Chapter 8. FIR Filter Design

Gao Xinbo
School of E.E., Xidian Univ.
xbgao@ieee.org
http://see.xidian.edu.cn/teach/matlabdsp/

Introduction

- In digital signal processing, there are two important types of systems:
  - Digital filters: perform signal filtering in the time domain
  - Spectrum analyzers: provide signal representation in the frequency domain

- In this and next chapter we will study several basic design algorithms for both FIR and IIR filters.
  - These designs are mostly of the frequency selective type
    - Multiband lowpass, highpass, bandpass and bandstop filters

Preliminaries

- The design of a digital filter is carried out in three steps:
  - Specifications: they are determined by the applications
  - Approximations: once the specifications are defined, we use various concepts and mathematics that we studied so far to come up with a filter description that approximates the given set of specifications. (in detail)
  - Implementation: The product of the above step is a filter description in the form of either a difference equation, or a system function $H(z)$, or an impulse response $h(n)$. From this description we implement the filter in hardware or through software on a computer.

- In FIR filter design we will also consider systems like differentiators or Hilbert transformers.
  - It is not frequency selective filters
  - Nevertheless follow the design techniques being considered.

- We first begin with some preliminary issues related to design philosophy and design specifications. These issues are applicable to both FIR and IIR filter designs.

- We will study FIR filter design algorithms in the rest of this chapter.
Specifications

- Specifications are required in the frequency-domain in terms of the desired magnitude and phase response of the filter. Generally a linear phase response in the passband is desirable.
  - In the case of FIR filters, it is possible to have exact linear phase.
  - In the case of IIR filters, a linear phase in the passband is not achievable.
  - Hence we will consider magnitude-only specifications.

Magnitude Specifications

- **Absolute** specifications
  - Provide a set of requirements on the magnitude response function $|H(e^{j\omega})|$.
  - Generally used for FIR filters.
- **Relative** specifications
  - Provide requirements in decibels (dB), given by $\text{dB scale} = -20 \log_{10} \left( \frac{|H(e^{j\omega})|}{|H(e^{j\omega})|_{\text{max}}} \right) \geq 0$
  - Used for both FIR and IIR filters.

Absolute specifications of a lowpass filter

Absolute Specifications

- Band $[0, w_p]$ is called the passband, and $\delta_1$ is the tolerance (or ripple) that we are willing to accept in the ideal passband response.
- Band $[w_p, w_s]$ is called the stopband, and $\delta_2$ is the corresponding tolerance (or ripple).
- Band $[w_s, \pi]$ is called the transition band, and there are no restriction on the magnitude response in this band.
Relative (DB) Specifications

\[ R_p = -20\log_0 \frac{1-\delta_1}{1+\delta_1} > 0 \quad \text{for passband} \]
\[ A_s = -20\log_0 \frac{\delta_2}{1+\delta_1} > 0 \quad \text{for stopband} \]

Problem statement

- Design a lowpass filter (i.e., obtain its system function \( H(z) \) or its difference equation) that has a passband \([0, w_p]\) with tolerance \( \delta_1 \) (or \( R_p \) in dB) and a stopband \([w_s, \pi]\) with tolerance \( \delta_2 \) (or \( A_s \) in dB)

Why we concentrate on a lowpass filter?

- The above specifications were given for a lowpass filter.
- Similar specifications can also be given for other types of frequency-selective filters such as highpass or bandpass.
- However, the most important design parameters are frequency-band tolerance (or ripples) and band-edge frequencies.
- Whether the given band is a passband or stopband is a relatively minor issue.
Design and implementational advantages of the FIR digital filter

- The phase response can be exactly linear;
- They are relatively easy to design since there are no stability problems.
- They are efficient to implement;
- The DFT can be used in their implementation.

Advantages of a linear-phase response

- Design problem contains only real arithmetic and not complex arithmetic;
- Linear-phase filter provide no delay distortion and only a fixed amount of delay;
- For the filter of length M (or order M-1) the number of operations are of the order of M/2 as we discussed in the linear phase implementation.

