

Lecture Notes 7

Stationary Random Processes

- Strict-Sense and Wide-Sense Stationarity
- Autocorrelation Function of a Stationary Process
- Power Spectral Density
- Continuity and Integration of Random Processes
- Stationary Ergodic Random Processes

Stationary Random Processes

- Stationarity refers to *time invariance* of some, or all, of the statistics of a random process, such as mean, autocorrelation, n -th-order distribution
- We define two types of stationarity: *strict sense* (SSS) and *wide sense* (WSS)
- A random process $X(t)$ (or X_n) is said to be SSS if *all* its finite order distributions are time invariant, i.e., the joint cdfs (pdfs, pmfs) of

$$X(t_1), X(t_2), \dots, X(t_k) \quad \text{and} \quad X(t_1 + \tau), X(t_2 + \tau), \dots, X(t_k + \tau)$$

are the same for all k , all t_1, t_2, \dots, t_k , and all time shifts τ

- So for a SSS process, the first-order distribution is independent of t , and the second-order distribution — the distribution of any two samples $X(t_1)$ and $X(t_2)$ — depends only on $\tau = t_2 - t_1$

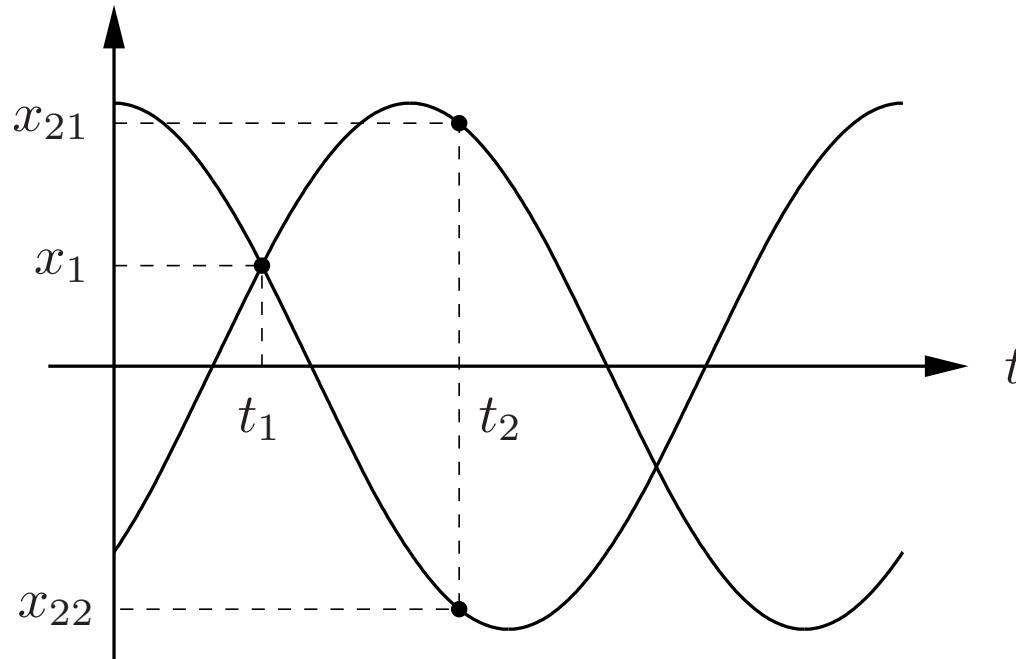
To see this, note that from the definition of stationarity, for any t , the joint distribution of $X(t_1)$ and $X(t_2)$ is the same as the joint distribution of $X(t_1 + (t - t_1)) = X(t)$ and $X(t_2 + (t - t_1)) = X(t + (t_2 - t_1))$

- Example: The random phase signal $X(t) = \alpha \cos(\omega t + \Theta)$ where $\Theta \in U[0, 2\pi]$ is SSS
 - We already know that the first order pdf is

$$f_{X(t)}(x) = \frac{1}{\pi\alpha\sqrt{1-(x/\alpha)^2}}, \quad -\alpha < x < +\alpha$$

which is independent of t , and is therefore stationary

- To find the second order pdf, note that if we are given the value of $X(t)$ at one point, say t_1 , there are (at most) two possible sample functions:



The second order pdf can thus be written as

$$\begin{aligned} f_{X(t_1),X(t_2)}(x_1, x_2) &= f_{X(t_1)}(x_1) f_{X(t_2)|X(t_1)}(x_2|x_1) \\ &= f_{X(t_1)}(x_1) \left(\frac{1}{2}\delta(x_2 - x_{21}) + \frac{1}{2}\delta(x_2 - x_{22}) \right), \end{aligned}$$

which depends only on $t_2 - t_1$, and thus the second order pdf is stationary

- Now if we know that $X(t_1) = x_1$ and $X(t_2) = x_2$, the sample path is totally determined (except when $x_1 = x_2 = 0$, where two paths may be possible), and thus all n -th order pdfs are stationary
- IID processes are SSS
- Random walk and Poisson processes are not SSS
- The Gauss-Markov process (as we defined it) is not SSS. However, if we set X_1 to the steady state distribution of X_n , it becomes SSS (see homework exercise)

Wide-Sense Stationary Random Processes

- A random process $X(t)$ is said to be *wide-sense stationary* (WSS) if its mean and autocorrelation functions are time invariant, i.e.,
 - $E(X(t)) = \mu$, independent of t
 - $R_X(t_1, t_2)$ is a function only of the time difference $t_2 - t_1$
 - $E[X(t)^2] < \infty$ (technical condition)
- Since $R_X(t_1, t_2) = R_X(t_2, t_1)$, for any wide sense stationary process $X(t)$, $R_X(t_1, t_2)$ is a function only of $|t_2 - t_1|$
- Clearly SSS \Rightarrow WSS. The converse is not necessarily true

- Example: Let

$$X(t) = \begin{cases} +\sin t & \text{with probability } \frac{1}{4} \\ -\sin t & \text{with probability } \frac{1}{4} \\ +\cos t & \text{with probability } \frac{1}{4} \\ -\cos t & \text{with probability } \frac{1}{4} \end{cases}$$

- $E(X(t)) = 0$ and $R_X(t_1, t_2) = \frac{1}{2} \cos(t_2 - t_1)$, thus $X(t)$ is WSS
- But $X(0)$ and $X(\frac{\pi}{4})$ do not have the same pmf (different ranges), so the first order pmf is not stationary, and the process is not SSS
- For Gaussian random processes, WSS \Rightarrow SSS, since the process is completely specified by its mean and autocorrelation functions
- Random walk is not WSS, since $R_X(n_1, n_2) = \min\{n_1, n_2\}$ is not time invariant; similarly Poisson process is not WSS

