DIGITAL FILTER DESIGN

Digital Filter Design

I. Introduction

- Why Filter
- Pole/Zero placement

II. IIR Filter Design

- Impulse Invariance
- Bilinear transform

III. FIR Filter Design

- Window methods
- Parks-McClellan

IV. Statistical Filter Design

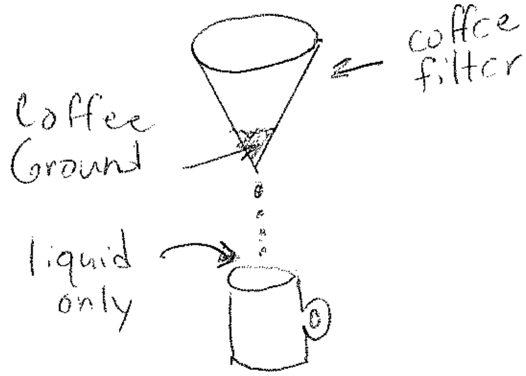
- Wiener Filters

V. Adaptive Filters

- LMS

WHY FILTERs?

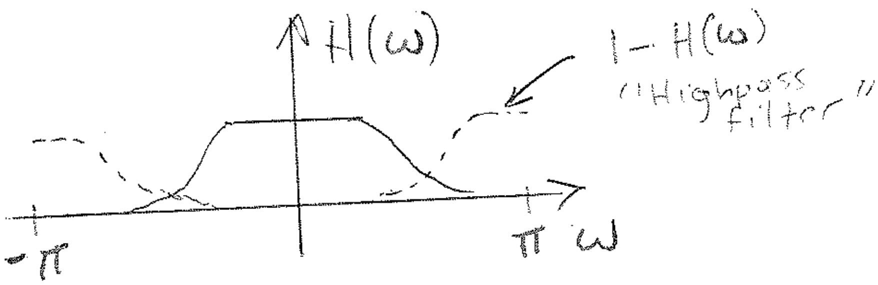
☆ To <u>separate</u> signal components.



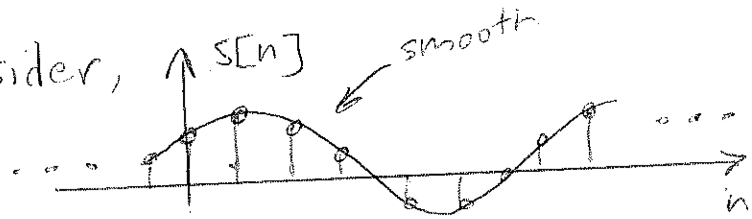
⇐ Separates components based on their particle size

Digital filters separate signal components based on their <u>frequency content</u>.

<u>Ex.</u> Lowpass filter

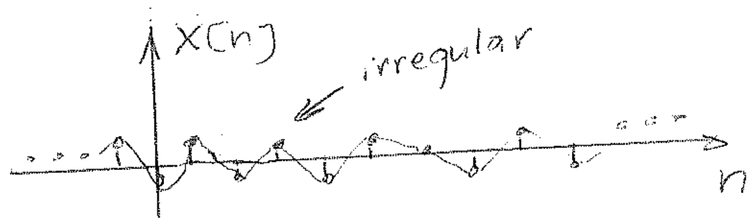


Consider,

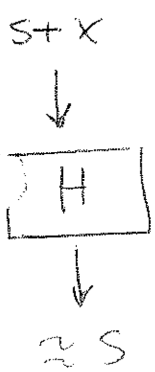


<u>signal</u>

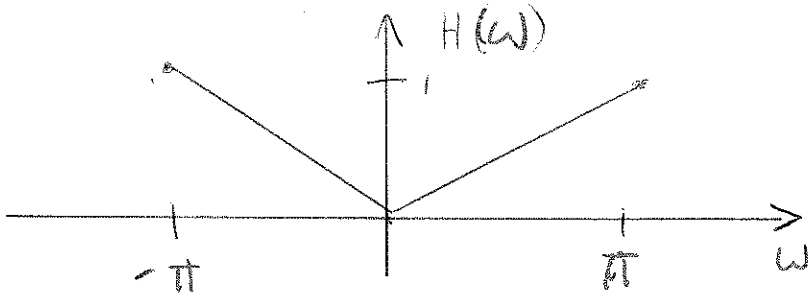
+



<u>noise</u>



ex "Differentiator"



Types of Filters :

Lowpass
Highpass
Bandpass
Bandstop

} All can be obtained from lowpass filters

IIR → Infinite Impulse Response

- h[n] infinite in extent
- H(z) has poles

FIR - Finite Impulse Response

- h[n] finite in extent
- H(z) has only zeros

Simple Filter Design by Pole-Zero Placement

Linear Difference Equation Filter:

$$y[n] = \sum_{k=0}^{M} b_k x[n-k] - \sum_{k=1}^{N} a_k y[n-k]$$

Transfer Function:

$$Y(z) = \sum_{k=0}^{M} b_k z^{-k} X(z) - \sum_{k=1}^{N} a_k z^{-k} Y(z)$$

$$\longrightarrow H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{N} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}}, \quad a_0 = 1$$

Factor numerator and denominator polynomials

$$\longrightarrow H(z) = \frac{b_0}{a_0} \frac{\prod_{k=1}^{M} (1 - c_k z^{-k})}{\prod_{k=1}^{N} (1 - d_k z^{-k})} = \frac{\text{"zeros"}}{\text{"poles"}}$$

$$H(\ ) = \frac{b_0}{a_0} \frac{\prod_{k=1}^{M} (1 - c_k e^{-j2\pi F k})}{\prod_{k=1}^{N} (1 - d_k e^{-j2\pi F k})}$$

☆ By cleverly placing poles and zeros we can obtain a desired frequency response

Ex. Remove a certain frequency component

For example, 60 Hz "line noise"

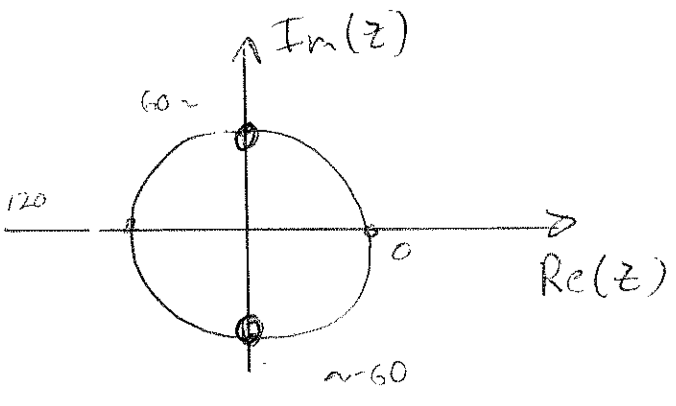
$$x(t) = s(t) + \sin(2\pi(60)t)$$

Filter $x(t)$ to remove 60 Hz component.

Sample $x(t)$ at $\quad = 240$ Hz.

$$\longrightarrow x[n] = s[n] + \sin(2\pi \tfrac{60}{\quad} n)$$

Digital Filter Design I :



$$= \frac{60}{240} =$$

Place zeros @

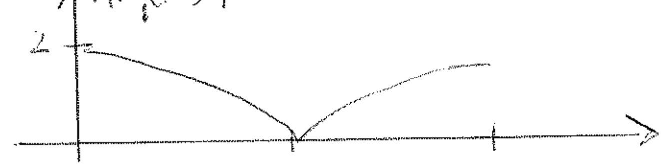
$= \quad , \quad =$

$$\longrightarrow z_1 = j \ , \ z_2 = -j$$

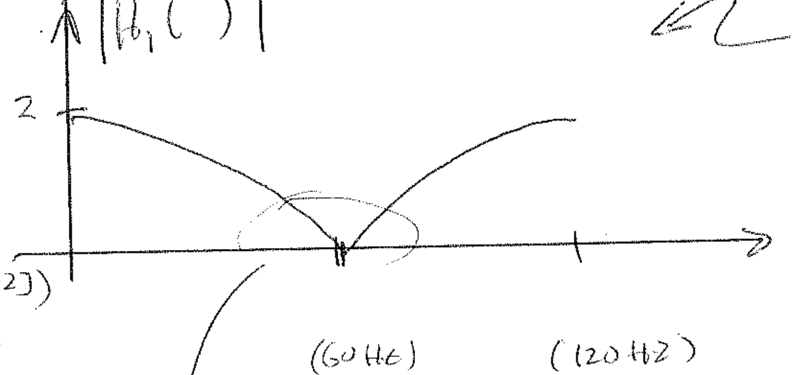
$$H_1(z) = (1 - jz^{-1})(1 + jz^{-1})$$

$$= 1 + z^{-2}$$

$$\longrightarrow h_1[n] = \delta[n] + \delta[n-2] \longrightarrow y[n] = x[n] + x[n-2]$$



H_1 does the job, but also attenuates frequencies close to 60 Hz

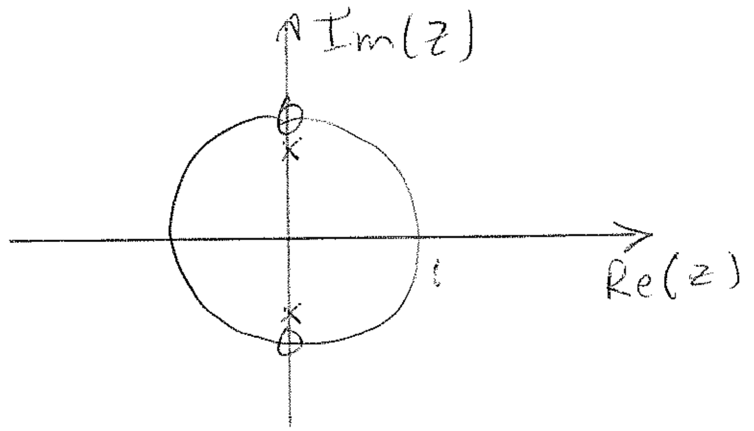


zeros are "dragging" frequency response down

reduces or "attenuates" other frequencies near 60 Hz

★ Note: DC Gain of $H_1(F)$ is 2. Therefore, use $h_1[n] = \frac{1}{2}(\delta[n] + \delta[n-2])$ to get dc gain 1.

Digital Filter Design II



Add poles close to zeros. Response still must go to zero at 60 Hz, but as we move away from 60 Hz pole offsets the effect of the zero.

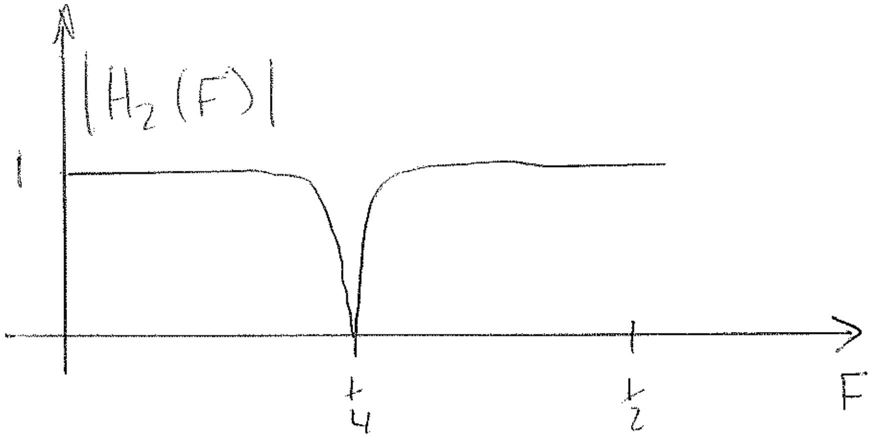
$P_1 = 0.9j$, $P_2 =$

$$H_2(z) = \frac{(1-jz^{-1})(1+jz^{-1})}{(1-0.9jz^{-1})(1+0.9jz^{-1})} = \frac{1+z^{-2}}{1+0.81z^{-2}}$$

$\longrightarrow$ $y[n] + 0.81\,y[n-2] = x[n] + x[n-2]$

$$\boxed{y[n] = x[n] + x[n-2] - 0.81\,y[n-2]}$$

Note: multiply $H(z)$ by $\frac{1.81}{2}$ to get unity gain at DC!



☆ Much closer to "ideal" notch at 60 Hz

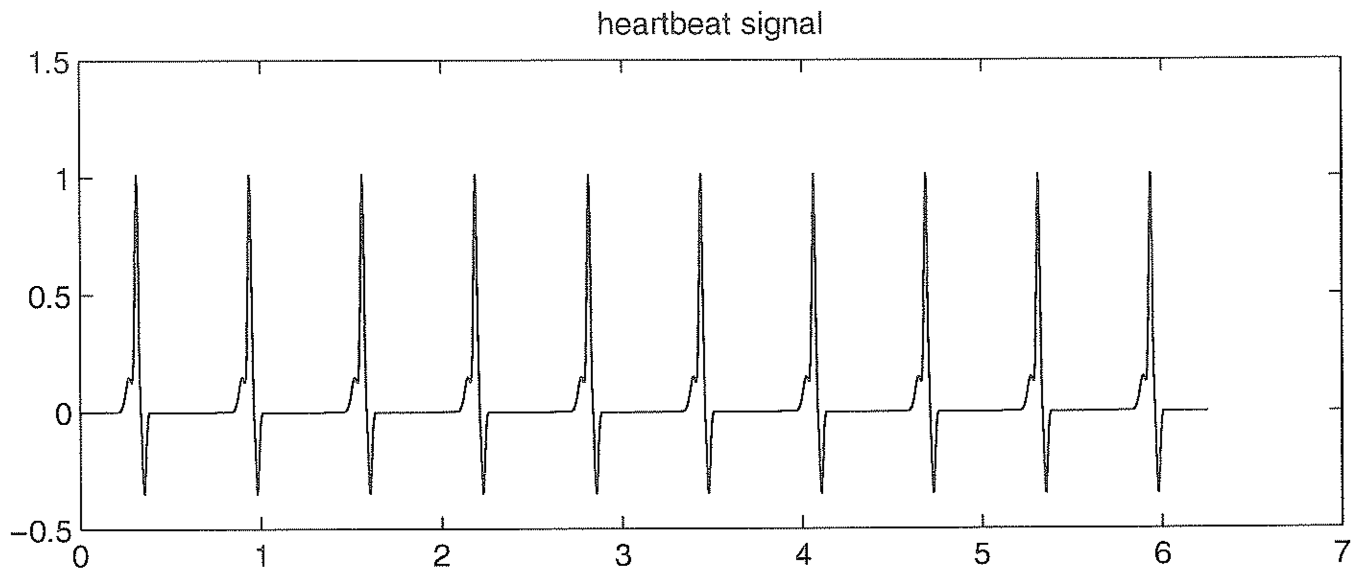
- $h_1[n] = \delta[n] + \delta[n-2]$ is FIR
- $h_2[n]$ impulse response $x[n] = \delta[n]$

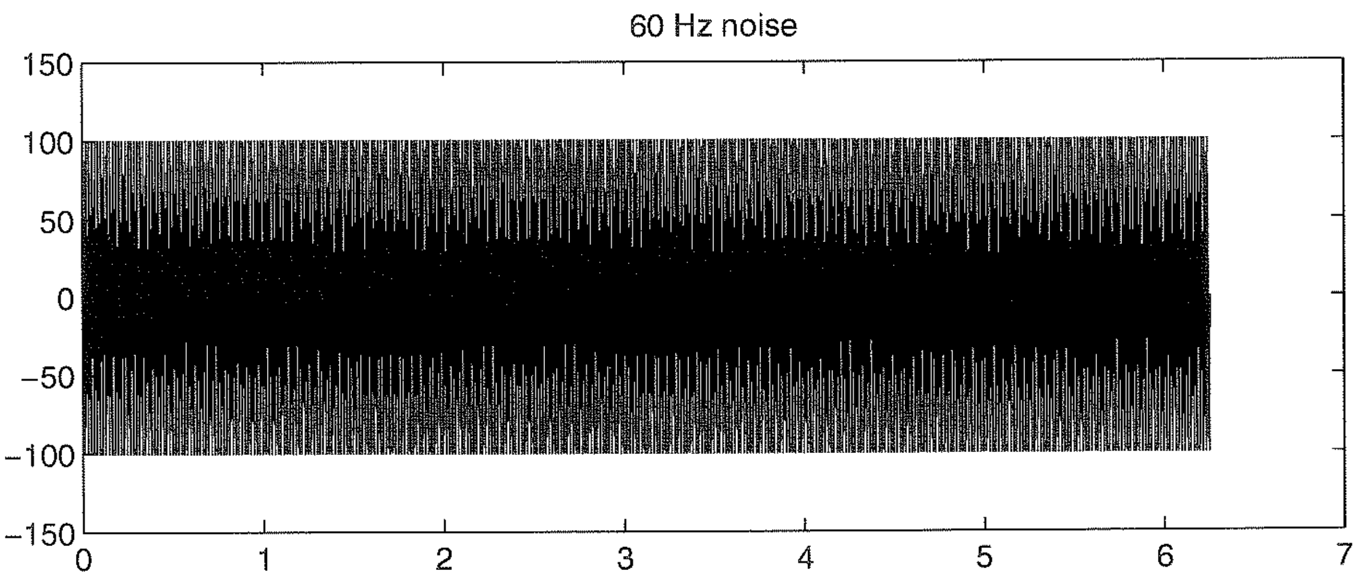
  $y[n] = x[n] + x[n-2] - 0.81\, y[n-2]$

