# Stationary Processes

Gonzalo Mateos
Dept. of ECE and Goergen Institute for Data Science
University of Rochester
gmateosb@ece.rochester.edu
http://www.ece.rochester.edu/~gmateosb/

November 30, 2018

# Stationary random processes

# Stationary random processes

- All joint probabilities invariant to time shifts, i.e., for any $s$

$$\mathrm{P}\left(X(t_1+s)\le x_1, X(t_2+s)\le x_2,\ldots,X(t_n+s)\le x_n\right) = \mathrm{P}\left(X(t_1)\le x_1, X(t_2)\le x_2,\ldots,X(t_n)\le x_n\right)$$

$\Rightarrow$ If above relation holds $X(t)$ is called strictly stationary (SS)

- First-order stationary $\Rightarrow$ probs. of single variables are shift invariant

$$\mathrm{P}\left(X(t+s)\le x\right) = \mathrm{P}\left(X(t)\le x\right)$$

- Second-order stationary $\Rightarrow$ joint probs. of pairs are shift invariant

$$\mathrm{P}\left(X(t_1+s)\le x_1, X(t_2+s)\le x_2\right) = \mathrm{P}\left(X(t_1)\le x_1, X(t_2)\le x_2\right)$$

# Pdfs and moments of stationary processes

- For SS process joint cdfs are shift invariant. Hence, pdfs also are

$$f_{X(t+s)}(x) = f_{X(t)}(x) = f_{X(0)}(x) := f_X(x)$$

- As a consequence, the mean of a SS process is constant

$$\mu(t) := \mathbb{E}\left[X(t)\right] = \int_{-\infty}^{\infty} x f_{X(t)}(x)dx = \int_{-\infty}^{\infty} x f_X(x)dx = \mu$$

- The variance of a SS process is also constant

$$\text{var}\left[X(t)\right] := \int_{-\infty}^{\infty} (x-\mu)^2 f_{X(t)}(x)dx = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)dx = \sigma^2$$

- The power (second moment) of a SS process is also constant

$$\mathbb{E}\left[X^2(t)\right] := \int_{-\infty}^{\infty} x^2 f_{X(t)}(x)dx = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \sigma^2 + \mu^2$$

# Joint pdfs of stationary processes

- Joint pdf of two values of a SS random process

$$f_{X(t_1)X(t_2)}(x_1, x_2) = f_{X(0)X(t_2-t_1)}(x_1, x_2)$$

  $\Rightarrow$ Used shift invariance for shift of $t_1$

  $\Rightarrow$ Note that $t_1 = 0 + t_1$ and $t_2 = (t_2 - t_1) + t_1$

- Result above true for any pair $t_1$, $t_2$

  $\Rightarrow$ Joint pdf depends only on time difference $s := t_2 - t_1$

- Writing $t_1 = t$ and $t_2 = t + s$ we equivalently have

$$f_{X(t)X(t+s)}(x_1, x_2) = f_{X(0)X(s)}(x_1, x_2) = f_X(x_1, x_2; s)$$

# Stationary processes and limit distributions

- Stationary processes follow the footsteps of limit distributions

- For Markov processes limit distributions exist under mild conditions
  - Limit distributions also exist for some non-Markov processes

- Process somewhat easier to analyze in the limit as $t \to \infty$

  $\Rightarrow$ Properties can be derived from the limit distribution

- Stationary process $\approx$ study of limit distribution

  $\Rightarrow$ Formally initialize at limit distribution

  $\Rightarrow$ In practice results true for time sufficiently large

- Deterministic linear systems $\Rightarrow$ transient + steady-state behavior

  $\Rightarrow$ Stationary systems akin to the study of steady-state

- But steady-state is in a probabilistic sense (probs., not realizations)

# Autocorrelation and wide-sense stationarity

Stationary random processes

Autocorrelation function and wide-sense stationary processes

Fourier transforms

Linear time-invariant systems

Power spectral density and linear filtering of random processes

The matched and Wiener filters

# Autocorrelation function

- From the definition of autocorrelation function we can write

$$R_X(t_1, t_2) = \mathbb{E}\left[X(t_1)X(t_2)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1)X(t_2)}(x_1, x_2)\, dx_1 dx_2$$

- For SS process $f_{X(t_1)X(t_2)}(\cdot)$ depends on time difference only

$$R_X(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X(0)X(t_2-t_1)}(x_1, x_2)\, dx_1 dx_2 = \mathbb{E}\left[X(0)X(t_2-t_1)\right]$$

$\Rightarrow$ $R_X(t_1, t_2)$ is a function of $s = t_2 - t_1$ only

$$R_X(t_1, t_2) = R_X(0, t_2 - t_1) := R_X(s)$$

- The autocorrelation function of a SS random process $X(t)$ is $R_X(s)$
  - $\Rightarrow$ Variable $s$ denotes a time difference / shift / lag
  - $\Rightarrow$ $R_X(s)$ specifies correlation between values $X(t)$ spaced $s$ in time

# Autocovariance function

- Similarly to autocorrelation, define the autocovariance function as

$$C_X(t_1, t_2) = \mathbb{E}\left[\big(X(t_1) - \mu(t_1)\big)\big(X(t_2) - \mu(t_2)\big)\right]$$

- Expand product to write $C_X(t_1, t_2)$ as

$$C_X(t_1, t_2) = \mathbb{E}\left[X(t_1)X(t_2)\right] + \mu(t_1)\mu(t_2) - \mathbb{E}\left[X(t_1)\right]\mu(t_2) - \mathbb{E}\left[X(t_2)\right]\mu(t_1)$$

- For SS process $\mu(t_1) = \mu(t_2) = \mu$ and $\mathbb{E}\left[X(t_1)X(t_2)\right] = R_X(t_2 - t_1)$

$$C_X(t_1, t_2) = R_X(t_2 - t_1) - \mu^2 = C_X(t_2 - t_1)$$

$\Rightarrow$ Autocovariance function depends only on the shift $s = t_2 - t_1$

- We will typically assume that $\mu = 0$ in which case

$$R_X(s) = C_X(s)$$

$\Rightarrow$ If $\mu \neq 0$ can study process $X(t) - \mu$ whose mean is null

# Wide-sense stationary processes

- **Def:** A process is wide-sense stationary (WSS) when its
  - $\Rightarrow$ Mean is constant $\Rightarrow \mu(t) = \mu$ for all $t$
  - $\Rightarrow$ Autocorrelation is shift invariant $\Rightarrow R_X(t_1, t_2) = R_X(t_2 - t_1)$
- Consequently, autocovariance of WSS process is also shift invariant

$$
\begin{aligned}
C_X(t_1, t_2) &= \mathbb{E}\left[X(t_1)X(t_2)\right] + \mu(t_1)\mu(t_2) - \mathbb{E}\left[X(t_1)\right]\mu(t_2) - \mathbb{E}\left[X(t_2)\right]\mu(t_1) \\
&= R_X(t_2 - t_1) - \mu^2
\end{aligned}
$$

- Most of the analysis of stationary processes is based on $R_X(t_2 - t_1)$
  - $\Rightarrow$ Thus, such analysis does not require SS, WSS suffices

