

Lecture Notes 9

Stationary Random Processes

- Strict-Sense and Wide-Sense Stationarity
- Autocorrelation Function of a Stationary Process
- Power Spectral Density
- Response of LTI System to WSS Process Input
- Linear Estimation: the Random Process Case

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 $X(t) = X(t_1 + (t - t_1))$ and $X(t_2 + (t - t_1)) = X(t + (t_2 - t_1))$

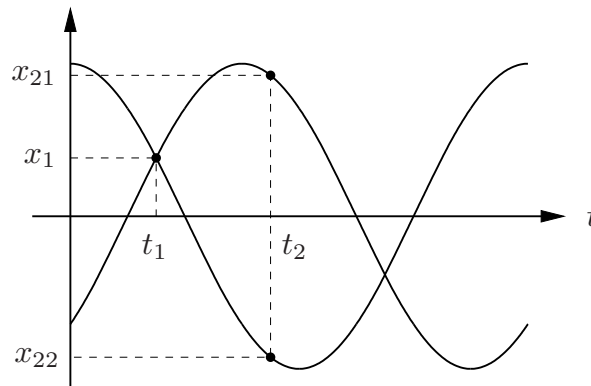
- Example: The random phase signal $X(t) = \alpha \cos(\omega t + \Theta)$ where $\Theta \in U[0, 2\pi]$ is SSS

- 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 are possible), and thus all n -th order pdfs are stationary
- IID processes are SSS
- Random walk process is not SSS (in fact, no independent increment process is 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 $t_2 - t_1$

- 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

Note: The technical condition $E[X(t)^2] < \infty$ is sometimes added to the definition of WSS. Although the condition holds for real-world processes and will be assumed to hold unless stated otherwise, it is not mathematically a part of stationarity. We will encounter a useful model which violates this condition — continuous time white noise,

- 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. In fact no independent increment process can be 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. $|R_X(\tau)| \leq R_X(0) = E[X^2(t)]$, the “average power” of $X(t)$

This can be shown using the Schwarz inequality. For any t ,

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

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)

- 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 any τ ,

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

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 $\mathbf{a} = (a_1, \dots, a_n)$,

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

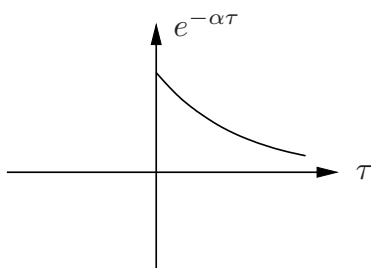
To see why this is necessary, recall that the correlation matrix for a random vector must be nonnegative definite, so if we take a set of n samples from the WSS random process, their correlation matrix must be nonnegative definite

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

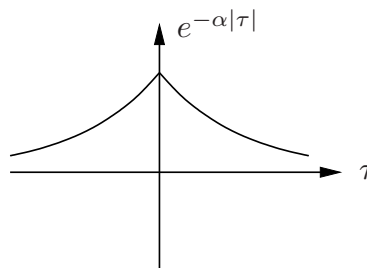
- The nonnegative definite condition may be difficult to verify directly. It turns out, however, to be equivalent to the condition that the Fourier transform of $R_X(\tau)$, which is called the *power spectral density* $S_X(f)$, is nonnegative for all frequencies f

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

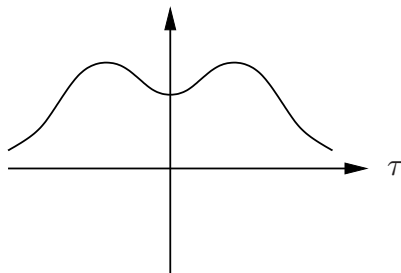
1.



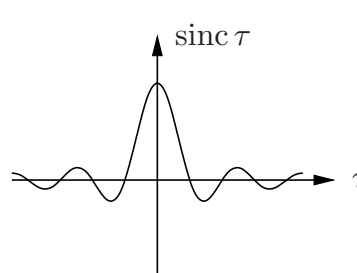
2.



3.

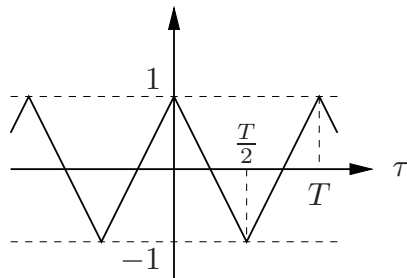


4.