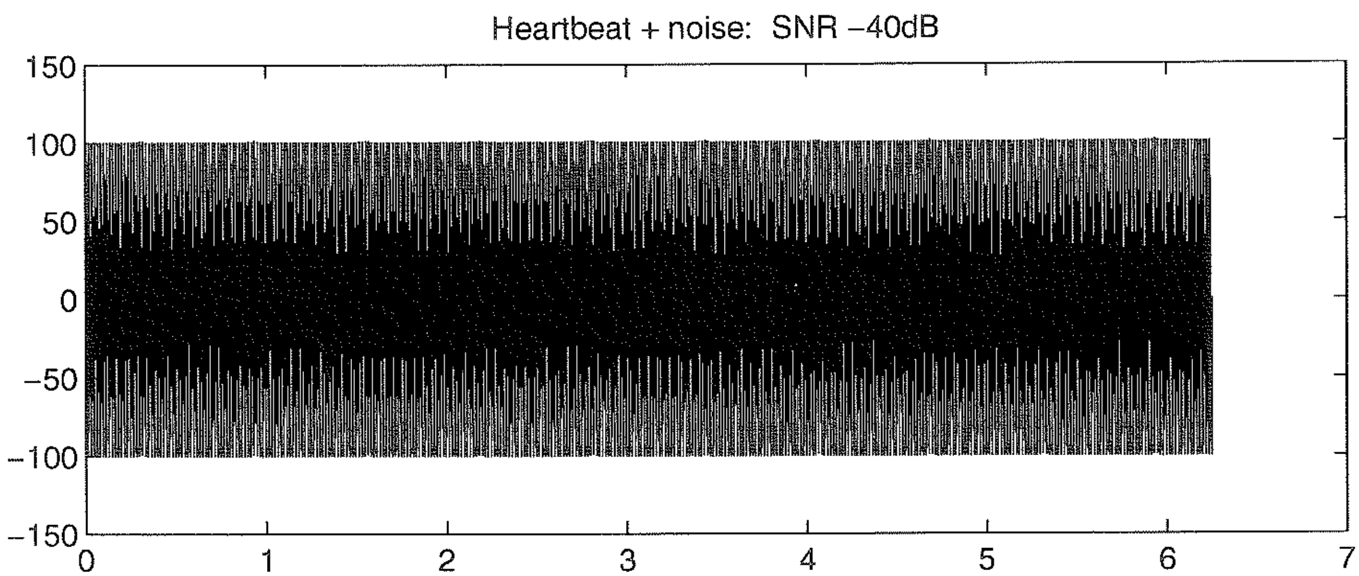
  IIR (feedback) ⟶
- Conjugate symmetry of poles/zeros guarantees real-valued digital filter

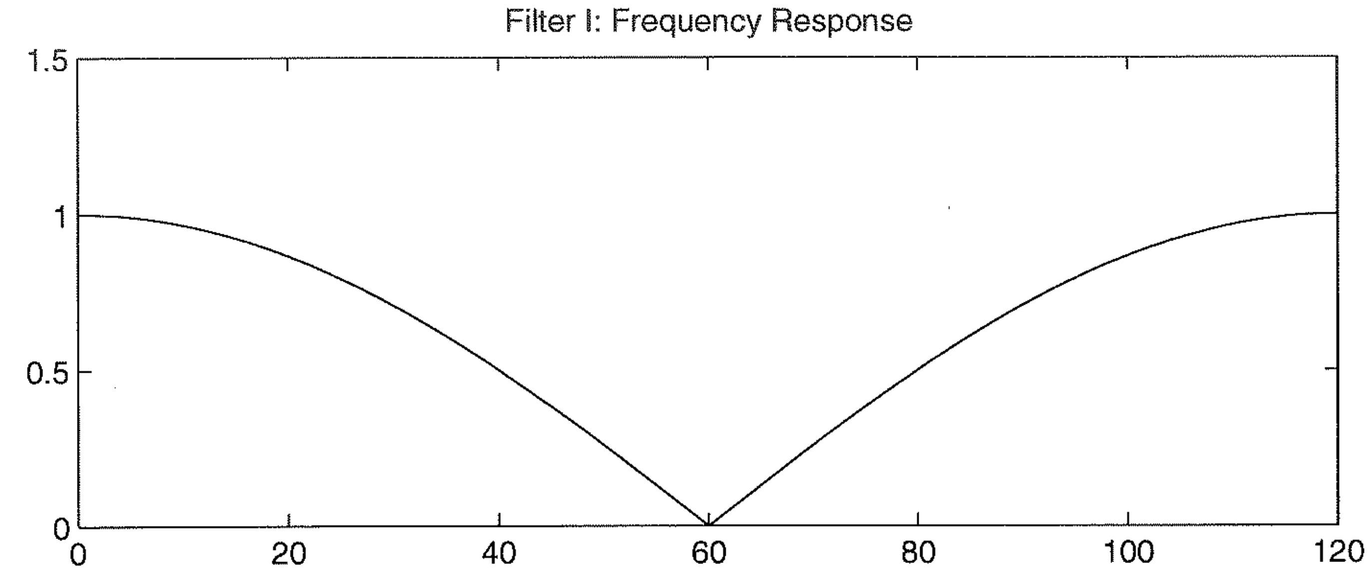
This simple approach of placing poles and zeros can work, but is is sort of a trial and error approach.

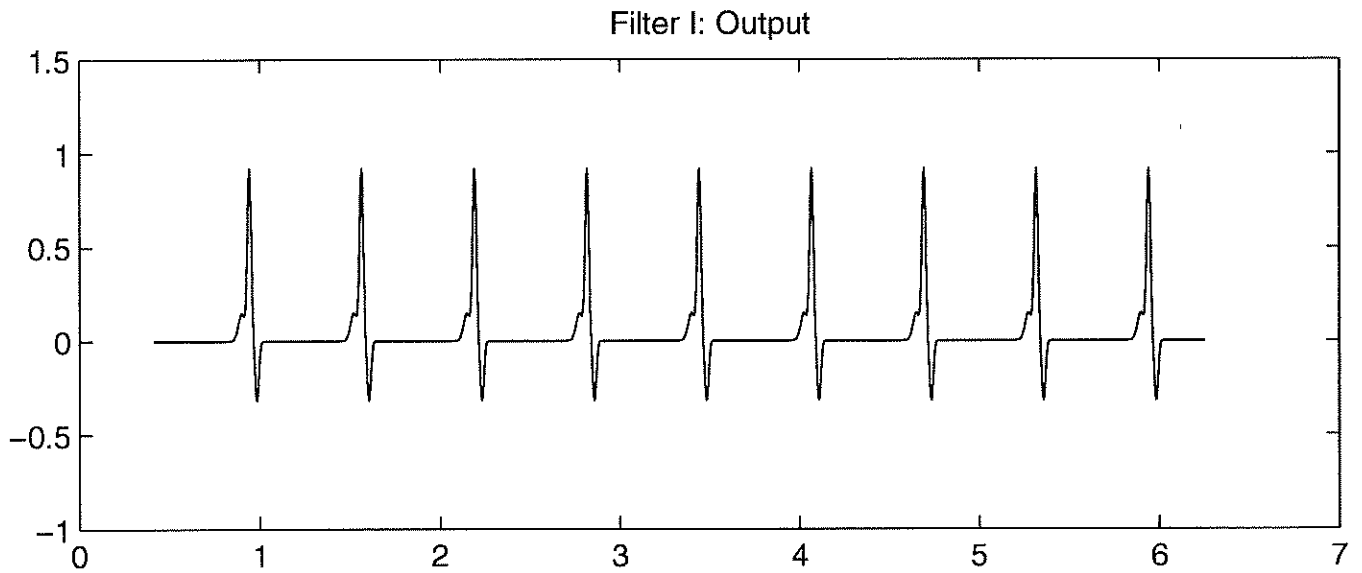
We can do better!

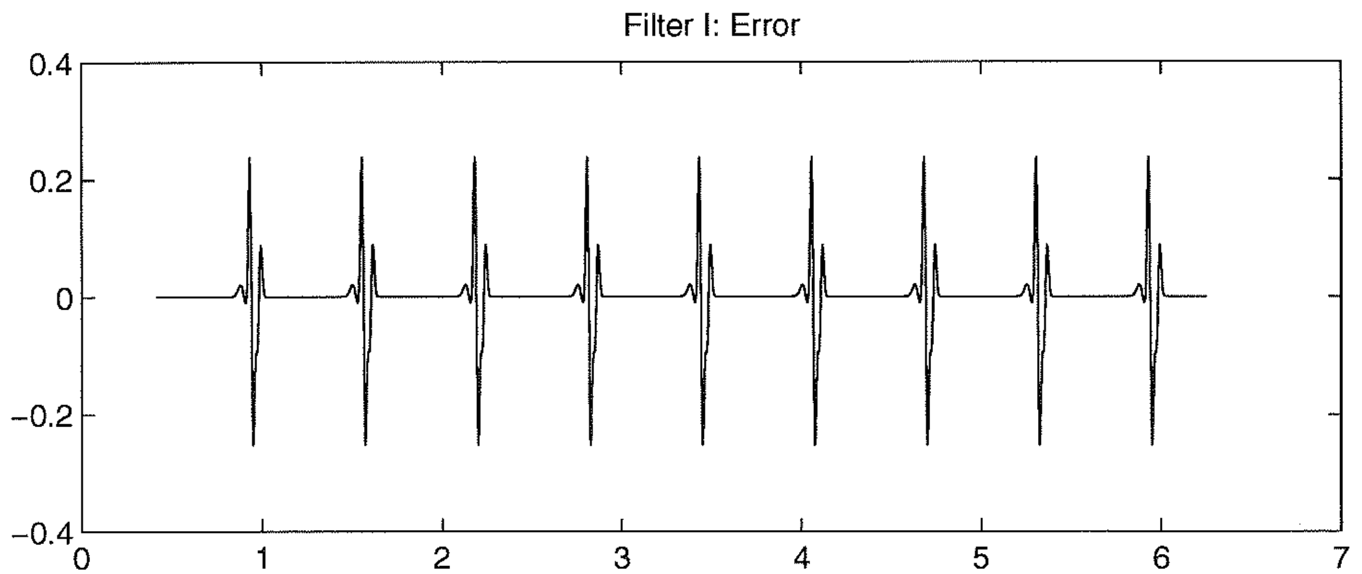

heartbeat signal
1.5
1
0.5
0
−0.5
0
1
2
3
4
5
6
7


60 Hz noise
150
100
50
0
−50
−100
−150
0
1
2
3
4
5
6
7

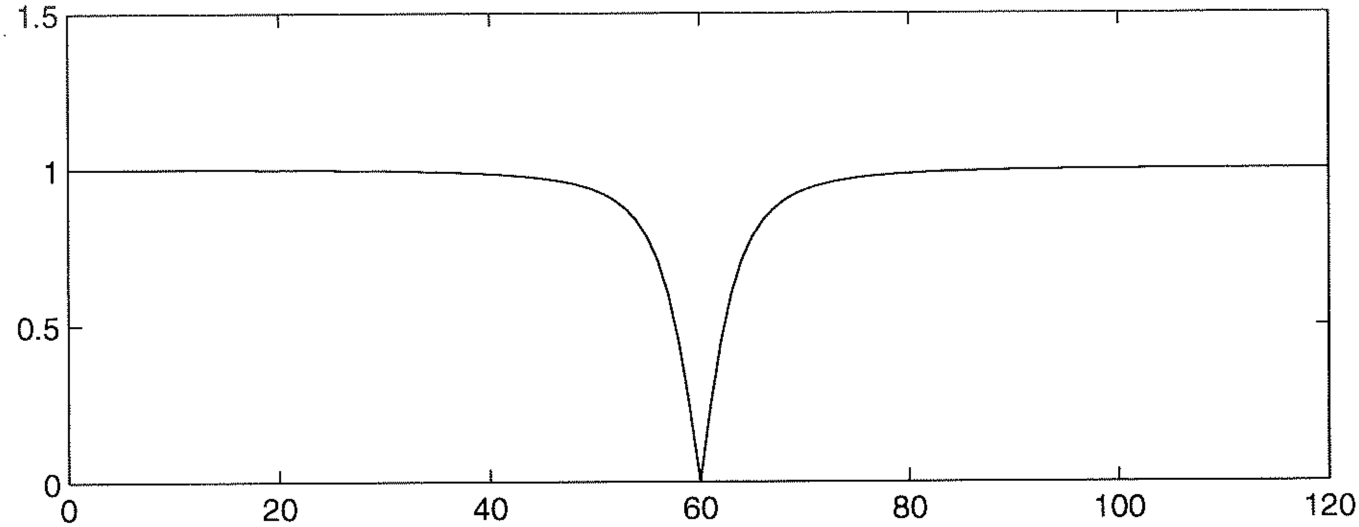

Heartbeat + noise: SNR −40dB
150
100
50
0
−50
−100
−150
0
1
2
3
4
5
6
7


Filter I: Frequency Response
1.5
1
0.5
0
0
20
40
60
80
100
120

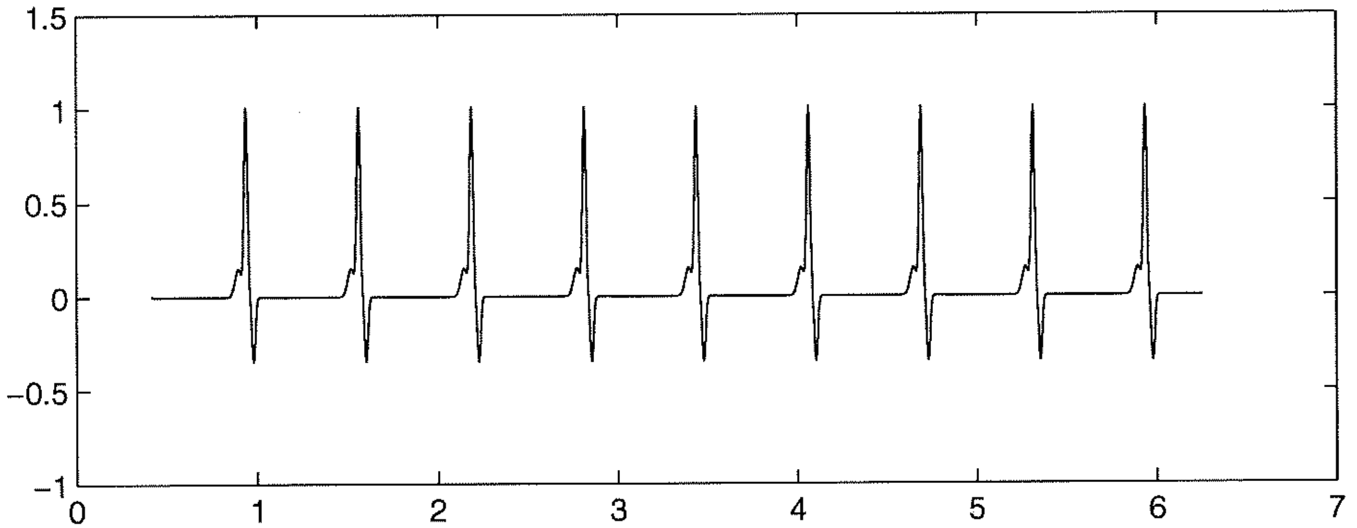

Filter I: Output
1.5
1
0.5
0
−0.5
−1
0
1
2
3
4
5
6
7


Filter I: Error
0.4
0.2
0
−0.2
−0.4
0
1
2
3
4
5
6
7

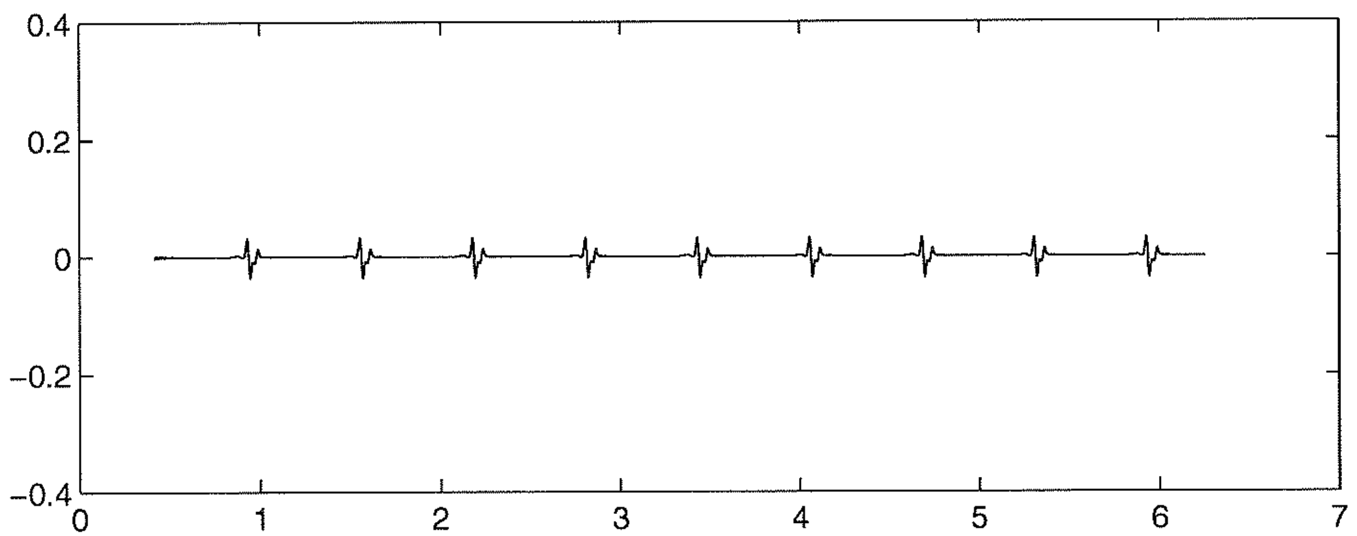
Filter II: Frequency Response



Filter II: Output



Filter II: Error

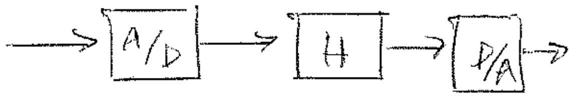


Given that we want to separate a signal into frequency components, we would clearly like to do the <u>BEST JOB POSSIBLE</u>