# Wide-sense stationarity versus strict stationarity

- SS processes have shift-invariant pdfs
  - $\Rightarrow$ Mean function is constant
  - $\Rightarrow$ Autocorrelation is shift-invariant
- Then, a SS process is also WSS
  - $\Rightarrow$ For that reason WSS is also called weak-sense stationary
- The opposite is obviously not true in general
- But if Gaussian, process determined by mean and autocorrelation
  - $\Rightarrow$ WSS implies SS for Gaussian process
- WSS and SS are equivalent for Gaussian processes (More coming)

# Gaussian wide-sense stationary process

- WSS Gaussian process $X(t)$ with mean 0 and autocorrelation $R(s)$
- The covariance matrix for $X(t_1+s), X(t_2+s), \ldots, X(t_n+s)$ is

$$\mathbf{C}(t_1+s,\ldots,t_n+s) = \begin{pmatrix} R(t_1+s, t_1+s) & R(t_1+s, t_2+s) & \ldots & R(t_1+s, t_n+s) \\ R(t_2+s, t_1+s) & R(t_2+s, t_2+s) & \ldots & R(t_2+s, t_n+s) \\ \vdots & \vdots & \ddots & \vdots \\ R(t_n+s, t_1+s) & R(t_n+s, t_2+s) & \ldots & R(t_n+s, t_n+s) \end{pmatrix}$$

- For WSS process, autocorrelations depend only on time differences

$$\color{red}{\mathbf{C}(t_1+s,\ldots,t_n+s)} = \begin{pmatrix} R(t_1-t_1) & R(t_2-t_1) & \ldots & R(t_n-t_1) \\ R(t_1-t_2) & R(t_2-t_2) & \ldots & R(t_n-t_2) \\ \vdots & \vdots & \ddots & \vdots \\ R(t_1-t_n) & R(t_2-t_n) & \ldots & R(t_n-t_n) \end{pmatrix} = \color{red}{\mathbf{C}(t_1,\ldots,t_n)}$$

$\Rightarrow$ Covariance matrices $\mathbf{C}(t_1,\ldots,t_n)$ are shift invariant

# Gaussian wide-sense stationary process (continued)

- The joint pdf of $X(t_1+s), X(t_2+s), \ldots, X(t_n+s)$ is

$$f_{X(t_1+s),\ldots,X(t_n+s)}(x_1,\ldots,x_n) = \mathcal{N}(\mathbf{0}, \mathbf{C}(t_1+s,\ldots,t_n+s); [x_1,\ldots,x_n]^T)$$

  $\Rightarrow$ Completely determined by $\mathbf{C}(t_1+s,\ldots,t_n+s)$

- Since covariance matrix is shift invariant can write

$$f_{X(t_1+s),\ldots,X(t_n+s)}(x_1,\ldots,x_n) = \mathcal{N}(\mathbf{0}, \mathbf{C}(t_1,\ldots,t_n); [x_1,\ldots,x_n]^T)$$

- Expression on the right is the pdf of $X(t_1), X(t_2), \ldots, X(t_n)$. Then

$$f_{X(t_1+s),\ldots,X(t_n+s)}(x_1,\ldots,x_n) = f_{X(t_1),\ldots,X(t_n)}(x_1,\ldots,x_n)$$

- Joint pdf of $X(t_1), X(t_2), \ldots, X(t_n)$ is shift invariant

  $\Rightarrow$ Proving that WSS is equivalent to SS for Gaussian processes

# Brownian motion and white Gaussian noise

Ex: Brownian motion $X(t)$ with variance parameter $\sigma^2$

$\Rightarrow$ Mean function is $\mu(t) = 0$ for all $t \geq 0$

$\Rightarrow$ Autocorrelation is $R_X(t_1, t_2) = \sigma^2 \min(t_1, t_2)$

- While the mean is constant, autocorrelation is not shift invariant

$\Rightarrow$ Brownian motion is not WSS (hence not SS)

Ex: White Gaussian noise $W(t)$ with variance parameter $\sigma^2$

$\Rightarrow$ Mean function is $\mu(t) = 0$ for all $t$

$\Rightarrow$ Autocorrelation is $R_W(t_1, t_2) = \sigma^2 \delta(t_2 - t_1)$

- The mean is constant and the autocorrelation is shift invariant

$\Rightarrow$ White Gaussian noise is WSS

$\Rightarrow$ Also SS because white Gaussian noise is a GP

# Properties of autocorrelation function

**For WSS processes:**

**(i)** The autocorrelation for $s = 0$ is the power of the process

$$R_X(0) = \mathbb{E}\left[X^2(t)\right] = \mathbb{E}\left[X(t)X(t+0)\right]$$

**(ii)** The autocorrelation function is symmetric $\Rightarrow R_X(s) = R_X(-s)$

Proof.
Commutative property of product and shift invariance of $R_X(t_1, t_2)$

$$\begin{aligned} R_X(s) &= R_X(t, t+s) \\ &= \mathbb{E}\left[X(t)X(t+s)\right] \\ &= \mathbb{E}\left[X(t+s)X(t)\right] \\ &= R_X(t+s, t) = R_X(-s) \end{aligned}$$

$\square$

# Properties of autocorrelation function (continued)

**For WSS processes:**

**(iii)** Maximum absolute value of the autocorrelation function is for $s = 0$

$$\left|R_X(s)\right| \leq R_X(0)$$

Proof.
Expand the square $\mathbb{E}\left[\left(X(t+s) \pm X(t)\right)^2\right]$

$$\mathbb{E}\left[\left(X(t+s) \pm X(t)\right)^2\right] = \mathbb{E}\left[X^2(t+s)\right] + \mathbb{E}\left[X^2(t)\right] \pm 2\mathbb{E}\left[X(t+s)X(t)\right]$$
$$= R_X(0) + R_X(0) \pm 2R_X(s)$$

Square $\mathbb{E}\left[\left(X(t+s) \pm X(t)\right)^2\right]$ is always nonnegative, then

$$0 \leq \mathbb{E}\left[\left(X(t+s) \pm X(t)\right)^2\right] = 2R_X(0) \pm 2R_X(s)$$

Rearranging terms $\Rightarrow R_X(0) \geq \mp R_X(s)$

□

# Fourier transforms

Stationary random processes

Autocorrelation function and wide-sense stationary processes

Fourier transforms

Linear time-invariant systems

Power spectral density and linear filtering of random processes

The matched and Wiener filters

# Definition of Fourier transform

- **Def:** The Fourier transform of a function (signal) $x(t)$ is

$$X(f) = \mathcal{F}\big(x(t)\big) := \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\, dt$$

- The complex exponential is (recall $j^2 = -1$)

$$\begin{aligned} e^{-j2\pi ft} &= \cos(-2\pi ft) + j\sin(-2\pi ft) \\ &= \cos(2\pi ft) - j\sin(2\pi ft) \\ &= 1\angle - 2\pi ft \end{aligned}$$

- The Fourier transform is complex valued
  - $\Rightarrow$ It has a real and a imaginary part (rectangular coordinates)
  - $\Rightarrow$ It has a magnitude and a phase (polar coordinates)
- Argument $f$ of $X(f)$ is referred to as frequency

# Examples

Ex: Fourier transform of a constant $x(t) = c$

$$\mathcal{F}(c) = \int_{-\infty}^{\infty} c e^{-j2\pi ft}\, dt = c\delta(f)$$

Ex: Fourier transform of scaled delta function $x(t) = c\delta(t)$

$$\mathcal{F}(c\delta(t)) = \int_{-\infty}^{\infty} c\delta(t) e^{-j2\pi ft}\, dt = c$$

