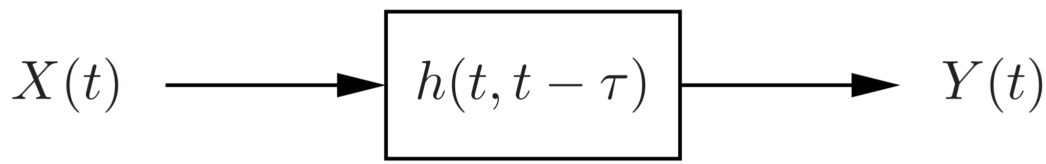
Lecture Notes 8
Random Processes in Linear Systems

- Linear System with Random process Input
- LTI System with WSS Process Input
- Process Linear Estimation
  - Infinite smoothing filter
  - Spectral Factorization
  - Wiener Filter

Linear System with Random Process Input

- Consider a linear system with (time-varying) impulse response $h(t, t-\tau)$ driven by a random process input $X(t)$



- The output of the system is

$$Y(t) = \int_{-\infty}^{\infty} h(t, t-\tau) X(\tau)\, d\tau$$

- We wish to specify the output random process $Y(t)$
- It is difficult to obtain a complete specification of the output process in general
- Important special case: If $X(t)$ is a GRP, the output process $Y(t)$ is also a GRP (since the integral above can be approximated by a sum and thus the output process is obtained via a linear transformation of $X(t)$)

- We focus on finding the mean and autocorrelation functions of $Y(t)$ in terms of the mean and autocorrelation functions of the input process $X(t)$ and the impulse response of the system $h(t, t-\tau)$

  We are also interested in finding the crosscorrelation function between $X(t)$ and $Y(t)$ defined as

$$R_{XY}(t_1, t_2) = \mathrm{E}(X(t_1)Y(t_2))$$

  Note that unlike $R_X(t_1, t_2)$, $R_{XY}(t_1, t_2)$ is not necessarily symmetric in t_1 and t_2. However, $R_{XY}(t_1, t_2) = R_{YX}(t_2, t_1)$

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

$$\mathrm{E}(Y(t)) = \mathrm{E}\left(\int_{-\infty}^{\infty} h(t, t-\tau)X(\tau)\, d\tau\right) = \int_{-\infty}^{\infty} h(t, t-\tau)\,\mathrm{E}(X(\tau))\, d\tau$$

- The crosscorrelation function between $Y(t)$ and $X(t)$ is

$$\begin{aligned} R_{YX}(t_1, t_2) &= \mathrm{E}(Y(t_1)X(t_2)) \\ &= \mathrm{E}\left(\int_{-\infty}^{\infty} h(t_1, t_1-\tau)X(\tau)X(t_2)\, d\tau\right) \\ &= \int_{-\infty}^{\infty} h(t_1, t_1-\tau)R_X(\tau, t_2)\, d\tau \end{aligned}$$

- The autocorrelation function of $Y(t)$ is

$$\begin{aligned} R_Y(t_1, t_2) &= \mathrm{E}(Y(t_1)Y(t_2)) \\ &= \mathrm{E}\left(\int_{-\infty}^{\infty} h(t_2, t_2 - \tau)X(\tau)Y(t_1)\, d\tau\right) \\ &= \int_{-\infty}^{\infty} h(t_2, t_2 - \tau)R_{YX}(t_1, \tau)\, d\tau \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t_2, t_2 - \tau_2)h(t_1, t_1 - \tau_1)R_X(\tau_1, \tau_2)\, d\tau_1\, d\tau_2 \end{aligned}$$

  The average power is

$$\mathrm{E}(Y^2(t)) = R_Y(t, t)$$

- Example (Integrator): Let $X(t)$ be a white noise process with autocorrelation function $R_X(\tau) = (N/2)\delta(\tau)$ and let the linear system be an ideal integrator, i.e.,

$$Y(t) = \int_0^t X(\tau)\, d\tau$$

  Find the mean and autocorrelation functions and the average power of the integrator output $Y(t)$, for $t > 0$

This example is motivated by several applications:

- Noise in an image sensor pixel: the white noise models the photodetector shot noise, which is integrated with the signal over a capacitor before sampling
- Noise in a voltage controlled oscillator (for phase locked loops)

• Solution: The mean is

$$\mathrm{E}(Y(t)) = \int_0^t \mathrm{E}(X(\tau))\, d\tau = 0$$

To obtain the autocorrelation function and average power for this case, we can specialize the previous results to

$$\begin{aligned} R_{YX}(t_1, t_2) &= \int_0^{t_1} \frac{N}{2}\delta(t_2 - \tau)\, d\tau \\ &= \begin{cases} \frac{N}{2}, & \text{for } t_2 \le t_1 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

$$R_Y(t_1,t_2) = \begin{cases} \frac{N}{2}t_2, & \text{for } t_2 \le t_1 \\ \frac{N}{2}t_1 & \text{otherwise} \end{cases}$$

$$= \frac{N}{2}\min\{t_1,t_2\}$$

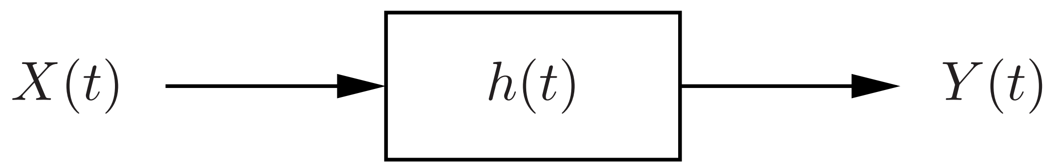
$$\mathrm{E}(Y^2(t)) = R_Y(t,t) = \frac{N}{2}t$$