$\Rightarrow$ **Filter Design / Engineering**

Usually we will have <u>specifications</u> that we need to meet.

<u>Example</u>: Lowpass filter
Specifications in CT.



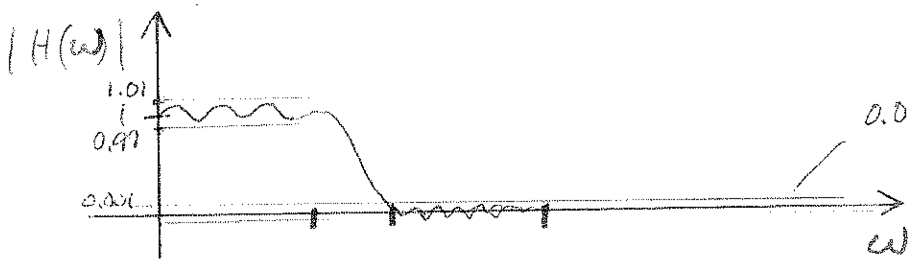
<u>Passband:</u> $|H_{eff}(\Omega)|$ within ± 0.01 of unity gain in the range

$$0 \leq f \leq 2\text{kHz} \longrightarrow 0 \leq \Omega \leq \_\_\_\_$$

<u>Stopband:</u> $|H_{eff}(\quad)| \leq 0.001$ in the range

$$f \geq 3\text{kHz} \longrightarrow \Omega \geq \_\_\_\_$$

Suppose $T = 10^{-4}\text{s} \rightarrow f_s = 10\text{kHz} \Rightarrow \Omega_s = \_\_\_\_$
Draw specs in digital freq.



IIR Filter Design By Impulse Invariance

★ Traditional approach to IIR Filter transforms CT filters (Butterworth, Chebyshev) to DT Filters. Why?

1. Continuous-time filter design is advanced, mature, well-understood.

2. We get elegant closed-form formulas.

Properties we desire of the transform:

1. It should map imaginary axis of the s-plane (Laplace) to the unit circle of the z-plane

   (Recall, evaluating Laplace transform $H(s)$ along $j\omega$-axis produces the CTFT)

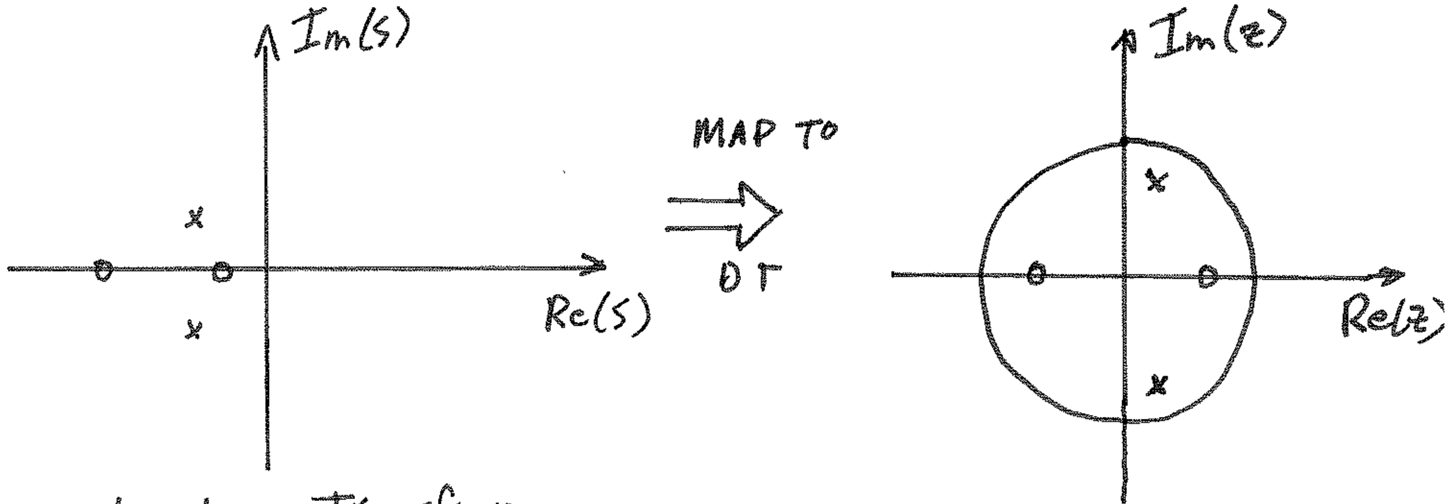
2. A stable CT filter should transform to a stable DT filter.

Design Steps:

1. Translate DT specs to CT filter
2. Design CT prototype filter
3. Transform filter to discrete-time

CT Filter

DT Filter



Laplace Transform
s-plane, LHP stability
$H_c(s)$, $H_c(\Omega)$

z-transform
z-plane, unit circle
$H(z)$, $H(\omega)$

$$h_c(t) \xrightarrow{\text{impulse invariance}} h[n] = h_c(nT)$$

$$H_c(s) \xrightarrow{\text{bilinear transform}} H(z) = H_c(s)\Big|_{s=2\frac{z-1}{z+1}}$$

Key: Convert from rectangular coordinates 
to polar coordinates 

Impulse Invariance Design

Easy - we simply sample the CT filter's impulse response.

$$h[n] = T_d \, h_c(nT_d)$$

← a sampling period. (need not be the same as T used in A/D and D/A)

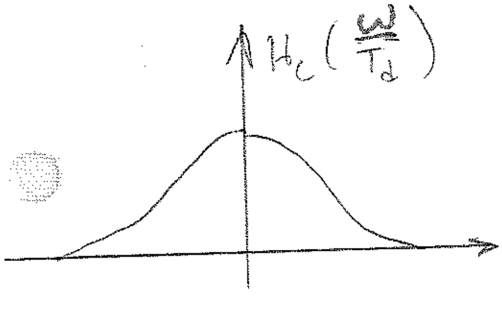
$h[n]$ is a sampled version of $h_c(t) \Rightarrow$ aliasing in frequency domain.

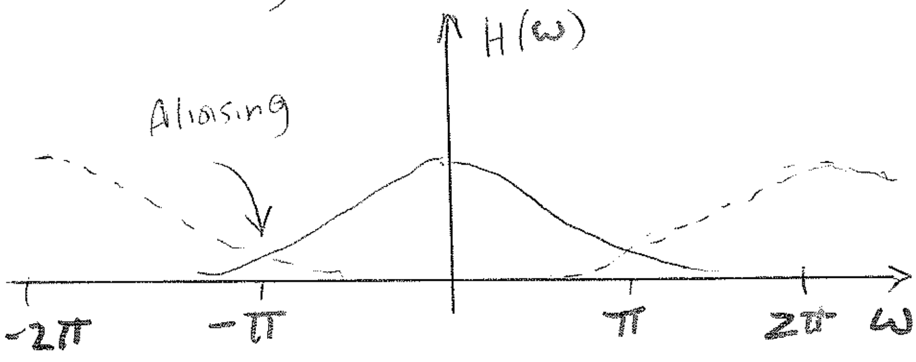
$$H(\omega) = \sum_{k=-\infty}^{\infty} H_c\left(\frac{\omega + 2\pi k}{T_d}\right)$$

If H_c is bandlimited, with $H_c(\Omega) = 0$, $|\Omega| \geq \frac{\pi}{T_d}$, then

$$H(\omega) = H_c\left(\frac{\omega}{T_d}\right)$$

However, this is impossible, since any practical CT filter cannot be perfectly bandlimited. ⇒ Aliasing





Aliasing $\Rightarrow$ Need to <u>over</u>design CT prototypes

- Given an CT transfer function $H_c(s)$, what is the corresponding DT transfer function?

Expand $H_c(s)$ in partial fractions

$$H_c(s) = \sum_{k=1}^{N} \frac{A_k}{s - s_k} \qquad \text{poles @ } s = s_k$$

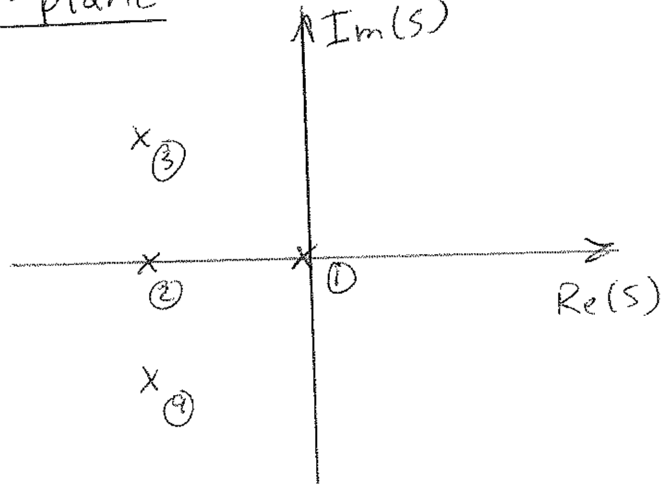
$$\Rightarrow h_c(t) = \sum_{k=1}^{N} A_k e^{s_k t} u(t) \qquad \left(\begin{array}{c}\text{Inverse}\\ \text{Laplace}\\ \text{Transform}\end{array}\right)$$

$$\Rightarrow h[n] = T_d \, h_c(nT_d) = T_d \sum_{k=1}^{N} A_k e^{s_k n T_d} u[n]$$

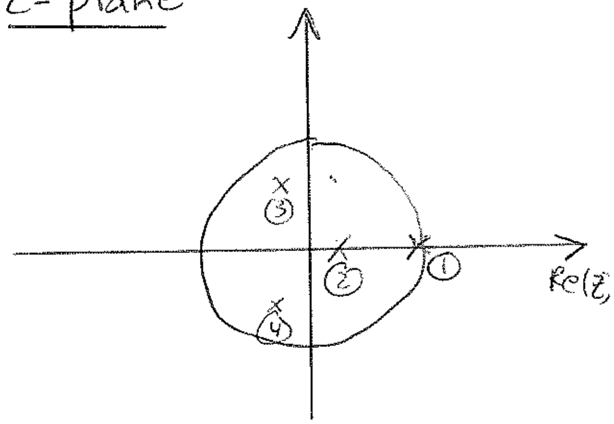
$$= T_d \sum_{k=1}^{N} A_k \underbrace{\left(e^{s_k T_d}\right)^n}_{\text{"}a^n\text{"}} u[n]$$

$$\Rightarrow \boxed{H(z) = \sum_{k=1}^{N} \frac{T_d A_k}{1 - e^{s_k T_d} z^{-1}}} \qquad \text{poles @ } z = e^{s_k T_d}$$

<u>S-plane</u>



<u>Z-plane</u>



IIR Filter Design by IMPULSE INVARIANCE

5 GOLDEN STEPS

1. Given specs for $H(\omega)$, choose T_d (arbitrary) & translate specs to $H_c(\omega/T_d)$ "=" $H(\omega)$

2. Design $H_c(s)$: Butterworth, Chebyshev, ...

3. Find $H_c(s)$ & do partial fractions (get poles, s_k)

4. Transform poles through

$$s_k \longrightarrow e^{s_k T_d}$$

s-plane   z-plane

$$\frac{A_k}{s - s_k} \longrightarrow \frac{T_d A_k}{1 - e^{s_k T_d} z^{-1}}$$

Compute $H(z)$

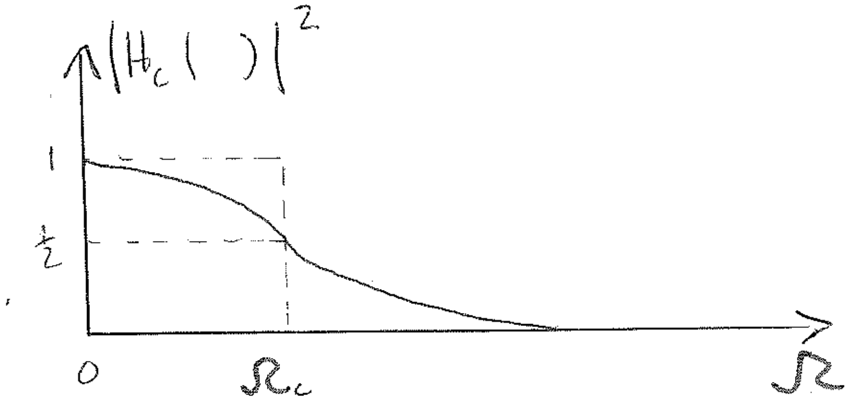
5. Check performance of $H(\omega)$. If not up to spec, redesign $H_c(s)$ in step 2 with more conservative specs.

Continuous-Time Butterworth Filters

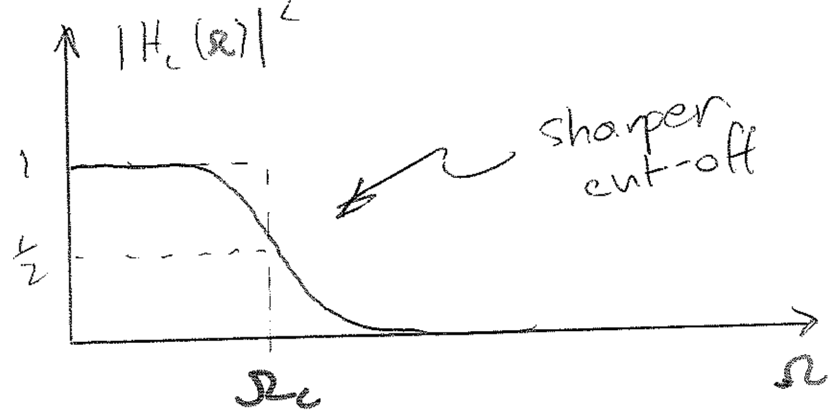
* Appendix B in OSB

- Magnitude response is <u>maximally flat</u> in the passband
- Magnitude response is <u>monotone</u>

$$|H_c(\Omega)|^2 = \frac{1}{1 + (\Omega/\Omega_c)^{2N}}, \quad \Omega_c \sim \text{half power frequency}$$



What happens as N increases?