Autocorrelation Function of WSS Processes

- Let $X(t)$ be a WSS process. Relabel $R_X(t_1, t_2)$ as $R_X(\tau)$ where $\tau = t_1 - t_2$
 1. $R_X(\tau)$ is real and even, i.e., $R_X(\tau) = R_X(-\tau)$ for every τ
 2. $|R_X(\tau)| \leq R_X(0) = \mathbb{E}[X^2(t)]$, the “average power” of $X(t)$

This can be shown as follows. For every t ,

$$\begin{aligned}(R_X(\tau))^2 &= [\mathbb{E}(X(t)X(t + \tau))]^2 \\ &\leq \mathbb{E}[X^2(t)] \mathbb{E}[X^2(t + \tau)] \quad \text{by Schwarz inequality} \\ &= (R_X(0))^2 \quad \text{by stationarity}\end{aligned}$$

- 3. If $R_X(T) = R_X(0)$ for some $T \neq 0$, then $R_X(\tau)$ is periodic with period T and so is $X(t)$ (with probability 1) !! That is,

$$R_X(\tau) = R_X(\tau + T), \quad X(t) = X(t + T) \text{ w.p.1 for every } \tau$$

- Example: The autocorrelation function for the periodic signal with random phase $X(t) = \alpha \cos(\omega t + \Theta)$ is $R_X(\tau) = \frac{\alpha^2}{2} \cos \omega \tau$ (also periodic)
- To prove property 3, we again use the Schwarz inequality: For every τ ,

$$\begin{aligned}
 [R_X(\tau) - R_X(\tau + T)]^2 &= [\mathbb{E}(X(t)(X(t + \tau) - X(t + \tau + T)))]^2 \\
 &\leq \mathbb{E}[X^2(t)] \mathbb{E}[(X(t + \tau) - X(t + \tau + T))^2] \\
 &= R_X(0)(2R_X(0) - 2R_X(T)) \\
 &= R_X(0)(2R_X(0) - 2R_X(0)) = 0
 \end{aligned}$$

Thus $R_X(\tau) = R_X(\tau + T)$ for all τ , i.e., $R_X(\tau)$ is periodic with period T

- The above properties of $R_X(\tau)$ are necessary but not sufficient for a function to qualify as an autocorrelation function for a WSS process

- The necessary and sufficient conditions for a function to be an autocorrelation function for a WSS process is that it be *real*, *even*, and *nonnegative definite*. By nonnegative definite we mean that for any n , any t_1, t_2, \dots, t_n and any real vector $\mathbf{a} = (a_1, \dots, a_n)$,

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j R(t_i - t_j) \geq 0$$

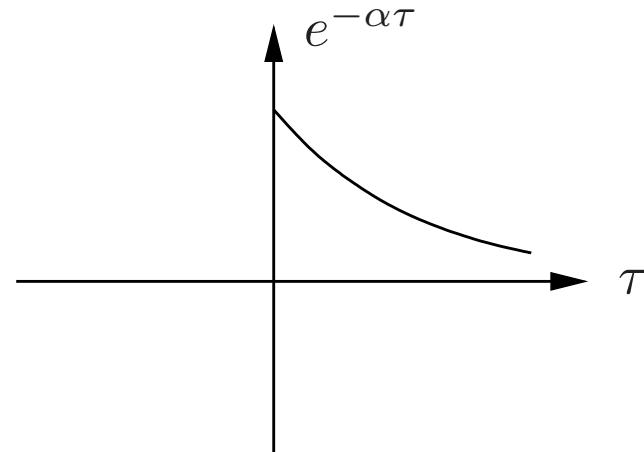
To see why this is necessary, recall that the correlation matrix for a random vector must be nonnegative definite, so if we take a set of n samples from the WSS random process, their correlation matrix must be nonnegative definite.

The condition is sufficient since such an $R(\tau)$ can specify a zero mean stationary Gaussian random process.

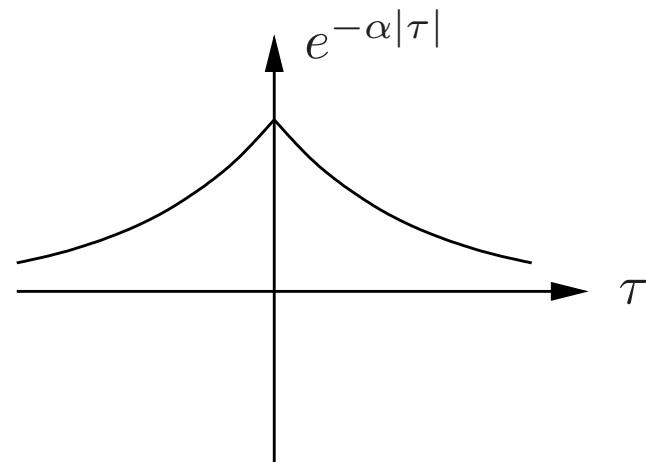
- The nonnegative definite condition may be difficult to verify directly. It turns out, however, to be equivalent to the condition that the Fourier transform of $R_X(\tau)$, which is called the *power spectral density* $S_X(f)$, is nonnegative for all frequencies f .

Which Functions Can Be an $R_X(\tau)$?

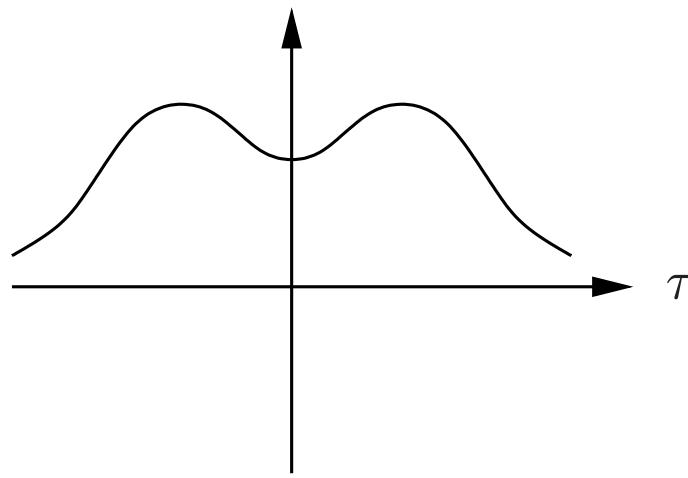
1.



