

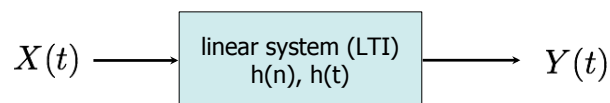
Random signal processing

Stochastic Processes for EE (EE2511)

Lecture 5

1

Summary Filtered WSS Process



$$\mu_X$$

$$\mu_Y = \mu_X H_0$$

$$R_X(k)$$

$$R_Y(k) = h(k) * h(-k) * R_X(k)$$

- Time continuous: $R_Y(\tau) = h(\tau) * h(-\tau) * R_X(\tau)$

2

What can we do with Stoch. Proc.?

- System identification
- Modeling of real signals (e.g. speech)
- Prediction (next week)

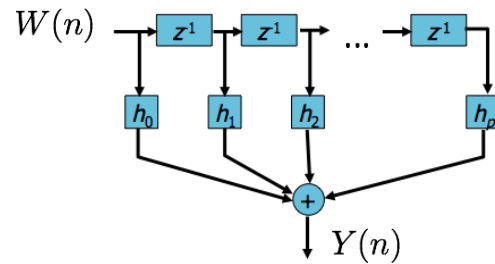
3

Today's Agenda

- Moving average (MA) and autoregressive (AR) processes
- Modeling of speech signals as an AR process
- Fourier Transform of the autocorrelation function: Power Spectral Density

Two typical LTI systems

- Tapped delay line or Finite Impulse Response Filter (FIR) system

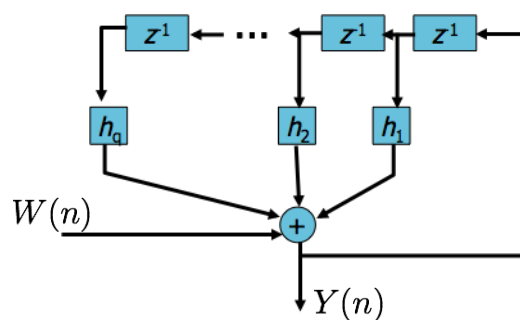


$$Y(n) = h_0 W(n) + h_1 W(n-1) + \dots + h_p W(n-p)$$

5

Two typical LTI systems

- Infinite Impulse Response (IIR) system



$$Y(n) = h_1 Y(n-1) + h_2 Y(n-2) + \dots + h_q Y(n-q) + W(n)$$

6

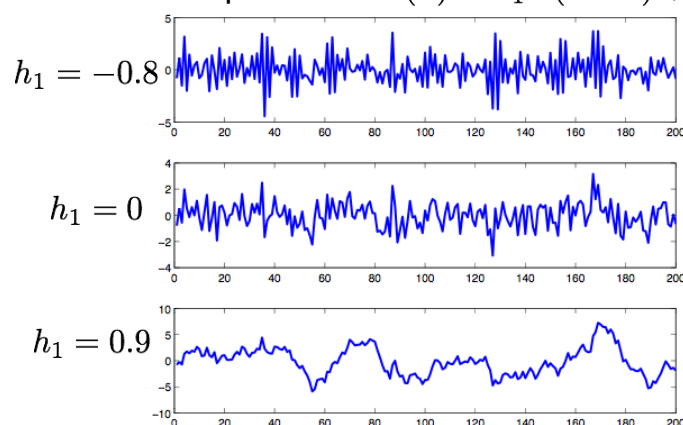
Terminology

If we input white noise into these LTI systems, then

- output of FIR filter is called a *moving average* (MA) process
- output of IIR filter is called an *autoregressive* (AR) process
- Combination of the two is called ARMA process
- Many speech processing systems (including speech compression in GSM) use AR model

Example of AR realizations

- First order AR process $Y(n) = h_1 Y(n-1) + W(n)$



Autocorrelation Functions (MA-1)

- First order MA process

$$Y(n) = h_0 W(n) + h_1 W(n-1)$$

- Autocorrelation function (*Method 1*)

$$\begin{aligned} R_Y(k) &= E[Y(n)Y(n+k)] \\ &= E[(h_0 W(n) + h_1 W(n-1))(h_0 W(n+k) + h_1 W(n+k-1))] \\ &= h_0^2 E[W(n)W(n+k)] + h_0 h_1 E[W(n)W(n+k-1)] \\ &\quad + h_0 h_1 E[W(n-1)W(n+k)] + h_1^2 E[W(n-1)W(n+k-1)] \\ &= (h_0^2 + h_1^2)R_W(k) + h_0 h_1 R_W(k+1) + h_0 h_1 R_W(k-1) \end{aligned}$$

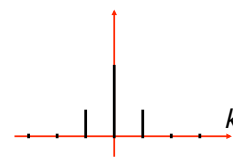
Autocorrelation Functions (MA-1)

- First order MA process

$$Y(n) = h_0 W(n) + h_1 W(n-1)$$

- Autocorrelation function (*Method 1*)

$$\begin{aligned} R_Y(k) &= E[Y(n)Y(n+k)] \\ &= (h_0^2 + h_1^2)R_W(k) + h_0 h_1 R_W(k+1) + h_0 h_1 R_W(k-1) \\ &= \begin{cases} (h_0^2 + h_1^2)\sigma_w^2 & k = 0 \\ h_0 h_1 \sigma_w^2 & k = 1 \\ h_0 h_1 \sigma_w^2 & k = -1 \\ 0 & \text{elsewhere} \end{cases} \end{aligned}$$

