10 FIR Digital Filter Design

FIR Digital Filter Design

- For IIR filters, it is necessary to ensure that the derived transfer function $G(z)$ is stable
- In the case of FIR filters, the stability is not an issue as the transfer function is a polynomial in z^{-1} and the stability is always guaranteed
- Unlike the IIR filter design problem, it is always possible to design FIR digital filters with exactly linear phase response

FIR Digital Filter Design

- FIR filter design does not have any connection with the analog filters
- The design of FIR filters is therefore based on a direct approximation of the specified magnitude response, with the usually added requirement that the phase response be linear
- A causal FIR transfer function $H(z)$ of length $N+1$ is

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

- The corresponding frequency response is

$$H(e^{j\omega})=\sum_{n=0}^{N}h[n]e^{-j\omega n}$$

FIR Digital Filter Design

- It has been shown that any finite duration sequence of length $N+1$ is completely characterized by $N+1$ samples of its discrete-time Fourier transform $X(e^{j\omega})$
- As a result, the design of an FIR filter of length can be accomplished by finding either the impulse response sequence $\{h[n]\}$ or $N+1$ samples of its frequency response $H(e^{j\omega})$
- To ensure the linear-phase design, the symmetry condition of the impulse response must be satisfied

$$h[n]=\pm h[N-n]$$

Basic Approaches to FIR Filter Design

- Basic approaches in designing FIR filters:
  1) Truncating the Fourier series representation of the desired frequency response
     => Window method
  2) Frequency sampling
     Length N FIR filter, N distinct equally spaced frequency samples of the desired frequency response constitute the N-point DFT of its impulse response
  3) Computer-aided design based on optimization

Truncating the Impulse Response

- Let $H_d(e^{j\omega})$ denote the desired frequency response function
- $H_d(e^{j\omega})$ is periodic function of ω with period 2π and can be expressed as a Fourier series

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

- The Fourier coefficients $\{h_d[n]\}$ are the impulse response samples

$$h_d[n]=\frac{1}{2\pi}\int_{-\pi}^{\pi}H_d(e^{j\omega})e^{j\omega n}d\omega,\quad -\infty\le n\le\infty$$

Truncating the Impulse Response

- Thus, given $H_d(e^{j\omega})$ we can compute $h_d[n]$ and the corresponding $H_d(z)$
- Usually, $H_d(e^{j\omega})$ is piecewise constant with ideal (or sharp) transitions between bands

  => $\{h_d[n]\}$ sequence is of infinite length and noncausal
- The objective is to find a finite-duration impulse response $\{h_t[n]\}$ of length $2M+1$ whose DTFT $H_t(e^{j\omega})$ approximates the desired DTFT $H_d(e^{j\omega})$

Truncating the Impulse Response

- Minimizing the integral squared error

$$\Phi = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|H_t(e^{j\omega}) - H_d(e^{j\omega})\right|^2 d\omega$$

where $H_t(e^{j\omega}) = \sum_{n=-M}^{M} h_t[n]e^{-j\omega n}$

- Using the Parseval's relation

$$\Phi = \sum_{n=-\infty}^{\infty}\left|h_t[n] - h_d[n]\right|^2 = \sum_{n=-M}^{M}\left|h_t[n] - h_d[n]\right|^2 + \sum_{n=-\infty}^{-M+1} h_d[n]^2 + \sum_{n=m+1}^{\infty} h_d[n]^2$$

- Now, Φ is minimum when $h_t[n] = h_d[n]$ for $-M \le n \le M$, i.e., the best finite-length approximation is obtained by truncating the impulse response

Truncating the Impulse Response

- A causal impulse response $h[n]$ can be obtained from $h_t[n]$ by delaying it with M samples

$$h[n] = h_t[n - M]$$

- $h[n]$ has the same magnitude response as $h_t[n]$ but its phase response has a linear phase shift of ωM radians
- The group delay of $h[n]$ is M samples

$$\tau(\omega) = -\frac{d}{d\omega}(-\omega M) = M$$

where the linear phase response is $\theta(\omega) = -\omega M$

Impulse Response of Ideal Lowpass Filter

- The ideal lowpass filter has a zero-phase frequency response

$$H_{LP}(e^{j\omega}) = \begin{cases} 1, & |\omega| \le \omega_c \\ 0, & \omega_c < |\omega| \le \pi \end{cases}$$