In the s-plane:

$$H_c(s)\,H_c(-s) = \frac{1}{1 + \left(s/j\Omega_c\right)^{2N}}$$

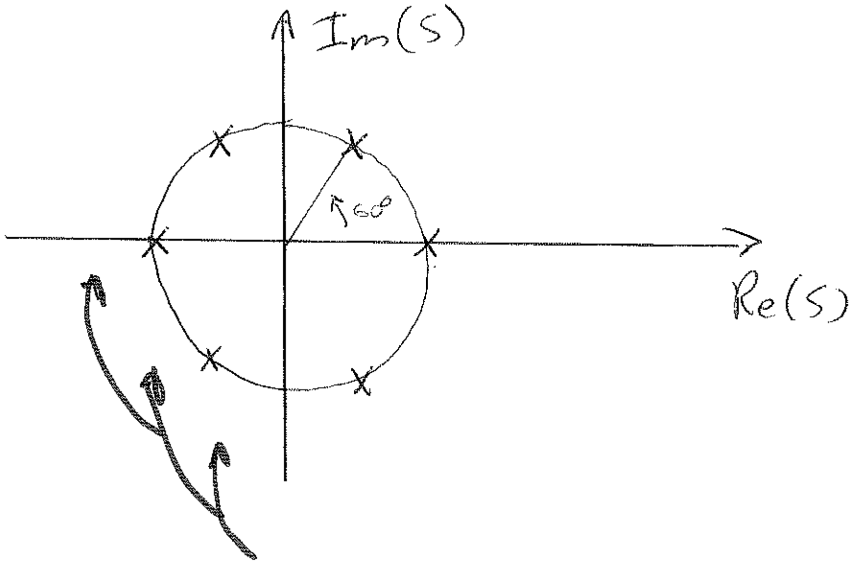
$\Rightarrow$ poles are at

$$s_k = (-1)^{\frac{1}{2N}}\,(j\Omega_c) = \Omega_c\, e^{j\frac{\pi}{2N}(2k+N-1)}$$

$k = 0, 1, \ldots, 2N-1$

$\Rightarrow$ poles lie on a circle of radius Ω_c

ex. $N = 3$ (Third order filter)



* For a stable, causal filter, we take only the left half plane poles.

* See Matlab function
butter(___, 's')

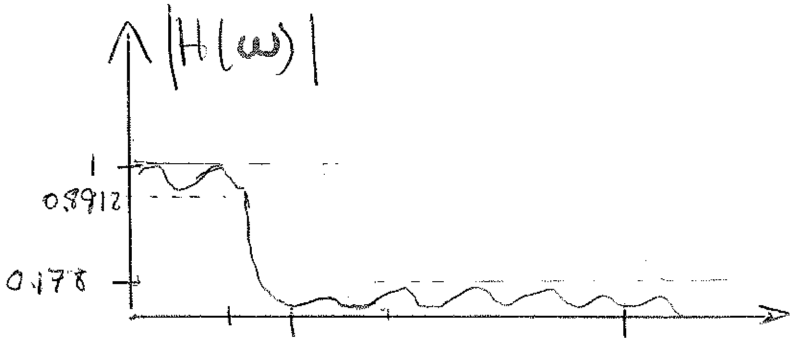
Design Example : Impulse Invariance

Specs on H( ):

Passband: $0.89125 \le |H(\omega)| \le 1,$

$0 \le |\omega| \le 0.2\pi$

Stopband: $|H(\omega)| \le 0.17783 \;,\; 0.3\pi \le |\omega| \le \pi$



STEP 1 : Set $T_d = 1$

$\Rightarrow$ Specs for $H_c(\Omega)$ are identical to those for $H(\omega)$.

Step 2: Design $H_c(\Omega)$. Choose Butterworth.

What are design parameters?

Ω_c, N

Choose to meet passband constraint exactly.

Since magnitude of Butterworth is monotonic if freq

(1) $|H_c(\Omega)| \geq 0.89125$

$$\rightarrow \frac{1}{1+\left(\frac{0.2\pi}{\Omega_c}\right)^{2N}} = (0.89129)^2$$

(using monotonicity property of Butterworth)

(2) $|H_c(\Omega)| \leq 0.17783$

$$\rightarrow \frac{1}{1+\left(\frac{0.3\pi}{\Omega_c}\right)^{2N}} = (0.17783)^2$$

(1) & (2) $\longrightarrow$ $N = 5.8858$, $\Omega_c = 0.70474$

$\rightarrow$ round N to nearest integer

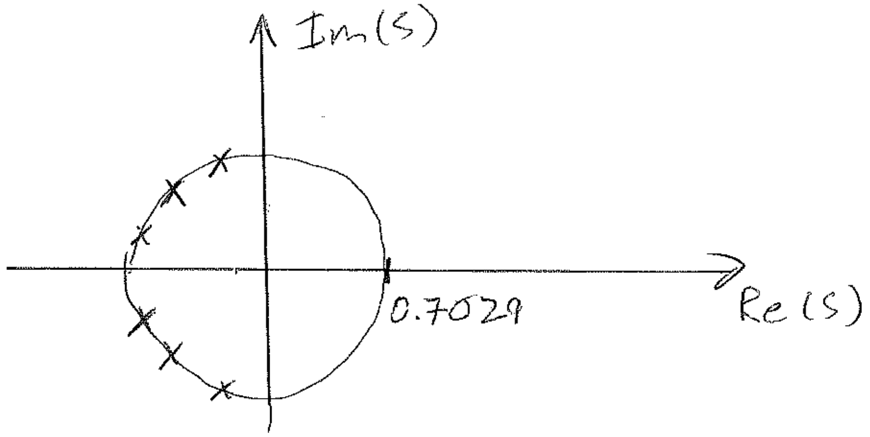
$\rightarrow$ $N = 6$ (3 pole pairs)

Pole pair 1: $s_1, s_2 = -0.182 \pm j(0.679)$

Pole pair 2: $s_3, s_4 = -0.497 \pm j(0.497)$

Pole pair 3: $s_5, s_6 = -0.679 \pm j(0.182)$

STEP 3: Where are the poles?



$$|s_k| = 0.7029$$

$$H_c(s) = \frac{0.1209}{(s^2+0.3640s+0.4945)(s^2+0.995+0.49)(s^2+1.385s+0.49)}$$

Step 4:

Partial fraction expansion and transform

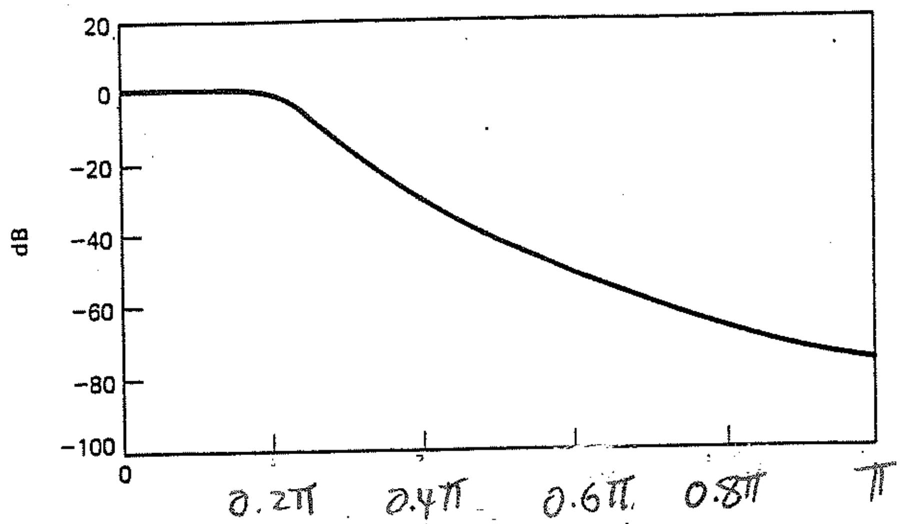
$$\frac{A_k}{s-s_k} \longrightarrow \frac{T_d A_k}{1-e^{s_k T_d} z^{-1}}$$

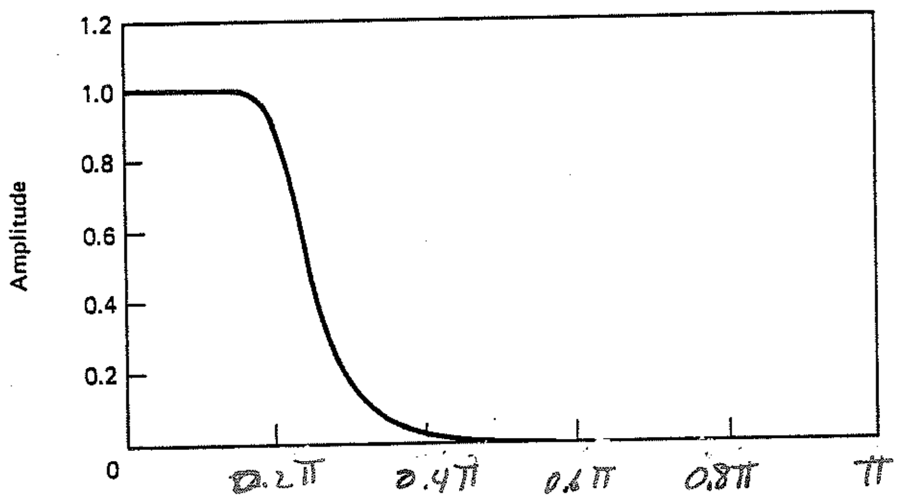
$\Longrightarrow$

$$H(z) = \frac{0.2871-0.4466z^{-1}}{1-1.2971z^{-1}+0.6949z^{-2}} + \frac{-2.1428+1.1455z^{-1}}{1-1.0691z^{-1}+0.3699z^{-2}}$$

$$+ \frac{1.857-0.6303z^{-1}}{1-0.9972z^{-1}+0.2570z^{-2}}$$

STEP 5: Plot H( ) and verify it meets specs.





Are specs met?

What if they weren't met?

Why wouldn't they by met?

Impulse Invariance

* Motivated by a desire to match the impulse response ⟹ if CT filter is bandlimited, then approximation will be close in the frequency domain.

Problem: Aliasing

⟹ Difficult to control errors in design process

Q: { How to design a highpass or band stop filter this way?

In practice, II design is not used very often due to these difficulties.

IIR FILTER DESIGN BY BILINEAR TRANSFORMATION

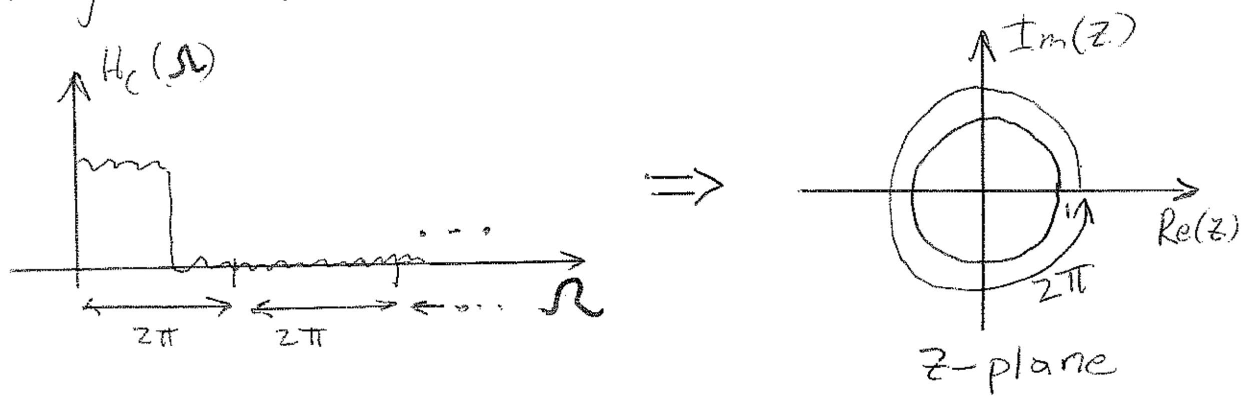
Recall: <u>Impulse Invariance design</u>

maps poles & zeros in s-plane $(\sim H_c(\Omega))$

poles & zeros in z-plane $(\sim H(\omega))$

problem: $H(\omega)$ is an <u>aliased</u> version of $H_c(\Omega)$.

Why: Because we get $H(\omega)$ by wrapping $H_c(\Omega)$ around the unit circle (an infinite # of times) and adding the result together.



The <u>Bilinear Transformation</u> avoids aliasing by compressing the <u>entire Ω-axis</u> ($j\Omega$-axis in s-plane) into one revolution of the unit circle, through a

<u>FREQUENCY AXIS WARPING</u>

BILINEAR TRANSFORMATION (BLT)	$s = 2\left(\frac{z-1}{z+1}\right)$
INVERSE BLT	$z = \frac{1+s/2}{1-s/2}$

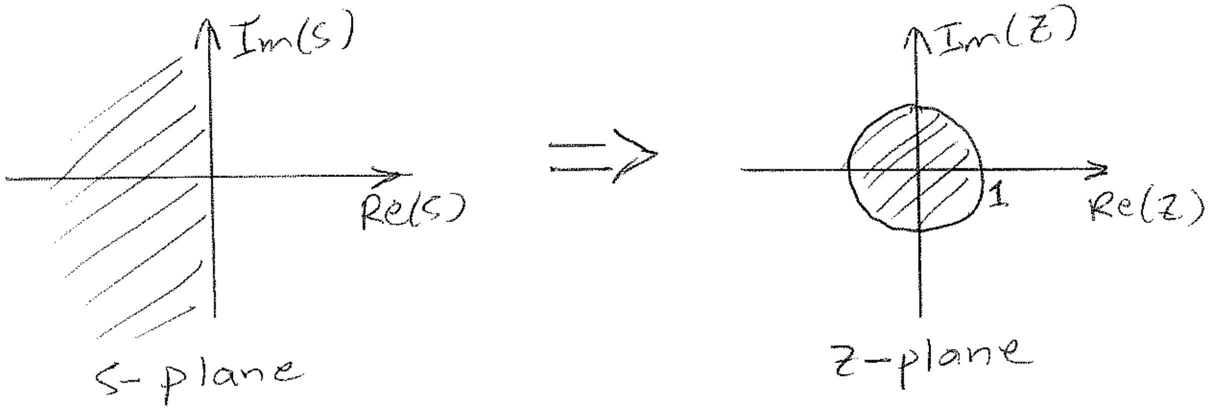
Substitute $s = \sigma + j\Omega$

$$z = \frac{1 + \sigma/2 + j\Omega/2}{1 - \sigma/2 + j\Omega/2}$$

Property 1: If $\sigma < 0$, then $|z| < 1$

If $\sigma > 0$, then $|z| > 1$.

$\Rightarrow$



Left-half-plane $\Rightarrow$ inside unit circle

stability $\Rightarrow$ stability.

Property 2:

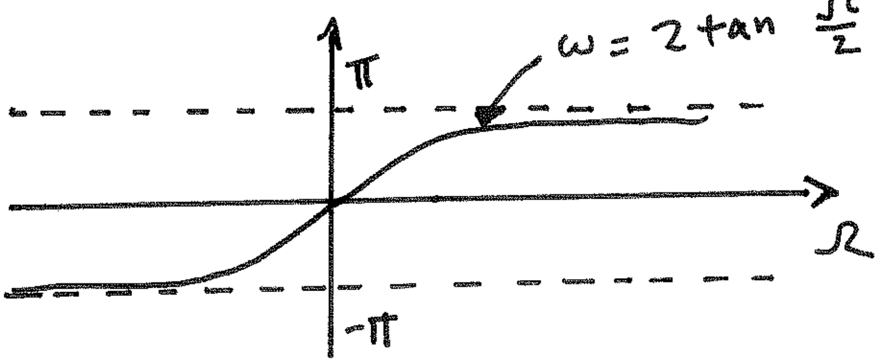
BLT Maps $j\Omega$-axis of s-plane to unit circle in z-plane.

Substitute $s = j\Omega$

$$\longrightarrow \quad z = \frac{1+j\Omega}{1-j\Omega} \quad \Rightarrow \quad e^{j\omega} = \frac{1+j\Omega}{1-j\Omega}$$

$$\longrightarrow \quad j\Omega = \frac{1-e^{-j\omega}}{1+e^{-j\omega}} = \frac{2je^{-j\omega/2}\sin \omega/2}{2e^{-j\omega/2}\cos \omega/2}$$

$$\Longrightarrow \quad \boxed{\Omega = 2\tan \omega/2}$$



★ maps entire $j\Omega$-axis $(-\infty, \infty)$ to 1-period ω-axis $[-\pi, \pi]$

KEY: Design a CT Filter $H_c(s)$ → $H(\omega) = H_c(2\tan\frac{\omega}{2})$

substitute $H(z) = H_c(s)\Big|_{s=2\left(\frac{z-1}{z+1}\right)}$

- Where does BLT map <u>poles</u> and <u>zeros</u>?