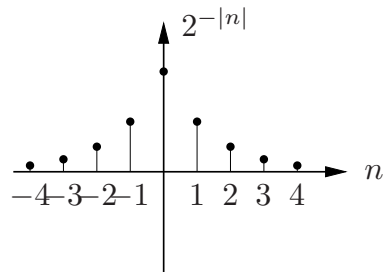


Which Functions can be an $R_X(\tau)$?

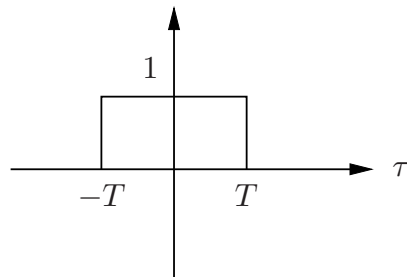
5.



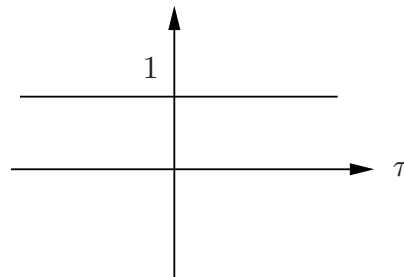
6.



7.

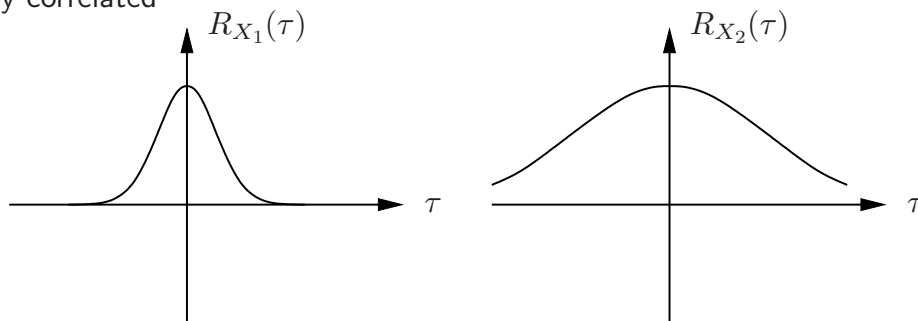


8.



Interpretation of Autocorrelation Function

- 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 Fourier transform (DFT) 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 DFT:

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

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

Properties of the Power Spectral Density

1. $S_X(f)$ is real and even, since the Fourier transform of a real and even function is real and even ($R_X(\tau)$ is real and even)
2. $\int_{-\infty}^{\infty} S_X(f) df = R_X(0) = E(X^2(t))$, 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

