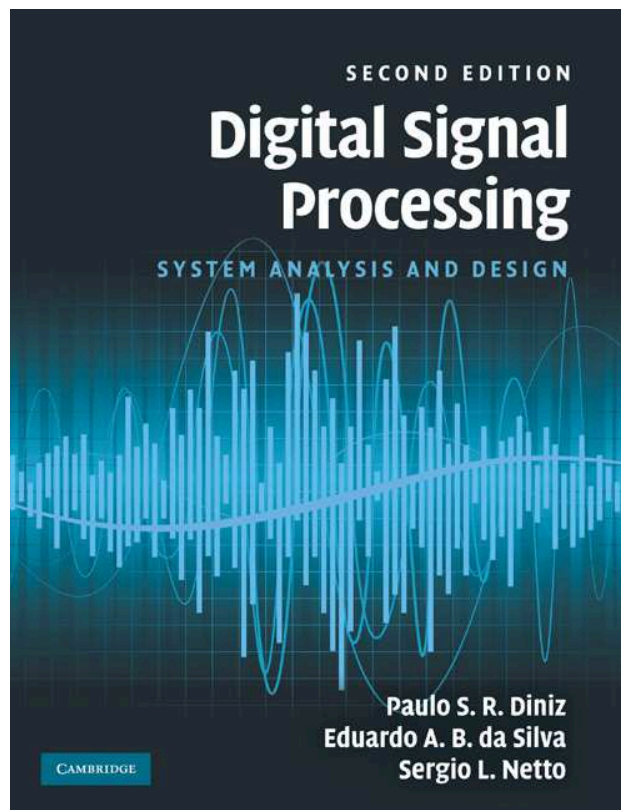


FIR Filter Approximations



Paulo S. R. Diniz

Eduardo A. B. da Silva

Sergio L. Netto

diniz,eduardo,sergioln@lps.ufrj.br

September 2010

Contents

- Ideal characteristics of standard filters
 - Lowpass, highpass, bandpass, and bandstop filters
 - Differentiators
 - Hilbert transformers
- FIR filter approximation by frequency sampling
- FIR filter approximation by window functions
 - Rectangular window
 - Triangular windows
 - Hamming and Hann windows
 - Blackman window
 - Kaiser window
 - Dolph-Chebyshev window

Contents (cont.)

- Maximally-flat FIR filter approximation
- FIR filter approximation by optimization
 - Weighted-least-squares method
 - Chebyshev method
 - WLS-Chebyshev method
- Do-it-yourself: FIR filter approximations

Ideal characteristics of standard filters

- As seen in Section 2.8, the relationship between $H(e^{j\omega})$ and $h(n)$ is given by the following pair of equations:

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n)e^{-j\omega n} \quad (2)$$

$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega \quad (3)$$

- In what follows, we determine $H(e^{j\omega})$ and $h(n)$ related to ideal standard filters.

Lowpass, highpass, bandpass, and bandstop filters

- The ideal magnitude responses of some standard digital filters are depicted below.

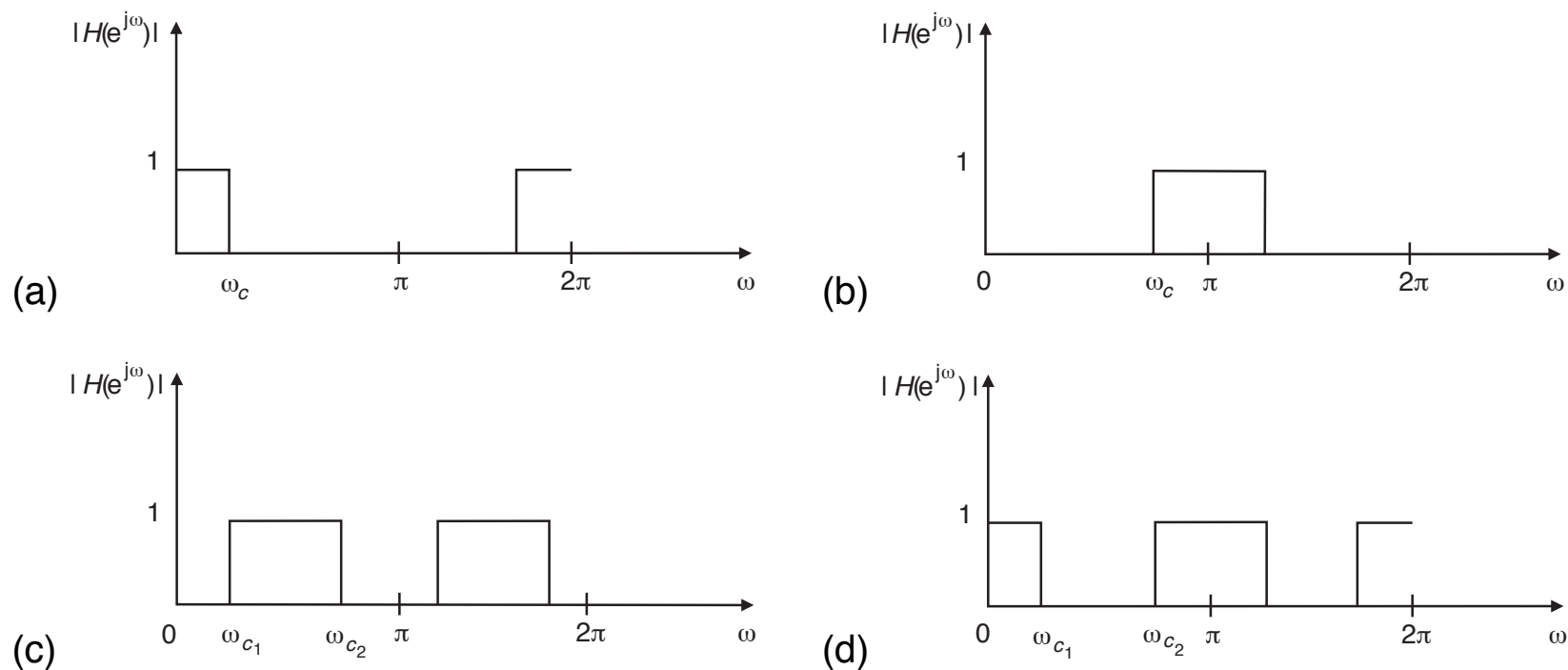


Figure 1: Ideal magnitude responses: (a) lowpass; (b) highpass; (c) bandpass; (d) bandstop filters.

Lowpass, highpass, bandpass, and bandstop filters

- For instance, the lowpass filter, as seen in Figure 1a, is described by

$$|H(e^{j\omega})| = \begin{cases} 1, & \text{for } |\omega| \leq \omega_c \\ 0, & \text{for } \omega_c < |\omega| \leq \pi \end{cases} \quad (4)$$

- Using (3), the impulse response for the ideal lowpass filter is

$$h(n) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{j\omega n} d\omega = \begin{cases} \frac{\omega_c}{\pi}, & \text{for } n = 0 \\ \frac{\sin(\omega_c n)}{\pi n}, & \text{for } n \neq 0 \end{cases} \quad (5)$$

- One should note that in the above inverse transform calculations we have supposed that the phase of the filter is zero.

Lowpass, highpass, bandpass, and bandstop filters

- From Section 4.2.3, we have that the phase of an FIR filter must be of the form $e^{-j\omega \frac{M}{2}}$, where M is an integer.
- Therefore, for M even, it suffices to shift the above impulse response by $\frac{M}{2}$ samples. However, for M odd, $\frac{M}{2}$ is not an integer, and the impulse response must be computed as

$$\begin{aligned}
 h(n) &= \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{-j\omega \frac{M}{2}} e^{j\omega n} d\omega \\
 &= \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{j\omega \left(n - \frac{M}{2}\right)} d\omega \\
 &= \frac{\sin \left[\omega_c \left(n - \frac{M}{2}\right) \right]}{\pi \left(n - \frac{M}{2}\right)} \tag{6}
 \end{aligned}$$

Lowpass, highpass, bandpass, and bandstop filters

- Likewise, for bandstop filters, the ideal magnitude response, depicted in Figure 1d, is given by

$$|H(e^{j\omega})| = \begin{cases} 1, & \text{for } 0 \leq |\omega| \leq \omega_{c_1} \\ 0, & \text{for } \omega_{c_1} < |\omega| < \omega_{c_2} \\ 1, & \text{for } \omega_{c_2} \leq |\omega| \leq \pi \end{cases} \quad (7)$$

- Then, using (3), the impulse response for such an ideal filter is

$$\begin{aligned} h(n) &= \frac{1}{2\pi} \left[\int_{-\omega_{c_1}}^{\omega_{c_1}} e^{j\omega n} d\omega + \int_{\omega_{c_2}}^{\pi} e^{j\omega n} d\omega + \int_{-\pi}^{-\omega_{c_2}} e^{j\omega n} d\omega \right] \\ &= \begin{cases} 1 + \frac{\omega_{c_1} - \omega_{c_2}}{\pi}, & \text{for } n = 0 \\ \frac{1}{\pi n} [\sin(\omega_{c_1} n) - \sin(\omega_{c_2} n)], & \text{for } n \neq 0 \end{cases} \quad (8) \end{aligned}$$

Differentiators

- An ideal discrete-time differentiator is a linear system that, when samples of a band-limited continuous signal are used as input, the output samples represent the derivative of the continuous signal.
- More precisely, given a continuous-time signal $x_a(t)$ band-limited to $[-\frac{\pi}{T}, \frac{\pi}{T})$, when its corresponding sampled version $x(n) = x_a(nT)$ is input to an ideal differentiator, it produces the output signal, $y(n)$, such that

$$y(n) = \left. \frac{dx_a(t)}{dt} \right|_{t=nT} \quad (9)$$

- If the Fourier transform of the continuous-time signal is denoted by $X_a(j\Omega)$, we have that the Fourier transform of its derivative is $j\Omega X_a(j\Omega)$.

Differentiators

- Therefore, an ideal discrete-time differentiator is characterized by a frequency response, up to a multiplicative constant, of the form

$$H(e^{j\omega}) = j\omega, \text{ for } -\pi \leq \omega < \pi \quad (10)$$

- The magnitude and phase responses of a differentiator are depicted in Figure 2.

Differentiators

- Using equation (3), the corresponding impulse response is given by

$$\begin{aligned}
 h(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} j\omega e^{j\omega n} d\omega \\
 &= \begin{cases} 0, & \text{for } n = 0 \\ \frac{1}{2\pi} \left[e^{j\omega n} \left(\frac{\omega}{n} - \frac{1}{jn^2} \right) \right] \Big|_{-\pi}^{\pi} = \frac{(-1)^n}{n}, & \text{for } n \neq 0 \end{cases}
 \end{aligned}
 \tag{11}$$

- One should note that if a differentiator is to be approximated by a linear-phase FIR filter, one should necessarily use either a Type-III or a Type-IV form.

Hilbert transformers

- These equations provide a relation between the Fourier transforms of the real and imaginary parts of a signal whose Fourier transform is null for $-\pi \leq \omega < 0$. It thus implies that the ideal Hilbert transformer has the following transfer function:

$$H(e^{j\omega}) = \begin{cases} -j, & \text{for } 0 \leq \omega < \pi \\ j, & \text{for } -\pi \leq \omega < 0 \end{cases} \quad (17)$$

- The magnitude and phase components of such a frequency response are depicted in Figure 3.

Hilbert transformers

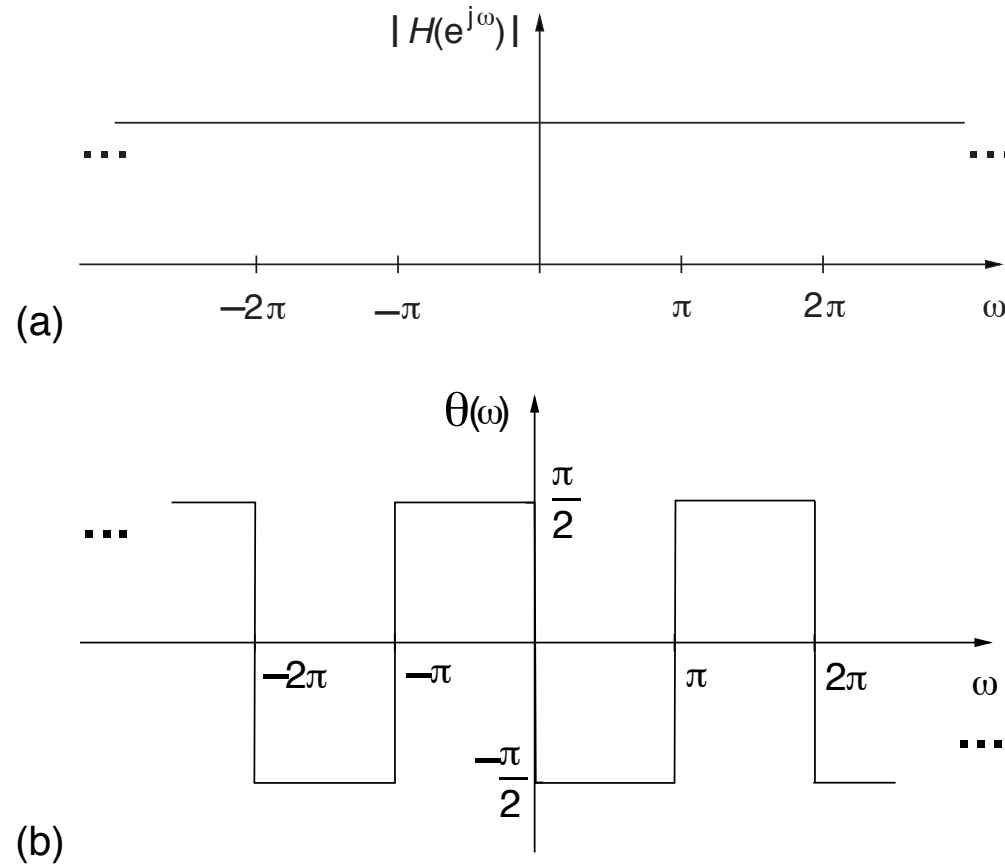


Figure 3: Characteristics of an ideal Hilbert transformer: (a) magnitude response; (b) phase response.