Ex: For a complex exponential $x(t) = e^{j2\pi f_0 t}$ with frequency $f_0$ we have

$$\mathcal{F}\left(e^{j2\pi f_0 t}\right) = \int_{-\infty}^{\infty} e^{j2\pi f_0 t} e^{-j2\pi ft}\, dt = \int_{-\infty}^{\infty} e^{-j2\pi (f-f_0)t}\, dt = \delta(f - f_0)$$

Ex: For a shifted delta $\delta(t - t_0)$ we have

$$\mathcal{F}(\delta(t - t_0)) = \int_{-\infty}^{\infty} \delta(t - t_0) e^{-j2\pi ft}\, dt = e^{-j2\pi f t_0}$$

$\Rightarrow$ Note the symmetry (duality) in the first two and last two transforms

# Fourier transform of a cosine

Ex: Fourier transform of a cosine $x(t) = \cos(2\pi f_0 t)$

- Begin noticing that we may write $\cos(2\pi f_0 t) = \frac{1}{2}e^{j2\pi f_0 t} + \frac{1}{2}e^{-j2\pi f_0 t}$
- Fourier transformation is a linear operation (integral), then

$$\mathcal{F}\big(\cos(2\pi f_0 t)\big) = \int_{-\infty}^{\infty} \left(\frac{1}{2}e^{j2\pi f_0 t} + \frac{1}{2}e^{-j2\pi f_0 t}\right) e^{-j2\pi f t}\, dt$$
$$= \frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)$$

$\Rightarrow$ A pair of delta functions at frequencies $f = \pm f_0$ (tones)

- Frequency of the cosine is $f_0$ $\Rightarrow$ "Justifies" the name frequency for $f$

# Inverse Fourier transform

- **Def:** The inverse Fourier transform of $X(f) = \mathcal{F}(x(t))$ is

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft}\, df$$

$\Rightarrow$ Exponent's sign changes with respect to Fourier transform

- We show next that $x(t)$ can be recovered from $X(f)$ as above
- First substitute $X(f)$ for its definition

$$\int_{-\infty}^{\infty} X(f) e^{j2\pi ft}\, df = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} x(u) e^{-j2\pi fu}\, du \right) e^{j2\pi ft}\, df$$

# Inverse Fourier transform (continued)

- Nested integral can be written as double integral

$$\int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,df = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(u)e^{-j2\pi fu}e^{j2\pi ft}\,du\,df$$

- Rewrite as nested integral with integration w.r.t. $f$ carried out first

$$\int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,df = \int_{-\infty}^{\infty} x(u)\left(\int_{-\infty}^{\infty} e^{-j2\pi f(t-u)}\,df\right)\,du$$

- Innermost integral is a delta function

$$\int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,df = \int_{-\infty}^{\infty} x(u)\delta(t-u)\,du = x(t)$$

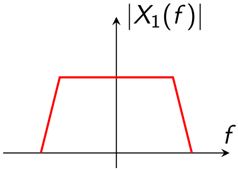
# Frequency components of a signal

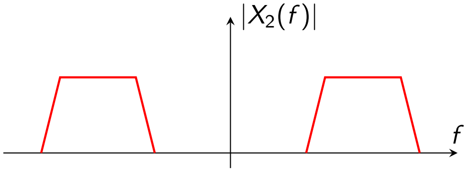
- Interpretation of Fourier transform through synthesis formula

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\, df \approx \Delta f \times \sum_{n=-\infty}^{\infty} X(f_n)e^{j2\pi f_n t}$$

$\Rightarrow$ Signal $x(t)$ as linear combination of complex exponentials

- $X(f)$ determines the weight of frequency $f$ in the signal $x(t)$




Ex: Signal on the left contains low frequencies (changes slowly in time)

Ex: Signal on the right contains high frequencies (changes fast in time)

# Linear time-invariant systems

Stationary random processes

Autocorrelation function and wide-sense stationary processes

Fourier transforms

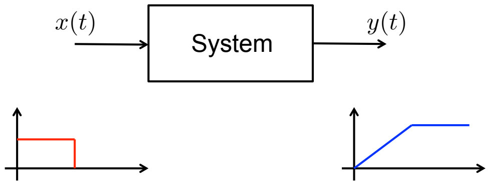
Linear time-invariant systems

Power spectral density and linear filtering of random processes

The matched and Wiener filters

# Systems

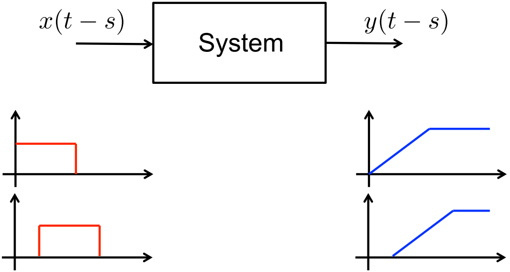
- **Def:** A system characterizes an input-output relationship
- This relation is between functions, not values
  - $\Rightarrow$ Each output value $y(t)$ depends on all input values $x(t)$
  - $\Rightarrow$ A mapping from the input signal to the output signal



# Time-invariant system

- **Def:** A system is time invariant if a delayed input yields a delayed output
- If input $x(t)$ yields output $y(t)$ then input $x(t-s)$ yields $y(t-s)$

  $\Rightarrow$ Think of output applied $s$ time units later

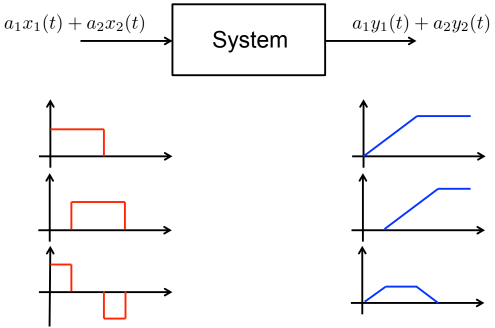


# Linear system

- **Def:** A system is linear if the output of a linear combination of inputs is the same linear combination of the respective outputs

- If input $x_1(t)$ yields output $y_1(t)$ and $x_2(t)$ yields $y_2(t)$, then

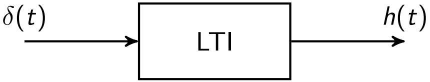
$$a_1x_1(t) + a_2x_2(t) \quad \Rightarrow \quad a_1y_1(t) + a_2y_2(t)$$



# Linear time-invariant system

- Linear + time-invariant system = linear time-invariant system (LTI)
- Denote as $h(t)$ the system's output when the input is $\delta(t)$

  $\Rightarrow$ $h(t)$ is the impulse response of the LTI system




1) Response to $\delta(t-u)$ $\Rightarrow$ $h(t-u)$ due to time invariance
2) Response to $x(u)\delta(t-u)$ $\Rightarrow$ $x(u)h(t-u)$ due to linearity
3) Reponse to $x(u_1)\delta(t-u_1) + x(u_2)\delta(t-u_2)$

$$\Rightarrow x(u_1)h(t-u_1) + x(u_2)h(t-u_2)$$

- Any function $x(t)$ can be written as

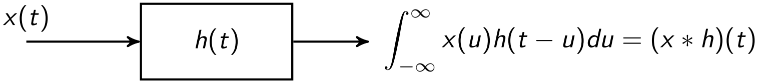
$$x(t) = \int_{-\infty}^{\infty} x(u)\delta(t-u)\, du$$

