# Chapter 9

## Filter Design

### 9.1 INTRODUCTION

This chapter considers the problem of designing a digital filter. The design process begins with the filter specifications, which may include constraints on the magnitude and/or phase of the frequency response, constraints on the unit sample response or step response of the filter, specification of the type of filter (e.g., FIR or IIR), and the filter order. Once the specifications have been defined, the next step is to find a set of filter coefficients that produce an acceptable filter. After the filter has been designed, the last step is to implement the system in hardware or software, quantizing the filter coefficients if necessary, and choosing an appropriate filter structure (Chap. 8).

### 9.2 FILTER SPECIFICATIONS

Before a filter can be designed, a set of filter specifications must be defined. For example, suppose that we would like to design a low-pass filter with a cutoff frequency $\omega_c$. The frequency response of an ideal low-pass filter with linear phase and a cutoff frequency $\omega_c$ is

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\alpha\omega} & |\omega| \le \omega_c \\ 0 & \omega_c < |\omega| \le \pi \end{cases}$$

which has a unit sample response

$$h_d(n) = \frac{\sin(n-\alpha)\omega_c}{\pi(n-\alpha)}$$

Because this filter is unrealizable (noncausal and unstable), it is necessary to relax the ideal constraints on the frequency response and allow some deviation from the ideal response. The specifications for a low-pass filter will typically have the form

$$\begin{aligned} 1-\delta_p < |H(e^{j\omega})| &\le 1+\delta_p \qquad 0 \le |\omega| < \omega_p \\ |H(e^{j\omega})| &\le \delta_s \qquad \omega_s \le |\omega| < \pi \end{aligned}$$

as illustrated in Fig. 9-1. Thus, the specifications include the passband cutoff frequency, $\omega_p$, the stopband cutoff frequency, $\omega_s$, the passband deviation, $\delta_p$, and the stopband deviation, $\delta_s$. The passband and stopband deviations

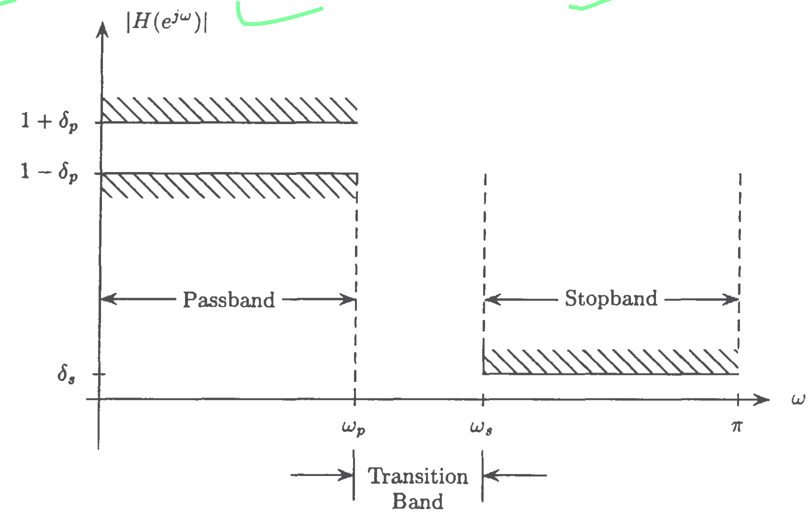



**Fig. 9-1.** Filter specifications for a low-pass filter.

are often given in decibels (dB) as follows:

$$\alpha_p = -20\log(1-\delta_p)$$

and

$$\alpha_s = -20\log(\delta_s)$$

The interval $[\omega_p, \omega_s]$ is called the *transition band*.

Once the filter specifications have been defined, the next step is to design a filter that meets these specifications.

## 9.3 FIR FILTER DESIGN

The frequency response of an $N$th-order causal FIR filter is

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

and the design of an FIR filter involves finding the coefficients $h(n)$ that result in a frequency response that satisfies a given set of filter specifications. FIR filters have two important advantages over IIR filters. First, they are guaranteed to be stable, even after the filter coefficients have been quantized. Second, they may be easily constrained to have (generalized) linear phase. Because FIR filters are generally designed to have linear phase, in the following we consider the design of linear phase FIR filters.