Hilbert transformers

- Using equation (3), the corresponding impulse response for the ideal Hilbert transformer is given by

$$h(n) = \frac{1}{2\pi} \left[\int_0^{\pi} -je^{j\omega n} d\omega + \int_{-\pi}^0 je^{j\omega n} d\omega \right] = \begin{cases} 0, & \text{for } n = 0 \\ \frac{1}{\pi n} [1 - (-1)^n], & \text{for } n \neq 0 \end{cases} \quad (18)$$

- By examining equation (17) we conclude, as in the case of the differentiator, that a Hilbert transformer must be approximated, when using a linear-phase FIR filter, by either a Type-III or Type-IV structure.

Summary

Table 1: Ideal frequency characteristics and corresponding impulse responses for low-pass, highpass, bandpass, and bandstop filters, as well as for differentiators and Hilbert transformers.

Filter type	Magnitude response $ H(e^{j\omega}) $	Impulse response $h(n)$
Lowpass	$\begin{cases} 1, & \text{for } 0 \leq \omega \leq \omega_c \\ 0, & \text{for } \omega_c < \omega \leq \pi \end{cases}$	$\begin{cases} \frac{\omega_c}{\pi}, & \text{for } n = 0 \\ \frac{1}{\pi n} \sin(\omega_c n), & \text{for } n \neq 0 \end{cases}$
Highpass	$\begin{cases} 0, & \text{for } 0 \leq \omega < \omega_c \\ 1, & \text{for } \omega_c \leq \omega \leq \pi \end{cases}$	$\begin{cases} 1 - \frac{\omega_c}{\pi}, & \text{for } n = 0 \\ -\frac{1}{\pi n} \sin(\omega_c n), & \text{for } n \neq 0 \end{cases}$

Summary

Filter type	Magnitude response $ H(e^{j\omega}) $	Impulse response $h(n)$
Bandpass	$\begin{cases} 0, & \text{for } 0 \leq \omega < \omega_{c_1} \\ 1, & \text{for } \omega_{c_1} \leq \omega \leq \omega_{c_2} \\ 0, & \text{for } \omega_{c_2} < \omega \leq \pi \end{cases}$	$\begin{cases} \frac{(\omega_{c_2} - \omega_{c_1})}{\pi}, & \text{for } n = 0 \\ \frac{1}{\pi n} [\sin(\omega_{c_2} n) - \sin(\omega_{c_1} n)], & \text{for } n \neq 0 \end{cases}$
Bandstop	$\begin{cases} 1, & \text{for } 0 \leq \omega \leq \omega_{c_1} \\ 0, & \text{for } \omega_{c_1} < \omega < \omega_{c_2} \\ 1, & \text{for } \omega_{c_2} \leq \omega \leq \pi \end{cases}$	$\begin{cases} 1 - \frac{(\omega_{c_2} - \omega_{c_1})}{\pi}, & \text{for } n = 0 \\ \frac{1}{\pi n} [\sin(\omega_{c_1} n) - \sin(\omega_{c_2} n)], & \text{for } n \neq 0 \end{cases}$

Summary

Filter type	Frequency response $H(e^{j\omega})$	Impulse response $h(n)$
Differentiator	$j\omega, \quad \text{for } -\pi \leq \omega < \pi$	$\begin{cases} 0, & \text{for } n = 0 \\ \frac{(-1)^n}{n}, & \text{for } n \neq 0 \end{cases}$
Hilbert transformer	$\begin{cases} -j, & \text{for } 0 \leq \omega < \pi \\ j, & \text{for } -\pi \leq \omega < 0 \end{cases}$	$\begin{cases} 0, & \text{for } n = 0 \\ \frac{1}{\pi n} [1 - (-1)^n], & \text{for } n \neq 0 \end{cases}$

FIR filter approximation by frequency sampling

- In general, the problem of FIR filter design is to find a finite-length impulse response $h(n)$, whose Fourier transform $H(e^{j\omega})$ approximates a given frequency response well enough.
- As seen in Section 3.2, one way of achieving such a goal is by noting that the DFT of a length- N sequence $h(n)$ corresponds to samples of its Fourier transform at the frequencies $\omega = \frac{2\pi k}{N}$, that is

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n} \quad (19)$$

and then

$$H(e^{j\frac{2\pi k}{N}}) = \sum_{n=0}^{N-1} h(n)e^{-j\frac{2\pi k n}{N}}, \text{ for } k = 0, 1, \dots, (N - 1) \quad (20)$$

FIR filter approximation by frequency sampling

- It is then natural to consider designing a length- N FIR filter by finding an $h(n)$ whose DFT corresponds exactly to samples of the desired frequency response.
- In other words, $h(n)$ can be determined by sampling the desired frequency response at the N points $e^{j\frac{2\pi}{N}k}$ and finding its inverse DFT. This method is generally referred to as the frequency sampling approach.
- More precisely, if the desired frequency response is given by $D(\omega)$, one must first find

$$A(k)e^{j\theta(k)} = D\left(\frac{\omega_s k}{N}\right), \text{ for } k = 0, 1, \dots, (N - 1) \quad (21)$$

where $A(k)$ and $\theta(k)$ are samples of the desired amplitude and phase responses, respectively.

FIR filter approximation by frequency sampling

- Type I: Even order M and symmetrical impulse response. In this case, the phase and amplitude responses must satisfy

$$\theta(k) = -\frac{\pi k M}{M+1}, \text{ for } 0 \leq k \leq M \quad (22)$$

$$A(k) = A(M - k + 1), \text{ for } 1 \leq k \leq \frac{M}{2} \quad (23)$$

and then, the impulse response is given by

$$h(n) = \frac{1}{M+1} \left[A(0) + 2 \sum_{k=1}^{\frac{M}{2}} (-1)^k A(k) \cos \frac{\pi k (1 + 2n)}{M+1} \right] \quad (24)$$

for $n = 0, 1, \dots, M$.

Table 2: Impulse responses for linear-phase FIR filters with frequency sampling approach.

Filter type	Impulse response $h(n)$, for $n = 0, 1, \dots, M$	Condition
Type I	$\frac{1}{M+1} \left[A(0) + 2 \sum_{k=1}^{\frac{M}{2}} (-1)^k A(k) \cos \frac{\pi k(1+2n)}{M+1} \right]$	
Type II	$\frac{1}{M+1} \left[A(0) + 2 \sum_{k=1}^{\frac{M-1}{2}} (-1)^k A(k) \cos \frac{\pi k(1+2n)}{M+1} \right]$	$A\left(\frac{M+1}{2}\right) = 0$
Type III	$\frac{2}{M+1} \sum_{k=1}^{\frac{M}{2}} (-1)^{k+1} A(k) \sin \frac{\pi k(1+2n)}{M+1}$	$A(0) = 0$
Type IV	$\frac{1}{M+1} \left[(-1)^{\frac{M+1}{2}+n} A\left(\frac{M+1}{2}\right) + 2 \sum_{k=1}^{\frac{M-1}{2}} (-1)^k A(k) \sin \frac{\pi k(1+2n)}{M+1} \right]$	$A(0) = 0$

Example 5.1

- Design a lowpass filter satisfying the specification below using the frequency sampling method:

$$\left. \begin{aligned} M &= 52 \\ \Omega_p &= 4.0 \text{ rad/s} \\ \Omega_r &= 4.2 \text{ rad/s} \\ \Omega_s &= 10.0 \text{ rad/s} \end{aligned} \right\} \quad (37)$$

- Obs.: Note that in this text, in general, the variable Ω represents an analog frequency, and the variable ω a digital frequency.

Example 5.1 - Solution

- We divide the $[0, \Omega_s]$ interval into $(M + 1) = 53$ sub-intervals of same length $\frac{\Omega_s}{M+1}$, each starting at $\Omega_k = \frac{\Omega_s}{M+1}k$, for $k = 0, 1, \dots, M$.
- According to the prescribed specifications, Ω_p and Ω_r lie close to the extremes

$$k_p = \left\lfloor (M + 1) \times \frac{\Omega_p}{\Omega_s} \right\rfloor = \lfloor 53 \times \frac{4}{10} \rfloor = 21 \quad (38)$$

$$k_r = \left\lfloor (M + 1) \times \frac{\Omega_r}{\Omega_s} \right\rfloor = \lfloor 53 \times \frac{4.2}{10} \rfloor = 22 \quad (39)$$

- Thus, we assign

$$A(k) = \begin{cases} 1, & \text{for } 0 \leq k \leq k_p \\ 0, & \text{for } k_r \leq k \leq \frac{M}{2} \end{cases} \quad (40)$$

Example 5.1 - Solution

- Then, one can employ the following MATLAB script, implementing the first row in Table 2, to design a Type I lowpass filter using the frequency sampling method:

```

M = 52; N = M+1;
Omega_p = 4; Omega_r = 4.2; Omega_s = 10;
kp = floor(N*Omega_p/Omega_s);
kr = floor(N*Omega_r/Omega_s);
A = [ones(1,kp+1) zeros(1,M/2-kr+1)];
k = 1:M/2;
for n=0:M,
    h(n+1) = A(1) +
2*sum((-1).^k.*A(k+1).*cos(pi.*k*(1+2*n)/N));
end;
h = h./N;
  
```

- Using this script, one ends up with the set of coefficients shown in Table 3.

Example 5.1 - Solution

- The corresponding magnitude response is shown in Figure 4.

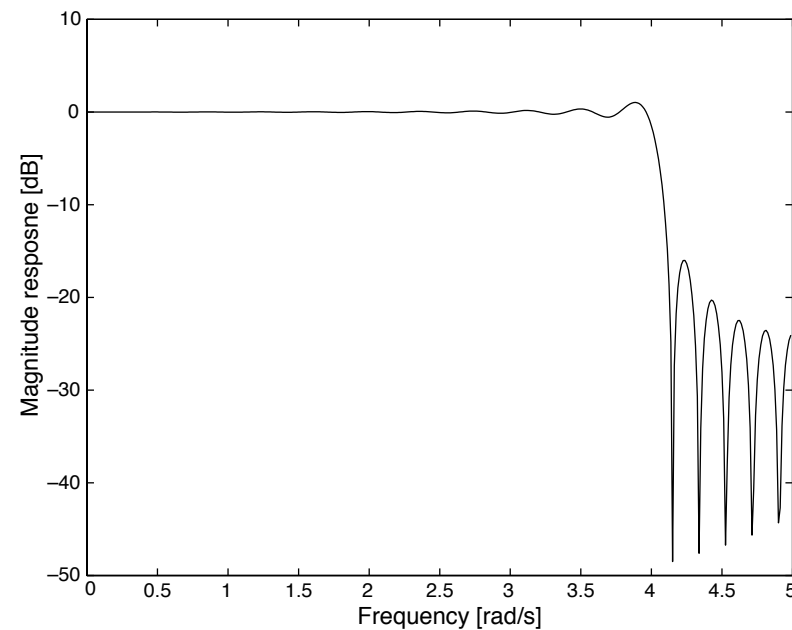


Figure 4: Magnitude response of the lowpass filter designed with the frequency sampling method.

FIR filter approximation by frequency sampling

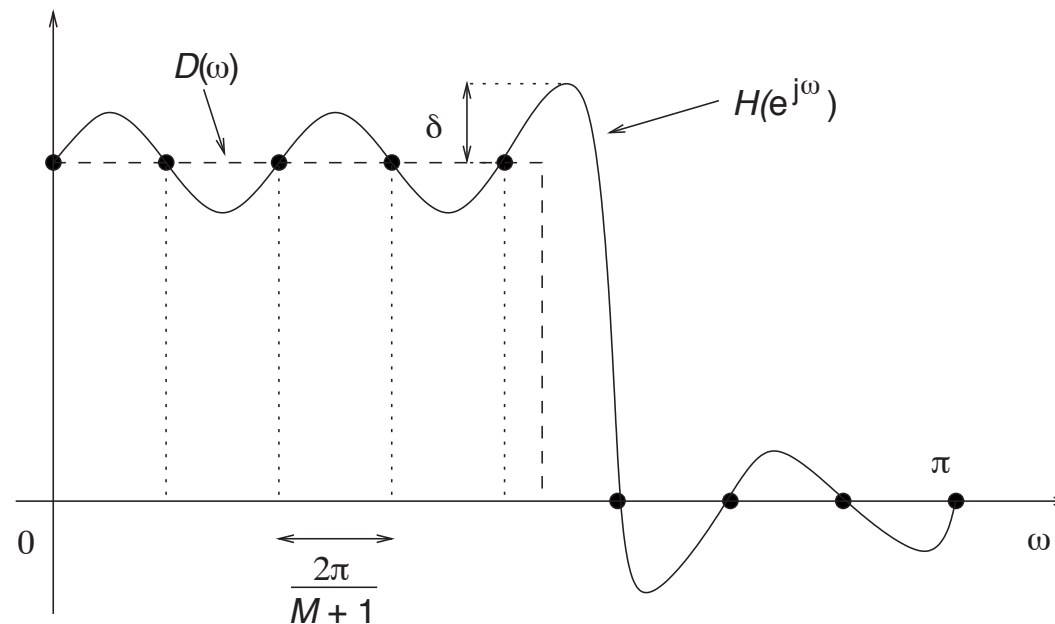


Figure 5: The desired magnitude response and the Fourier transform of $h(n)$ coincide only at the frequencies $\frac{2\pi k}{M+1}$, when using the frequency sampling approximation method.