- Thus, the output of a LTI with impulse response $h(t)$ to input $x(t)$ is

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)\, du = (x * h)(t)$$

- The above integral is called the convolution of $x(t)$ and $h(t)$

  $\Rightarrow$ It is a "product" between signals, denoted as $(x * h)(t)$



# Fourier transform of output

- The Fourier transform $Y(f)$ of the output $y(t)$ is given by

$$Y(f) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} x(u)h(t-u)\,du \right) e^{-j2\pi ft}\,dt$$

- Write nested integral as double integral & change variable $t \rightarrow u+v$

$$Y(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(u)h(v)e^{-j2\pi f(u+v)}\,dv\,du$$

- Write $e^{-j2\pi f(u+v)} = e^{-j2\pi fu}e^{-j2\pi fv}$ and reorder terms to obtain

$$Y(f) = \left( \int_{-\infty}^{\infty} x(u)e^{-j2\pi fu}\,du \right) \left( \int_{-\infty}^{\infty} h(v)e^{-j2\pi fv}\,dv \right)$$

- The factors on the right are the Fourier transforms of $x(t)$ and $h(t)$

# Frequency response of linear time-invariant system

- **Def:** The frequency response of a LTI system is

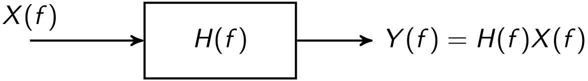
$$H(f) := \mathcal{F}(h(t)) = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}\, dt$$

$\Rightarrow$ Fourier transform of the impulse response $h(t)$

- Input signal with spectrum $X(f)$, LTI system with freq. response $H(f)$

  $\Rightarrow$ We established that the spectrum $Y(f)$ of the output is

$$Y(f) = H(f)X(f)$$



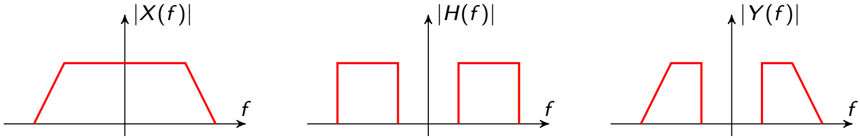
# More on frequency response

- Frequency components of input get "scaled" by $H(f)$
  - Since $H(f)$ is complex, scaling is a complex number
  - Represents a scaling part (amplitude) and a phase shift (argument)

- Effect of LTI on input easier to analyze

  $\Rightarrow$ "Usual product" instead of convolution

$$X(f) \longrightarrow \boxed{H(f)} \longrightarrow Y(f) = H(f)X(f)$$



# Power spectral density and linear filtering

Stationary random processes

Autocorrelation function and wide-sense stationary processes

Fourier transforms

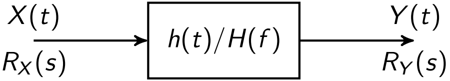
Linear time-invariant systems

Power spectral density and linear filtering of random processes

The matched and Wiener filters

# Linear filters

- Linear filter (system) with $\Rightarrow$ impulse response $h(t)$
  $\Rightarrow$ frequency response $H(f)$
- Input to filter is wide-sense stationary (WSS) random process $X(t)$
  $\Rightarrow$ Process has zero mean and autocorrelation function $R_X(s)$
- Output is obviously another random process $Y(t)$
- Describe $Y(t)$ in terms of $\Rightarrow$ properties of $X(t)$
  $\Rightarrow$ filter's impulse and/or frequency response
- Q: Is $Y(t)$ WSS? Mean of $Y(t)$? Autocorrelation function of $Y(t)$?
  $\Rightarrow$ Easier and more enlightening in the frequency domain

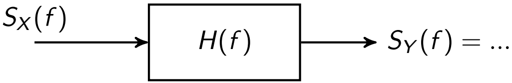


# Power spectral density

- **Def:** The power spectral density (PSD) of a WSS random process is the Fourier transform of the autocorrelation function

$$S_X(f) = \mathcal{F}\big(R_X(s)\big) = \int_{-\infty}^{\infty} R_X(s) e^{-j2\pi fs} \, ds$$

- Does $S_X(f)$ carry information about frequency components of $X(t)$?
  $\Rightarrow$ Not clear, $S_X(f)$ is Fourier transform of $R_X(s)$, not $X(t)$
- But yes. We'll see $S_X(f)$ describes spectrum of $X(t)$ in some sense
- Q: Can we relate PSDs at the input and output of a linear filter?



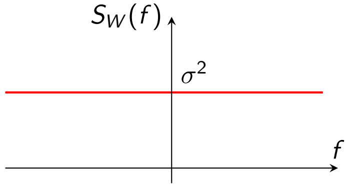
# Example: Power spectral density of white noise

- Autocorrelation of white noise $W(t)$ is $\Rightarrow R_W(s) = \sigma^2\delta(s)$
- PSD of white noise is Fourier transform of $R_W(s)$

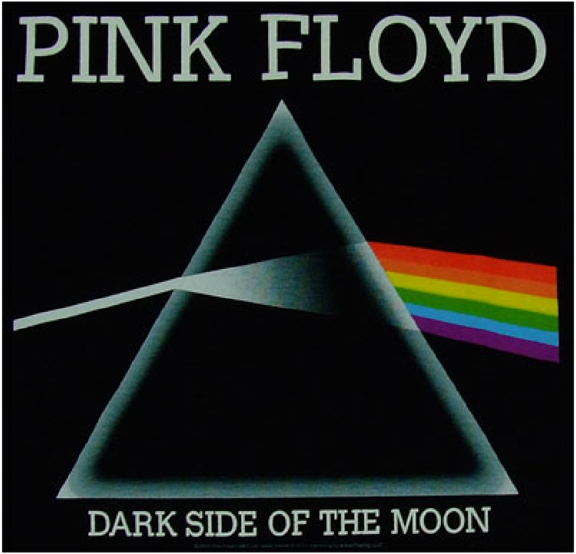
$$S_W(f) = \int_{-\infty}^{\infty} \sigma^2\delta(s)e^{-j2\pi fs}\, ds = \sigma^2$$

$\Rightarrow$ PSD of white noise is constant for all frequencies

- That's why it's white $\Rightarrow$ Contains all frequencies in equal measure



# Dark side of the moon



# Power of a process

- The power of WSS process $X(t)$ is its (constant) second moment

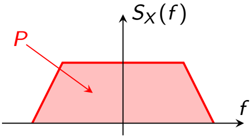
$$P = \mathbb{E}\left[X^2(t)\right] = R_X(0)$$

- Use expression for inverse Fourier transform evaluated at $t = 0$

$$R_X(s) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f s}\, df \;\Rightarrow\; R_X(0) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f 0}\, df$$

- Since $e^0 = 1$, can write $R_X(0)$ and therefore process' power as

$$P = \int_{-\infty}^{\infty} S_X(f)\, df$$




$\Rightarrow$ Area under PSD is the power of the process

# Mean of filter's output

- Q: If input $X(t)$ to a LTI filter is WSS, is output $Y(t)$ WSS as well?