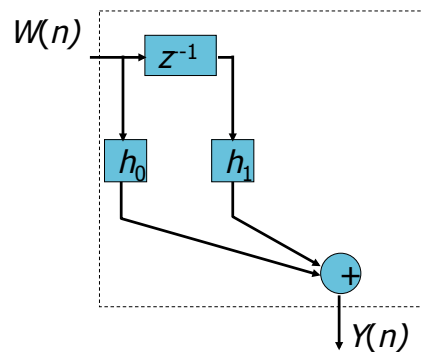


Autocorrelation Functions (MA-1)

- First order MA process

$$Y(n) = h_0 W(n) + h_1 W(n-1)$$

- Autocorrelation function (*Method 2*)



Linear system with impulse response:
 $\dots, 0, 0, 0, h_0, h_1, 0, 0, 0, \dots$

Autocorrelation Functions (MA-1)

- First order MA process

$$Y(n) = h_0 W(n) + h_1 W(n-1)$$

- Autocorrelation function (*Method 2*)

$$Y(n) = (\dots, 0, 0, h_0, h_1, 0, 0, \dots) * W(n) = h(n) * W(n)$$

$$\begin{aligned} R_Y(k) &= h(k) * h(-k) * R_W(k) \\ &= (\dots, 0, 0, h_0 h_1, h_0^2 + h_1^2, h_0 h_1, 0, 0, \dots) * \sigma_W^2 \delta(k) \end{aligned}$$

Autocorrelation Functions (MA-1)

Autocorrelation function (*Method 2*)

$$R_Y(k) = h(k) * h(-k) * R_W(k)$$

$$\begin{aligned} f(k) &= h(k) * h(-k) \\ &= \sum_n h(n)h(-(k-n)) \\ &= h(0)h(-k) + h(1)h(-k+1) \\ R_Y(k) &= \sum_n (h(0)h(-n) + h(1)h(-n+1)) \sigma_W^2 \delta(k-n) \\ &= \sigma_W^2 h(0)h(-k) + \sigma_W^2 h(1)h(-k+1) \\ &= \begin{cases} (h(0)^2 + h(1)^2) \sigma_W^2 & \text{if } k = 0 \\ h(0)h(1) \sigma_W^2 & \text{if } k = 1 \\ h(0)h(1) \sigma_W^2 & \text{if } k = -1 \end{cases} \end{aligned}$$



Autocorrelation Functions (AR-1)

- First order AR process

$$Y(n) = h_1 Y(n-1) + W(n)$$



Autocorrelation Functions (AR-1)

- First order AR process

$$Y(n) = h_1 Y(n-1) + W(n)$$

- Autocorrelation function (*Method 1*)

$$\begin{aligned}
 R_Y(k) &= E[Y(n)Y(n+k)] \\
 &= E[Y(n)(h_1 Y(n+k-1) + W(n+k))] \\
 &= h_1 E[Y(n)Y(n+k-1)] + E[Y(n)W(n+k)] \\
 &= h_1 R_Y(k-1) + E[Y(n)]E[W(n+k)] \quad \text{independent} \\
 &= h_1 R_Y(k-1) \quad (k > 0)
 \end{aligned}$$

(Similarly for $k < 0$:

$$R_Y(k) = h_1 R_Y(k+1) \quad (k < 0)$$

TU Delft

Autocorrelation Functions (AR-1)

- First order AR process

$$Y(n) = h_1 Y(n-1) + W(n)$$

- Autocorrelation function (*Method 1*)

$$\begin{aligned}
 R_Y(0) &= E[(h_1 Y(n-1) + W(n))(h_1 Y(n-1) + W(n))] \\
 &= h_1^2 E[Y(n-1)^2] + 2h_1 E[Y(n-1)W(n)] + E[W(n)^2] \\
 &= h_1^2 R_Y(0) + 2h_1 E[Y(n-1)]E[W(n)] + R_W(0) \\
 &= h_1^2 R_Y(0) + \sigma_w^2
 \end{aligned}$$

↓

$$R_Y(0) = \frac{1}{1-h_1^2} \sigma_w^2$$

TU Delft

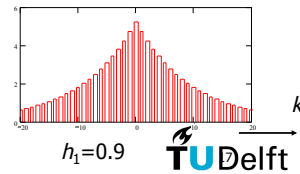
Autocorrelation Functions (AR-1)

- First order AR process

$$Y(n) = h_1 Y(n-1) + W(n)$$

- Autocorrelation function (*Method 1*)

$$\left. \begin{aligned} R_Y(k) &= h_1 R_Y(k-1) & k > 0 \\ R_Y(k) &= h_1 R_Y(k+1) & k < 0 \\ R_Y(0) &= \frac{1}{1-h_1^2} \sigma_w^2 \end{aligned} \right\} \Rightarrow R_Y(k) = \frac{\sigma_w^2}{1-h_1^2} h_1^{|k|}$$

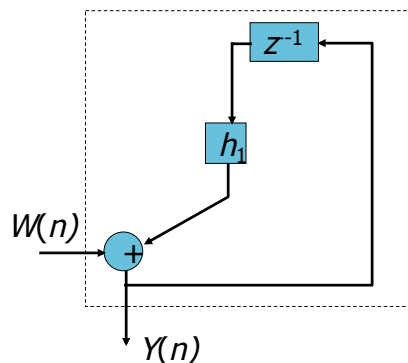


Autocorrelation Functions (AR-1)

