time you take multiple measurements of a single process at different times and average them together to calculate an estimate of the mean of signal you are assuming that the process being observed is in some way ergodic.

We call two processes jointly ergodic if they are individually ergodic and if

$$\mathcal{R}_{XY}\left(\tau\right) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{x}(t)\,\mathrm{y}(t+\tau)\mathrm{d}t = \mathrm{R}_{XY}\left(\tau\right) \tag{23}$$

3.1 Mean Ergodic Processes

A process with a mean value $\overline{X}$ which is not dependent on t is called mean ergodic or ergodic in the mean if its statistical average, $\overline{X} = \mathrm{E}\left[X\right]$, equals the time average, $\overline{x} = \mathrm{A}\left[\mathrm{x}(t)\right]$ of any sample function $\mathrm{x}(t)$ with probability 1.

Assume:

1. $\mathrm{E}[\mathrm{X}(t)] = \overline{X} < \infty$
2. $\mathrm{X}(t)$ is bounded, which means that for all sample functions $\mathrm{x}(t)$, $|\,\mathrm{x}(t)| < \infty$ for all t.
3. $\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{E}[|\,\mathrm{X}(t)|] < \infty$
4. $\mathrm{E}\left[|\mathrm{X}(t)|^2\right] = \mathrm{R}_{XX}(t,t) = \mathrm{E}\left[\mathrm{X}(t)^2\right] < \infty$. A random process that satisfies this is called a regular process.

The first three properties are required to allow us to exchange statistical average and time average integrals for these random processes.
Define a random variable A_x from $\mathrm{X}(t)$ as

$$A_x = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{X}(t)\mathrm{d}t \tag{24}$$

$$\begin{aligned} \overline{A}_x &= \mathrm{E}[A_x] = \mathrm{E}\left\{\lim_{T\to\infty} \int_{-T}^{T} \frac{1}{2T}\mathrm{X}(t)\mathrm{d}t\right\} \\ &= \lim_{T\to\infty} \int_{-T}^{T} \frac{1}{2T} \mathrm{E}[\mathrm{X}(t)]\,\mathrm{d}t \\ &= \lim_{T\to\infty} \int_{-T}^{T} \frac{1}{2T}\overline{\mathrm{X}}\mathrm{d}t = \overline{X} \end{aligned} \tag{25}$$

We now use Chebychev's Inequality for a random variable X:

$$\mathrm{P}\left[\left|X - \overline{X}\right| < \varepsilon\right] \geq 1 - \frac{{\sigma_X}^2}{\varepsilon^2}$$

for positive ε, ${\sigma_X}^2 = \mathrm{Var}(X)$, and $\overline{X} = \mathrm{E}[X]$. In our case, $X = A_x$:

$$\mathrm{P}\left\{\left|A_X - \overline{A}_X\right| < \varepsilon\right\} \geq 1 - \frac{\mathrm{Var}(A_X)}{\varepsilon^2} \text{ for } \varepsilon > 0 \tag{26}$$