  $\Rightarrow$ Check first that mean $\mu_Y(t)$ of filter's output $Y(t)$ is constant

- Recall that for any time $t$, filter's output is

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

- The mean function $\mu_Y(t)$ of the process $Y(t)$ is

$$\mu_Y(t) = \mathbb{E}\left[Y(t)\right] = \mathbb{E}\left[\int_{-\infty}^{\infty} h(u)X(t-u)\,du\right]$$

- Expectation is linear and $X(t)$ is WSS, thus

$$\mu_Y(t) = \int_{-\infty}^{\infty} h(u)\mathbb{E}\left[X(t-u)\right]\,du = \mu_X \int_{-\infty}^{\infty} h(u)\,du = \mu_Y$$

# Autocorrelation of filter's output

- Compute autocorrelation function $R_Y(t, t+s)$ of filter's output $Y(t)$

  $\Rightarrow$ Check that $R_Y(t, t+s) = R_Y(s)$, only function of $s$

- Start noting that for any times $t$ and $s$, filter's output is

$$Y(t) = \int_{-\infty}^{\infty} h(u_1) X(t-u_1)\, du_1, \quad Y(t+s) = \int_{-\infty}^{\infty} h(u_2) X(t+s-u_2)\, du_2$$

- The autocorrelation function $R_Y(t, t+s)$ of the process $Y(t)$ is

$$R_Y(t, t+s) = \mathbb{E}\left[Y(t)Y(t+s)\right]$$

- Substituting $Y(t)$ and $Y(t+s)$ by their convolution forms

$$R_Y(t, t+s) = \mathbb{E}\left[\int_{-\infty}^{\infty} h(u_1) X(t-u_1)\, du_1 \int_{-\infty}^{\infty} h(u_2) X(t+s-u_2)\, du_2\right]$$

# Autocorrelation of filter's output (continued)

- Product of integrals is double integral of product

$$R_Y(t, t+s) = \mathbb{E}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(u_1)X(t-u_1)h(u_2)X(t+s-u_2)\,du_1du_2\right]$$

- Exchange order of integral and expectation

$$R_Y(t, t+s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(u_1)\mathbb{E}\Big[X(t-u_1)X(t+s-u_2)\Big]h(u_2)\,du_1du_2$$

- Expectation in the integral is autocorrelation function of input $X(t)$

$$\mathbb{E}\Big[X(t-u_1)X(t+s-u_2)\Big] = R_X\Big(t+s-u_2-(t-u_1)\Big) = R_X\big(s-u_2+u_1\big)$$

- Which upon substitution in expression for $R_Y(t, t+s)$ yields

$$R_Y(t, t+s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(u_1)R_X\big(s-u_2+u_1\big)h(u_2)\,du_1du_2 = R_Y(s)$$

# Jointly wide-sense stationary processes

- **Def:** Two WSS processes $X(t)$ and $Y(t)$ are said jointly WSS if

$$R_{XY}(t, t+s) := \mathbb{E}\left[X(t)Y(t+s)\right] = R_{XY}(s)$$

$\Rightarrow$ The cross-correlation function is shift-invariant

- If input to filter $X(t)$ is WSS, showed output $Y(t)$ also WSS
- Also jointly WSS since the input-output cross-correlation is

$$R_{XY}(t, t+s) = \mathbb{E}\left[X(t)\int_{-\infty}^{\infty} h(u)X(t+s-u)\,du\right]$$
$$= \int_{-\infty}^{\infty} h(u)R_X(s-u)\,du = R_{XY}(s)$$

$\Rightarrow$ Cross-correlation given by convolution $R_{XY}(s) = h(s) * R_X(s)$

# Autocorrelation of filter's output as convolution

- Going back to the autocorrelation of $Y(t)$, recall we found

$$R_Y(s) = \int_{-\infty}^{\infty} h(u_2) \left[ \int_{-\infty}^{\infty} h(u_1) R_X(s - u_2 + u_1)\, du_1 \right] du_2$$

- Inner integral is cross-correlation $R_{XY}(u_2 - s)$

$$R_Y(s) = \int_{-\infty}^{\infty} h(u_2) R_{XY}(u_2 - s) du_2$$

- Noting that $R_{XY}(u_2 - s) = R_{XY}(-(s - u_2))$

$$R_Y(s) = \int_{-\infty}^{\infty} h(u_2) R_{XY}(-(s - u_2)) du_2$$

- Autocorrelation given by convolution $R_Y(s) = h(s) * R_{XY}(-s)$

  $\Rightarrow$ Recall $R_Y(s) = R_Y(-s)$, hence also $R_Y(s) = h(-s) * R_{XY}(s)$

# Power spectral density of filter's output

- Power spectral density of $Y(t)$ is Fourier transform of $R_Y(s)$

$$S_Y(f) = \mathcal{F}\big(R_Y(s)\big) = \int_{-\infty}^{\infty} R_Y(s) e^{-j2\pi fs}\, ds$$

- Substituting $R_Y(s)$ for its value

$$S_Y(f) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(u_1) R_X\big(s - u_2 + u_1\big) h(u_2)\, du_1 du_2 \right) e^{-j2\pi fs}\, ds$$

- Change variable $s$ by variable $v = s - u_2 + u_1$ $(dv = ds)$

$$S_Y(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(u_1) R_X(v) h(u_2) e^{-j2\pi f(v+u_2-u_1)}\, du_1 du_2 dv$$

- Rewrite exponential as $e^{-j2\pi f(v+u_2-u_1)} = e^{-j2\pi fv} e^{-j2\pi fu_2} e^{+j2\pi fu_1}$

# Power spectral density of filter's output (continued)

- Write triple integral as product of three integrals

$$S_Y(f) = \int_{-\infty}^{\infty} h(u_1) e^{j2\pi f u_1}\, du_1 \int_{-\infty}^{\infty} R_X(v) e^{-j2\pi f v}\, dv \int_{-\infty}^{\infty} h(u_2) e^{-j2\pi f u_2}\, du_2$$

- Integrals are Fourier transforms

$$S_Y(f) = \mathcal{F}\big(h(-u_1)\big) \times \mathcal{F}\big(R_X(v)\big) \times \mathcal{F}\big(h(u_2)\big)$$

- Note definitions of $\Rightarrow$ $X(t)$'s PSD $\Rightarrow S_X(f) = \mathcal{F}\big(R_X(s)\big)$
  $\Rightarrow$ Filter's frequency response $\Rightarrow H(f) := \mathcal{F}\big(h(t)\big)$
  Also note that $\Rightarrow H^*(f) := \mathcal{F}\big(h(-t)\big)$

- Latter three observations yield (also use $H^*(f)H(f) = \big|H(f)\big|^2$)

$$S_Y(f) = H^*(f) S_X(f) H(f) = \big|H(f)\big|^2 S_X(f)$$

$\Rightarrow$ Key identity relating the input and output PSDs

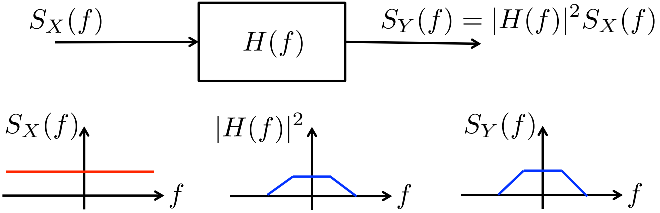
# Example: White noise filtering

Ex: Input process $X(t) = W(t) =$ white Gaussian noise with variance $\sigma^2$

$\Rightarrow$ Filter with frequency response $H(f)$. Q: PSD of output $Y(t)$?