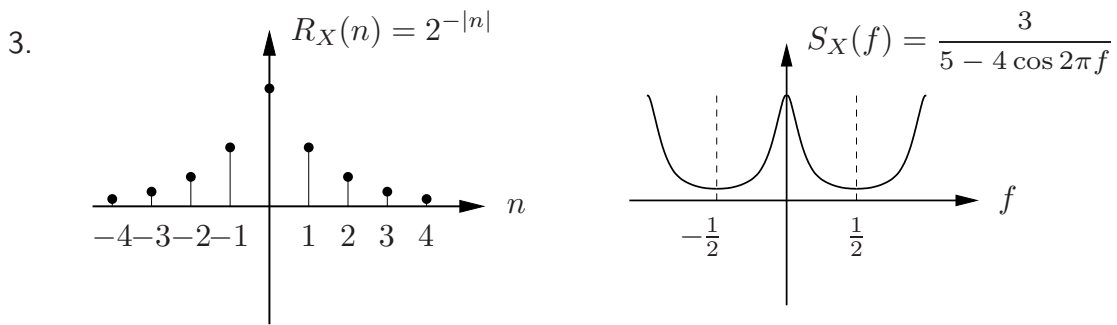
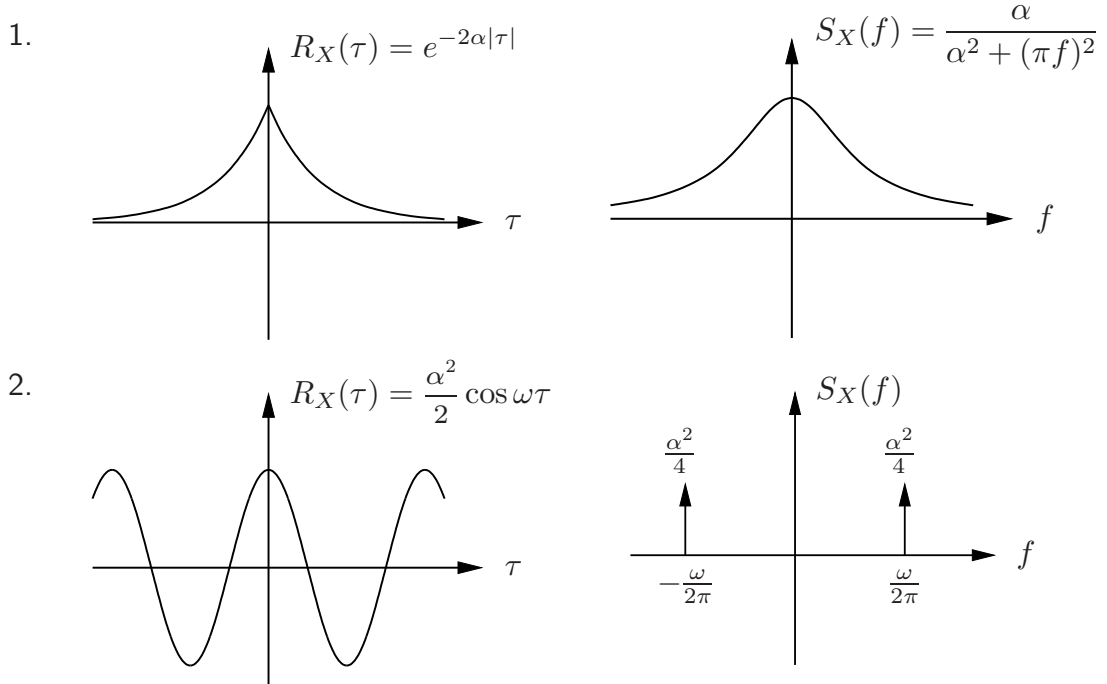
$$\int_{-f_2}^{-f_1} S_X(f) df + \int_{f_1}^{f_2} S_X(f) df = 2 \int_{f_1}^{f_2} S_X(f) df$$

(we will show this soon)

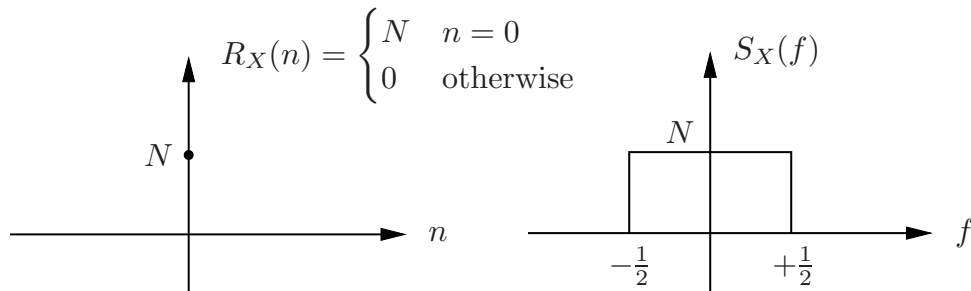
- From (3) it follows that $S_X(f) \geq 0$. Why?
- In general, a function $S(f)$ is a psd if and only if it is real, even, nonnegative, and