Properties of Linear-phase FIR Filter

- Let h(n), n=0,1,…,M-1 be the impulse response of length (or duration) M. Then the system function is

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

Which has (M-1) poles at the origin (trivial poles) and M-1 zeros located anywhere in the z-plane. The frequency response function is

\[
H(e^{j\omega}) = \sum_{n=0}^{M-1} h(n)e^{-j\omega n}, \quad -\pi < \omega \leq \pi
\]

Impulse Response h(n)

- We impose a linear-phase constraint

\[
\angle H(e^{\alpha}) = -\alpha \omega, \quad -\pi < \omega \leq \pi
\]

Where \(\alpha\) is a constant phase delay. Then we know from Ch6 that \(h(n)\) must be symmetric, that is

\[
h(n) = h(M-1-n), \quad 0 \leq n \leq M-1, \quad \text{with} \quad \alpha = \frac{M-1}{2}
\]

Hence \(h(n)\) is symmetric about \(\alpha\), which is the index of symmetry. There are two possible types of symmetry:
A second type of linear-phase FIR filter

- The phase response satisfy the condition
  \[ \angle H(e^{j\omega}) = \beta - \alpha \omega, \quad -\pi < \omega \leq \pi \]
  Which is a straight line but not through the origin. In this case \( \alpha \) is not a constant phase delay, but
  \[ \frac{d\angle H(e^{j\omega})}{d\omega} = -\alpha \]
  Is constant, which is the group delay. Therefore \( \alpha \) is a constant group delay. In this case, as a group, frequencies are delayed at a constant rate.

---

A second type of linear-phase FIR filter

- For this type of linear phase one can show that
  \[ h(n) = -h(M - 1 - n), \quad 0 \leq n \leq M - 1, \quad \text{with } \alpha = \frac{M - 1}{2}, \quad \beta = \pm \frac{\pi}{2} \]
  This means that the impulse response \( h(n) \) is antisymmetric. The index of symmetry is still \( \alpha = (M - 1)/2 \).
  Once again we have two possible types, one for \( M \) odd and one for \( M \) even.
  Figure P230

---

Frequency Response \( H(e^{j\omega}) \)

- When the case of symmetry and anti-symmetry are combined with odd and even \( M \), we obtain four types of linear phase FIR filters. Frequency response functions for each of these types have some peculiar expressions and shapes. To study these response, we write \( H(e^{j\omega}) \) as
  \[ H(e^{j\omega}) = H_1(e^{j(\beta - \alpha \omega)}) e^{j\beta \omega}, \quad \beta = \pm \frac{\pi}{2}, \quad \alpha = \frac{M - 1}{2} \]
  \( H_1(e^{j\omega}) \) is an amplitude response function and not a magnitude response function.
  The phase response associated with the magnitude response is a discontinuous function, while that associated with the amplitude response is a continuous linear function.
Type-1 Linear-phase FIR filter: Symmetrical impulse response, M odd
In this case, $\beta=0$, $\alpha=(M-1)/2$ is an integer, and $h(n)=h(M-1-n)$, $0\leq n\leq M-1$. Then

$$H(e^{j\omega}) = \left[ \sum_{n=0}^{M-1/2} a(n) \cos \omega n \right] e^{-j(M-1)/2}$$

- $a(0)=H(M-1)/2$: the middle sample
- $a(n)=2H(M-1)/2$, $1\leq n \leq M-3/2$
- $H_f(w) = \sum_{n=0}^{(M-1)/2} a(n) \cos \omega n$

Note that $H_r(\pi)=0$, hence we cannot use this type for designing a lowpass filter or a highpass filter.

Type-2 Linear-phase FIR filter: Symmetrical impulse response, M even
In this case, $\beta=0$, $h(n)=h(M-1-n)$, $0\leq n\leq M-1$, but $\alpha=(M-1)/2$ is not an integer, and then

$$H(e^{j\omega}) = \left[ \sum_{n=0}^{M/2} b(n) \cos \left( n - \frac{1}{2} \right) \right] e^{-j(M-1)/2}$$

- $b(n)=2H(M-1)/2 - n$ for $n=1,2,\ldots,M/2$
- $H_f(w) = \sum_{n=0}^{M/2} b(n) \cos \left( w - \frac{1}{2} \right)$

Note that $H_r(\pi)=0$, hence we cannot use this type for designing a lowpass or bandstop filter.