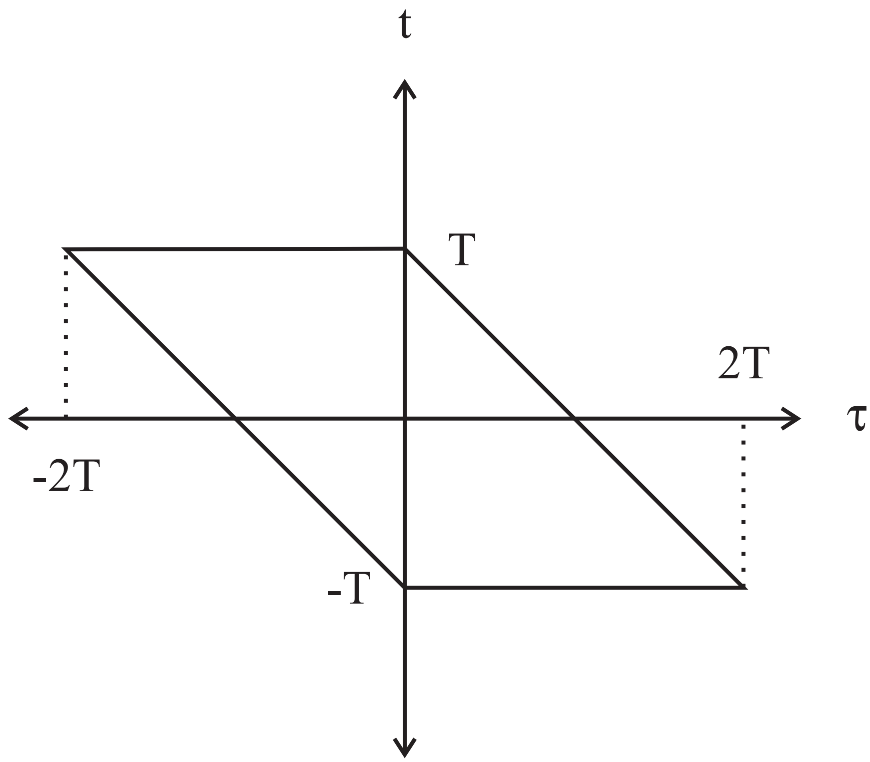
For our random process to be mean ergodic, we need the probability in (26) to be 1 no matter the value of ε is. For this to be true, we need the $\mathrm{Var}(A_X) = 0$. To see what the requirements for this condition are, we calculate this variance as

$$\begin{aligned} \mathrm{Var}(A_x) &= \mathrm{E}\left[\left(A_x - \overline{A}_x\right)^2\right] \\ &= \mathrm{E}\left\{\left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{X}(t) - \overline{\mathrm{X}}\mathrm{d}t\right]^2\right\} \\ &= \mathrm{E}\left\{\lim_{T\to\infty} \frac{1}{(2T)^2} \int_{-T}^{T}\int_{-T}^{T} \left[\mathrm{X}(t) - \overline{X}(t)\right]\left[\mathrm{X}(t_1) - \overline{X}(t_1)\right]\mathrm{d}t\mathrm{d}t_1\right\} \\ & \quad \text{defining } \mathrm{C}_{XX}(t_1,t_2) = \mathrm{E}\{[\mathrm{X}(t_1) - \mathrm{E}[\mathrm{X}(t_1)]][\mathrm{X}(t_2) - \mathrm{E}[\mathrm{X}(t_2)]]\} \\ &= \lim_{T\to\infty} \frac{1}{(2T)^2} \int_{-T}^{T}\int_{-T}^{T} \mathrm{C}_{XX}(t,t_1)\,\mathrm{d}t\mathrm{d}t_1 \end{aligned} \tag{27}$$

For the random process X(t) to be mean ergodic, we need the integral in (27) to be 0. If X(t) is a Wide Sense Stationary process, then $\mathrm{C}_{XX}(t, t_1) = \mathrm{R}_{XX}(t, t_1) - \overline{X}^2 = \mathrm{R}_{XX}(t_1 - t) - \overline{X}^2$. We can then rewrite the integral from (27) as

$$\begin{aligned} \mathrm{Var}(A_x) &= \lim_{T\to\infty} \frac{1}{(2T)^2} \int_{-T}^{T} \int_{-T}^{T} \mathrm{C}_{XX}(t_1 - t)\, \mathrm{d}t\mathrm{d}t_1 \\ &= \lim_{T\to\infty} \left(\frac{1}{2T}\right)^2 \int_{-T}^{T} \int_{-T-t}^{T-t} \mathrm{C}_{XX}(\tau)\, \mathrm{d}\tau\mathrm{d}t \text{ using } \tau = t_1 - t \end{aligned} \tag{28}$$

The integration region of (28) is shown below.



By swapping the order of the integrals, the variance is rewritten as

$$\mathrm{Var}(A_X) = \lim_{T\to\infty} \frac{1}{2T} \int_{-2T}^{2T} \left(1 - \frac{|\tau|}{2T}\right) \mathrm{C}_{XX}(\tau)\, \mathrm{d}\tau \tag{29}$$

Noting that $\mathrm{C}_{XX}(-\tau) = \mathrm{R}_{XX}(-\tau) - \overline{X}^2 = \mathrm{R}_{XX}(\tau) - \overline{X}^2 = \mathrm{C}_{XX}(\tau)$, the variance is bounded by

$$\mathrm{Var}(A_X) < \lim_{T\to\infty} \frac{1}{2T} \int_{-2T}^{2T} |\mathrm{C}_{XX}(\tau)|\, \mathrm{d}\tau \tag{30}$$

This bound goes to zero if

1. $\mathrm{C}_{XX}(0) < \infty$
2. $\lim_{\tau\to\infty} \mathrm{C}_{XX}(\tau) = 0$.
3. $\int_{-\infty}^{\infty} |\mathrm{C}_{XX}(\tau)|\, \mathrm{d}\tau < \infty$.