Note that the average power grows linearly with t (as for the random walk)

- If in addition $X(t)$ is a GRP, then $Y(t)$ is also a GRP and is referred to as the Wiener process

LTI System with WSS Process Input

- Consider a linear time invariant (LTI) system with real 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 $Y(t) = X(t) * h(t) = \int_{-\infty}^{\infty} X(\tau)h(t-\tau)d\tau$

- It turns out (not surprisingly) that if the system is stable, i.e., $\left| \int_{-\infty}^{\infty} h(t)\, dt \right| = |H(0)| < \infty$, then $X(t)$ 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.,

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

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

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

Again $R_{XY}(\tau)$ is not necessarily even. However,

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

- Example: Let $\Theta \sim \mathrm{U}[0, 2\pi]$. 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

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

- 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

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

  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

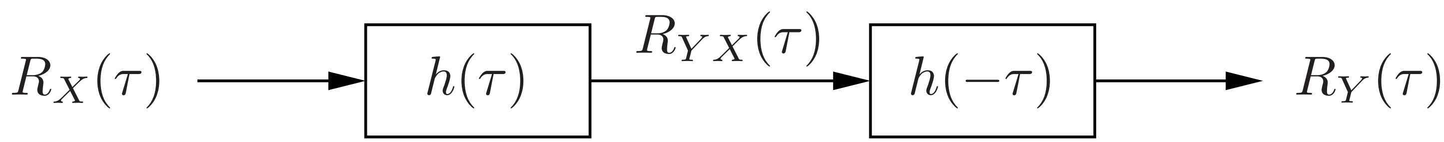
$$\begin{aligned} R_{XY}(t_1, t_2) &= \mathrm{E}\left[X(t_1)(X(t_2) + Z(t_2))\right] \\ &= \mathrm{E}\big(X(t_1)X(t_2)\big) + \mathrm{E}\big(X(t_1)Z(t_2)\big) \\ &= R_X(t_1, t_2) + 0 \quad (X(t),\, Z(t) \text{ zero mean, uncorrelated}) \\ &= R_X(\tau) \end{aligned}$$

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

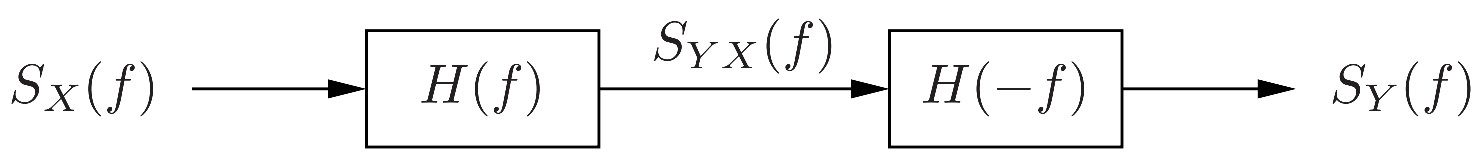
Output Mean, Autocorrelation, and PSD

Theorem: Let $X(t)$, $t \in \mathbf{R}$, be a WSS process input to a stable LTI system with real impulse response $h(t)$ and transfer function $H(f)$. Then the input $X(t)$ and output $Y(t)$ are jointly WSS with:

1. $\mathrm{E}(Y(t)) = H(0)\,\mathrm{E}(X(t))$
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. $S_Y(f) = |H(f)|^2 S_X(f)$



Remark: For a discrete time WSS process $X(n)$ and a stable LTI system $h(n)$, $X(n)$ and the output process $Y(n)$ are jointly WSS and we can similarly find $R_Y(n), \ldots$

Proof: Note that here the LTI system is in steady state

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

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

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

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

$$= \int_{-\infty}^{\infty} h(\alpha) R_X(\tau - \alpha)\, d\alpha$$

$$= h(\tau) * R_X(\tau)$$

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

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

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

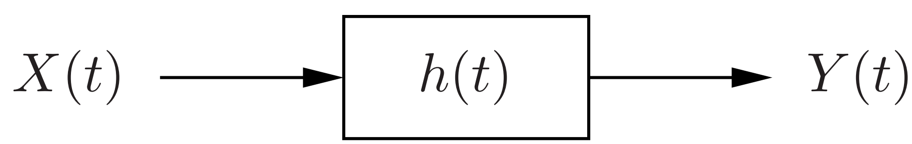
5. Follows by taking the Fourier transform of $R_Y(\tau)$