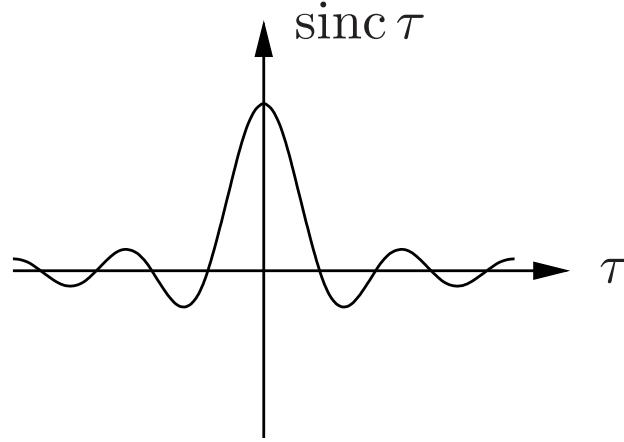
2.



3.

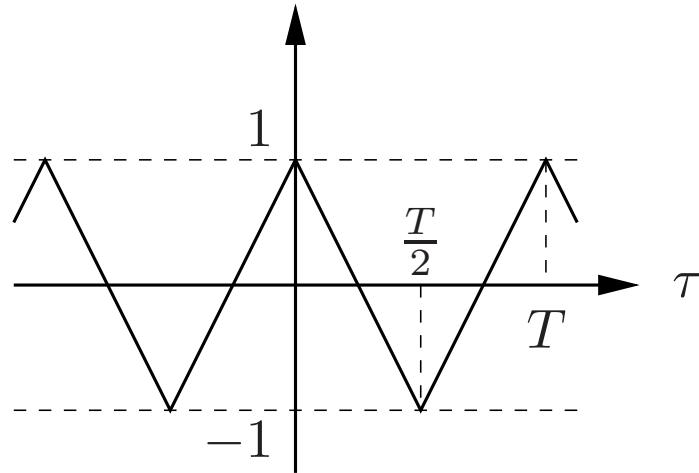


4.

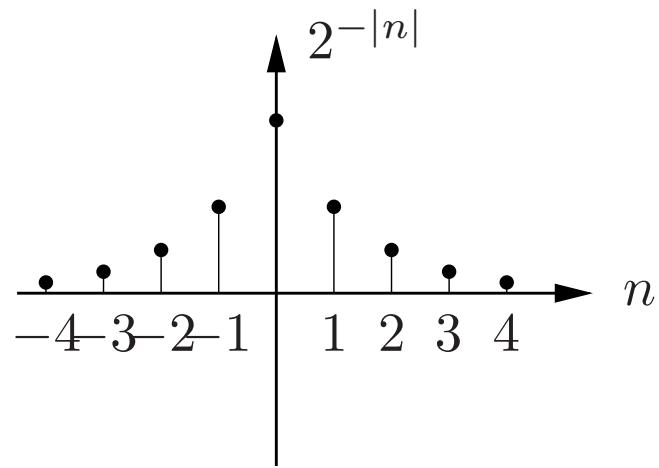


Which Functions can be an $R_X(\tau)$?

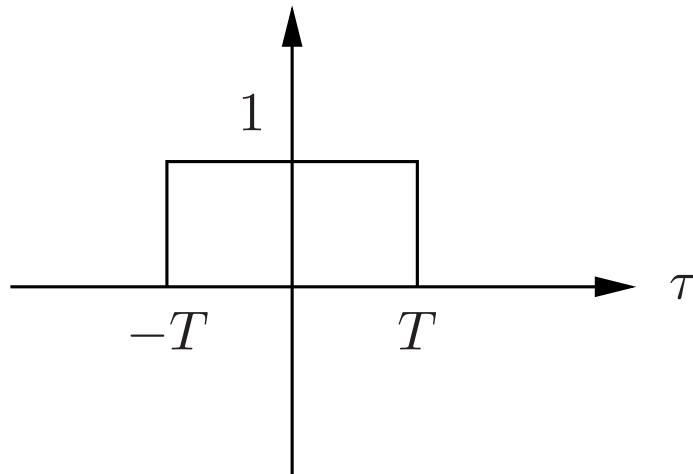
5.



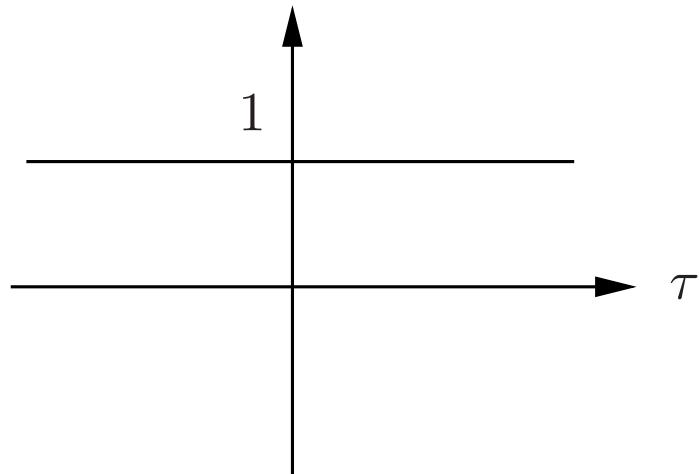
6.



7.