These conditions cause $\mathrm{Var}(A_X) = 0$ and thus are sufficient conditions for $\mathrm{X}(t)$ to be mean ergodic.
Example: A wide sense stationary random process, $\mathrm{X}(t)$, has the autocorrelation function of $\mathrm{R}_{XX}(\tau) = \mathrm{e}^{-\alpha\tau^2}$ and a mean of $\mathrm{E}[\mathrm{X}(t)] = 0$ for all t. This process is also mean ergodic since $\mathrm{C}_{XX}(0) = \mathrm{R}_{XX}(0) = 1 < \infty$, $\lim_{\tau\to\infty} \mathrm{C}_{XX}(\tau) = 0$, and $\int_{-\infty}^{\infty} |\mathrm{C}_{XX}(\tau)|\,\mathrm{d}\tau = \int_{-\infty}^{\infty} \mathrm{e}^{-\alpha\tau^2}\,\mathrm{d}\tau = \frac{\sqrt{\pi}}{\sqrt{\alpha}} < \infty$ using Equation (C-51) of the textbook.

3.1.1 Mean Ergodic Sequences

The procedure for defining sufficient conditions for mean ergodic processes can also be extended to random sequences. Define $\mathrm{R}_{XX}[m] = \mathrm{E}\{\mathrm{X}[n]\,\mathrm{X}[n+m]\}$ and $\mathrm{C}_{XX}[m] = \mathrm{R}_{XX}[m] - \overline{X}^2$, with $\mathrm{X}[n]$ being a wide sense stationary random sequence. A random sequence is called mean ergodic if

$$A_X = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \mathrm{X}[n] = \overline{X} \text{ with probability 1} \tag{31}$$

We proceed, as in the previous section, by calculating the variance of A_X and finding conditions for making its variance 0:

$$\begin{aligned}
\mathrm{Var}(A_X) &= \mathrm{E}\left\{\lim_{N\to\infty}\left[\frac{1}{2N+1}\sum_{m=-N}^{N}\mathrm{X}[m] - \overline{X}\right]\left[\frac{1}{2N+1}\sum_{n=-N}^{N}\mathrm{X}[n] - \overline{X}\right]\right\} \\
&= \mathrm{E}\left\{\lim_{N\to\infty}\left(\frac{1}{2N+1}\right)^2 \sum_{m=-N}^{N}\sum_{n=-N}^{N}\left[\mathrm{X}[m] - \overline{X}\right]\left[\mathrm{X}[n] - \overline{X}\right]\right\} \\
&= \lim_{N\to\infty}\left(\frac{1}{2N+1}\right)^2 \sum_{m=-N}^{N}\sum_{n=-N}^{N}\mathrm{E}\left\{\left[\mathrm{X}[m] - \overline{X}\right]\left[\mathrm{X}[n] - \overline{X}\right]\right\} \\
&= \lim_{N\to\infty}\left(\frac{1}{2N+1}\right)^2 \sum_{m=-N}^{N}\sum_{n=-N}^{N}\mathrm{C}_{XX}[n-m] \\
&= \lim_{N\to\infty}\left(\frac{1}{2N+1}\right)^2 \sum_{m=-N}^{N}\sum_{k=-N-m}^{N-m}\mathrm{C}_{XX}[k] \\
&= \lim_{N\to\infty}\frac{1}{2N+1} \sum_{k=-2N}^{2N}\mathrm{C}_{XX}[k]\left(1 - \frac{|k|}{2N+1}\right)
\end{aligned} \tag{32}$$

If in the limit as $N \to \infty$, this variance sum goes to zero, then the random sequence is mean ergodic.