FIR filter approximation with window functions

- For all ideal filters analyzed in Section 5.2, the impulse responses obtained from equation (3) have infinite duration, which leads to non-realizable FIR filters.
- A straightforward way to overcome this limitation is to define a finite-length auxiliary sequence $h'(n)$, yielding a filter of order M , as

$$h'(n) = \begin{cases} h(n), & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (48)$$

assuming that M is even.

FIR filter approximation with window functions

- The resulting transfer function is written as

$$H'(z) = h(0) + \sum_{n=1}^{\frac{M}{2}} (h(-n)z^n + h(n)z^{-n}) \quad (49)$$

- This is still a noncausal function which we can make causal by multiplying it by $z^{-\frac{M}{2}}$, without either distorting the filter magnitude response or destroying the linear-phase property.
- The example below highlights some of the impacts that the truncation of the impulse response in equations (48) and (49) has on the filter frequency response.

Example 5.2

- Design a bandstop filter satisfying the specification below:

$$\left. \begin{aligned} M &= 50 \\ \Omega_{c_1} &= \frac{\pi}{4} \text{ rad/s} \\ \Omega_{c_2} &= \frac{\pi}{2} \text{ rad/s} \\ \Omega_s &= 2\pi \text{ rad/s} \end{aligned} \right\} \quad (50)$$

Example 5.2 - Solution

- Applying equations (48) and (49) to the corresponding bandstop equations in Table 1, one may use the script:

```
M = 50;
```

```
wc1 = pi/4; wc2 = pi/2; ws = 2*pi;
```

```
n = 1:M/2;
```

```
h0 = 1 - (wc2 - wc1)/pi;
```

```
haux = (sin(wc1.*n) - sin(wc2.*n))./(pi.*n);
```

```
h = [fliplr(haux) h0 aux];
```

to obtain the filter coefficients listed in Table 4 (only half of them are listed as the others can be found using $h(n) = h(50 - n)$).

- The resulting magnitude response is depicted in Figure 6.

Example 5.2 - Solution

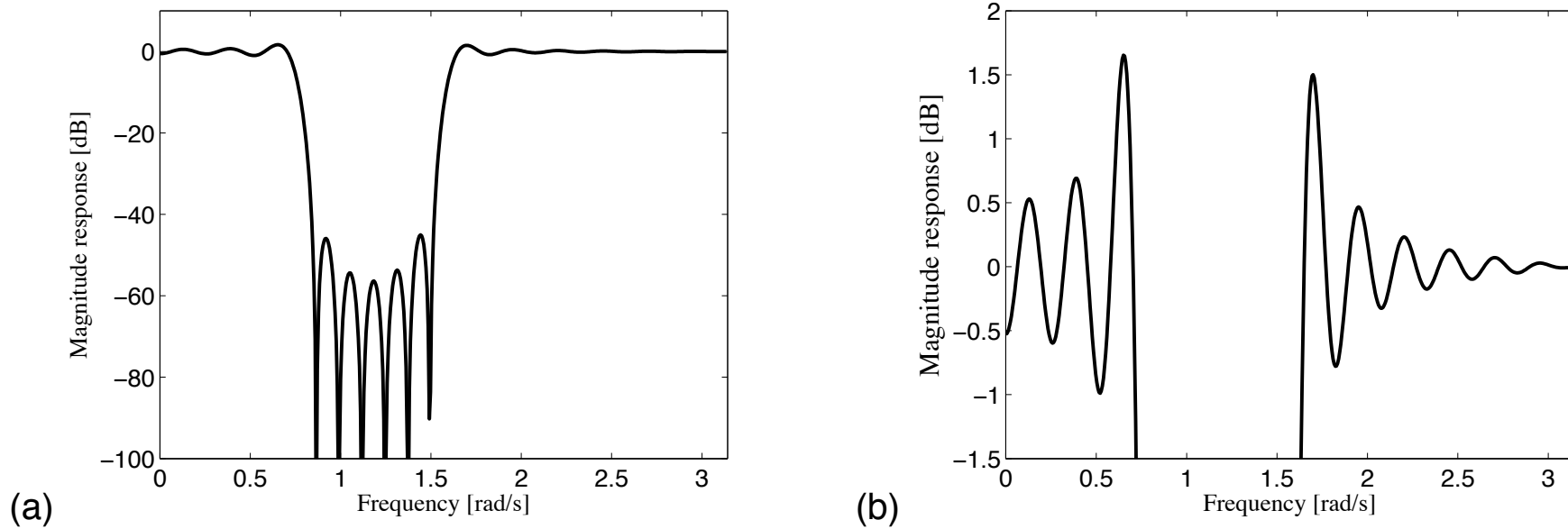


Figure 6: Bandstop filter: (a) magnitude response; (b) passband detail.

FIR filter approximation with window functions

- The ripple seen in Figure 6 close to the band edges is due to the slow convergence of the Fourier series $h(n)$ when approximating functions presenting discontinuities, such as the ideal responses seen in Figure 1.
- This implies that large amplitude ripples in the magnitude response appear close to the edges whenever an infinite-length $h(n)$ is truncated to generate a finite-length filter. These ripples are commonly referred to as Gibbs' oscillations.
- It can be shown that Gibbs' oscillations possess the property that their amplitudes do not decrease even when the filter order M is increased dramatically. This severely limits the practical usefulness of equations (48) and (49) in FIR design, because the maximum deviation from the ideal magnitude response can not be minimized by increasing the filter length.

FIR filter approximation with window functions

- Although we can not remove the ripples introduced by the poor convergence of the Fourier series, we can still attempt to control their amplitude by multiplying the impulse response $h(n)$ by a window function $w(n)$.
- The window $w(n)$ must be designed such that it introduces minimum deviation from the ideal frequency response. The coefficients of the resulting impulse response $h'(n)$ become

$$h'(n) = h(n)w(n) \quad (51)$$

- In the frequency domain, such a multiplication corresponds to a periodic convolution operation between the frequency responses of the ideal filter, $H(e^{j\omega})$, and of the window function, $W(e^{j\omega})$, that is

$$H'(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega'}) W(e^{j(\omega-\omega')}) d\omega' \quad (52)$$

FIR filter approximation with window functions

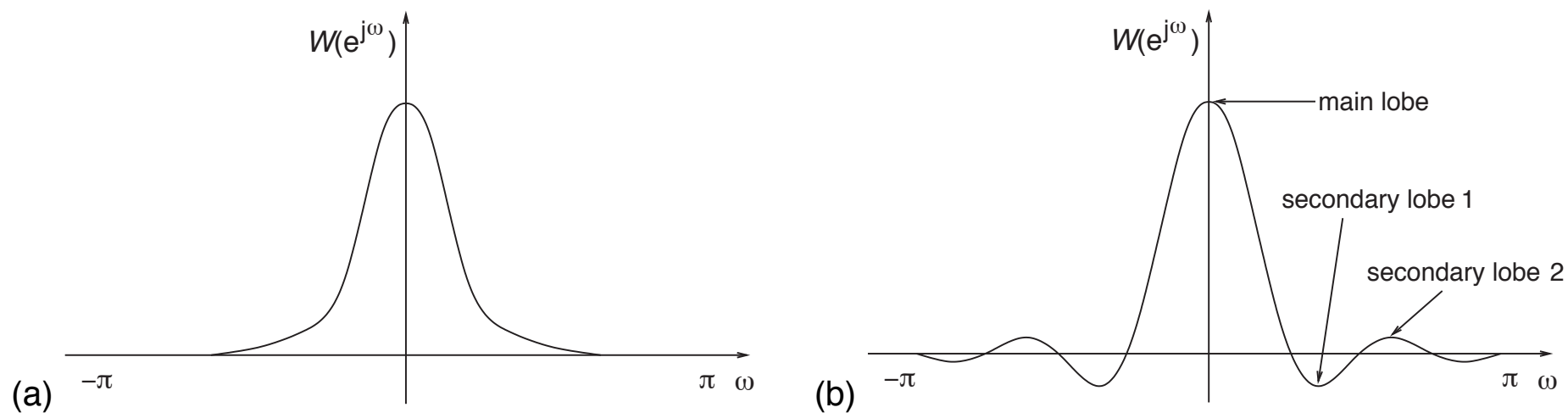


Figure 7: Magnitude responses of a window function: (a) ideal case; (b) practical case.

Rectangular window

- A simple truncation of the impulse response as described in equation (48) can be interpreted as the product between the ideal $h(n)$ and a window given by

$$w_r(n) = \begin{cases} 1, & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (53)$$

- Note that if we want to truncate the impulse responses in Table 1 using the above equation, and still keep the linear-phase property, the resulting truncated sequences would have to be either symmetric or antisymmetric around $n = 0$.
- This implies that, for those cases, M would have to be even (Type-I and Type-III filters, as seen in Subsection 4.2.3).
- For the case of M odd, the solution would be to shift $h(n)$ so that it is causal and apply a window different from zero from $n = 0$ to $n = M - 1$. This solution, however, is not commonly used in practice.

Rectangular window

- From equation (53), the frequency response of the rectangular window is given by

$$\begin{aligned}W_r(e^{j\omega}) &= \sum_{n=-\frac{M}{2}}^{\frac{M}{2}} e^{-j\omega n} \\&= \frac{e^{j\omega \frac{M}{2}} - e^{-j\omega \frac{M}{2}} e^{-j\omega}}{1 - e^{-j\omega}} \\&= e^{-j\frac{\omega}{2}} \frac{\left[e^{j\omega \left(\frac{M+1}{2}\right)} - e^{-j\omega \left(\frac{M+1}{2}\right)} \right]}{1 - e^{-j\omega}} \\&= \frac{\sin \left[\omega \left(\frac{M+1}{2}\right) \right]}{\sin \left(\frac{\omega}{2} \right)}\end{aligned}\tag{54}$$

Triangular windows

- The main problem associated with the rectangular window is the presence of ripples near the band edges of the resulting filter, which are caused by the existence of sidelobes in the frequency response of the window.
- Such a problem is due to the inherent discontinuity of the rectangular window in the time domain. One way to reduce such a discontinuity is to employ a triangular-shaped window, which will present only small discontinuities near its edges.
- The standard triangular window is defined as

$$w_t(n) = \begin{cases} -\frac{2|n|}{M+2} + 1, & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (55)$$

Triangular windows

- A small variation of such a window is called the Bartlett window and is defined by

$$w_{tB}(n) = \begin{cases} -\frac{2|n|}{M} + 1, & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (56)$$

- These two triangular-type window functions are closely related. Their main difference lies in the fact that the Bartlett window presents one null element at each of its extremities. In that manner, an M th-order Bartlett window can be obtained by juxtaposing one zero at each extremity of the $(M - 2)$ th-order standard triangular window.
- In some cases, an even greater reduction of the sidelobes is necessary, and then more complex window functions should be used, such as the ones described below.

Hamming and Hann windows

- The generalized Hamming window is defined as

$$w_H(n) = \begin{cases} \alpha + (1 - \alpha) \cos\left(\frac{2\pi n}{M}\right), & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (57)$$

with $0 \leq \alpha \leq 1$.

- This generalized window is referred to as the Hamming window when $\alpha = 0.54$, and for $\alpha = 0.5$, it is known as the Hann or Hanning window.
- The frequency response for the general Hamming window can be expressed based on the frequency response of the rectangular window.

Hamming and Hann windows

- We first write equation (57) as

$$w_H(n) = w_r(n) \left[\alpha + (1 - \alpha) \cos \left(\frac{2\pi n}{M} \right) \right] \quad (58)$$

- By transforming the above equation to the frequency domain, clearly the frequency response of the generalized Hamming window results from the periodic convolution between $W_r(e^{j\omega})$ and three impulse functions as

$$W_H(e^{j\omega}) = W_r(e^{j\omega}) * \left[\alpha \delta(\omega) + \left(\frac{1-\alpha}{2} \right) \delta \left(\omega - \frac{2\pi}{M} \right) + \left(\frac{1-\alpha}{2} \right) \delta \left(\omega + \frac{2\pi}{M} \right) \right] \quad (59)$$

- And then

$$W_H(e^{j\omega}) = \alpha W_r(e^{j\omega}) + \left(\frac{1-\alpha}{2} \right) W_r \left(e^{j(\omega - \frac{2\pi}{M})} \right) + \left(\frac{1-\alpha}{2} \right) W_r \left(e^{j(\omega + \frac{2\pi}{M})} \right) \quad (60)$$

Hamming and Hann windows

- From this equation, one notices that $W_H(e^{j\omega})$ is composed of three versions of the rectangular spectrum $W_r(e^{j\omega})$: the main component, $\alpha W_r(e^{j\omega})$, centered at $\omega = 0$, and two additional ones with smaller amplitudes, centered at $\omega = \pm 2\pi/M$, that reduce the secondary lobe of the main component.