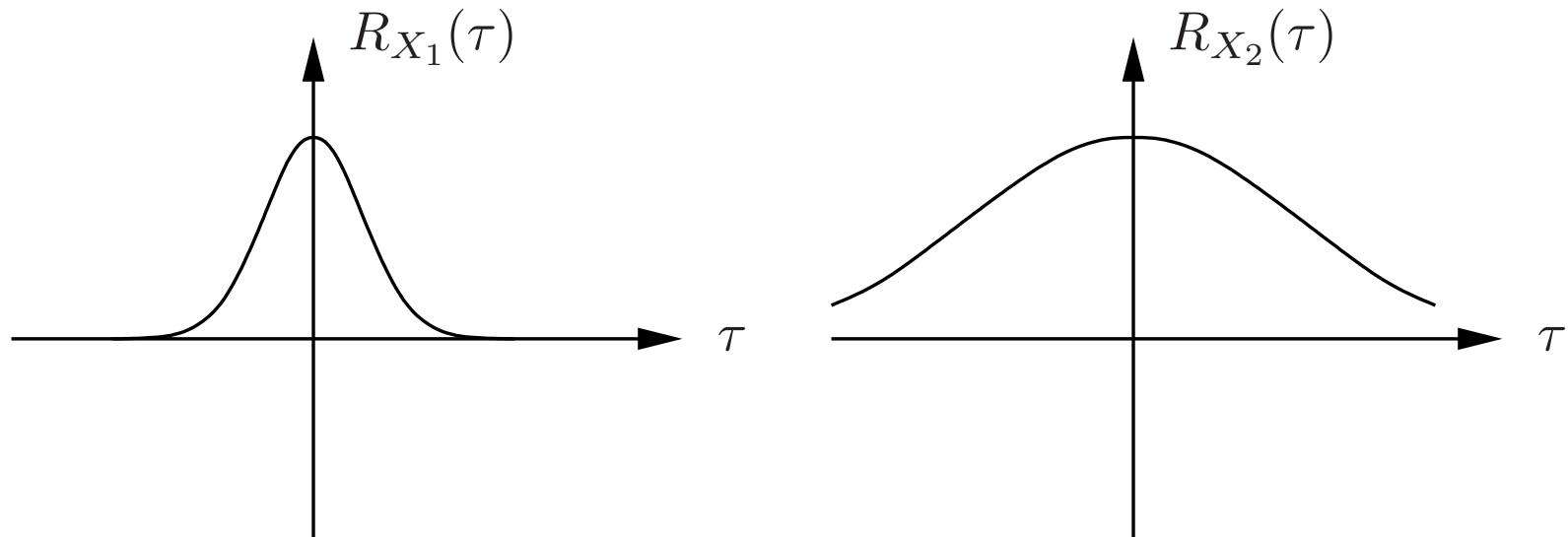


8.



Interpretation of Autocorrelation Function

- Let $X(t)$ be WSS with zero mean. If $R_X(\tau)$ drops quickly with τ , this means that samples become uncorrelated quickly as we increase τ . Conversely, if $R_X(\tau)$ drops slowly with τ , samples are highly correlated



- So $R_X(\tau)$ is a measure of the rate of change of $X(t)$ with time t , i.e., *the frequency response* of $X(t)$
- It turns out that this is not just an intuitive interpretation—the Fourier transform of $R_X(\tau)$ (the power spectral density) is in fact the average power density of $X(t)$ over frequency

Power Spectral Density

- The *power spectral density* (psd) of a WSS random process $X(t)$ is the Fourier transform of $R_X(\tau)$:

$$S_X(f) = \mathcal{F}(R_X(\tau)) = \int_{-\infty}^{\infty} R_X(\tau) e^{-i2\pi\tau f} d\tau$$

- For a discrete time process X_n , the power spectral density is the discrete-time Fourier transform (DTFT) of the sequence $R_X(n)$:

$$S_X(f) = \sum_{n=-\infty}^{\infty} R_X(n) e^{-i2\pi n f}, \quad |f| < \frac{1}{2}$$

- $R_X(\tau)$ (or $R_X(n)$) can be recovered from $S_X(f)$ by taking the inverse Fourier transform or inverse DTFT:

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) e^{i2\pi\tau f} df$$

$$R_X(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_X(f) e^{i2\pi n f} df$$

Properties of the Power Spectral Density

1. $S_X(f)$ is real and even, since the Fourier transform of the real and even function $R_X(\tau)$ is real and even
2. $\int_{-\infty}^{\infty} S_X(f) df = R_X(0) = \text{E}(X^2(t))$, the average power of $X(t)$, i.e., the area under S_X is the average power
3. $S_X(f)$ is *the average power density*, i.e., the average power of $X(t)$ in the frequency band $[f_1, f_2]$ is

$$\int_{-f_2}^{-f_1} S_X(f) df + \int_{f_1}^{f_2} S_X(f) df = 2 \int_{f_1}^{f_2} S_X(f) df$$

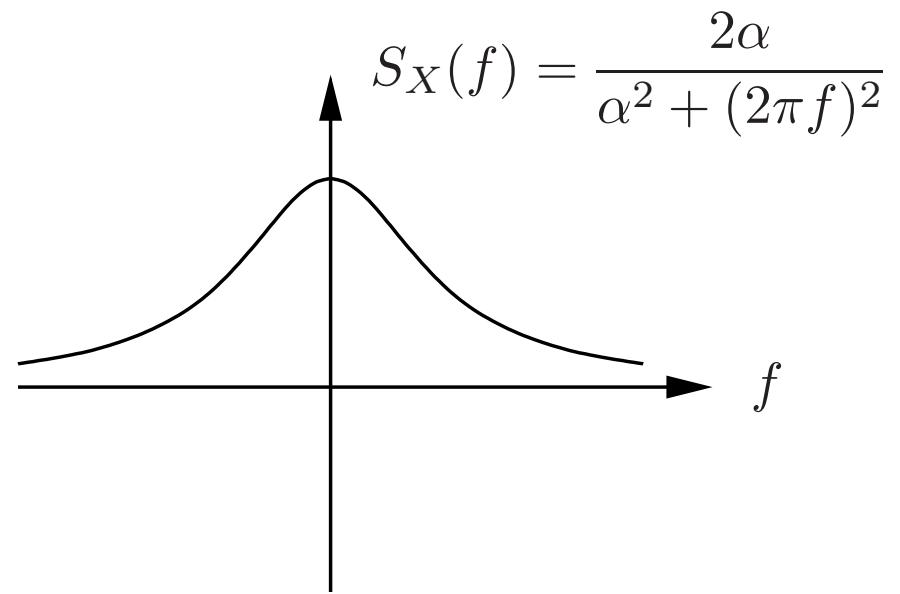
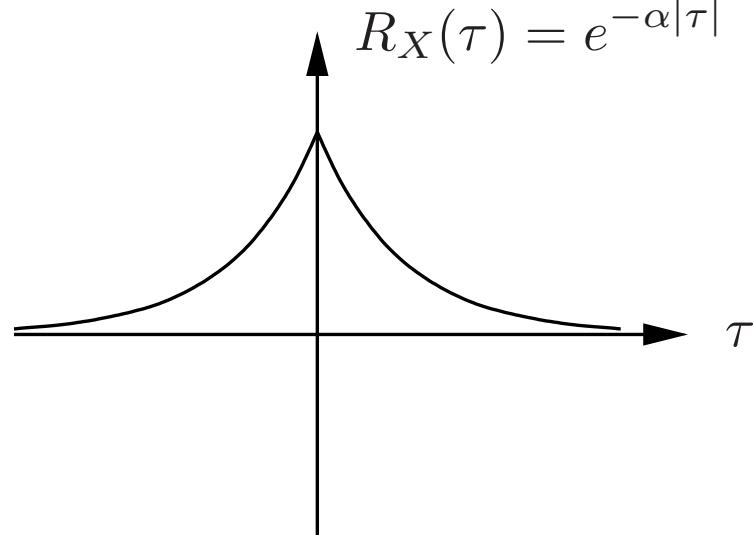
(we will show this soon)

