Lecture Notes 7
Stationary Random Processes

- Strict-Sense and Wide-Sense Stationarity
- Autocorrelation Function of a Stationary Process
- Power Spectral Density
- 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

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

  are the same for all k, all $t_1, t_2, \ldots, t_k$, 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 $\tau = t_2 - t_1$

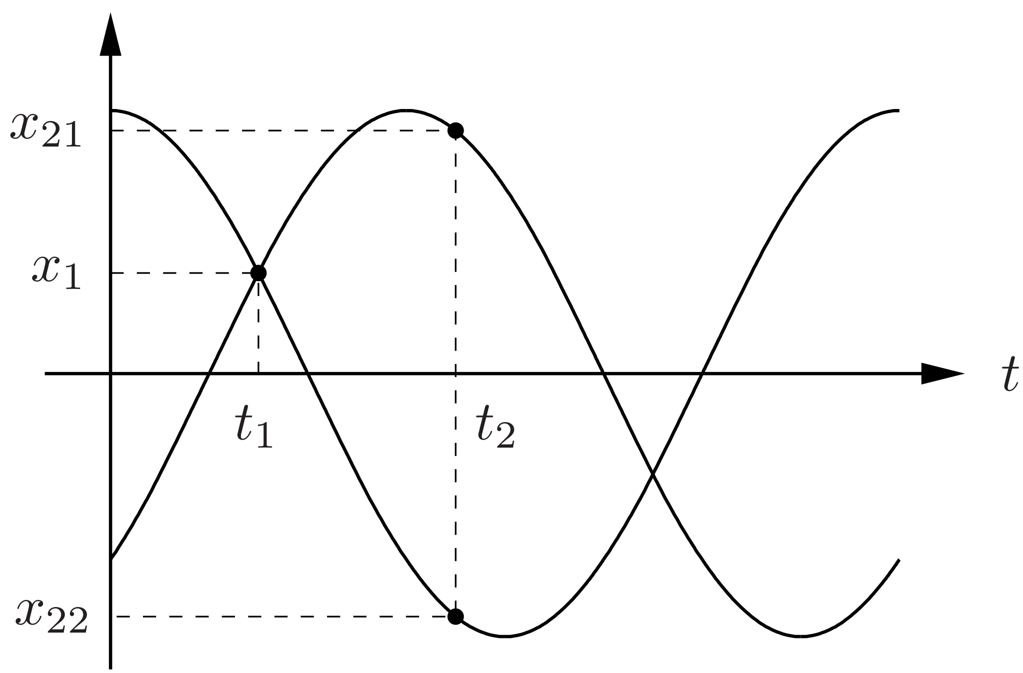
  To see this, note that from the definition of stationarity, for any t, the joint distribution of $X(t_1)$ and $X(t_2)$ is the same as the joint distribution of $X(t_1 + (t - t_1)) = X(t)$ and $X(t_2 + (t - t_1)) = X(t + (t_2 - t_1))$

- Example: The random phase signal $X(t) = \alpha\cos(\omega t + \Theta)$ where $\Theta \in \mathrm{U}[0, 2\pi]$ 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

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

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.,
  - $\mathrm{E}(X(t)) = \mu$, independent of t
  - $R_X(t_1, t_2)$ is a function only of the time difference $t_2 - t_1$
  - $\mathrm{E}[X(t)^2] < \infty$ (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

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

  - $\mathrm{E}(X(t)) = 0$ 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. $|R_X(\tau)| \le R_X(0) = \mathrm{E}[X^2(t)]$, the "average power" of $X(t)$

   This can be shown as follows. For every t,

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

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

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

- 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,

$$\begin{aligned}
\big[R_X(\tau) - R_X(\tau+T)\big]^2 &= \big[\,\mathrm{E}\,(X(t)(X(t+\tau) - X(t+\tau+T)))\,\big]^2 \\
&\le \mathrm{E}[X^2(t)]\,\mathrm{E}\left[(X(t+\tau) - X(t+\tau+T))^2\right] \\
&= R_X(0)(2R_X(0) - 2R_X(T)) \\
&= R_X(0)(2R_X(0) - 2R_X(0)) \;=\; 0
\end{aligned}$$

  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 $t_1, t_2, \ldots, t_n$ and any real vector $\mathbf{a} = (a_1, \ldots, a_n)$,

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

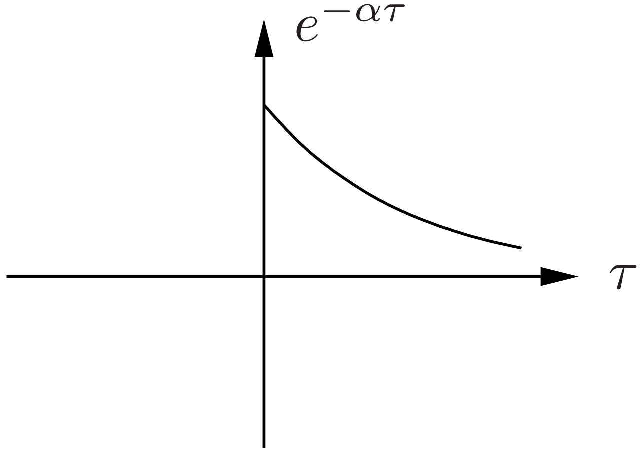
  To see why this is necessary, recall that the correlation matrix for a random vector must be nonnegative definite, so if we take a set of n samples from the WSS random process, their correlation matrix must be nonnegative definite