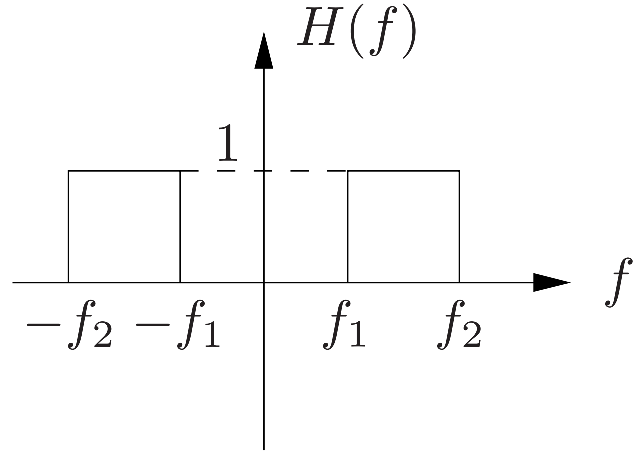
$S_X(f)$ is the Power Spectral Density

- 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





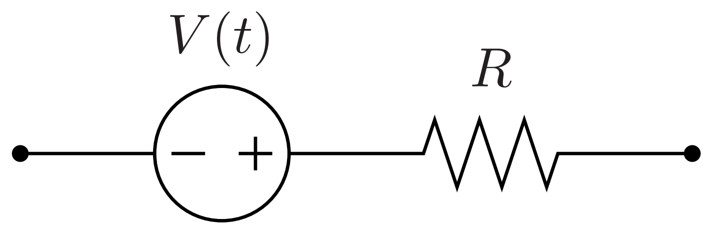
- Now the average power of $X(t)$ in the band $[f_1, f_2]$ is

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

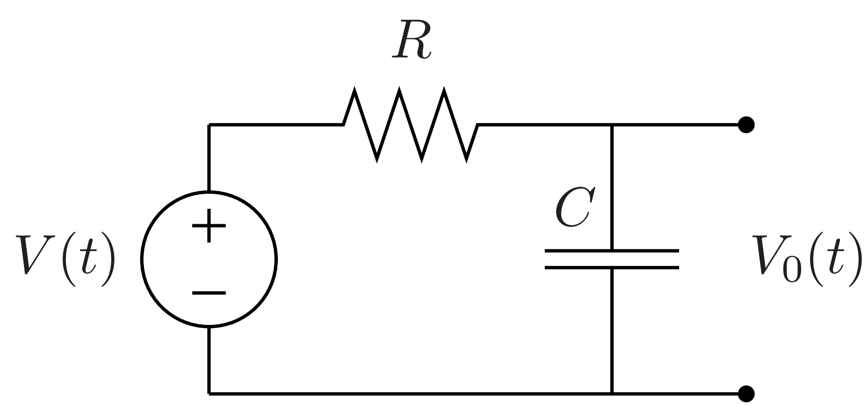
- This also shows that $S_X(f) \geq 0$ for all f

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 $S_V(f) = 2kTR\ \mathrm{V}^2/\mathrm{Hz}$ for all f, where k is Boltzmann'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

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

- Thus the average output power is

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

which is independent of R!

Autoregressive Moving Average Process

- Let X_n, $-\infty < n < \infty$, be a discrete time white noise process with average power N

  The autoregressive moving average (ARMA) process of order (p, q), Y_n, $-\infty < n < \infty$, is defined as

  $$Y_n = -\sum_{k=1}^{p} \alpha_k Y_{n-k} + \sum_{l=0}^{q} \beta_l X_{n-l}$$

  where $\beta_0 = 1$, $\alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_q$ are fixed parameters

- This process can be viewed as the output of an LTI system with transfer function

  $$H(f) = \frac{1 + \sum_{l=1}^{q} \beta_l e^{-i2\pi f l}}{1 + \sum_{k=1}^{p} \alpha_k e^{-i2\pi f k}}, \quad |f| < \frac{1}{2}$$

  Therefore, the PSD of Y_n is $S_Y(f) = |H(f)|^2 N$ for $|f| < 1/2$

- Moving average (MA) process of order q: Let $\alpha_1 = \cdots = \alpha_p = 0$, then Y_n is simply a weighted sum of the $q+1$ most recent X_n samples with weights $(1, \beta_1, \ldots, \beta_q)$, i.e.,

$$Y_n = \sum_{l=0}^{q} \beta_l X_{n-l}, \text{ and the transfer function of the LTI system is}$$

$$H(f) = 1 + \sum_{l=1}^{q} \beta_l e^{-i2\pi f l}, \quad |f| < \frac{1}{2}$$

- Communication channel with intersymbol interference: The X_n process represents the transmitted information symbols and $1, \beta_1, \ldots, \beta_q$ are the coefficients of the channel impulse response

  The process Y_n is the interference-impaired received symbols

- Autoregressive (AR) process of order p: Let $\beta_1 = \cdots = \beta_n = 0$. Then

$$Y_n = -\sum_{k=1}^{p} \alpha_k Y_{n-k} + X_n, \text{ and the transfer function of the LTI system is}$$

$$H(f) = \frac{1}{1 + \sum_{k=1}^{p} \alpha_k e^{-i2\pi f k}}, \quad |f| < \frac{1}{2}$$