- From property 3, it follows that $S_X(f) \geq 0$. Why?
- In general, a function $S(f)$ is a psd if and only if it is real, even, nonnegative, and

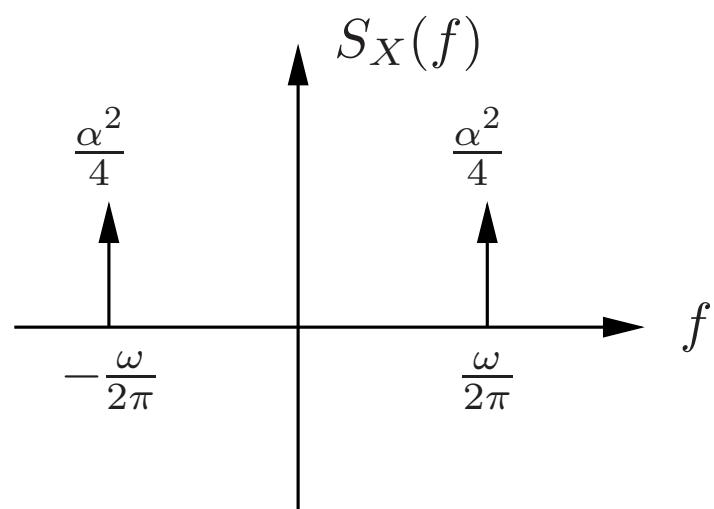
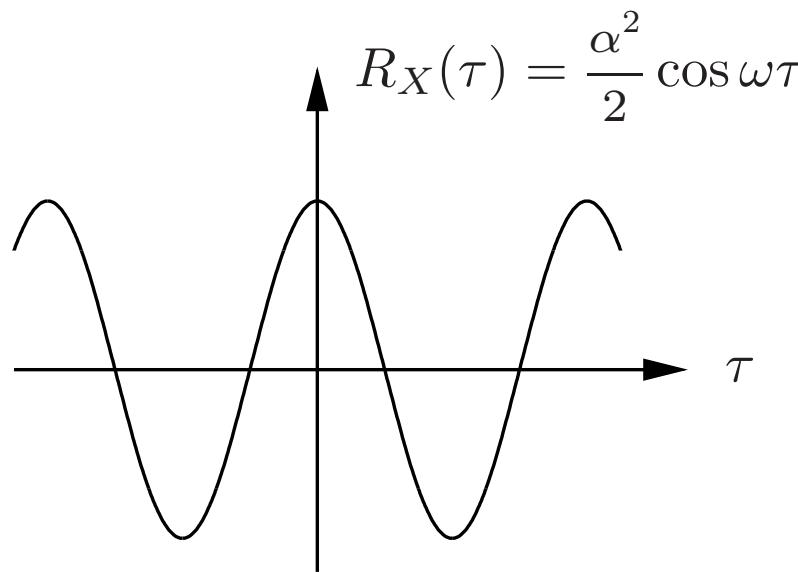
$$\int_{-\infty}^{\infty} S(f) df < \infty$$

Examples

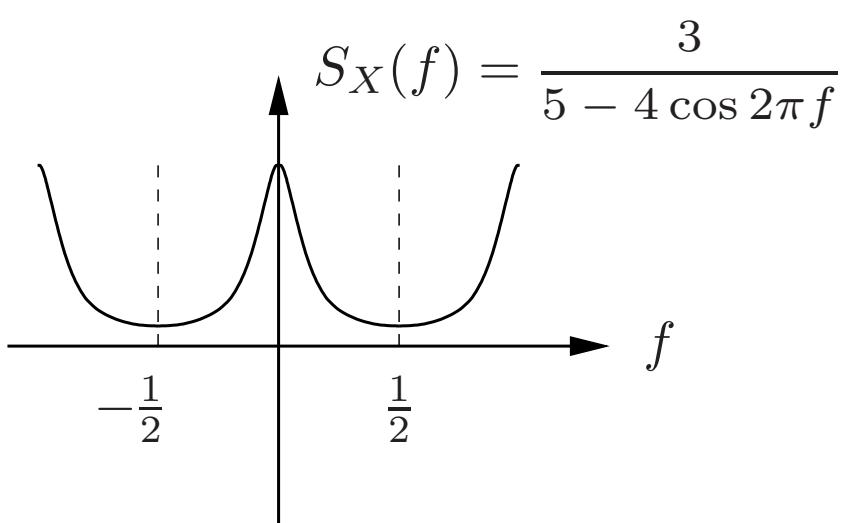
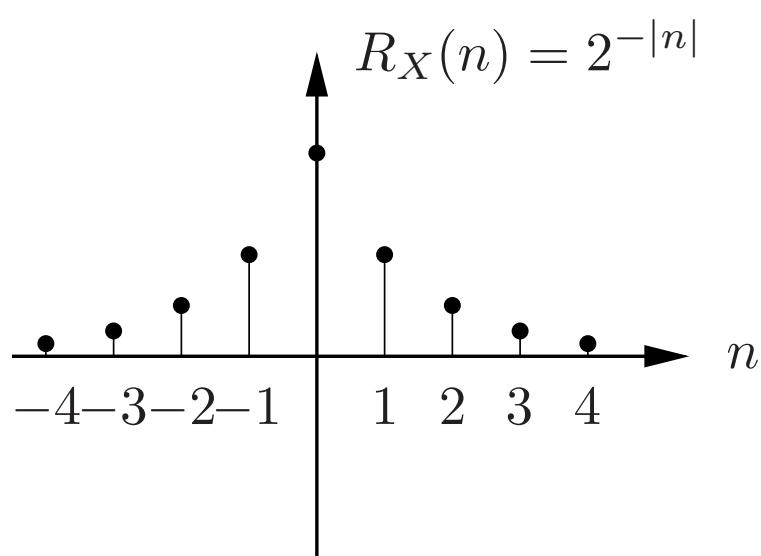
1.