Type-3 Linear-phase FIR filter: Antisymmetric impulse response, M odd
In this case, $\beta=\pi/2$, $\alpha=(M-1)/2$ is an integer, and $h(n)=-h(M-1-n)$, $0\leq n\leq M-1$. Then

$$H(e^{j\omega}) = \left[ \sum_{n=0}^{(M-1)/2} c(n) \sin \omega n \right] e^{j\omega (M-1)/2}$$

- $c(n)=2H(M-1)/2 - n$ for $n=1,2,\ldots,M-1/2$
- $H_f(w) = \sum_{n=0}^{(M-1)/2} c(n) \sin \omega n$

$H_r(\omega)=H_r(\pi)=0$, hence this type of filter is not suitable for designing a lowpass or highpass filter.

Type-4 Linear-phase FIR filter: Antisymmetric impulse response, M even
This case is similar to Type-2. We have

$$H(e^{j\omega}) = \left[ \sum_{n=0}^{M/2} d(n) \sin \left( n - \frac{1}{2} \right) \right] e^{j\omega (M-1)/2}$$

- $d(n)=2H(M-1)/2 - n$ for $n=1,2,\ldots,M/2$
- $H_f(w) = \sum_{n=0}^{M/2} d(n) \sin \left( w - \frac{1}{2} \right)$

$H_r(\omega)=0$ and $\exp(j\pi/2)=j$. Hence this type of filter is also suitable for designing digital Hilbert transforms and differentiators.
Matlab Implementation

- Hr-type 1
- Hr-type 2
- Hr-type 3
- Hr-type 4

Zeros quadruplet for linear-phase filters

- For real sequence, zeros are in **conjugates**;
- For symmetry sequence, zeros are in **mirror**;
  - Substitute \( q = z^{-1} \), the polynomial coefficients for \( q \) are in reverse order of polynomial of \( z \).
  - Since coefficients \( h(n) \) are symmetry, reverse in order do not change the coefficients vector.
  - If \( z_i \) is a root of the polynomial, thus \( p_i = z_i^{-1} \) is also a root.

Mirror zeros for symmetry coef.

- If \( z_k \) satisfy polynomial:
  \[
  h_0 + h_1 z_k^{-1} + h_2 z_k^{-2} + \ldots + h_{M-2} z_k^{-M+2} + h_{M-1} z_k^{-M+1} = 0
  \]
  where \( h_{M-1} = h_0, h_{M-2} = h_1, \ldots \)
  Then \( r_k = z_k^{-1} \) satisfy the same equation
  \[
  h_0 + h_1 r_k + h_2 r_k^2 + \ldots + h_{M-2} r_k^{M-2} + h_{M-1} r_k^{M-1} = 0
  \]
  \[
  = h_0 z_k^{-M+1} + h_1 z_k^{-M+2} + \ldots + h_{M-2} z_k^{-2} + h_{M-1} z_k^{-1} + h_0 =
  = z_k^{-M-1}(h_0 + h_1 z_k^{-1} + \ldots + h_{M-2} z_k^{-M+2} + h_{M-1} z_k^{-M+1}) = 0
  \]
Windows Design Techniques

Basic idea: choose a proper ideal frequency-selective filter (which always has a noncausal, infinite-duration impulse response) and then truncate (or window) its impulse response to obtain a linear-phase and causal FIR filter.

- Appropriate windowing function
- Appropriate ideal filter
- An ideal LPF of bandwidth \( w_c \) is given by

\[
H_f(e^{j\omega}) = \begin{cases} 
1 - e^{-j\omega \alpha}, & |\omega| \leq w_c \\
0, & w_c < |\omega| \leq \pi 
\end{cases}
\]

Where \( w_c \) is also called the cutoff frequency, \( \alpha \) is called the sample delay.

Windows Design Techniques

\[ h_c(n) = F^{-1}[H_f(e^{j\omega})] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_f(e^{j\omega}) e^{j\omega n} d\omega = \frac{\sin(w_c(\pi-\alpha))}{\pi(\pi-\alpha)} 
\]

Note that \( h_c(n) \) is symmetric with respect to \( \alpha \), a fact useful for linear-phase FIR filter.

To obtain a causal and linear-phase FIR filter \( h(n) \) of length \( M \), we must have

\[ h(n) = \begin{cases} 
h_c(n), & 0 \leq n \leq M-1 \\
0, & \text{elsewhere} 
\end{cases}, \quad \text{and} \quad \alpha = \frac{M-1}{2}
\]

This operation is called “windowing”.