### *9.3.1 Linear Phase FIR Design Using Windows*

Let $h_d(n)$ be the unit sample response of an ideal frequency selective filter with linear phase,

$$H_d(e^{j\omega}) = A(e^{j\omega})e^{-j(\alpha\omega-\beta)}$$

Because $h_d(n)$ will generally be infinite in length, it is necessary to find an FIR approximation to $H_d(e^{j\omega})$. With the window design method, the filter is designed by windowing the unit sample response,

$$h(n) = h_d(n)w(n)$$

where $w(n)$ is a finite-length window that is equal to zero outside the interval $0 \le n \le N$ and is symmetric about its midpoint:

$$w(n) = w(N-n)$$

The effect of the window on the frequency response may be seen from the complex convolution theorem,

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

Thus, the ideal frequency response is *smoothed* by the discrete-time Fourier transform of the window, $W(e^{j\omega})$.

There are many different types of windows that may be used in the window design method, a few of which are listed in Table 9-1.

How well the frequency response of a filter designed with the window design method approximates a desired response, $H_d(e^{j\omega})$, is determined by two factors (see Fig. 9-2):

1. The width of the main lobe of $W(e^{j\omega})$.
2. The peak side-lobe amplitude of $W(e^{j\omega})$.

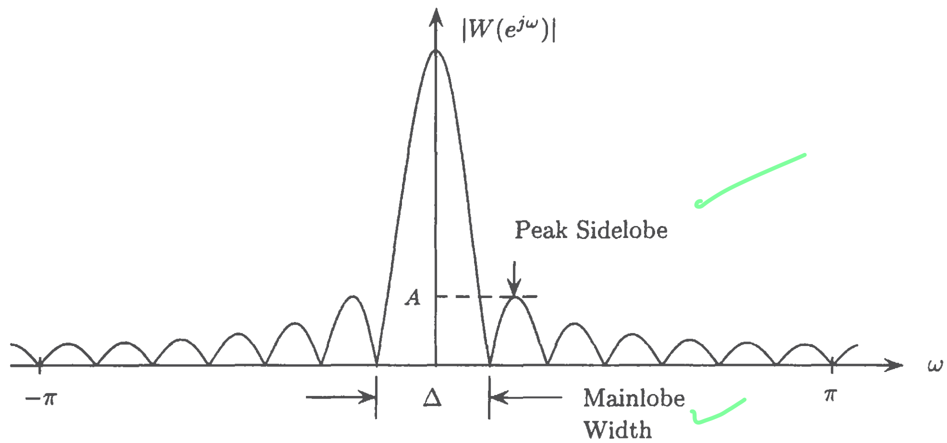



**Fig. 9-2.** The DTFT of a typical window, which is characterized by the width of its main lobe, $\Delta$, and the peak amplitude of its side lobes, $A$, relative to the amplitude of $W(e^{j\omega})$ at $\omega = 0$.

Ideally, the main-lobe width should be narrow, and the side-lobe amplitude should be small. However, for a fixed-length window, these cannot be minimized independently. Some general properties of windows are as follows:

1. As the length $N$ of the window increases, the width of the main lobe decreases, which results in a decrease in the transition width between passbands and stopbands. This relationship is given approximately by

$$N\Delta f = c \tag{9.1}$$

where $\Delta f$ is the transition width, and $c$ is a parameter that depends on the window.

2. The peak side-lobe amplitude of the window is determined by the shape of the window, and it is essentially independent of the window length.
3. If the window shape is changed to decrease the side-lobe amplitude, the width of the main lobe will generally increase.

Listed in Table 9.2 are the side-lobe amplitudes of several windows along with the approximate transition width and stopband attenuation that results when the given window is used to design an $N$th-order low-pass filter.