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

Examples

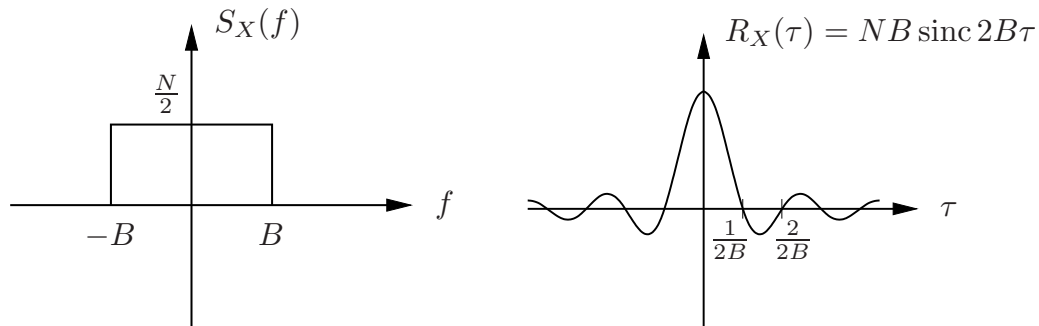


4. Discrete time white noise process: $X_1, X_2, \dots, X_n, \dots$ zero mean, uncorrelated, with common variance N

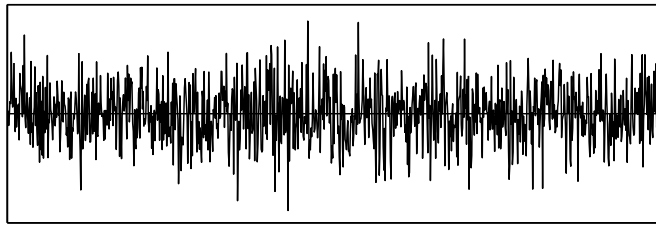


If X_n is also a GRP, then we obtain a discrete time WGN process

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: Consider a process continuous time random process with $E(X(t)) = 0$ and

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

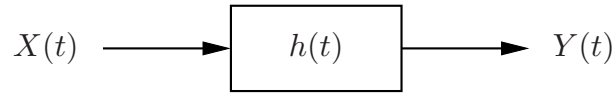
$$R_X(\tau) = \frac{N}{2} \delta(\tau)$$

This process is not physical (it has infinite power). 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
 - It plays a similar role in random processes to the role of a point mass in physics and delta function in EE
 - It will prove to be a useful and simple model for analyzing the output of a linear filter driven by a process with flat psd over a frequency range much larger than the filter bandwidth.
 - Thermal noise and shot noise are well modeled as white Gaussian noise, since they have very flat psd over very wide band (GHz)

Response of LTI System To WSS Process Input

- Consider a linear time invariant (LTI) system with impulse response $h(t)$ and transfer function $H(f) = \mathcal{F}(h(t))$, driven by WSS process $X(t)$, $-\infty < t < \infty$



- We want to characterize its output $Y(t) = X(t) * h(t) = \int_{-\infty}^{\infty} X(\tau)h(t - \tau)d\tau$
- Assuming that the system is in *steady state*, it turns out (not surprisingly) that the output is also a WSS process. In fact, $X(t)$ and $Y(t)$ are *jointly WSS*, which means that:
 - $X(t)$ and $Y(t)$ are WSS, and
 - Their *crosscorrelation function* $R_{XY}(t_1, t_2)$ is time invariant, i.e.,

$$R_{XY}(t_1, t_2) = E(X(t_1)Y(t_2)) = R_{XY}(t_1 + \tau, t_2 + \tau) \quad \text{for all } \tau$$

- Relabel $R_{XY}(t_1, t_2)$ for jointly WSS $X(t), Y(t)$ as $R_{XY}(\tau)$, where $\tau = t_1 - t_2$

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

Note that unlike $R_X(\tau)$, $R_{XY}(\tau)$ is not necessarily even. However,

$$R_{XY}(\tau) = R_{YX}(-\tau)$$

- Example: Let $\Theta \sim U[0, 2\pi]$. Consider two processes

$$X(t) = \alpha \cos(\omega t + \Theta) \quad \text{and} \quad Y(t) = \alpha \sin(\omega t + \Theta)$$

These processes are jointly WSS, since each is WSS (in fact SSS) and

$$\begin{aligned} R_{XY}(t_1, t_2) &= E [\alpha^2 \cos(\omega t_1 + \Theta) \sin(\omega t_2 + \Theta)] \\ &= \frac{\alpha^2}{4\pi} \int_0^{2\pi} [\sin(\omega(t_1 + t_2) + 2\theta) - \sin(\omega(t_1 - t_2))] d\theta \\ &= -\frac{\alpha^2}{2} \sin(\omega(t_1 - t_2)) \end{aligned}$$