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

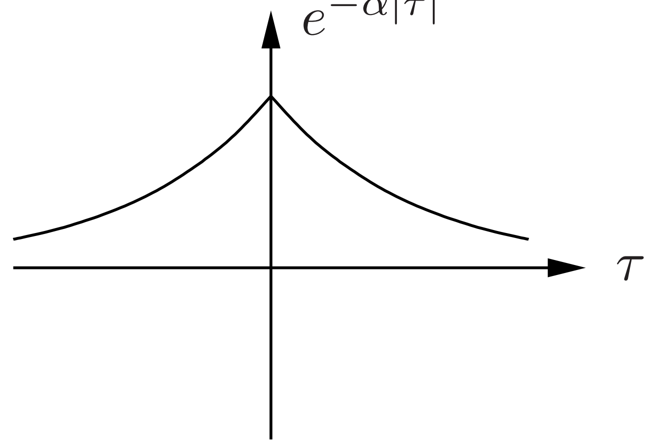
- The nonnegative definite condition may be difficult to verify directly. It turns out, however, to be equivalent to the condition that the Fourier transform of $R_X(\tau)$, which is called the power spectral density $S_X(f)$, is nonnegative for all frequencies f

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

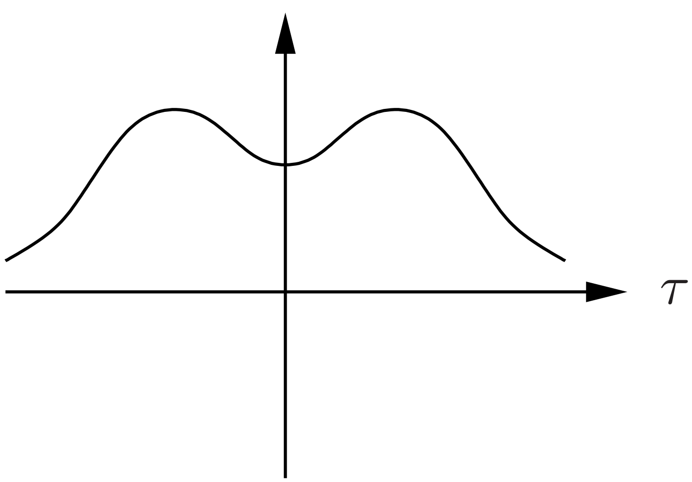
1.



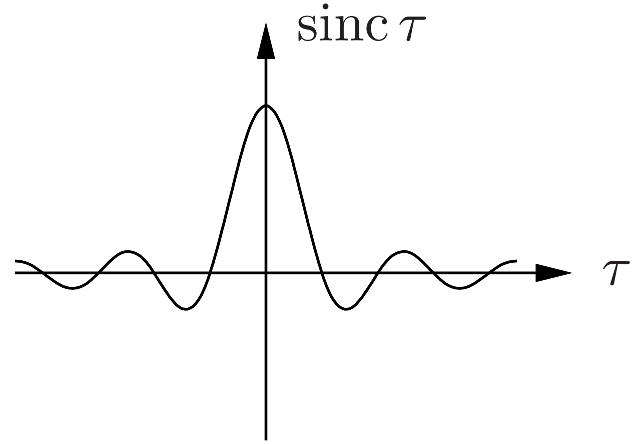
2.



3.

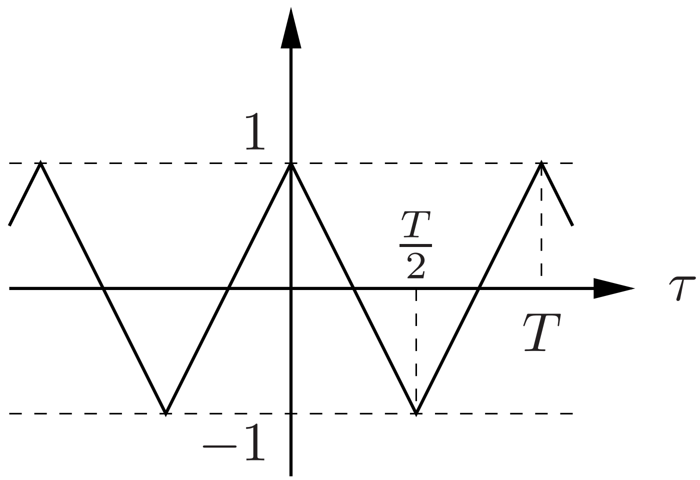


4.