- Modeling the human speech generation process: The process X_n is generated by the vocal cords. The vocal tract is modeled as a series of coupled lossless acoustic tubes parameterized by $(\alpha_1, \ldots, \alpha_p)$

  The process Y_n is the uttered speech signal after it passes through the vocal tract

- For $p = 1$, we obtain the first-order autoregressive process

$$Y_n = -\alpha_1 Y_{n-1} + X_n,$$

$$H(f) = \frac{1}{1 + \alpha_1 e^{-i2\pi f}}, \quad |f| < \frac{1}{2}$$

$$h(n) = (-\alpha_1)^n \, u(n)$$

  This transfer function is stable iff $\sum_{n=-\infty}^{\infty} |h(n)| < \infty$, i.e., iff $|\alpha_1| < 1$

  If X_n is Gaussian, we obtain a stationary version of the Gauss–Markov process discussed in Lecture Notes 6 with $\alpha = -\alpha_1$

Sampling Theorem for Bandlimited WSS Processes

- Recall the Nyquist sampling theorem for bandlimited deterministic signals:
  - Let $x(t)$ be a signal with Fourier transform $X(f)$ such that $X(f) = 0$ for $f \notin [-B, B]$
  - We sample the signal at rate $1/T$ to obtain the sampled signal

    $$y_n = x(nT) \quad \text{for } n = \ldots, -2, -1, 0, 1, 2, \ldots$$

    The Fourier transform of the sequence y_n,

    $$Y(f) = \sum_{n=-\infty}^{\infty} X(f - n/T),$$

    is periodic with period $1/T$
  - To recover the signal, we pass y_n through an ideal low pass filter of bandwidth $1/T$. The Fourier transform of the reconstructed signal is

    $$\hat{X}(f) = Y(f) \cdot \sqcap(fT)$$

  - Hence if the sampling rate $1/T \geq 2B$, $\hat{X}(f) = X(f)$ and the signal can be reconstructed perfectly from its samples

- It turns out that a similar result holds for sampling of bandlimited WSS random processes
- Sampling theorem for WSS processes:
  - Let $X(t)$ be a continuous time WSS process with zero mean and autocorrelation function $R_X(\tau)$ and PSD $S_X(f) = 0$ for $f \notin [-B, B]$
  - We sample $X(t)$ at rate $1/T$ to obtain the sampled (discrete time) process $Y_n = X(nT)$ with

$$\mu_Y(n) = \mathrm{E}(Y_n) = \mathrm{E}(X(nT)) = 0,$$
$$R_Y(n_1, n_2) = \mathrm{E}(Y_{n_1} Y_{n_2}) = \mathrm{E}(X(n_1 T) X(n_2 T)) = R_X((n_1 - n_2)T)$$

Hence Y_n is WSS with zero mean and autocorrelation function

$$R_Y(n) = R_X(nT)$$

The PSD of Y_n,

$$S_Y(f) = \sum_{n=-\infty}^{\infty} S_X(f - n/T),$$

is periodic with period $1/T$

- As for the deterministic signal case, to reconstruct the RP $X(t)$, we pass Y_n through an ideal low pass filter

  The resulting reconstruction process $\hat{X}(t)$ is WSS with PSD

$$S_{\hat{X}}(f) = S_Y(f)|\sqcap(fT)|^2$$

- Hence if the sampling rate $1/T \geq 2B$, $S_{\hat{X}}(f) = S_X(f)$

- We show that this implies that the reconstruction process $\hat{X}(t) = X(t)$ for every t with probability one. Specifically, we show that if $1/T \geq 2B$,

$$\mathrm{E}\left[(X(t) - \hat{X}(t))^2\right] = 0 \quad \text{for every } t$$

- Proof: We know that if $1/T \geq 2B$, $S_{\hat{X}}(f) = S_X(f)$, which implies that $R_{\hat{X}}(\tau) = R_X(\tau)$

  Moreover,

$$R_{\hat{X}X}(\tau) = \operatorname{sinc}\left(\frac{\tau}{T}\right) * R_X(\tau)$$

  Now, consider

$$\mathrm{E}\left[(X(t) - \hat{X}(t))^2\right] = R_X(0) + R_{\hat{X}}(0) - 2R_{\hat{X}X}(0) = 2R_X(0) - 2R_X(0) = 0$$

  Hence, $\hat{X}(t) = X(t)$ w.p.1 for every t

Process Linear Estimation

- 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 $t - a \le \alpha \le t + b$ $(-a \le b)$ and wish to find the MMSE linear estimate of the signal $X(t)$, i.e., $\hat{X}(t)$ such that the $\mathrm{MSE} = \mathrm{E}\left[(X(t) - \hat{X}(t))^2\right]$ 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

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

or

$$\mathrm{E}\left[(X(t) - \hat{X}(t)) Y(t-\tau)\right] = 0, \quad -b \le \tau \le a$$

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

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

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 $a, b \to \infty$
  - Filtering: when $a \to \infty$ and $b = 0$ (Wiener–Hopf equations)