- PSD of input $\Rightarrow S_W(f) = \sigma^2$
- PSD of output $\Rightarrow S_Y(f) = \left|H(f)\right|^2 S_W(f) = \left|H(f)\right|^2 \sigma^2$

$\Rightarrow$ Output's spectrum is filter's frequency response scaled by $\sigma^2$




Ex: System identification $\Rightarrow$ LTI system with unknown response

- White noise input $\Rightarrow$ PSD of output is frequency response of filter

# Interpretation of power spectral density

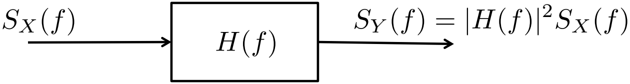
- Consider a narrowband filter with frequency response centered at $f_0$

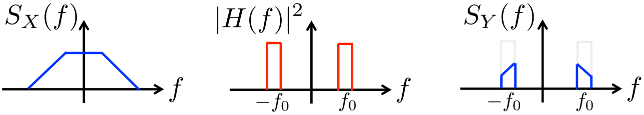
$$H(f) = 1 \quad \text{for: } f_0 - h/2 \le f \le f_0 + h/2$$
$$-f_0 - h/2 \le f \le -f_0 + h/2$$

- Input is WSS process with PSD $S_X(f)$. Output's power $P_Y$ is

$$P_Y = \int_{-\infty}^{\infty} S_Y(f)\,df = \int_{-\infty}^{\infty} \left|H(f)\right|^2 S_X(f)\,df \approx h\Big(S_X(f_0) + S_X(-f_0)\Big)$$

$\Rightarrow$ $S_X(f)$ is the power density the process $X(t)$ contains at frequency $f$



# Properties of power spectral density

**For WSS processes:**

**(i)** The power spectral density is a real-valued function

Proof.
Recall that $R_X(s) = R_X(-s)$ and $e^{j\theta} = \cos(\theta) + j\sin(\theta)$

$$
\begin{aligned}
S_X(f) &= \int_{-\infty}^{\infty} R_X(s) e^{-j2\pi fs}\, ds \\
&= \int_{-\infty}^{\infty} R_X(s)\cos(-2\pi fs)\, ds + j\int_{-\infty}^{\infty} R_X(-s)\sin(-2\pi fs)\, ds \\
&= \int_{-\infty}^{\infty} R_X(s)\cos(2\pi fs)\, ds
\end{aligned}
$$

Gray integral vanishes since $R_X(-s)\sin(-2\pi fs) = -R_X(s)\sin(2\pi fs)$ $\square$

**(ii)** The power spectral density is an even function, i.e., $S_X(f) = S_X(-f)$

# Properties of power spectral density (continued)

**For WSS processes:**

**(iii)** The power spectral density is a non-negative function, i.e., $S_X(f) \geq 0$

Proof.
Pass WSS $X(t)$ through narrowband filter centered at $f_0$

$$H(f) = 1 \quad \text{for: } f_0 - h/2 \leq f \leq f_0 + h/2$$
$$-f_0 - h/2 \leq f \leq -f_0 + h/2$$

For $h \to 0$, output's power $P_Y$ can be approximated as

$$0 \leq P_Y = \int_{-\infty}^{\infty} \left|H(f)\right|^2 S_X(f)\, df$$
$$\approx h\Big(S_X(f_0) + S_X(-f_0)\Big) = 2hS_X(f_0)$$

Since $f_0$ is arbitrary and $P_Y \geq 0 \;\Rightarrow S_X(f) \geq 0$ $\square$

# Example: Interference rejection filter

Ex: WSS signal $S(t)$ corrupted by additive, independent interference

$$I(t) = A\cos(2\pi f_0 t + \theta), \ \ \theta \sim \text{Uniform}(0, 2\pi)$$

$\Rightarrow$ Randomly phased sinusoidal interference $I(t)$ (fixed $A$, $f_0 > 0$)

- Corrupted signal $X(t) = S(t) + I(t)$. Q: Filter out interference?
- Sinusoidal interference has period $T = 1/f_0$. Use differencing filter

$$Y(t) = X(t) - X(t - T)$$

$\Rightarrow$ Difference $I(t) - I(t - T) = 0$ for all $t$

- Wish to determine the PSD of the output $S_Y(f) = \left|H(f)\right|^2 S_X(f)$

# Differencing filter

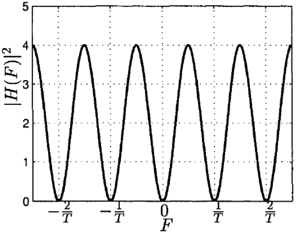
- The differencing filter is an LTI system with impulse response

$$Y(t) = X(t) - X(t-T) \;\Rightarrow\; h(t) = \delta(t) - \delta(t-T)$$

- By taking the Fourier transform, the frequency response becomes

$$H(f) = \int_{-\infty}^{\infty} (\delta(t) - \delta(t-T)) e^{-j2\pi ft} dt = 1 - e^{-j2\pi fT}$$

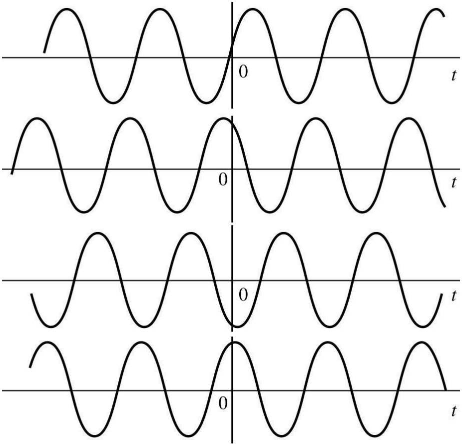
- The magnitude-squared of $H(f)$ is $|H(f)|^2 = 2 - 2\cos(2\pi fT)$




$\Rightarrow$ As expected, it exhibits zeros at multiples of $f = 1/T = f_0$

## Randomly phased sinusoid

- Interference $I(t) = A\cos(2\pi f_0 t + \theta)$, with $\theta \sim \text{Uniform}(0, 2\pi)$
  - $\Rightarrow$ Once $\theta$ is drawn, process realization specified for all $t$




- Above are four different sample paths of $I(t)$

# Randomly phased sinusoid is wide-sense stationary

- Q: Is $I(t)$ a wide-sense stationary process?
  $\Rightarrow$ Compute $\mu_I(t)$ and $R_I(t_1, t_2)$ and check
- Cosine integral over a cycle vanishes, hence

$$\mu_I(t) = \mathbb{E}\left[I(t)\right] = \int_0^{2\pi} A\cos(2\pi f_0 t + \theta)\frac{1}{2\pi}d\theta = 0$$

- Use $\cos(\theta_1)\cos(\theta_2) = (\cos(\theta_1 + \theta_2) + \cos(\theta_1 - \theta_2))/2$ to obtain

$$\begin{aligned} R_I(t_1, t_2) &= A^2\mathbb{E}\left[\cos(2\pi f_0 t_1 + \theta)\cos(2\pi f_0 t_2 + \theta)\right] \\ &= \frac{A^2}{2}\cos(2\pi f_0(t_2 - t_1)) + \frac{A^2}{2}\mathbb{E}\left[\cos(2\pi f_0(t_1 + t_2) + 2\theta)\right] \\ &= \frac{A^2}{2}\cos(2\pi f_0(t_2 - t_1)) \end{aligned}$$

- Thus $I(t)$ is WSS with PSD given by

$$S_I(f) = \mathcal{F}\left(R_I(s)\right) = \frac{A^2}{4}\delta(f - f_0) + \frac{A^2}{4}\delta(f + f_0)$$

