## ESE 5310: Digital Signal Processing

Lecture 18: March 23, 2023 Optimal Filter Design



Penn ESE 5310 Spring 2023 - Khanna



- Window method
  - Design Filters heuristically using windowed sinc functions
  - Choose order and window type
  - Check DTFT to see if filter specs are met
- Optimal design
  - Design a filter h[n] with  $H(e^{j\omega})$
  - Approximate H<sub>d</sub>(e<sup>jω</sup>) with some optimality criteria or satisfies specs.



### (mathematical) optimization problem

minimize 
$$f_0(x)$$
  
subject to  $f_i(x) \le b_i, \quad i = 1, \dots, m$ 

- $x = (x_1, \ldots, x_n)$ : optimization variables
- $f_0: \mathbf{R}^n \to \mathbf{R}$ : objective function
- $f_i: \mathbf{R}^n \to \mathbf{R}, i = 1, \dots, m$ : constraint functions

**optimal solution**  $x^{\star}$  has smallest value of  $f_0$  among all vectors that satisfy the constraints



#### general optimization problem

- very difficult to solve
- methods involve some compromise, e.g., very long computation time, or not always finding the solution

exceptions: certain problem classes can be solved efficiently and reliably

- least-squares problems
- linear programming problems
- convex optimization problems



minimize  $||Ax - b||_2^2$ 

### solving least-squares problems

- analytical solution:  $x^{\star} = (A^T A)^{-1} A^T b$
- reliable and efficient algorithms and software
- computation time proportional to  $n^2k$  ( $A \in \mathbf{R}^{k \times n}$ ); less if structured
- a mature technology

#### using least-squares

- least-squares problems are easy to recognize
- a few standard techniques increase flexibility (*e.g.*, including weights, adding regularization terms)



$$\begin{array}{ll} \mathsf{minimize} & c^T x\\ \mathsf{subject to} & a_i^T x \leq b_i, \quad i=1,\ldots,m \end{array}$$

### solving linear programs

- no analytical formula for solution
- reliable and efficient algorithms and software
- computation time proportional to  $n^2m$  if  $m \ge n$ ; less with structure
- a mature technology

### using linear programming

- not as easy to recognize as least-squares problems
- a few standard tricks used to convert problems into linear programs (e.g., problems involving ℓ<sub>1</sub>- or ℓ<sub>∞</sub>-norms, piecewise-linear functions)



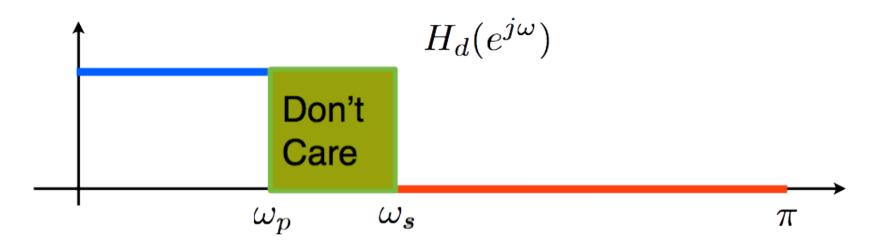
$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i, \quad i=1,\ldots,m \end{array}$$

• objective and constraint functions are convex:

$$f_i(\alpha x + \beta y) \le \alpha f_i(x) + \beta f_i(y)$$

 $\text{if } \alpha + \beta = 1 \text{, } \alpha \geq 0 \text{, } \beta \geq 0$ 

• includes least-squares problems and linear programs as special cases



□ Least Squares:

minimize 
$$\int_{\omega \in \text{care}} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega$$

□ Variation: Weighted Least Squares:

minimize

$$\int_{-\pi}^{\pi} W(\omega) |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega$$

Design Through Optimization

## □ Idea: Sample/discretize the frequency response

$$H(e^{j\omega}) \Rightarrow H(e^{j\omega_k})$$

• Sample points are fixed  $\omega_k = k \frac{\pi}{P}$ 

$$-\pi \leq \omega_1 < \cdots < \omega_p \leq \pi$$

- □ M+1 is the filter order
- $\square P >> M + 1 (rule of thumb P=15M)$
- Yields a (good) approximation of the original problem



□ Target: Design M+1=2N+1 filter

□ First design non-causal  $\tilde{H}(e^{j\omega})$  and hence  $\tilde{h}[n]$ 



- □ Target: Design M+1=2N+1 filter
- □ First design non-causal  $\tilde{H}(e^{j\omega})$  and hence  $\tilde{h}[n]$
- □ Then, shift to make causal

$$\begin{split} h[n] &= \tilde{h}[n-M/2] \\ H(e^{j\omega}) &= e^{-j\frac{M}{2}}\tilde{H}(e^{j\omega}) \end{split}$$



$$\tilde{h} = \left[\tilde{h}[-N], \tilde{h}[-N+1], \cdots, \tilde{h}[N]\right]^T$$

$$b = \left[H_d(e^{j\omega_1}), \cdots, H_d(e^{j\omega_P})\right]^T$$

$$A = \begin{bmatrix} e^{-j\omega_{1}(-N)} & \cdots & e^{-j\omega_{1}(+N)} \\ \vdots \\ e^{-j\omega_{P}(-N)} & \cdots & e^{-j\omega_{P}(+N)} \end{bmatrix}$$
$$\operatorname{argmin}_{\tilde{h}} ||A\tilde{h} - b||_{2}^{2}$$



Solution: 
$$\begin{aligned} \arg \min_{\tilde{h}} \ ||A\tilde{h}-b||_2^2 \\ \tilde{h} = (A^*A)^{-1}A^*b \end{aligned}$$

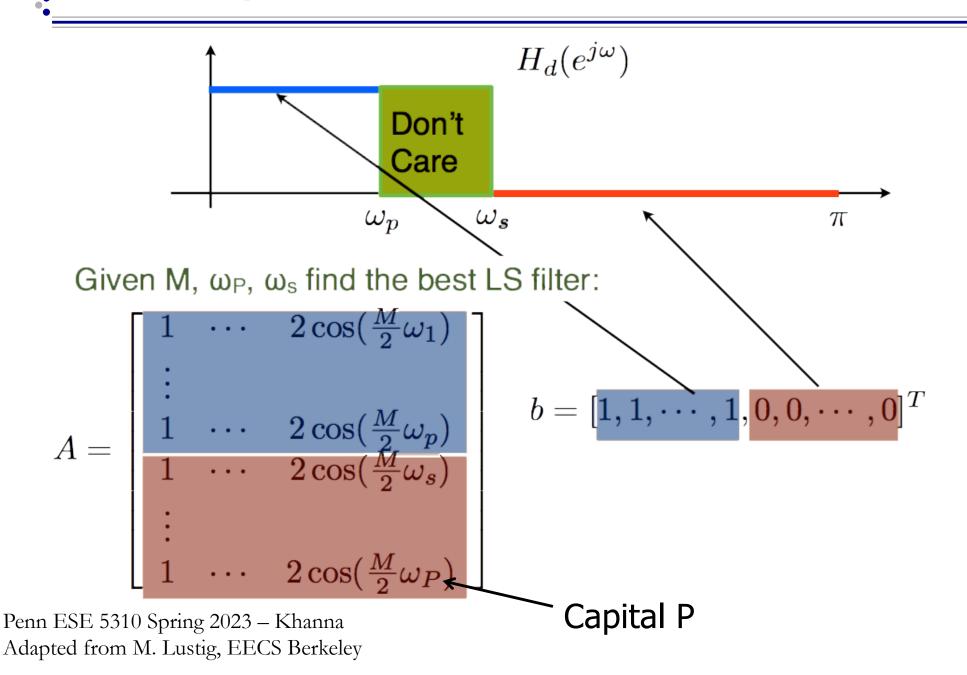
- Result will generally be non-symmetric and complex valued.
- However, if  $\tilde{H}(e^{j\omega})$  is real,  $\tilde{h}[n]$  should have symmetry!

# • Design of Linear-Phase L.P Filter

- Suppose:
  - $\tilde{H}(e^{j\omega})$  is real-symmetric
  - M is even (M+1 length)
- **Then:** 
  - $\tilde{h}[n]$  is real-symmetric around midpoint
- So:

$$\begin{split} \tilde{H}(e^{j\omega}) &= \tilde{h}[0] + \tilde{h}[1]e^{-j\omega} + \tilde{h}[-1]e^{+j\omega} \\ &+ \tilde{h}[2]e^{-j2\omega} + \tilde{h}[-2]e^{+j2\omega} \cdots \\ &= \tilde{h}[0] + 2\cos(\omega)\tilde{h}[1] + 2\cos(2\omega)\tilde{h}[2] + \cdots \end{split}$$

# Least-Squares Linear Phase Filter



# Least-Squares Linear Phase Filter

Given M,  $\omega_P$ ,  $\omega_s$  find the best LS filter:

$$A = \begin{bmatrix} 1 & \cdots & 2\cos(\frac{M}{2}\omega_{1}) \\ \vdots & & \\ 1 & \cdots & 2\cos(\frac{M}{2}\omega_{p}) \\ 1 & \cdots & 2\cos(\frac{M}{2}\omega_{s}) \\ \vdots & & \\ 1 & \cdots & 2\cos(\frac{M}{2}\omega_{P}) \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = [\tilde{h}[0], \cdots, \tilde{h}[\frac{M}{2}]]^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[0], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots, \tilde{h}[\frac{M}{2}] \end{bmatrix}^{T} = (A^{*}A)^{-1}A^{*}b \\ \tilde{h}_{+} = \begin{bmatrix} \tilde{h}[n], \cdots,$$



LS has no preference for pass band or stop band
Use weighting of LS to change ratio

want to solve the discrete version of:

minimize 
$$\int_{-\pi}^{\pi} W(\omega) |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega$$

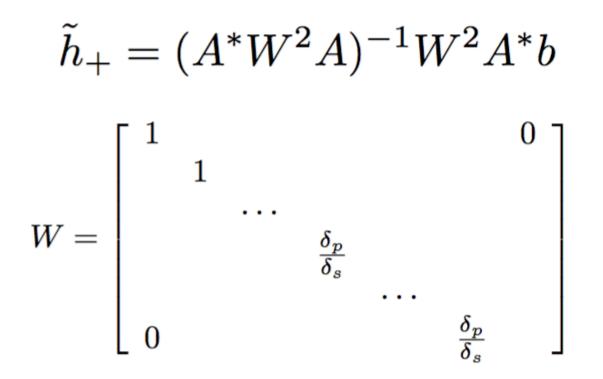
where  $W(\omega)$  is  $\delta p$  in the pass band and  $\delta s$  in stop band

Similarly:  $W(\omega)$  is 1 in the pass band and  $\delta p/\delta s$  in stop band



$$\operatorname{argmin}_{\tilde{h}_{+}} (A\tilde{h}_{+} - b)^* W^2 (A\tilde{h}_{+} - b)$$

Solution:





□ Chebychev Design (min-max)

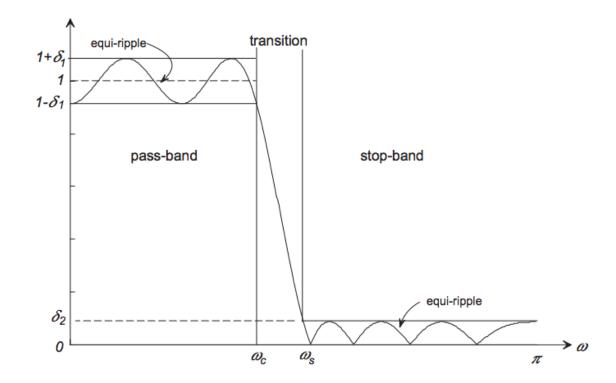
minimize<sub> $\omega \in care$ </sub> max  $|H(e^{j\omega}) - H_d(e^{j\omega})|$ 

- Parks-McClellan algorithm equiripple
- Also known as Remez exchange algorithms (signal.remez)
- Can also use convex optimization



- □ Allows for multiple pass- and stop-bands.
- Is an equi-ripple design in the pass- and stop-bands, but allows independent weighting of the ripple in each band.
- □ Allows specification of the band edges.





• For the low-pass filter shown above the specifications are

$$\begin{array}{rcl} 1 - \delta_1 &< & H(\mathrm{e}^{\mathrm{j}\,\omega}) &< & 1 + \delta_1 & \quad \text{in the pass-band } 0 < \omega \leq \omega_c \\ -\delta_2 &< & H(\mathrm{e}^{\mathrm{j}\,\omega}) &< & \delta_2 & \quad \text{in the stop-band } \omega_s < \omega \leq \pi. \end{array}$$



 Need to determine M+1 (length of the filter) and the filter coefficients {h<sub>n</sub>}



- Need to determine M+1 (length of the filter) and the filter coefficients {h<sub>n</sub>}
- If we assume M even and even symmetry FIR filter (Type I), then

$$A_e(e^{j\omega}) = h_e[0] + \sum_{n=1}^{L} 2h_e[n]\cos(\omega n).$$

$$H(e^{j\omega}) = A_e(e^{j\omega})e^{-j\omega M/2}$$



### **Reformulate:**

$$A_{e}(e^{j\omega}) = h_{e}[0] + \sum_{n=1}^{L} 2h_{e}[n]\cos(\omega n).$$

• ... to fitting a polynomial

$$A_e(e^{j\omega}) = \sum_{k=0}^L a_k(\cos\omega)^k,$$

$$A_e(e^{j\omega}) = P(x)|_{x=\cos\omega}, \qquad P(x) = \sum_{k=0}^L a_k x^k.$$

Penn ESE 5310 Spring 2023 - Khanna



Define approximation error function

$$E(\omega) = W(\omega)[H_d(e^{j\omega}) - A_e(e^{j\omega})],$$

$$W(e^{j\omega}) = \begin{cases} \delta_2/\delta_1 & \text{in the pass-band} \\ 1 & \text{in the stop-band} \\ 0 & \text{in the transition band.} \end{cases}$$



Define approximation error function

$$E(\omega) = W(\omega)[H_d(e^{j\omega}) - A_e(e^{j\omega})],$$

### □ Apply min-max or Chebyshev criteria

$$\min_{\{h_e[n]:0\leq n\leq L\}} \Big(\max_{\omega\in F} |E(\omega)|\Big),$$

Penn ESE 5310 Spring 2023 - Khanna

Parks-McClellan – Alternation Theorem

The algorithm uses Chebyshev's alternation theorem to recognize the optimal solution.

Define the error E(x) as above, namely

$$E(e^{j\omega}) = W(e^{j\omega}) \left( H_d(e^{j\omega}) - H(e^{j\omega}) \right)$$

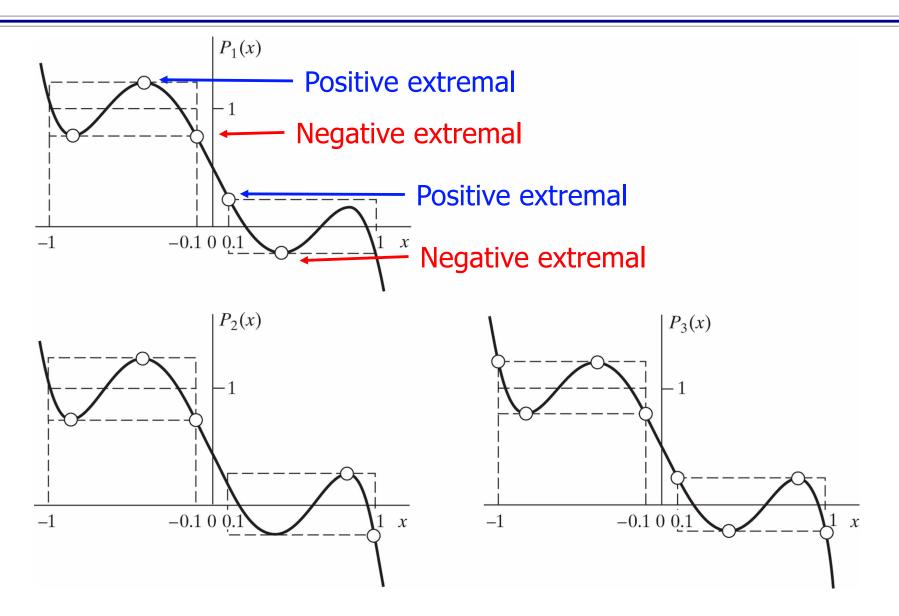
and the maximum error as

$$||E(e^{j\omega})||_{\infty} = \operatorname{argmax}_{x \in \Omega} |E(e^{j\omega})|$$

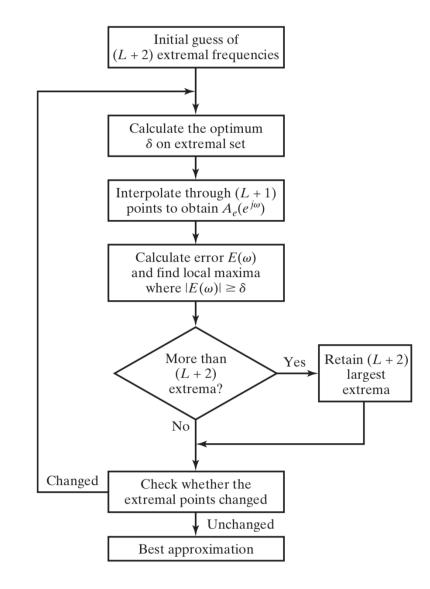
A necessary and sufficient condition that  $H(e^{j\omega})$  is the unique *L*th-order polynomial minimizing  $||E(e^{j\omega})||_{\infty}$  is that  $E(e^{j\omega})$  exhibit at least L + 2extremal frequencies, or "alternations", that is there must exist at least L+2values of  $\omega$ ,  $\omega_k \in \Omega$ ,  $k = [0, 1, \ldots, L+1]$ , such that  $\omega_0 < \omega_1 < \ldots < \omega_{L+1}$ , and such that

$$E(e^{j\omega_k}) = -E(e^{j\omega_{k+1}}) = \pm \left( \|E(e^{j\omega})\|_{\infty} \right).$$

## Alternation Theorem Example – 5<sup>th</sup> order







Penn ESE 531 Spring 2019 - Khanna



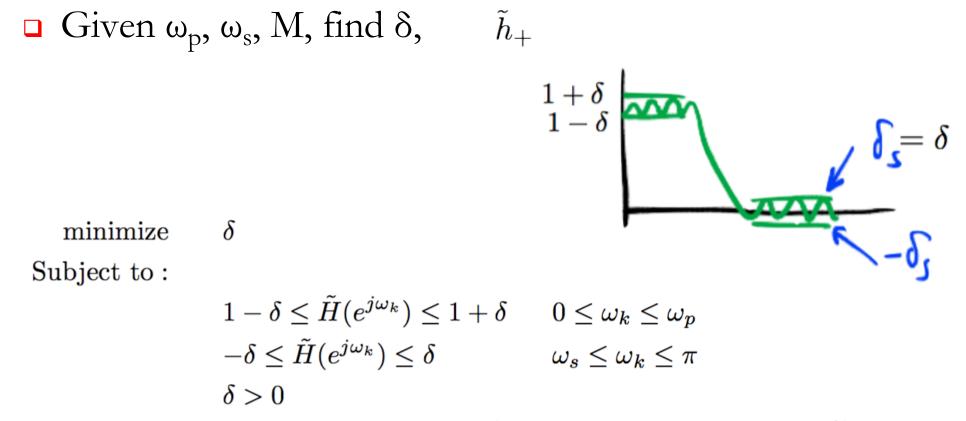
- Constraints:
  - min-max pass-band ripple

$$1 - \delta_p \le |H(e^{j\omega})| \le 1 + \delta_p, \qquad 0 \le w \le \omega_p$$

min-max stop-band ripple

$$|H(e^{j\omega})| \le \delta_s, \qquad \omega_s \le w \le \pi$$





Formulation is a linear program with solution δ, *h*<sub>+</sub>
 A well studied class of problems with good solvers



 $\begin{array}{ll} \text{minimize} & \delta\\ \text{subject to}: & \\ & 1-\delta \leq \end{array}$ 

$$1 - \delta \preceq A_p \tilde{h}_+ \preceq 1 + \delta$$
$$-\delta \preceq A_s \tilde{h}_+ \preceq \delta$$
$$\delta > 0$$

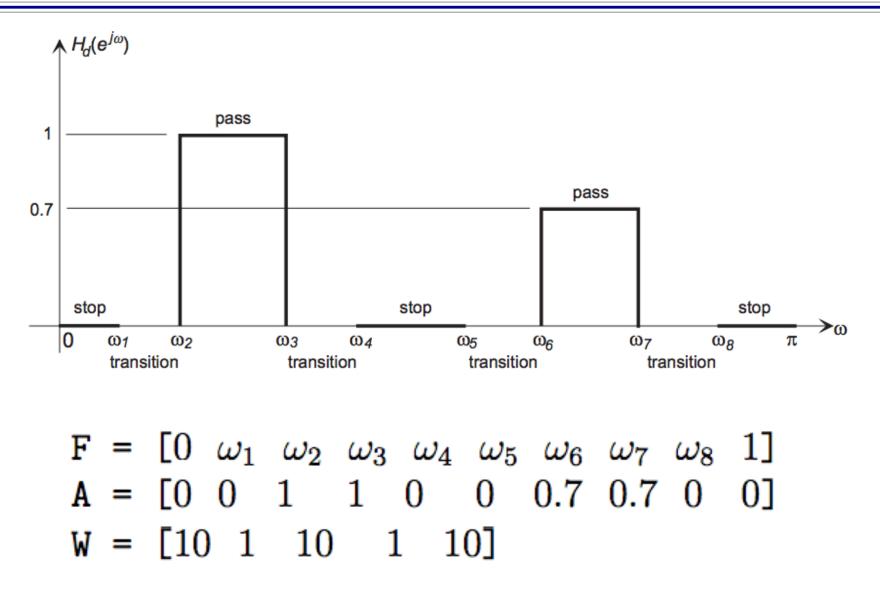
$$A_{p} = \begin{bmatrix} 1 & 2\cos(\omega_{1}) & \cdots & 2\cos(\frac{M}{2}\omega_{1}) \\ \vdots & & \\ 1 & 2\cos(\omega_{p}) & \cdots & 2\cos(\frac{M}{2}\omega_{p}) \end{bmatrix}$$
$$A_{s} = \begin{bmatrix} 1 & 2\cos(\omega_{s}) & \cdots & 2\cos(\frac{M}{2}\omega_{s}) \\ \vdots & & \\ 1 & 2\cos(\omega_{P}) & \cdots & 2\cos(\frac{M}{2}\omega_{P}) \checkmark \end{bmatrix}$$
capital P

# MATLAB Parks-McClellan Function

### b = firpm(M,F,A,W)

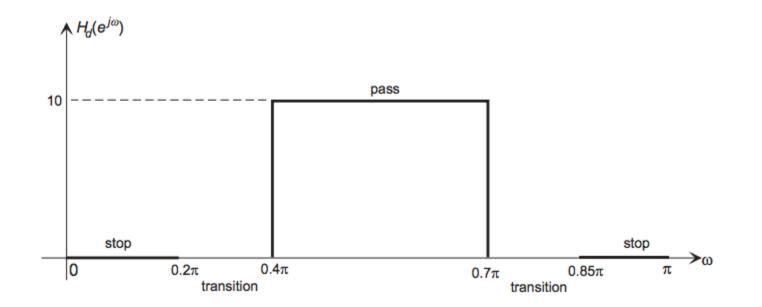
- b is the array of filter coefficients (impulse response)
- M is the filter order (M+1 is the length of the filter),
- **F** is a vector of band edge frequencies in ascending order
- A is a set of filter gains at the band edges
- W is an optional set of relative weights to be applied to each of the bands





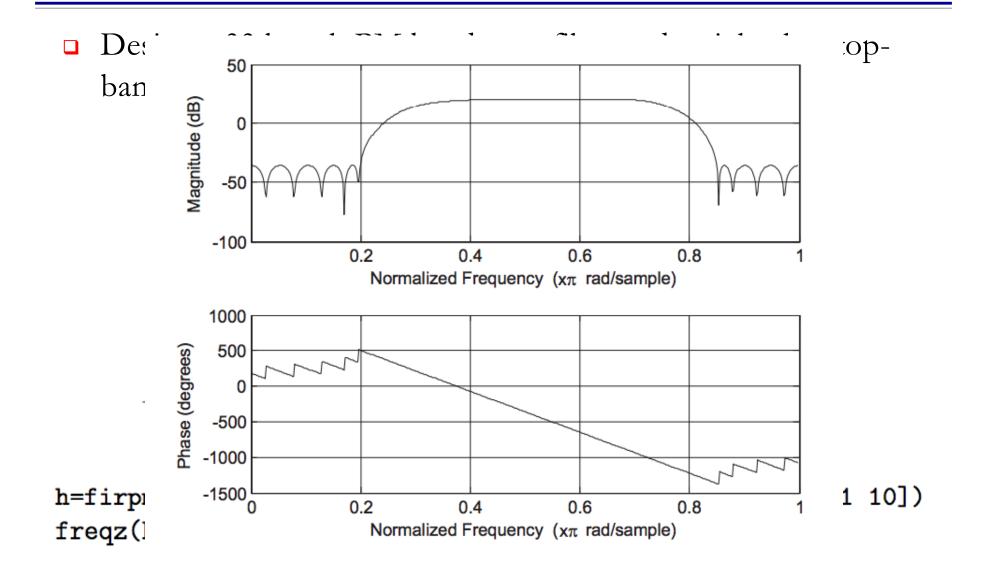


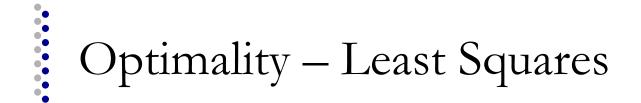
Design a 33 length PM ???? filter and weight the stop-band ripple 10x more than the pass-band ripple

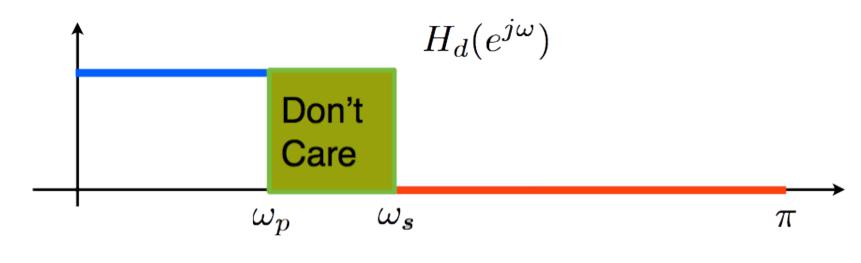


h=firpm(32,[0 0.2 0.4 0.7 0.85 1],[0 0 10 10 0 0],[10 1 10])
freqz(h,1)









□ Least Squares:

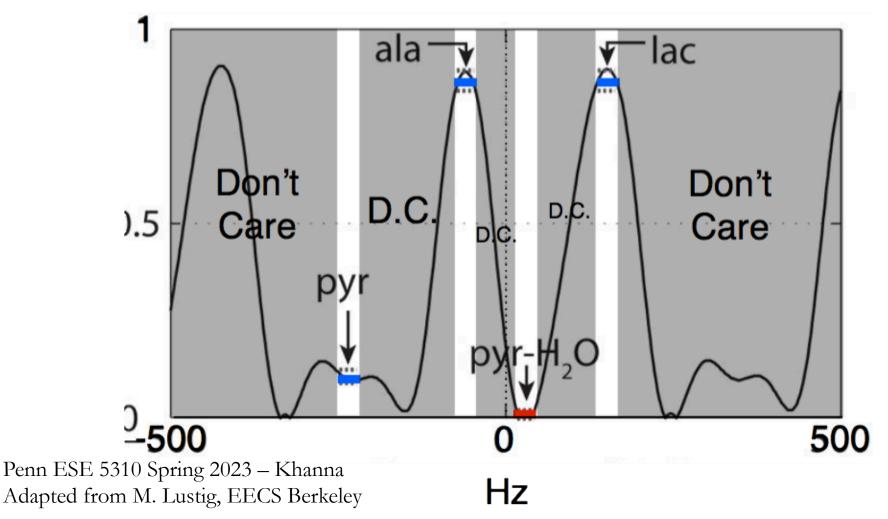
minimize 
$$\int_{\omega \in \text{care}} |H(e^{j\omega}) - H_d(e^{j\omega})|^2 d\omega$$

Parks-McClellan

$$\min_{\{h_e[n]:0\leq n\leq L\}}\Big(\max_{\omega\in F}|E(\omega)|\Big),$$

# Example of Complex Filter

- Larson et. al, "Multiband Excitation Pulses for Hyperpolarized 13C Dynamic Chemical Shift Imaging" JMR 2008;194(1):121-127
- Need to design length 11 filter with following frequency response:





- Many tools and Solvers
- **D** Tools:
  - CVX (Matlab) <u>http://cvxr.com/cvx/</u>
  - CVXOPT, CVXMOD (Python)
- Optimization Engines:
  - Sedumi (Free)
  - MOSEK (commercial)

Using CVX (in Matlab)

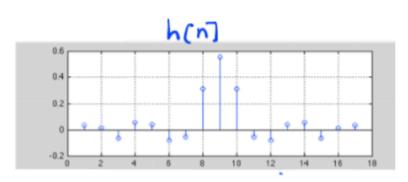
M = 16; wp = 0.5\*pi; ws = 0.6\*pi; MM = 15\*M; w = linspace(0,pi,MM);

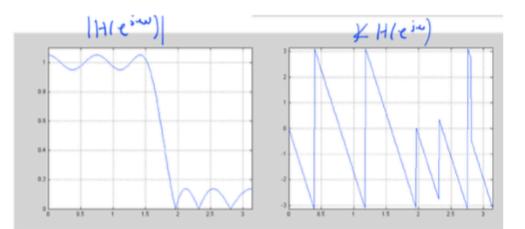
idxp = find(w <=wp); idxs = find(w >=ws);

Ap = [ones(length(idxp),1) 2\*cos(kron(w(idxp)', [1:M/2]))];As = [ones(length(idxs),1) 2\*cos(kron(w(idxs)', [1:M/2]))];

```
% optimization
cvx_begin
variable hh(M/2+1,1);
variable d(1,1);
```

```
minimize(d)
subject to
    Ap*hh <=1+d;
    Ap*hh >=1-d;
    As*hh < d;
    As*hh > -d;
    ds>0;
cvx_end
h = [hh(end:-1:1); hh(2:end)];
```







- Project1 out now
  - Due Tuesday 3/28
- Project 1 Submission
  - 2 assignments on Gradescope
  - Project 1 Report
    - Late penalty judged off this
  - Project 1 code
    - Can submit multiple files any format
- Project questions?