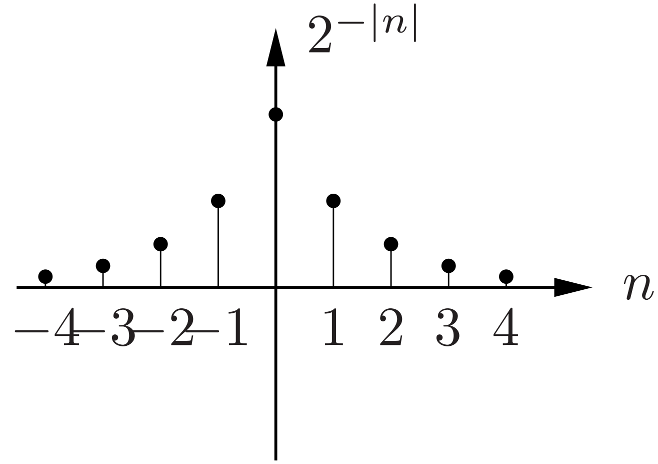


Which Functions can be an $R_X(\tau)$?

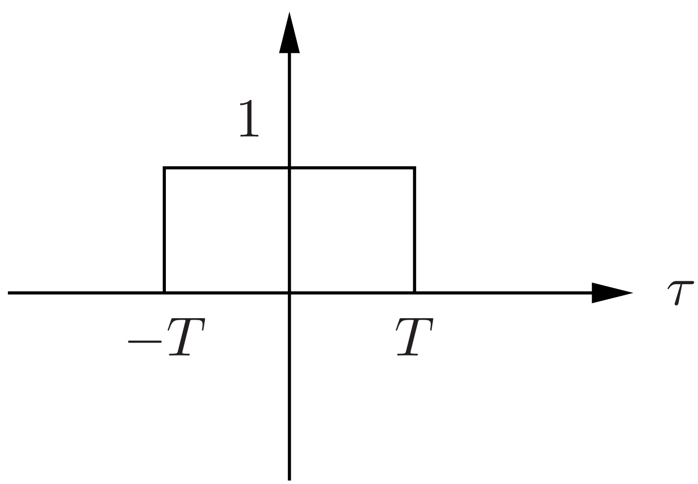
5.



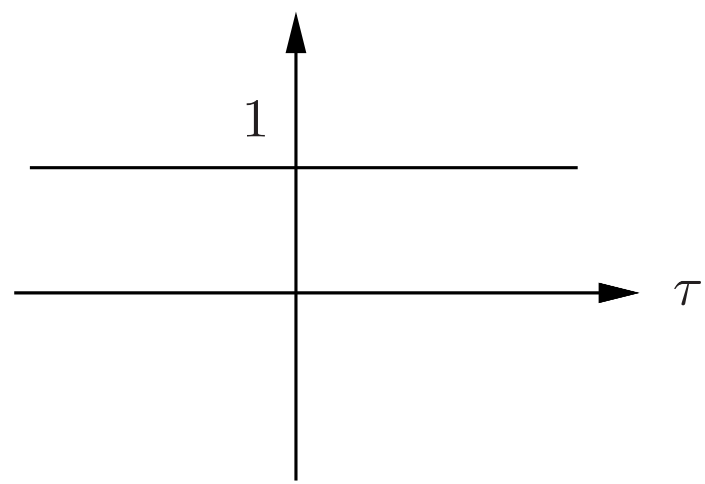
6.



7.