- First order AR process

$$Y(n) = h_1 Y(n-1) + W(n)$$

- Autocorrelation function (*Method 2*)



Linear system with impulse response:

$$\begin{aligned} &\dots, 0, 1, h_1, h_1^2, h_1^3, \dots \\ &= \dots, 0, 0, 0, h_1^n \end{aligned}$$

Autocorrelation Functions (AR-1)

- First order AR process

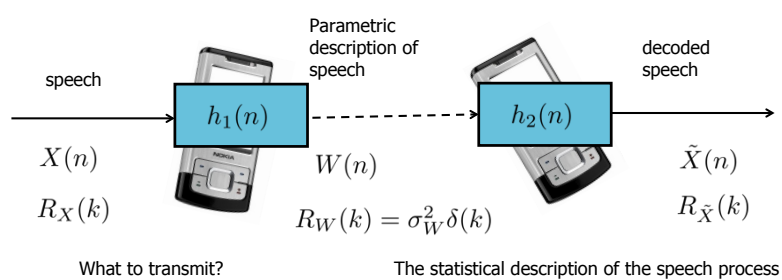
$$Y(n) = h_1 Y(n-1) + W(n)$$

- Autocorrelation function (*Method 2*)

$$Y(n) = (\dots, 0, 0, h_1^k) * W(n) = h(n) * W(n)$$

$$\begin{aligned} R_Y(k) &= h(k) * h(-k) * R_W(k) \\ &= (\dots, h_1^3, h_1^2, h_1, 1, h_1, h_1^2, h_1^3, \dots) * \sigma_W^2 \delta(k) \\ &= \frac{\sigma_W^2}{1 - h_1^2} h_1^{|k|} \end{aligned}$$

What About Speech?!?



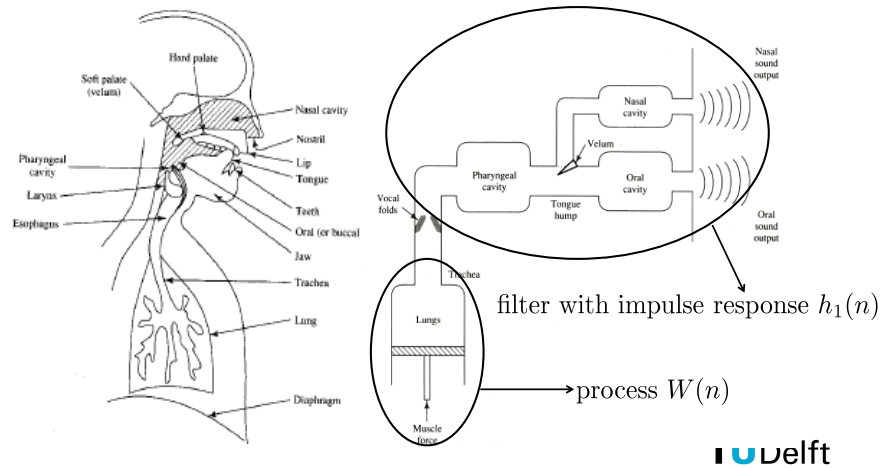
Choose $h_1(n)$ such that the output is uncorrelated with minimum variance.

$$W(n) = h_1(n) * X(n)$$

Then transmit $h_1(n)$ (= inverse filter of $h_2(n)$) and σ_W^2

Speech Production Model

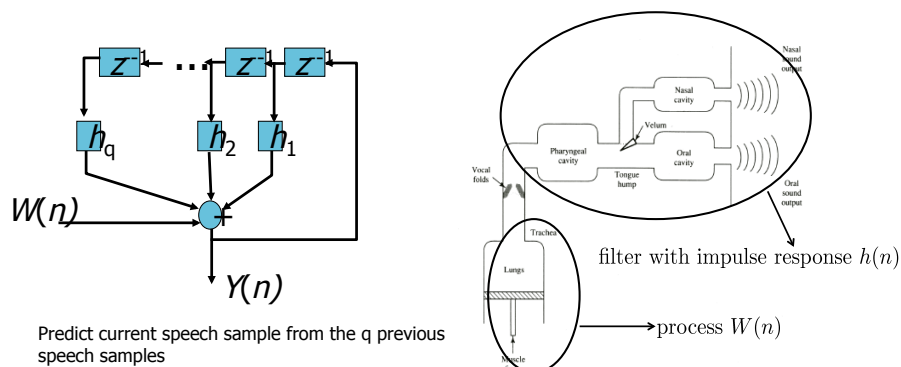
What does $h_1(n)$ model? It is a model of the speech production process.