**Table 9-1 Some Common Windows**

| | |
|---|---|
| Rectangular | $w(n) = \begin{cases} 1 & 0 \le n \le N \\ 0 & \text{else} \end{cases}$ |
| Hanning[1] | $w(n) = \begin{cases} 0.5 - 0.5\cos\left(\frac{2\pi n}{N}\right) & 0 \le n \le N \\ 0 & \text{else} \end{cases}$ |
| Hamming | $w(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N}\right) & 0 \le n \le N \\ 0 & \text{else} \end{cases}$ |
| Blackman | $w(n) = \begin{cases} 0.42 - 0.5\cos\left(\frac{2\pi n}{N}\right) + 0.08\cos\left(\frac{4\pi n}{N}\right) & 0 \le n \le N \\ 0 & \text{else} \end{cases}$ |

[1] In the literature, this window is also called a Hann window or a von Hann window.

**Table 9-2 The Peak Side-Lobe Amplitude of Some Common Windows and the Approximate Transition Width and Stopband Attenuation of an $N$th-Order Low-Pass Filter Designed Using the Given Window.**

| Window | Side-Lobe Amplitude (dB) | Transition Width ($\Delta f$) | Stopband Attenuation (dB) |
|---|---|---|---|
| Rectangular | −13 | $0.9/N$ | −21 |
| Hanning | −31 | $3.1/N$ | −44 |
| Hamming | −41 | $3.3/N$ | −53 |
| Blackman | −57 | $5.5/N$ | −74 |

**EXAMPLE 9.3.1** Suppose that we would like to design an FIR linear phase low-pass filter according to the following specifications:

$$0.99 \le |H(e^{j\omega})| \le 1.01 \qquad 0 \le |\omega| \le 0.19\pi$$
$$|H(e^{j\omega})| \le 0.01 \qquad 0.21\pi \le |\omega| \le \pi$$

For a stopband attenuation of $20\log(0.01) = -40$ dB, we may use a Hanning window. Although we could also use a Hamming or a Blackman window, these windows would overdesign the filter and produce a larger stopband attenuation at the expense of an increase in the transition width. Because the specification calls for a transition width of $\Delta\omega = \omega_s - \omega_p = 0.02\pi$, or $\Delta f = 0.01$, with

$$N\,\Delta f = 3.1$$

for a Hanning window (see Table 9.2), an estimate of the required filter order is

$$N = \frac{3.1}{\Delta f} = 310$$

The last step is to find the unit sample response of the ideal low-pass filter that is to be windowed. With a cutoff frequency of $\omega_c = (\omega_s + \omega_p)/2 = 0.2\pi$, and a delay of $\alpha = N/2 = 155$, the unit sample response is

$$h_d(n) = \frac{\sin[0.2\pi(n-155)]}{(n-155)\pi}$$

In addition to the windows listed in Table 9-1, Kaiser developed a *family* of windows that are defined by

$$w(n) = \frac{I_0[\beta(1-[(n-\alpha)/\alpha]^2)^{1/2}]}{I_0(\beta)} \qquad 0 \le n \le N$$

where $\alpha = N/2$, and $I_0(\cdot)$ is a zeroth-order modified Bessel function of the first kind, which may be easily generated using the power series expansion

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

The parameter $\beta$ determines the shape of the window and thus controls the trade-off between main-lobe width and side-lobe amplitude. A *Kaiser window* is nearly optimum in the sense of having the most energy in its main lobe for a given side-lobe amplitude. Table 9-3 illustrates the effect of changing the parameter $\beta$.

There are two empirically derived relationships for the Kaiser window that facilitate the use of these windows to design FIR filters. The first relates the stopband ripple of a low-pass filter, $\alpha_s = -20\log(\delta_s)$, to the parameter $\beta$,

$$\beta = \begin{cases} 0.1102(\alpha_s - 8.7) & \alpha_s > 50 \\ 0.5842(\alpha_s - 21)^{0.4} + 0.07886(\alpha_s - 21) & 21 \le \alpha_s \le 50 \\ 0.0 & \alpha_s < 21 \end{cases}$$