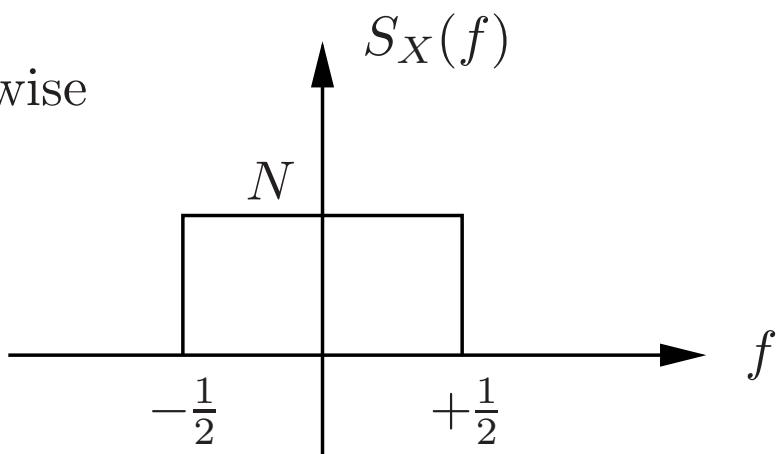
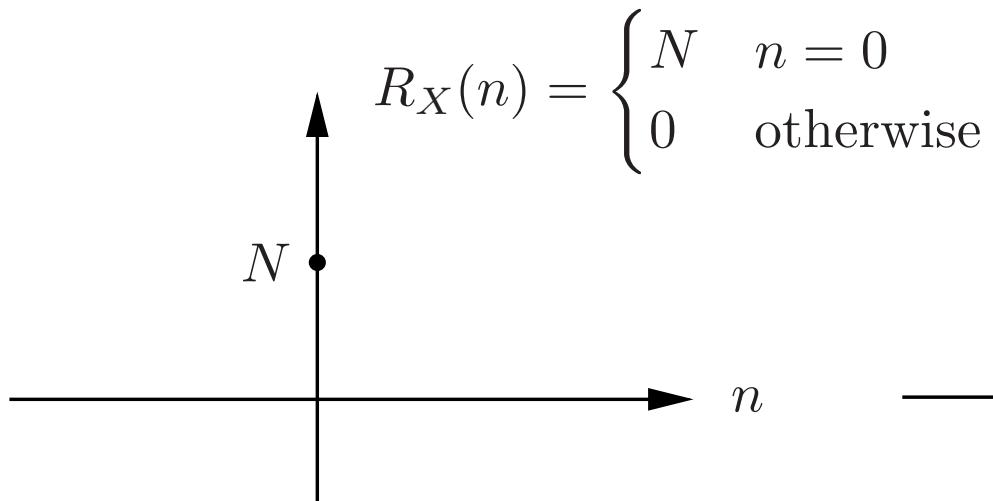
2.



3.

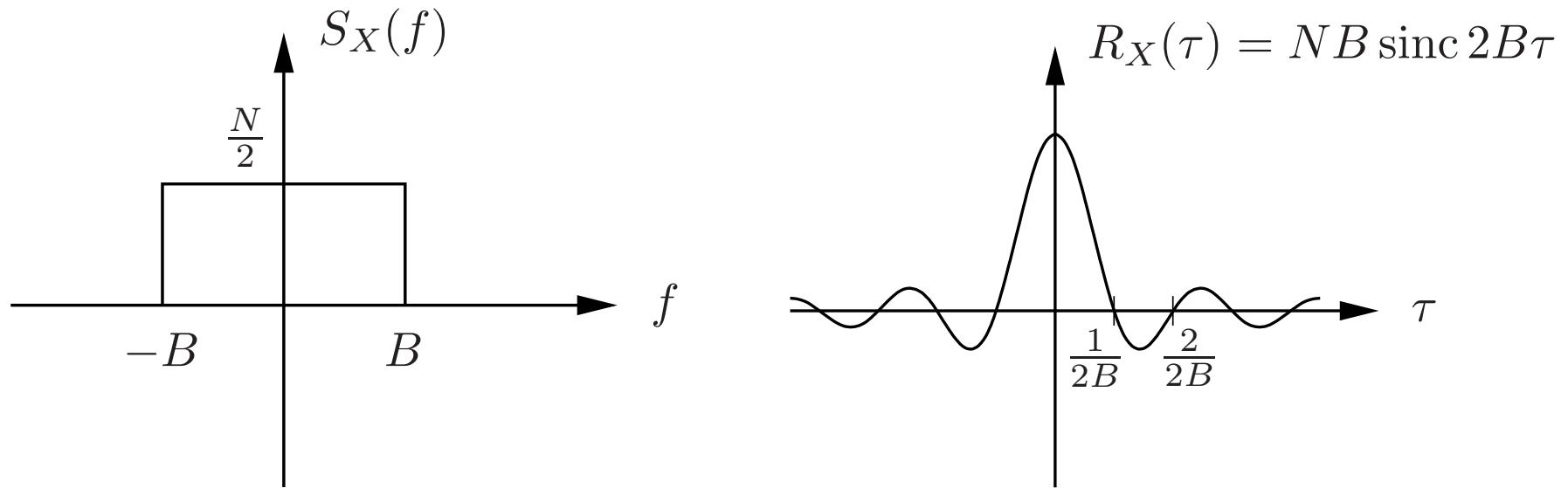


4. *Discrete time white noise process:* $X_1, X_2, \dots, X_n, \dots$ zero mean, uncorrelated, with average power N



If X_n is also a GRP, then we obtain a *discrete time WGN process*

5. *Bandlimited white noise process*: WSS zero mean process $X(t)$ with



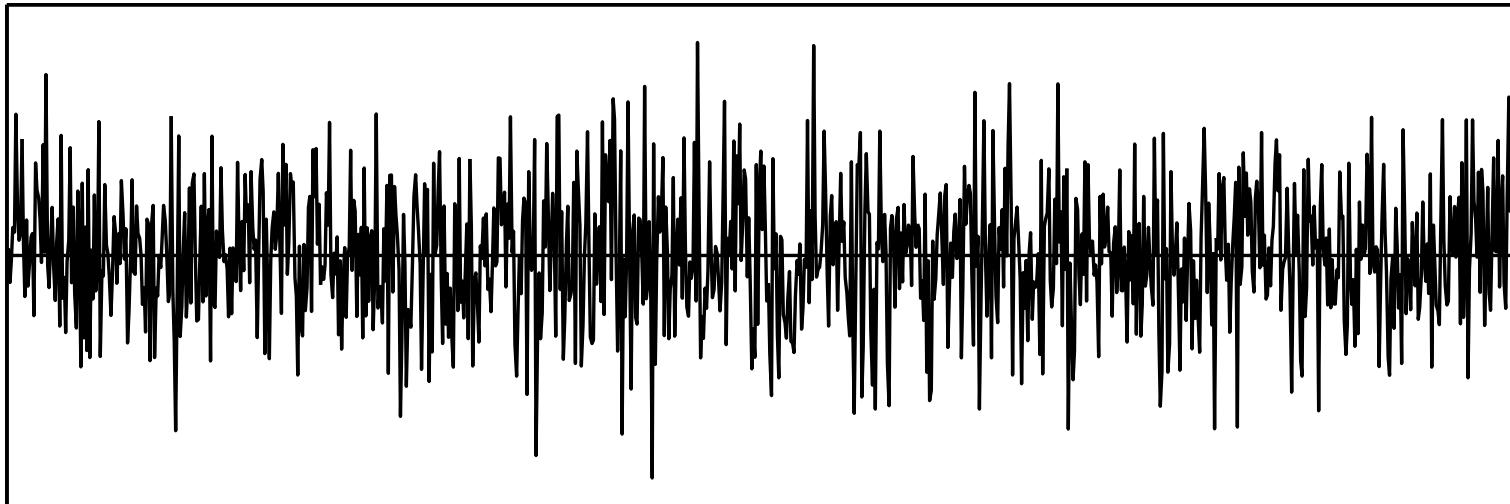
For any t , the samples $X\left(t \pm \frac{n}{2B}\right)$ for $n = 0, 1, 2, \dots$ are uncorrelated