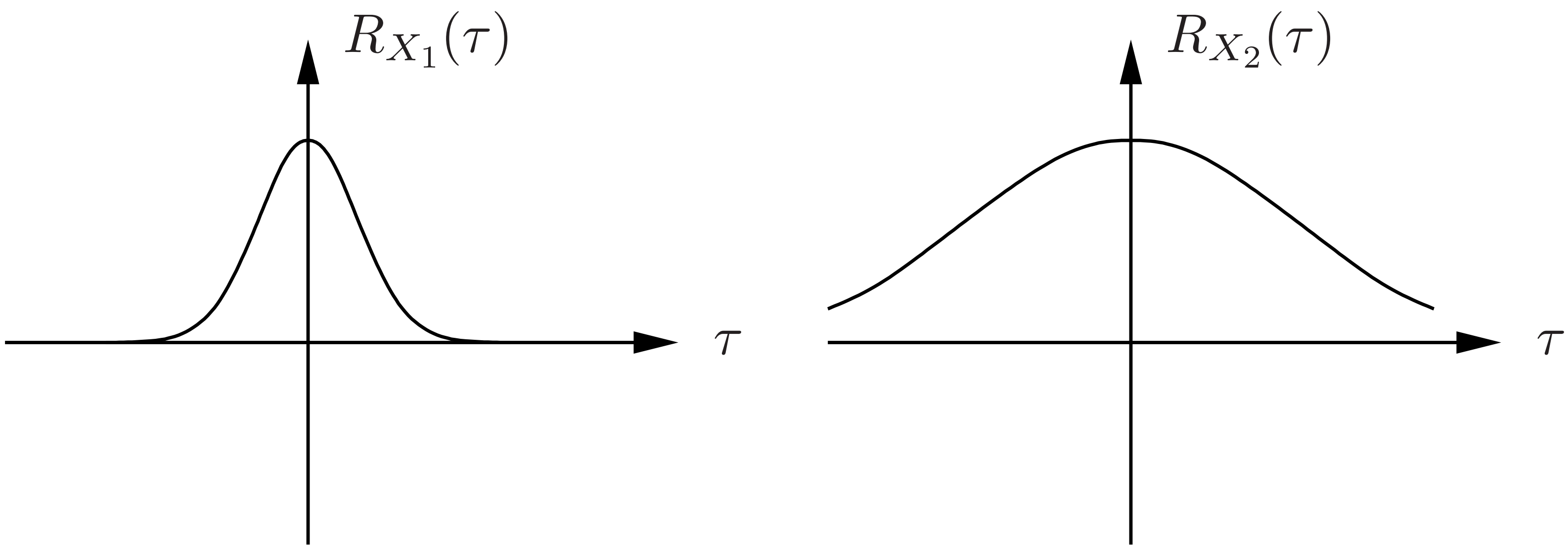


8.



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 $R_X(\tau)$:

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

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

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

- $R_X(\tau)$ (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. $\int_{-\infty}^{\infty} S_X(f)df = R_X(0) = \mathrm{E}(X^2(t))$, 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)

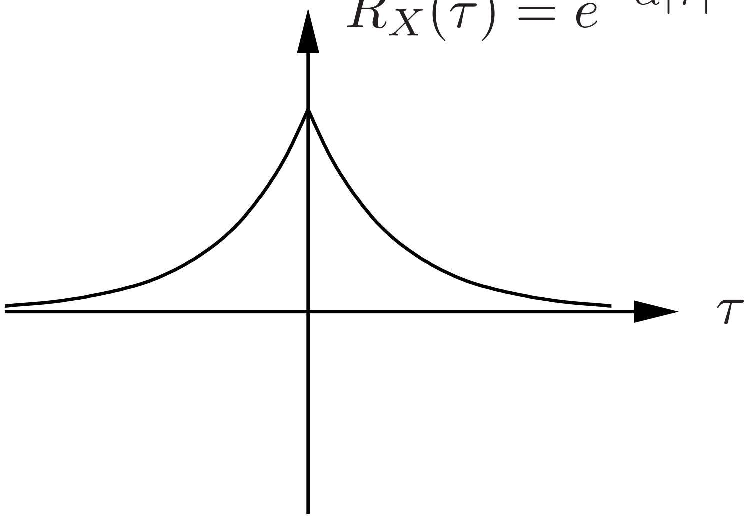
- From property 3, it follows that $S_X(f) \geq 0$. Why?

- In general, a function $S(f)$ is a psd if and only if it is real, even, nonnegative, and

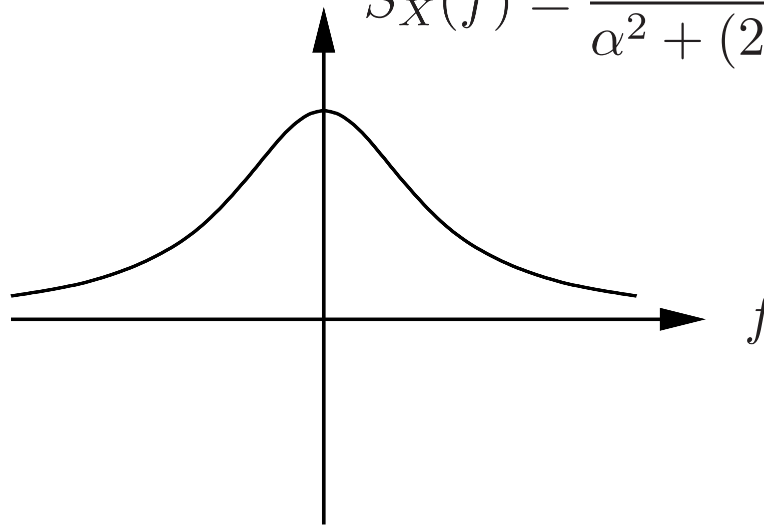
$$\int_{-\infty}^{\infty} S(f)\, df < \infty$$