**Table 9-3 Characteristics of the Kaiser Window as a Function of $\beta$**

| Parameter $\beta$ | Side Lobe (dB) | Transition Width ($N\Delta f$) | Stopband Attenuation (dB) |
|---|---|---|---|
| 2.0 | −19 | 1.5 | −29 |
| 3.0 | −24 | 2.0 | −37 |
| 4.0 | −30 | 2.6 | −45 |
| 5.0 | −37 | 3.2 | −54 |
| 6.0 | −44 | 3.8 | −63 |
| 7.0 | −51 | 4.5 | −72 |
| 8.0 | −59 | 5.1 | −81 |
| 9.0 | −67 | 5.7 | −90 |
| 10.0 | −74 | 6.4 | −99 |

The second relates $N$ to the transition width $\Delta f$ and the stopband attenuation $\alpha_s$,

$$N = \frac{\alpha_s - 7.95}{14.36\Delta f} \qquad \alpha_s \geq 21 \tag{9.2}$$

Note that if $\alpha_s < 21$ dB, a rectangular window may be used ($\beta = 0$), and $N = 0.9/\Delta f$.

**EXAMPLE 9.3.2** Suppose that we would like to design a low-pass filter with a cutoff frequency $\omega_c = \pi/4$, a transition width $\Delta\omega = 0.02\pi$, and a stopband ripple $\delta_s = 0.01$. Because $\alpha_s = -20\log(0.01) = -40$, the Kaiser window parameter is

$$\beta = 0.5842(40 - 21)^{0.4} + 0.07886(40 - 21) = 3.4$$

With $\Delta f = \Delta\omega/2\pi = 0.01$, we have

$$N = \frac{40 - 7.95}{14.36 \cdot (0.01)} = 224$$

Therefore,

$$h(n) = h_d(n)w(n)$$

where

$$h_d(n) = \frac{\sin[(n - 112)\pi/4]}{(n - 112)\pi}$$

is the unit sample response of the ideal low-pass filter.

Although it is simple to design a filter using the window design method, there are some limitations with this method. First, it is necessary to find a closed-form expression for $h_d(n)$ (or it must be approximated using a very long DFT). Second, for a frequency selective filter, the transition widths between frequency bands, and the ripples within these bands, will be approximately the same. As a result, the window design method requires that the filter be designed to the tightest tolerances in all of the bands by selecting the smallest transition width and the smallest ripple. Finally, window design filters are not, in general, *optimum* in the sense that they do not have the smallest possible ripple for a given filter order and a given set of cutoff frequencies.

### *9.3.2 Frequency Sampling Filter Design*

Another method for FIR filter design is the frequency sampling approach. In this approach, the desired frequency response, $H_d(e^{j\omega})$, is first uniformly sampled at $N$ equally spaced points between 0 and $2\pi$:

$$H(k) = H_d\left(e^{j2\pi k/N}\right) \qquad k = 0, 1, \ldots, N-1$$

These frequency samples constitute an $N$-point DFT, whose inverse is an FIR filter of order $N-1$:

$$h(n) = \frac{1}{N} \sum_{k=0}^{N-1} H(k) e^{j2\pi nk/N} \qquad 0 \le n \le N-1$$

The relationship between $h(n)$ and $h_d(n)$ (see Chap. 3) is

$$h(n) = \sum_{k=-\infty}^{\infty} h_d(n+kN) \qquad 0 \le n \le N-1$$

Although the frequency samples match the ideal frequency response exactly, there is no control on how the samples are *interpolated* between the samples. Because filters designed with the frequency sampling method are not generally very good, this method is often modified by introducing one or more *transition samples* as illustrated in Fig. 9-3. These transition samples are optimized in an iterative manner to maximize the stopband attenuation or minimize the passband ripple.