- We define the *cross power spectral density* for jointly WSS processes $X(t), Y(t)$ as

$$S_{XY}(f) = \mathcal{F}(R_{XY}(\tau))$$

- Example: Let $Y(t) = X(t) + Z(t)$, where $X(t)$ and $Z(t)$ are zero mean uncorrelated WSS processes. Show that $Y(t)$ and $X(t)$ are jointly WSS, and find $R_{XY}(\tau)$ (in terms of R_X and R_Z) and $S_{XY}(f)$ (in terms of S_X and S_Z)

Solution: First, we show that $Y(t)$ is WSS, since it is zero mean and

$$\begin{aligned} R_Y(t_1, t_2) &= E [(X(t_1) + Z(t_1))(X(t_2) + Z(t_2))] \\ &= E (X(t_1)X(t_2)) + E (Z(t_1)Z(t_2)) \\ &\quad (X(t), Z(t) \text{ zero mean, uncorrelated}) \\ &= R_X(\tau) + R_Z(\tau) \end{aligned}$$

Taking the Fourier transform of both sides, $S_Y(f) = S_X(f) + S_Z(f)$

To show that $Y(t)$ and $X(t)$ are jointly WSS, we need to show that their crosscorrelation function is time invariant

$$\begin{aligned} R_{XY}(t_1, t_2) &= E [X(t_1)(X(t_2) + Z(t_2))] \\ &= E (X(t_1)X(t_2)) + E (X(t_1)Z(t_2)) \\ &= R_X(t_1, t_2) + 0 \quad (X(t), Z(t) \text{ zero mean, uncorrelated}) \\ &= R_X(\tau) \end{aligned}$$

Taking the Fourier transform, $S_{XY}(f) = S_X(f)$

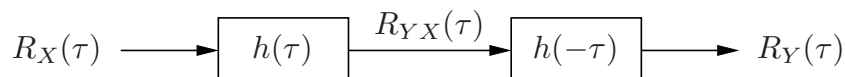
Output Mean, Autocorrelation, and PSD

Theorem: Let $X(t)$, $t \in \mathbf{R}$, be a WSS process input to a LTIS with impulse response $h(t)$ and transfer function $H(f)$. If the system is *stable*, i.e.,

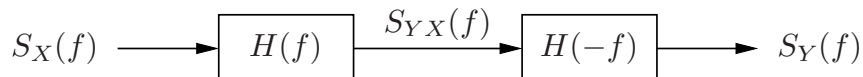
$$\left| \int_{-\infty}^{\infty} h(t) dt \right| = |H(0)| < \infty,$$

then the input $X(t)$ and output $Y(t)$ are jointly WSS with:

1. $E(Y(t)) = H(0) E(X(t))$
2. $R_{YX}(\tau) = h(\tau) * R_X(\tau)$
3. $R_Y(\tau) = h(\tau) * R_X(\tau) * h(-\tau)$



4. $S_{YX}(f) = H(f)S_X(f)$
5. $S_Y(f) = |H(f)|^2 S_X(f)$



Proof: Note that here the LTIS is in steady state

1. To find the mean of $Y(t)$, consider

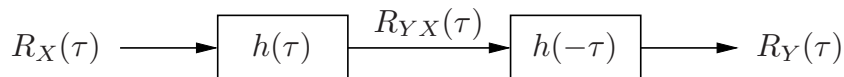
$$\begin{aligned} E(Y(t)) &= E\left(\int_{-\infty}^{\infty} X(\tau)h(t-\tau) d\tau\right) \\ &= \int_{-\infty}^{\infty} E(X(\tau)) h(t-\tau) d\tau \\ &= E(X(t)) \int_{-\infty}^{\infty} h(t-\tau) d\tau = E(X(t))H(0) \end{aligned}$$

2. To find the crosscorrelation function between $Y(t)$ and $X(t)$, consider

$$\begin{aligned} R_{YX}(\tau) &= E(Y(t+\tau)X(t)) \\ &= E\left(\int_{-\infty}^{\infty} h(\alpha)X(t+\tau-\alpha)X(t) d\alpha\right) \\ &= \int_{-\infty}^{\infty} h(\alpha)R_X(\tau-\alpha) d\alpha \\ &= h(\tau) * R_X(\tau) \end{aligned}$$