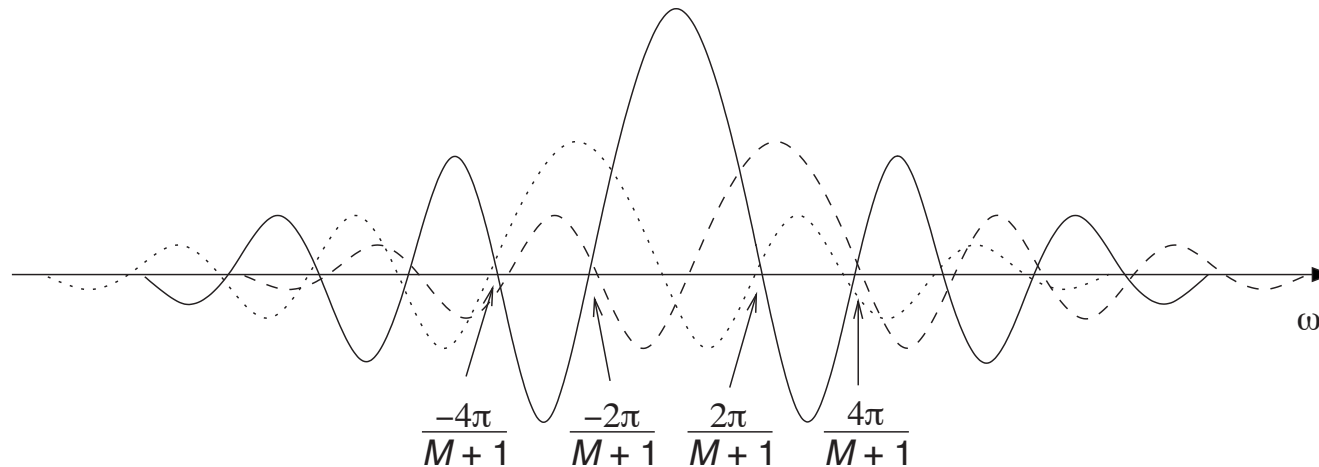


Figure 8: The three components of the generalized Hamming window combine to reduce the resulting secondary lobes. (Solid line – $\alpha W_r(e^{j\omega})$; dashed line – $\frac{1-\alpha}{2} W_r(e^{j(\omega - \frac{\pi}{M})})$; dotted line – $\frac{1-\alpha}{2} W_r(e^{j(\omega + \frac{\pi}{M})})$.)

Hamming and Hann windows

- The main characteristics of the generalized Hamming window are:
 - All three $W_r(e^{j\omega})$ components have zeros close to $\omega = \pm \frac{4\pi}{M+1}$. Hence, the main-lobe total width is $\frac{8\pi}{M+1}$.
 - When $\alpha = 0.54$, the main-lobe total energy is approximately 99.96% of the window total energy.
 - The transition band of the Hamming window is larger than the transition band of the rectangular window, due to its wider main lobe.
 - The ratio between the amplitudes of the main and secondary lobes of the Hamming window is much larger than for the rectangular window.
 - The stopband attenuation for the Hamming window is larger than the attenuation for the rectangular window.

Blackman window

- The Blackman window is defined as

$$w_B(n) = \begin{cases} 0.42 + 0.5 \cos\left(\frac{2\pi n}{M}\right) + 0.08 \cos\left(\frac{4\pi n}{M}\right), & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (61)$$

- Compared to the Hamming window function, the Blackman window introduces a second cosine term to further reduce the effects of the secondary lobes of $W_r(e^{j\omega})$.
- The Blackman window is characterized by the following issues:
 - The main-lobe width is approximately $\frac{12\pi}{M+1}$, which is wider than that for the previous windows.
 - The passband ripples are smaller than in the previous windows.
 - The stopband attenuation is larger than in the previous windows.

Example 5.3

- Design a bandstop filter satisfying the specification below using the rectangular, Hamming, Hann, and Blackman windows:

$$\left. \begin{aligned} M &= 80 \\ \Omega_{p_1} &= 2000 \text{ rad/s} \\ \Omega_{p_2} &= 4000 \text{ rad/s} \\ \Omega_s &= 10\,000 \text{ rad/s} \end{aligned} \right\} \quad (62)$$

Example 5.3 - Solution

- This time, filter specifications are given in the analog frequency domain. Hence, one must first normalize Ω_{p_1} and Ω_{p_2} before employing a script similar to the one given in Example 5.2:

```
M = 80;
```

```
Omega_c1 = 2000; Omega_c2 = 4000; Omega_s = 10000;
```

```
wc1 = Omega_c1*2*pi/Omega_s; wc2 =
```

```
Omega_c2*2*pi/Omega_s;
```

```
n = 1:M/2;
```

```
h0 = 1 - (wc2 - wc1)/pi;
```

```
haux = (sin(wc1.*n) - sin(wc2.*n))./(pi.*n);
```

```
h = [fliplr(haux) h0 haux];
```

to obtain the impulse response using the rectangular window.

Example 5.3 - Solution

- For the other three windows, one must multiply sample-by-sample $h(n)$ above by the corresponding window obtained with the MATLAB commands `hamming(M+1);`, `hanning(M+1);`, and `blackman(M+1);`.
- The resulting impulse responses are shown in Tables 5–8, where only the filter coefficients for $0 \leq n \leq 40$ are given, since the remaining coefficients can be obtained as $h(n) = h(80 - n)$.
- The magnitude responses associated to the four impulse responses listed in Tables 5–8 are depicted in Figure 9. The reader should notice the compromise between the transition bandwidth and the ripple in the passband and stopband when going from the rectangular to the Blackman window, that is, as the ripple decreases, the width of the transition band increases accordingly.

Example 5.3 - Solution

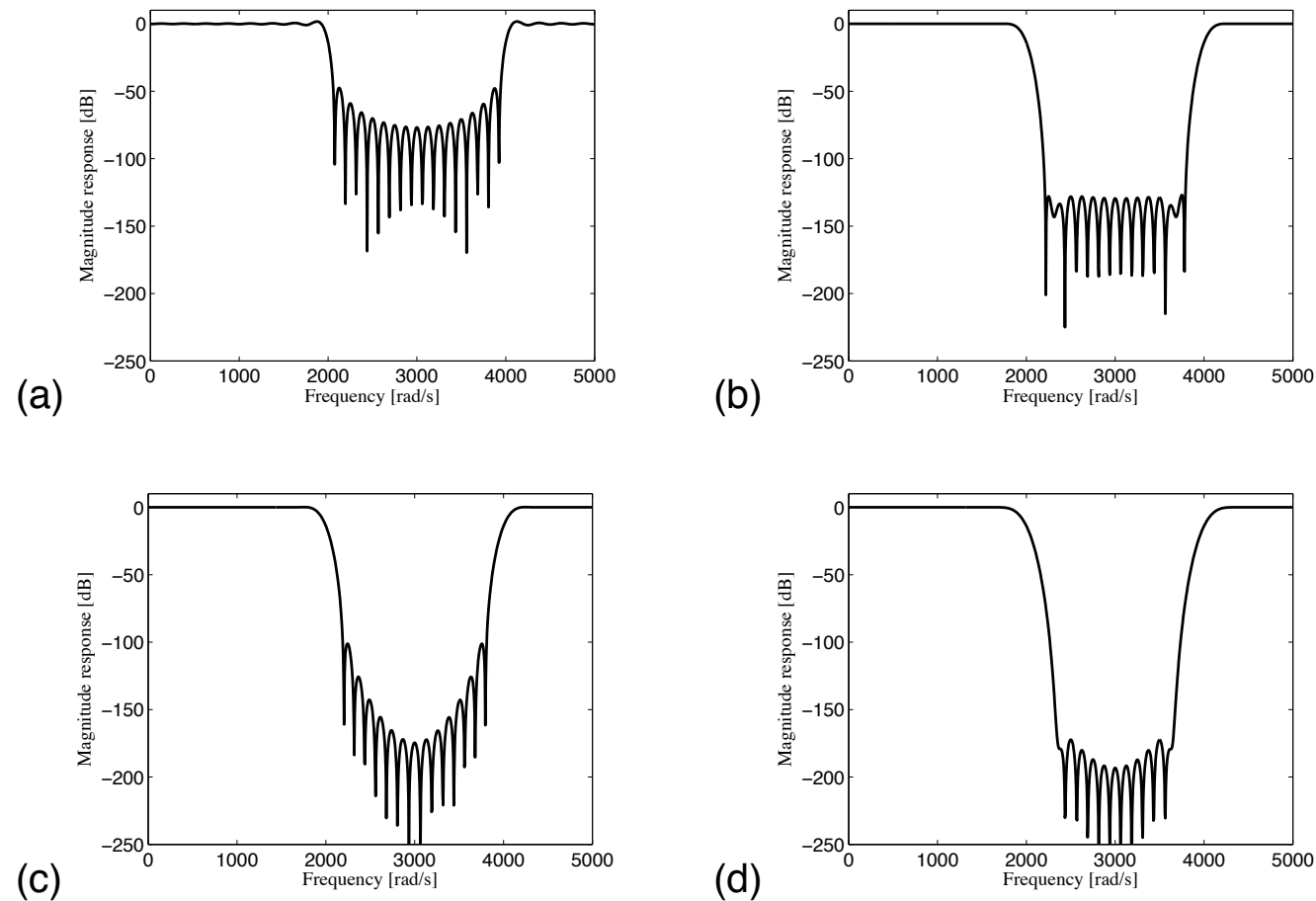


Figure 9: Magnitude responses when using: (a) rectangular; (b) Hamming; (c) Hann; (d) Blackman windows.

Kaiser window

- All the window functions seen so far allow us to control the transition band through a proper choice of the filter order M .
- However, no control can be achieved over the passband and stopband ripples, which makes these windows of little use when designing filters with prescribed frequency specifications, such as that depicted in Figure 10.
- Such problems are overcome with the Kaiser and Dolph-Chebyshev windows, presented in this and in the next subsections.

Kaiser window

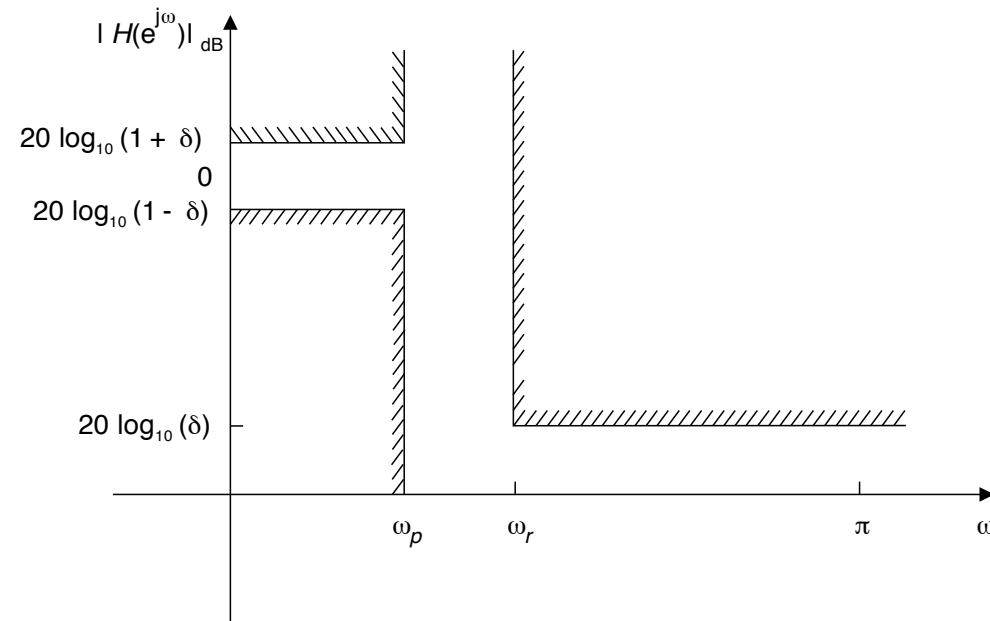


Figure 10: Typical specification of a lowpass filter. The specifications are in terms of the digital frequency $\omega = 2\pi \frac{\Omega}{\Omega_s} = \Omega T$.

Kaiser window

- As seen earlier in this section, the ideal window should be a finite-duration function such that most of its spectral energy is concentrated around $|\omega| = 0$, quickly decaying when $|\omega|$ increases.
- There is a family of continuous-time functions, called the prolate spheroidal functions, which are optimal for achieving these properties.