Given an s-plane zero or pole s_k, it is mapped to

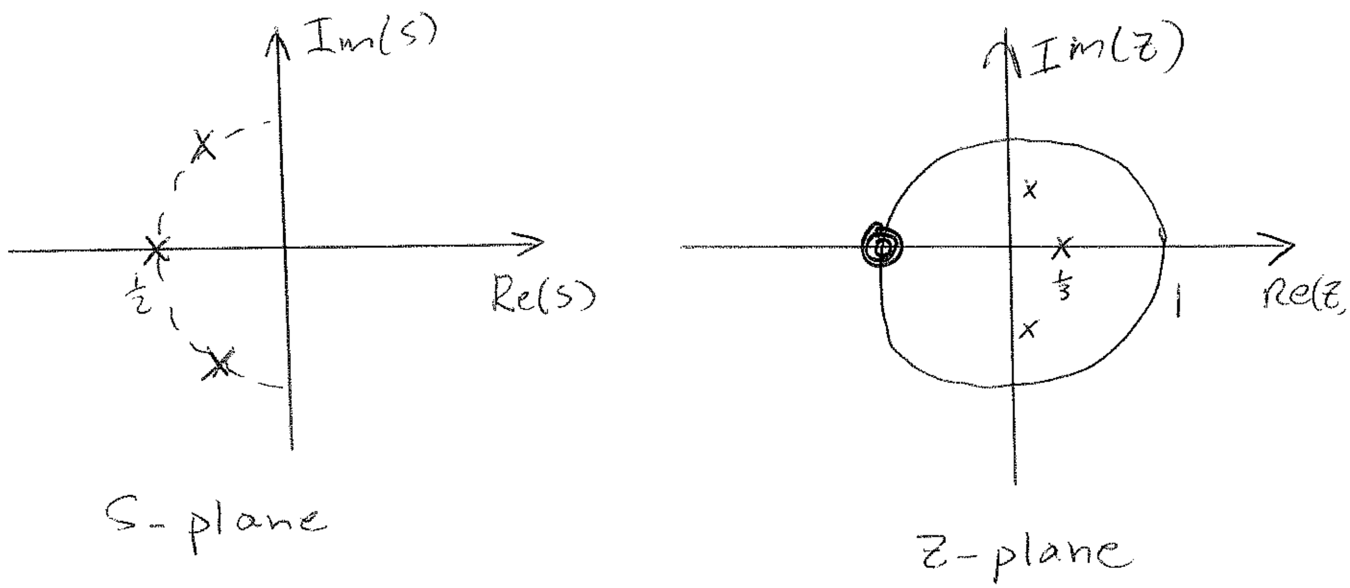
$$z_k = \frac{1 + s_k/2}{1 - s_k/2}$$

<u>ex.</u>

$$H_c(s) = \frac{1}{(s-s_1)(s-s_2)(s-s_3)}$$

will be mapped to

$$H(z) = \frac{\text{Const.}\cdot(z+1)^3}{\left(z-\left(\frac{1+s_1/2}{1-s_1/2}\right)\right)\left(z-\left(\frac{1+s_2/2}{1-s_2/2}\right)\right)\left(z-\left(\frac{1+s_3/2}{1-s_3/2}\right)\right)}$$



ex. Consider a single pole:

$$H_c(s) = \frac{1}{(s - s_i)}$$

$$H(z) = H_c(s)\Big|_{s = 2\left(\frac{z-1}{z+1}\right)}$$

$$= \frac{1}{\left(2\frac{z-1}{z+1} - s_i\right)}$$

$$= \frac{z+1}{\left(2(z-1) - s_i(z+1)\right)}$$

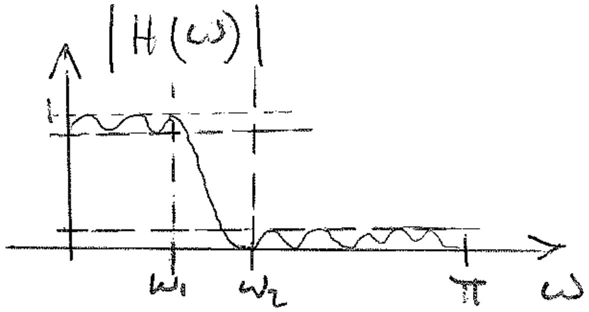
$$= \frac{(z+1)}{(2-s_k)z - (2+s_k)}$$

$$= \frac{\left(\frac{1}{2-s_k}\right)(z+1)}{z - \left(\frac{1+s_k/2}{1-s_k/2}\right)}$$

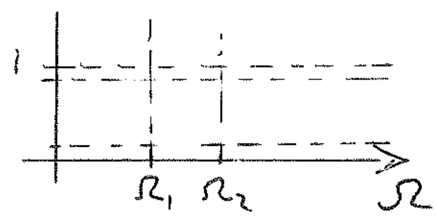
IIR FILTER DESIGN BY BLT

4 Golden Steps

DT specs



① Prewarp to obtain CT specs



② Design CT filter to meet these specs

③ Get $H_c(s)$

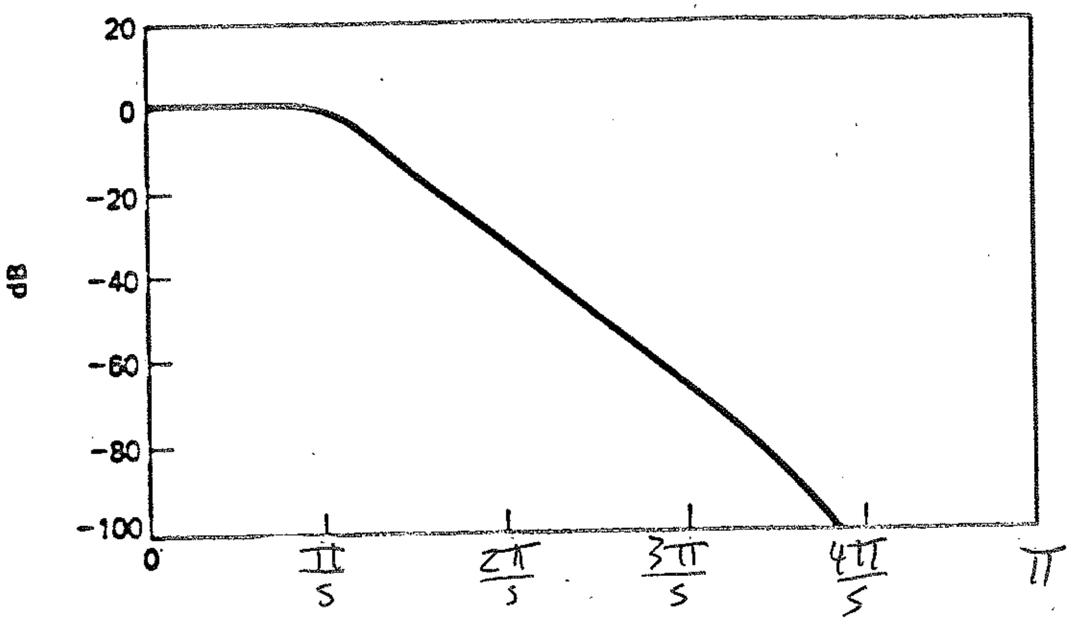
④ BLT : $H(z) = H_c(s)\Big|_{s = 2\left(\frac{z-1}{z+1}\right)}$

or re-map poles & zeros

pole or zero → $z_k = \dfrac{1 + s_k/2}{1 - s_k/2}$ ← pole or zero

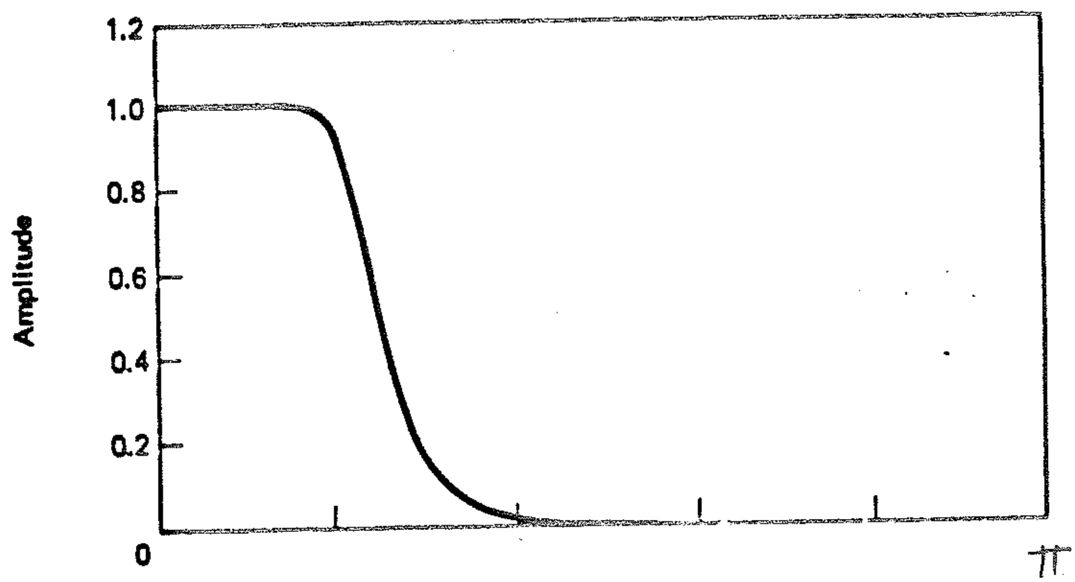
Ex. Butterworth Design in II example using BLT instead.

Resulting Filter:



Digital Freq.

* Compare with II design and be able to explain the differences



Digital freq.

Advantages of BLT Design

1. Simple <u>1-to-1</u> mapping between s and z-planes avoids aliasing

2. Can be used for bandpass and highpass designs as well as lowpass.

3. Don't necessarily need to use partial fraction expansion. Simply transform

$$s \longmapsto 2\left(\frac{z-1}{z+1}\right)$$

IIR FILTERS

- Impulse Invariance
- Bilinear Transformation

Advantage:

Simple designs base on CT Filters.
Frequency response (magnitude) requirements easily met.

Disadvantage:

Nonlinear phase

Ex. $y[n] = \frac{1}{2} y[n-1] + x[n]$

$$\rightarrow Y(z) = \frac{1}{2} z^{-1} Y(z) + X(z)$$

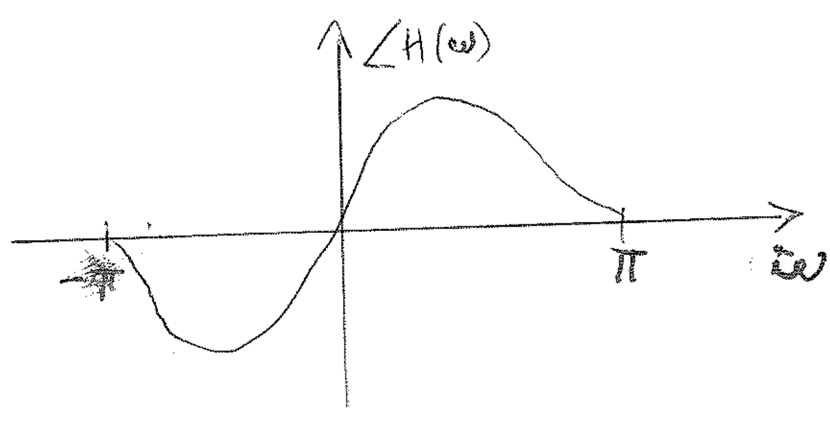
$$\rightarrow H(z) = \frac{1}{1 - \frac{1}{2} z^{-1}}$$

$$\rightarrow \boxed{H(\omega) = \frac{1}{1 - \frac{1}{2} e^{-j\omega}}}$$

$$H(\omega) = \frac{1}{(1 - \frac{1}{2}\cos\omega) + j\frac{1}{2}\sin\omega}$$

$$= \frac{(1 - \frac{1}{2}\cos\omega) - j\frac{1}{2}\sin\omega}{|1 - \frac{1}{2}\cos\omega|^2 + |\frac{1}{2}\sin\omega|^2}$$

$$\angle H(\omega) = \tan^{-1}\left(\frac{\frac{1}{2}\sin\omega}{1 - \frac{1}{2}\cos\omega}\right)$$



Phase Response is Important

Q: What does an LTI filter <u>do</u> to a signal?

A: In frequency domain

$$X \rightarrow \boxed{H} \rightarrow Y$$

$$Y(\omega) = H(\omega) \cdot X(\omega)$$

$$= \left[|H(\omega)| \cdot e^{j \angle H(\omega)} \right] \cdot X(\omega)$$

$|H(\omega)|$ ↑ attenuates or accentuates <u>amplitude</u> of frequency component of X at frequency ω.

$e^{j \angle H(\omega)}$ ↑ shifts the <u>phase</u> of frequency component of X at frequency ω.

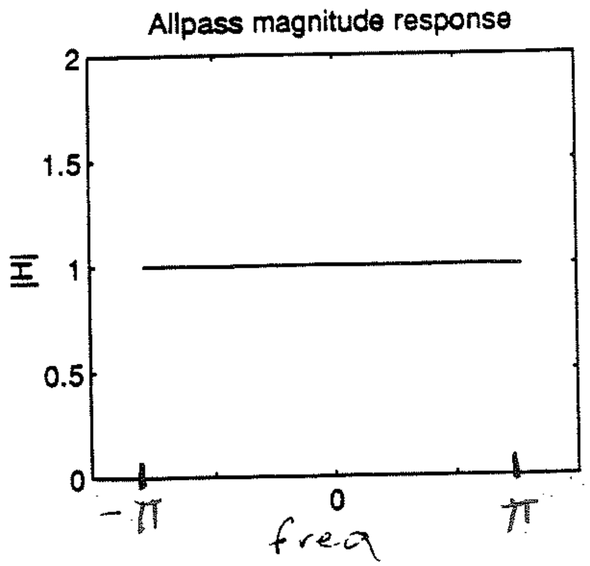
☆ IIR filter designs <u>only</u> attack magnitude of frequency response.

We can break things into 3 cases:

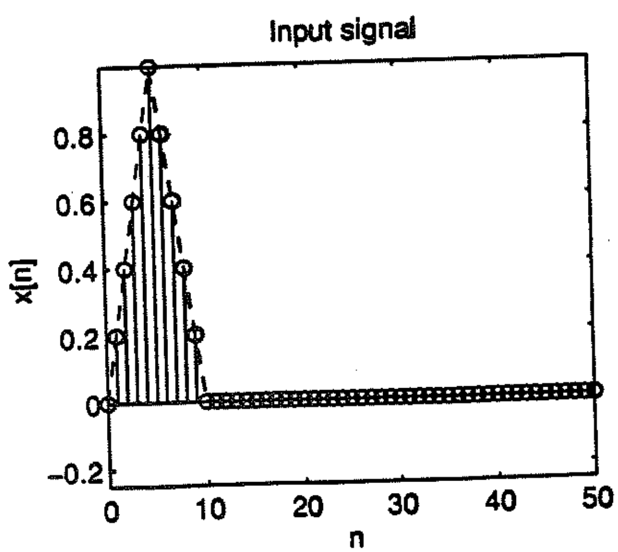
1. <u>zerophase</u> : $\angle H(\omega) = 0$
2. <u>linearphase</u> : $\angle H(\omega) = n_0 \omega$
3. <u>nonlinearphase</u> : Everything else

To see what the various phase types can do to signals, lets look at an example of an

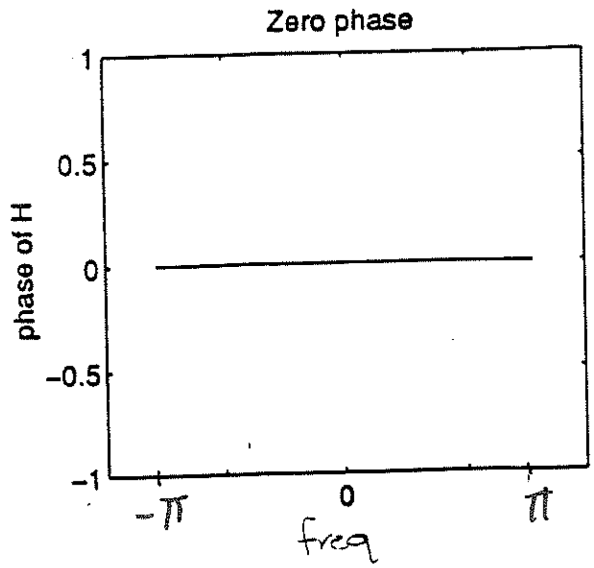
ALL PASS SYSTEM → $|H(\omega)| = 1$ for all

