ECE 4213/5213

MODULE 10

Index

FIR Filter Design

- Many FIR design techniques start with an estimate of the filter order.
- The general procedure goes in a loop:

1. Estimate order N
2. Design filter
3. Does it meet the spec?

NO

4. $N = N + 1$

YES

Done.

- If the original estimate for N <u>does</u> meet the spec, then you might also want to try decreasing the order by 1 to see if that will still meet the spec.
    - This is because the formulas for estimating the order are all approximations

- A widely used general formula for estimating the order of an FIR filter: the "Kaiser" formula (due to Jim Kaiser):

$$N \approx \frac{-20\log_{10}\left[\sqrt{\delta_p \delta_s}\right] - 13}{14.6\,|\omega_s - \omega_p| \,/\, (2\pi)} \qquad (*)$$

- Many specific FIR filter design methods also have their own specialized formulas for estimating the order.
  → In that case, you should use the formula that is specialized for the technique instead of the general formula above.

NOTE: For the Kaiser window design method, there is a second specialized "Kaiser" formula for estimating the order (see p. 10.20)

→ For a Kaiser Window design, you should always use that second formula instead of the one above.

NOTE: For a normalized spec, use $\delta_p = \frac{1}{2\sqrt{1+\varepsilon^2}}$ and $\delta_s = \frac{1}{A}$ in the formula $(*)$ above.
(This is customary)

- A more accurate but slightly more complicated approach for a normalized spec is to (see the figures on notes pages 9.1 and 9.2) divide everything on the vertical axis of the un-normalized spec (p. 9.1) by $1+\delta_p$ to obtain equivalency with the normalized spec (p. 9.2). This gives:

$$\underbrace{\frac{1-\delta_p}{1+\delta_p}}_{\text{p. 9.1}} = \underbrace{\frac{1}{\sqrt{1+\varepsilon^2}}}_{\text{p. 9.2}} \Rightarrow \delta_p = \frac{\frac{1}{\sqrt{1+\varepsilon^2}} - 1}{\frac{1}{\sqrt{1+\varepsilon^2}} + 1}$$

It then follows that:

$$\underbrace{\frac{\delta_s}{1+\delta_p}}_{\text{p. 9.1}} = \underbrace{\frac{1}{A}}_{\text{p. 9.2}} \Rightarrow \delta_s = \frac{1+\delta_p}{A}$$

→ Given a normalized spec, these values for δ_p and δ_s can then be used in the Kaiser formula (*) on p. 10.2.

Windowed FIR Filter Design

- Also called "the window method."
- Begins with a "desired" filter

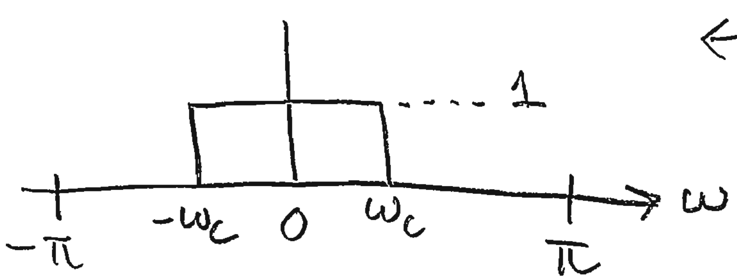
$$D(e^{j\omega}) \xleftrightarrow{\text{DTFT}} d[n].$$

Matlab:

fir1

- Often, $D(e^{j\omega})$ is one of the ideal filters
  - They are <u>not</u> realizable in practice
    - Cannot be causal
    - Cannot be stable
- we will usually assume that the desired filter is an ideal low-pass filter with cuttoff frequency ω_c.
- In this case, we will often refer to $D(e^{j\omega})$ as $H_{LP}(e^{j\omega})$ and $d[n]$ as $h_{LP}[n]$ ... where "LP" stands for "Low-Pass".
- So the desired (but unrealizable) filter will usually be:

$$H_{LP}(e^{j\omega}) \quad \xleftrightarrow{\text{DTFT}} \quad h_{LP}[n] = \frac{\sin \omega_c n}{\pi n}$$



Main idea of the window method:

- Multiply the infinite-length desired impulse response $d[n]$ (usually denoted $h_{LP}[n]$) by a finite-length window function $w[n]$ to "chop it off" and make it finite length.