- The corresponding impulse response coefficients

$$h_{LP}[n] = \frac{\sin \omega_c n}{\pi n}, \quad -\infty < n < \infty$$

- $h_{LP}[n]$ is doubly infinite, not absolutely summable, and therefore unrealizable

Impulse Response of Ideal Lowpass Filter

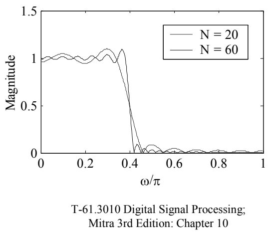
- Truncating to range $-M \le n \le M$ and delaying with M samples yields the causal FIR lowpass filter

$$h'_{LP}[n] = \begin{cases} \dfrac{\sin(\omega_c(n-M))}{\pi(n-M)}, & 0 \le n \le N-1 \\ 0 & \text{otherwise} \end{cases}$$

- The truncation of the impulse response coefficients of the ideal filters exhibit an oscillatory behavior in the respective magnitude responses
- This is commonly referred to as the ***Gibbs phenomenon***

Gibbs Phenomenon

- Gibbs phenomenon - Oscillatory behavior in the magnitude responses of causal FIR filters obtained by truncating the impulse response coefficients of ideal filters



Gibbs Phenomenon

- As can be seen, as the length of the lowpass filter is increased, the number of ripples in both passband and stopband increases, with a corresponding decrease in the ripple widths
- Height of the largest ripples remain the same independent of length
- Similar oscillatory behavior observed in the magnitude responses of the truncated versions of other types of ideal filters

Explanation of the Gibbs Phenomenon

- Truncation of $h_d[n]$ can be expressed by ***windowing operation***, i.e., by multiplying the $h_d[n]$ sequence with a finite-length sequence $w[n]$

$$h_t[n] = h_d[n]w[n]$$

where $w[n]$ is a window function

- For a rectangular window

$$w_R[n] = \begin{cases} 1, & 0 \le n \le M \\ 0, & \text{otherwise} \end{cases}$$

- The Gibbs phenomenon can be explained in the frequency domain by the convolution theorem

Explanation of the Gibbs Phenomenon

- Multiplication in the time domain corresponds to convolution in the frequency domain

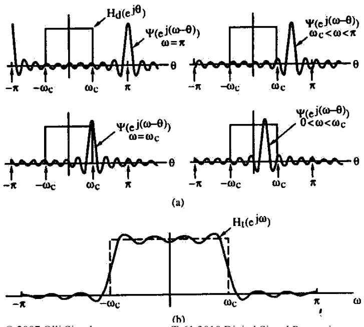
$$H_t(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_d(e^{j\varphi})\Psi(e^{j(\omega-\varphi)})d\varphi$$

where $H_t(e^{j\omega}) = \mathcal{F}\{h_t[n]\}$

$\Psi(e^{j\omega}) = \mathcal{F}\{w[n]\}$

- $H_t(e^{j\omega})$ is obtained by a periodic continuous convolution of the frequency response $H_d(e^{j\omega})$ with the Fourier transform $\Psi(e^{j\omega})$ of the window

Illustration of the Windowing

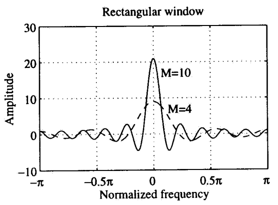


$$w[n] = w_R[n]$$

$$\Psi_R(e^{j\omega}) = \sum_{n=-M}^{M} e^{-j\omega n}$$

$$= \frac{\sin([2M+1]\omega/2)}{\sin(\omega/2)}$$

Illustration of the Windowing



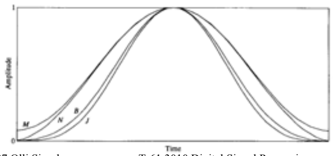
- The frequency response $\Psi(e^{j\omega})$ has a narrow ***main lobe*** centered at $\omega=0$
- All the other ripples in the frequency response are called ***sidelobes***
- The main lobe is characterized by its width $4\pi/(2M+1)$ defined by the first zero crossings on both sides of $\omega=0$
- As M increases the width of the main lobe decreases
- The area under each lobe remains constant, while the width of each lobe decreases with increasing M

Gibbs Phenomenon

- Rectangular window has an abrupt transition to zero outside the range $-M \le n \le M$, which results in Gibbs phenomenon in $H_t(e^{j\omega})$
- Gibbs phenomenon can be reduced either:
  1. Using a window that tapers smoothly to zero at each end, or
  2. Providing a smooth transition from passband to stopband in the magnitude specifications