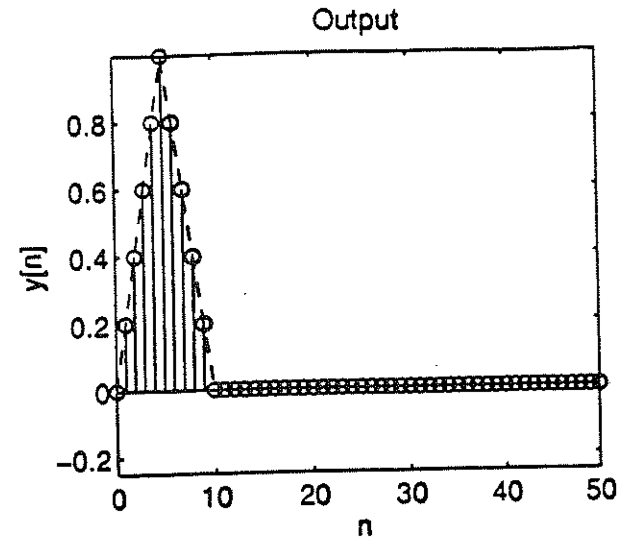


and a simple pulse signal $x(n)$

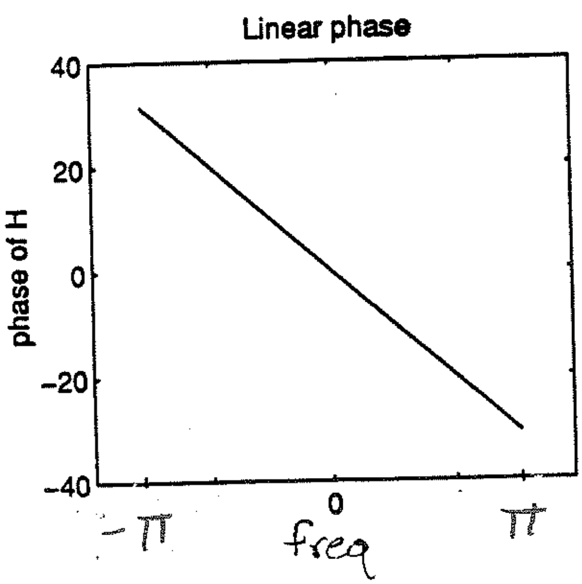


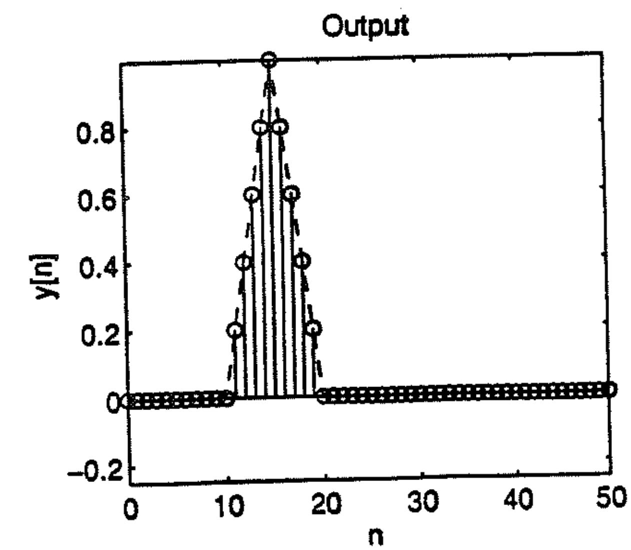
Zero phase: $\measuredangle H(\omega) = 0 \Rightarrow H(\omega) = 1$





Linear phase: $\measuredangle H(\omega) = -20\pi \Rightarrow H(\omega) = e^{-j10\omega}$





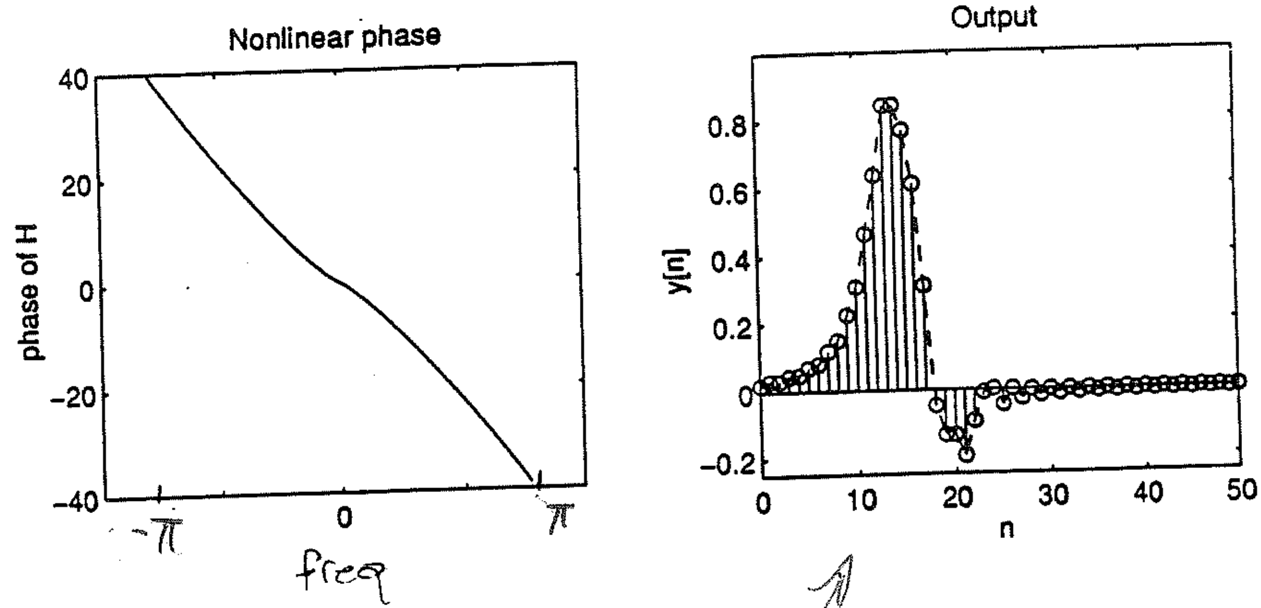
← simple delay of $\frac{20\pi}{2\pi} = 10$

Nonlinear Phase:

← a "mild" nonlinearity

$$\angle H(\omega) = -10|\omega|^{1.2}\,\mathrm{sign}(\omega)$$

$$\Rightarrow H(\omega) = e^{-j10|\omega|^{1.2}\mathrm{sign}(\omega)}$$



Distorted pulse!

Bottom Line: Nonlinear phase responce distorts output waveform

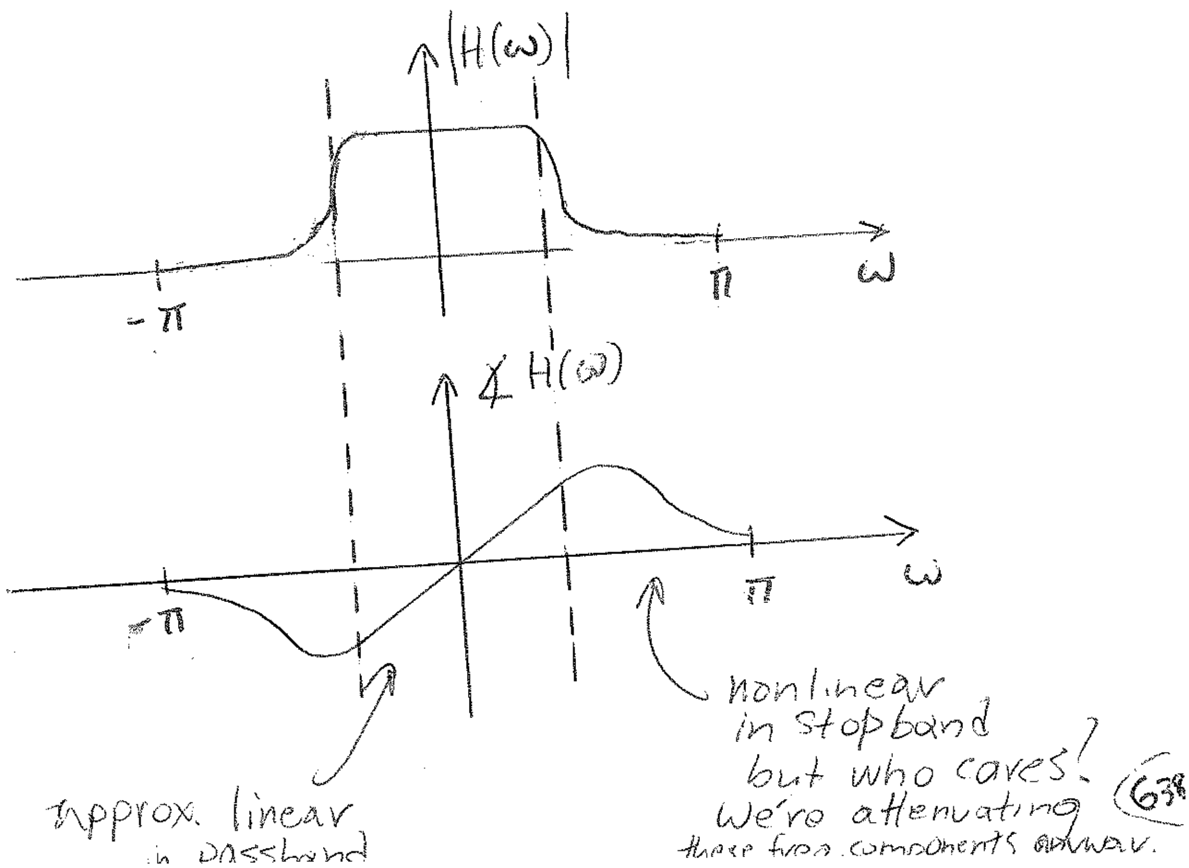
Linear Phase ⟹ simple delay

Nonlinear Phase ⟹ distortion of waveform

★ IIR Filters have nonlinear phase and therefore can distort waveforms

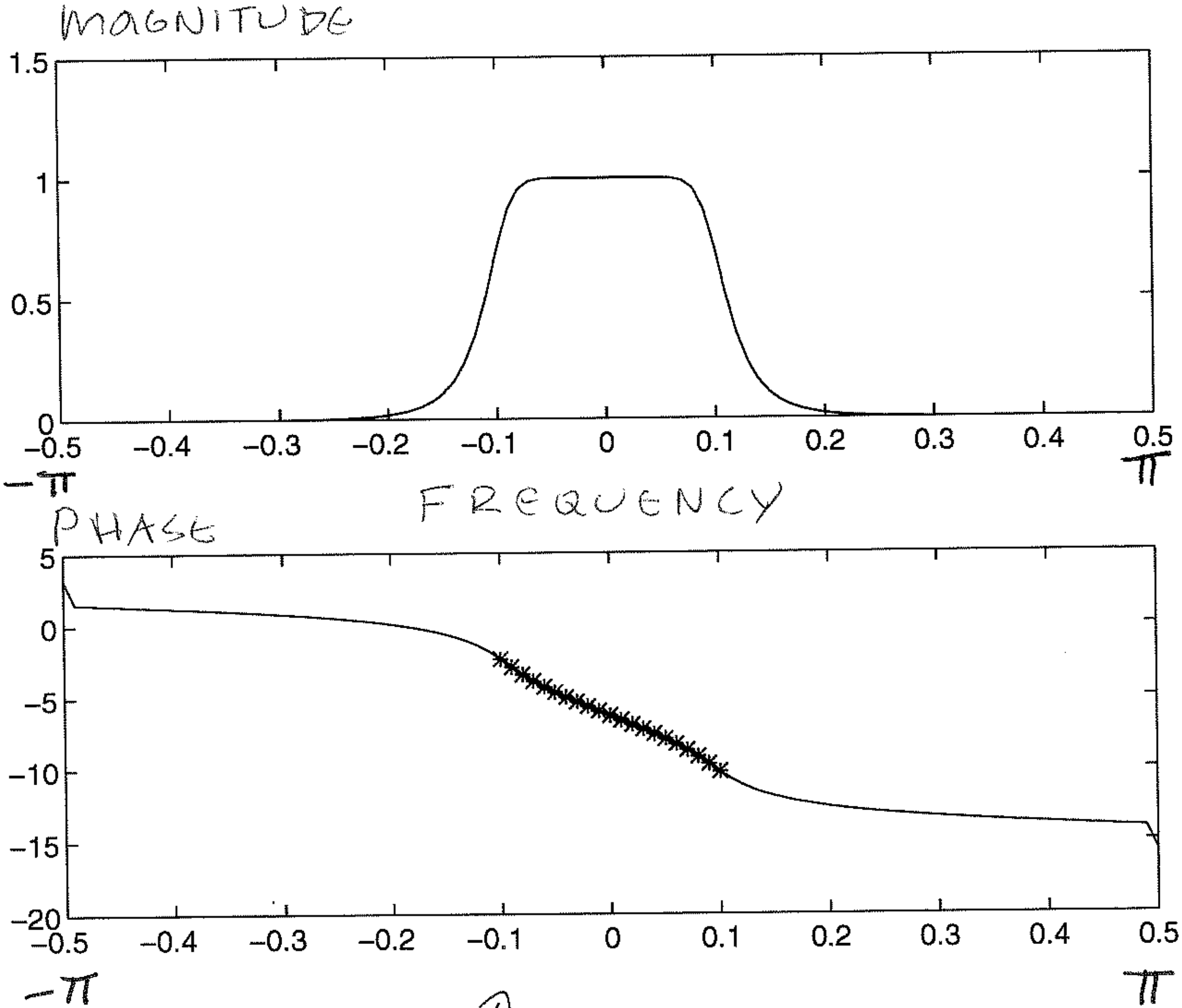
★ Distortion is often very minimal in passband (although hard to design for this)

ex.



638

Ex. <u>6-pole Butterworth LOWPASS IIR Filter</u>

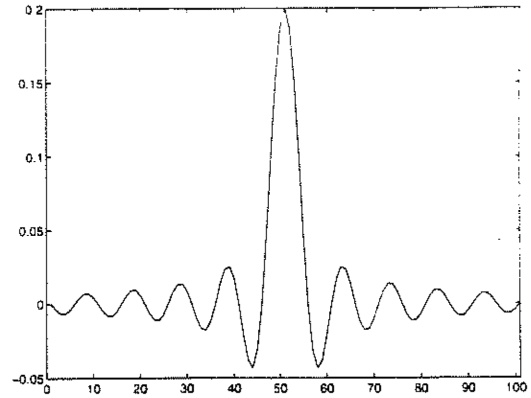


<u>Note</u>: phase is close to linear in passband

FIR LOWPASS FILTER

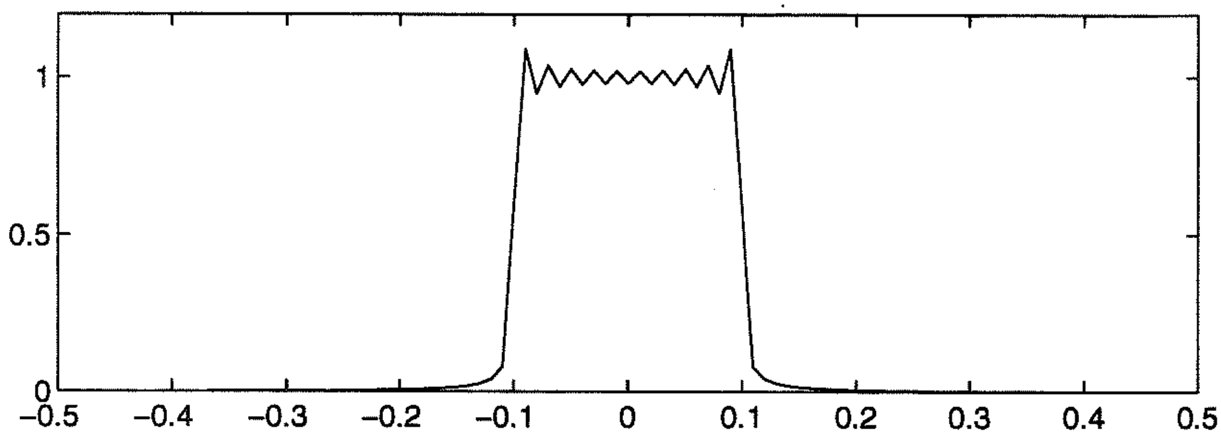
Shifted and truncated sinc

h[n] →



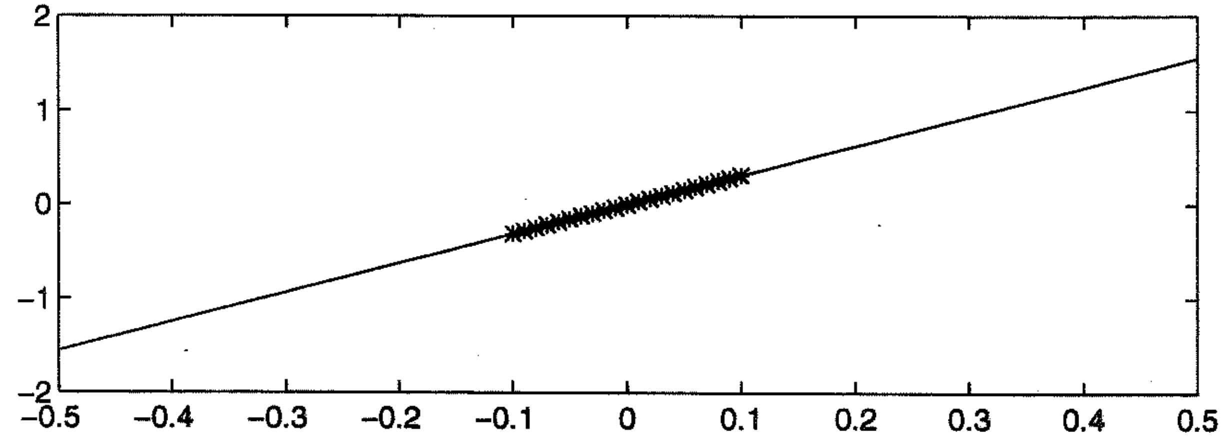
MAGNITUDE

← Note: Gibb's phenomena



$-\pi$ FREQ π

PHASE



$-\pi$ π

Note: Linear phase

FIR FILTER DESIGN

FIR Filter Design

- Does not involve mapping CT prototypes to DT
- Easy to design FIR filters with <u>linear phase</u>.
- Good numerical properties