Examples

- Examples 7.4
- Examples 7.5
- Examples 7.6
- Examples 7.7

Impulse Response

\[ a(n) \] coefficients

Type-I Amplitude Response

Windows Design Techniques

Basic idea: choose a proper ideal frequency-selective filter (which always has a noncausal, infinite-duration impulse response) and then truncate (or window) its impulse response to obtain a linear-phase and causal FIR filter.

- Appropriate windowing function
- Appropriate ideal filter
- An ideal LPF of bandwidth \( w_c \pi \) is given by

\[
H_f(e^{j\omega}) = \begin{cases} 
1 - e^{-j\omega \alpha}, & |\omega| \leq w_c \\
0, & w_c < |\omega| \leq \pi 
\end{cases}
\]

Where \( w_c \) is also called the cutoff frequency, \( \alpha \) is called the sample delay.
Windowing

\[ h(n) = h_r(n)w(n) \]

\[ w(n) = \begin{cases} 
\text{some symmetric function with respect to } \alpha & \text{over } 0 \leq n \leq M - 1 \\
0, \text{ otherwise} 
\end{cases} \]

Depending on how we define \( w(n) \) above, we obtain different window design. For example,

\[ W(n) = R_d(n), \text{ rectangular window} \]

\[ H(e^{j\alpha}) = H_d(e^{j\alpha}) \otimes W(e^{j\alpha}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(e^{j\alpha})H_d(e^{j(\alpha-\beta)})d\beta \]

This is shown pictorially in Fig.7.8.

Basic Window Design Idea

- For the given filter specifications choose the filter length \( M \) and a window function \( w(n) \) for the narrowest main lobe width and the smallest side lobe attenuation possible.

- From the observation 4 above we note that the passband tolerance \( \delta_1 \) and the stopband tolerance \( \delta_2 \) can not be specified independently. We generally take care of \( \delta_2 \) alone, which results in \( \delta_2 = \delta_1 \).

Rectangular Window

This is the simplest window function but provides the worst performance from the viewpoint of stopband attenuation.

\[ w(n) = \begin{cases} 
1, & 0 \leq n \leq M - 1 \\
0, & \text{otherwise} 
\end{cases} \]

\[ W(e^{j\alpha}) = \frac{\sin(M\alpha)}{M\sin(\alpha)} \Rightarrow W_d(n) = \frac{\sin(M\alpha)}{\sin(\alpha)} \]

\[ H_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_d(\alpha)d\alpha = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_d(\alpha)d\alpha, \quad M >> 1 \]

Figure 7.9

Observations:

- Since the window \( w(n) \) has a finite length equal to \( M \), its response has a peaky main lobe whose width is proportional to \( 1/M \), and its side lobes of smaller heights.

- The periodic convolution produces a smeared version of the ideal response \( H_d(e^{j\alpha}) \).

- The main lobe produces a transition band in \( H(e^{j\alpha}) \) whose width is responsible for the transition width. This width is then proportional to \( 1/M \). the wider the main lobe, the wider will be the transition width.

- The side lobes produce ripples that have similar shapes in both the passband and stopband.
Observations:

- The amplitude response $W_r(w)$ has the first zero at $w_1 = \frac{2\pi}{M}$. Hence the width of the main lobe is $2w_1 = \frac{4\pi}{M}$. Therefore the approximate transition bandwidth is $\frac{4\pi}{M}$.

- The magnitude of the first side lobe (the peak side lobe magnitude) is approximately at $w = \frac{3\pi}{M}$ and is given by $|W_r(3\pi/M)| = \frac{2M}{3\pi}$, for $M \gg 1$. Comparing this with the main lobe amplitude, which is equal to $M$, the peak side lobe magnitude is $\frac{2}{3\pi} = 21.11\% = 13$ dB of the main lobe amplitude.

Two main problems

- The minimum stopband attenuation of 21 dB is insufficient in practical applications.

- The rectangular windowing being a direct truncation of the infinite length $h_d(n)$, it suffers from the Gibbs phenomenon.

Observations

- The accumulated amplitude response has the first side lobe magnitude at 21 dB. This results in the minimum stopband attenuation of 21 dB irrespective of the window length $M$.

- Using the minimum stopband attenuation, the transition bandwidth can be accurately computed. This computed exact transition bandwidth is $w_s - w_p = \frac{1.8\pi}{M}$, which is about half the approximate bandwidth of $\frac{4\pi}{M}$.