6. *White noise process*: If we let $B \rightarrow \infty$ in the previous example, we obtain a *white noise process*, which has

$$S_X(f) = \frac{N}{2} \quad \text{for all } f$$

$$R_X(\tau) = \frac{N}{2} \delta(\tau)$$

If, in addition, $X(t)$ is a GRP, then we obtain the famous *white Gaussian noise (WGN) process*



- Remarks on white noise:
 - For a white noise process, all samples are uncorrelated
 - The process is not physically realizable, since it has infinite power
 - However, it plays a similar role in random processes to point mass in physics and delta function in linear systems
 - Thermal noise and shot noise are well modeled as white Gaussian noise, since they have very flat psd over very wide band (GHz)

Continuity and Integration of Random Processes

- We are all familiar with the definitions of continuity and integration for deterministic functions as limits
- Using the notions of convergence discussed in Lecture Notes 5, we can define these notions for random processes. We focus only on m.s. convergence
- *Continuity*: A process $X(t)$ is said to be *mean square continuous* if for every t

$$\lim_{s \rightarrow t} \mathbb{E}[(X(s) - X(t))^2] = 0$$

- The continuity of $X(t)$ depends only on its autocorrelation function $R_X(t_1, t_2)$
In fact, the following statements are all equivalent:
 1. $R_X(t_1, t_2)$ is continuous at all points of the form (t, t)
 2. $X(t)$ is m.s. continuous
 3. $R_X(t_1, t_2)$ is continuous in t_1, t_2

Proof:

1. 1 implies 2: Since if $R_X(t_1, t_2)$ is continuous at all points (t, t) ,

$$\mathbb{E}[(X(t) - X(s))^2] = R_X(t, t) + R_X(s, s) - 2R_X(s, t) \rightarrow 0 \text{ as } s \rightarrow t$$

2. 2 implies 3: Consider

$$\begin{aligned} R_X(s_1, s_2) &= \mathbb{E}[X(s_1)X(s_2)] \\ &= \mathbb{E}[(X(t_1) + (X(s_1) - X(t_1)))(X(t_2) + (X(s_2) - X(t_2)))] \\ &= R_X(t_1, t_2) + \mathbb{E}[X(t_1)(X(s_2) - X(t_2))] + \mathbb{E}[X(t_2)(X(s_1) - X(t_1))] \\ &\quad + \mathbb{E}[(X(s_1) - X(t_1))(X(s_2) - X(t_2))] \\ &\leq R_X(t_1, t_2) + \sqrt{\mathbb{E}[X^2(t_1)] \mathbb{E}[(X(s_2) - X(t_2))^2]} \\ &\quad + \sqrt{\mathbb{E}[X^2(t_2)] \mathbb{E}[(X(s_1) - X(t_1))^2]} \\ &\quad + \sqrt{\mathbb{E}[(X(s_1) - X(t_1))^2] \mathbb{E}[(X(s_2) - X(t_2))^2]} \text{ Schwartz inequality} \\ &\rightarrow R_X(t_1, t_2) \text{ as } s_1 \rightarrow t_1 \text{ and } s_2 \rightarrow t_2 \end{aligned}$$

since $X(t)$ is m.s. continuous

3. Since 3 implies 1, we are done

- Example: The Poisson process $N(t)$ with rate $\lambda > 0$ is m.s. continuous, since its autocorrelation function,

$$R_N(t_1, t_2) = \lambda \min\{t_1, t_2\} + \lambda^2 t_1 t_2$$

is a continuous function

- *Integration:* Let $X(t)$ be a RP and $h(t)$ be a function. We can define the integral

$$\int_a^b h(t)X(t)dt$$

as the limit of a sum (as in Riemann integral of a deterministic function) in m.s.

Let $\Delta > 0$ such that $b - a = n\Delta$ and

$a \leq \tau_1 \leq a + \Delta \leq \tau_2 \leq a + 2\Delta \leq \dots \leq \tau_{n-1} \leq a + (n-1)\Delta \leq \tau_n \leq a + n\Delta = b$,
then the corresponding Riemann sum is

$$\sum_{i=1}^{n-1} h(\tau_i)X(\tau_i)\Delta$$

The above integral then exists if this sum has a limit in m.s. as $\Delta \rightarrow 0$

Moreover, if the random integral exists for all a, b , then we can define

$$\int_{-\infty}^{\infty} h(t)X(t)dt = \lim_{a,b \rightarrow \infty} \int_a^b h(t)X(t)dt \quad \text{in m.s.}$$

- Fact: The existence of the m.s. integral depends only on R_X and h
More specifically, the above integral exists iff

$$\int_a^b \int_a^b R_X(t_1, t_2)h(t_1)h(t_2)dt_1 dt_2$$

exists (in the normal sense)

- Remark: We are skipping several mathematical details here. In what follows, we use the above fact to justify the existence of integrals involving random processes and in interchanging expectation and integration

Stationary Ergodic Random processes

- Let $X(t)$ be SSS or only WSS
- Ergodicity of $X(t)$ means that certain time averages converge to their respective statistical averages
- *Mean ergodic process*: Let $X(t)$ be a WSS and m.s. continuous RP with mean μ_X

To estimate the mean of $X(t)$, we form the *time average*

$$\bar{X}(t) = \frac{1}{t} \int_0^t X(\tau) d\tau$$

The RP $X(t)$ is said to be *mean ergodic* if $\bar{X}(t) \rightarrow \mu_X$ as $t \rightarrow \infty$ in m.s.