Examples

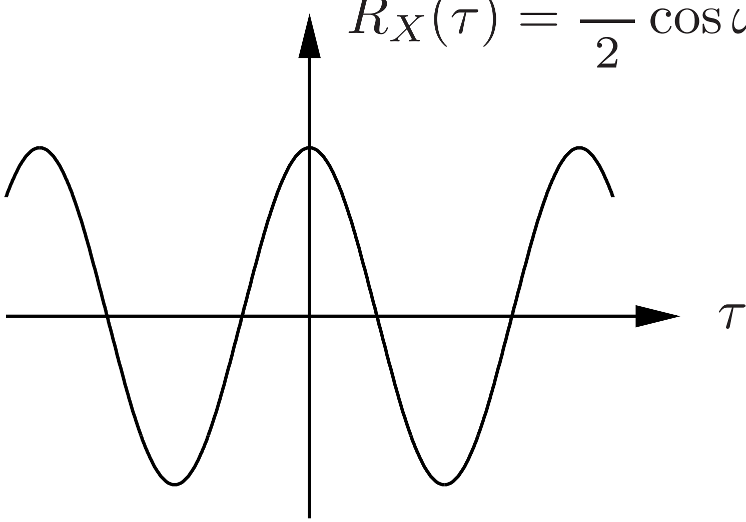
1. $R_X(\tau) = e^{-\alpha|\tau|}$



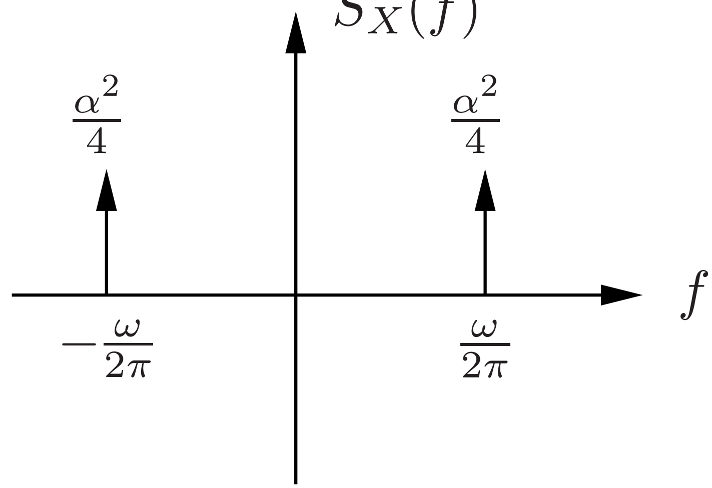
$$S_X(f) = \frac{2\alpha}{\alpha^2 + (2\pi f)^2}$$



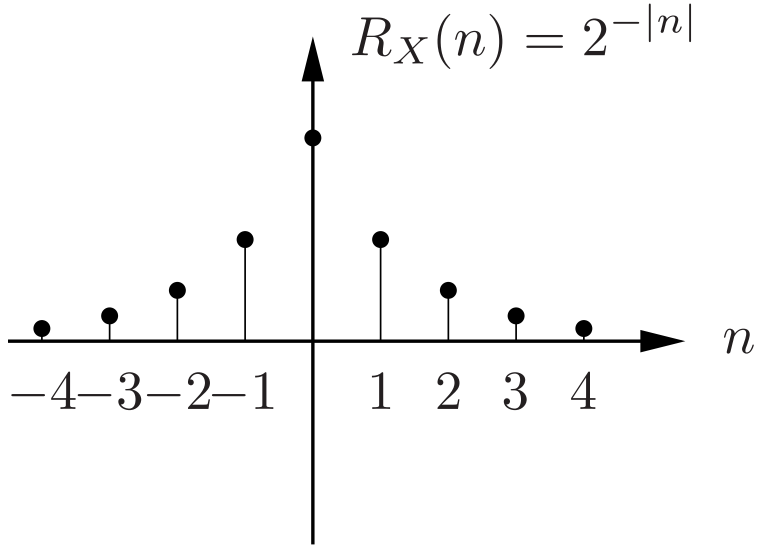
2. $R_X(\tau) = \frac{\alpha^2}{2}\cos\omega\tau$