# Power spectral density of filter's output

- Since $S(t)$ and $I(t)$ are independent and $\mu_I(t) = 0$

$$\begin{aligned} R_X(s) &= \mathbb{E}\left[(S(t) + I(t))(S(t+s) + I(t+s))\right] \\ &= R_S(s) + R_I(s) \end{aligned}$$

$\Rightarrow$ Also $S_X(f) = S_S(f) + S_I(f)$

- Therefore the PSD of the filter output $Y(t)$ is

$$\begin{aligned} S_Y(f) = |H(f)|^2 S_X(f) &= |H(f)|^2 (S_S(f) + S_I(f)) \\ &= 2\left(1 - \cos(2\pi f T)\right)(S_S(f) + S_I(f)) \end{aligned}$$

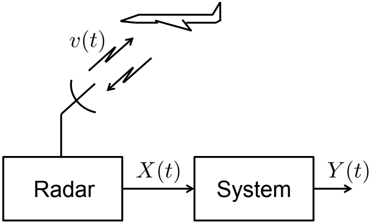
- Filter annihilates the tones in $S_I(f) = \frac{A^2}{4}\delta(f - f_0) + \frac{A^2}{4}\delta(f + f_0)$, so

$$S_Y(f) = 2\left(1 - \cos(2\pi f T)\right) S_S(f)$$

$\Rightarrow$ Unfortunately, the signal PSD has also been modified

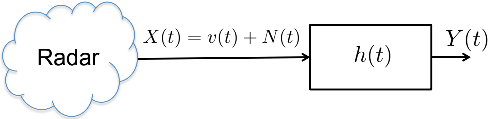
# The matched and Wiener filters

# A simple model of a radar system




- Air-traffic control system sends out a known radar pulse $v(t)$
- No plane in radar's range $\Rightarrow$ Radar output $X(t) = N(t)$ is noise
  $\Rightarrow$ Noise is zero-mean WSS process $N(t)$, with PSD $S_N(f)$
- Plane in range $\Rightarrow$ Reflected pulse in output $X(t) = v(t) + N(t)$
- Q: System to decide whether $X(t) = v(t) + N(t)$ or $X(t) = N(t)$?

# Filter design criterion




- Filter radar output $X(t)$ with LTI system $h(t)$. System output is

$$Y(t) = \int_{-\infty}^{\infty} h(t-s)[v(s) + N(s)]ds = v_0(t) + N_0(t)$$

- Filtered signal (radar pulse) and noise related components

$$v_0(t) = \int_{-\infty}^{\infty} h(t-s)v(s)ds, \quad N_0(t) = \int_{-\infty}^{\infty} h(t-s)N(s)ds$$

- Design filter to maximize output signal-to-noise ratio (SNR) at $t_0$

$$SNR = \frac{v_0^2(t_0)}{\mathbb{E}\left[N_0^2(t_0)\right]}$$

# Filtered signal and noise components

- The filtered noise power $\mathbb{E}\left[N_0^2(t_0)\right]$ is given by

$$\mathbb{E}\left[N_0^2(t_0)\right] = \int_{-\infty}^{\infty} S_{N_0}(f)df = \int_{-\infty}^{\infty} |H(f)|^2 S_N(f)df$$

- If $V(f) = \mathcal{F}(v(t))$, filtered radar pulse at time $t_0$

$$v_0(t_0) = \int_{-\infty}^{\infty} H(f)V(f)e^{j2\pi f t_0}df$$

- Multiply and divide by $\sqrt{S_N(f)}$, use complex conjugation

$$\begin{aligned} v_0(t_0) &= \int_{-\infty}^{\infty} H(f)\sqrt{S_N(f)}\frac{V(f)e^{j2\pi f t_0}}{\sqrt{S_N(f)}}df \\ &= \int_{-\infty}^{\infty} H(f)\sqrt{S_N(f)}\left[\frac{V^*(f)e^{-j2\pi f t_0}}{\sqrt{S_N(f)}}\right]^* df \end{aligned}$$

# Cauchy-Schwarz inequality

- The Cauchy-Schwarz inequality for complex functions $f$ and $g$ states

$$\left|\int_{-\infty}^{\infty} f(t)g^*(t)dt\right|^2 \leq \int_{-\infty}^{\infty} |f(t)|^2 dt \int_{-\infty}^{\infty} |g(t)|^2 dt$$

$\Rightarrow$ Equality is attained if and only if $f(t) = \alpha g(t)$

- Recall the filtered signal component at time $t_0$

$$v_0(t_0) = \int_{-\infty}^{\infty} H(f)\sqrt{S_N(f)} \left[\frac{V^*(f)e^{-j2\pi f t_0}}{\sqrt{S_N(f)}}\right]^* df$$

- Use the Cauchy-Schwarz inequality to obtain the upper-bound

$$|v_0(t_0)|^2 \leq \int_{-\infty}^{\infty} |H(f)|^2 S_N(f) df \int_{-\infty}^{\infty} \frac{|V(f)|^2}{S_N(f)} df$$

# The matched filter

- Since $\mathbb{E}\left[N_0^2(t_0)\right] = \int_{-\infty}^{\infty} |H(f)|^2 S_N(f) df$, bound SNR

$$SNR = \frac{|v_0(t_0)|^2}{\mathbb{E}\left[N_0^2(t_0)\right]} \leq \frac{\mathbb{E}\left[N_0^2(t_0)\right] \int_{-\infty}^{\infty} \frac{|V(f)|^2}{S_N(f)} df}{\mathbb{E}\left[N_0^2(t_0)\right]} = \int_{-\infty}^{\infty} \frac{|V(f)|^2}{S_N(f)} df$$

- The maximum SNR is attained when

$$H(f)\sqrt{S_N(f)} = \alpha \frac{V^*(f)e^{-j2\pi f t_0}}{\sqrt{S_N(f)}}$$

- The sought matched filter has frequency response

$$H(f) = \alpha \frac{V^*(f)e^{-j2\pi f t_0}}{S_N(f)}$$

$\Rightarrow$ $H(f)$ is "matched" to the known radar pulse and noise PSD

# Example: Matched filter for white noise

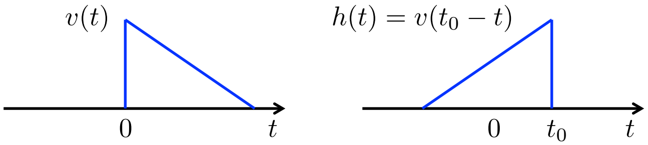
Ex: Suppose noise $N(t)$ is white, with PSD $S_N(f) = \sigma^2$. Let $\alpha = \sigma^2$

- The frequency response of the matched filter simplifies to

$$H(f) = V^*(f)e^{-j2\pi f t_0}$$

- The inverse Fourier transform of $H(f)$ yields the impulse response

$$h(t) = v(t_0 - t)$$




- Simply a time-reversed and translated copy of the radar pulse $v(t)$

# Analysis of matched filter output

- PSD of filtered noise is $S_{N_0}(f) = |H(f)|^2 S_N(f)$. For matched filter

$$S_{N_0}(f) = \frac{|\alpha V(f)|^2}{S_N^2(f)} S_N(f) = \frac{|\alpha V(f)|^2}{S_N(f)}$$