TU Delft

Speech Production Model

Typically, speech is seen as an AR-process:



$$Y(n) = h_1 Y(n-1) + h_2 Y(n-2) + \dots + h_q Y(n-q) + W(n)$$

TU Delft

Estimation of the AR Coefficients (1)

- The signal model should match the real data as well as possible

- Real data $Y(n)$

- Modeled data $\hat{Y}(n) = \sum_{k=1}^q h_k Y(n-k)$ (linear prediction)

- Difference $Y(n) - \hat{Y}(n) = Y(n) - \sum_{k=1}^q h_k Y(n-k) = W(n)$

- Minimize the difference by choosing the AR coefficients optimally



Estimation of the AR Coefficients (2)

- Difference is quantified as variance

$$\sigma_w^2 = E[(Y(n) - \hat{Y}(n))^2] = E[(Y(n) - \sum_{k=1}^q h_k Y(n-k))^2]$$

- Minimize this difference:

$$\min_{h_i} \sigma_w^2 \quad i = 1, 2, \dots, q$$

$$\Rightarrow \frac{\partial}{\partial h_i} \sigma_w^2 = 0 \quad i = 1, 2, \dots, q$$

$$\frac{\partial}{\partial h_i} E[(Y(n) - \sum_{k=1}^q h_k Y(n-k))^2] = 0$$



Estimation of the AR Coefficients (3)

- Solution of minimization problem:

$$\begin{aligned}\frac{\partial}{\partial h_i} E[(Y(n) - \sum_{k=1}^q h_k Y(n-k))^2] &= 0 \quad i=1,2,\dots,q \\ E[(Y(n) - \sum_{k=1}^q h_k Y(n-k))(-2) \left\{ \frac{\partial}{\partial h_i} \sum_{k=1}^q h_k Y(n-k) \right\}] &= 0 \\ E[(Y(n) - \sum_{k=1}^q h_k Y(n-k))Y(n-i)] &= 0 \\ R_Y(i) &= \sum_{k=1}^q h_k R_Y(i-k)\end{aligned}$$

Estimation of the AR Coefficients (4)

- Yule Walker or normal equations

$$R_Y(i) = \sum_{k=1}^q h_k R_Y(i-k) \quad i = 1, 2, \dots, q$$

$$\begin{pmatrix} R_Y(1) \\ R_Y(2) \\ \vdots \\ R_Y(q) \end{pmatrix} = \begin{pmatrix} R_Y(0) & R_Y(1) & \dots & R_Y(q-1) \\ R_Y(1) & R_Y(0) & R_Y(1) & \vdots \\ \vdots & R_Y(1) & \ddots & R_Y(1) \\ R_Y(q-1) & \dots & R_Y(1) & R_Y(0) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_q \end{pmatrix}$$

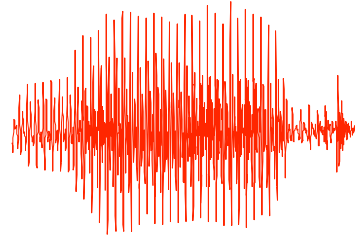
- Variance of excitation noise:

$$\sigma_w^2 = E[(Y(n) - \sum_{k=1}^q h_k Y(n-k))^2] = R_Y(0) - \sum_{k=1}^q h_k R_Y(k)$$

How to Synthesize speech? (1)

- Estimate the correlation function

$$\tilde{R}_Y(k) = \frac{1}{N-k} \sum_{n=1}^{N-k} Y(n)Y(n+k)$$



- Compute filter coefficients

$$\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_q \end{pmatrix} = \begin{pmatrix} \tilde{R}_Y(0) & \tilde{R}_Y(1) & \cdots & \tilde{R}_Y(q-1) \\ \tilde{R}_Y(1) & \tilde{R}_Y(0) & \tilde{R}_Y(1) & \vdots \\ \vdots & \tilde{R}_Y(1) & \ddots & \tilde{R}_Y(1) \\ \tilde{R}_Y(q-1) & \cdots & \tilde{R}_Y(1) & \tilde{R}_Y(0) \end{pmatrix}^{-1} \begin{pmatrix} \tilde{R}_Y(1) \\ \tilde{R}_Y(2) \\ \vdots \\ \tilde{R}_Y(q) \end{pmatrix}$$

How to Synthesize speech? (2)

- Variance of excitation noise:

$$\sigma_w^2 = \tilde{R}_Y(0) - \sum_{k=1}^q h_k \tilde{R}_Y(k)$$

- Perform filtering:

$$Y(n) = h_1 Y(n-1) + h_2 Y(n-2) + \dots + h_q Y(n-q) + W(n)$$

If we apply z-transform: $H(z) = \frac{1}{1 - \sum_{q=1}^Q h_q z^{-q}}$

All-pole!

Voiced vs. Unvoiced Speech

Does $W(n)$ really get uncorrelated in practice?

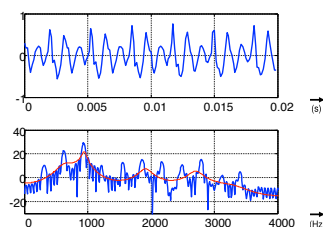
This depends on

- the model order used in $H_1(z^{-1})$
- the type of speech sound:
 - voiced - excitation still contains the long term correlation that originates from the vocal cords
 - unvoiced

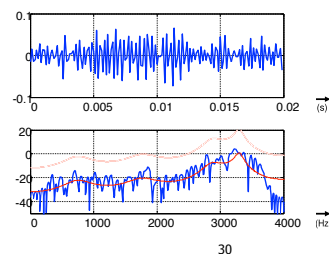
Voiced vs. Unvoiced Speech

- Voiced: Air pushed through the glottis which oscillates, generating quasi-periodic puffs of air (e.g. vowels /a/, /i/, etc.)
- Unvoiced: Air forced through constriction somewhere along the vocal tract (e.g. /s/, /f/).

voiced:

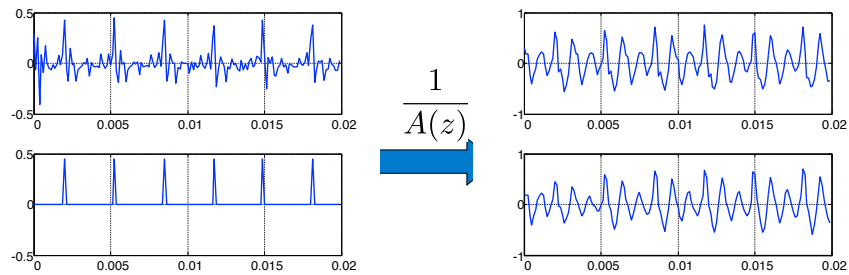


unvoiced:



Example: Speech Coding

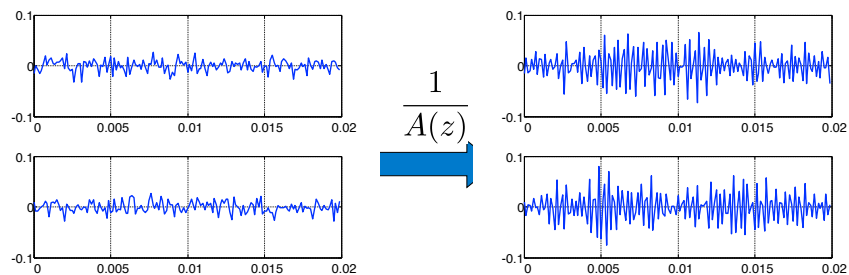
voiced speech:



31

Example: Speech Coding

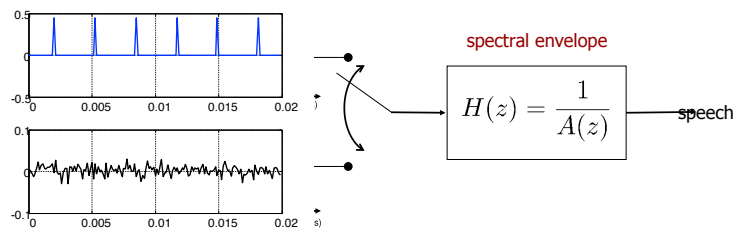
unvoiced speech:



32

Vocoder

- Discrete-time linear source-filter model of speech production

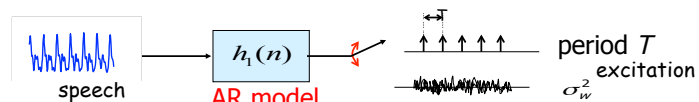


Speech can be synthesized if we know:

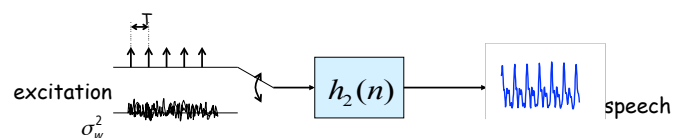
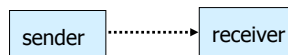
1. spectral envelope
2. pitch period T_0 (can be computed using the auto-correlation function)

33

Demonstration - Speech Coding



1. Determine filter $h(n)$ using Yule – Walker equations and error variance
2. Determine whether speech is unvoiced or voiced
3. Transmission/Quantization of $h(n)$ and voiced/unvoiced information



4. Synthesis of speech based on received filter $h(n)$ and voiced/unvoiced information.

Demo Vocoder

English female speech, $f_s = 8$ kHz, mono, 8 bit/sample



- original



- modelled



- modelled as unvoiced speech

German male speech, $f_s = 8$ kHz, mono, 8 bit/sample



- original



- modelled



- modelled as unvoiced speech

35

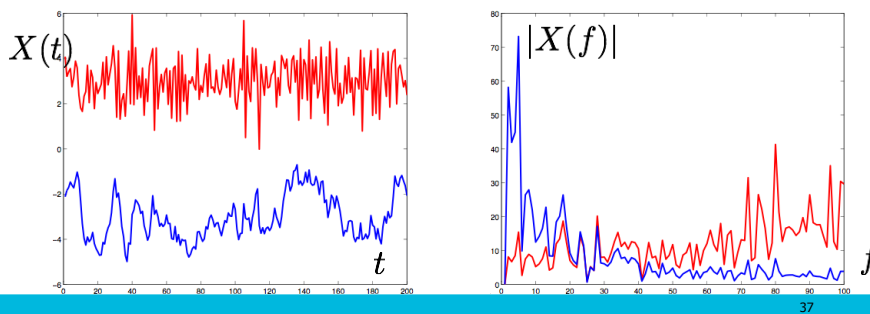
Usage of Fourier transforms The Power Spectral Density

- (needed for system identification)

36

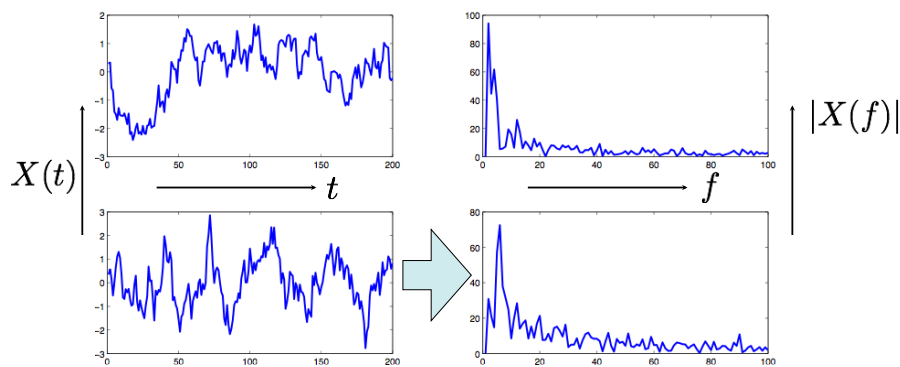
Usual tool in Signal Processing

- Fourier transforms are used to
 - describe deterministic signals (functions) and linear systems
 - analyze and design signals and linear systems



What about random signals?