Bartlett Window

- Bartlett suggested a more gradual transition in the form of a triangular window

$$ v(n) = \begin{cases} \frac{2n}{M-1} & 0 \leq n \leq \frac{M-1}{2} \\ \frac{2n}{M-1} - 1 & \frac{M-1}{2} < n \leq M-1 \\ 0 & \text{otherwise} \end{cases} $$

Figure 7.11
Hanning Window

This is a raised cosine window function.

\[
w(n) = \begin{cases} 
0.5 \left[ 1 - \cos \left( \frac{2\pi n}{M-1} \right) \right] & 0 \leq n \leq M-1 \\
0, & \text{otherwise} 
\end{cases}
\]

Hamming Window:

\[
w(n) = \begin{cases} 
0.54 - 0.46 \cos \left( \frac{2\pi n}{M-1} \right) & 0 \leq n \leq M-1 \\
0, & \text{otherwise} 
\end{cases}
\]

Blackman Window

\[
w(n) = \begin{cases} 
0.42 - 0.5 \cos \left( \frac{2\pi n}{M-1} \right) + 0.08 \cos \left( \frac{4\pi n}{M-1} \right) & 0 \leq n \leq M-1 \\
0, & \text{otherwise} 
\end{cases}
\]

Kaiser Window:

\[
w(n) = \frac{I_1(\beta \sqrt{1 - \left( 1 - \left( \frac{2\pi n}{M-1} \right) \right)^2})}{I_1(\beta)}, 
\] 

\]

\[I_1(\beta) \text{ is the modified zero-order Bessel function}
\]

Kaiser Window

If \( \beta = 5.658 \), then the transition width is equal to 7.8\( \pi \)/M, and the minimum stopband attenuation is equal to 60dB.

If \( \beta = 4.538 \), then the transition width is equal to 5.8\( \pi \)/M, and the minimum stopband attenuation is equal to 50dB.

KAISER HAS DEVELOPED EMPIRICAL DESIGN EQUATIONS.

Given \( w, n, R, \) and \( A \),

\[\text{Store transition width} = \Delta f = \frac{R - r}{2\pi}\]

\[\text{filter order} \ M = \frac{A - 7.95}{1.5855} + 1\]

\[\text{parameter} \ \beta = \begin{cases} 
0.1102(4 - A)^{1.8}, & A \geq 50 \\
0.0842(A - 21)^{1.3} + 0.0796(A - 21), & 21 < A < 50 
\end{cases}\]

取Kaiser窗时设定\( \beta \), 再用kaiserord函数求得M
Matlab Implementation

- \( W = \text{boxcar}(M) \): rectangular window
- \( W = \text{triang}(M) \): bartlett window
- \( W = \text{hanning}(M) \)
- \( W = \text{hamming}(M) \)
- \( W = \text{blackman}(M) \)
- \( W = \text{kaiser}(M, \beta) \)

Examples

\[
H(z) = \sum_{n=0}^{M-1} h(n)z^{-n} = \frac{1-z^{-M}}{1-z^{-1}} \sum_{k=1}^{M-1} H(k)
\]

\[
H(e^{j\omega}) = \frac{1-e^{-j\omega M}}{M} \sum_{k=0}^{M-1} e^{-j\omega k} H(k)
\]

\[
H(k) = H(e^{j2\pi k/M}) = \begin{cases} 
H(0), & k = 0 \\
H^*(M-k), & k = 1, 2, \ldots, M-1 
\end{cases}
\]

\[
H(k) = H(2\pi k/M) = \begin{cases} 
H(0), & k = 0 \\
|H| \frac{2\pi k}{\pi}, & k = 1, 2, \ldots, M-1 
\end{cases}
\]

Frequency Sampling Design Techniques

- In this design approach we use the fact that the system function \( H(z) \) can be obtained from the samples \( H(k) \) of the frequency response \( H(e^{j\omega}) \).
- This design technique fits nicely with the frequency sampling structure that we discussed in Ch6.