Window Functions

- Symmetric window functions are used in FIR filter design in order to guarantee the linear phase response
- Smoother behavior at the cutoff frequency is obtained by using different cosine-type functions instead of the rectangular window



- Hamming (M)
- Hanning (N)
- Blackman (B)
- 30 dB/octave (J)

Window Functions

- Various window functions:

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

$$\text{Hamming}: \quad w[n] = 0.54 + 0.46\cos\left(\frac{2\pi n}{2M+1}\right), \quad -M \le n \le M$$

$$\text{Blackman}: \quad w[n] = 0.42 + 0.5\cos\left(\frac{2\pi n}{2M+1}\right) + 0.08\cos\left(\frac{4\pi n}{2M+1}\right), \quad -M \le n \le M$$

$$\text{Bartlett}: \quad w[n] = 1 - \frac{|n|}{M+1}, \quad -M \le n \le M$$

Window Functions

- Parameters to be compared:
  1) Main lobe width
  2) Relative sidelobe level (the largest sidelobe

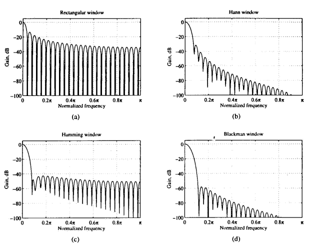
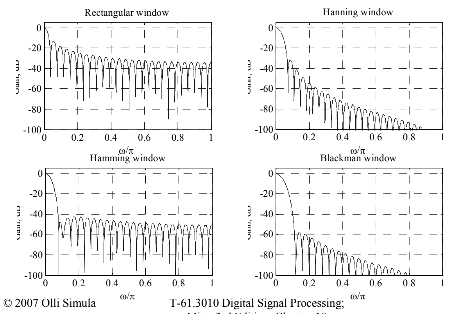


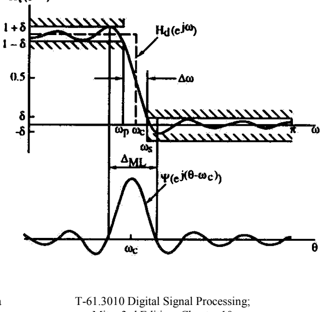
Figure 7.20 Gain response of the fixed window functions.

Fixed Window Functions

- Plots of magnitudes of the DTFTs of these windows for M = 25 are shown below:



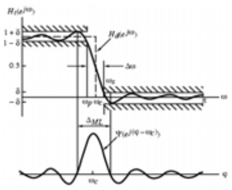
Lowpass Filter Design by Windowing



Fixed Window Functions

- Magnitude spectrum of each window characterized by a main lobe centered at $\omega = 0$ followed by a series of sidelobes with decreasing amplitudes
- Parameters predicting the performance of a window in filter design are:
  - **Main lobe width** Δ_{ML}
    given by the distance between zero crossings on both sides of main lobe
  - **Relative sidelobe level** A_{sl}
    given by the difference in dB between amplitudes of largest sidelobe and main lobe

Fixed Window Functions



- Observe $H_t(e^{j(\omega_c+\Delta\omega)}) + H_t(e^{j(\omega_c-\Delta\omega)}) \cong 1$
- Thus, $H_t(e^{j\omega_c}) \cong 0.5$
- Passband and stopband ripples are the same

Fixed Window Functions

- Distance between the locations of the maximum passband deviation and minimum stopband value is approximately $\approx\Delta_{ML}$
- Width of transition band is

$$\Delta\omega = \omega_s - \omega_p < \Delta_{ML}$$

Fixed Window Functions

- To ensure a fast transition from passband to stopband, window should have a very small main lobe width
- To reduce the passband and stopband ripple δ, the area under the sidelobes should be very small
- Unfortunately, these two requirements are contradictory

Fixed Window Functions

- In the case of rectangular, Hann, Hamming, and Blackman windows, the value of ripple does not depend on filter length or cutoff frequency ω_c, and is essentially constant
- In addition, the transition width is inversily proportional to the window length, i.e.,

$$\Delta\omega \approx \frac{c}{M}$$

where *c* is a constant for most practical purposes

Fixed Window Functions

Table 7.3 Properties of Some Fixed Window Functions.