Infinite Smoothing Filter

- When $a, b \to \infty$, 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

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

To evaluate the second term, consider

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

Therefore

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

and the minimum MSE is

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

- Example (Additive White Noise Channel): Let $X(t)$ and $Z(t)$ be zero mean uncorrelated WSS processes with

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

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

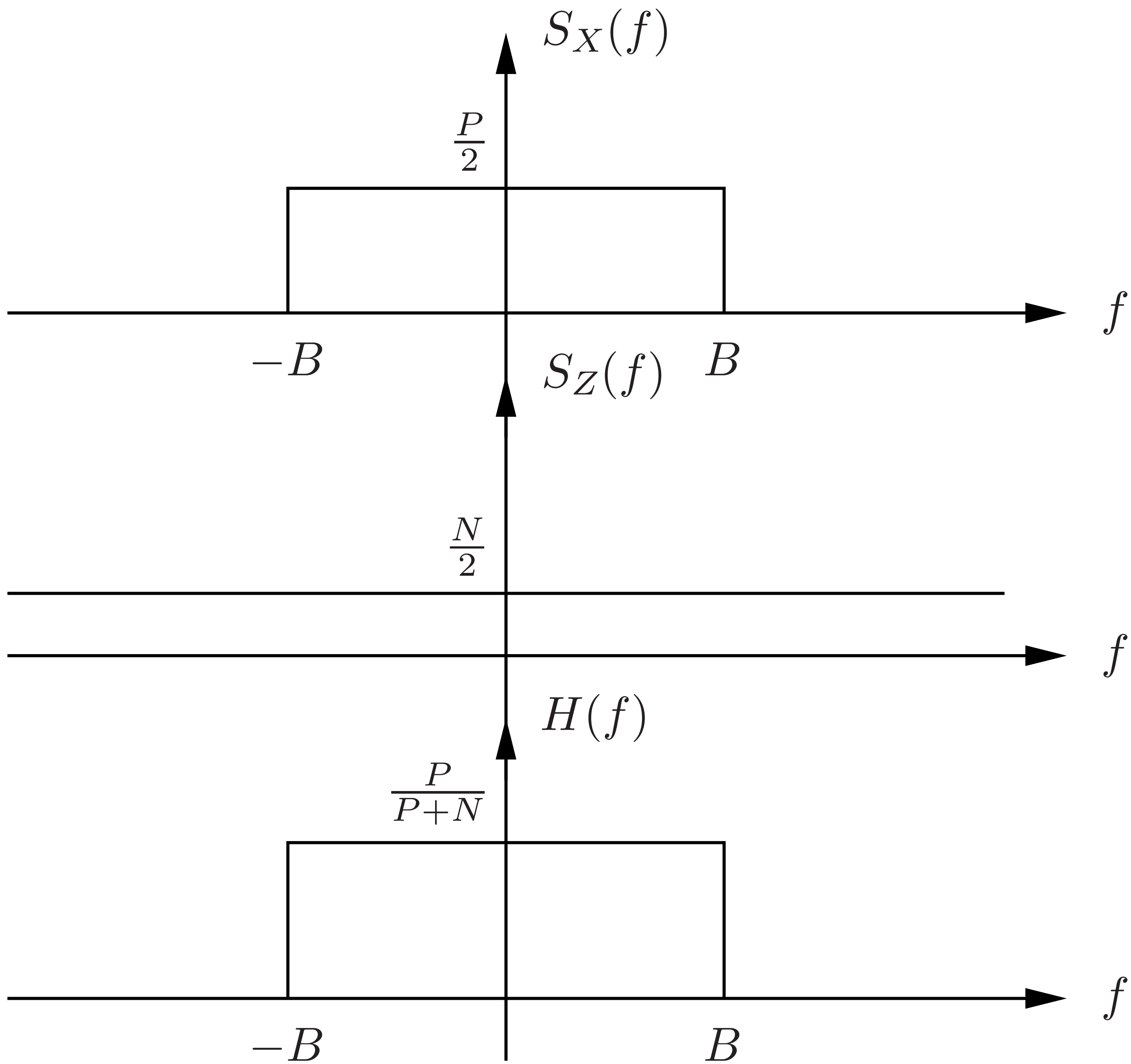
Here the signal X is bandlimited white noise, and Z is white noise

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

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

The MMSE is given by

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

Spectral Factorization

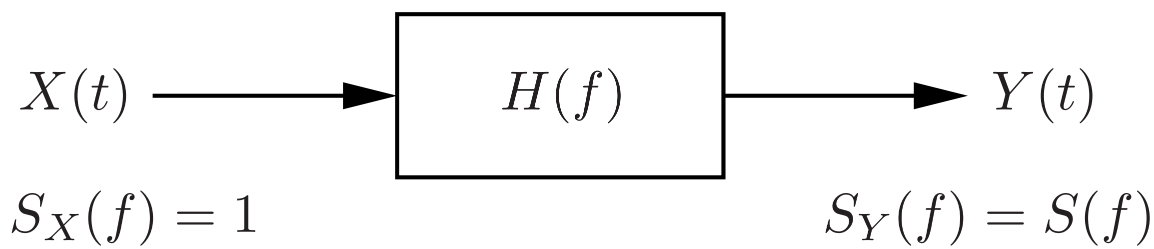
- It can be shown that the power spectral density $S_X(f)$ of WSS process $X(t)$ has a square root, i.e., a transfer function $H(f)$ such that