Phase for Type 1 & 2

\[
\angle H(k) = \begin{cases} 
\left( \frac{M-1}{2}, \frac{2\pi}{M} \right) & k = 0, \ldots, \frac{M-1}{2} \\
\left( \frac{M-1}{2}, \frac{2\pi(M-k)}{M} \right) & k = \frac{M-1}{2} + 1, \ldots, M-1 
\end{cases}
\]

Phase for Type 3 & 4

\[
\angle H(k) = \begin{cases} 
\left( \pm \frac{\pi}{2}, \frac{M-1}{2} \right) \left( \frac{2\pi}{M} \right) & k = 0, \ldots, \frac{M-1}{2} \\
\left( \pm \frac{\pi}{2}, -\frac{2\pi(M-k)}{M} \right) & k = \frac{M-1}{2} + 1, \ldots, M-1 
\end{cases}
\]
Frequency Sampling Design Techniques

- **Basic Idea:**
  - Given the ideal lowpass filter \( H_d(e^{j\omega}) \), choose the filter length \( M \) and then sample \( H_d(e^{j\omega}) \) at \( M \) equispaced frequencies between 0 and 2\( \pi \). The actual response is the interpolation of the samples is given by

\[
H(z) = \sum_{k=0}^{M-1} H(k) \frac{1-z^{-M}}{1-z^{-1} e^{-j2\pi{k}/M}}
\]

\[
h(n) = IDFT[H(k)]
\]

From Fig. 7.25, we observe the following

- The approximation error—that is difference between the ideal and the actual response—is zero at the sampled frequencies.
- The approximation error at all other frequencies depends on the shape of the ideal response; that is, the sharper the ideal response, the larger the approximation error.
- The error is larger near the band edge and smaller within the band.

Two Design Approaches

- **Naïve design method:** Use the basic idea literally and provide no constraints on the approximation error, that is, we accept whatever error we get from the design.

- **Optimum design method:** try to minimize error in the stop band by varying values of the transition band samples.

Naive design methods

- In this method we set \( H(k) = H_d(e^{j2\pi{k}/M}) \), \( k=0,\ldots, M-1 \) and use (7.35) through (7.39) to obtain the impulse response \( h(n) \).
- Example 7.14
- The naive design method is seldom used in practice.
The minimum stopband attenuation is about 16dB, which is clearly unacceptable. If we increase M, then there will be samples in the transition band, for which we do not precisely know the frequency response.

**Optimum design method**

- To obtain more attenuation, we will have to increase M and make transition band samples free samples—that is, we vary their values to obtain the largest attenuation for the given M and the transition width.
- This is an optimization problem and it is solved using linear programming techniques.
- Example 7.15, 7.16, 7.17, 7.18, 7.19, 7.20

**Comments**

- This method is superior in that by varying one sample we can get a much better design.
- In practice the transition bandwidth is generally small, containing either one or two samples. Hence we need to optimize at most two samples to obtain the largest minimum stopband attenuation.
- This is also equivalent to minimizing the maximum side lobe magnitudes in absolute sense. Hence this optimization problem is also called a minimax problem.
- The detailed algorithm is ignored.

**Optimal Equiripple Design Technique**

- **Disadvantages** of the window design and the frequency sampling design
  - We cannot specify the band frequencies \( w_p \) and \( w_s \) precisely in the design;
  - We cannot specify both \( \delta_1 \) and \( \delta_2 \) ripple factors simultaneously
  - The approximation error—that is, the difference between the ideal response and the actual response—is not uniformly distributed over the band intervals.
Techniques to eliminate the above three problems

For linear-phase FIR filter it is possible to derive a set of conditions for which it can be proved that the design solution is optimal in the sense of minimizing the maximum approximation error (sometimes called the minimax or the Chebyshev error).

Filters that have this property are called equiripple filter because the approximation error is uniformly distributed in both the passband and the stopband.

This results in low-order filter.

Development of the Minimax Problem

The frequency response of the four cases of linear-phase FIR filters can be written in the form

\[ H(e^{jw}) = e^{j\alpha} e^{-j\beta w} H_r(w) \]

Where the values for \( \beta \) and the expressions for \( H_r(w) \) are given in Table 7.2 (P.278)

Using simple trigonometric identities, each expression for \( H_r(w) \) above can be written as a product of a fixed function of \( w \) \( (Q(w)) \) and a function that is a sum of cosines \( (P(w)) \).

\[ \sum_{n=0}^{L} a_n \cos nw \]

Table 7.3 P.279

Chebyshev approximation problem

The purpose of this analysis is to have a common form for \( H_r(w) \) across all four cases. It makes the problem formulation much easier.

To formulate our problem as a Chebyshev approximation problem, we have to define the desired amplitude response \( H_r(w) \) and a weighting function \( W(w) \), both defined over passbands and stopbands.