Similarly for a discrete RP, the time average (same as sample average) is

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_n$$

and the RP is mean ergodic if $\bar{X}_n \rightarrow \mu_X$ in m.s.

- Example: Let X_n be a WSS process with $C_X(n) = 0$ for $n \neq 0$, i.e., the X_i s are uncorrelated, then X_n is mean ergodic
- The process does not need to have uncorrelated samples for it to be mean ergodic, however
- Whether a WSS process is mean ergodic again depends only on its autocorrelation function

By definition, mean ergodicity means that

$$\lim_{t \rightarrow \infty} E[(\bar{X}(t) - \mu_X)^2] \rightarrow 0$$

Since $E(\bar{X}(t)) = \mu_X$, the condition for mean ergodicity is the same as

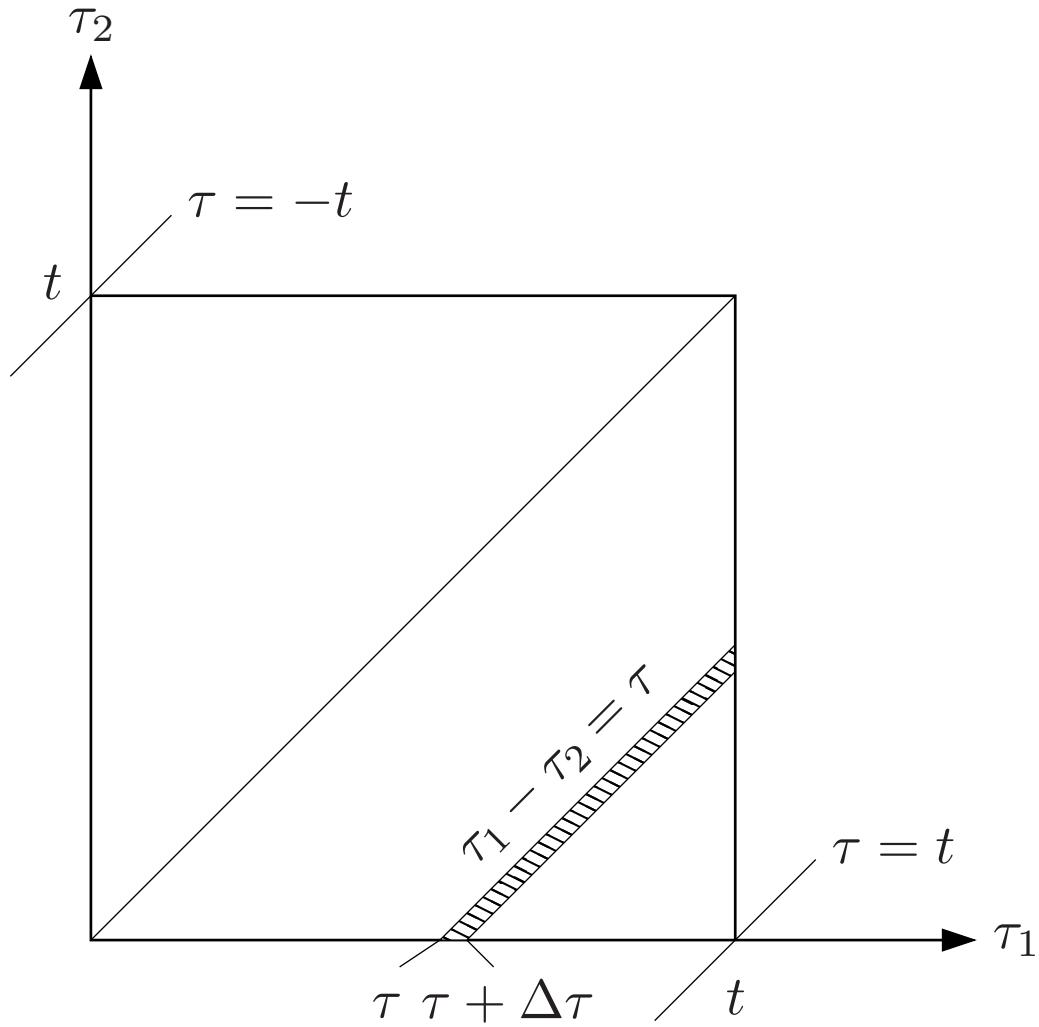
$$\lim_{t \rightarrow \infty} \text{Var}(\bar{X}(t)) = 0$$

Now, consider

$$\begin{aligned}\mathbb{E}(\bar{X}^2(t)) &= \mathbb{E}\left[\left(\frac{1}{t}\int_0^t X(\tau)d\tau\right)^2\right] \\ &= \mathbb{E}\left(\frac{1}{t^2}\int_0^t \int_0^t X(\tau_1)X(\tau_2)d\tau_1d\tau_2\right) \\ &= \frac{1}{t^2}\int_0^t \int_0^t R_X(\tau_1, \tau_2)d\tau_1d\tau_2 \\ &= \frac{1}{t^2}\int_0^t \int_0^t R_X(\tau_1 - \tau_2)d\tau_1d\tau_2\end{aligned}$$

From figure below, this double integral reduces to the single integral

$$\mathbb{E}(\bar{X}^2(t)) = \frac{2}{t^2}\int_0^t (t - \tau)R_X(\tau)d\tau$$



- Hence, a WSS process $X(t)$ is mean ergodic iff

$$\lim_{t \rightarrow \infty} \frac{2}{t^2} \int_0^t (t - \tau) R_X(\tau) d\tau = \mu_X^2$$

- Example: Let $X(t)$ be a WSS with zero mean and $R_X(\tau) = e^{-|\tau|}$

Evaluating the condition on mean ergodicity, we obtain

$$\frac{2}{t^2} \int_0^t (t - \tau) R_X(\tau) d\tau = \frac{2}{t^2} (e^{-t} + t - 1),$$

which $\rightarrow 0$ as $t \rightarrow \infty$. Hence $X(t)$ is mean ergodic

- Example: Consider the coin with random bias P example in Lecture Notes 5. The random process X_1, X_2, \dots is stationary

However, it is not mean ergodic, since $\bar{X}_n \rightarrow P$ in m.s.

- Remarks:

- The process in the above example can be viewed as a *mixture* of IID Bernoulli(p) processes, each of which is stationary ergodic (it turns out that every stationary process is a mixture of stationary ergodic processes)
- Ergodicity can be defined for general (not necessarily stationary) processes (this is beyond the scope of this course, however)