$$S_X(f) = H(f)H^*(f) = |H(f)|^2$$

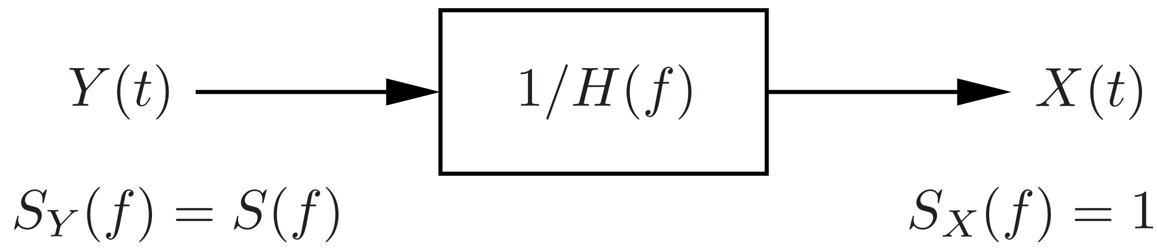
  This is similar to the square root of a covariance (correlation) matrix for a random vector discussed in Lecture notes 4

- As for the random vector case, the square root of a PSD, $H(f)$, and its inverse $1/H(f)$ can be used for coloring and whitening of WSS processes, e.g.,

  ◦ Coloring:



◦ Whitening:



Here $X(t)$ is the innovation process of $Y(t)$

- It turns out that under certain conditions, the PSD $S(f)$ of a WSS process has a causal square root, that is, $S^+(f)$ such that $S(f) = S^+(f)S^-(f)$, where $S^-(f) = (S^+(f))^*$ is an anticausal filter (note the similarity to the square root for correlation matrix via Cholesky decomposition)

- In particular, if $S(f)$ is a rational PSD for a continuous time WSS process, i.e.,

$$S(f) = c\,\frac{(2\pi i f + a_1)(2\pi i f + a_2)\ldots(2\pi i f + a_m)}{(2\pi i f + b_1)(2\pi i f + b_2)\ldots(2\pi i f + b_n)},$$

  then it can be factorized into product of causal and anticausal square roots

  Proof: Since $S(f)$ is real and nonnegative, $S^*(f) = S(f)$, if the denominator has factor $(2\pi i f + b)$, $\text{Re}(b) > 0$, then it must have factor $(-2\pi i f + b^*)$. Similarly, if numerator has factor $(2\pi i f + a)$, $\text{Re}(a) > 0$, then it must have factor $(-2\pi i f + a^*)$

Then we can express any rational PSD as $S(f) = S^+(f)S^-(f)$, where $S^+(f)$ is a causal square root that consists of the f factors and $S^-(f)$ is an anti-causal square root consisting of the $-f$ factors

- Example: Consider the PSD

$$S(f) = \frac{4\pi^2 f^2 + 3}{4\pi^2 f^2 + 1}$$

The causal square root of $S(f)$ is

$$S^+(f) = \frac{i2\pi f + \sqrt{3}}{i2\pi f + 1} \text{ and } S^-(f) = \frac{-i2\pi f + \sqrt{3}}{-i2\pi f + 1}$$

The corresponding impulse responses are:

$$h^+(t) = \delta(t) + (\sqrt{3} - 1)e^{-t}u(t), \;\; h^-(t) = \delta(t) + (\sqrt{3} - 1)e^{t}u(-t)$$

- For a discrete time WSS process a rational PSD is of the form

$$S(f) = c\,\frac{(a_1 - e^{-i2\pi f})(a_1^* - e^{i2\pi f})\ldots(a_m - e^{-i2\pi f})(a_m^* - e^{i2\pi f})}{(b_1 - e^{-i2\pi f})(b_1^* - e^{i2\pi f})\ldots(b_m - e^{-i2\pi f})(b_m^* - e^{i2\pi f})}$$

and can be expressed also as $S(f) = S^+(f)S^-(f)$ (the $e^{-i2\pi f}$ terms are causal and the $e^{i2\pi f}$ terms are noncausal)

- Example: Consider the PSD for a discrete time process

$$S(f) = \frac{3}{5 - 4\cos(2\pi f)}$$

The causal square root is

$$S^+(f) = \frac{\sqrt{3}}{2 - e^{-i2\pi f}} \text{ and } S^-(f) = \frac{\sqrt{3}}{2 - e^{i2\pi f}}$$

Spectral factorization theorem: In general, a PSD $S(f)$ has a causal square root if it satisfies the Paley-Wiener condition

$$\int_{-\infty}^{\infty} \log \frac{S(f)}{1 + 4\pi^2 f} \, df > -\infty \quad \text{for continuous time process}$$

$$\int_{-1/2}^{1/2} \log S(f) \, df > -\infty \quad \text{for discrete time process}$$