Kaiser window

- Such functions, although very difficult to implement in practice, can be effectively approximated with the hyperbolic-sine $I_0(\cdot)$ functions as

$$w(t) = \begin{cases} \frac{I_0 \left[\beta \sqrt{1 - \left(\frac{t}{\tau}\right)^2} \right]}{I_0(\beta)}, & \text{for } |t| \leq \tau \\ 0, & \text{for } |t| > \tau \end{cases} \quad (63)$$

where $\beta = \Omega_a \tau$ and $I_0(x)$ is the zeroth-order modified Bessel function of the first kind, which can be efficiently determined through its series expansion given by

$$I_0(x) = 1 + \sum_{k=1}^{\infty} \left[\frac{\left(\frac{x}{2}\right)^k}{k!} \right]^2 \quad (64)$$

Kaiser window

- The Fourier transform of $w(t)$ is given by

$$W(\Omega) = \frac{2\tau \sin \left[\beta \sqrt{\left(\frac{\Omega}{\Omega_a}\right)^2 - 1} \right]}{\beta I_0(\beta) \sqrt{\left(\frac{\Omega}{\Omega_a}\right)^2 - 1}} \quad (65)$$

- The Kaiser window is derived from equation (63) by making the transformation to the discrete-time domain given by $\tau \rightarrow \frac{M}{2}T$ and $t \rightarrow nT$. The window is then described by

$$w_K(n) = \begin{cases} \frac{I_0 \left[\beta \sqrt{1 - \left(\frac{2n}{M}\right)^2} \right]}{I_0(\beta)}, & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (66)$$

- Since the functions given by equation (65) tend to be highly concentrated around $|\Omega| = 0$, we can assume that $W(\Omega) \approx 0$, for $|\Omega| \geq \frac{\Omega_s}{2}$.

Kaiser window

- Therefore, we can approximate the frequency response for the Kaiser window by

$$W_K(e^{j\omega}) \approx \frac{1}{T} W\left(\frac{\omega}{T}\right) \quad (67)$$

where $W(\Omega)$ is given by equation (65) when τ is replaced by $\frac{M}{2}T$.

- This yields

$$W_K(e^{j\omega}) \approx \frac{M \sin \left[\beta \sqrt{\left(\frac{\omega}{\omega_a}\right)^2 - 1} \right]}{\beta I_0(\beta) \sqrt{\left(\frac{\omega}{\omega_a}\right)^2 - 1}} \quad (68)$$

where $\omega_a = \Omega_a T$ and $\beta = \Omega_a \tau = \frac{\omega_a}{T} \frac{M}{2} T = \omega_a \frac{M}{2}$.

- The main advantage of the Kaiser window appears in the design of FIR digital filters with prescribed specifications, such as that depicted in Figure 10, where the parameter β is used to control some filter characteristics.

Kaiser window

- The overall procedure for designing FIR filters using the Kaiser window is as follows:
 - (i) From the ideal frequency response that the filter is supposed to approximate, determine the impulse response $h(n)$ using Table 1. If the filter is either lowpass or highpass, one should make $\Omega_c = \frac{\Omega_p + \Omega_r}{2}$. The case of bandpass and bandstop filters is dealt with later in this subsection.
 - (ii) Given the maximum passband ripple in dB, A_p , and the minimum stopband attenuation in dB, A_r , determine the corresponding ripples

$$\delta_p = \frac{10^{0.05A_p} - 1}{10^{0.05A_p} + 1} \quad (69)$$

$$\delta_r = 10^{-0.05A_r} \quad (70)$$

Kaiser window

- (cont.)
 - (iii) As with all other window functions, the Kaiser window can only be used to design filters that present the same passband and stopband ripples. Therefore, in order to satisfy the prescribed specifications, one should use $\delta = \min\{\delta_p, \delta_r\}$.
 - (iv) Compute the resulting passband ripple and stopband attenuation in dB using

$$A_p = 20 \log \frac{1 + \delta}{1 - \delta} \quad (71)$$

$$A_r = -20 \log \delta \quad (72)$$

- (v) Given the passband and stopband edges, Ω_p and Ω_r , respectively, compute the transition bandwidth $T_r = (\Omega_r - \Omega_p)$.

Kaiser window

- (cont.)
 - (vi) Compute β using

$$\beta = \begin{cases} 0, & \text{for } A_r \leq 21 \\ 0.5842(A_r - 21)^{0.4} + 0.07886(A_r - 21), & \text{for } 21 < A_r \leq 50 \\ 0.1102(A_r - 8.7), & \text{for } 50 < A_r \end{cases} \quad (73)$$

This empirical formula was devised by Kaiser based on the behavior of the function $W(\Omega)$ in equation (65).

- (vii) Defining the normalized window length $D = \frac{T_r M}{\Omega_s}$, where Ω_s is the sampling frequency, we have that:

$$D = \begin{cases} 0.9222, & \text{for } A_r \leq 21 \\ \frac{(A_r - 7.95)}{14.36}, & \text{for } 21 < A_r \end{cases} \quad (74)$$

Kaiser window

- (cont.)
 - (viii) Having computed D using equation (74), we can determine the filter order M as the smallest even number that satisfies

$$M \geq \frac{\Omega_s D}{T_r} \quad (75)$$

One should remember that T_r must be in the same units as Ω_s .

- (ix) With M and β determined, we compute the window $w_K(n)$ using equation (66). We are now ready to form the sequence $h'(n) = w_K(n)h(n)$, where $h(n)$ is the ideal filter impulse response computed in step (i).
- (x) The desired transfer function is then given by

$$H(z) = z^{-\frac{M}{2}} \mathcal{Z}\{h'(n)\} \quad (76)$$

Kaiser window

- The above procedure applies to lowpass filters (see Figure 10) as well as highpass filters. If the filter is either bandpass or bandstop, we must include the following reasoning in step (i) above:

- 1. Compute the narrower transition band

$$T_r = \pm \min\{|\Omega_{r_1} - \Omega_{p_1}|, |\Omega_{p_2} - \Omega_{r_2}|\} \quad (77)$$

Notice that T_r is negative for bandpass filters and positive for bandstop filters.

- 2. Determine the two central frequencies as

$$\Omega_{c_1} = \left(\Omega_{p_1} + \frac{T_r}{2} \right) \quad (78)$$

$$\Omega_{c_2} = \left(\Omega_{p_2} - \frac{T_r}{2} \right) \quad (79)$$

Kaiser window

- A typical magnitude specification for a bandstop filter is depicted in Figure 11.

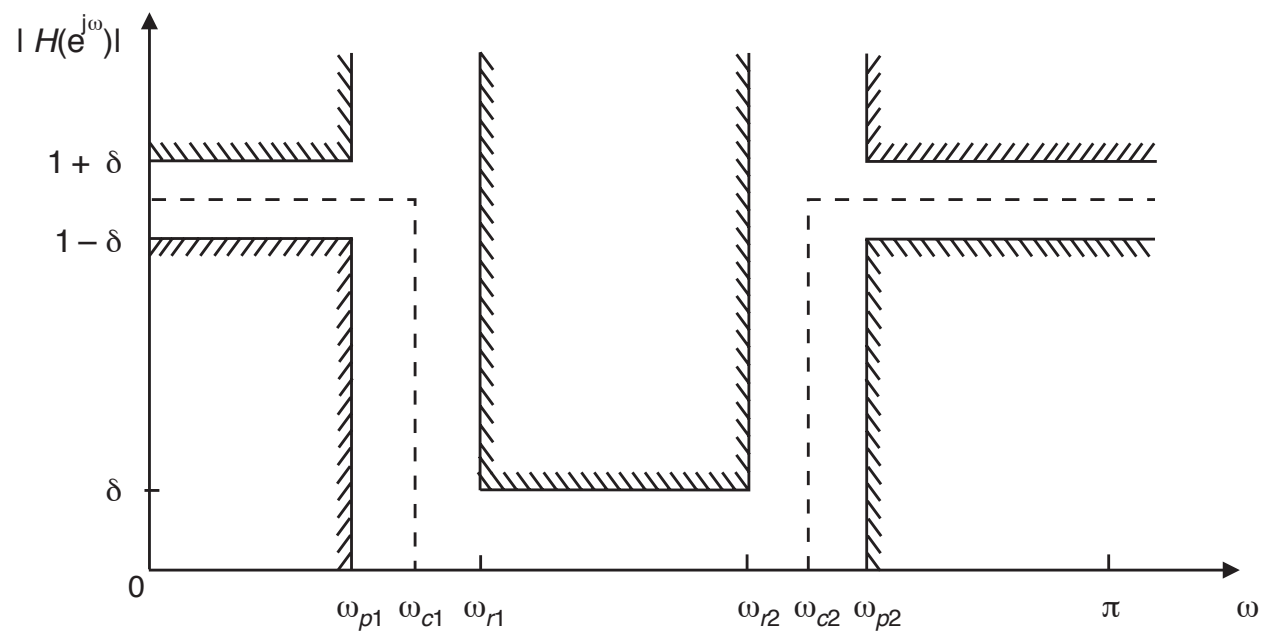


Figure 11: Typical specification of a bandstop filter.

Example 5.4

- Design a bandstop filter satisfying the specification below using the Kaiser window:

$$\left. \begin{aligned} A_p &= 1.0 \text{ dB} \\ A_r &= 45 \text{ dB} \\ \Omega_{p_1} &= 800 \text{ Hz} \\ \Omega_{r_1} &= 950 \text{ Hz} \\ \Omega_{r_2} &= 1050 \text{ Hz} \\ \Omega_{p_2} &= 1200 \text{ Hz} \\ \Omega_s &= 6000 \text{ Hz} \end{aligned} \right\} \quad (80)$$

Example 5.4 - Solution

- Following the procedure described above, the resulting filter is obtained as follows (note that in the FIR design procedure described above, the parameters of the Kaiser window depend only on the ratio of the analog frequencies in the filter specification to the sampling frequency; therefore, the frequencies can be entered in the formula in hertz, as long as the sampling frequency Ω_s is also in hertz):
 - (i) From equations (77)–(79), we have that

$$T_r = + \min\{(950 - 800), (1200 - 1050)\} = 150 \text{ Hz} \quad (81)$$

$$\Omega_{c_1} = 800 + 75 = 875 \text{ Hz} \quad (82)$$

$$\Omega_{c_2} = 1200 - 75 = 1125 \text{ Hz} \quad (83)$$

Example 5.4 - Solution

- (cont.)
 - (ii) From equations (69) and (70),

$$\delta_p = \frac{10^{0.05} - 1}{10^{0.05} + 1} = 0.0575 \quad (84)$$

$$\delta_r = 10^{-0.05 \times 45} = 0.00562 \quad (85)$$

$$(86)$$

- (iii) $\delta = \min\{0.0575, 0.00562\} = 0.00562$

- (iv) From equations (71) and (72),

$$A_p = 20 \log \frac{1 + 0.00562}{1 - 0.00562} = 0.0977 \text{ dB} \quad (87)$$

$$A_r = -20 \log 0.00562 = 45 \text{ dB} \quad (88)$$

Example 5.4 - Solution

- (cont.)
 - (v) T_r has already been computed as 150 Hz in step (i).
 - (vi) From equation (73), since $A_r = 45$ dB, then

$$\beta = 0.5842(45 - 21)^{0.4} + 0.07886(45 - 21) = 3.9754327 \quad (89)$$

- (vii) From equation (74), since $A_r = 45$ dB, then

$$D = \frac{(45 - 7.95)}{14.36} = 2.5800835 \quad (90)$$

- (viii) Since the sampling period is $T = \frac{1}{6000}$ s, we have, from equation (75),

$$M \geq \frac{6000 \times 2.5800835}{150} = 103.20334 \Rightarrow M = 104 \quad (91)$$

Example 5.4 - Solution

- This whole procedure is implemented by a simple MATLAB script:

```
Ap = 1; Ar = 45;
Omega_p1 = 800; Omega_r1 = 950;
Omega_r2 = 1050; Omega_p2 = 1200;
Omega_s = 6000;
delta_p = (10^(0.05*Ap) - 1)/(10^(0.05*Ap) + 1);
delta_r = 10^(-0.05*Ar);
F = [Omega_p1 Omega_r1 Omega_r2 Omega_p2];
A = [1 0 1];
ripples = [delta_p delta_r delta_p];
[M,Wn,beta,FILTYPE] =
kaiserord(F,A,ripples,Omega_s);
which yields as outputs beta = 3.9754 and M = 104, as determined above.
```

Example 5.4 - Solution

- In this short script, the auxiliary vectors `A` and `ripples` specify the desired gain and allowed ripple, respectively, in each filter band.
- The Kaiser window coefficients are determined by:

```
kaiser_win = kaiser(M+1,beta);
```

and are shown in Figure 12 along with the associated magnitude response.