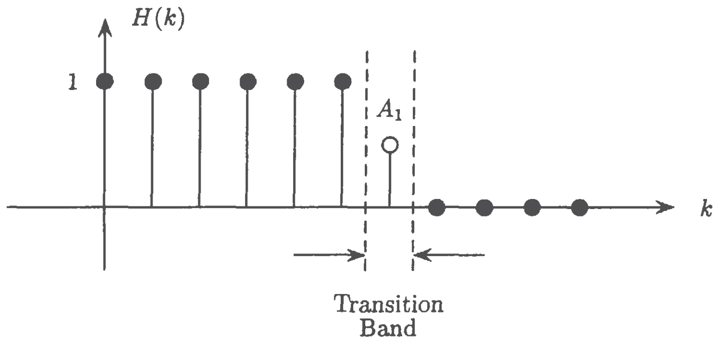



**Fig. 9-3.** Introducing a transition sample with an amplitude of $A_1$ in the frequency sampling method.

### *9.3.3 Equiripple Linear Phase Filters*

The design of an FIR low-pass filter using the window design technique is simple and generally results in a filter with relatively good performance. However, in two respects, these filters are not optimal:

1. First, the passband and stopband deviations, $\delta_p$ and $\delta_s$, are approximately equal. Although it is common to require $\delta_s$ to be much smaller than $\delta_p$, these parameters cannot be independently controlled in the window design method. Therefore, with the window design method, it is necessary to *overdesign* the filter in the passband in order to satisfy the stricter requirements in the stopband.
2. Second, for most windows, the ripple is not uniform in either the passband or the stopband and generally decreases when moving away from the transition band. Allowing the ripple to be uniformly distributed over the entire band would produce a smaller *peak ripple*.

An equiripple linear phase filter, on the other hand, is optimal in the sense that the magnitude of the ripple is minimized in all bands of interest for a given filter order, $N$. In the following discussion, we consider the design of a type I linear phase filter. The results may be easily modified to design other types of linear phase filters.

The frequency response of an FIR linear phase filter may be written as

$$H(e^{j\omega}) = A(e^{j\omega}) e^{-j\alpha\omega} \tag{9.3}$$

where the amplitude, $A(e^{j\omega})$, is a real-valued function of $\omega$. For a type I linear phase filter,

$$h(n) = h(N - n)$$

where $N$ is an even integer. The symmetry of $h(n)$ allows the frequency response to be expressed as

$$A(e^{j\omega}) = \sum_{k=0}^{L} a(k)\cos(k\omega) \tag{9.4}$$

where $L = N/2$ and

$$a(0) = h\left(\frac{N}{2}\right)$$

$$a(k) = h\left(k + \frac{N}{2}\right) \qquad k = 1, 2, \ldots, \frac{N}{2}$$

The terms $\cos(k\omega)$ may be expressed as a sum of powers of $\cos\omega$ in the form

$$\cos(k\omega) = T_k(\cos\omega)$$

where $T_k(x)$ is a $k$th-order Chebyshev polynomial [see Eq. (*9.9*)]. Therefore, Eq. (*9.4*) may be written as

$$A(e^{j\omega}) = \sum_{k=0}^{L} \alpha(k)(\cos\omega)^k$$

Thus, $A(e^{j\omega})$ is an $L$th-order polynomial in $\cos\omega$.

With $A_d(e^{j\omega})$ a desired amplitude, and $W(e^{j\omega})$ a positive weighting function, let

$$E(e^{j\omega}) = W(e^{j\omega})[A_d(e^{j\omega}) - A(e^{j\omega})]$$

be a weighted approximation error. The equiripple filter design problem thus involves finding the coefficients $a(k)$ that minimize the maximum absolute value of $E(e^{j\omega})$ over a set of frequencies, $\mathcal{F}$,

$$\min_{a(k)} \left\{ \max_{\omega\in\mathcal{F}} |E(e^{j\omega})| \right\}$$

For example, to design a low-pass filter, the set $\mathcal{F}$ will be the frequencies in the passband, $[0, \omega_p]$, and the stopband, $[\omega_s, \pi]$, as illustrated in Fig. 9-4. The transition band, $(\omega_p, \omega_s)$, is a *don't care* region, and it is not