- Fourier transforms of random signals are rather useless: the result depends on the particular realization



Power Spectral Density

- Instead of considering the Fourier transform of one realization, we should look at 'average behavior' in the Fourier domain

- **Power Spectral Density (PSD)**

$$S_X(f) = F\{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) d\tau$$

$$S_X(f) = F\{R_X(k)\} = \sum_{k=-\infty}^{\infty} R_X(k) \exp(-j2\pi fk)$$

PSD only exists for WSS random processes!

39

Actually...

- The original definition of PSD is a bit more ugly:

$$S_X(f) = \lim_{T \rightarrow \infty} \frac{1}{2T} E\left[|X_T(f)|^2\right]$$

- It highlights the fact that we are computing the average power in a frequency, but for (in principle) infinitely long signals
- The Wiener-Khintchine theorem shows the equivalence with

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) d\tau$$

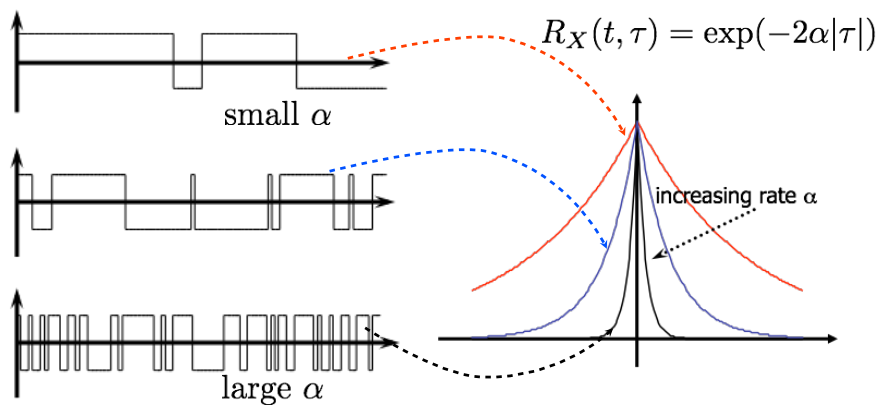
40

Heuristic interpretation

- (Amplitude component of) Fourier transform of a deterministic signal gives the strength of a particular complex exponential (sine-cosine combination)
- PSD of a random process/signal gives the **average power** carried by a particular complex exponential
- PSD can be calculated from the autocorrelation function of the WSS random process

41

Example Random telegraph



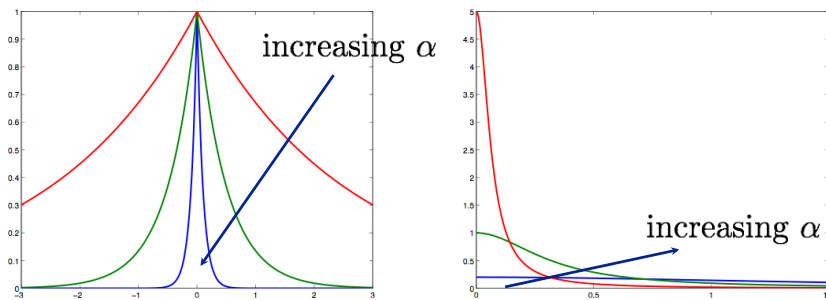
- Realizations of three **different** random processes

42

Example Random Telegraph

$$R_X(t, \tau) = \exp(-2\alpha|\tau|)$$

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



43

From PSD to autocorrelation funct.

- Continuous time:

$$S_X(f) = F\{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) d\tau$$

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) \exp(j2\pi f\tau) df$$

- Discrete time:

$$S_X(f) = F\{R_X(k)\} = \sum_{k=-\infty}^{\infty} R_X(k) \exp(-j2\pi fk)$$

$$R_X(k) = \int_{-1/2}^{1/2} S_X(f) \exp(j2\pi fk) df$$

44

Properties of PSD

- Because we are using real-valued signals, the power spectral density function is symmetric around $f=0$

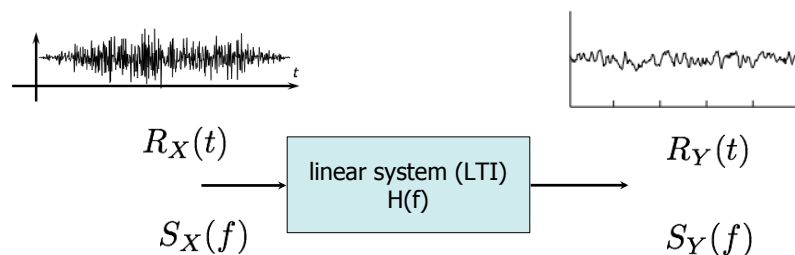
$$S_X(-f) = S_X(f)$$

- The function expresses a **density**, and is therefore always non-negative
- Integral over power spectral density is **average power**

$$\int_{-\infty}^{\infty} S_X(f) df = E[X(t)^2] = R_X(0)$$

45

Filtered WSS Process



- The power spectral density of $Y(t)$ is given by

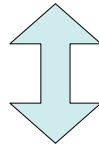
$$S_Y(f) = |H(f)|^2 S_X(f)$$

46

Result can easily be understood

- Since

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



Fourier transform

- then

$$\begin{aligned} S_Y(f) &= H(f)H^*(f)S_X(f) \\ &= |H(f)|^2 S_X(f) \end{aligned}$$