We will look at two methods for FIR filter design

1. Windowing ← just truncate ideal filter impulse response to make it FIR.

2. Optimal Design ← Designs optimize specified criteria (Parks-McClellan method)

FIR Filters

$$H(z) = \sum_{n=0}^{N-1} h[n] z^{-n}$$

$$= h[0] + h[1] z^{-1} + \cdots + h[N-1] z^{-(N-1)}$$

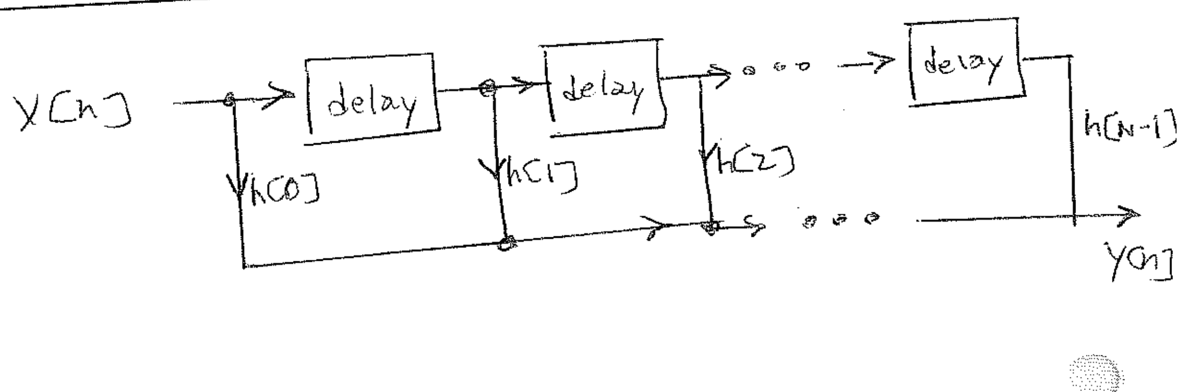
↳ polynomial in z^{-1} (not rational)

$h[n]$ ~ impulse response of FIR filter

$$y[n] = \sum_{k=0}^{N-1} h[k]\, x[n-k]$$

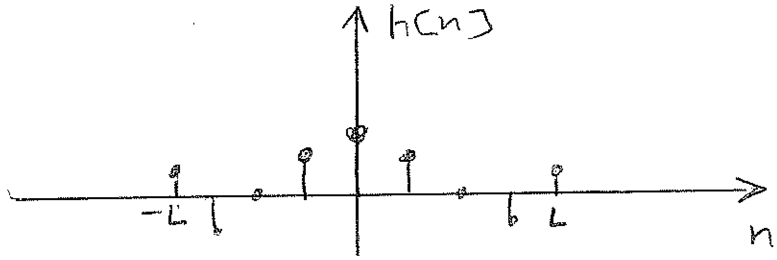
- no feedback!
- function of present and past inputs only

Implementation:



Symmetric FIR Filters

Consider a symmetric impulse response



$$H(\omega) = \sum_{n=-L}^{L} h[n] e^{-j\omega n}$$

$$= h[0] + \sum_{n=1}^{L} h[n] \left(e^{-j\omega n} + e^{j\omega n}\right)$$

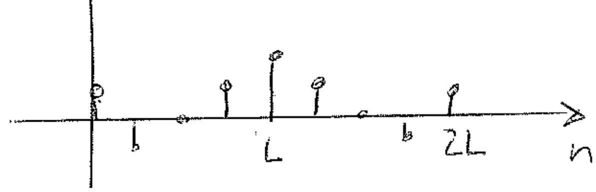
$$= h[n] + \sum_{n=1}^{K} h[n]\, 2\cos(\omega n)$$

← pure real

$\Rightarrow$ zero phase!

Problem: Noncausal!

Solution: Shift $h[n]$ to get causality



Note: $\tilde{H}(\omega) = \left(h[n] + \sum_{n=1}^{L} h[n]\, 2\cos(\omega n) \right) e^{-j\omega L}$

$$|H(\omega)| = |\tilde{H}(\omega)|$$

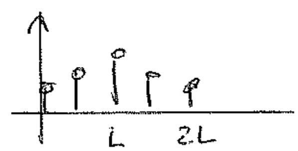
$$\angle H(\omega) = 0 \;, \quad \angle \tilde{H}(\omega) = \angle e^{-j\omega L} = -L$$

linear phase
- pure delay
- no distortion

There are four ways, in general, to construct a linear phase filter.

- two kinds of lengths : even / odd
- two kinds of symmetry : even / odd

TYPE I — EVEN symmetry + ODD length $2L+1$

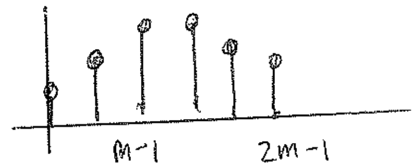


$$|H(\omega)| = h[0] + 2\sum_{n=1}^{L} h[n] \cos(\omega n)$$

$$\angle H(\omega) = -L$$

TYPE 2

EVEN symmetry + EVEN length $2M$

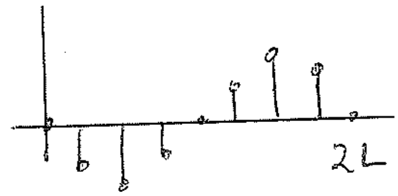


$$|H(\omega)| = 2 \sum_{k=1}^{M} h[k] \cos\left(\omega\left(k - \tfrac{1}{2}\right)\right)$$

$$\measuredangle H(\omega) = \measuredangle e^{-j\omega M} = -M$$

TYPE 3

ODD symmetry + ODD length $2L+1$

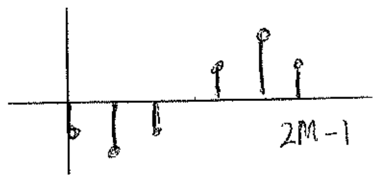


$$|H(\omega)| = 2 \sum_{k=1}^{L} h[k] \sin(\omega k)$$

$$\measuredangle H(\omega) = -L$$

TYPE 4

ODD symmetry + EVEN length $2M$



$$|H(\omega)| = 2 \sum_{k=1}^{M} h[k] \sin\left(\omega\left(k - \tfrac{1}{2}\right)\right)$$

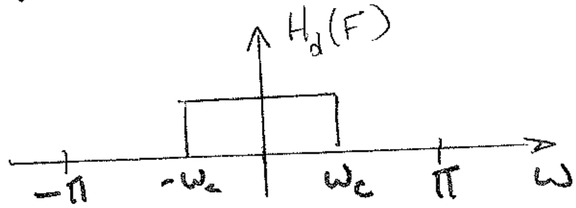
$$\measuredangle H(\omega) = -M$$

FIR FILTER DESIGN - WINDOW METHOD

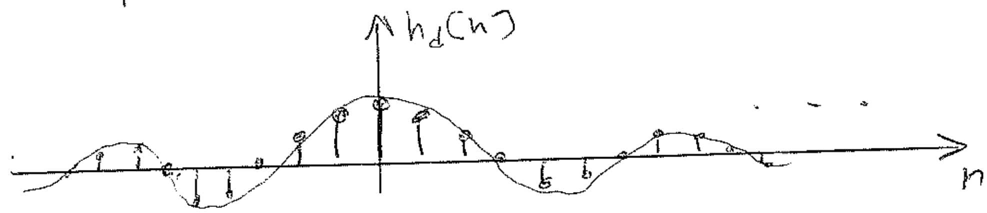
★ LO - TECH !

Simple motivation : Say we want to design a lowpass filter.

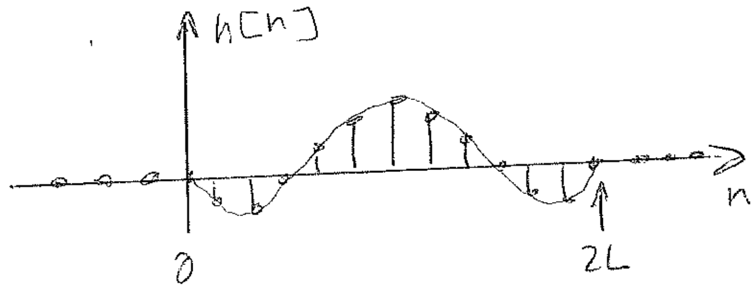
1. Start with the ideal filter response $H_d(\omega)$



2. Compute IDTFT of $H_d(\omega) \longrightarrow h_d[n]$



3. $h_d[n]$ is infinite extent, so truncate and shift to make it causal



MAJOR DRAWBACK

Truncation introduces Gibbs' phenomena!

To see this, we can view

$$h[n] = \begin{cases} h_d[n-L] & , 0 \le n \le N-1 \\ 0 & , \text{otherwise} \end{cases}$$

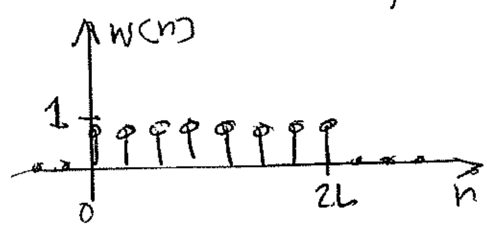
$$= h_d[n-L] \cdot w[n]$$

↑ rectangular window

$$w[n] = \begin{cases} 1 & , 0 \le n \le 2L \\ 0 & , \text{otherwise} \end{cases}$$

$$W(\omega) = e^{-j\omega\frac{(2L+2)}{2}} \frac{\sin\left(\omega\left(\frac{2L+1}{2}\right)\right)}{\sin \omega/2}$$

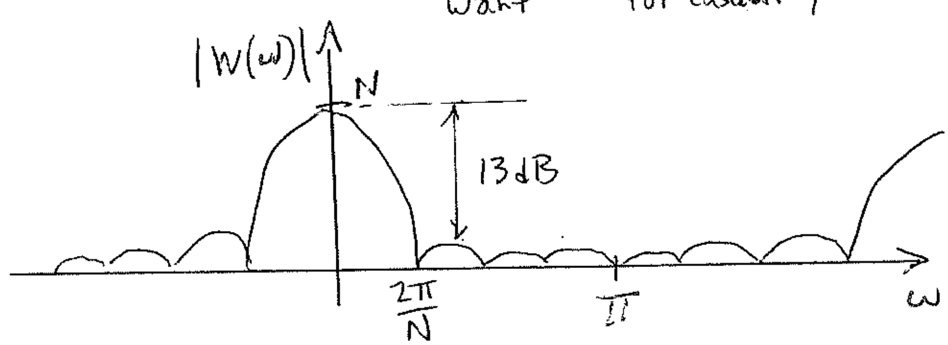
↗ Sinc function



$\Rightarrow$

$$H(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_d(\omega)\, e^{-j\lambda L}\, W(\omega-\lambda)\, d\lambda$$

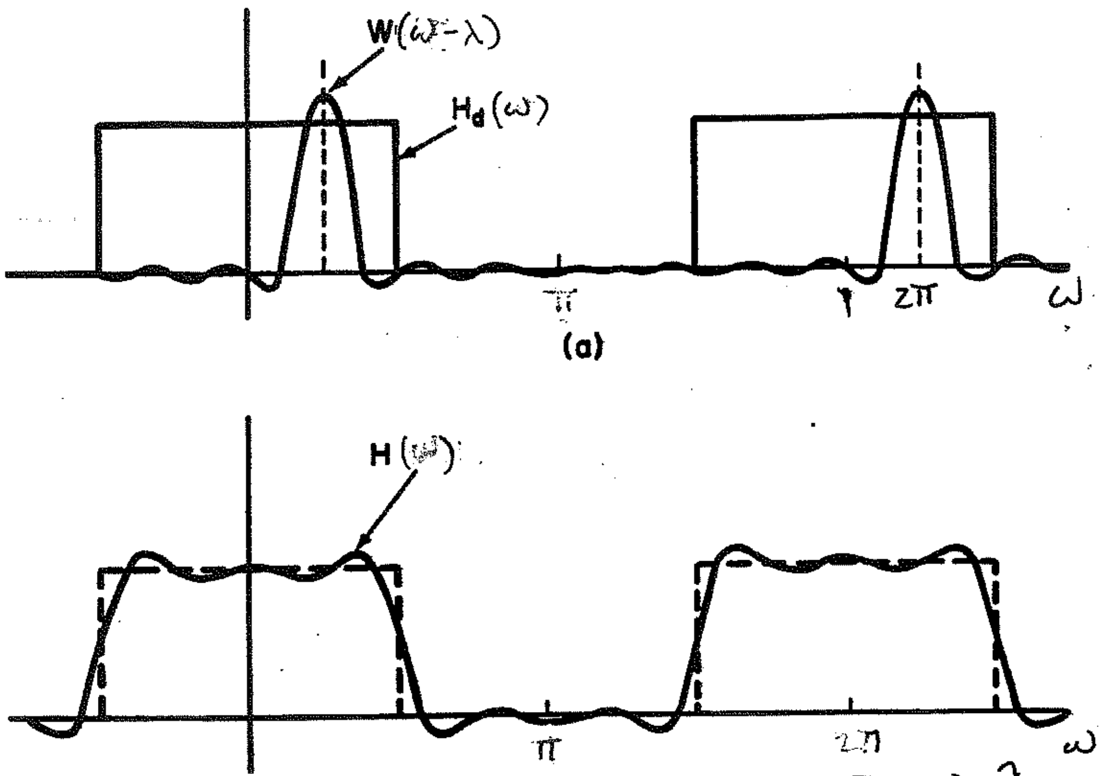
- $H_d(\omega)$ ↑ What we want
- $e^{-j\lambda L}$ ↑ delay for causality
- $W(\omega-\lambda)$ ↑ "smears" or "blurs" desired freq. response



Convolution with $W(\omega)$ has two detrimental effects:

1. main lobe width smears out sharp edges in $H_d(\omega)$.

2. Sidelobes add ripples (Gibb's phenomena)



Why not just make $W(\omega) = \delta(\omega)$?

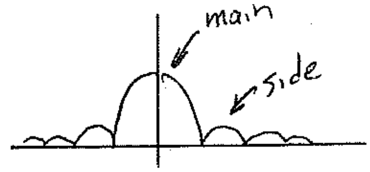
→ $w[n] = 1$ for all n

→ <u>no windowing (truncation) at all!</u>

Not very useful.

Every finite length window has DTFT with

1. Main lobe of some width
2. sidelobes



We want to minimize mainlobe width and minimize sidelobe height.