Example 5.4 - Solution

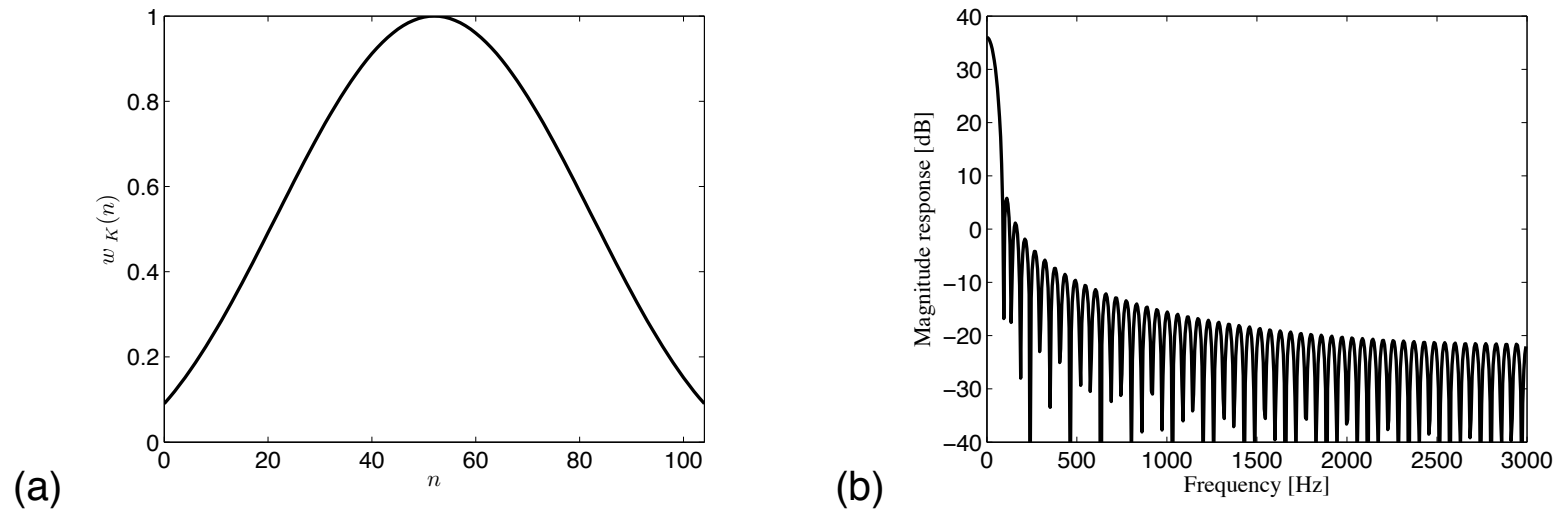


Figure 12: Kaiser window: (a) window function; (b) magnitude response.

Example 5.4 - Solution

- The desired filter is obtained using the `fir1` command, as exemplified by:

```
h = fir1(M,Wn,FILTYPE,kaiser_win,'noscale');
```

where the `noscale` flag avoids the unitary gain at the first passband center imposed by MATLAB.
- The designed filter characteristics are summarized in Table 9.

Example 5.4 - Solution

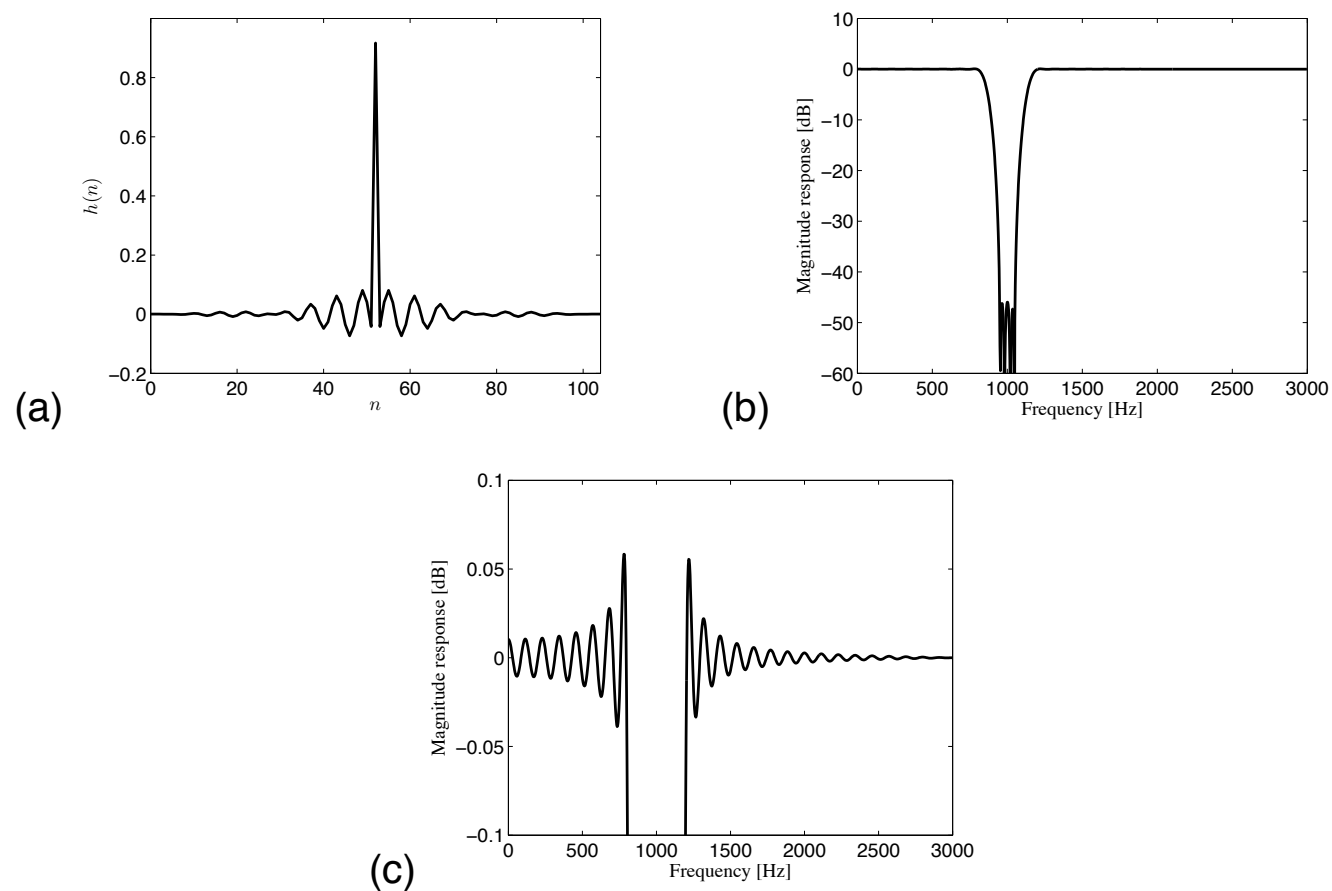


Figure 13: Resulting bandstop filter: (a) impulse response; (b) magnitude response; (c) passband detail.

Dolph-Chebyshev window

- Based on the M th-order Chebyshev polynomial given by

$$C_M(x) = \begin{cases} \cos [M \cos^{-1}(x)], & \text{for } |x| \leq 1 \\ \cosh [M \cosh^{-1}(x)], & \text{for } |x| > 1 \end{cases} \quad (92)$$

the Dolph-Chebyshev window is defined as

$$w_{DC}(n) = \begin{cases} \frac{1}{M+1} \left\{ \frac{1}{r} + 2 \sum_{i=1}^{\frac{M}{2}} C_M \left[x_0 \cos \left(\frac{i\pi}{M+1} \right) \right] \cos \left(\frac{2ni\pi}{M+1} \right) \right\}, & \text{for } |n| \leq \frac{M}{2} \\ 0, & \text{for } |n| > \frac{M}{2} \end{cases} \quad (93)$$

where r is the ripple ratio defined as $r = \frac{\delta_r}{\delta_p}$ and

$$x_0 = \cosh \left[\frac{1}{M} \cosh^{-1} \left(\frac{1}{r} \right) \right] \quad (94)$$

Dolph-Chebyshev window

- The procedure for designing FIR filters using the Dolph-Chebyshev window is very similar to the one for the Kaiser window:
 - (i) Perform steps (i) and (ii) of the Kaiser procedure.
 - (ii) Determine $r = \frac{\delta_r}{\delta_p}$.
 - (iii) Perform steps (iii)–(v) and (vii)–(viii) of the Kaiser procedure, to determine the filter order M . In step (vii), however, as the stopband attenuation achieved with the Dolph-Chebyshev window is typically 1 to 4 dB higher than that obtained using the Kaiser window, one should compute D for the Dolph-Chebyshev window using equation (5.74) with A_r replaced by $A_r + 2.5$.
 - (iv) With r and M determined, compute x_0 from equation (94), and then compute the window coefficients from equation (93).
 - (v) We are now ready to form the sequence $h'(n) = w_{DC}(n)h(n)$, where $h(n)$ is the ideal filter impulse response computed in step (86).
 - (vi) Perform step (x) of the Kaiser procedure to determine the resulting FIR filter.

Dolph-Chebyshev window

- Overall, the Dolph-Chebyshev window is characterized by:
 - The main-lobe width, and consequently the resulting filter transition band, can be controlled by varying M .
 - The ripple ratio is controlled through an independent parameter r .
 - All secondary lobes have the same amplitude. Therefore, the stopband of the resulting filter is equiripple.

Maximally flat FIR filter approximation

- Maximally flat approximations should be employed when a signal must be preserved with minimal error around the zero frequency or when a monotone frequency response is necessary.
- FIR filters with a maximally flat frequency response at $\omega = 0$ and $\omega = \pi$ were introduced by Herrmann. We consider here, following the standard literature on the subject, the lowpass Type-I FIR filter of even order M and symmetric impulse response.
- In this case, the frequency response of a maximally flat FIR filter is determined in such a way that $H(e^{j\omega}) - 1$ has $2L$ zeros at $\omega = 0$, and $H(e^{j\omega})$ has $2K$ zeros at $\omega = \pi$.
- To achieve a maximally flat response, the filter order M must satisfy $M = (2K + 2L - 2)$. Thus, the first $2L - 1$ derivatives of $H(e^{j\omega})$ are zero at $\omega = 0$, and the first $2K - 1$ derivatives of $H(e^{j\omega})$ are zero at $\omega = \pi$.

Maximally flat FIR filter approximation

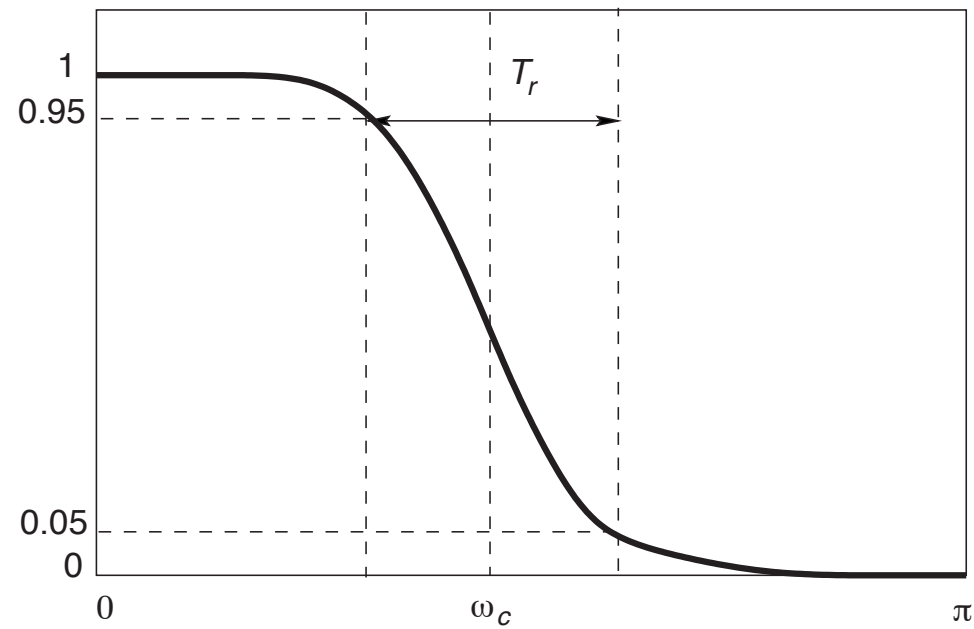


Figure 14: Typical specification of a maximally flat lowpass FIR filter.

Example 5.5

- Design a maximally flat lowpass filter satisfying the specification below:

$$\left. \begin{aligned} \Omega_c &= 0.3\pi \text{ rad/s} \\ T_r &= 0.2\pi \text{ rad/s} \\ \Omega_s &= 2\pi \text{ rad/s} \end{aligned} \right\} \quad (108)$$

Example 5.5 - Solution

Table 11: Coefficients of the lowpass filter designed with the maximally flat method.