47

Autocorrelation Functions (AR-1)

- First order AR process

$$Y(n) = h_1 Y(n-1) + W(n)$$

- Autocorrelation function

$$\begin{aligned} \textcircled{1} \quad R_Y(k) &= E[Y(n)Y(n+k)] \\ &= E[Y(n)(h_1 Y(n+k-1) + W(n+k))] \\ &= \dots \end{aligned}$$

$$\begin{aligned} \textcircled{2} \quad Y(n) &= (\dots, 0, 0, h_1^k) * W(n) = h(n) * W(n) \\ R_Y(k) &= h(k) * h(-k) * R_W(k) \\ &= (\dots, h_1^3, h_1^2, h_1, 1, h_1, h_1^2, h_1^3, \dots) * \sigma_W^2 \delta(k) = \frac{\sigma_W^2}{1 - h_1^2} h_1^{|k|} \end{aligned}$$

Autocorrelation Functions (AR-1)

- First order AR process

$$Y(n) = h_1 Y(n-1) + W(n)$$

- Impulse response

$$h(n) = \begin{cases} 0 & n < 0 \\ h_1^n & n \geq 0 \end{cases}$$

- Frequency response

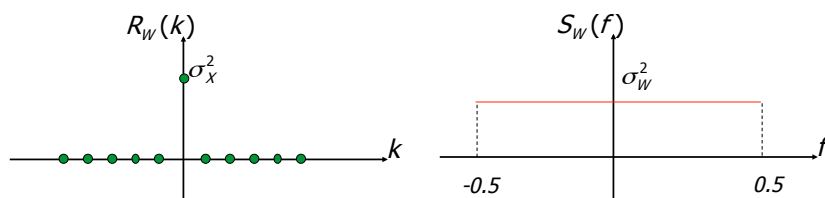
$$H(f) = \frac{1}{1 - h_1 \exp(-j2\pi f)}$$

Autocorrelation Functions (AR-1)

- Power Spectral Density of Output $Y(k)$

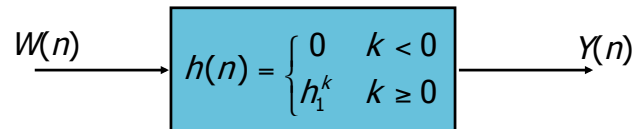
$$S_Y(f) = |H(f)|^2 S_W(f)$$

- Power Spectral Density of Input $W(k)$



$$S_Y(f) = |H(f)|^2 \sigma_W^2 = \frac{\sigma_W^2}{|1 - h_1 \exp(-j2\pi f)|^2} \Rightarrow R_Y(k)$$

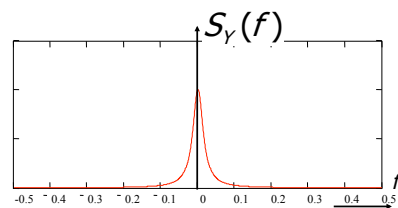
Autocorrelation Functions (AR-1)



- Power Spectral Density

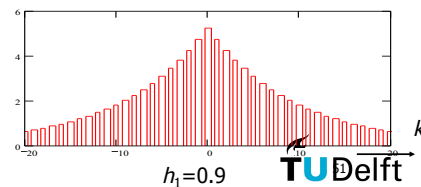
$$S_Y(f) = |H(f)|^2 \sigma_W^2$$

$$= \frac{\sigma_W^2}{|1 - h_1 \exp(-j2\pi f)|^2}$$

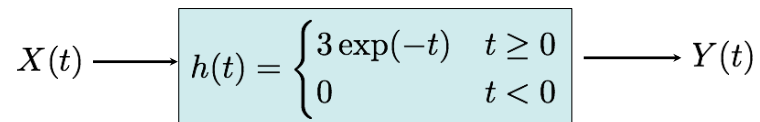


- Autocorrelation Function

$$R_Y(k) = \frac{\sigma_W^2}{1 - h_1^2} h_1^{|k|}$$



Example (time continuous)



$$R_X(\tau) = 4 + 3\delta(\tau)$$

- What is $R_Y(\tau)$ =?