- These condition are not always satisfied. For example they are not satisfied for bandlimited processes

- Remark: We assume throughout that $\mathcal{F}^{-1}[S^+(f)]$ is real; hence $S^-(f) = S^+(-f)$

Wiener Filter

- Again let $X(t)$ and $Y(t)$ be jointly WSS random processes. consider the linear estimation of process $X(t)$ from observations $Y(\alpha)$, $t-a \le \alpha \le t+b$

- When $a \to \infty$ and $b = 0$, the equations for the MMSE linear estimate, called Wiener–Hopf equations, are

$$R_{XY}(\tau) = \int_0^\infty h(\alpha) R_Y(\tau - \alpha)\, d\alpha\,, \quad 0 \le \tau < \infty$$
$$= \int_{-\infty}^\infty h(\alpha) R_Y(\tau - \alpha)\, d\alpha\,, \quad 0 \le \tau < \infty$$

  where $h(t)$ is a causal impulse response

- Notation: A real-valued function $h(t)$ can be expressed as

$$h(t) = [h(t)]_+ + [h(t)]_-,$$

  where $[h(t)]_+ = h(t)$ for $t \ge 0$ and $[h(t)]_+ = 0$ for $t < 0$ is the positive (causal) part of $h(t)$, and $[h(t)]_- = h(t) - [h(t)]_+$ is the negative (anticausal) part

Taking the Fourier transform, we have

$$H(f) = [H(f)]_+ + [H(f)]_-,$$

where $[H(f)]_+$ and $[H(f)]_-$ are the FT of the positive and negative parts of $h(t)$, respectively

Example: Let

$$S(f) = \frac{4\pi^2 f^2 + 3}{4\pi^2 f^2 + 1}$$

We can write

$$S(f) = \frac{i2\pi f + 2}{i2\pi f + 1} + \frac{1}{-i2\pi f + 1}$$

The first term is $[S(f)]_+$ and the second is $[S(f)]_-$. The corresponding impulse responses are

$$[R(t)]_+ = \delta(t) + e^{-t}u(t)$$

$$[R(t)]_- = e^{t}u(-t)$$

Compare to the causal square root factors

- Now, back to the linear estimation problem. First assume that the observation process $Y(\tau)$ is white, i.e., $R_Y(\tau) = \delta(\tau)$, then the Wiener–Hopf equations reduce to

$$R_{XY}(\tau) = h(\tau)\,, \quad 0 \le \tau < \infty,$$

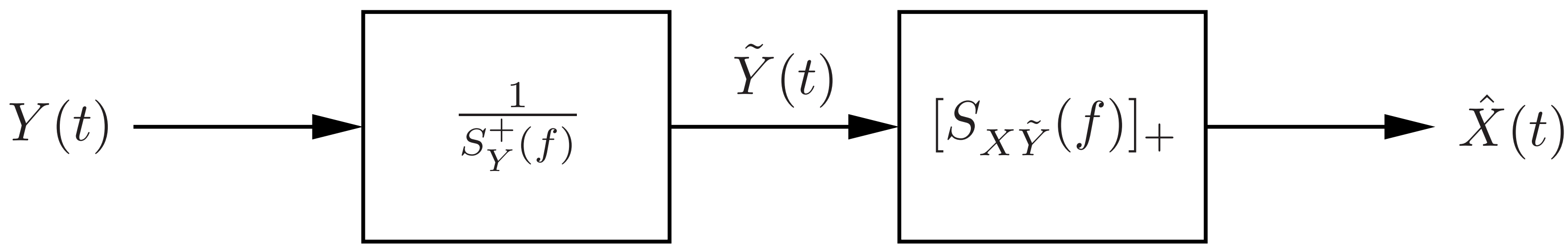
i.e., $h(\tau) = [R_{XY}(\tau)]_+$

and the corresponding transfer function is

$$H(f) = \int_0^\infty R_{XY}(\tau) e^{-2\pi i f \tau}\, d\tau,$$

i.e., $H(f) = [S_{XY}(f)]_+$

- For general $S_Y(f)$ with causal square root $S_Y^+(f)$, we first whiten the process to obtain $\tilde{Y}(\tau)$ with $R_{\tilde{Y}}(\tau) = \delta(\tau)$, then convolve with $[R_{X\tilde{Y}}(\tau)]_+$



- Now to find $R_{X\tilde{Y}}(\tau)$, let $g(t) = \mathcal{F}^{-1}[1/S_Y^+(f)]$ and consider

$$\begin{aligned} R_{X\tilde{Y}}(\tau) &= \mathrm{E}(X(t+\tau)\tilde{Y}(t)) \\ &= \mathrm{E}(X(t+\tau)Y(t) * g(t)) \\ &= R_{XY}(\tau) * g(-\tau) \end{aligned}$$