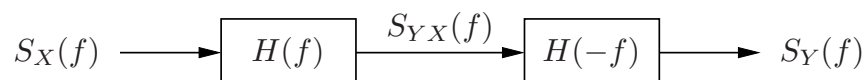
3. To find the autocorrelation function of $Y(t)$, consider

$$\begin{aligned} R_Y(\tau) &= E(Y(t+\tau)Y(t)) \\ &= E\left(Y(t+\tau) \int_{-\infty}^{\infty} h(\alpha)X(t-\alpha) d\alpha\right) \\ &= \int_{-\infty}^{\infty} h(\alpha)R_{YX}(\tau+\alpha) d\alpha \\ &= R_{YX}(\tau) * h(-\tau) \end{aligned}$$



4. Follows by taking the Fourier transform of $R_{YX}(\tau)$

5. Follows by taking the Fourier transform of $R_Y(\tau)$

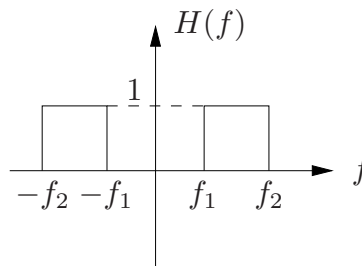
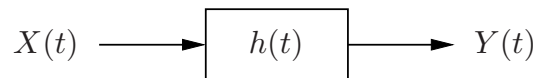


$S_X(f)$ is the Power Spectral Density

- We can use the above results to show that $S_X(f)$ is indeed the power spectral density of $X(t)$; i.e., the average power in any frequency band $[f_1, f_2]$ is

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

- To show this we pass $X(t)$ through an ideal band-pass filter



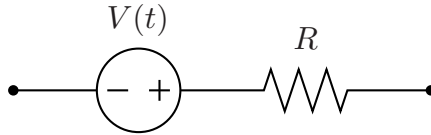
- Now the average power of $X(t)$ in the band $[f_1, f_2]$ is given by

$$\begin{aligned} E(Y^2(t)) &= \int_{-\infty}^{\infty} S_Y(f) df \\ &= \int_{-\infty}^{\infty} |H(f)|^2 S_X(f) df \\ &= \int_{-f_2}^{-f_1} S_X(f) df + \int_{f_1}^{f_2} S_X(f) df \\ &= 2 \int_{f_1}^{f_2} S_X(f) df \end{aligned}$$

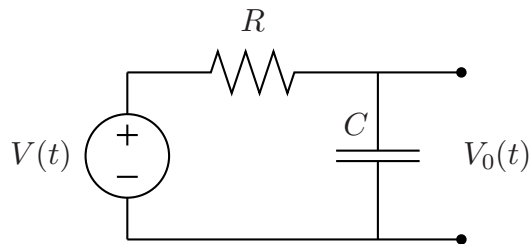
- This also shows that $S_X(f) \geq 0$ for all f
- Suppose that $X(t)$ is white noise and $|H(f)| = 0$ for $|f| \geq B$. Then the output psd will be the same as if the input were bandlimited white noise. Hence so far as the output psd is concerned, white noise and bandlimited white noise are equivalent.
- If White noise is put into the above ideal band-pass filter, the output is called bandpass white noise.

KT/C Noise

- The noise in a resistor R (in ohms) due to thermal noise is modeled as a WGN voltage source $V(t)$ in series with R . The psd of $V(t)$ is $S_V(f) = 2kTR \text{ V}^2/\text{Hz}$ for all f , where k is Boltzman's constant and T is the temperature in degrees K



- Now let's find the average output noise power for an RC circuit



- First we find the transfer function for the circuit

$$H(f) = \frac{1}{1 + i2\pi fRC} \Rightarrow |H(f)|^2 = \frac{1}{1 + (2\pi fRC)^2}$$

- Now we write the output psd in terms of the input psd as

$$S_{V_o} = S_V(f)|H(f)|^2 = 2kTR \frac{1}{1 + (2\pi fRC)^2}, \quad -\infty < f < \infty$$

- Thus the average output power is

$$\begin{aligned} E(V_o^2(t)) &= \int_{-\infty}^{\infty} S_{V_o}(f) df \\ &= \frac{2kTR}{2\pi RC} \int_{-\infty}^{\infty} \frac{1}{1 + (2\pi fRC)^2} d(2\pi fRC) \\ &= \frac{kT}{\pi C} \int_{-\infty}^{\infty} \frac{1}{1 + x^2} dx \\ &= \frac{kT}{\pi C} \arctan x \Big|_{-\infty}^{+\infty} = \frac{kT}{\pi C} \pi = \frac{kT}{C}, \end{aligned}$$

which is independent of R !