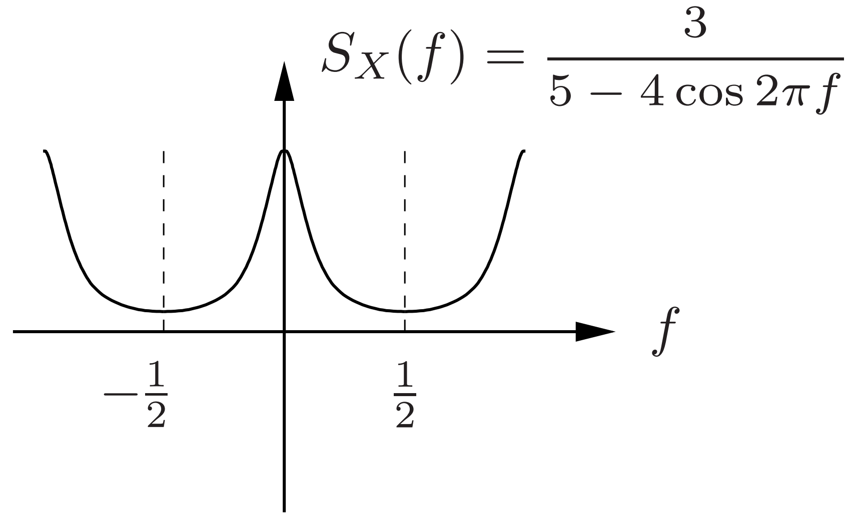


$S_X(f)$



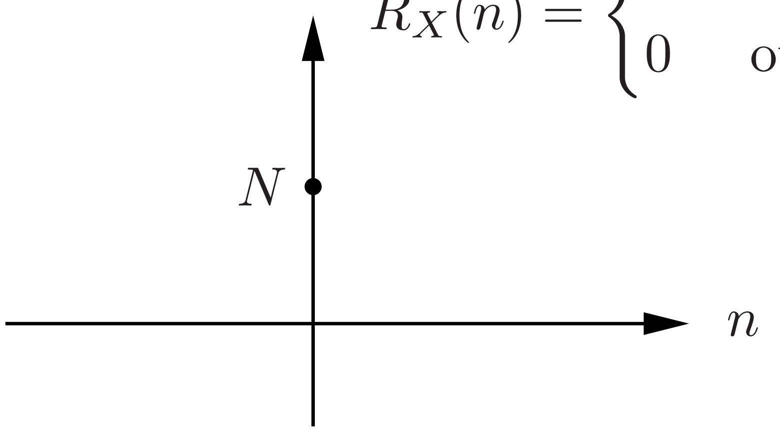
3.

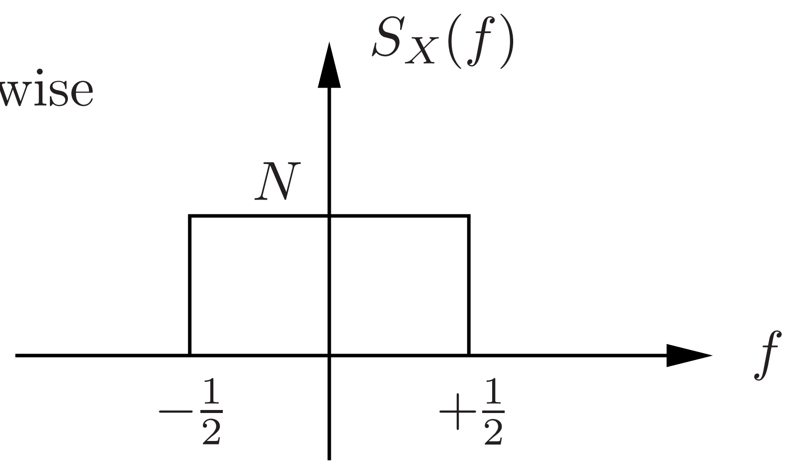




4. Discrete time white noise process: $X_1, X_2, \ldots, X_n, \ldots$ zero mean, uncorrelated, with average power N

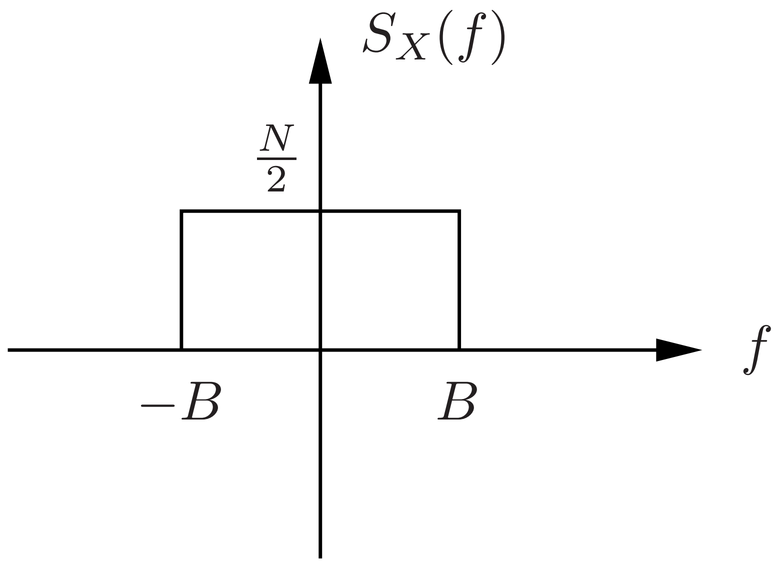
$$R_X(n) = \begin{cases} N & n = 0 \\ 0 & \text{otherwise} \end{cases}$$