  Taking the Fourier Transform we have

$$S_{X\tilde{Y}}(f) = \frac{S_{XY}(f)}{S_Y^-(f)}$$

  Hence,

$$[S_{X\tilde{Y}}(f)]_+ = \left[\frac{S_{XY}(f)}{S_Y^-(f)}\right]_+$$

  The transfer function of the Wiener filter is then given by

$$H(f) = \frac{1}{S_Y^+(f)} \left[\frac{S_{XY}(f)}{S_Y^-(f)}\right]_+$$

- To find the MMSE, we follow similar steps to the infinite smoothing case to obtain

$$\text{MSE} = \int_{-\infty}^{\infty} \left(S_X(f) - \left| \left[\frac{S_{XY}(f)}{S_Y^-(f)} \right]_+ \right|^2 \right) df$$

- Example: Consider a continuous-time RP $X(t)$ with

$$S_X(f) = \frac{2}{1 + 4\pi^2 f^2}$$

and the noisy observation $Y(t) = X(t) + Z(t)$, where $Z(t)$ is white noise uncorrelated with $X(t)$ with $S_Z(f) = 1$

To compute the Wiener filter, we first factor the PSD

$$S_Y(f) = S_X(f) + S_Z(f) = \frac{4\pi^2 f^2 + 3}{4\pi^2 f^2 + 1}$$

to obtain

$$S_Y^+(f) = \frac{i2\pi f + \sqrt{3}}{i2\pi f + 1},$$

$$S_Y^-(f) = \frac{-i2\pi f + \sqrt{3}}{-i2\pi f + 1}$$

The crosspower spectral density $S_{XY}(f) = S_X(f)$, hence

$$\frac{S_{XY}(f)}{S_Y^-(f)} = \frac{2}{1+4\pi^2 f^2} \cdot \frac{-i2\pi f + 1}{-i2\pi f + \sqrt{3}}$$

$$= \frac{2}{(i2\pi f + 1)(-i2\pi f + \sqrt{3})}$$

$$= \frac{\sqrt{3}-1}{i2\pi f + 1} + \frac{\sqrt{3}-1}{-i2\pi f + \sqrt{3}}$$

The first term is causal and the second term is anticausal. Therefore,

$$\left[\frac{S_{XY}(f)}{S_Y^-(f)}\right]_+ = \frac{\sqrt{3}-1}{i2\pi f + 1}$$

Hence, the Wiener filter is

$$H(f) = \frac{\sqrt{3}-1}{i2\pi f + \sqrt{3}}$$

$$h(t) = (\sqrt{3}-1)\; e^{-\sqrt{3}t}\; u(t)$$

The MSE is

$$\text{MSE} = \int_{-\infty}^{\infty} \frac{2\sqrt{3}-2}{1+4\pi^2 f^2}\, df = \sqrt{3}-1$$

- Example: Consider a discrete-time RP $X(t)$ with

$$S_X(f) = \frac{3}{5-4\cos(2\pi f)}$$

and the noisy observation $Y(t) = X(t) + Z(t)$, where $Z(t)$ is white noise, independent of $X(t)$, with $S_Z(f) = 1$

Again we factor the PSD

$$S_Y(f) = S_X(f) + S_Z(f) = \frac{8-4\cos(2\pi f)}{5-4\cos(2\pi f)}$$

to obtain

$$S_Y^+(f) = \sqrt{4-2\sqrt{3}} \cdot \frac{2+\sqrt{3}-e^{-i2\pi f}}{2-e^{-i2\pi f}}$$

$$S_Y^-(f) = \sqrt{4-2\sqrt{3}} \cdot \frac{2+\sqrt{3}-e^{i2\pi f}}{2-e^{i2\pi f}}$$

The crosspower spectral density is $S_{XY}(f) = S_X(f)$ and

$$\frac{S_{XY}(f)}{S_Y^-(f)} = \frac{3}{\sqrt{4-2\sqrt{3}}} \cdot \frac{1}{(2-e^{-i2\pi f})(2+\sqrt{3}-e^{-i2\pi f})}$$

$$= \frac{\sqrt{12-6\sqrt{3}}}{2-e^{-i2\pi f}} + \frac{\sqrt{3-3\sqrt{3}/2}}{2+\sqrt{3}-e^{i2\pi f}}$$

The first term is causal and the second term is anticausal. Therefore

$$\left[\frac{S_{XY}(f)}{S_Y^-(f)}\right]_+ = \frac{\sqrt{12-6\sqrt{3}}}{2-e^{-i2\pi f}}$$

Hence, the Wiener filter is

$$H(f) = \frac{\sqrt{3}}{2+\sqrt{3}+e^{-i2\pi f}}$$

$$h(n) = (2\sqrt{3}-3)\ (2-\sqrt{3})^n\ u(n)$$

Wiener Filter Versus Kalman Filter

- Both the Kalman and Wiener filters are MMSE linear estimates of a process from causal observations
- There are several differences, however

Kalman filter	Wiener filter
X_n, Y_n state space model	X_n, Y_n jointly WSS
not necessarily WSS	
time domain filter	frequency domain filter
recursive	non-recursive

- Remarks:
  - A continuous time counterpart to the Kalman filter exists and is known as the Kalman-Bucy filter
  - If X_n, Y_n are jointly WSS and can be described via a state space model, then the Kalman filter gives a recursive way to compute the Wiener filter