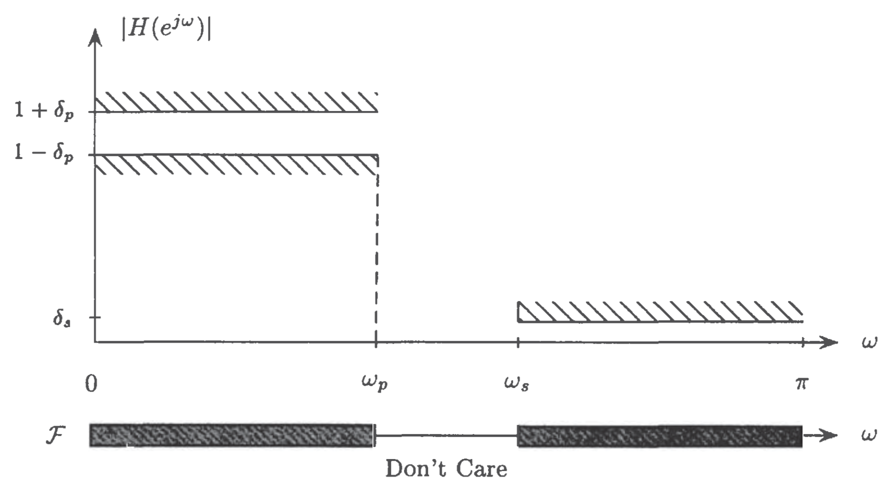



**Fig. 9-4.** The set $R$ in the equiripple filter design problem, consisting of the passband $[0, \omega_p]$ and the stopband $[\omega_p, \omega_s]$. The transition band $(\omega_p, \omega_s)$ is a don't care region.

considered in the minimization of the weighted error. The solution to this optimization problem is given in the *alternation theorem*, which is as follows:

**Alternation Theorem:** Let $\mathcal{F}$ be a union of closed subsets over the interval $[0, \pi]$. For a positive weighting function $W(e^{j\omega})$, a necessary and sufficient condition for

$$A(e^{j\omega}) = \sum_{k=0}^{L} a(k)\cos(k\omega)$$

to be the unique function that minimizes the maximum value of the weighted error $|E(e^{j\omega})|$ over the set $\mathcal{F}$ is that the $E(e^{j\omega})$ have at least $L+2$ *alternations*. That is to say, there must be at least $L+2$ *extremal frequencies*,

$$\omega_0 < \omega_1 < \cdots < \omega_{L+1}$$

over the set $\mathcal{F}$ such that

$$E(e^{j\omega_k}) = -E(e^{j\omega_{k+1}}) \qquad k = 0, 1, \ldots, L$$

and

$$|E(e^{j\omega_k})| = \max_{\omega\in\mathcal{F}} |E(e^{j\omega})| \qquad k = 0, 1, \ldots, L+1$$

Thus, the alternation theorem states that the optimum filter is equiripple. Although the alternation theorem specifies the minimum number of extremal frequencies (or ripples) that the optimum filter must have, it may have more. For example, a low-pass filter may have either $L+2$ or $L+3$ extremal frequencies. A low-pass filter with $L+3$ extrema is called an *extraripple filter*.

From the alternation theorem, it follows that

$$W(e^{j\omega_k})[A_d(e^{j\omega_k}) - A(e^{j\omega_k})] = (-1)^k\epsilon \qquad k = 0, 1, \ldots, L+1$$

where

$$\epsilon = \pm\max_{\omega\in\mathcal{F}} |E(e^{j\omega})|$$

is the maximum absolute weighted error. These equations may be written in matrix form in terms of the unknowns $a(0), \ldots, a(L)$ and $\epsilon$ as follows:

$$\begin{bmatrix} 1 & \cos(\omega_0) & \cdots & \cos(L\omega_0) & 1/W(e^{j\omega_0}) \\ 1 & \cos(\omega_1) & \cdots & \cos(L\omega_1) & -1/W(e^{j\omega_1}) \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & \cos(\omega_L) & \cdots & \cos(L\omega_L) & (-1)^L/W(e^{j\omega_L}) \\ 1 & \cos(\omega_{L+1}) & \cdots & \cos(L\omega_{L+1}) & (-1)^{L+1}/W(e^{j\omega_{L+1}}) \end{bmatrix} \begin{bmatrix} a(0) \\ a(1) \\ \vdots \\ a(L) \\ \epsilon \end{bmatrix} = \begin{bmatrix} A_d(e^{j\omega_0}) \\ A_d(e^{j\omega_1}) \\ \vdots \\ A_d(e^{j\omega_L}) \\ A_d(e^{j\omega_{L+1}}) \end{bmatrix} \qquad (9.5)$$