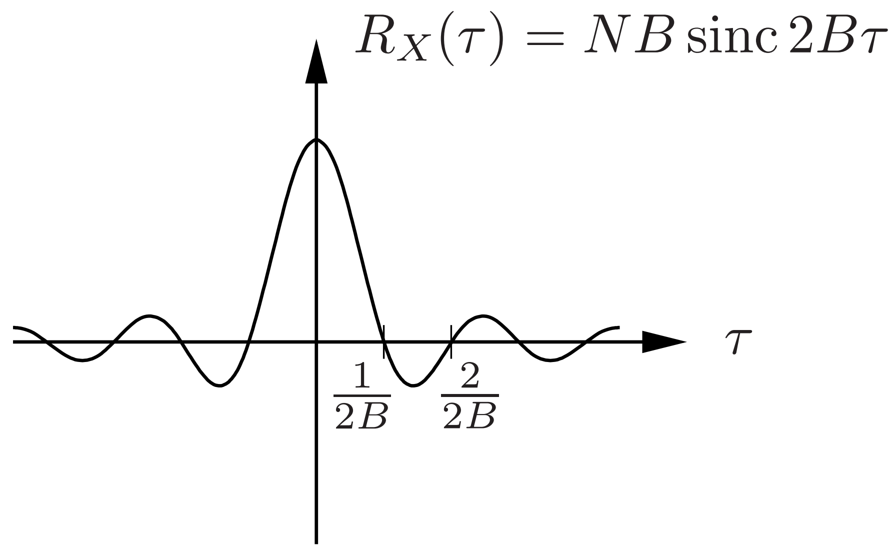




If X_n is also a GRP, then we obtain a discrete time WGN process

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





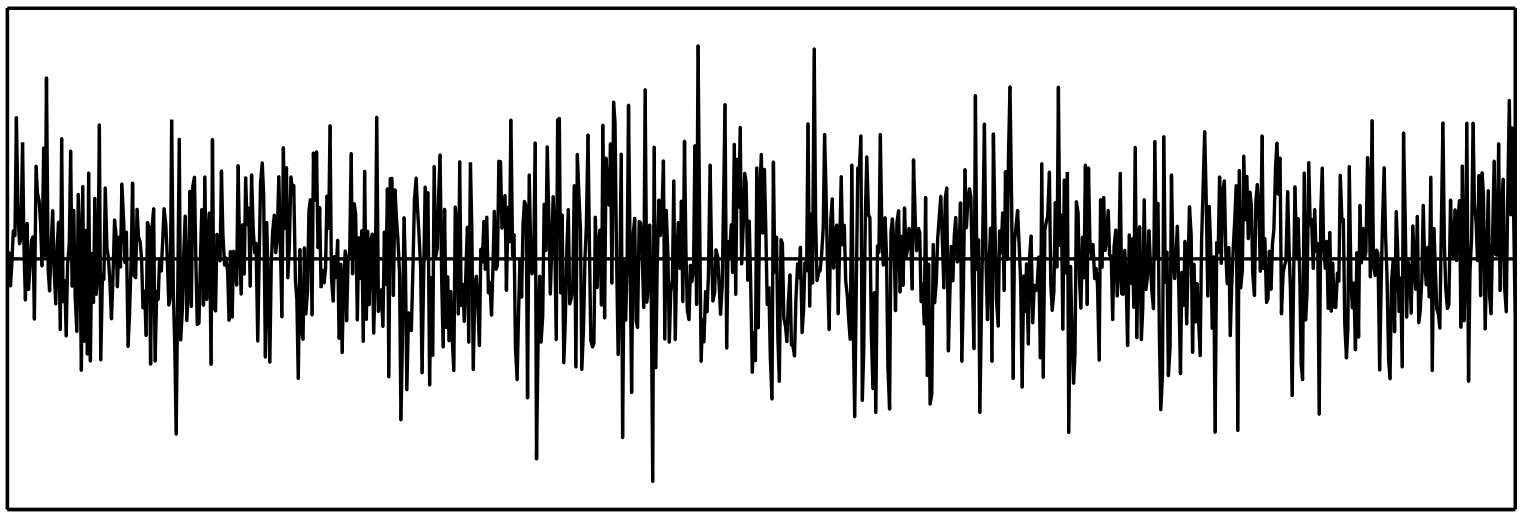
For any t, the samples $X\left(t \pm \frac{n}{2B}\right)$ for $n = 0, 1, 2, \ldots$ are uncorrelated

6. White noise process: If we let $B \to \infty$ 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)

Stationary Ergodic Random processes

- Ergodicity refers to certain time averages of random processes converging to their respective statistical averages
- We focus only on mean ergodicity of WSS processes
- Let X_n, $n = 1, 2, \ldots$, be a discrete time WSS process with mean μ and autocorrelation function $R_X(n)$
- To estimate the mean of X_n, we form the sample mean

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

- The process X_n is said to be mean ergodic if:

$$\bar{X}_n \to \mu \quad \text{in mean square,}$$

  i.e., $\lim_{n\to\infty} \mathrm{E}[(\bar{X}_n - \mu)^2] = 0$

- Since $\mathrm{E}(\bar{X}_n) = \mu$, this condition is equivalent to:

$$\mathrm{Var}(\bar{X}_n) \to 0 \text{ as } n \to \infty$$

- We can express this condition in terms of $C_X(n) = R_X(n) - \mu^2$ as follows

$$\begin{aligned}
\mathrm{Var}(\bar{X}_n) &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}[(X_i-\mu)(X_j-\mu)] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} C_X(i-j) \\
&= \frac{1}{n}C_X(0) + \frac{2}{n^2}\sum_{i=1}^{n-1}(n-i)C_X(i)
\end{aligned}$$

  Since by definition, $C_X(0) < \infty$, the condition for mean ergodicity is:

$$\frac{2}{n^2}\sum_{i=1}^{n-1}(n-i)C_X(i) \to 0$$

- Example: Let X_n a WSS process with $C_X(n) = 0$ for $n \neq 0$, then the process is mean ergodic

- The process does not need to have uncorrelated samples for it to be mean ergodic, however (see stationary Gauss-Markov process problem in HW7)

- Not every WSS process is mean ergodic

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

  However, it is not mean ergodic, since $\bar{X}_n \to P$ 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 SSS 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)

Mean Ergodicity for Continuous-time WSS Processes

- Let $X(t)$ be WSS process with mean μ and autocorrelation function $R_X(\tau)$
- To estimate the mean, we form the time average

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

But what does this integral mean?