- Inverse Fourier transform yields autocorrelation function of $N_0(t)$

$$R_{N_0}(s) = \int_{-\infty}^{\infty} \frac{|\alpha V(f)|^2}{S_N(f)} e^{j2\pi fs} df$$

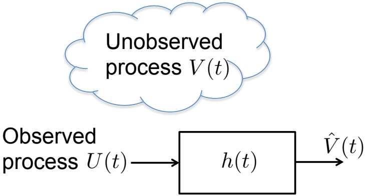
- The matched filter signal output is

$$v_0(t) = \int_{-\infty}^{\infty} H(f)V(f)e^{j2\pi ft} df = \int_{-\infty}^{\infty} \frac{\alpha |V(f)|^2}{S_N(f)} e^{j2\pi f(t-t_0)} df$$

- Last two equations imply that $v_0(t) = (1/\alpha) R_{N_0}(t - t_0)$

  $\Rightarrow$ Matched filter signal output $\propto$ shifted autocorrelation

# Linear estimation



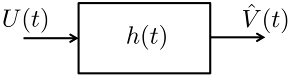

- Estimate unobserved process $V(t)$ from correlated process $U(t)$
  - $\Rightarrow$ Zero mean $U(t)$ and $V(t)$
  - $\Rightarrow$ Known (cross-) PSDs $S_U(f)$ and $S_{VU}(f)$

Ex: Say $U(t) = V(t) + W(t)$, with $W(t)$ a white noise process

- Restrict attention to linear estimators

$$\hat{V}(t) = \int_{-\infty}^{\infty} h(s)U(t-s)ds$$

# Filter design criterion




- Criterion is mean-square error (MSE) minimization, i.e, find

$$\min_{h} \mathbb{E}\left[|V(t) - \hat{V}(t)|^2\right], \quad \text{s. to } \hat{V}(t) = \int_{-\infty}^{\infty} h(s)U(t-s)ds$$

- Suppose $\tilde{h}(t)$ is any other impulse response such that

$$\tilde{V}(t) = \int_{-\infty}^{\infty} \tilde{h}(s)U(t-s)ds$$

$\Rightarrow$ MSE-sense optimality of filter $h(t)$ means

$$\mathbb{E}\left[|V(t) - \hat{V}(t)|^2\right] \leq \mathbb{E}\left[|V(t) - \tilde{V}(t)|^2\right]$$

# Orthogonality principle

Theorem

*If for every linear filter $\tilde{h}(t)$ it holds*

$$\mathbb{E}\left[(V(t)-\hat{V}(t))\int_{-\infty}^{\infty}\tilde{h}(s)U(t-s)ds\right]=0$$

*then $h(t)$ is the MSE-sense optimal filter.*

- Orthogonality principle implicitly characterizes the optimal filter $h(t)$
- Condition must hold for all $\tilde{h}$, in particular for $h-\tilde{h}$ implying

$$\mathbb{E}\left[(V(t)-\hat{V}(t))(\hat{V}(t)-\tilde{V}(t))\right]=0$$

$\Rightarrow$ Recall this identity, we will use it next

# Orthogonality principle (proof)

Proof.

- The MSE for an arbitrary filter $\tilde{h}(t)$ can be written as

$$\mathbb{E}\left[|V(t)-\tilde{V}(t)|^2\right]=\mathbb{E}\left[|(V(t)-\hat{V}(t))+(\hat{V}(t)-\tilde{V}(t))|^2\right]$$

- Expand the squares, use linearity of expectation

$$\begin{aligned}\mathbb{E}\left[|V(t)-\tilde{V}(t)|^2\right]&=\mathbb{E}\left[|V(t)-\hat{V}(t)|^2\right]+\mathbb{E}\left[|\hat{V}(t)-\tilde{V}(t)|^2\right]\\&+2\mathbb{E}\left[(V(t)-\hat{V}(t))(\hat{V}(t)-\tilde{V}(t))\right]\end{aligned}$$

- But $\mathbb{E}\left[(V(t)-\hat{V}(t))(\hat{V}(t)-\tilde{V}(t))\right]=0$ by assumption, hence

$$\begin{aligned}\mathbb{E}\left[|V(t)-\tilde{V}(t)|^2\right]&=\mathbb{E}\left[|V(t)-\hat{V}(t)|^2\right]+\mathbb{E}\left[|\hat{V}(t)-\tilde{V}(t)|^2\right]\\&\geq\mathbb{E}\left[|V(t)-\hat{V}(t)|^2\right]\end{aligned}$$

□

# Leveraging the orthogonality principle

- If $h(t)$ is optimum, for any $\tilde{h}(t)$ orthogonality principle implies

$$0 = \mathbb{E}\left[(V(t) - \hat{V}(t)) \int_{-\infty}^{\infty} \tilde{h}(s)U(t-s)ds\right]$$
$$= \mathbb{E}\left[\int_{-\infty}^{\infty} \tilde{h}(s)(V(t) - \hat{V}(t))U(t-s)ds\right]$$

- Interchange order of expectation and integration, $\tilde{h}(t)$ deterministic

$$\int_{-\infty}^{\infty} \tilde{h}(s)\mathbb{E}\left[(V(t) - \hat{V}(t))U(t-s)\right] ds = 0$$

- Recall definitions of cross-correlation functions $R_{VU}(s)$ and $R_{\hat{V}U}(s)$

$$\int_{-\infty}^{\infty} \tilde{h}(s)(R_{VU}(s) - R_{\hat{V}U}(s))ds = 0$$

# Matching cross-correlations condition

- For arbitrary $\tilde{h}(t)$, orthogonality principle requires

$$\int_{-\infty}^{\infty} \tilde{h}(s)(R_{VU}(s) - R_{\hat{V}U}(s))ds = 0$$

- In particular, select $\tilde{h}(t) = R_{VU}(t) - R_{\hat{V}U}(t)$ to get

$$\int_{-\infty}^{\infty} (R_{VU}(s) - R_{\hat{V}U}(s))^2 ds = 0$$

$\Rightarrow$ Above integral vanishes if and only if $R_{VU}(s) = R_{\hat{V}U}(s)$

- At the optimum, cross-correlations $R_{VU}(s)$ and $R_{\hat{V}U}(s)$ coincide

$\Rightarrow$ Reasonable, since MSE is a second-order cost function

# The Wiener filter

- Best filter yields estimates $\hat{V}(t)$ for which $R_{VU}(s) = R_{\hat{V}U}(s)$
- Since $\hat{V}(t)$ is the output of the LTI system $h(t)$, with input $U(t)$

$$R_{\hat{V}U}(s) = \int_{-\infty}^{\infty} h(t)R_U(s-t)dt = h(s) * R_U(s)$$

- Taking Fourier transforms

$$S_{\hat{V}U}(f) = H(f)S_U(f) = S_{VU}(f)$$

$\Rightarrow$ The optimal Wiener filter has frequency response

$$H(f) = \frac{S_{VU}(f)}{S_U(f)}$$

# Glossary

- Strict stationarity
- Shift invariance
- Power of a process
- Limit distribution
- Mean function
- Autocorrelation function
- Wide-sense stationarity
- Fourier transform
- Frequency components
- Linear time-invariant system
- Impulse response
- Convolution
- Frequency response
- Power spectral density
- Joint wide-sense stationarity
- Cross-correlation function
- System identification
- Signal-to-noise ratio
- Cauchy-Schwarz inequality
- Matched filter
- Linear estimation
- Mean-square error
- Orthogonality principle
- Wiener filter