Recall from our discussions on windowing (E-30 to E-33) these are conflicting requirements, leading to

mainlobe width / sidelobe trade-off

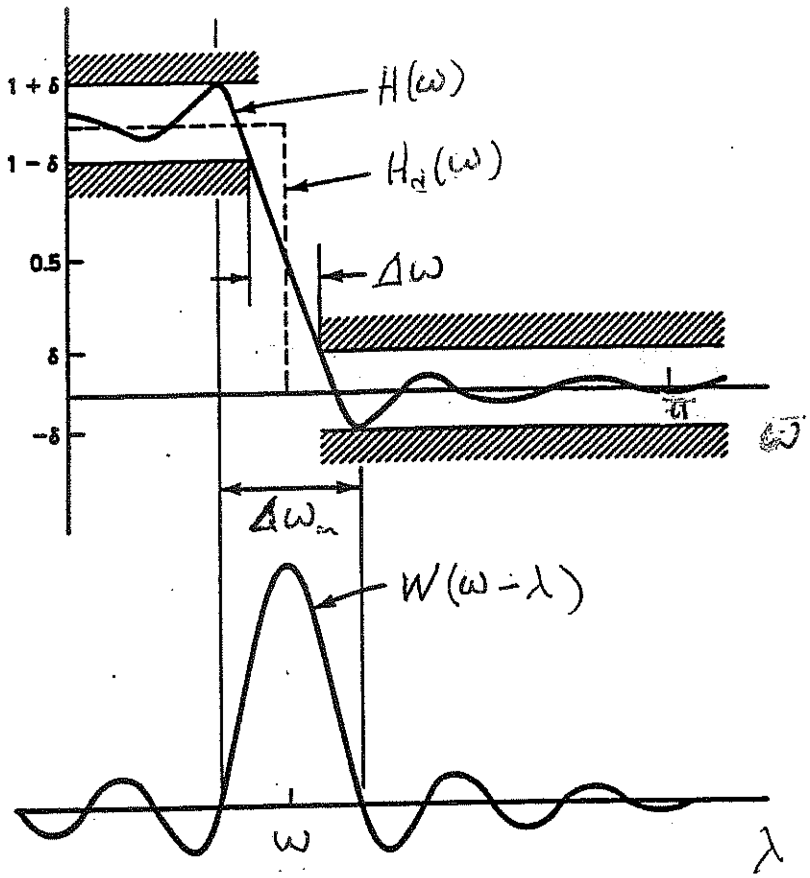
↑ implies engineering

Unfortunately, this trade-off means that there really isn't <u>a</u> best window.

The choice of window is more or less a subjective one:

Do you want a better main lobe or better sidelobes?

Example Design



Bottom Line:

1. Window designs are not optimal in any sense (save the rectangular & L^2) ← we will show this in the next section
2. Need a closed form expression for $H_d(\omega)$
3. Window smearing makes placement of transition points difficult (need to iterate design process)
4. Easy, but kind of stupid ⇒ don't use window design.
5. Ripple is the same in passband & stopband

OPTIMAL FIR FILTER DESIGN

Objective: Design the "best" filter for a given problem

$\Rightarrow$ optimizotion

Needed: a numerical quantity or criterion that measures performance.

Sensible Measure: Error in frequency response.

How "close" is $H(\omega)$ to $H_d(\omega)$?

How can we measure close?

Criterion 1: (minimize Energy of Error)

$$\text{minimize} \int_{-\pi}^{\pi} |H(\omega) - H_d(\omega)|^2 \, d\omega$$

Criterion 2: (minimize maximum error)

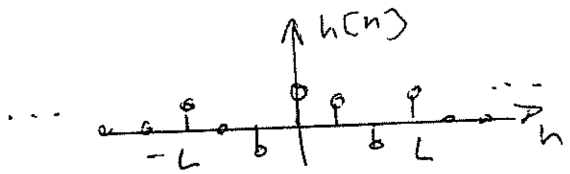
$$\text{minimize} \max_{-\frac{1}{2} \le \;\le \frac{1}{2}} |H_d(\omega) - H(\omega)|$$

★ Choose FIR Filter $h[n]$ that minimizes criteron 1 or 2..

CRITERION 1 : Minimize Energy of Error

To start with:

- $h[n]$ is fixed length $2L+1$
- centered at origin
  (we can take care of shift (for causality) later)



Objective:

Find $h[n]$ to minimize

$$\int_{-\pi}^{\pi} |H_d(\omega) - H(\omega)|^2$$

$$H_d(\omega) = \sum_{n=-\infty}^{\infty} h_d[n] e^{-j\omega n}$$

$$H(\omega) = \sum_{n=-\infty}^{\infty} h[n] e^{-j\omega n}$$

$\longrightarrow$

$$H_d(\omega) - H(\omega) = \sum_{n=-\infty}^{\infty} (h_d[n] - h[n]) e^{-j\omega n}$$

$$= \sum_{n=-\infty}^{\infty} d[n] e^{-j\omega n}$$

$$d[n] = h_d[n] - h[n]$$

So we see,

the FIR filter that minimizes the energy of the error is

$$h[n] = \begin{cases} h_d[n], & -L \leq n \leq L \\ 0, & \text{otherwise} \end{cases}$$

NOTE:

1. We can shift $h[n]$ above to get causality (and linear phase)

2. This shows that the optimal filter minimizing the energy criterion is nothing but the rectangular window truncation!

→ We have already seen this is not such a great filter design.

⇒ maybe energy criterion isn't such a hot idea!

CRITERION 2 : "Minimize maximum error"

minimax design

Parks-McClellan Algorithm

$$\min_{h} \max_{\omega \in \text{passband} \cup \text{stopband}} |H_d(\omega) - H(\omega)|$$

h ← N tap filter coeff.

$\omega \in$ passband ∪ stopband ← frequencies in passband or stopband but not transition band

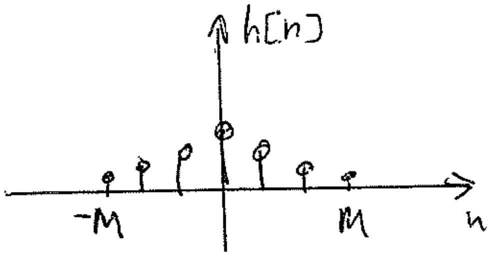
- IDEA makes good engineering sense (minimize the max error, make response $H(\cdot)$ meet specs to within $\pm\delta$ at all frequencies with δ as small as possible)
- Results in good designs (they work well in real-world DSP problems)
- Also, spreads error out over passband and stopband rather than concentrating at the band edges like a window design.

Let's consider Type 1 (symmetric, odd-length) FIR filters. Recall that the DTFT of a Type I filter is

$$H(\omega) = h[0] + 2\sum_{k=1}^{M} h[k]\cos(k\omega)$$

$$\doteq \sum_{k=0}^{M} \alpha_k \cos(k\omega)$$

Chebyshev Polynomial



DTFT ⟷ $H(\omega) = \sum_{k=0}^{M} \alpha_k \cos(k\omega)$

Note: Noncausal.

We will match desired magnitude response, then shift to get causality.

Simple relationship between α_k's and $h[n]$.

$$\left(\begin{array}{l} h[0] = \alpha_0 \\ h[k] = \tfrac{1}{2}\alpha_k \\ \quad k > 0. \end{array} \right)$$

Objective: Choose α_k's to minimize

$$\max_{\omega \in \text{Passband} \cup \text{Stopband}} \left| H_d(\omega) - \sum_{k=0}^{M} \alpha_k \cos(k\omega) \right|$$

Question: How?

The Alternation Theorem: (from mathematical Approximation Theory)

Consider the weighted approximation error

$$\epsilon(\omega) = W(\omega) \left| H_d(\omega) - \sum_{k=0}^{M} \alpha_k \cos(k\omega) \right|$$

$W(F)$ represents a weighting function that can be used to select different error bounds in passband and stopband.

The Alternation Theorem tells us that there are at least $M+2$ frequencies, $_k$, $k = 1, 2, \ldots, M+2$, called extremal frequencies where

① The error alternates between two equal maxima and minima (extrema)

$$\epsilon(\omega_k) = -\epsilon(\omega_{k+1}) \quad k = 1, 2, \ldots, M+1$$

② the error at the frequencies F_k equals the maximum absolute error

$$|\epsilon(\omega_k)| = \max |\epsilon(\omega)|, \quad k = 1, 2, \ldots, M+2$$

Now,

$$\int_{-\pi}^{\pi} |H_d(\omega) - H(\omega)|^2 \, dF$$

$$= \int_{-\pi}^{\pi} \left| \sum_{n=-\infty}^{\infty} d[n] e^{-j\omega n} \right|^2 d\omega$$

conj.

$$= \int_{-\pi}^{\pi} \left(\sum_{n=-\infty}^{\infty} d(n) e^{-j\omega n} \right) \left(\sum_{m=-\infty}^{\infty} d[m] e^{-j\omega m} \right)^* d\omega$$

$$= \int_{-\pi}^{\pi} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} d[n] d(m) e^{-j\omega(n-m)} \, d\omega$$

$$= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} d[n] d[m] \underbrace{\int_{-\pi}^{\pi} e^{-j\omega(n-m)} \, d\omega}$$

$$= \begin{cases} 1, & n=m \\ 0, & n \neq m \end{cases}$$

because of periodicity

$$= \sum_{n=-\infty}^{\infty} d^2[n]$$

$$= \sum_{n=-\infty}^{\infty} |h_d[n] - h[n]|^2$$

$$= \sum_{n=-L}^{L} |h_d[n] - h[n]|^2 + \sum_{\substack{n<-L \\ n>L}} |h_d[n]|^2$$

↑ this term is minimum (0) when $h[n] = h_d[n]$

↑ can't do anything about is term

Alternation Thm ⇒ there exist at least $M+2$ extremal frequencies where we achieve the maximum error.

⇒ let's find the extremal frequencies and choose α_k to mimize the error at these points.

⇒ Requires iterative methods

Most popular method is the

Parks - McClellan Algorithm

which uses the

Remez Exchange Algorithm

Resulting FIR filter design is called an

Equiripple Approximation

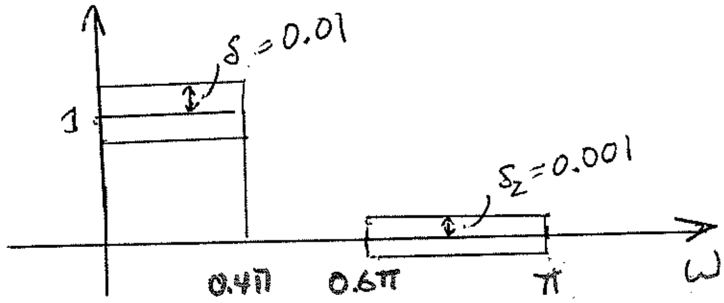
Example — LOWPASS DESIGN

SPECS: $0.99 \le |H(\omega)| \le 1.01$, $|\omega| \le 0.4\pi$

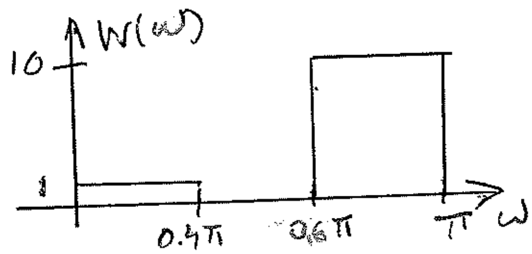
$|H(\omega)| \le 0.001$, $0.6\pi \le |\omega| \le \pi$

$\Longrightarrow \delta_1 = 0.01$, $\delta_2 = 0.001$, $\omega_p = 0.4\pi$

$\omega_s = 0.6\pi$ $\qquad K = \dfrac{\delta_1}{\delta_2} = 10$



Note: $\delta_1 \ne \delta_2 \Longrightarrow$ use weight function & $K = 10$



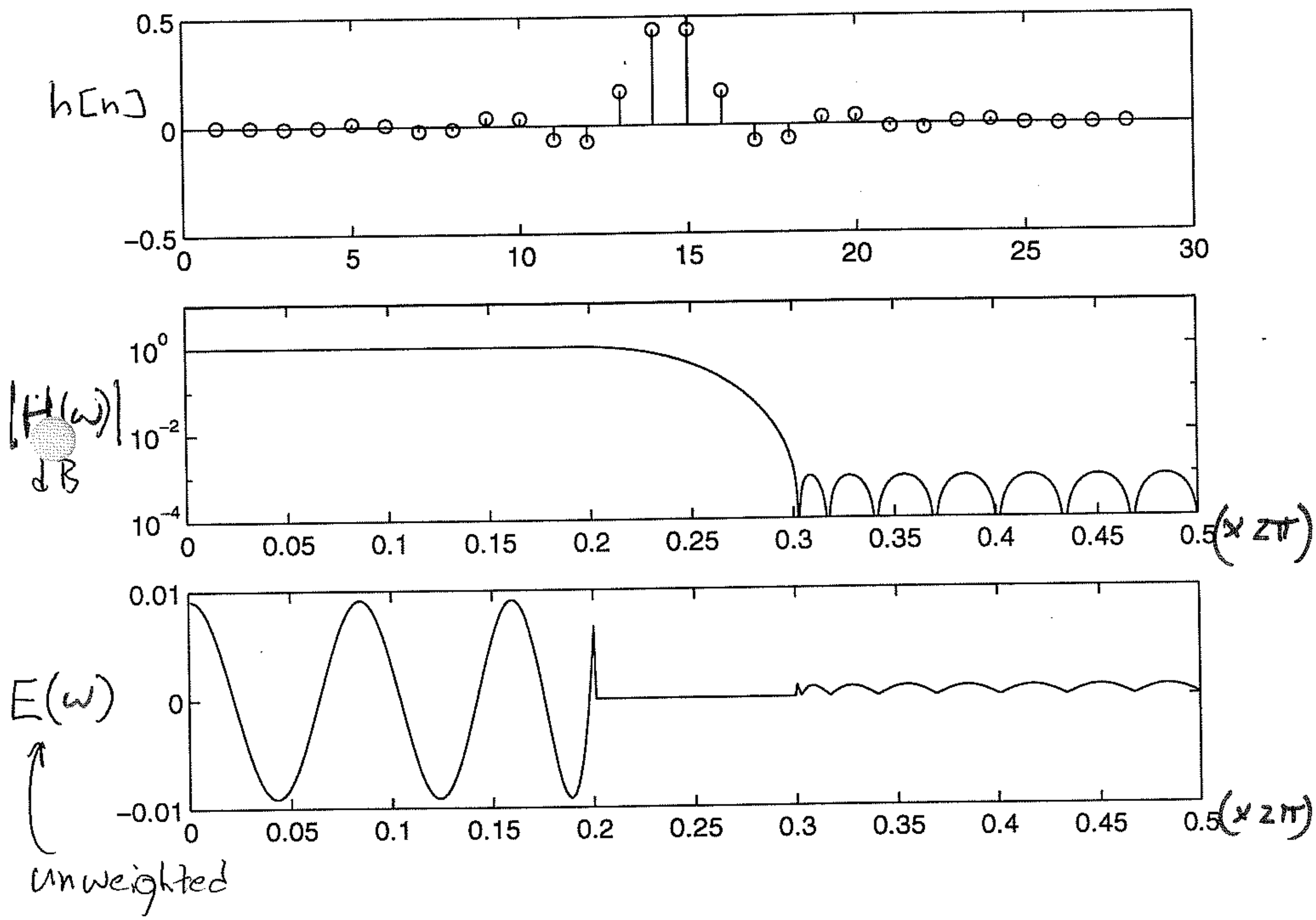
Matlab Command:

h = remez(N-1 , f , m , w)

- N-1 (Note -1): filter length N
- f: $[0 \;\; \frac{\omega_p}{\pi} \;\; \frac{\omega_s}{\pi} \;\; 1] = [0 \;\; 0.4 \;\; 0.6 \;\; 1]$
- m: desired values of freq response at freqs in f: [1 1 0 0]
- w: band weight [1 10]

☆ Increase N until δ_1 & δ_2 specs are met h = remez(N-1, f, m, w)

Resulting Design



How the Remez Exchange Works

① Start with trial set of frequencies
$\omega_1, \omega_2, \ldots, \omega_{M+2}$

② Force satisfaction of the alternation conditions at $\omega_1 \ldots \omega_{M+1}$

Set $\quad \epsilon(\omega_k) = \rho$
$\quad\quad \epsilon(\omega_{k+1}) = -\rho \quad k = 1, \ldots, M+1$

(we don't know max error, so $\rho > 0$ is picked arbitrarily)

③ Now we have M+1 unknown coefficients α_k, and unknown true value of ρ. We have M+2 equations:

$$(-1)^k \rho = W(\omega_k) \,|\, H_d(\omega_k) - H(\omega_k) |$$
$$k = 1, 2, \ldots, M+2$$

→ solve for α_k's and ρ.

④ Compute true extremal frequencies for new α_k's. These may not satisfy alternation condition so go back and repeat ② and ③

(Repeat until convergence)

☆ Check out built-in Matlab command remez