Example (time continuous)

$$X(t) \longrightarrow h(t) = \begin{cases} 3 \exp(-t) & t \geq 0 \\ 0 & t < 0 \end{cases} \longrightarrow Y(t)$$

$$R_X(\tau) = 4 + 3\delta(\tau)$$

- then $S_X(f) = 4\delta(f) + 3$ and $H(f) = \frac{3}{1 + j(2\pi f)}$ (pg 413. book)
 - and $S_Y(f) = |H(f)|^2 S_X(f) = \frac{9}{1 + (2\pi f)^2} (4\delta(f) + 3)$
- $$= 36\delta(f) + \frac{27}{1 + (2\pi f)^2}$$
- Inverse Fourier transform
- so $R_Y(\tau) = 36 + \frac{27}{2} \exp(-|\tau|)$

53

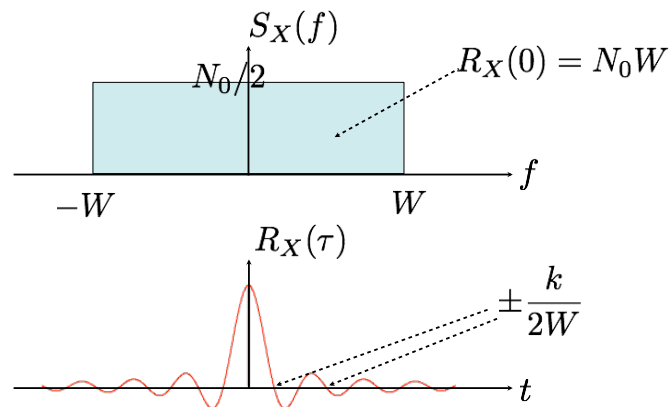
Using the PSD

- White noise process (again)
- Cross-correlation (again)
- System identification

54

White noise signal/process

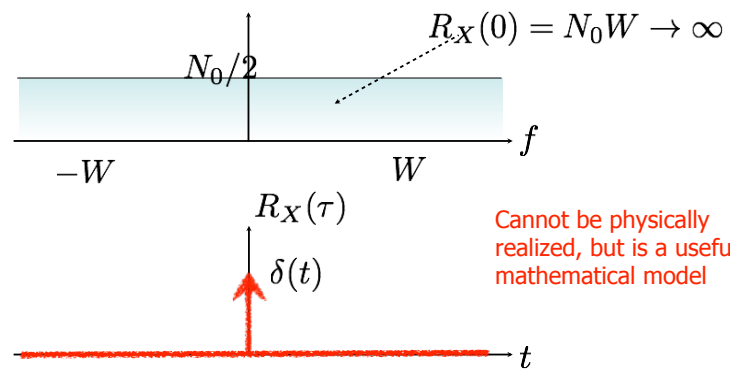
- Bandlimited white noise process with $\mu_X = 0$



55

White noise signal/process

- Let $W \rightarrow \infty$
White noise or uncorrelated noise process

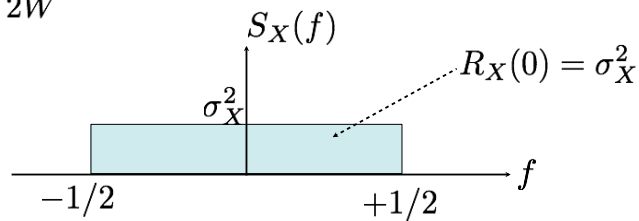


Cannot be physically realized, but is a useful mathematical model

56

White noise signal/process

- If we sample the time continuous white noise process at $f_s = 2W$

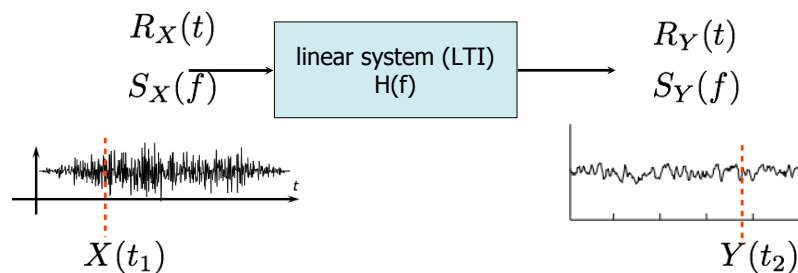


- This time discrete process **can** be realized:

$$R_X(k) = \begin{cases} \sigma_X^2 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

57

Cross correlation function



- What is the stochastic relation between $X(t_1)$ and $Y(t_2)$?

58

Cross correlation function

- Cross correlation function for a jointly WSS random process is:

$$\begin{aligned} R_{XY}(\tau) &= E[X(t)Y(t + \tau)] \\ &= R_{YX}(-\tau) \end{aligned}$$

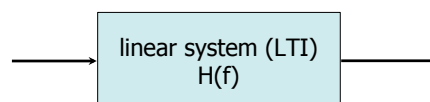
- We get in frequency domain:

$$\begin{aligned} R_{XY}(\tau) &= h(\tau) * R_X(\tau) \\ S_{XY}(f) &= H(f) S_X(f) \end{aligned}$$

cross power spectral density

59

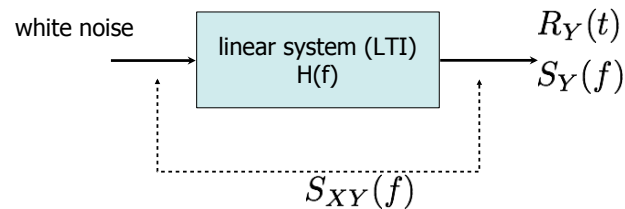
Application: system identification



- Find the impulse (or frequency) response $H(f)$ of the system

60

Application: system identification



- When feeding white noise into a linear system, we obtain

$$S_{XY}(f) = H(f)S_X(f) = \frac{N_0}{2}H(f)$$

61

Covered Today

- Chapter 11
- Key terms
 - Moving average, autoregressive processes
 - Autocorrelation for infinite impulse response
 - Power spectral density
 - Average power
 - White Gaussian noise process
 - Cross power spectral density
 - System identification

46