Type of window	Main lobe width Δ_{ML}	Relative side-lobe level A_{sl}	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	$5.56\pi/M$

Fixed Window Functions

- Filter Design Steps:

(1) Set

$$\omega_c = (\omega_p + \omega_s)/2$$

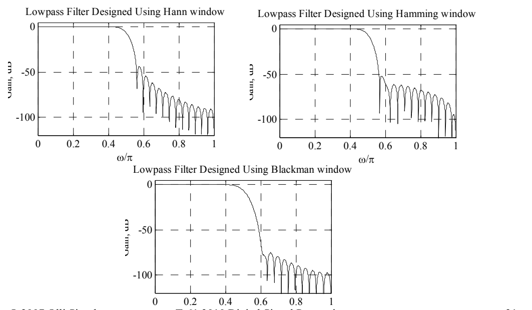
(2) Choose window based on specified α_s

(3) Estimate M using

$$\Delta\omega \approx \frac{c}{M}$$

FIR Filter Design Example

- Lowpass filter of length 51 and $\omega_c = \pi / 2$



FIR Filter Design Example

- An increase in the main lobe width is associated with an increase in the width of the transition band
- A decrease in the sidelobe amplitude results in an increase in the stopband attenuation
- Several windows have been developed that provide control over the ripple δ by means of an additional parameter characterizing the window

⇨ **Adjustable window functions**

Adjustable Window Functions

- **Dolph-Chebyshev Window:**

$$w[n] = \frac{1}{2M+1}\left[\frac{1}{\gamma} + 2\sum_{k=1}^{M} T_k\left(\beta \cos\frac{k}{2M+1}\right)\cos\frac{2nk\pi}{2M+1}\right], \quad -M \le n \le M$$

where

$$\gamma = \frac{amplitude\ of\ sidelobe}{main\ lobe\ amplitude}, \quad \beta = \cosh\left(\frac{1}{2M}\cosh^{-1}\frac{1}{\gamma}\right)$$

and

$$T_\ell(x) = \begin{cases} \cos(\ell \cos^{-1} x), & |x| \le 1 \\ \cosh(\ell \cosh^{-1} x), & |x| > 1 \end{cases}$$

Adjustable Window Functions

- Dolph-Chebyshev window can be designed with any specified relative sidelobe level while the main lobe width adjusted by choosing length appropriately
- Filter order is estimated using

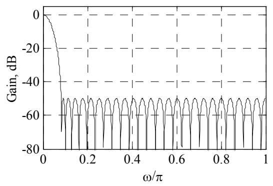
$$N = \frac{2.056\alpha_s - 16.4}{2.85(\Delta\omega)}$$

where $\Delta\omega$ is the normalized transition bandwidth, e.g, for a lowpass filter

$$\Delta\omega = \omega_s - \omega_p$$

Adjustable Window Functions

- Gain response of a Dolph-Chebyshev window of length 51 and relative sidelobe level of 50 dB is shown below



Adjustable Window Functions

Properties of Dolph-Chebyshev window:

- All sidelobes are of equal height
- Stopband approximation error of filters designed have essentially equiripple behavior
- For a given window length, it has the smallest main lobe width compared to other windows resulting in filters with the smallest transition band

Adjustable Window Functions

- **Kaiser Window :**

$$w[n] = \frac{I_0\{\beta\sqrt{1-(n/M)^2}\}}{I_0(\beta)}, \quad -M \le n \le M$$

where β is an adjustable parameter and is the modified zeroth-order Bessel function of the first kind:
$$I_0(u) = 1 + \sum_{r=1}^{\infty}\left[\frac{(u/2)^r}{r!}\right]^2$$

- Note $I_0(u) > 0$ for $u > 0$
- In practice
$$I_0(u) \cong 1 + \sum_{r=1}^{20}\left[\frac{(u/2)^r}{r!}\right]^2$$

Adjustable Window Functions

- Parameter β controls the minimum stopband attenuation of the windowed filter response
- β is estimated using

$$\beta = \begin{cases} 0.1102(\alpha_s - 8.7), & \text{for } \alpha_s > 50 \\ 0.5842(\alpha_s - 21)^{0.4} + 0.07886(\alpha_s - 21), & \text{for } 21 \le \alpha_s \le 50 \\ 0, & \text{for } \alpha_s < 21 \end{cases}$$