<u>d(0) to d(5)</u>
$d(0) = 1$
$d(1) = 27$
$d(2) = 378$
$d(3) = 3\,654$
$d(4) = 27\,405$
$d(5) = 169\,911$
<u>$d(6) = 906\,192$</u>

Example 5.5 - Solution

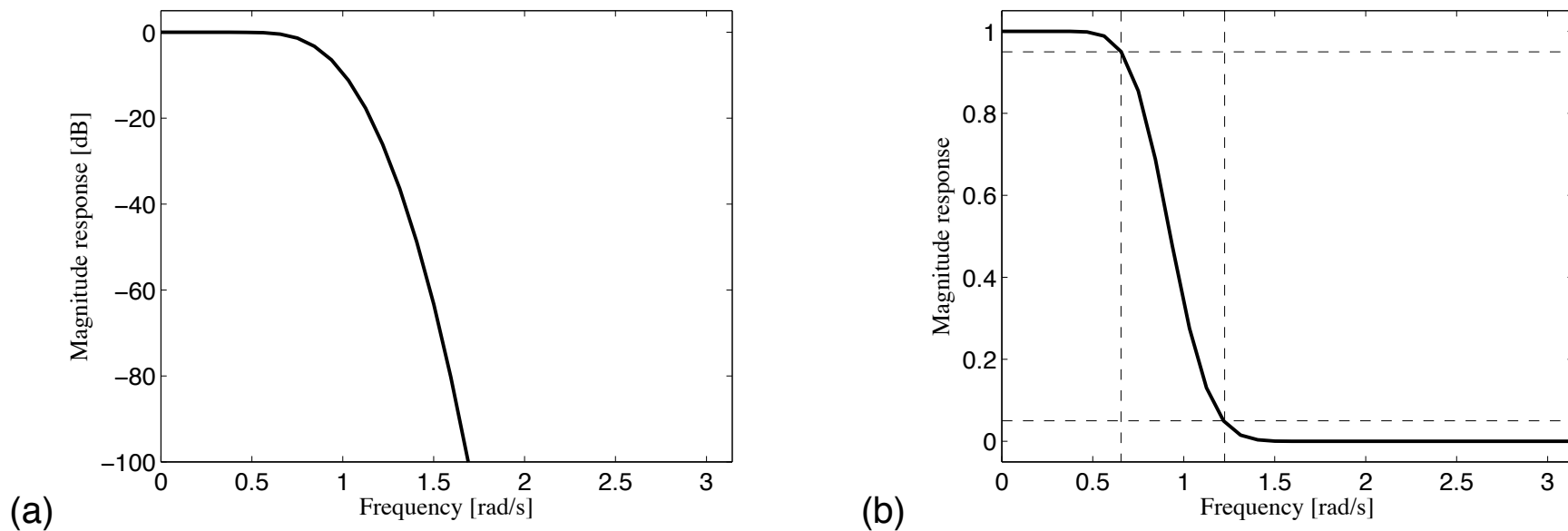


Figure 15: magnitude response of maximally flat lowpass FIR filter: (a) linear scale; (b) dB scale.

FIR filter approximation by optimization

- The window method seen in Section 5.4 has a very straightforward design procedure for approximating the desired magnitude response. However, the window method is not efficient for designing, for example, FIR filters with different ripples in the passband and stopband, or nonsymmetric bandpass or bandstop filters.
- To fill this gap, in this section we present several numerical algorithms for designing more general FIR digital filters.
- In many signal processing systems, filters with linear or zero phase are required. Unfortunately, filters designed to have zero phase are not causal; this can be a problem in applications where very little processing delay is permissible.
- Also, nonlinear phase causes distortion in the processed signal, which can be very perceptible in applications like data transmission, image processing, and so on.

FIR filter approximation by optimization

- One of the major advantages of using an FIR system instead of a causal IIR system is that FIR systems can be designed with exact linear phase.
- As seen in Subsection 4.2.3, there are four distinct cases where an FIR filter presents linear phase. To present general algorithms for designing linear-phase FIR filters, a unified representation of these four cases is necessary.
- We define an auxiliary function $P(\omega)$ as

$$P(\omega) = \sum_{l=0}^L p(l) \cos(\omega l) \quad (109)$$

where $L + 1$ is the number of cosine functions in the expression of $H(e^{j\omega})$.

FIR filter approximation by optimization

- Based on this function, we can express the frequency response of the four types of linear-phase FIR filters as:
 - Type I: Even order M and symmetric impulse response. We get

$$\begin{aligned}
 H(e^{j\omega}) &= e^{-j\omega \frac{M}{2}} \sum_{m=0}^{\frac{M}{2}} a(m) \cos(\omega m) \\
 &= e^{-j\omega \frac{M}{2}} \sum_{l=0}^{\frac{M}{2}} p(l) \cos(\omega l) \\
 &= e^{-j\omega \frac{M}{2}} P(\omega)
 \end{aligned} \tag{110}$$

with

$$a(m) = p(m), \text{ for } m = 0, 1, \dots, L \tag{111}$$

where $L = \frac{M}{2}$.

FIR filter approximation by optimization

- (cont.)
 - Type II: Odd order M and symmetric impulse response. In this case, we have

$$H(e^{j\omega}) = e^{-j\omega \frac{M}{2}} \sum_{m=1}^{\frac{M+1}{2}} b(m) \cos \left[\omega \left(m - \frac{1}{2} \right) \right] \quad (112)$$

Using

$$b(m) = \begin{cases} p(0) + \frac{1}{2}p(1), & \text{for } m = 1 \\ \frac{1}{2} (p(m-1) + p(m)), & \text{for } m = 2, 3, \dots, L \\ \frac{1}{2}p(L), & \text{for } m = L + 1 \end{cases} \quad (113)$$

FIR filter approximation by optimization

- With $L = \frac{M-1}{2}$, then $H(e^{j\omega})$ can be written in the form

$$H(e^{j\omega}) = e^{-j\omega \frac{M}{2}} \cos\left(\frac{\omega}{2}\right) P(\omega) \quad (114)$$

using the trigonometric identity

$$2 \cos\left(\frac{\omega}{2}\right) \cos(\omega m) = \cos\left[\omega\left(m + \frac{1}{2}\right)\right] + \cos\left[\omega\left(m - \frac{1}{2}\right)\right] \quad (115)$$

- The complete algebraic development is left as an exercise to the interested reader.

FIR filter approximation by optimization

- (cont.)
 - Type III: Even order M and antisymmetric impulse response. In this case, we have

$$H(e^{j\omega}) = e^{-j(\omega \frac{M}{2} - \frac{\pi}{2})} \sum_{m=1}^{\frac{M}{2}} c(m) \sin(\omega m) \quad (116)$$

and then, by substituting

$$c(m) = \begin{cases} p(0) - \frac{1}{2}p(2), & \text{for } m = 1 \\ \frac{1}{2}(p(m-1) - p(m+1)), & \text{for } m = 2, 3, \dots, (L-1) \\ \frac{1}{2}p(m-1), & \text{for } m = L, L+1 \end{cases} \quad (117)$$

FIR filter approximation by optimization

- With $L = \frac{M}{2} - 1$, equation (116) can be written as

$$H(e^{j\omega}) = e^{-j(\omega \frac{M}{2} - \frac{\pi}{2})} \sin \omega P(\omega) \quad (118)$$

using, in this case, the identity

$$\sin \omega \cos (\omega m) = \sin [\omega (m + 1)] - \sin [\omega (m - 1)] \quad (119)$$

- Once again, the algebraic proof is left as an exercise at the end of this chapter.

FIR filter approximation by optimization

- (cont.)
 - Type IV: Odd order M and antisymmetric impulse response. We have that

$$H(e^{j\omega}) = e^{-j(\omega \frac{M}{2} - \frac{\pi}{2})} \sum_{m=1}^{\frac{M+1}{2}} d(m) \sin \left[\omega \left(m - \frac{1}{2} \right) \right] \quad (120)$$

By substituting

$$d(m) = \begin{cases} p(0) - \frac{1}{2}p(1), & \text{for } m = 1 \\ \frac{1}{2}(p(m-1) - p(m)), & \text{for } m = 2, 3, \dots, L \\ \frac{1}{2}p(L), & \text{for } m = L + 1 \end{cases} \quad (121)$$

FIR filter approximation by optimization

- With $L = \frac{M-1}{2}$, then $H(e^{j\omega})$ can be written as

$$H(e^{j\omega}) = e^{-j(\omega \frac{M}{2} - \frac{\pi}{2})} \sin\left(\frac{\omega}{2}\right) P(\omega) \quad (122)$$

using the identity

$$2 \sin\left(\frac{\omega}{2}\right) \cos(\omega m) = \sin\left[\omega\left(m + \frac{1}{2}\right)\right] - \sin\left[\omega\left(m - \frac{1}{2}\right)\right] \quad (123)$$

FIR filter approximation by optimization

- Equations (110), (114), (118), and (122) indicate that we can write the frequency response for any linear-phase FIR filter as

$$H(e^{j\omega}) = e^{-j(\alpha\omega - \beta)} Q(\omega)P(\omega) = e^{-j(\alpha\omega - \beta)} A(\omega) \quad (124)$$

where $A(\omega) = Q(\omega)P(\omega)$, $\alpha = \frac{M}{2}$, and for each case we have that:

- Type I: $\beta = 0$ and $Q(\omega) = 1$
- Type II: $\beta = 0$ and $Q(\omega) = \cos(\frac{\omega}{2})$
- Type III: $\beta = \frac{\pi}{2}$ and $Q(\omega) = \sin(\omega)$
- Type IV: $\beta = \frac{\pi}{2}$ and $Q(\omega) = \sin(\frac{\omega}{2})$.

FIR filter approximation by optimization

- Let $D(\omega)$ be the desired amplitude response. We define the weighted error function as

$$E(\omega) = W(\omega)(D(\omega) - A(\omega)) \quad (125)$$

- We can then write $E(\omega)$ as

$$\begin{aligned} E(\omega) &= W(\omega)(D(\omega) - Q(\omega)P(\omega)) \\ &= W(\omega)Q(\omega) \left(\frac{D(\omega)}{Q(\omega)} - P(\omega) \right) \end{aligned} \quad (126)$$

for all $0 \leq \omega \leq \pi$, as $Q(\omega)$ is independent of the coefficients for each ω .

FIR filter approximation by optimization

- Defining

$$W_q(\omega) = W(\omega)Q(\omega) \quad (127)$$

$$D_q(\omega) = \frac{D(\omega)}{Q(\omega)} \quad (128)$$

the error function can be rewritten as

$$E(\omega) = W_q(\omega)(D_q(\omega) - P(\omega)) \quad (129)$$

and one can formulate the optimization problem for approximating linear-phase FIR filters as:

- *Determine the set of coefficients $p(l)$ that minimizes some objective function of the weighted error $E(\omega)$ over a set of prescribed frequency bands.*

FIR filter approximation by optimization

- To solve such a problem numerically, we evaluate the weighted error function on a dense frequency grid with $0 \leq \omega_i \leq \pi$, for $i = 1, 2, \dots, KM$, where M is the filter order, obtaining a good discrete approximation of $E(\omega)$.
- For most practical purposes, using $8 \leq K \leq 16$ is recommended. Points associated with the transition bands can be disregarded, and the remaining frequencies should be linearly redistributed in the passbands and stopbands to include their corresponding edges.

FIR filter approximation by optimization

- Thus, the following equation results

$$\mathbf{e} = \mathbf{W}_q (\mathbf{d}_q - \mathbf{U}\mathbf{p}) \quad (130)$$

where

$$\mathbf{e} = [E(\omega_1) E(\omega_2) \cdots E(\omega_{\overline{KM}})]^T \quad (131)$$

$$\mathbf{W}_q = \text{diag} [W_q(\omega_1) W_q(\omega_2) \cdots W_q(\omega_{\overline{KM}})] \quad (132)$$

$$\mathbf{d}_q = [D_q(\omega_1) D_q(\omega_2) \cdots D_q(\omega_{\overline{KM}})]^T \quad (133)$$

$$\mathbf{U} = \begin{bmatrix} 1 & \cos(\omega_1) & \cos(2\omega_1) & \cdots & \cos(L\omega_1) \\ 1 & \cos(\omega_2) & \cos(2\omega_2) & \cdots & \cos(L\omega_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cos(\omega_{\overline{KM}}) & \cos(2\omega_{\overline{KM}}) & \cdots & \cos(L\omega_{\overline{KM}}) \end{bmatrix} \quad (134)$$

$$\mathbf{p} = [p(0) p(1) \cdots p(L)]^T \quad (135)$$

FIR filter approximation by optimization

- With $\overline{KM} \leq KM$, as the original frequencies in the transition band were discarded.
- For the four standard types of filters, namely lowpass, highpass, bandpass, and bandstop, as well as differentiators and Hilbert transformers, the definitions of $W(\omega)$ and $D(\omega)$ are summarized in Table 12.

FIR filter approximation by optimization

Table 12: Weight functions and ideal magnitude responses for basic lowpass, highpass, bandpass, and bandstop filters, as well as differentiators and Hilbert transformers.