Linear Estimation: the Random Process Case

- Let $X(t)$ and $Y(t)$ be zero mean jointly WSS processes with known autocorrelation and crosscorrelation functions $R_X(\tau)$, $R_Y(\tau)$, and $R_{XY}(\tau)$
- We observe the random process $Y(\alpha)$ for $t - a \leq \alpha \leq t + b$ ($-a \leq b$) and wish to find the MMSE linear estimate of the signal $X(t)$, i.e., $\hat{X}(t)$ such that the $\text{MSE} = \text{E} [(X(t) - \hat{X}(t))^2]$ is minimized
- The linear estimate is of the form

$$\hat{X}(t) = \int_{-b}^a h(\tau) Y(t - \tau) d\tau$$

- By the orthogonality principle, the MMSE linear estimate must satisfy

$$(X(t) - \hat{X}(t)) \perp Y(t - \tau), \quad -b \leq \tau \leq a$$

or

$$\text{E} [(X(t) - \hat{X}(t))Y(t - \tau)] = 0, \quad -b \leq \tau \leq a$$

Thus, for $-b \leq \tau \leq a$, we must have

$$\begin{aligned} R_{XY}(\tau) &= \text{E} [X(t)Y(t - \tau)] = \text{E} [\hat{X}(t)Y(t - \tau)] \\ &= \text{E} \left(\int_{-b}^a h(\alpha) Y(t - \alpha) Y(t - \tau) d\alpha \right) \\ &= \int_{-b}^a h(\alpha) R_Y(\tau - \alpha) d\alpha \end{aligned}$$

So, to find $h(\alpha)$ we need to solve an infinite set of integral equations

- Solving these equations analytically is not possible in general. However, it can be done for two important special cases:
 - *Infinite smoothing*: when $a, b \rightarrow \infty$
 - *Filtering*: when $a \rightarrow \infty$ and $b = 0$

We discuss only the first case. (The second case, which leads to the *Wiener-Hopf equations* from which the famous *Wiener filter* is derived, is covered in EE 378.)

Infinite Smoothing

- When $a, b \rightarrow \infty$, the integral equations for the MMSE linear estimate become

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} h(\alpha) R_Y(\tau - \alpha) d\alpha, \quad -\infty < \tau < +\infty$$

In other words,

$$R_{XY}(\tau) = h(\tau) * R_Y(\tau)$$

- The Fourier transform convolution theorem gives the transfer function for the optimal *infinite smoothing filter*:

$$S_{XY}(f) = H(f)S_Y(f) \Rightarrow H(f) = \frac{S_{XY}(f)}{S_Y(f)}$$

- The minimum MSE is given by

$$\begin{aligned} \text{MSE} &= \text{E} [(X(t) - \hat{X}(t))^2] \\ &= \text{E} [(X(t) - \hat{X}(t))X(t)] - \text{E} [(X(t) - \hat{X}(t))\hat{X}(t)] \\ &= \text{E} [(X(t) - \hat{X}(t))X(t)] \quad (\text{by orthogonality}) \\ &= \text{E} [(X(t))^2] - \text{E} [\hat{X}(t)X(t)] \end{aligned}$$

To evaluate the second term, consider

$$\begin{aligned} R_{X\hat{X}}(\tau) &= \text{E}(X(t + \tau)\hat{X}(t)) \\ &= \text{E}\left(X(t + \tau) \int_{-\infty}^{\infty} h(\alpha)Y(t - \alpha) d\alpha\right) \\ &= \int_{-\infty}^{\infty} h(\alpha)R_{XY}(\tau + \alpha) d\alpha = R_{XY}(\tau) * h(-\tau) \end{aligned}$$

Therefore

$$\text{E}(\hat{X}(t)X(t)) = R_{X\hat{X}}(0) = \int_{-\infty}^{\infty} H(-f)S_{XY}(f) df = \int_{-\infty}^{\infty} \frac{|S_{XY}(f)|^2}{S_Y(f)} df,$$

and the minimum MSE is given by

$$\begin{aligned}
 \mathbb{E} [(X(t) - \hat{X}(t))^2] &= \mathbb{E} [(X(t)^2)] - \mathbb{E} (\hat{X}(t)X(t)) \\
 &= \int_{-\infty}^{\infty} S_X(f) df - \int_{-\infty}^{\infty} \frac{|S_{XY}(f)|^2}{S_Y(f)} df \\
 &= \int_{-\infty}^{\infty} \left(S_X(f) - \frac{|S_{XY}(f)|^2}{S_Y(f)} \right) df
 \end{aligned}$$

Example: Additive White Noise Channel

- Let $X(t)$ and $Z(t)$ be zero mean uncorrelated WSS processes with

$$S_X(f) = \begin{cases} \frac{P}{2} & |f| \leq B \\ 0 & \text{otherwise} \end{cases}$$

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

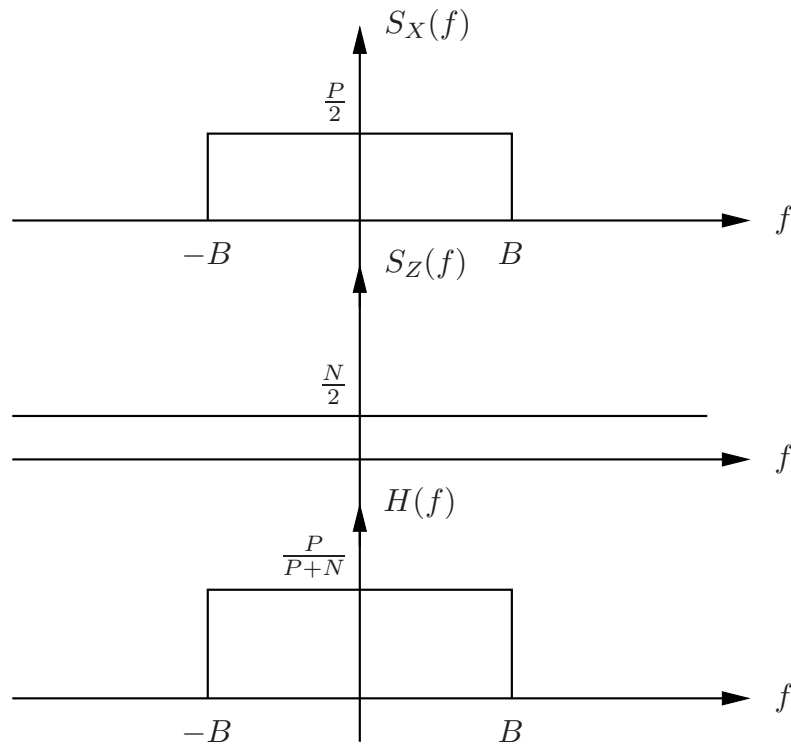
Here the signal X is bandlimited white noise, and Z is white noise

- Find the optimal infinite smoothing filter for estimating $X(t)$ given

$$Y(\tau) = X(\tau) + Z(\tau), \quad -\infty < \tau < +\infty$$

and the MSE for the estimate produced by this filter

The power spectral densities of X and Z are shown below



- The transfer function of the optimal infinite smoothing filter is given by

$$\begin{aligned}
 H(f) &= \frac{S_{XY}(f)}{S_Y(f)} \\
 &= \frac{S_X(f)}{S_X(f) + S_Z(f)} \\
 &= \begin{cases} \frac{P}{P+N} & |f| \leq B \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

The MMSE is given by

$$\begin{aligned}
 \text{MSE} &= \int_{-\infty}^{\infty} S_X(f) df - \int_{-\infty}^{\infty} \frac{|S_{XY}(f)|^2}{S_Y(f)} df \\
 &= \int_{-B}^{+B} \frac{P}{2} df - \int_{-B}^{+B} \frac{(P/2)^2}{P/2 + N/2} df \\
 &= PB - \frac{P^2/4}{(P+N)/2} 2B \\
 &= \frac{NPB}{N+P}
 \end{aligned}$$