Given the extremal frequencies, these equations may be solved for $a(0), \ldots, a(L)$ and $\epsilon$. To find the extremal frequencies, there is an efficient iterative procedure known as the Parks-McClellan algorithm, which involves the following steps:

1. Guess an initial set of extremal frequencies.
2. Find $\epsilon$ by solving Eq. (9.5). The value of $\epsilon$ has been shown to be

$$\epsilon = \frac{\sum_{k=0}^{L+1} b(k)D(e^{j\omega_k})}{\sum_{k=0}^{L+1}(-1)^k b(k)/W(e^{j\omega_k})}$$

where
$$b(k) = \prod_{i=1, i \neq k}^{L+1} \frac{1}{\cos(\omega_k) - \cos(\omega_i)}$$

3. Evaluate the weighted error function over the set $\mathcal{F}$ by interpolating between the extremal frequencies using the Lagrange interpolation formula.
4. Select a new set of extremal frequencies by choosing the $L + 2$ frequencies for which the interpolated error function is maximum.
5. If the extremal frequencies have changed, repeat the iteration from step 2.

A design formula that may be used to estimate the equiripple filter order for a low-pass filter with a transition width $\Delta f$, passband ripple $\delta_p$, and stopband ripple $\delta_s$ is

$$N = \frac{-10\log(\delta_p \delta_s) - 13}{14.6\Delta f} \tag{9.6}$$

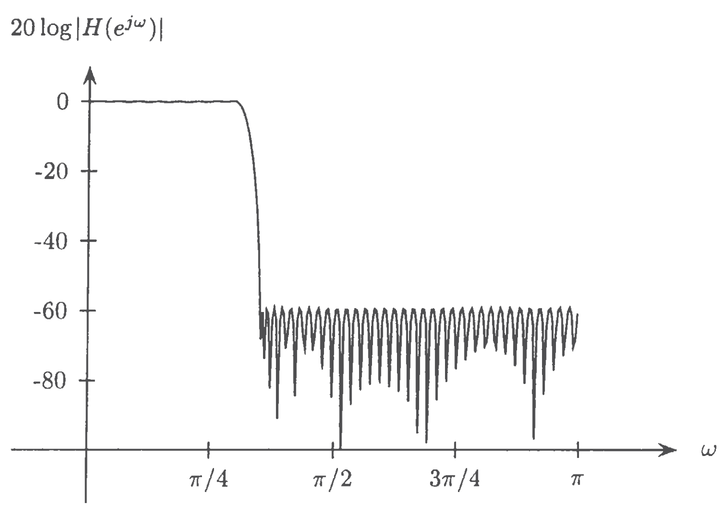
**EXAMPLE 9.3.3** Suppose that we would like to design an equiripple low-pass filter with a passband cutoff frequency $\omega_p = 0.3\pi$, a stopband cutoff frequency $\omega_s = 0.35\pi$, a passband ripple of $\delta_p = 0.01$, and a stopband ripple of $\delta_s = 0.001$. Estimating the filter using Eq. (9.6), we find

$$N = \frac{-10\log(\delta_p \delta_s) - 13}{14.6\Delta f} = 102$$

Because we want the ripple in the stopband to be 10 times smaller than the ripple in the passband, the error must be weighted using the weighting function

$$W(e^{j\omega}) = \begin{cases} 1 & 0 \leq |\omega| \leq 0.3\pi \\ 10 & 0.35\pi \leq |\omega| \leq \pi \end{cases}$$

Using the Parks-McClellan algorithm to design the filter, we obtain a filter with the frequency response magnitude shown below.




## 9.4 IIR FILTER DESIGN

There are two general approaches used to design IIR digital filters. The most common is to design an analog IIR filter and then map it into an equivalent digital filter because the art of analog filter design is highly advanced. Therefore, it is prudent to consider optimal ways for mapping these filters into the discrete-time domain. Furthermore, because there are powerful design procedures that facilitate the design of analog filters, this approach