→ This gives a finite-length practical impulse response $h_1[n] = w[n]\,d[n]$.

- $h_1[n]$ can then be shifted as needed to get an FIR filter $h[n] = h_1[n-n_0] = w[n-n_0]\,d[n-n_0]$ that is both causal and stable (choose n_0 as needed to make $h[n]=0 \ \forall \ n<0$).

NOTE:

- In Oppenheim & Schafer (Section 7.5) the desired frequency response is called $H_d(e^{j\omega})$ (7.53) and the desired impulse response is called $h_d[n]$ (7.54). But we will use the notation above on this page instead... for agreement with old tests and old handouts.

- Also, in Oppenheim & Schafer, the window functions are defined for $0 \le n \le M$, making "M" the filter order and "$M+1$" the length of $h[n]$.

- We are going to define things in a little bit different way... we will use "M" to indicate the half-length of the window and "N" to indicate the order of the filter.
  - This will make things simpler!
  - But slightly less optimal... because our procedure will always result in a filter order N that is even.

  - Our windows $w[n]$ will be defined for $-M \le n \le M$, where $M \in \mathbb{Z}$ is an integer.
  - Thus,
    - → The length of $w[n]$ will be $2M+1$.
    - → The length of $h_1[n] = w[n]\,d[n]$ will also be $2M+1$.
    - → The filter order will be $N = 2M$.
    - → The shift amount needed to make the filter causal ("n_0" on p. 10.5) will be $n_0 = M$.
    - → The length of $h[n] = h_1[n-M]$ will also be $2M+1$ with order $N = 2M$.

The Effect of Windowing:

- Since $h_1[n] = w[n]\,d[n]$, the DTFT frequency convolution property tells us that:

  → the frequency response of the designed practical filter, prior to shifting by M to make it causal, will be given by

$$H_1(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(e^{j\theta})\,D(e^{j(\omega-\theta)})\,d\theta$$

$$= \frac{1}{2\pi} W(e^{j\omega}) * D(e^{j\omega}).$$

⇒ In the windowed FIR filter design method, the desired frequency response $D(e^{j\omega})$ gets <u>convolved</u> with $W(e^{j\omega})$, the DTFT of the window function.

- This means that the frequency response magnitude $|H_1(e^{j\omega})|$ of the designed filter (prior to shifting to make it causal) will be <u>distorted</u> compared to the desired magnitude response $|D(e^{j\omega})|$.
- Shifting the designed impulse response $h_1[n]$ to get a causal filter $h[n] = h_1[n-M]$ will also introduce nonzero phase.

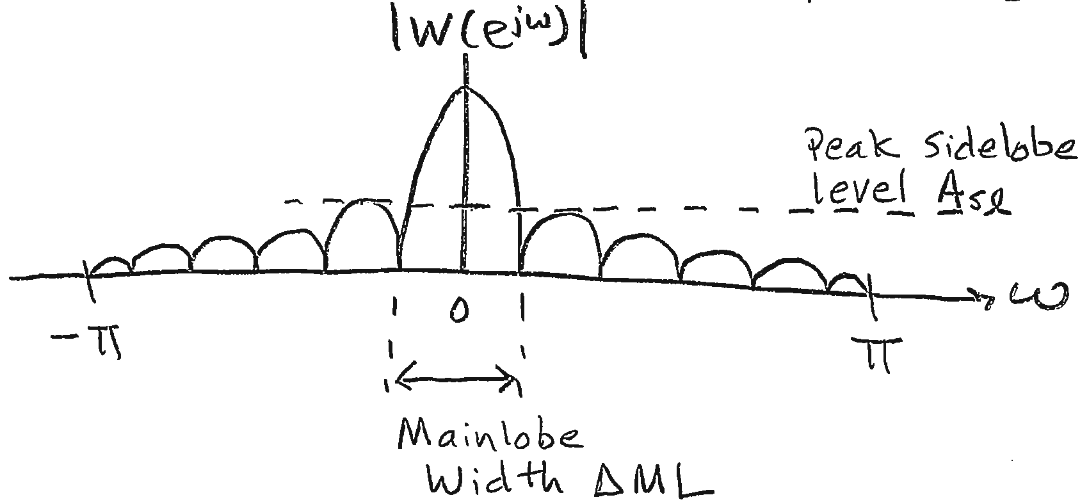
  → But if $d[n]$ and $w[n]$ both have symmetry (usually the case), ⟶

→ Then the final designed filter (causal and stable) $h[n] = h_1[n-M]$ will have linear phase or generalized linear phase.

- Thus, in designing a window function $w[n]$, we want:
  1. $w[n]$ to be symmetric about $n=0$ so that $h[n] = h_1[n-M]$ will have linear phase or generalized linear phase,
  2. $w[n]$ to have finite-length $2M+1$ so that the designed filter can be realizable and stable... ideally we want $2M+1$ to be as small as possible to minimize the delay and the implementation cost.
  3. $W(e^{j\omega})$ to be skinny and "impulse like" so that the convolution on page 10.7 causes as little distortion as possible.
- But this is <u>hard</u>!
  - Because the uncertainty principle tells us that 2 and 3 are <u>conflicting goals</u>.
- Many people became world famous by figuring out how to design a good window function $w[n]$.

  e.g. Hann, Hamming, Blackman, Kaiser, and others.

- The main window functions $w[n]$ that are used today all have DTFT magnitudes that are "sync-like":



- For good agreement between $H_1(e^{j\omega})$ and $D(e^{j\omega})$, we need:

    - ΔML as narrow as possible
    - A_{sl} as low as possible

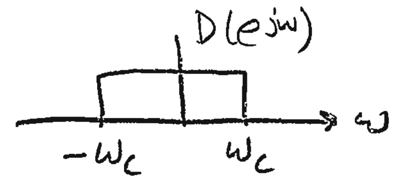
  → But these goals conflict with each other

  → They also conflict with the goal of keeping the length (and hence the order) as small as possible.

- The different window functions (Hann, Hamming, Blackman, etc) trade off between these conflicting goals differently.

- In all cases, the nature of the distortion in $H_1(e^{j\omega})$ relative to $D(e^{j\omega})$ will be two-fold:

(A) Instead of having the sharp transition of the ideal filter



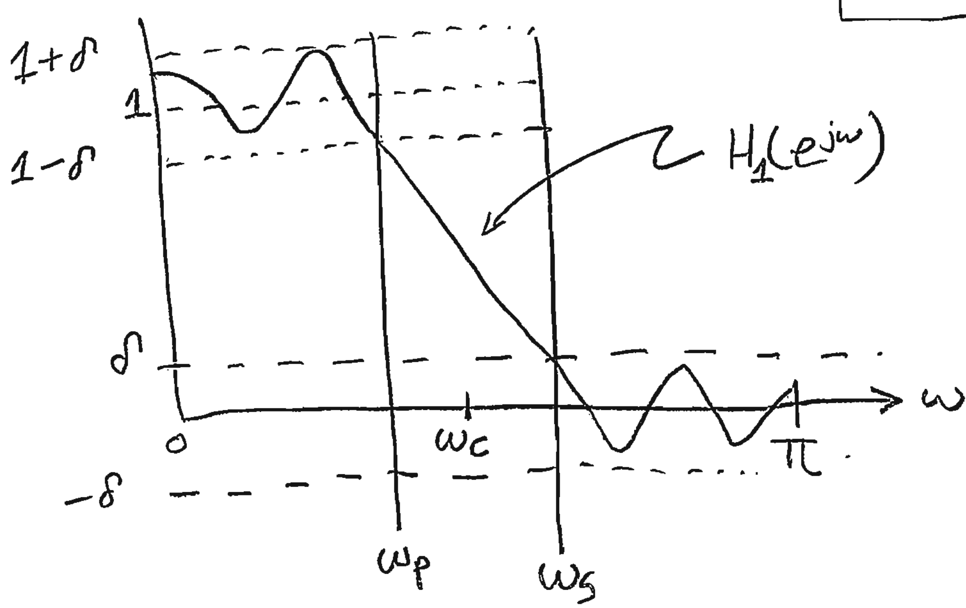
)

the designed practical filter $H_1(e^{j\omega})$ [prior to shifting to make it causal] will have a nonzero transition band going from a passband edge frequency ω_p to a stopband edge frequency ω_s.

(B) The passband and stopband of $H_1(e^{j\omega})$ will both show rippling that results from convolution of $D(e^{j\omega})$ with $W(e^{j\omega})$ [as shown on page 10.9 and indicated by the convolution on page 10.7].

EX;

Usually,
$\delta_s = \delta_p \equiv$ "δ"



- The transition bandwidth is defined by $\Delta\omega = |\omega_s - \omega_p|$.

- The filter $h_1[n] = w[n]d[n]$ will generally have zero phase or generalized zero phase. But when we shift it to make it causal, the final designed filter $h[n] = h_1[n-M]$ will have linear phase or generalized linear phase, as we said on p. 10.8.

Overview of the Window Design Method for FIR Filters

- The common window functions $w[n]$ are given on page 10.14. They are all nonzero for $-M \le n \le M$.
- There is also a table that gives the main properties of the window functions. It is similar to Table 7.2 in the Oppenheim & Schafer book.

- Given a digital filter design spec $\delta_p, \delta_s, \omega_p, \omega_s$:
  - It is assumed that $\delta_p = \delta_s$. If they are different, then use $\delta_s = \min(\delta_p, \delta_s)$. [use the more stringent spec]
  - Let $\alpha_s = -20 \log_{10} \delta_s$.
  - Let $\Delta\omega = \omega_s - \omega_p$.
  - Use the table to find which window functions can satisfy the spec by providing at least α_s dB of stopband attenuation.
  - For each of the window functions that can satisfy the stopband spec, use the formula in the last column of the table to solve for M.

$\longrightarrow$

- The formula from the table will not usually give an integer! It will be a floating point number!
  - → But our M <u>must</u> be an integer... because it is the half-length of $w[n]$.
  - → The formula gives the <u>minimum</u> value that can satisfy the spec... M must be <u>at least</u> as big.
  - → Therefore, you <u>must</u> <u>round</u> <u>up</u> to the next integer.
  - → For example, if the formula tells you that M must be at least 2.0001, then you have to use $\underline{\underline{M=3}}$.
- Choose the window function that satisfies the stopband spec with the smallest M. The table is arranged so that this will be the the first window (highest in the table) that can meet the spec.
  - Using the solved value of M for the chosen window,
    - → The filter order is $N = 2M$
    - → The length of $w[n]$, $h_1[n]$, and $h[n]$ is $2M+1$
- Let $\omega_c = \dfrac{\omega_p + \omega_s}{2}$. This puts ω_c for the ideal filter in the center of the transition band of the designed filter.
- Let $h_{LP}[n] = \dfrac{\sin(\omega_c n)}{\pi n}$.
  - → This is the impulse response $d[n]$ of the desired ideal low-pass filter.

→

- Let $h_1[n] = w[n]\, h_{LP}[n]$, $\quad -M \le n \le M$

($h_1[n]$ is <u>zero</u> outside of $-M \le n \le M$)

→ This is a stable low pass FIR that meets the spec (at least within the approximations inherent in the table) and has zero phase or generalized zero phase.

→ But it is <u>not</u> causal.

- Shift to make it causal. This gives the final designed filter

$$h[n] = h_1[n-M] = w[n-M]\, h_{LP}[n-M], \quad 0 \le n \le 2M.$$

→ causal ($h[n]=0 \ \forall\ n<0$)
→ stable ($\sum_{n=-\infty}^{\infty} |h[n]| < \infty$)
→ linear phase or generalized linear phase (provided $w[n]$ and $h_{LP}[n]$ have symmetry)

→ order: $N = 2M$

→ Length of $h[n]$: $2M+1$

- The next page (10.14) gives:
→ The common window functions
→ Table of the main properties of the window functions
→ Concise version of the design steps.

Common Window Functions for FIR filter design:

$$\textit{Rectangular:} \quad w[n] = 1, \qquad -M \leq n \leq M,$$

$$\textit{Hann:} \quad w[n] = \frac{1}{2}\left[1 + \cos\left(\frac{\pi n}{M}\right)\right], \qquad -M \leq n \leq M,$$

$$\textit{Hamming:} \quad w[n] = 0.54 + 0.46\cos\left(\frac{\pi n}{M}\right), \qquad -M \leq n \leq M,$$

$$\textit{Blackman:} \quad w[n] = 0.42 + 0.5\cos\left(\frac{\pi n}{M}\right) + 0.08\cos\left(\frac{2\pi n}{M}\right), \qquad -M \leq n \leq M.$$

Main Properties of the Window Functions:

Type of Window	**Main Lobe Width Δ_{ML}**	**Relative Sidelobe Level $A_{s\ell}$**	**Minimum Stopband Attenuation**	**Transition Bandwidth $\Delta\omega$**
Rectangular	$4\pi/(2M+1)$	13.3 dB	20.9 dB	$0.92\pi/M$
Hann	$8\pi/(2M+1)$	31.5 dB	43.9 dB	$3.11\pi/M$
Hamming	$8\pi/(2M+1)$	42.7 dB	54.5 dB	$3.32\pi/M$
Blackman	$12\pi/(2M+1)$	58.1 dB	75.3 dB	$5.56\pi/M$

Design Steps:

1. Convert the minimum stopband attenuation spec δ_s to dB using the formula $\alpha_s = -20\log_{10}\delta_s$.
2. Look in column 4 (Minimum Stopband Attenuation) of the table above to determine which window functions can provide at least α_s dB of stopband attenuation.
3. Let $\Delta\omega = \omega_s - \omega_p$. Use the last column of the table to figure out which window function $w[n]$ can meet the stopband spec with the smallest value M.
   - To do this, set $\Delta\omega$ equal to the formula in the last column of the table and solve for M. M must be an integer and you must always round up. For example, 2.001 means $M = 3$.
   - The order of your filter will be $2M$.
   - The length of $h[n]$ will be $2M+1$.
4. Let $\omega_c = \dfrac{\omega_p + \omega_s}{2}$.
5. Let $h_{LP}[n] = \dfrac{\sin(\omega_c n)}{\pi n}$.
6. For the window function $w[n]$ that meets the stopband spec with the smallest M, compute the "centered" impulse response $h_1[n] = w[n]h_{LP}[n]$, $-M \leq n \leq M$.
7. Shift it right by M to make it causal: $h[n] = h_1[n-M] = w[n-M]h_{LP}[n-M]$, $0 \leq n \leq 2M$.

EXAMPLE: WINDOWED FIR FILTER DESIGN

Use the window design method to design a causal lowpass FIR digital filter that meets the following specifications:

Passband Edge Freq.	$\omega_p = 0.3\pi$ rad/sample
Stopband Edge Freq.	$\omega_s = 0.35\pi$ rad/sample
Max. Passband Ripple	$\delta_p = 0.01$
Max. Stopband Ripple	$\delta_s = 0.01$

Give the filter impulse response $h[n]$.

① $\alpha_s = -20 \log_{10} \delta_s = 40$ dB

② The Hann, Hamming, and Blackman windows can all provide more than 40 dB of stopband attenuation.

→ We know that Hann will meet the spec with the smallest M (and therefore the smallest <u>order</u>) because: of the three that can meet the stopband spec, Hann is highest in the table. But let's go ahead and compute the "M" for all three windows to show this...

③ $\Delta\omega = \omega_s - \omega_p = 0.35\pi - 0.30\pi = 0.05\pi$

Rectangular: can <u>not</u> provide 40 dB of stopband attenuation.

Hann: $\Delta\omega = 0.05\pi = \frac{3.11\pi}{M} \Rightarrow M = \left\lceil \frac{3.11}{0.05} \right\rceil = \lceil 62.2 \rceil = 63$ ✓

Hamming: $\Delta\omega = 0.05\pi = \frac{3.32\pi}{M} \Rightarrow M = \left\lceil \frac{3.32}{0.05} \right\rceil = \lceil 66.4 \rceil = 67$

Blackman: $\Delta\omega = 0.05\pi = \frac{5.56\pi}{M} \Rightarrow M = \left\lceil \frac{5.56}{0.05} \right\rceil = \lceil 111.2 \rceil = 112$

⟶

→ Hann meets the spec with the smallest M.

⇒ Choose Hann → $M = 63$.

$\text{Order} = N = 2M = 126$

Length of $h[n]$ = length of $h_1[n] = 2M+1 = 127$

④ $\omega_c = \dfrac{\omega_p + \omega_2}{2} = \dfrac{0.30\pi + 0.35\pi}{2} = 0.325\pi$

⑤ $h_{LP}[n] = \dfrac{\sin \omega_c n}{\pi n} = \dfrac{\sin(0.325\pi n)}{\pi n}$

⑥ $h_1[n] = w[n]\, h_{LP}[n] = \frac{1}{2}\left[1 + \cos\left(\frac{\pi n}{M}\right)\right] \dfrac{\sin \omega_c n}{\pi n}$

$= \frac{1}{2}\left[1 + \cos\left(\frac{\pi n}{63}\right)\right] \dfrac{\sin(0.325\pi n)}{\pi n}, \quad -63 \le n \le 63.$

⑦ Shift to make causal:

$h[n] = h_1[n-M] = \frac{1}{2}\left[1 + \cos\left(\frac{\pi(n-63)}{63}\right)\right] \dfrac{\sin[0.325\pi(n-63)]}{\pi(n-63)},$

$0 \le n \le 126$

$$\boxed{h[n] = \frac{1}{2}\left[1 + \cos\left(\frac{\pi(n-63)}{63}\right)\right] \frac{\sin[0.325\pi(n-63)]}{\pi(n-63)}, \quad 0 \le n \le 126}$$

- After designing the filter, you should follow the procedure on p. 10.1 to verify that the designed filter meets the spec.

→ If not, then you can

Ⓐ increase M by one and repeat steps ④-⑦. This will make the transition band narrower.

-or-

Ⓑ switch to the next lower window type in the table. This will increase the stopband attenuation (and increase the order).

- But on a test or exam, you don't have Matlab and you don't have time... so you don't have to do this.

→ i.e., on a test or exam the procedure shown in the previous example completes the problem.

NOTE : as we said on p. 9.12, when you are doing a numerical optimization to design a filter that minimizes the figure of merit $\mathcal{E}$ with $p=2$ (i.e., least squares optimization), the filter that minimizes the mean squared error (MSE) between $H_1(e^{j\omega})$ and $D(e^{j\omega})$

$$\mathcal{E} = \int_{-\pi}^{\pi} \left| D(e^{j\omega}) - H_1(e^{j\omega}) \right|^2 d\omega$$

is always an FIR filter truncated by the rectangular window $w[n]$. ⟶

→ However, this filter will have significant passband ripple (a.k.a. "Gibbs phenomena") caused by the convolution of $D(e^{j\omega})$ with $W(e^{j\omega})$.

→ As the order N increases, the ripple becomes more localized at the transition band edge ω_p.

⇒ But the peak of the ripple will never decrease... it remains the same no matter how large you make M $(N=2M)$.

- For this reason, the optimal MSE filter is often not used.

<u>NOTE</u>: the FIR Window Design Method as we have described it here can be extended in a straightforward way to directly design other filter types like high-pass, bandpass, bandstop...

- The main idea is:
  - instead of using $h_{LP}[n] = \frac{\sin \omega_c n}{\pi n}$ for the desired ideal filter impulse response $d[n]$,
  - you set $d[n]$ equal to one of the other ideal filter impulse responses given on pages 7.1 – 7.2.
  - $\Delta\omega$ and ω_c must then be modified as appropriate, but the procedure is otherwise straightforward.

Kaiser Window FIR Filter Design

- The window functions $w[n]$ given on p. 10.14 are called "fixed windows" because they do not have any adjustable parameters.

- There are other window types called "adjustable windows" that have parameters that can be tuned to improve the design.
  - These include the Dolph-Chebyshev window and the Kaiser window.
  - Here, we will restrict our attention to the Kaiser window because it is the most widely used adjustable window function.

- The Kaiser window function:

$$w[n] = \frac{I_0\left(\beta\sqrt{1-(n/M)^2}\right)}{I_0(\beta)}, \quad -M \leq n \leq M \qquad (*)$$

where:

- β, M are designed parameters
- $I_0(x)$ is a zeroth-order modified Bessel function:

$$I_0(x) = 1 + \sum_{r=1}^{\infty}\left[\frac{(x/2)^r}{r!}\right]^2$$

$\rightarrow$ In practice, it is customory to retain only the first 20 terms of the sum.

- To use a Kaiser window, you must design M and β.
  - Let $\alpha_s = -20\log_{10}\delta_s$
  - Let $\Delta\omega = |\omega_s - \omega_p|$
  - Estimate the filter order using the special Kaiser formula

$$N = \frac{\alpha_s - 8}{2.285\,\Delta\omega}$$

$\rightarrow$ Then $M = \frac{N-1}{2}$.

- Design β according to the Table below:

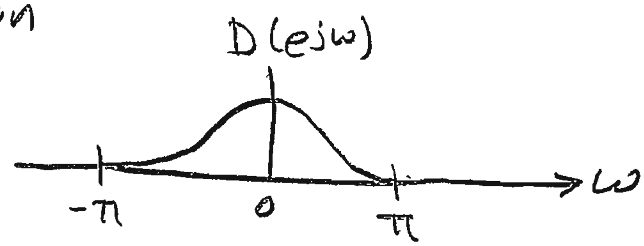
α_s	β
$\alpha_s > 50$	$0.1102(\alpha_s - 8.7)$
$21 \leq \alpha_s \leq 50$	$0.5842(\alpha_s - 21)^{0.4} + 0.07886(\alpha_s - 21)$
$\alpha_s < 21$	0

- plug in the designed values of M and β to the window function (*) on p. 10.19.

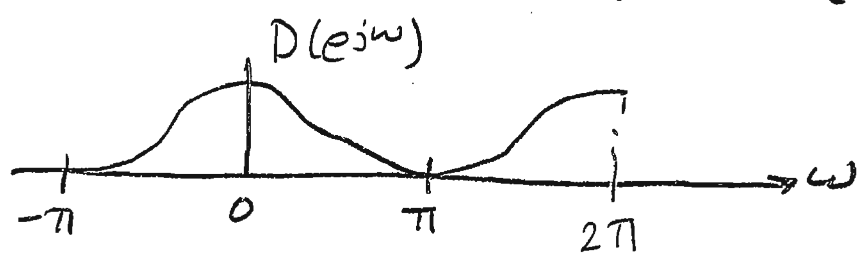
- Follow steps 4-7 on p. 10.14 to complete the filter design.

FIR Filter Design by Frequency Sampling

- This technique is widely used in image processing.
- Also used in 1D signal processing, albeit to a lesser extent than in image processing (which is 2D).
- Start with a desired frequency response curve $D(e^{j\omega})$ defined from $\omega = -\pi$ to $\omega = \pi$.
  - This is often obtained by setting $D(e^{j\omega})$ from $-\pi$ to π equal to a known, desirable analog function like a Gaussian



- Extend $D(e^{j\omega})$ periodically from $\omega = \pi$ to $\omega = 2\pi$, e.g.



- Define a <u>DFT</u> $H[k]$ by sampling $D(e^{j\omega})$ at N equally spaced frequencies from $\omega = 0$ to $\omega = 2\pi$:

$$H[k] = D\left(e^{j2\pi k/N}\right), \quad k = 0, 1, \ldots, N-1$$

- Obtain the FIR filter impulse response $h[n]$ by inverse DFT:

$$h[n] = \frac{1}{N}\sum_{k=0}^{N-1} H[k]\, e^{j2\pi nk/N}, \quad 0 \le n \le N-1$$

- The length of $h[n]$ is N.
- The filter order is $N-1$.

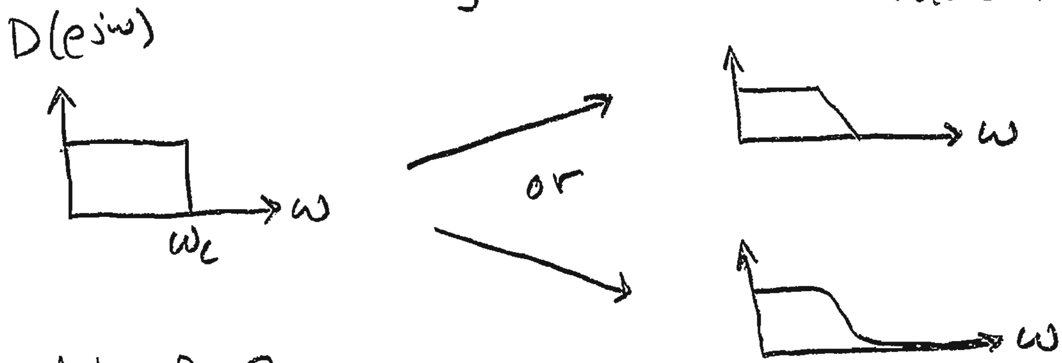
→ If $D(e^{j\omega})$ has a reasonably smooth transition band and if there are "enough" samples on the curve so that the samples have a shape that looks like $D(e^{j\omega})$, then the approach usually works well.

- There will generally be aliasing in $h[n]$ that results from the sampling of $D(e^{j\omega})$. Specifically,

$$h[n] = \sum_{k=-\infty}^{\infty} d[n+kN], \quad 0 \le n \le N-1.$$

→ But this is typically not a concern if $d[n]$ decays towards zero rapidly for $|n| > N$.

- If the transition band of $D(e^{j\omega})$ is very abrupt, as with an ideal low-pass filter, the aliasing can be reduced by inserting a smoother transition band:



- Matlab: fir2

Impulse Invariant Design

- Start with a desirable analog filter $H_a(s)$ having impulse response $h_a(t)$.

- Main idea: sample $h_a(t)$ on a finite interval to get $h[n]$:

$$h[n] = h_a(nT), \quad 0 \leq n \leq N-1$$

length of $h[n]$: N
order of filter: $N-1$

  - As in windowing, there will be distortion of the desired frequency response due to truncation (unless $h_a(t)$ just happens to be finite length... usually not the case).

  - As in frequency sampling, there will generally be aliasing. Not too much of a concern as long as $h_a(t)$ decays rapidly towards zero as $t \to NT$ and as long as there are enough samples so that the shape of $h[n]$ looks like $h_a(t)$.

  - It can be shown that

$$H(e^{j\omega}) = \frac{1}{T} \sum_{k=-\infty}^{\infty} H_a\left(\frac{\omega}{T} - \frac{2\pi k}{T}\right).$$

- Matlab: impinvar

Parks-McClellan Algorithm

- It is a numerical technique for designing equirriple linear phase FIR filters. They have:
  - equal ripple through the stopband
  - equal but different ripple through the passband
  - linear phase.
- This is almost always implemented on the computer.
- You will get some experience with it in the homework (Chapter 7 of the Lab Manual).

- Matlab: firpm