- Integration of RP: Let $X(t)$ be a RP and $h(t)$ be a function. 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 $a \le \tau_1 \le a + \Delta \le \tau_2 \le a + 2\Delta \le \cdots \le \tau_{n-1} \le a + (n-1)\Delta \le \tau_n \le a + n\Delta = b$, then the above integral exists if the Riemann sum

$$\sum_{i=1}^{n-1} h(\tau_i)X(\tau_i)\Delta \quad \text{has a limit in m.s.}$$

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

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

Key fact: The existence of the m.s. integral depends only on R_X and h

More specifically, the above integral exists iff

$$\int_a^b \int_a^b R_X(t_1, t_2)h(t_1)h(t_2)dt_1dt_2$$

exists (in the normal sense)

- Now let's go back to mean ergodicity for continuous-time WSS process

  By definition, mean ergodicity means that

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

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

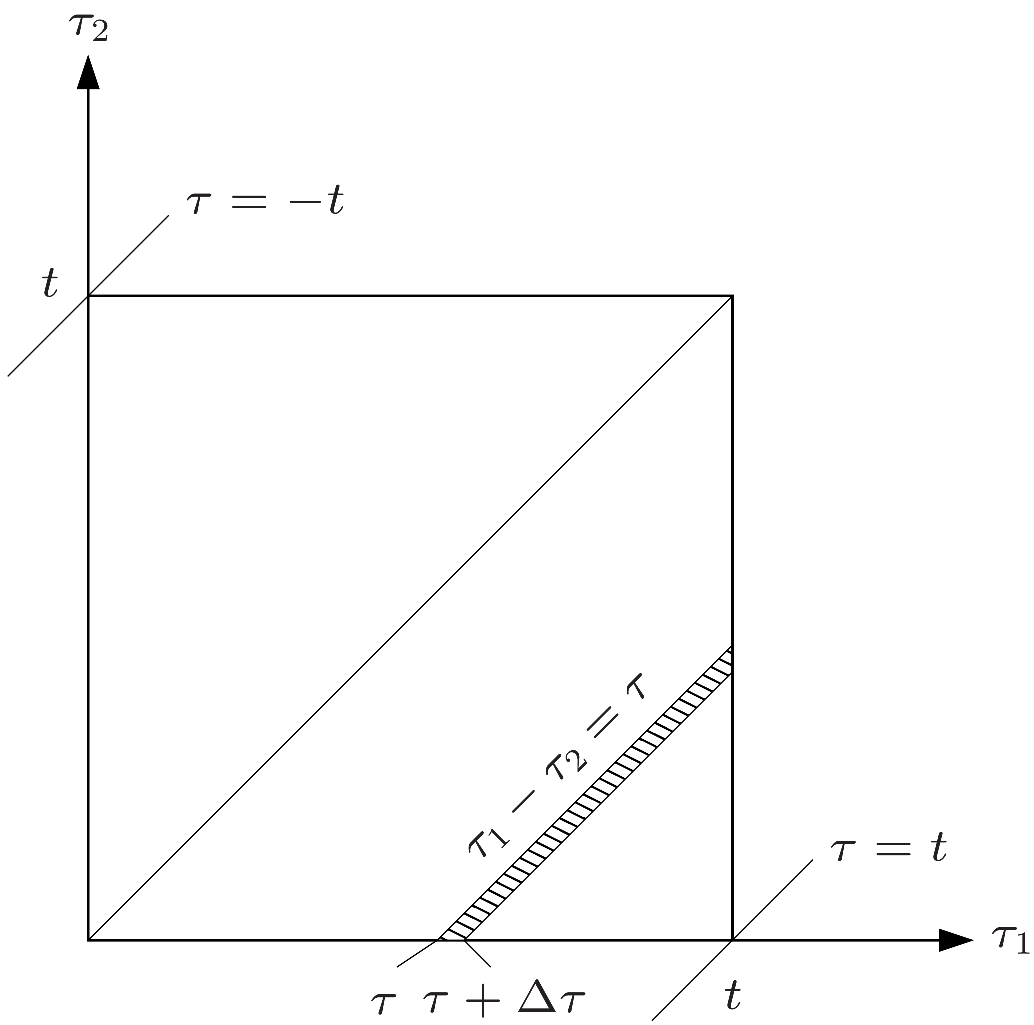
  $$\lim_{t\to\infty} \mathrm{Var}(\bar{X}(t)) = 0$$

Now, consider

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

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

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



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

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

- 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 $\to 0$ as $t \to \infty$. Hence $X(t)$ is mean ergodic