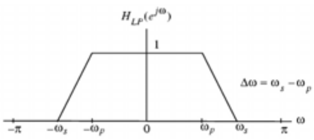
- Filter order is estimated using

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

where $\Delta\omega$ is the normalized transition bandwidth

Impulse Responses of FIR Filters with a Smooth Transition

- First-order spline passband-to-stopband transition



$$\omega_c = (\omega_p + \omega_s)/2$$

$$\Delta\omega = \omega_s - \omega_p$$

$$h_{LP}[n] = \begin{cases} \omega_c/\pi, & n = 0 \\ \frac{2\sin(\Delta\omega n/2)}{\Delta\omega n}\cdot\frac{\sin(\omega_c n)}{\pi n} & |n| > 0 \end{cases}$$

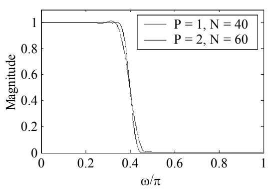
Impulse Responses of FIR Filters with a Smooth Transition

- P^{th}-order spline passband-to-stopband transition

$$h_{LP}[n] = \begin{cases} \omega_c/\pi, & n = 0 \\ \left(\frac{2\sin(\Delta\omega n/2P)}{\Delta\omega n/2P}\right)^P\cdot\frac{\sin(\omega_c n)}{\pi n} & |n| > 0 \end{cases}$$

Lowpass FIR Filter Design Example

- Example



Computer-Aided Design of FIR Filters

- Let $|H(e^{j\omega})|$ denote the the magnitude response of $H(z)$ designed to approximate the desired magnitude response $D(z)$:

$$E(\omega) = P(\omega)\,\left[\,|H(e^{j\omega})| - D(\omega)\right]$$

- The design is based on minimizing the weighted error function $E(z)$
- Minimax criterion minimizes the peak absolute value of the weighted error (Parks-McClellan algorithm):

$$E = \max_{0\le\omega\le\pi}|E(\omega)|$$

Design of Equiripple Linear-Phase FIR Filters

- The linear-phase FIR filter obtained by minimizing the peak absolute value of

$$\varepsilon = \max_{\omega \in R} |\mathcal{E}(\omega)|$$

is usually called the **equiripple FIR filter**

- After ε is minimized, the weighted error function $\mathcal{E}(\omega)$ exhibits an equiripple behavior in the frequency range R

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 10 43 Copyright © 2001, S. K. Mitra

Design of Equiripple Linear-Phase FIR Filters

- The general form of frequency response of a causal linear-phase FIR filter of length $2M+1$:

$$H(e^{j\omega}) = e^{-jM\omega} e^{j\beta} \breve{H}(\omega)$$

where the amplitude response $\breve{H}(\omega)$ is a real function of ω

- Weighted error function is given by

$$\mathcal{E}(\omega) = W(\omega)[\breve{H}(\omega) - D(\omega)]$$

where $D(\omega)$ is the desired amplitude response and $W(\omega)$ is a positive weighting function

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 10 44 Copyright © 2001, S. K. Mitra

Design of Equiripple Linear-Phase FIR Filters

- **Parks-McClellan Algorithm**

  Based on iteratively adjusting the coefficients of $\breve{H}(\omega)$ until the peak absolute value of $\mathcal{E}(\omega)$ is minimized

- If peak absolute value of $\mathcal{E}(\omega)$ in a band $\omega_a \le w \le \omega_b$ is ε_0, then the absolute error satisfies

$$|\breve{H}(\omega) - D(\omega)| \le \frac{\varepsilon_o}{|W(\omega)|}, \quad \omega_a \le \omega \le \omega_b$$

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 10 45 Copyright © 2001, S. K. Mitra

Design of Equiripple Linear-Phase FIR Filters

- For filter design,

$$D(\omega) = \begin{cases} 1, & \text{in the passband} \\ 0, & \text{in the stopband} \end{cases}$$

- $\breve{H}(\omega)$ is required to satisfy the above desired response with a ripple of $\pm\delta_p$ in the passband and a ripple of δ_s in the stopband

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 10 46 Copyright © 2001, S. K. Mitra

Design of Equiripple Linear-Phase FIR Filters

- Thus, weighting function can be chosen either as

$$W(\omega) = \begin{cases} 1, & \text{in the passband} \\ \delta_p / \delta_s, & \text{in the stopband} \end{cases}$$

or

$$W(\omega) = \begin{cases} \delta_s / \delta_p, & \text{in the passband} \\ 1, & \text{in the stopband} \end{cases}$$

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 10 47 Copyright © 2001, S. K. Mitra