Chebyshev approximation problem

The weighting function is necessary so that we can have an independent control over \( \delta_1 \) and \( \delta_2 \). The weighted error is defined as

\[ E(w) = W(w)[H_r(w) - H_{\delta}(w)] \]

\[ E(w) = W(w)[H_{\delta}(w) - Q(w)P(w)] \]

\[ = W(w)Q(w)H_{\delta}(w) - P(w)] \]

The common form of \( E(w) \)

\[ E(w) = \tilde{W}(w)[\tilde{H}_{\delta}(w) - P(w)] \]
Problem Statement

- Determine the set of coefficients $a(n)$ or $b(-n)$ or $c(-n)$ or $d(-n)$ [or equivalently $a(n)$ or $b(n)$ or $c(n)$ or $d(n)$] to minimize the maximum absolute value of $E(w)$ over the passband and stopband.

$$\min_{\text{over \ coeff}} \left[ \max_{\text{w}} |E(w)| \right]$$

Now we succeeded in specifying the exact $w_p, w_s, \Delta_1, \Delta_2$. In addition, the error can now be distributed uniformly in both the passband and stopband.

Constraint on the Number of Extrema

- How many local maxima and minima exist in the error function $E(w)$ for a given $M$-point filter?

- Conclusion: the error function $E(w)$ has at most $(L+3)$ extrema in $S$.

Alternation Theorem

- Let $S$ be any closed subset of the closed interval $[0, \pi]$. In order that $P(w)$ be the unique minimax approximation to $H^*(w)$ on $S$, it is necessary and sufficient that the error function $E(w)$ exhibit at least $(L+2)$ "alternations" or extremal frequencies in $S$; that is, there must exist $(L+2)$ frequencies $w_i$ in $S$ such that

$$E(w_i) = -E(w_{i+1}) = \pm \max_{S} |E(w)| = \pm \delta$$

$$\forall w_k < w_i < \cdots < w_{L+1} \in S$$

Parks-McClellan Algorithm

- It was solved by Remez.

1. Estimate the filter length order $M$ by (7.48)
2. Guess extrema frequencies $w_i \ (i=1:L+2)$
3. Find an $L$th polynomial that fits these points
4. Determine new $w_i$’s by interpolation of the polynomial
5. Iteration from beginning
6. Determine $a(t)$ and $E_{\text{max}}$ by $\min(\max(E(w)))$

These steps were integrated in function `remez`
Equiripple design function `remez`

- General: `[h]=remez(N,f,n,weights,ftype)`
- When `weight=1` everywhere, and `ftype` is not Hilbert filter or differentiator
- `[h]=remez(N,f,m)`
  - `h` is the filter coefficients with length `M=N+1`
  - `N` denotes the order of the filter
  - `f` -- an array denotes band edges in units of `\pi`.
  - `m` -- desired magnitude response at each `f`

Remez equiripple design examples

- Example 7.23: LP filter Design
  - compare with `window` design (ex7.8)
  - and `freq.sampling` design (ex7.14,7.15,7.16)
- Example 7.24: BP filter Design
  - compare with `window` (ex7.10) design
  - and `freq.sampling` (ex7.17) design
- Example 7.25: HP filter Design
- Example 7.26: “Staircase” filter Design

Other Examples of equiripple

- Ex7.27: digital differentiator design using `remez` function
- Ex7.28: digital Hilbert Transformer design using `remez` function

\[
R_p = -20 \log_{10} \frac{1 - \delta^2_1}{1 + \delta^2_1} \Rightarrow \delta_1 = \frac{10^{R_p/20} - 1}{10^{R_p/20} + 1}
\]

\[
A_s = -20 \log_{10} \frac{\delta^2_2}{1 + \delta^2_2} \Rightarrow \delta_2 = (1 + \delta_1) \cdot 10^{-A_s/20}
\]

\[
M = -20 \log_{10} \frac{\sqrt{\delta^2_1 \delta^2_2}}{14.6 \Delta f}, \quad \Delta f = \frac{W_s - W_p}{2\pi}
\]
Readings and exercises

- Readings: for Dec.3 to Dec.5 (2 classes)
  - Chinese text book: pp.195~216, 220~222

- Exercises:
  1. Complete the program of example 2, p7.2
  2. p7.3, p7.5, p7.7, 7.14, P7.19