Filter type	Weight function $W(\omega)$	Ideal amplitude response $D(\omega)$
Lowpass	$\begin{cases} 1, & \text{for } 0 \leq \omega \leq \omega_p \\ \frac{\delta_p}{\delta_r}, & \text{for } \omega_r \leq \omega \leq \pi \end{cases}$	$\begin{cases} 1, & \text{for } 0 \leq \omega \leq \omega_p \\ 0, & \text{for } \omega_r \leq \omega \leq \pi \end{cases}$
Highpass	$\begin{cases} \frac{\delta_p}{\delta_r}, & \text{for } 0 \leq \omega \leq \omega_r \\ 1, & \text{for } \omega_p \leq \omega \leq \pi \end{cases}$	$\begin{cases} 0, & \text{for } 0 \leq \omega \leq \omega_r \\ 1, & \text{for } \omega_p \leq \omega \leq \pi \end{cases}$

FIR filter approximation by optimization

Filter type	Weight function $W(\omega)$	Ideal amplitude response $D(\omega)$
Bandpass	$\begin{cases} \frac{\delta_p}{\delta_r}, & \text{for } 0 \leq \omega \leq \omega_{r_1} \\ 1, & \text{for } \omega_{p_1} \leq \omega \leq \omega_{p_2} \\ \frac{\delta_p}{\delta_r}, & \text{for } \omega_{r_2} \leq \omega \leq \pi \end{cases}$	$\begin{cases} 0, & \text{for } 0 \leq \omega \leq \omega_{r_1} \\ 1, & \text{for } \omega_{p_1} \leq \omega \leq \omega_{p_2} \\ 0, & \text{for } \omega_{r_2} \leq \omega \leq \pi \end{cases}$
Bandstop	$\begin{cases} 1, & \text{for } 0 \leq \omega \leq \omega_{p_1} \\ \frac{\delta_p}{\delta_r}, & \text{for } \omega_{r_1} \leq \omega \leq \omega_{r_2} \\ 1, & \text{for } \omega_{p_2} \leq \omega \leq \pi \end{cases}$	$\begin{cases} 1, & \text{for } 0 \leq \omega \leq \omega_{p_1} \\ 0, & \text{for } \omega_{r_1} \leq \omega \leq \omega_{r_2} \\ 1, & \text{for } \omega_{p_2} \leq \omega \leq \pi \end{cases}$

FIR filter approximation by optimization

- It is important to remember all design constraints, due to the magnitude and phase characteristics of the four linear-phase filter types, as summarized below, where a ‘Yes’ entry indicates that the corresponding filter structure is suitable to implement the given filter type.

Filter type	Type I	Type II	Type III	Type IV
Lowpass	Yes	Yes	No	No
Highpass	Yes	No	No	Yes
Bandpass	Yes	Yes	Yes	Yes
Bandstop	Yes	No	No	No
Differentiator	No	No	Yes	Yes
Hilbert transformer	No	No	Yes	Yes

Weighted-least-squares method

- In the weighted-least-squares (WLS) approach, the idea is to minimize the square of the energy of the error function $E(\omega)$, that is

$$\min_{\mathbf{p}} \{ \|E(\omega)\|_2^2 \} = \min_{\mathbf{p}} \left\{ \int_0^\pi |E(\omega)|^2 d\omega \right\} \quad (136)$$

- For a discrete set of frequencies, this objective function is approximated by (see equations (130)–(135))

$$\|E(\omega)\|_2^2 \approx \frac{1}{KM} \sum_{k=1}^{\overline{KM}} |E(\omega_k)|^2 = \frac{1}{KM} \mathbf{e}^T \mathbf{e} \quad (137)$$

since in these equations \mathbf{e} is a real vector.

Weighted-least-squares method

- Using equation (130), and noting that \mathbf{W}_q is diagonal, we can write that

$$\begin{aligned}
 \mathbf{e}^T \mathbf{e} &= (\mathbf{d}_q^T - \mathbf{p}^T \mathbf{U}^T) \mathbf{W}_q^T \mathbf{W}_q (\mathbf{d}_q - \mathbf{U} \mathbf{p}) \\
 &= (\mathbf{d}_q^T - \mathbf{p}^T \mathbf{U}^T) \mathbf{W}_q^2 (\mathbf{d}_q - \mathbf{U} \mathbf{p}) \\
 &= \mathbf{d}_q^T \mathbf{W}_q^2 \mathbf{d}_q - \mathbf{d}_q^T \mathbf{W}_q^2 \mathbf{U} \mathbf{p} - \mathbf{p}^T \mathbf{U}^T \mathbf{W}_q^2 \mathbf{d}_q + \mathbf{p}^T \mathbf{U}^T \mathbf{W}_q^2 \mathbf{U} \mathbf{p} \\
 &= \mathbf{d}_q^T \mathbf{W}_q^2 \mathbf{d}_q - 2\mathbf{p}^T \mathbf{U}^T \mathbf{W}_q^2 \mathbf{d}_q + \mathbf{p}^T \mathbf{U}^T \mathbf{W}_q^2 \mathbf{U} \mathbf{p}
 \end{aligned} \tag{138}$$

because $\mathbf{d}_q^T \mathbf{W}_q^2 \mathbf{U} \mathbf{p} = \mathbf{p}^T \mathbf{U}^T \mathbf{W}_q^2 \mathbf{d}_q$, since these two terms are scalar.

- The minimization of such a functional is achieved by calculating its gradient vector with respect to the coefficient vector and equating it to zero.

Weighted-least-squares method

- Since

$$\nabla_{\mathbf{x}}\{\mathbf{Ax}\} = \mathbf{A}^T \quad (139)$$

$$\nabla_{\mathbf{x}}\{\mathbf{x}^T\mathbf{Ax}\} = (\mathbf{A} + \mathbf{A}^T)\mathbf{x} \quad (140)$$

this yields

$$\nabla_{\mathbf{p}}\{\mathbf{e}^T\mathbf{e}\} = -2\mathbf{U}^T\mathbf{W}_q^2\mathbf{d}_q + 2\mathbf{U}^T\mathbf{W}_q^2\mathbf{U}\mathbf{p}^* = \mathbf{0} \quad (141)$$

which implies that

$$\mathbf{p}^* = (\mathbf{U}^T\mathbf{W}_q^2\mathbf{U})^{-1} \mathbf{U}^T\mathbf{W}_q^2\mathbf{d}_q \quad (142)$$

- It can be shown that when the weight function $W(\omega)$ is made constant, the WLS approach is equivalent to the rectangular window presented in the previous section, and so suffers from the same problem of Gibbs' oscillations near the band edges.
- When $W(\omega)$ is not constant, the oscillations still occur but their energies will vary from band to band.

Example 5.6

- Design a Hilbert transformer of order $M = 5$ using the WLS approach by choosing an appropriate grid of only 3 frequencies. Obtain \mathbf{p}^* and the filter transfer function.

Example 5.6 - Solution

- For the odd order $M = 5$, the FIR Hilbert transformer should be of Type IV and the number of coefficients of \mathbf{p} is $(L + 1)$, where $L = \frac{M-1}{2} = 2$.
- According to equations (129) and (130), the response error is

$$\mathbf{e} = \begin{bmatrix} \sin\left(\frac{\omega_1}{2}\right) & 0 & 0 \\ 0 & \sin\left(\frac{\omega_2}{2}\right) & 0 \\ 0 & 0 & \sin\left(\frac{\omega_3}{2}\right) \end{bmatrix} \left\{ \begin{bmatrix} 1 \\ \frac{1}{\sin\left(\frac{\omega_1}{2}\right)} \\ 1 \\ \frac{1}{\sin\left(\frac{\omega_2}{2}\right)} \\ 1 \\ \frac{1}{\sin\left(\frac{\omega_3}{2}\right)} \end{bmatrix} - \begin{bmatrix} 1 & \cos \omega_1 & \cos 2\omega_1 \\ 1 & \cos \omega_2 & \cos 2\omega_2 \\ 1 & \cos \omega_3 & \cos 2\omega_3 \end{bmatrix} \begin{bmatrix} p(0) \\ p(1) \\ p(2) \end{bmatrix} \right\} \quad (143)$$

Example 5.6 - Solution

- We then form the frequency grid within the range defined in Table 12, such as

$$\left. \begin{aligned} \omega_1 &= \frac{\pi}{3} \\ \omega_2 &= \frac{\pi}{2} \\ \omega_2 &= \frac{2\pi}{3} \end{aligned} \right\} \quad (144)$$

in such a way that the error vector becomes

$$\mathbf{e} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 & -1 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} p(0) \\ p(1) \\ p(2) \end{bmatrix} \right\} \quad (145)$$

Example 5.6 - Solution

- The WLS solution requires the following matrix:

$$\begin{aligned}
 \mathbf{U}^T \mathbf{W}_q^2 \mathbf{U} &= \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 & -1 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \\
 &= \frac{1}{4} \begin{bmatrix} 6 & -1 & -4 \\ -1 & 1 & \frac{1}{2} \\ -4 & \frac{1}{2} & 3 \end{bmatrix} \tag{146}
 \end{aligned}$$

whose inverse is

$$(\mathbf{U}^T \mathbf{W}_q^2 \mathbf{U})^{-1} = \frac{1}{3} \begin{bmatrix} 22 & 8 & 28 \\ 8 & 16 & 8 \\ 28 & 8 & 40 \end{bmatrix} \tag{147}$$

Example 5.6 - Solution

- Then, the vector \mathbf{p}^* is computed as follows

$$\begin{aligned}
 \mathbf{p}^* &= (\mathbf{U}^T \mathbf{W}_q^2 \mathbf{U})^{-1} \mathbf{U}^T \mathbf{W}_q^2 \mathbf{d}_q \\
 &= \frac{1}{3} \begin{bmatrix} 22 & 8 & 28 \\ 8 & 16 & 8 \\ 28 & 8 & 40 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & \frac{\sqrt{3}}{2} \end{bmatrix}^2 \begin{bmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} 1.7405 \\ 0.8453 \\ 0.3263 \end{bmatrix} \tag{148}
 \end{aligned}$$

Example 5.6 - Solution

- According to equations (121), we then have

$$\left. \begin{aligned}
 d(1) &= p(0) - \frac{1}{2}p(1) = 2h(2) = 1.31785 \\
 d(2) &= \frac{1}{2}(p(1) - p(2)) = 2h(1) = 0.2595 \\
 d(3) &= \frac{1}{2}p(2) = 2h(0) = 0.16315
 \end{aligned} \right\} \quad (149)$$

and the overall transfer function is given by

$$H(z) = 0.0816 + 0.1298z^{-1} + 0.6589z^{-2} - 0.6589z^{-3} - 0.1298z^{-4} - 0.0816z^{-5} \quad (150)$$

Example 5.6 - Solution

- If a Hilbert filter of same order is designed with the MATLAB `firls` command, which uses a uniform sampling to determine the frequency grid, the resulting transfer function is

$$\bar{H}(z) = -0.0828 - 0.1853z^{-1} - 0.6277z^{-2} + 0.6277z^{-3} + 0.1853z^{-4} + 0.0828z^{-5} \quad (151)$$

- As can be observed in Figure 16, $H(z)$ and $\bar{H}(z)$ have very similar magnitude responses, with the differences arising from the nonuniform frequency grid employed in the $H(z)$ design for didactic purposes.

Example 5.6 - Solution

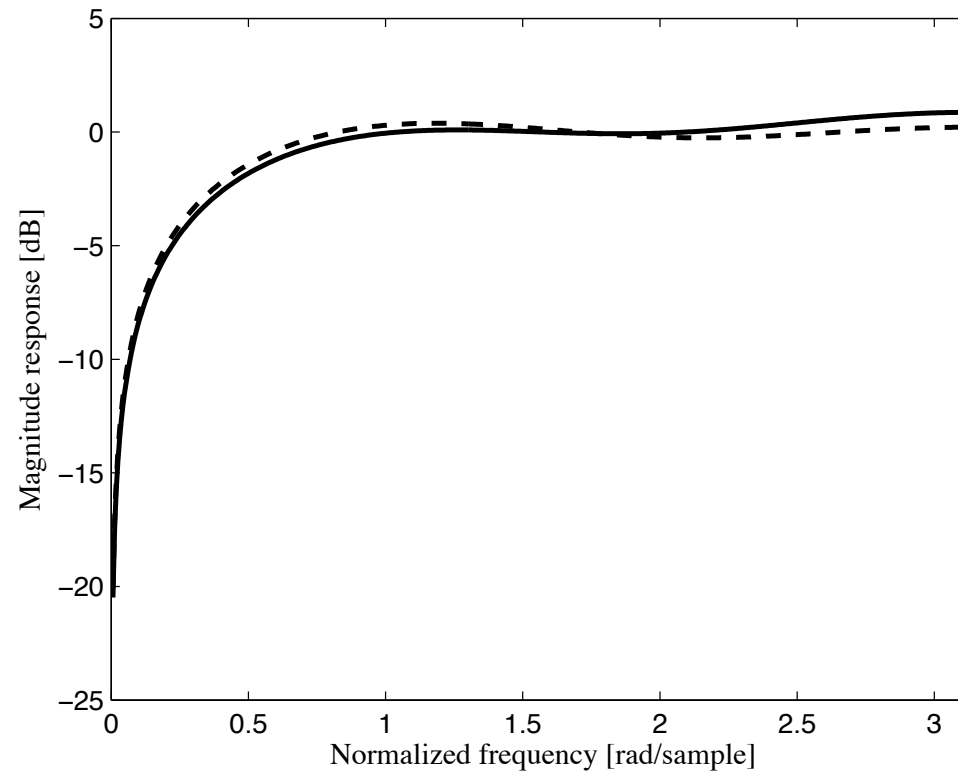


Figure 16: Magnitude responses of Hilbert transformers in Example 5.6: step-by-step $H(z)$ design (solid line) and MATLAB $\overline{H}(z)$ design (dashed line).

Chebyshev method

- In the Chebyshev optimization design approach, the idea is to minimize the maximum absolute value of the error function $E(\omega)$.
- Mathematically, such scheme is described by

$$\min_{\mathbf{p}} \{ \|E(\omega)\|_{\infty} \} = \min_{\mathbf{p}} \left\{ \max_{\omega \in F} \{ |E(\omega)| \} \right\} \quad (152)$$

where F is the set of prescribed frequency bands. This problem can be solved with the help of the following important theorem: