

80509 LINEAR DIGITAL FILTERING I

PART III: Design of Linear-Phase FIR Filters

- 1) Different types of linear-phase FIR filters and their characteristics as well as their use in filtering.
 - 2) Several design techniques.
- What to read for the examination ?:
- 1) Different types of linear-phase FIR filters: how to express their frequency response in a simple form; their phase and group delays as well as their implementation using the coefficient symmetry; characteristics of the impulse responses.
 - 2) Basic idea of designing FIR filters using windowing; basic idea of using the Remez (MPR) algorithm for designing FIR filters in the minimax sense. Advantages and disadvantages of these design techniques; Characteristics of the resulting filters.
 - 3) Pages 1–162; 193–221; and 234–250 (as well as the appendices).

DESIGN OF FINITE IMPULSE RESPONSE (FIR) FILTERS

1. Why FIR filters?
2. Conditions for linear phase
3. Design of linear-phase FIR filters by windowing
4. ●Design of linear-phase FIR filters in the least-mean-square sense
5. ●Design of maximally-flat linear-phase FIR filters
6. ●Some simple linear-phase FIR filter designs
7. Design of linear-phase FIR filters in the minimax sense
8. ●Design of minimum-phase FIR filters
9. ●Design of FIR filters with constraints in the time or frequency domains
10. ●Design of linear-phase FIR filters using periodic subfilters as basic building blocks
11. ●●Design of linear-phase FIR filters using identical subfilters as basic building blocks
12. ●●Design of linear-phase FIR filters using multirate and complementary filtering
 - To be considered in the course “Digital Linear Filtering II”.
 - To be considered in the course “System Level DSP Algorithms”.

SECTION 1: WHY FIR FILTERS?

The main advantages of the FIR filter designs over their IIR equivalents are the following:

1. FIR filters with exactly linear phase can be easily designed.
2. There exist computationally efficient realizations for implementing FIR filters. These include both nonrecursive and recursive realizations.
3. FIR filters realized nonrecursively are inherently stable and free of limit cycle oscillations when implemented on a finite-wordlength digital system.
4. Excellent design methods are available for various kinds of FIR filters with **arbitrary** specifications.
5. The output noise due to multiplication roundoff errors in an FIR filter is usually very low and the sensitivity to variations in the filter coefficients is also low.

DISADVANTAGES

- FIR filter designs require, especially in applications demanding narrow transition bands, considerably more arithmetic operations and hardware components, such as multipliers, adders, and delay elements than do comparable IIR filters.
- As the transition bandwidth is made narrower, the filter order, and correspondingly the arithmetic complexity, increases inversely proportionally to this width.
- This makes the implementation of narrow transition band FIR filters very costly.
- The cost of implementation of an FIR filter can, however, be reduced by using multiplier-efficient realizations, fast convolution algorithms, and multirate filtering.
 - These topics will be considered in the courses “Digital Linear Filtering II” and “System Level DSP Algorithms”.

EXAMPLE

- $\omega_p = 0.4\pi$, $\delta_p = 0.01$ ($A_p = 0.17$ dB), $\delta_s = 0.0001$ ($A_s = 80$ dB).

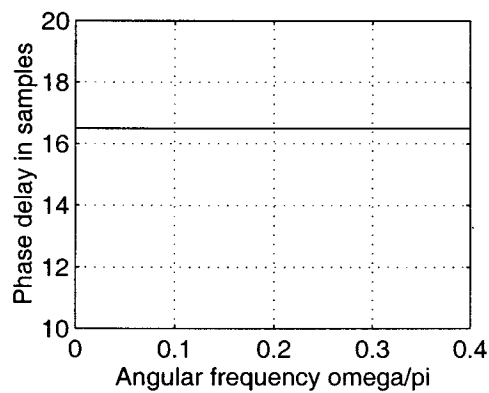
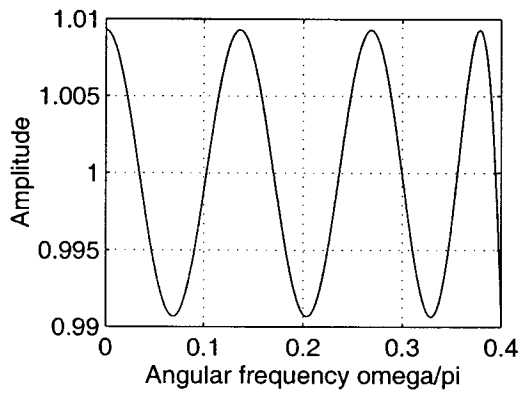
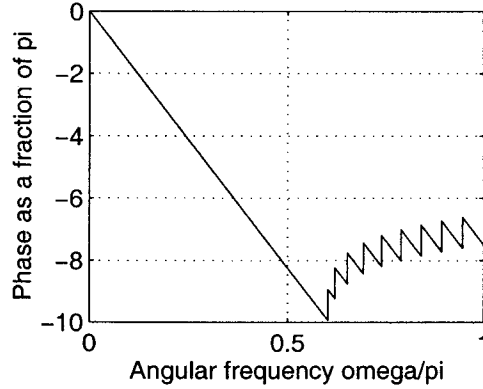
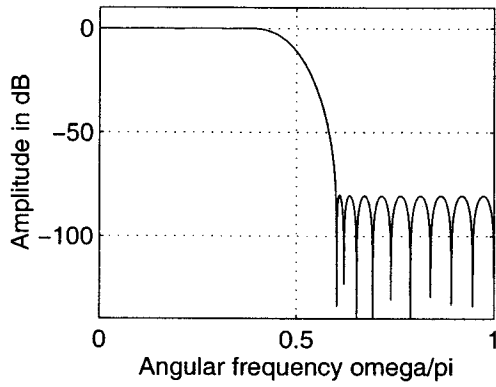
Case I: $\omega_s = 0.6\pi$, Elliptic filter: order 7, Linear-phase FIR filter: order 33

Case II: $\omega_s = 0.402\pi$, Elliptic filter: order 18, Linear-phase FIR filter: order 3138

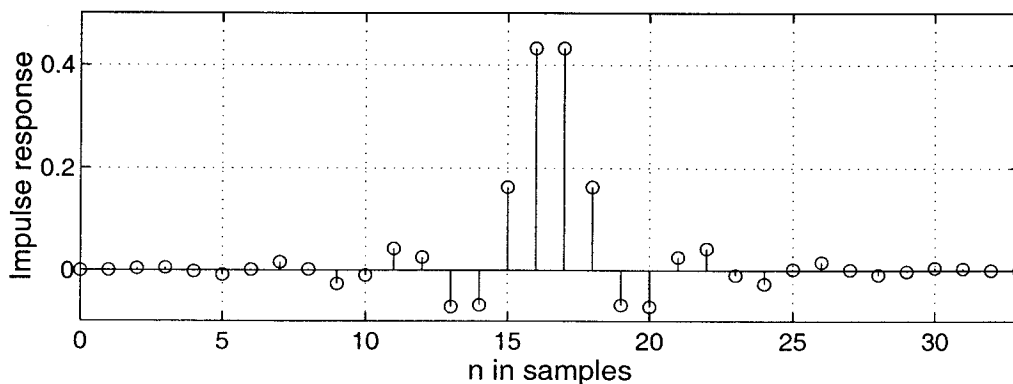
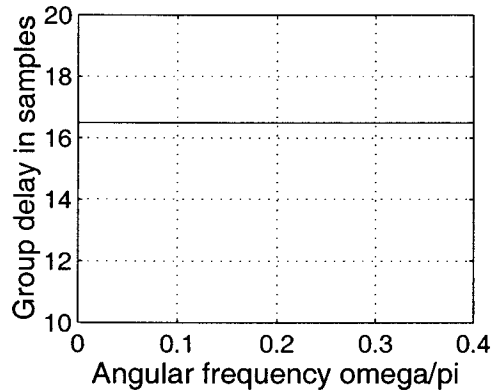
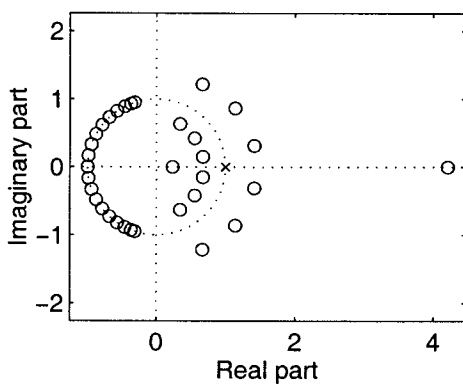
- **In Case I**, the linear-phase FIR filter is very attractive to implement compared to its IIR counterpart:
 - For the signal processor implementation, the code length is shorter.
 - For the hardware or VLSI implementation, only 17 multipliers are required if the coefficient symmetry is exploited (to be considered later).
 - The finite wordlength effects are much milder.
 - The phase is exactly linear.

- The above two desired properties make the FIR filter more attractive compared to the IIR implementation that requires 7 multipliers when implemented as a parallel connection of two allpass filters (lattice wave digital filters) to be considered in the course “Digital Linear Filtering II”.
- For the optimally scaled cascade-form structure using first- and second-order direct-form II blocks (to be considered later), the elliptic filter requires 19 or 15 multipliers, depending on the implementation.
- The next two transparencies show the characteristics of the FIR and IIR filter designs in Case I (the elliptic filter has been designed such that the passband criteria are just met).

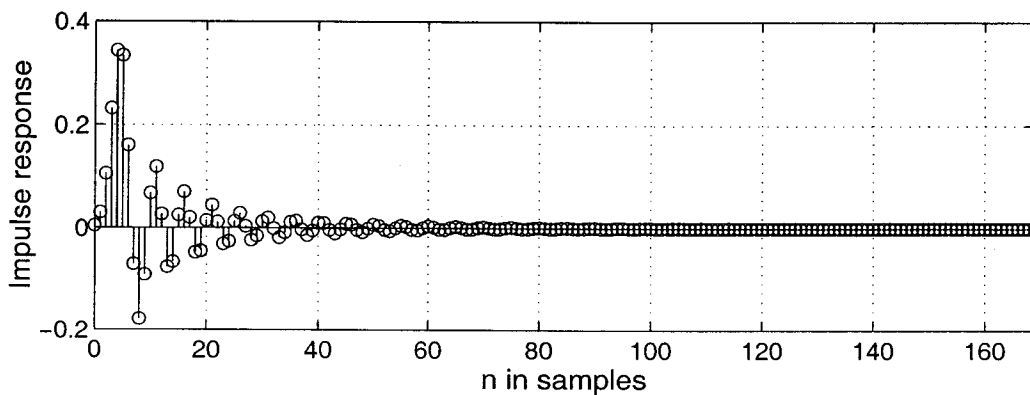
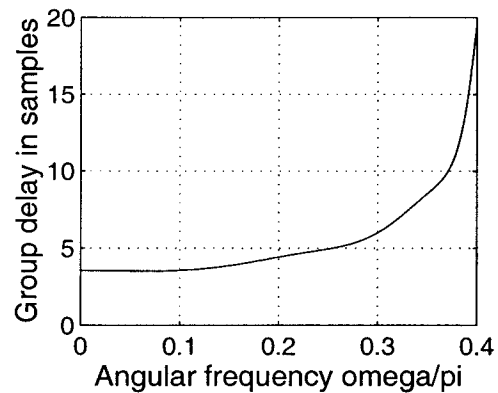
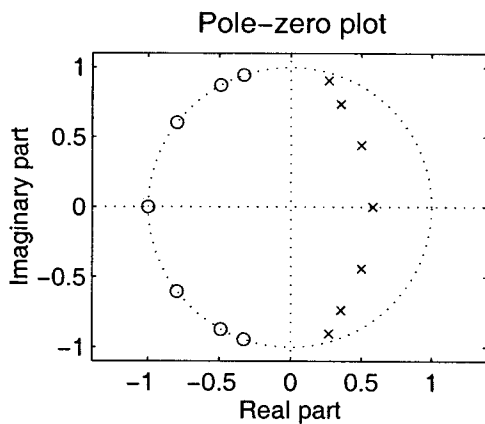
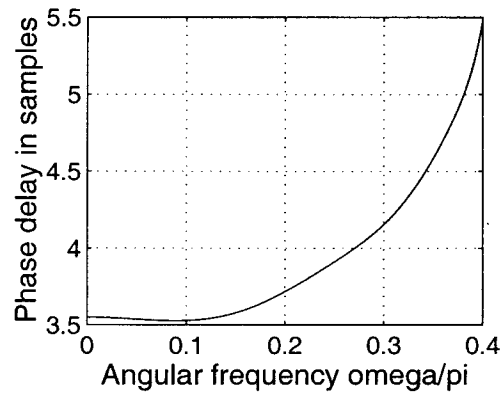
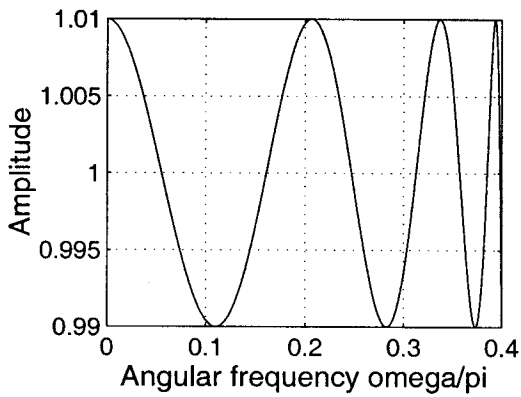
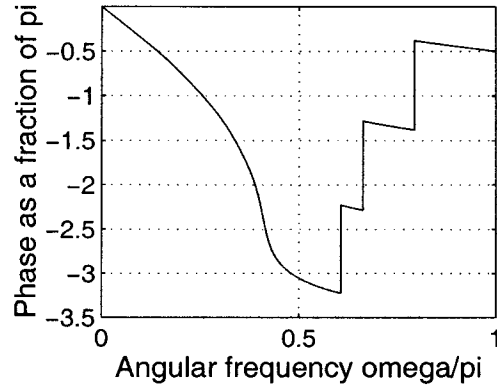
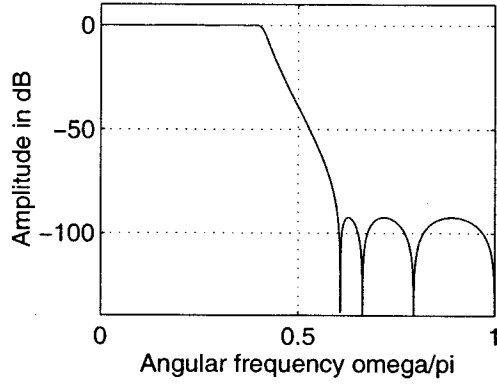
Characteristics for the linear-phase FIR filter of order 33 meeting the Case I criteria



Zero plot

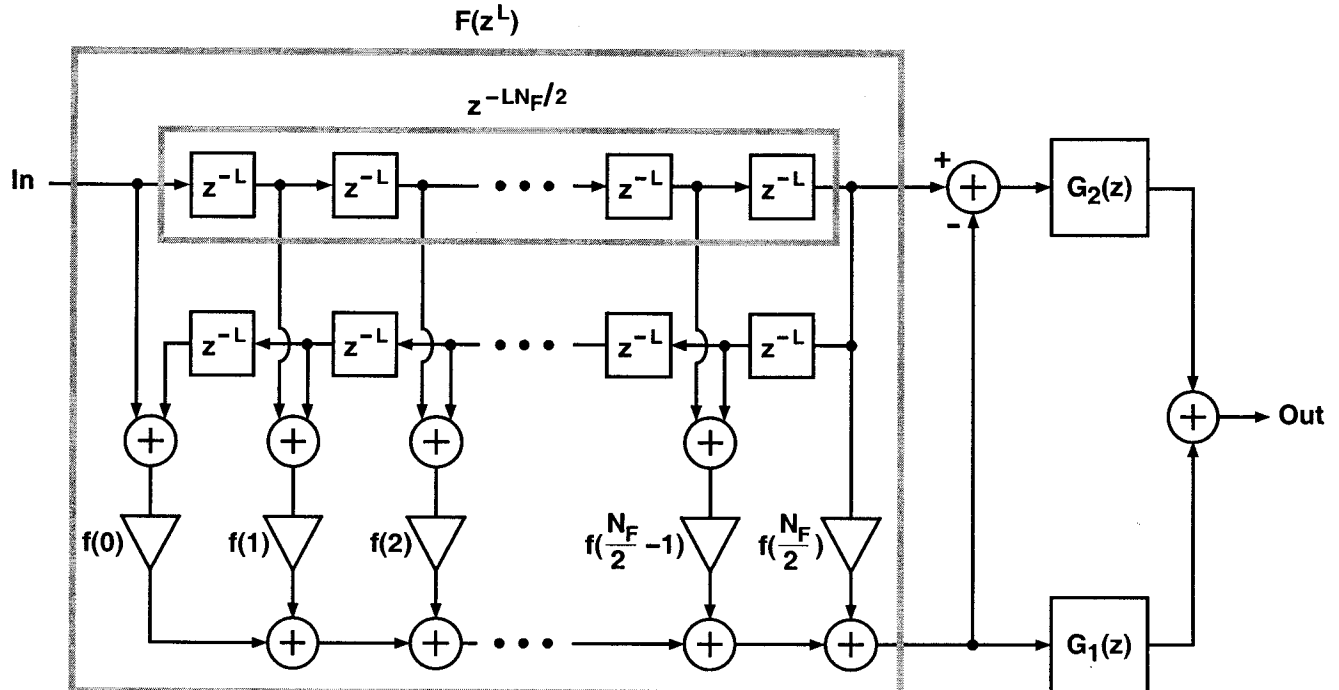


Characteristics for the elliptic filter of order 7 meeting the Case I criteria



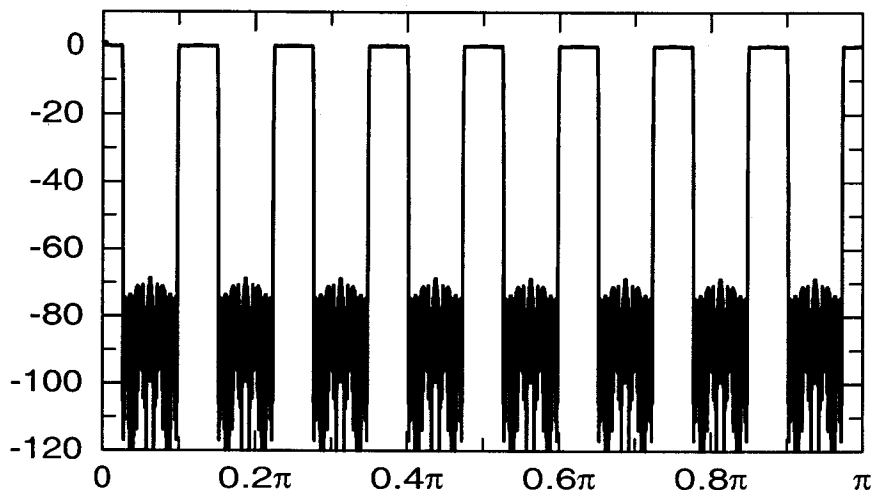
- **In Case II**, the linear-phase FIR filter is not worth implementing in the conventional manner:
 - It is on the borderline whether it is possible to design this filter with the minimum order. The order has been estimated.
 - This is why the filter characteristics for the conventional FIR filter are not shown for Case II.
- If the coefficient symmetry is exploited, 1570 multipliers are required compared to 19 multipliers required by an elliptic filter of order 19 when implemented as a parallel connection of two all-pass filters.
 - For this implementation, the order must be odd and is **attractive if the phase linearity is not needed**.
 - **The elliptic filter cannot be used if the phase linearity is of importance.**
- The following two transparencies show one alternative to reduce the number of multipliers of a linear-phase FIR filter to meet the Case II criteria.

An approach for reducing the number of multipliers in Case II

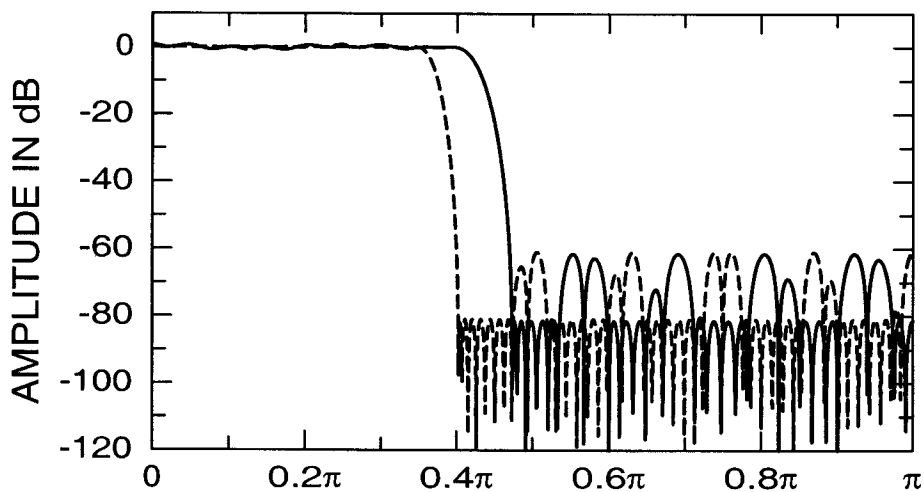


- $L = 16$, $F(z)$ of order 200, $G_1(z)$ of order 84, and $G_2(z)$ of order 122
- When exploiting the coefficient symmetry, the number of multipliers reduces from 1570 to 206 at the expense of an increase in the filter order (number of delays) from 3138 to 3322.
- This approach will be considered in more details in the course “Digital Linear Filtering II”.

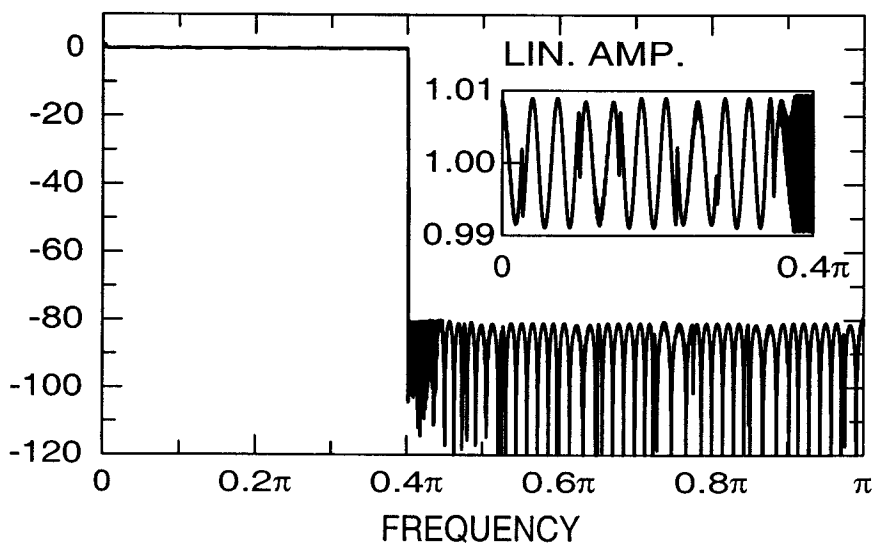
Frequency responses: (a) for $F(z^L)$; solid and dashed lines in (b) for $G_1(z)$ and $G_2(z)$; (c) for overall filter



(a)



(b)



(c)

SECTION 2: CONDITIONS FOR LINEAR-PHASE

- The transfer function of an FIR filter of order N is

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

- The corresponding frequency response is given by

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

- Here, N is the **order** of the filter that is equal to the highest power of z^{-1} , that is, the number of delays.
- The number of impulse response coefficient values $h[0], h[1], \dots, h[N]$ is $N + 1$. This is called the **length** of the filter.
- **Exact linearity** of phase is achieved if we can write $H(e^{j\omega})$ in terms of the **phase term** $\phi(\omega)$

and the **zero-phase frequency response** $H(\omega)$ as

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$\phi(\omega) = \alpha\omega + \beta.$$

- $H(\omega)$ is what is left when the phase term is taken away.
- As we shall see later, in all the linear phase cases to be considered, $H(\omega)$ is a real and continuous function of ω that can take on both positive and negative values.
- After introducing the four basic linear-phase cases, the relations between the amplitude and phase responses and the zero-phase frequency response and the phase term will be given.

FOUR LINEAR-PHASE TYPES

- The four linear-phase cases are the following:

Type I: N is even and $h[N - n] = h[n]$ for all n

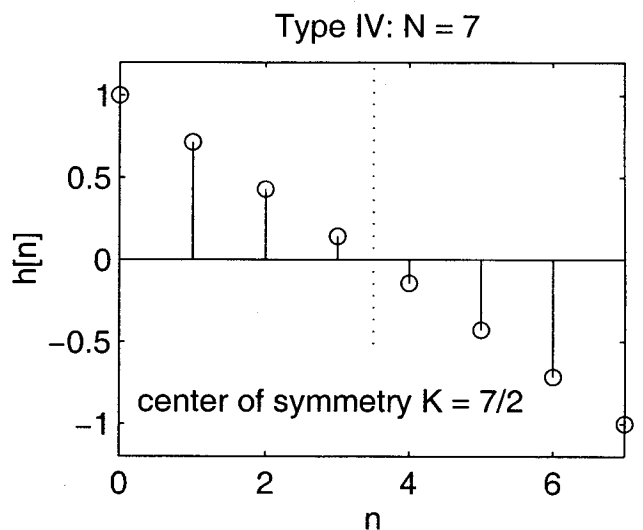
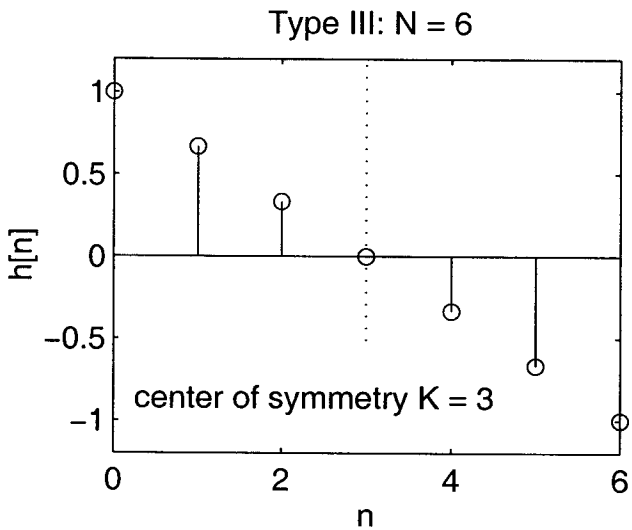
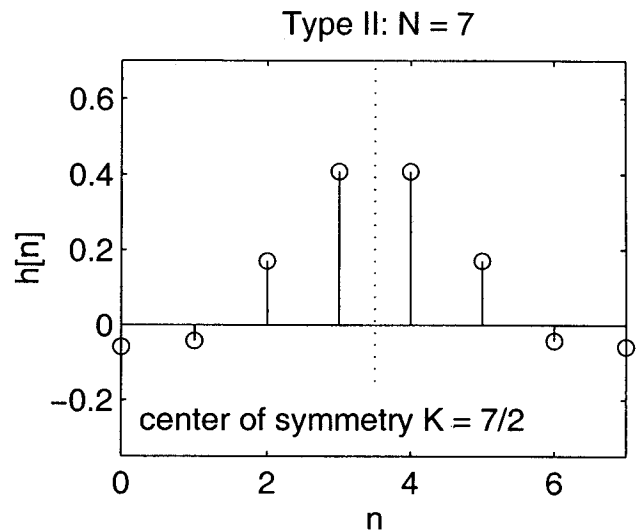
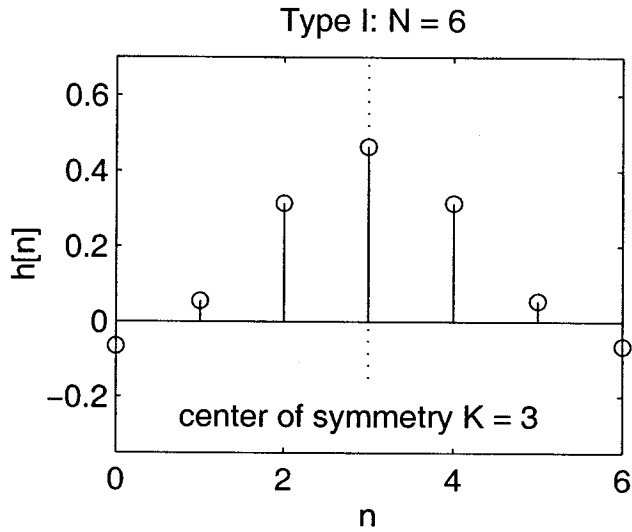
Type II: N is odd and $h[N - n] = h[n]$ for all n

Type III: N is even and $h[N - n] = -h[n]$ for all n

Type IV: N is odd and $h[N - n] = -h[n]$ for all n

- The following transparency gives typical (short) impulse responses for these four linear-phase types.

TYPICAL IMPULSE RESPONSES FOR THE FOUR LINEAR-PHASE FIR FILTER TYPES



CHARACTERISTICS OF THE IMPULSE RESPONSES FOR LINEAR-PHASE FILTERS

- As seen from the previous transparency, the impulse responses for the four different linear-phase cases are characterized by the following facts:

Type I: There exists a **lonely center impulse response sample** at the **center of symmetry** $n = K = N/2$ since N is **even**.

- The other impulse response values are related via $h[K - k] = h[K + k]$ for $k = 1, 2, \dots, N/2$, that is, $h[0] = h[N]$, $h[1] = h[N - 1]$, \dots , $h[K - 1] = h[N/2 - 1] = h[K + 1] = h[N/2 + 1]$.
- Hence, the impulse response is **symmetric** around $n = K = N/2$.
- Because of the symmetry, the number of distinct impulse response values is $1 + N/2$.

Type II: There exists **no lonely impulse response sample** at the **center of symmetry** $n = K = N/2$ since N is **odd**.

- The impulse response values are related via $h[K - (n + 1/2)] = h[K + (n + 1/2)]$ for $k = 0, 1, \dots, (N - 1)/2$, that is, $h[0] = h[N]$, $h[1] = h[N - 1]$, \dots , $h[K - 1/2] = h[(N - 1)/2] = h[K + 1/2] = h[(N + 1)/2]$.
- Hence, the impulse response is **symmetric** around $n = K = N/2$.
- The number of distinct impulse response values is $(N + 1)/2$.

Type III: There exists a **lonely center impulse response sample** at the **center of symmetry** $n = K = N/2$ since N is **even**.

- The condition $h[N - n] = -h[n]$ implies that $h[N/2] = -h[N - N/2] = -h[N]$. This means that the value of $h[K] = h[N/2]$ is restricted to be zero.
- The other impulse response values are related via $h[K - k] = -h[K + k]$ for $k = 1, 2, \dots, N/2$, that is, $h[0] = -h[N]$, $h[1] = -h[N - 1]$, \dots , $h[K - 1] = h[N/2 - 1] = -h[K + 1] = -h[N/2 + 1]$.

- Hence, the impulse response is **antisymmetric** around $n = K = N/2$.
- The number of distinct impulse response values is $N/2$ ($h[N/2] = 0$).

Type IV: There exists **no lonely impulse response sample** at the **center of symmetry** $n = K = N/2$ since N is odd.

- The impulse response values are related via $h[K - (n + 1/2)] = -h[K + (n + 1/2)]$ for $k = 0, 1, \dots, (N - 1)/2$, that is, $h[0] = -h[N]$, $h[1] = -h[N - 1]$, \dots , $h[K - 1/2] = h[(N - 1)/2] = -h[K + 1/2] = -h[(N + 1)/2]$
- Hence, the impulse response is **symmetric** around $n = K = N/2$.
- The number of distinct impulse response values is $(N + 1)/2$.

Introductory Example for Type I: $N = 6$

- Since $h[6] = h[0]$, $h[5] = h[1]$, and $h[4] = h[2]$, the transfer function is expressible as

$$\begin{aligned} H(z) &= \sum_{n=0}^6 h[n]z^{-n} \\ &= h[0](1 + z^{-6}) + h[1](z^{-1} + z^{-5}) \\ &\quad + h[2](z^{-2} + z^{-4}) + h[3]z^{-3} \\ &= z^{-3}\{h[0](z^3 + z^{-3}) + h[1](z^2 + z^{-2}) \\ &\quad + h[2](z^1 + z^{-1}) + h[3]\}. \end{aligned}$$

- By using the substitution $z = e^{j\omega}$ and exploiting the identity $(e^{jk\omega} + e^{-jk\omega}) = 2\cos(k\omega)$, the frequency response can be written as

$$\begin{aligned} H(e^{j\omega}) &= e^{-j3\omega}\{h[0](e^{j3\omega} + e^{-j3\omega}) + h[1](e^{j2\omega} + e^{-j2\omega}) \\ &\quad h[2](e^{j\omega} + e^{-j\omega}) + h[3]\} \\ &= e^{-j3\omega}\{2h[0]\cos(3\omega) + 2h[1]\cos(2\omega) \\ &\quad + 2h[2]\cos(\omega) + h[3]\}. \end{aligned}$$

- This $H(e^{j\omega})$ can be expressed as

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where the **zero-phase frequency response** $H(\omega)$ (the phase term is kicked out) and the **phase term** $\phi(\omega)$ are given by

$$H(\omega) = 2h[0]\cos(3\omega) + 2h[1]\cos(2\omega) + 2h[2]\cos(\omega) + h[3]$$

and

$$\phi(\omega) = -3\omega.$$

- $H(\omega)$ taking **both positive and negative values** and $\phi(\omega)$ for $h[n] = 1$ for $n = 0, 1, \dots, 6$ are shown in the following transparency.
- Alternatively, $H(e^{j\omega})$ is expressible as

$$H(e^{j\omega}) = |H(e^{j\omega})|e^{j\arg H(e^{j\omega})}$$

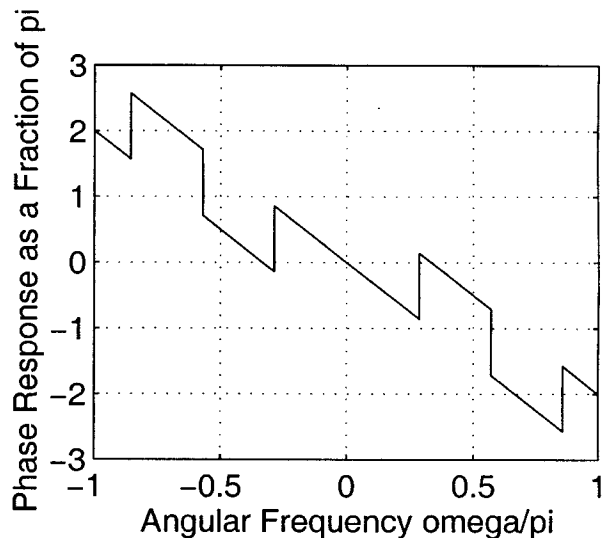
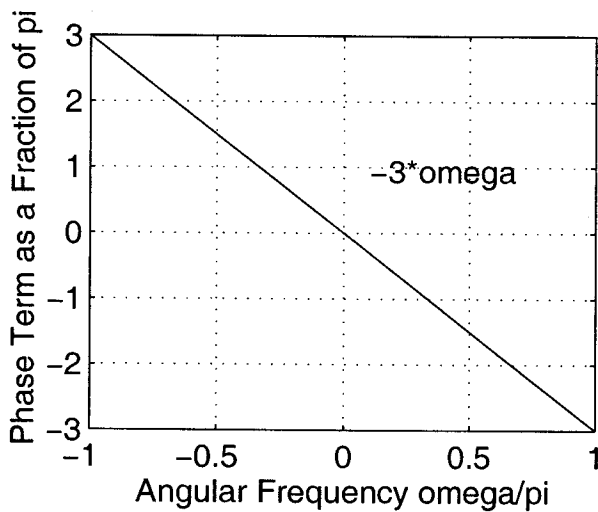
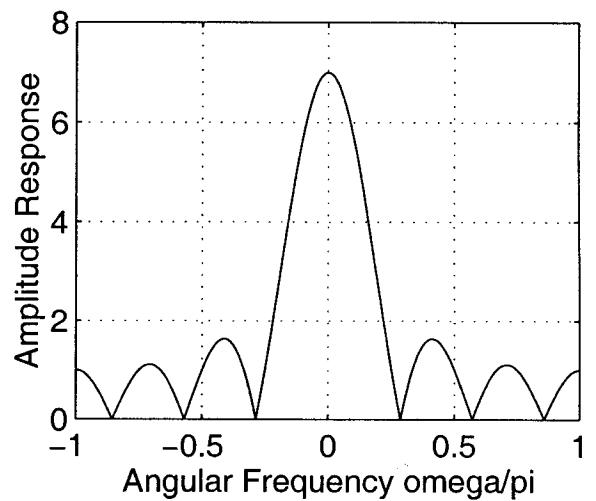
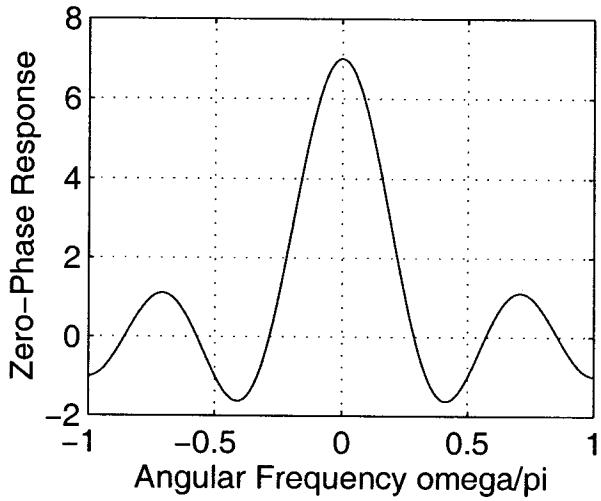
where the **amplitude response** $|H(e^{j\omega})|$ and the **phase response** $\arg H(e^{j\omega})$ are related to $H(\omega)$ and $\phi(\omega)$ via

$$|H(e^{j\omega})| = |H(\omega)|$$

$$\arg H(e^{j\omega}) = \begin{cases} \phi(\omega) & \text{for } H(\omega) \geq 0 \\ \phi(\omega) + \pi & \text{for } H(\omega) < 0 \text{ and } \omega > 0 \\ \phi(\omega) - \pi & \text{for } H(\omega) < 0 \text{ and } \omega \leq 0. \end{cases}$$

- These are also shown in the next transparency.
- Note that the above definitions make the amplitude response even around $\omega = 0$ and the phase response odd around this point, as is desired.
- Note also that equally well we can select jumps of $-\pi$ and $+\pi$ in the definition of the phase response.

Characteristics for a Type I linear phase filter with $N = 6$ and $h[n] = 1$ for $n = 0, 1, \dots, 6$



General Type I Linear-Phase Filter of Order N

- In the general case, $h[N/2]$ is a lonely impulse response sample, whereas $h[N - n] = h[n]$ for $n = 0, 1, \dots, N/2 - 1$.

- Therefore, the transfer function is expressible as

$$\begin{aligned}
 H(z) &= \sum_{m=0}^N h[m]z^{-m} \\
 &= h[N/2]z^{-N/2} + \sum_{m=0}^{N/2-1} h[m](z^{-m} + z^{-(N-m)}) \\
 &= h[N/2]z^{-N/2} + \sum_{m=0}^{N/2-1} h[m]z^{-N/2}(z^{N/2-m} + z^{-(N/2-m)}) \\
 &= z^{-N/2} \left\{ h[N/2] + \sum_{m=0}^{N/2-1} h[m](z^{N/2-m} + z^{-(N/2-m)}) \right\} \\
 &= z^{-N/2} \left\{ h[N/2] + \sum_{n=1}^{N/2} h[N/2 - n](z^n + z^{-n}) \right\}.
 \end{aligned}$$

- The last expression has been generated by using the substitution $m = N/2 - n$ in the summation.

- By exploiting the identity $(e^{jn\omega} + e^{-jn\omega}) = 2 \cos(n\omega)$, the frequency response takes the following form:

$$\begin{aligned} H(e^{j\omega}) &= e^{-jN\omega/2} \left\{ h[N/2] + \sum_{n=1}^{N/2} h[N/2 - n] (e^{jn\omega} + e^{-jn\omega}) \right\} \\ &= e^{-jN\omega/2} \left\{ h[N/2] + \sum_{n=1}^{N/2} h[N/2 - n] [2 \cos(n\omega)] \right\}. \end{aligned}$$

- This $H(e^{j\omega})$ can be expressed in the following desired form:

$$H(e^{j\omega}) = H(\omega) e^{j\phi(\omega)},$$

where

$$H(\omega) = h[N/2] + \sum_{n=1}^{N/2} h[N/2 - n] [2 \cos(n\omega)]$$

and

$$\phi(\omega) = -N\omega/2.$$

Introductory Example for Type II: $N = 7$

- Since $h[7] = h[0]$, $h[6] = h[1]$, $h[5] = h[2]$, and $h[4] = h[3]$, the transfer function is expressible as

$$\begin{aligned}
 H(z) &= \sum_{n=0}^7 h[n]z^{-n} \\
 &= h[0](1 + z^{-7}) + h[1](z^{-1} + z^{-6}) \\
 &\quad + h[2](z^{-2} + z^{-5}) + h[3](z^{-3} + z^{-4}) \\
 &= z^{-7/2} \{ h[0](z^{7/2} + z^{-7/2}) + h[1](z^{5/2} + z^{-5/2}) \\
 &\quad + h[2](z^{3/2} + z^{-3/2}) + h[3](z^{1/2} + z^{-1/2}) \}.
 \end{aligned}$$

- By exploiting the identity $(e^{jk\omega/2} + e^{-jk\omega/2}) = 2 \cos(k\omega/2)$, the corresponding frequency response can be written as

$$\begin{aligned}
 H(e^{j\omega}) &= e^{-j7\omega/2} \{ h[0](e^{j7\omega/2} + e^{-j7\omega/2}) \\
 &\quad + h[1](e^{j5\omega/2} + e^{-j5\omega/2}) + h[2](e^{j3\omega/2} + e^{-j3\omega/2}) \\
 &\quad + h[3](e^{j\omega/2} + e^{-j\omega/2}) \} \\
 &= e^{-j7\omega/2} \{ 2h[0] \cos(7\omega/2) + 2h[1] \cos(5\omega/2) \\
 &\quad + 2h[2] \cos(3\omega/2) + 2h[3] \cos(\omega/2) \}.
 \end{aligned}$$

- Alternatively, $H(e^{j\omega})$ can be expressed in the following desired form:

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$\begin{aligned} H(\omega) = & 2h[0] \cos(7\omega/2) + 2h[1] \cos(5\omega/2) \\ & + 2h[2] \cos(3\omega/2) + 2h[3] \cos(\omega/2) \end{aligned}$$

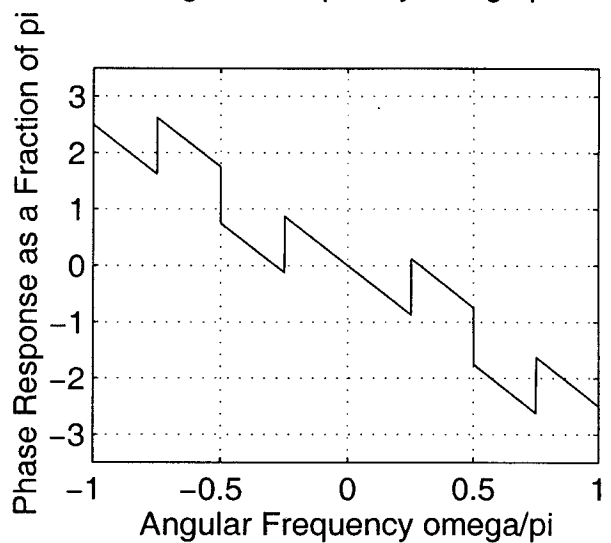
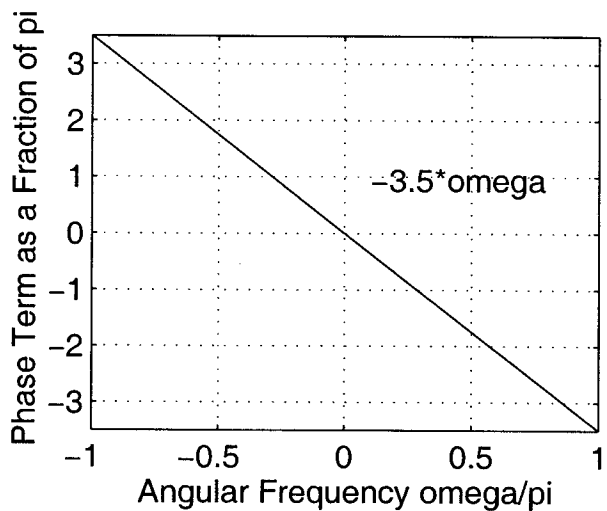
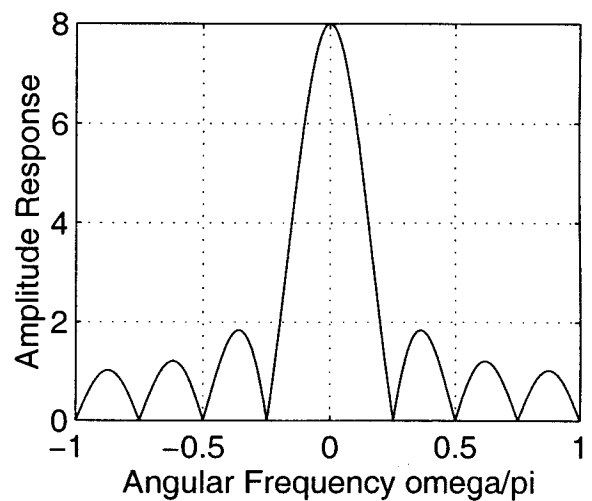
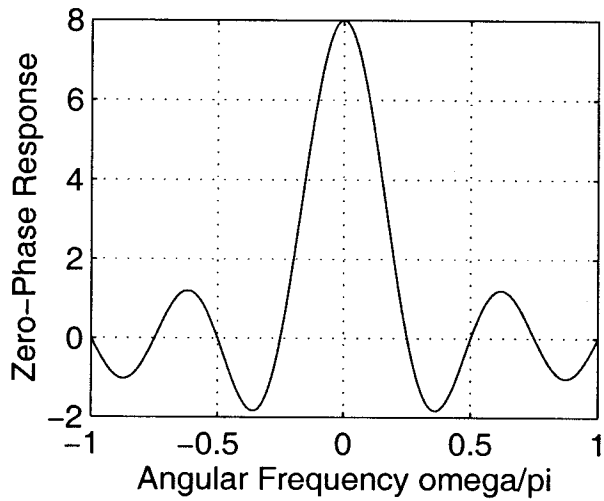
and

$$\phi(\omega) = -7\omega/2.$$

- $H(\omega)$ and $\phi(\omega)$ for $h[n] = 1$ for $n = 0, 1, \dots, 7$ are shown in the following transparency.
- It gives also the amplitude and the phase responses that are given by

$$\begin{aligned} |H(e^{j\omega})| &= |H(\omega)| \\ \arg H(e^{j\omega}) &= \begin{cases} \phi(\omega) & \text{for } H(\omega) > 0 \\ \phi(\omega) + \pi & \text{for } H(\omega) \leq 0 \text{ and } \omega > 0 \\ \phi(\omega) - \pi & \text{for } H(\omega) < 0 \text{ and } \omega \leq 0. \end{cases} \end{aligned}$$

Characteristics for a Type II linear phase filter with $N = 7$ and $h[n] = 1$ for $n = 0, 1, \dots, 7$



General Type II Linear-Phase Filter of Order N

- In the general case, $h[N - n] = h[n]$ for $n = 0, 1, \dots, (N - 1)/2$.
- Therefore, the transfer function is expressible as

$$\begin{aligned}
 H(z) &= \sum_{m=0}^N h[m]z^{-m} \\
 &= \sum_{m=0}^{(N-1)/2} h[m](z^{-m} + z^{-(N-m)}) \\
 &= \sum_{m=0}^{(N-1)/2} h[m]z^{-N/2}(z^{N/2-m} + z^{-(N/2-m)}) \\
 &= z^{-N/2} \left\{ \sum_{m=0}^{(N-1)/2} h[m](z^{N/2-m} + z^{-(N/2-m)}) \right\} \\
 &= z^{-N/2} \left\{ \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n](z^{(n+1)/2} + z^{-(n+1)/2}) \right\}.
 \end{aligned}$$

- The last expression has been generated by using the substitution $m = (N - 1)/2 - n$ in the summation.

- By exploiting the identity $(e^{j(n+1)\omega/2} + e^{-j(n+1)\omega/2}) = 2 \cos[(n+1)\omega/2]$, the frequency response can be written as

$$\begin{aligned}
 H(e^{j\omega}) &= e^{-jN\omega/2} \left\{ \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n] (e^{j(n+1)\omega/2} + e^{-j(n+1)\omega/2}) \right\} \\
 &= e^{-jN\omega/2} \left\{ \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n] [2 \cos[(n+1)\omega/2]] \right\}.
 \end{aligned}$$

- This $H(e^{j\omega})$ can be expressed in the following desired form:

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$H(\omega) = \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n] [2 \cos[(n+1)\omega/2]]$$

and

$$\phi(\omega) = -N\omega/2.$$

Introductory Example for Type III: $N = 6$

- Since $h[6] = -h[0]$, $h[5] = -h[1]$, $h[4] = -h[2]$, and $h[3] = 0$, the transfer function is expressible as

$$\begin{aligned} H(z) &= \sum_{n=0}^6 h[n]z^{-n} \\ &= h[0](1 - z^{-6}) + h[1](z^{-1} - z^{-5}) \\ &\quad + h[2](z^{-2} - z^{-4}) \\ &= z^{-3}\{h[0](z^3 - z^{-3}) + h[1](z^2 - z^{-2}) \\ &\quad + h[2](z^1 - z^{-1})\}. \end{aligned}$$

- By exploiting the identities $(e^{jk\omega} - e^{-jk\omega}) = 2j \sin(k\omega)$ and $j = e^{j\pi/2}$, the frequency response can be written as

$$\begin{aligned} H(e^{j\omega}) &= e^{-j3\omega}\{h[0](e^{j3\omega} - e^{-j3\omega}) + h[1](e^{j2\omega} - e^{-j2\omega}) \\ &\quad + h[2](e^{j\omega} - e^{-j\omega})\} \\ &= je^{-j3\omega}\{2h[0] \sin(3\omega) + 2h[1] \sin(2\omega) \\ &\quad + 2h[2] \sin(\omega)\} \\ &= e^{j[\pi/2-3\omega]}\{2h[0] \sin(3\omega) + 2h[1] \sin(2\omega) \\ &\quad + 2h[2] \sin(\omega)\}. \end{aligned}$$

- Alternatively, $H(e^{j\omega})$ can be expressed in the following desired form:

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$H(\omega) = 2h[0] \sin(3\omega) + 2h[1] \sin(\omega) + 2h[2] \sin(\omega)$$

and

$$\phi(\omega) = \pi/2 - 3\omega.$$

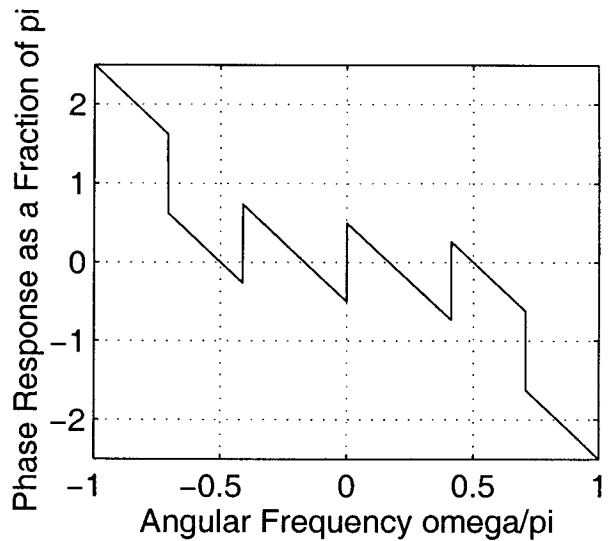
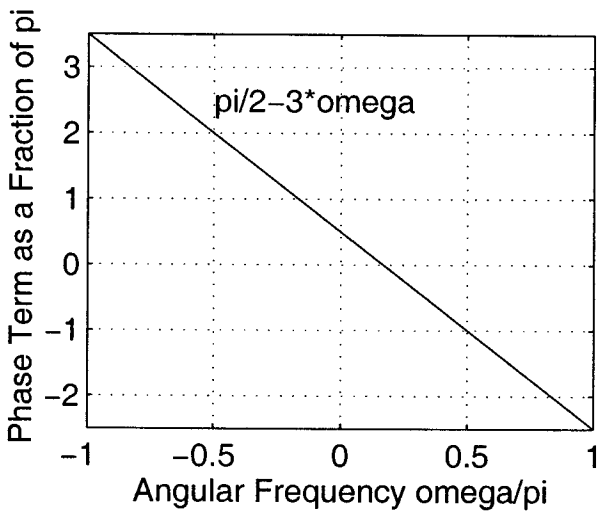
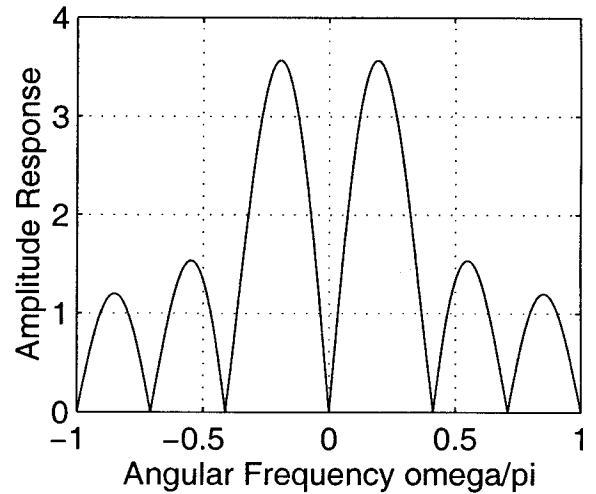
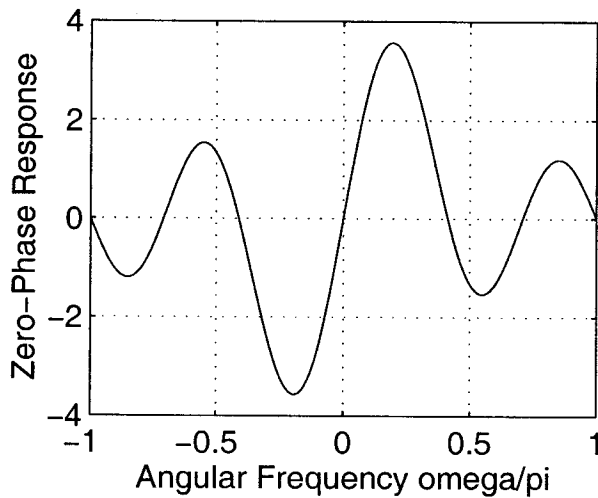
- $H(\omega)$ and $\phi(\omega)$ for $h[0] = 1$, $h[1] = 2/3$, $h[2] = 1/3$, $h[3] = 0$, $h[4] = -1/3$, $h[5] = -2/3$, and $h[6] = -1$ are shown in the next transparency.
- This transparency shows also the amplitude and the phase responses as given by

$$|H(e^{j\omega})| = |H(\omega)|$$

$$\arg H(e^{j\omega}) = \begin{cases} \phi(\omega) & \text{for } H(\omega) \geq 0 \\ \phi(\omega) + \pi & \text{for } H(\omega) < 0 \text{ and } \omega > 0 \\ \phi(\omega) - \pi & \text{for } H(\omega) < 0 \text{ and } \omega \leq 0. \end{cases}$$

- Note that for the phase response $\arg H(e^{j\omega})$ there is a jump of π at $\omega = 0$ (from $-\pi/2$ to $\pi/2$).

Characteristics for a Type III linear phase filter with $N = 6$ and $h[0] = 1, h[1] = 2/3, h[2] = 1/3, h[3] = 0, h[4] = -1/3, h[5] = -2/3,$ and $h[6] = -1$



General Type III Linear-Phase Filter of Order N

- In the general case, $h[N/2] = 0$ and $h[N - n] = -h[n]$ for $n = 0, 1, \dots, N/2 - 1$.
- Therefore, the transfer function is expressible as

$$\begin{aligned}
 H(z) &= \sum_{m=0}^N h[m]z^{-m} \\
 &= \sum_{m=0}^{N/2-1} h[m](z^{-m} - z^{-(N-m)}) \\
 &= \sum_{m=0}^{N/2-1} h[m]z^{-N/2}(z^{N/2-m} - z^{-(N/2-m)}) \\
 &= z^{-N/2} \sum_{m=0}^{N/2-1} h[m](z^{N/2-m} - z^{-(N/2-m)}) \\
 &= z^{-N/2} \left\{ \sum_{n=0}^{N/2-1} h[N/2 - 1 - n](z^{n+1} - z^{-(n+1)}) \right\}.
 \end{aligned}$$

- The last expression has been generated by using the substitution $m = N/2 - 1 - n$ in the summation.

- By exploiting the identities $(e^{j(n+1)\omega} - e^{-j(n+1)\omega}) = 2j \sin[(n+1)\omega]$ and $j = e^{j\pi/2}$, the frequency response can be written as

$$\begin{aligned}
 H(e^{j\omega}) &= \\
 &= e^{-jN\omega/2} \sum_{n=0}^{N/2-1} h[N/2 - 1 - n](e^{j(n+1)\omega} - e^{-j(n+1)\omega}) \\
 &= je^{-jN\omega/2} \left\{ \sum_{n=0}^{N/2-1} h[N/2 - 1 - n][2 \sin[(n+1)\omega]] \right\}. \\
 &= e^{j[\pi/2 - N\omega/2]} \left\{ \sum_{n=0}^{N/2-1} h[N/2 - 1 - n][2 \sin[(n+1)\omega]] \right\}.
 \end{aligned}$$

- This $H(e^{j\omega})$ can be expressed in the following desired form:

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$H(\omega) = \sum_{n=0}^{N/2-1} h[N/2 - 1 - n][2 \sin[(n+1)\omega]]$$

and

$$\phi(\omega) = \pi/2 - N\omega/2.$$

Introductory Example for Type IV: $N = 7$

- Since $h[7] = -h[0]$, $h[6] = -h[1]$, $h[5] = -h[2]$, and $h[4] = -h[3]$, the transfer function is given by

$$\begin{aligned}
 H(z) &= \sum_{n=0}^7 h[n]z^{-n} \\
 &= h[0](1 - z^{-7}) + h[1](z^{-1} - z^{-6}) \\
 &\quad + h[2](z^{-2} - z^{-5}) + h[3](z^{-3} - z^{-4}) \\
 &= z^{-7/2} \{ h[0](z^{7/2} - z^{-7/2}) + h[1](z^{5/2} - z^{-5/2}) \\
 &\quad + h[2](z^{3/2} - z^{-3/2}) + h[3](z^{1/2} - z^{-1/2}) \}.
 \end{aligned}$$

- By exploiting the identities $(e^{jk\omega/2} - e^{-jk\omega/2}) = 2j \sin(k\omega/2)$ and $j = e^{j\pi/2}$, the frequency response can be written as

$$\begin{aligned}
 H(e^{j\omega}) &= e^{-j7\omega/2} \{ h[0](e^{j7\omega/2} - e^{-j7\omega/2}) \\
 &\quad + h[1](e^{j5\omega/2} - e^{-j5\omega/2}) + h[2](e^{j3\omega/2} - e^{-j3\omega/2}) \\
 &\quad + h[3](e^{j\omega/2} - e^{-j\omega/2}) \} \\
 &= je^{-j7\omega/2} \{ 2h[0] \sin(7\omega/2) + 2h[1] \sin(5\omega/2) \\
 &\quad + h[2] \sin(3\omega/2) + h[3] \sin(\omega/2) \} \\
 &= e^{j[\pi/2 - j7\omega/2]} \{ 2h[0] \sin(7\omega/2) + 2h[1] \sin(5\omega/2) \\
 &\quad + 2h[2] \sin(3\omega/2) + 2h[3] \sin(\omega/2) \}.
 \end{aligned}$$

- This $H(e^{j\omega})$ can be expressed in the following desired form:

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$\begin{aligned} H(\omega) = & 2h[0] \sin(7\omega/2) + 2h[1] \sin(5\omega/2) \\ & + 2h[2] \sin(3\omega/2) + 2h[3] \sin(\omega/2) \end{aligned}$$

and

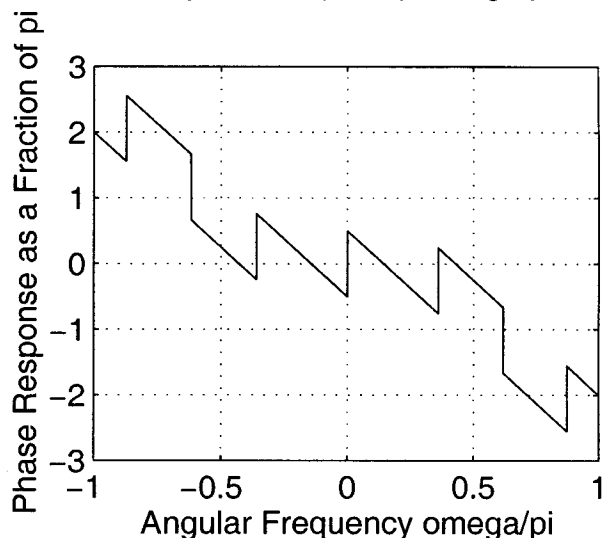
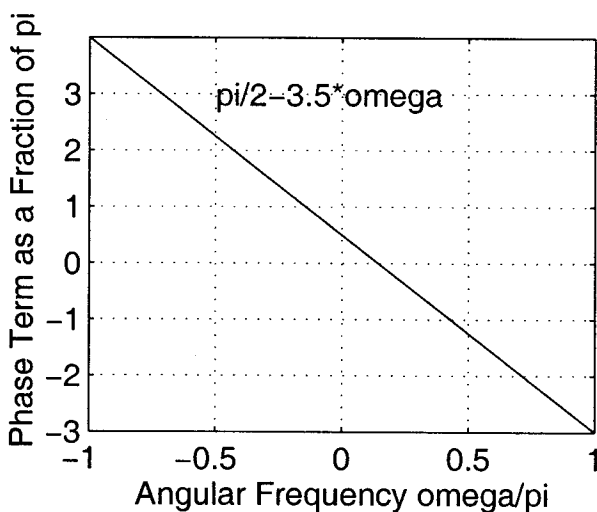
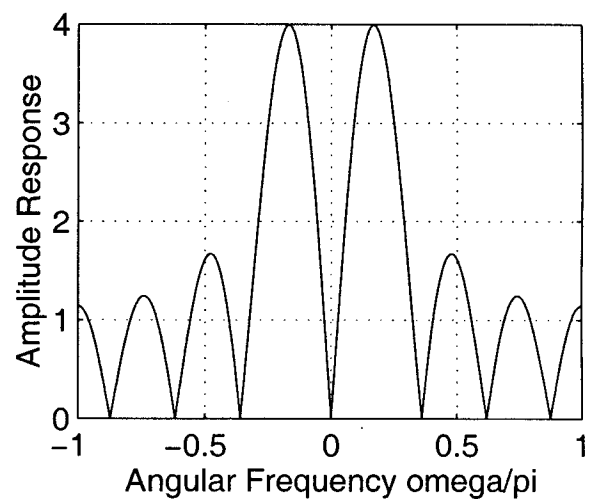
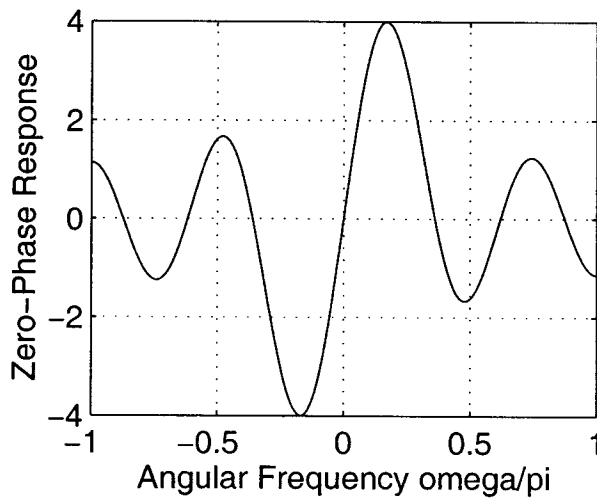
$$\phi(\omega) = \pi/2 - 7\omega/2.$$

- $H(\omega)$ and $\phi(\omega)$ for $h[0] = 1$, $h[1] = 5/7$, $h[2] = 3/7$, $h[3] = 1/7$, $h[4] = -1/7$, $h[5] = -3/7$, $h[6] = -5/7$, and $h[7] = -1$ are shown in the next transparency.
- This transparency shows also the amplitude and the phase responses as given by

$$\begin{aligned} |H(e^{j\omega})| &= |H(\omega)| \\ \arg H(e^{j\omega}) &= \begin{cases} \phi(\omega) & \text{for } H(\omega) \geq 0 \\ \phi(\omega) + \pi & \text{for } H(\omega) < 0 \text{ and } \omega > 0 \\ \phi(\omega) - \pi & \text{for } H(\omega) < 0 \text{ and } \omega \leq 0. \end{cases} \end{aligned}$$

- Note that, like for Type III, for the phase response $\arg H(e^{j\omega})$, there is a jump of π at $\omega = 0$.

Characteristics for a Type IV linear phase filter with $N = 7$ and $h[0] = 1$, $h[1] = 5/7$, $h[2] = 3/7$, $h[3] = 1/7$, $h[4] = -1/7$, $h[5] = -3/7$, $h[6] = -5/7$, and $h[7] = -1$



General Type IV Linear-Phase Filter of Order N

- In the general case, $h[N - n] = h[n]$ for $n = 0, 1, \dots, (N - 1)/2$.
- Therefore, the transfer function is expressible as

$$\begin{aligned}
 H(z) &= \sum_{m=0}^N h[m]z^{-m} \\
 &= \sum_{m=0}^{(N-1)/2} h[m](z^{-m} - z^{-(N-m)}) \\
 &= \sum_{m=0}^{(N-1)/2} h[m]z^{-N/2}(z^{N/2-m} - z^{-(N/2-m)}) \\
 &= z^{-N/2} \left\{ \sum_{m=0}^{(N-1)/2} h[m](z^{N/2-m} - z^{-(N/2-m)}) \right\} \\
 &= z^{-N/2} \left\{ \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n](z^{(n+1)/2} - z^{-(n+1)/2}) \right\}.
 \end{aligned}$$

- The last expression has been generated by using the substitution $m = (N - 1)/2 - n$ in the summation.

- By exploiting the identities $(e^{j(n+1)\omega/2} - e^{-j(n+1)\omega/2}) = 2j \sin[(n+1)\omega/2]$, the frequency response can be written as

$$\begin{aligned}
 & H(e^{j\omega}) \\
 &= e^{-jN\omega/2} \left\{ \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n] (e^{j(n+1)\omega/2} - e^{-j(n+1)\omega/2}) \right\} \\
 &= j e^{-jN\omega/2} \left\{ \sum_{n=1}^{(N-1)/2} h[(N-1)/2 - n] [2 \sin[(n+1)\omega/2]] \right\} \\
 &= e^{j[\pi/2 - jN\omega/2]} \left\{ \sum_{n=1}^{(N-1)/2} h[(N-1)/2 - n] [2 \sin[(n+1)\omega/2]] \right\}.
 \end{aligned}$$

- This $H(e^{j\omega})$ can be expressed in the following desired form:

$$H(e^{j\omega}) = H(\omega) e^{j\phi(\omega)},$$

where

$$H(\omega) = \sum_{n=1}^{(N-1)/2} h[(N-1)/2 - n] [2 \sin[(n+1)\omega/2]]$$

and

$$\phi(\omega) = \pi/2 - N\omega/2.$$

COMMON REPRESENTATION FORM FOR THE FOUR LINEAR-PHASE TYPES

- By combining the above results, $H(e^{j\omega})$ can be expressed in the following common form:

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$H(\omega) = \begin{cases} h[N/2] + \sum_{n=1}^{N/2} h[N/2 - n][2 \cos n\omega] & \text{for Type I} \\ \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n][2 \cos[(n + 1/2)\omega]] & \text{for Type II} \\ \sum_{n=0}^{N/2-1} h[N/2 - 1 - n][2 \sin[(n + 1)\omega]] & \text{for Type III} \\ \sum_{n=0}^{(N-1)/2} h[(N-1)/2 - n][2 \sin[(n + 1/2)\omega]] & \text{for Type IV.} \end{cases}$$

and

$$\phi(\omega) = \begin{cases} -N\omega/2 & \text{for Types I and II} \\ \pi/2 - N\omega/2 & \text{for Types III and IV.} \end{cases}$$

ALTERNATIVE REPRESENTATION FORM FOR $H(\omega)$

$$H(\omega) = \sum_{n=0}^M b[n] \text{trig}(\omega, n),$$

where

$$\text{trig}(\omega, n) = \begin{cases} 1 & \text{for Type I; } n = 0 \\ 2 \cos n\omega & \text{for Type I; } n > 0 \\ 2 \cos[(n + 1/2)\omega] & \text{for Type II} \\ 2 \sin(n + 1)\omega & \text{for Type III} \\ 2 \sin[(n + 1/2)\omega] & \text{for Type IV,} \end{cases}$$

$$b[n] = \begin{cases} h[N/2 - n] & \text{for Type I} \\ h[(N - 1)/2 - n] & \text{for Types II and IV} \\ h[N/2 - 1 - n] & \text{for Type III,} \end{cases}$$

and

$$M = \begin{cases} N/2 & \text{for Type I} \\ (N - 1)/2 & \text{for Type II} \\ (N - 2)/2 & \text{for Type III} \\ (N - 1)/2 & \text{for Type IV.} \end{cases}$$

- This form is useful when designing the four linear-phase filter types in the least-mean-square sense or when using linear programming.

EXPRESSION OF THE TRANSFER FUNCTION AS A CASCADE OF A COMMON TERM AND A FIXED TERM

- According to the previous considerations, the overall transfer function can be expressed as

$$H(z) = \begin{cases} h[N/2]z^{-N} + \sum_{n=0}^{(N/2)-1} h[n][z^{-n} + z^{-(N-n)}] & \text{for Type I} \\ \sum_{n=0}^{(N-1)/2} h[n][z^{-n} + z^{-(N-n)}] & \text{for Type II} \\ \sum_{n=0}^{N/2-1} h[n][z^{-n} - z^{-(N-n)}] & \text{for Type III} \\ \sum_{n=0}^{(N-1)/2} h[n][z^{-n} - z^{-(N-n)}] & \text{for Type IV.} \end{cases}$$

- What we are now interested in is to check whether these filter types have fixed zeros at $z = 1$ or at $z = -1$.
- This can be found out by substituting $z = 1$ or $z = -1$ in the above equations and checking whether $H(1)$ or $H(-1)$ is zero or not.

Type I: These filters do not necessarily have any zeros at $z = 1$ or at $z = -1$

- The $h[n]$'s can be selected such that $H(1)$ and $H(-1)$ are not equal to zero.

Type II: These filters have **at least one zero at $z = -1$.**

- Because N is odd, each of the terms $[z^{-n} + z^{-(N-n)}]$ is zero at $z = -1$ (For n even [odd], $N - n$ is odd [even]).

Type III: These filters have **at least one zero at $z = 1$ and $z = -1$.**

- $h[N/2] = 0$ is absent. Because N is odd, each of the terms $[z^{-n} - z^{-(N-n)}]$ is zero at $z = \pm 1$ (For n even [odd], $N - n$ is even [odd]).

Type IV: These filters have **at least one zero at $z = 1$.**

- Because N is odd, each of the terms $[z^{-n} + z^{-(N-n)}]$ is zero at $z = 1$ (For n even [odd], $N - n$ is odd [even]).

COMMON REPRESENTATION FORM

- Based on the above, the overall transfer function is expressible as a cascade of a fixed term $F(z)$ and an adjustable Type I term as follows:

$$H(z) = F(z)G(z),$$

where

$$F(z) = \begin{cases} 1 & \text{for Type I} \\ [1 + z^{-1}]/2 & \text{for Type II} \\ [1 - z^{-2}]/2 & \text{for Type III} \\ [1 - z^{-1}]/2 & \text{for Type IV} \end{cases}$$

and

$$G(z) = \sum_{n=0}^{2M} g[n]z^{-n}$$

with

$$g[2M - n] = g[n] \text{ for all } n$$

and

$$M = \begin{cases} N/2 & \text{for Type I} \\ (N - 1)/2 & \text{for Type II} \\ (N - 2)/2 & \text{for Type III} \\ (N - 1)/2 & \text{for Type IV.} \end{cases}$$

- In the equation of the previous transparency, $F(z)$ contains the fixed zeros.
- The $g[n]$'s for $n = 0, 1, \dots, M$ contain the adjustable parameters of $H(z)$ for all the four linear-phase filter types.
- The next two transparencies show the relations between the impulse response coefficients $h[n]$ of the overall filter transfer function $H(z)$ and the $g[n]$'s of $G(z)$.

Determination of the $h[n]$'s in terms of the $g[n]$'s

Type I: $F(z) \equiv 1$.

- Therefore, $h[n] \equiv g[n]$ for $n = 0, 1, \dots, N$.

Type II: $F(z) = (1 + z^{-1})/2$.

- Therefore, $h[n] = 1/2(g[n] + g[n - 1])$.
- This means that $h[0] = g[0]/2$, $h[n] = 1/2(g[n] + g[n - 1])$ for $n = 1, 2, \dots, N - 1$, and $h[N] = g[N - 1]/2$.

Type III: $F(z) = (1 - z^{-2})/2$.

- Therefore, $h[n] = 1/2(g[n] - g[n - 2])$.
- This means that $h[0] = g[0]/2$, $h[1] = g[1]/2$, $h[n] = 1/2(g[n] + g[n - 1])$ for $n = 2, 3, \dots, N - 2$, $h[N - 1] = -g[N - 3]/2$, and $h[N] = -g[N - 2]/2$.

Type IV: $F(z) = (1 - z^{-1})/2$.

- Therefore, $h[n] = 1/2(g[n] - g[n - 1])$.
- This means $h[0] = g[0]/2$, $h[n] = 1/2(g[n] - g[n - 1])$ for $n = 1, 2, \dots, N - 1$, and $h[N] = -g[N - 1]/2$.

Determination of the $g[n]$'s in terms of the $h[n]$'s for Types II, III, and IV

- **Type II:** $H(z) = [(1 + z^{-1})/2]G(z)$
 $\Rightarrow G(z) = 2H(z)/[1 + z^{-1}]$
 $\Rightarrow g[n] = -g[n - 1] + 2h[n], \quad g[n] = 0, n < 0.$
- **Type III:** $H(z) = [(1 - z^{-2})/2]G(z)$
 $\Rightarrow G(z) = 2H(z)/[1 - z^{-2}]$
 $\Rightarrow g[n] = g[n - 2] + 2h[n], \quad g[n] = 0, n < 0.$
- **Type IV:** $H(z) = [(1 - z^{-1})/2]G(z)$
 $\Rightarrow G(z) = 2H(z)/[1 - z^{-1}]$
 $\Rightarrow g[n] = g[n - 1] + 2h[n], \quad g[n] = 0, n < 0.$

EXAMPLES

- To illustrate the use of the formulas of the previous transparency, we consider the introductory examples of pages 25, 30, and 35.

Type II: $N = 7$ and $h[n] = 1$ for $n = 0, 1, \dots, 7$.

- Using $g[n] = -g[n - 1] + 2h[n]$, $g[n] = 0$ for $n < 0$, we obtain $g[0] = -g[-1] + 2h[0] = 2$, $g[1] = -g[0] + 2h[1] = 0$, $g[2] = -g[1] + 2h[2] = 2$, $g[3] = -g[2] + 2h[3] = 0$, $g[4] = -g[3] + 2h[4] = 2$, $g[5] = -g[4] + 2h[5] = 0$, $g[6] = -g[5] + 2h[6] = 2$, $g[7] = -g[6] + 2h[7] = 0$, and, in general, $g[n] = 0$ for $n > 6$.

Type III: $N = 6$ and $h[0] = 1$, $h[1] = 2/3$, $h[2] = 1/3$, $h[3] = 0$, $h[4] = -1/3$, $h[5] = -2/3$, and $h[6] = -1$

- Using $g[n] = g[n - 2] + 2h[n]$, $g[n] = 0$ for $n < 0$, we obtain $g[0] = g[-2] + 2h[0] = 6/3$, $g[1] = g[-1] + 2h[1] = 4/3$, $g[2] = g[0] + 2h[2] = 8/3$, $g[3] = g[1] + 2h[3] = 4/3$, $g[4] = g[2] + 2h[4] = 6/3$, $g[5] = g[3] + 2h[5] = 0$, $g[6] = g[4] + 2h[6] = 0$, and, in general, $g[n] = 0$ for $n > 4$.

Type IV: $N = 7$ and $h[0] = 1$, $h[1] = 5/7$, $h[2] = 3/7$, $h[3] = 1/7$, $h[4] = -1/7$, $h[5] = -3/7$, $h[6] = -5/7$, and $h[7] = -1$

- Using $g[n] = g[n - 1] + 2h[n]$, $g[n] = 0$ for $n < 0$, we obtain $g[0] = g[-1] + 2h[0] = 14/7$, $g[1] = g[0] + 2h[1] = 24/7$, $g[2] = g[1] + 2h[2] = 30/7$, $g[3] = g[2] + 2h[3] = 32/7$, $g[4] = g[3] + 2h[4] = 30/7$, $g[5] = g[4] + 2h[5] = 24/7$, $g[6] = g[5] + 2h[6] = 14/7$, $g[7] = g[6] + 2h[7] = 0$, and, in general, $g[n] = 0$ for $n > 6$.

FREQUENCY RESPONSE FOR THE CASCADE FORM

- Using the substitution $z = e^{j\omega}$ for $H(z) = F(z)G(z)$ as given on transparency 42, we obtain after some manipulations (left as an exercise)

$$H(e^{j\omega}) = F(e^{j\omega})G(e^{j\omega})$$

where

$$F(e^{j\omega}) = \begin{cases} 1 & \text{for Type I} \\ e^{j(-\omega/2)} \cos(\omega/2) & \text{for Type II} \\ e^{j(\pi/2-\omega)} \sin \omega & \text{for Type III} \\ e^{j(\pi/2-\omega/2)} \sin(\omega/2) & \text{for Type IV} \end{cases}$$

and

$$G(e^{j\omega}) = e^{-jM\omega} \left\{ g[M] + \sum_{n=1}^M g[M-n][2 \cos(n\omega)] \right\}.$$

EXPRESSION IN TERMS OF THE ZERO-PHASE FREQUENCY RESPONSE AND PHASE TERM

$$H(e^{j\omega}) = H(\omega)e^{j\phi(\omega)},$$

where

$$H(\omega) = F(\omega)G(\omega)$$

and

$$\phi(\omega) = \begin{cases} -N\omega/2 & \text{for Types I and II} \\ \pi/2 - N\omega/2 & \text{for Types III and IV} \end{cases}$$

with

$$F(\omega) = \begin{cases} 1 & \text{for Type I} \\ \cos(\omega/2) & \text{for Type II} \\ \sin \omega & \text{for Type III} \\ \sin(\omega/2) & \text{for Type IV,} \end{cases}$$

$$G(\omega) = \sum_{n=0}^M a[n] \cos n\omega,$$

and

$$a[n] = \begin{cases} g[M], & n = 0 \\ 2g[M - n], & n \neq 0. \end{cases}$$

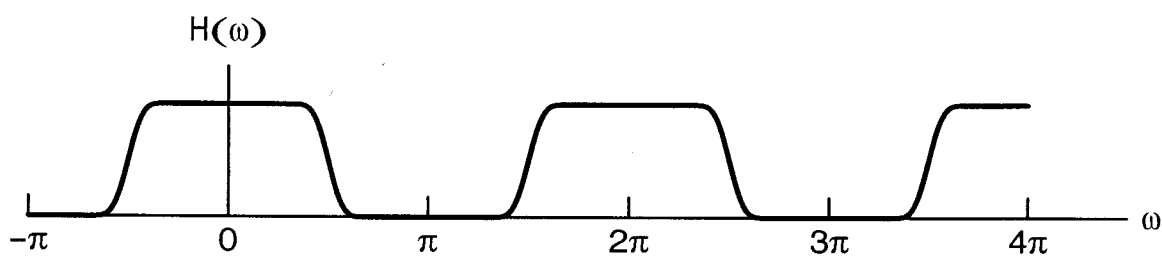
- The above expression form for $G(\omega)$ is very useful when designing linear-phase FIR filters in the minimax sense to meet the given criteria.

ZERO-PHASE RESPONSES FOR DIFFERENT FILTER TYPES

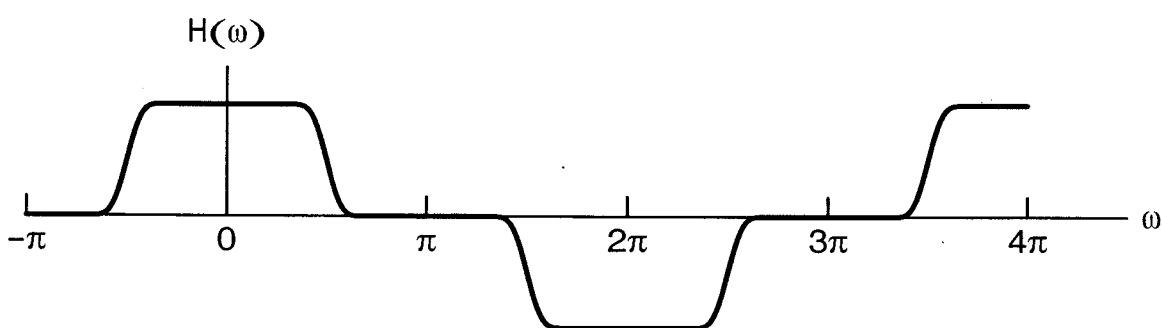
- **Type I:** $H(\omega)$ is even about $\omega = 0$ and $\omega = \pi$ and the periodicity is 2π (see the following transparency).
- **Type II:** The fixed term $F(\omega) = \cos(\omega/2)$ generates a zero for $H(\omega)$ at $\omega = \pi$, making it odd about this point. The periodicity is 4π (see the following transparency).
- **Type III:** The fixed term $F(\omega) = \sin\omega$ gives a zero at both $\omega = 0$ and $\omega = \pi$, making $H(\omega)$ odd about these points. The periodicity is 2π (see the following transparency).
- **Type IV:** The fixed term $F(\omega) = \sin(\omega/2)$ generates a zero at $\omega = 0$, making $H(\omega)$ is odd about $\omega = 0$. The periodicity is 4π (see the following transparency).

EXAMPLE ZERO-PHASE FREQUENCY RESPONSES

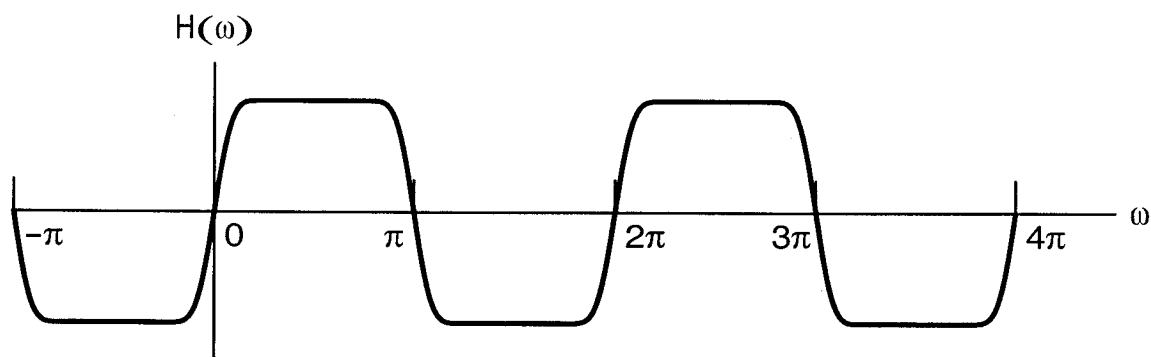
Type I



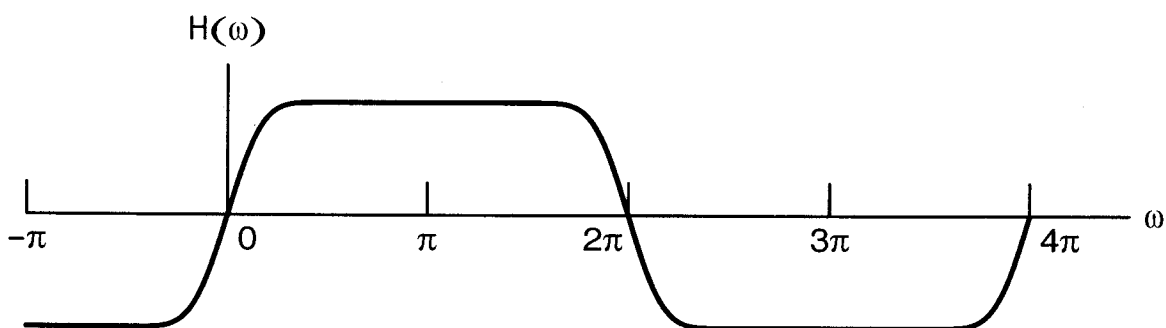
Type II



Type III



Type IV



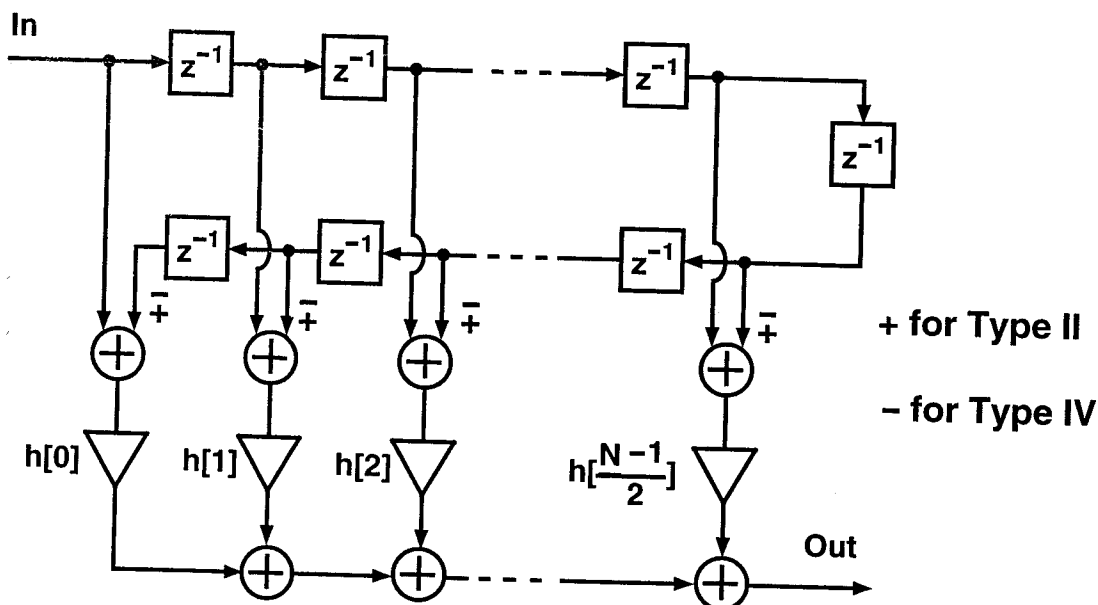
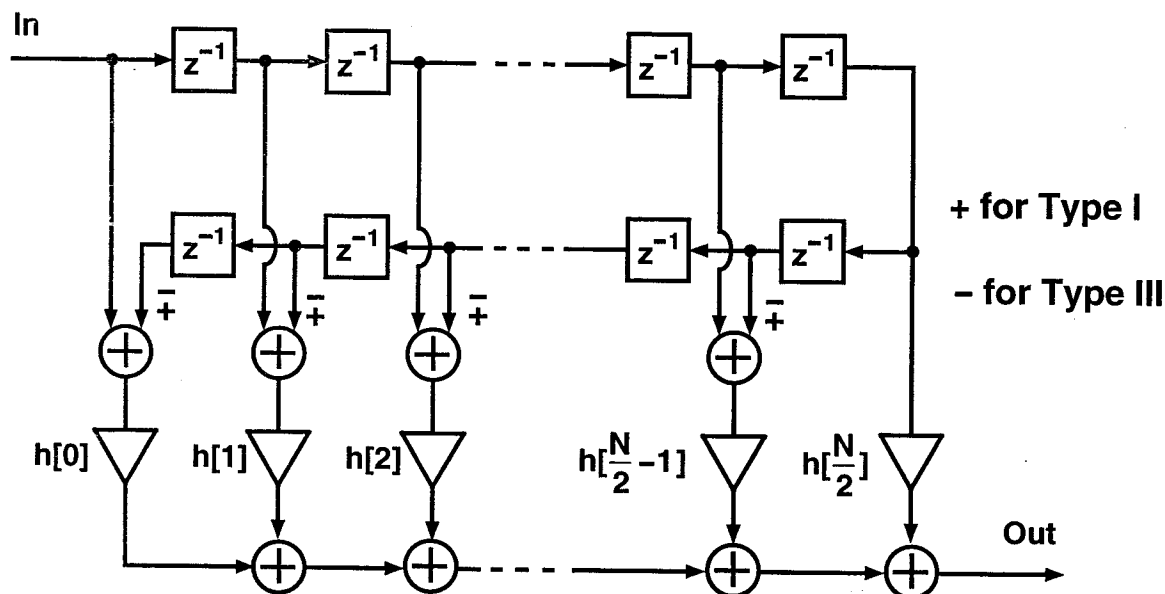
USE OF THE DIFFERENT TYPES

- Type I and II filters are used for conventional filtering applications as, in these cases, the delay caused for sinusoidal signals, $-\phi(\omega)/\omega = N/2$, is independent of the frequency ω .
- Type III and IV filters have an additional 90-degree phase shift and they are most suitable for realizing such filters as differentiators and Hilbert transformers.
- For Type III and IV filters, the delay caused for sinusoidal signals depends on the frequency: $-\phi(\omega)/\omega = N/2 - (\pi/2)/\omega$. However, the group delay, $-d\phi(\omega)/d\omega$, is a constant (equal to $N/2$ in all the cases).

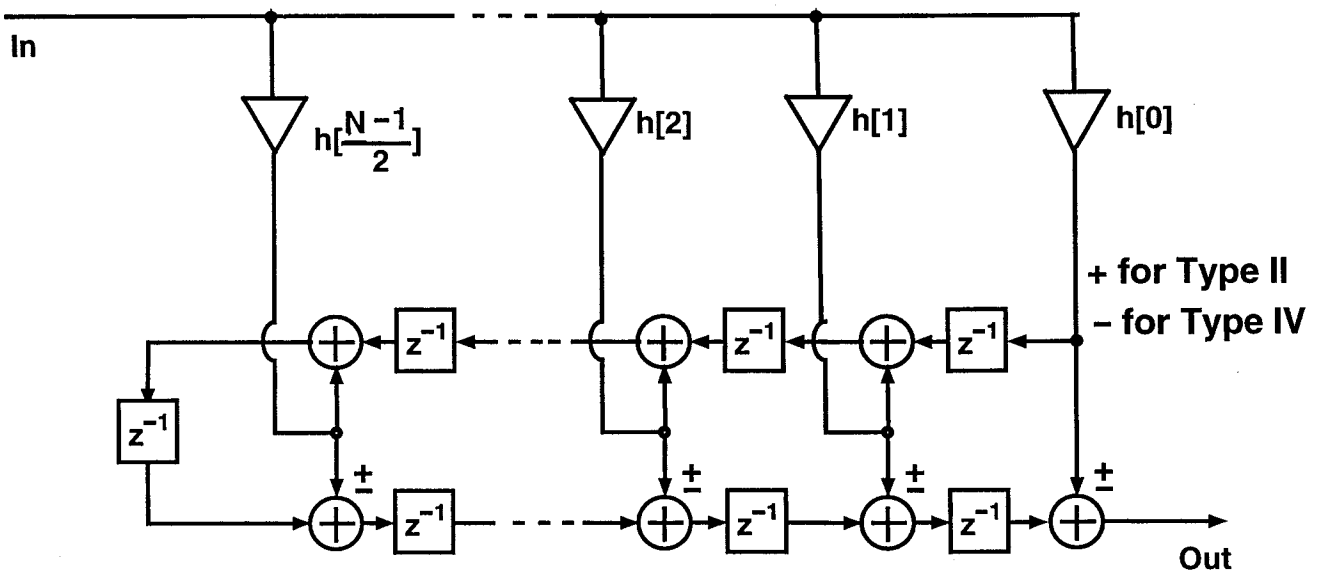
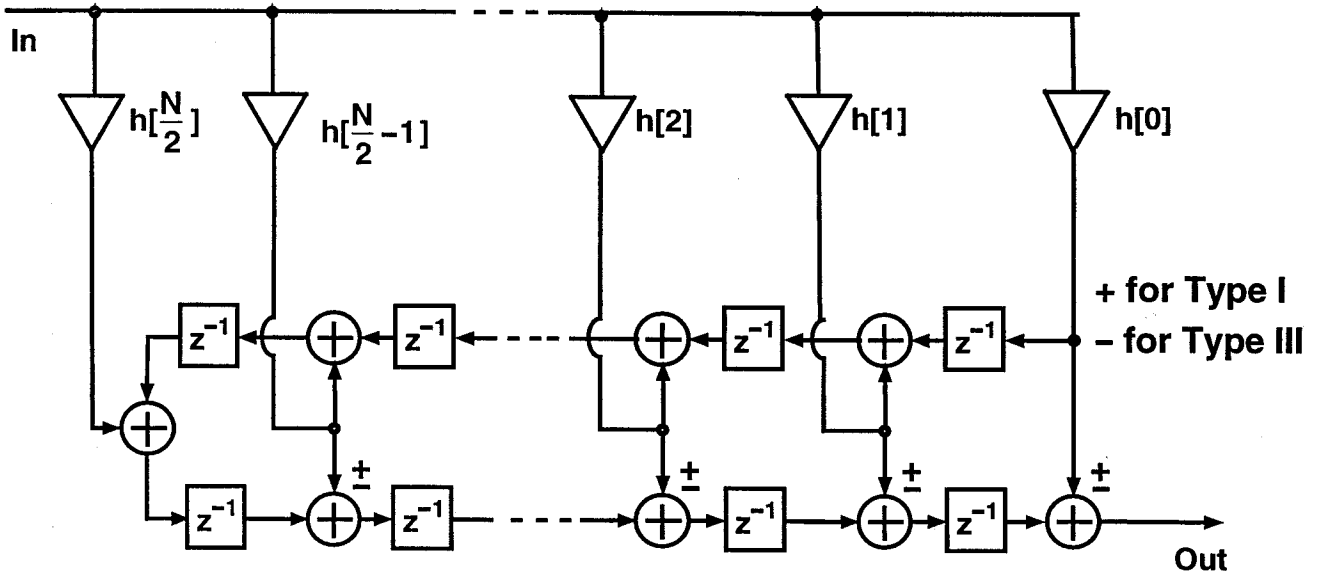
EFFICIENT IMPLEMENTATIONS: STRUCTURES EXPLOITING THE COEFFICIENT SYMMETRY

- Because of the symmetry in the filter coefficients, all the four linear phase filter types can be implemented such that the multipliers needed in the actual implementation is approximately half the filter order N .
- The following two transparencies give the direct-form structures and the transposed direct-form structures exploiting the coefficient symmetry for all the linear-phase filter types.
- For Type I and II filters, the number of multipliers needed is $1 + N/2$ and $(N + 1)/2$, respectively.
- For Type III and IV filters, the number of multipliers needed is $N/2$ ($h[N/2] = 0$) and $(N + 1)/2$, respectively.

EFFICIENT IMPLEMENTATIONS: DIRECT-FORM STRUCTURES EXPLOITING THE COEFFICIENT SYMMETRY



EFFICIENT IMPLEMENTATIONS: TRANSPOSED STRUCTURES EXPLOITING THE COEFFICIENT SYMMETRY



ZERO LOCATIONS FOR THE COMMON ADJUSTABLE TYPE I TRANSFER FUNCTION $G(z)$

1. Since $g[2M - n] = g[n]$ with $2M$ being the order of $G(z)$,

$$G(z) = g[M]z^{-M} + \sum_{n=1}^{M-1} g[n][z^{-n} + z^{-(2M-n)}],$$

and

$$\begin{aligned} G(z^{-1}) &= g[M]z^M + \sum_{n=1}^{M-1} h[n][z^n + z^{(n-2M)}] \\ &= z^{2M} \left\{ g[M]z^{-M} + \sum_{n=1}^{M-1} g[n][z^{-n} + z^{-(2M-n)}] \right\} \\ &= z^{2M} G(z). \end{aligned}$$

- This means that $G(z)$ and $G(z^{-1})$ have identical zeros and the zeros of $G(z)$ thus occur in reciprocal pairs (If there is a zero at $z = re^{j\phi}$, then there exists also zero at $z = 1/(re^{j\phi}) = (1/r)e^{-j\phi}$).
2. The coefficients of $G(z)$ are real so that the zeros are either real or occur in complex conjugate pairs (If there is a zero at $z = re^{j\phi}$, then there exists also zero at $z = re^{-j\phi}$).

- From the above facts, it follows that $G(z)$ is expressible as

$$G(z) = g[0]G_1(z)G_2(z)G_3(z),$$

where

$$\begin{aligned} G_1(z) = & \prod_{i=1}^{N_1} (1 - [2(r_i + \frac{1}{r_i}) \cos \theta_i]z^{-1} \\ & + [r_i^2 + \frac{1}{r_i^2} + 4 \cos^2 \theta_i]z^{-2} \\ & - [2(r_i + \frac{1}{r_i}) \cos \theta_i]z^{-3} + z^{-4}) \end{aligned}$$

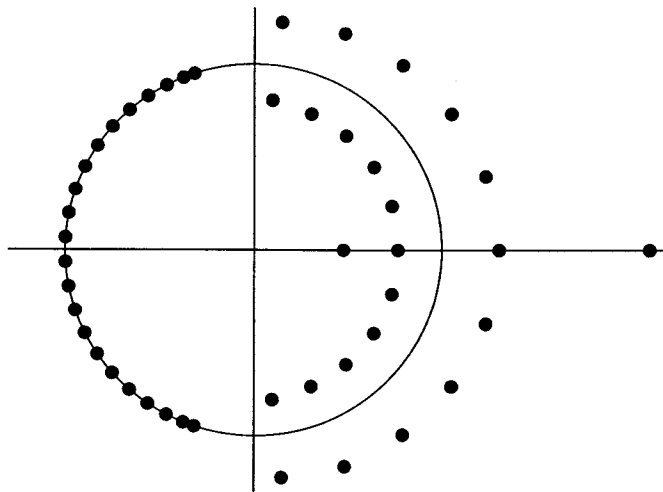
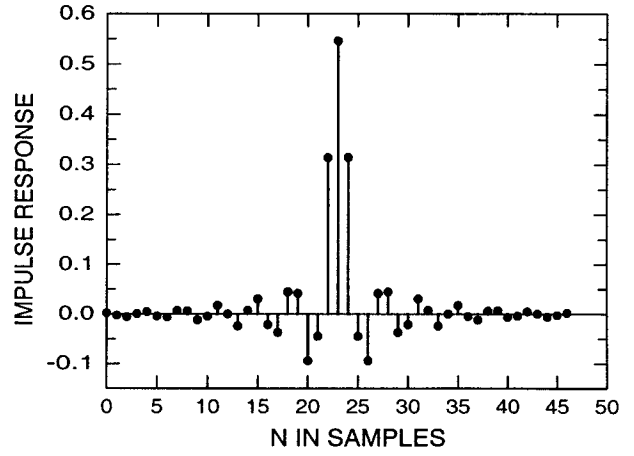
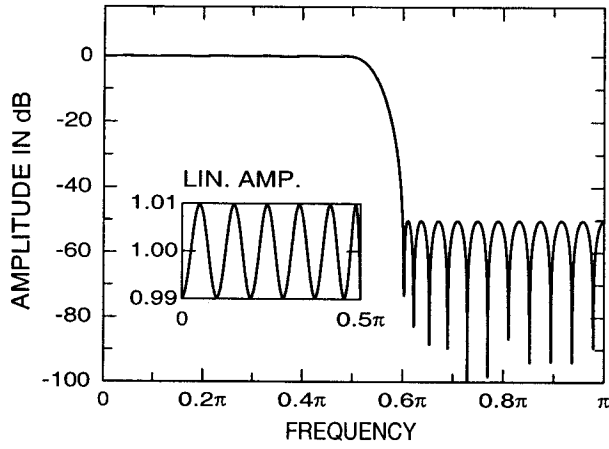
$$G_2(z) = \prod_{i=1}^{N_2} (1 - [2 \cos \hat{\theta}_i]z^{-1} + z^{-2})$$

$$G_3(z) = \prod_{i=1}^{N_3} (1 - [\hat{r}_i + \frac{1}{\hat{r}_i}]z^{-1} + z^{-2}),$$

1. $G_1(z)$ contains the zeros occurring in quadruplets, that is, in complex conjugate and mirror-image pairs off the unit circle at $z = r_i e^{\pm j\theta_i}$, $(1/r_i) e^{\pm j\theta_i}$ for $i = 1, 2, \dots, N_1$.
2. $G_2(z)$ contains the zeros occurring in complex conjugate pairs on the unit circle at $z = e^{\pm j\hat{\theta}_i}$ for $i = 1, 2, \dots, N_2$.

- The reciprocal zero of $z = e^{j\theta}$ is $z = 1/e^{j\theta} = e^{-j\theta}$, that is, it is simultaneously the complex conjugate of the zero at $z = e^{j\theta}$.
3. $G_3(z)$ contains the zeros occurring in reciprocal pairs on the real axis at $z = \hat{r}_i, 1/\hat{r}_i$ for $i = 1, 2, \dots, N_3$.
 4. If $G(z)$ possesses a zero at $z = 1$ or at $z = -1$, then it follows from the symmetry of $G(z)$ and the fact that $G(z)$ is of even order that the number of zeros at this point must be even.
- The figure of the next transparency shows several characteristics for a Type I filter of order 46 [$H(z) \equiv G(z)$] including the zero plot.

Some Characteristics of a Type I Filter of order $N = 46$



DIFFERENCE BETWEEN THE LINEAR-PHASE CASES

1. Type I designs have an even number or no zeros at $z = 1$ and at $z = -1$.
 2. Type II designs have either an even number or no zeros at $z = 1$, and an odd number of zeros at $z = -1$.
 3. Type III designs have an odd number of zeros at $z = 1$ and at $z = -1$.
 4. Type IV designs have an odd number of zeros at $z = 1$, and either an even number or no zeros $z = -1$.
- Zero plots will be given for several filters in the following examples after introducing one more useful representation form for the zero-phase frequency response $G(\omega)$.

POLYNOMIAL REPRESENTATION FORM FOR TYPE I FILTERS

$$H(\omega) \equiv G(\omega) = \sum_{n=0}^M a[n] \cos n\omega \quad (A)$$

- Using the identity

$$\cos n\omega = T_n(\cos \omega), \quad (B)$$

where $T_n(x)$ is the n -th degree Chebyshev polynomial (see the next transparency), defined by

$$T_n(x) = \begin{cases} \cos(n \cos^{-1} x) & \text{for } |x| \leq 1 \\ \cosh(n \cosh^{-1} x) & \text{for } |x| > 1, \end{cases}$$

we get

$$G(\omega) = \sum_{n=0}^M \alpha[n] \cos^n \omega. \quad (C)$$

- Chebyshev polynomials can be conveniently generated by using the following recursion formulas:

$$T_0[x] = 1$$

$$T_1[x] = x$$

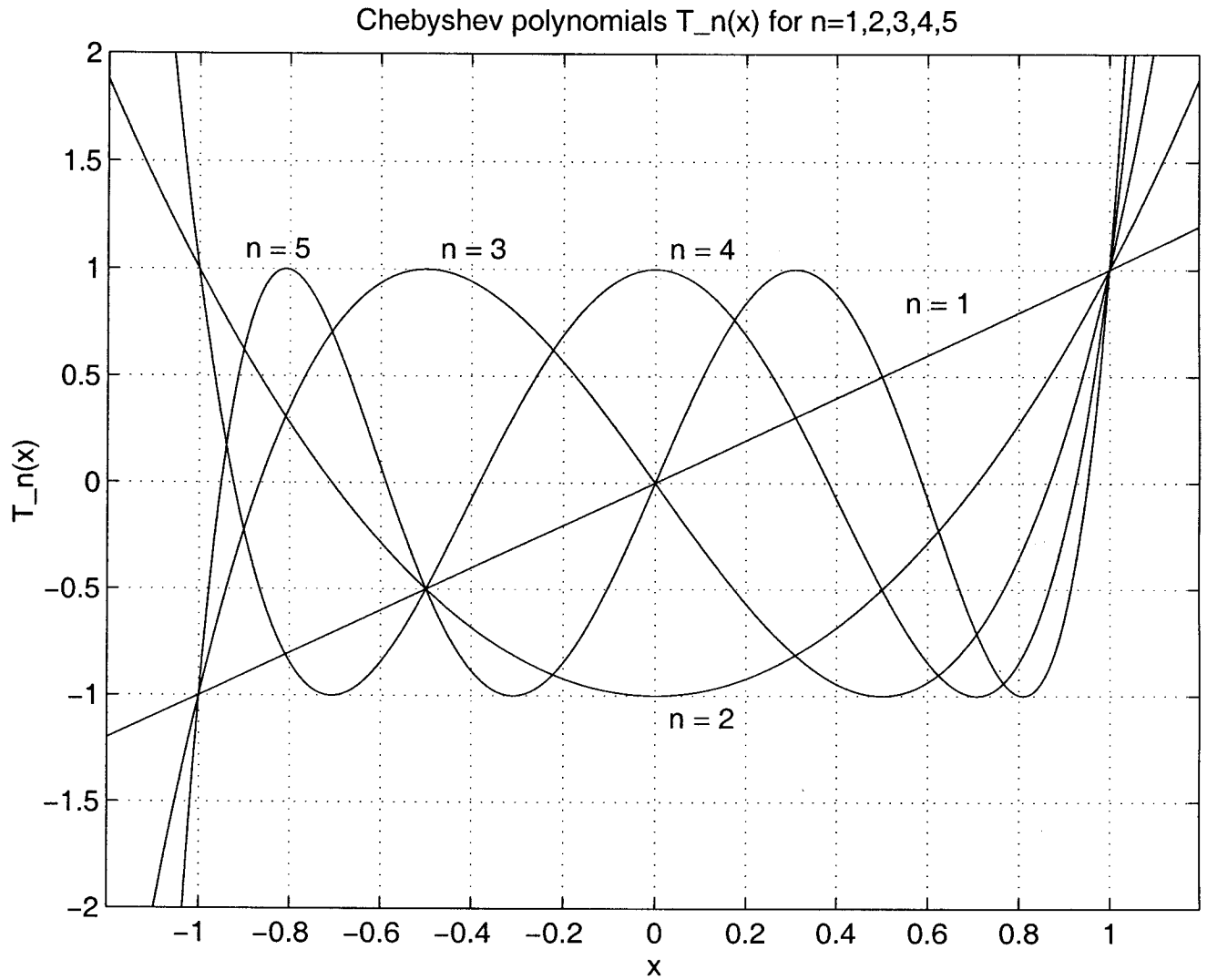
$$T_n[x] = 2xT_{n-1}[x] - T_{n-2}[x].$$

- Later on, the above form is used in designing maximally-flat linear-phase FIR filters as well as some simple FIR filters.

Characteristics of Chebyshev Polynomials

- The first six Chebyshev polynomials are $T_0(x) = 1$, $T_1(x) = x$, $T_2(x) = 2x^2 - 1$, $T_3(x) = 4x^3 - 3x$, $T_4(x) = 8x^4 - 8x^2 + 1$, and $T_5(x) = 16x^5 - 20x^3 + 5x$.
- If $M = 5$ in Equation (A), then the above equations and the identity of Equation (B) shown that $G(\omega)$ can be expressed in the form of Equation (C) with $\alpha[5] = 16a[5]$, $\alpha[4] = 8a[4]$, $\alpha[3] = -20a[5] + 4a[3]$, $\alpha[2] = -8a[4] + 2a[2]$, $\alpha[1] = 5a[5] - 3a[3] + a[1]$, and $\alpha[0] = a[4] - a[2] + a[0]$.
- The figure of the next transparency shows the responses of $T_n(x)$ for $n = 1, 2, \dots, 5$.
- It is observed that $T_n(x)$ oscillates in the interval $[-1, 1]$ between ± 1 achieving these values at $n + 1$ points such that the value at $x = 1$ [$x = -1$] is 1 [$(-1)^n$].
- This attractive behavior is utilized in Section 6 for generating FIR filters whose response oscillates within $\pm\delta_s$ ($1 \pm \delta_p$) in the stopband (passband).

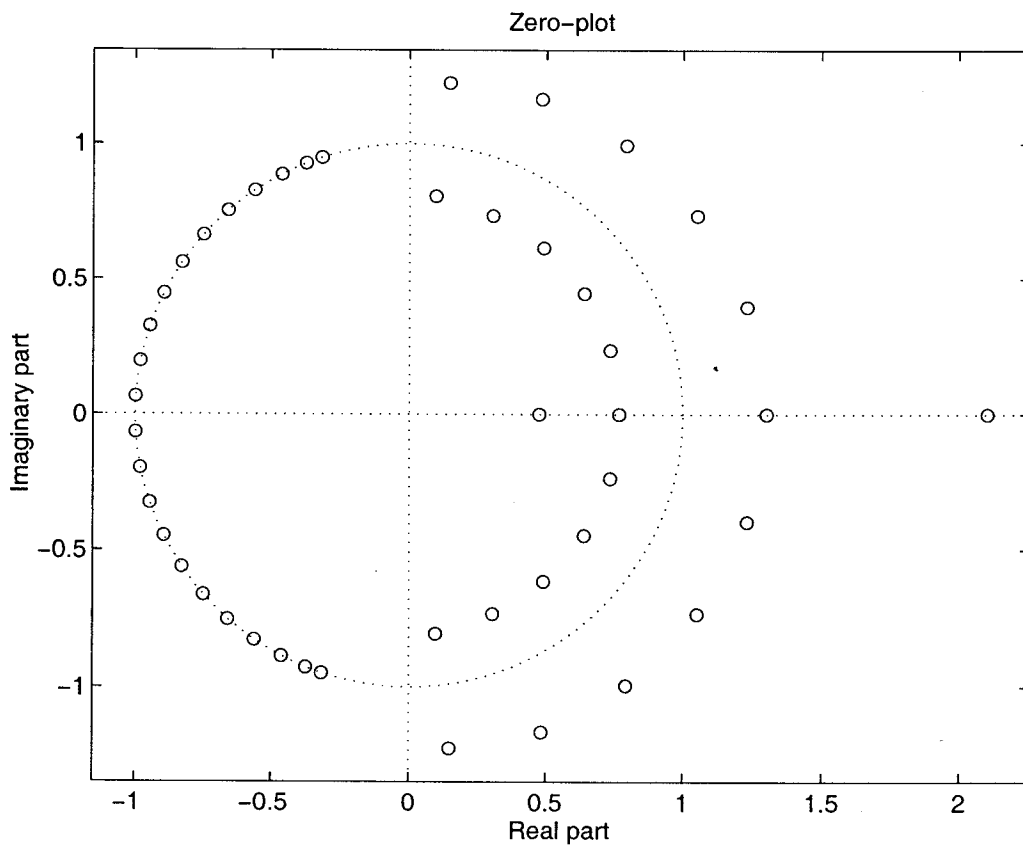
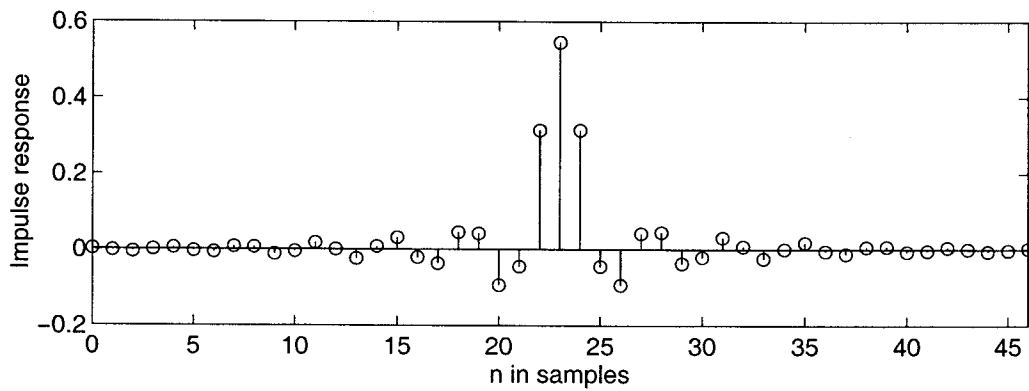
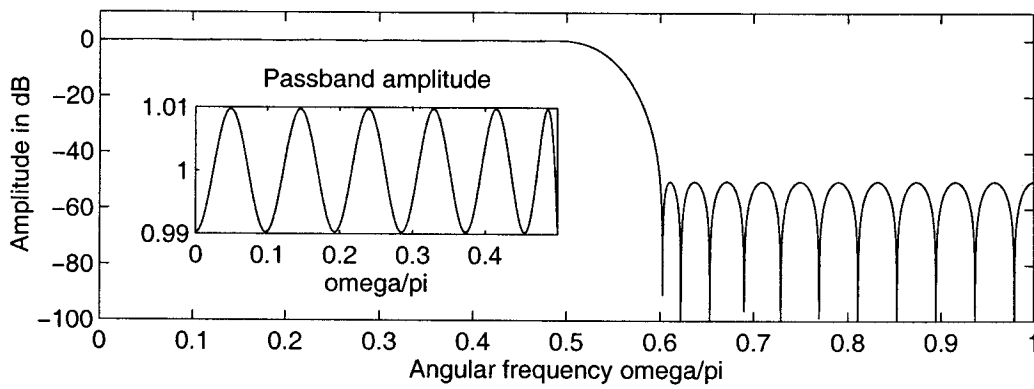
Chebyshev Polynomials $T_n(x)$ for $n = 1, 2, 3, 4, 5$



SOME TYPICAL LINEAR-PHASE FIR FILTERS

- In the following, six typical linear-phase FIR filters will be given.
- In each case, the impulse response, the zero plot as well as the amplitude response are given.
- The phase response is not given as it is known after fixing the filter type and the filter order.
- Also the weighted error function and the matlab code for generating the filter in the minimax sense are given for each case.
- This error function as well as the matlab code should become clear after reading Section 7: Design of linear-phase FIR filters in the minimax sense.
- This is why it is not worth trying to understand these topics at this point. Please return back to the following six examples after reading Section 7.

EXAMPLE 1: Type I lowpass filter, $N = 46$



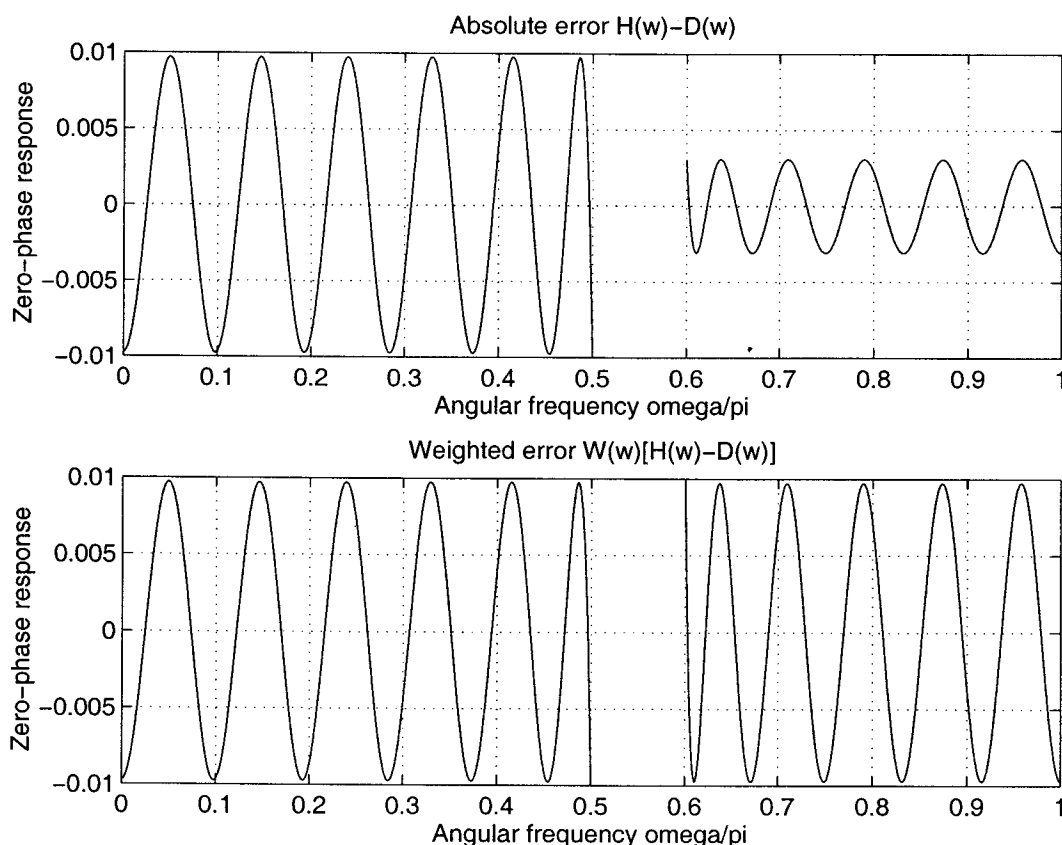
EXAMPLE 1: Error function

- The filter coefficients have been determined to minimize on $[0, 0.5\pi] \cup [0.6\pi, \pi]$ the peak absolute value of

$$E(\omega) = W(\omega)[H(\omega) - D(\omega)],$$

where $D(\omega) = 1$ and $W(\omega) = 1$ on $[0, 0.5\pi]$ and $D(\omega) = 0$ and $W(\omega) = \sqrt{10}$ on $[0.6\pi, \pi]$.

- $N = 46$ is the minimum order to make this quantity less than 0.01 (passband ripple). The stopband ripple is then $0.01/\sqrt{10}$ (50 dB).

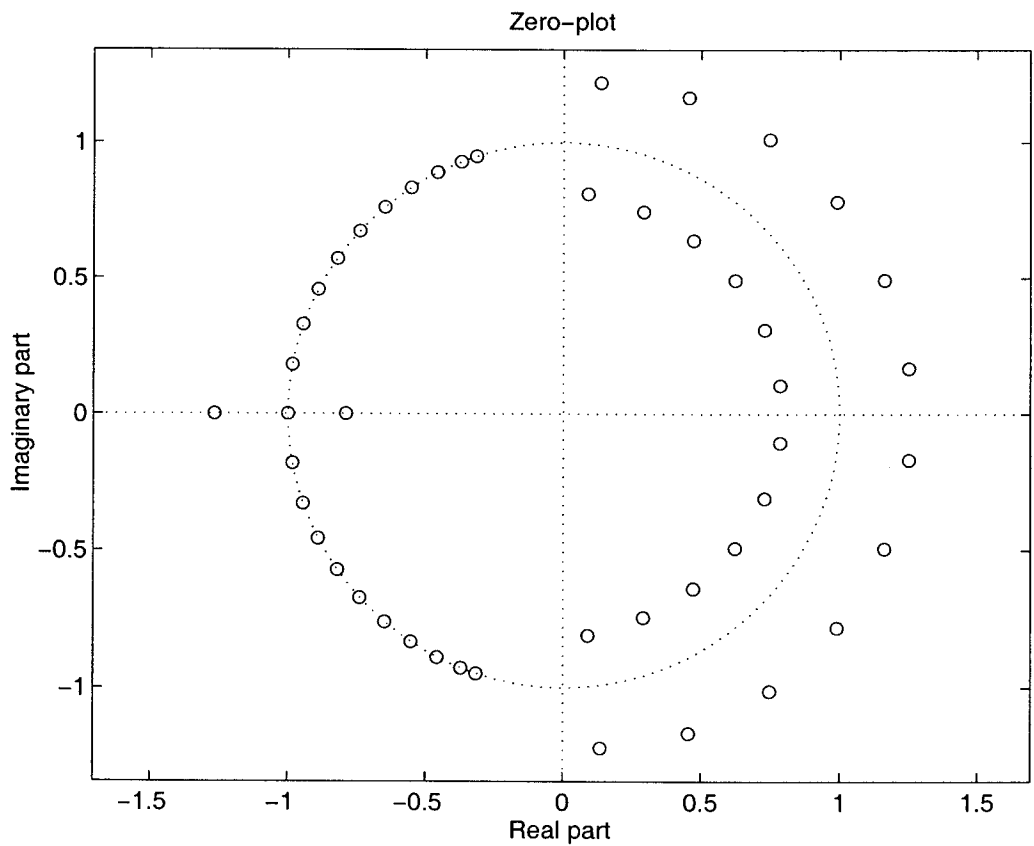
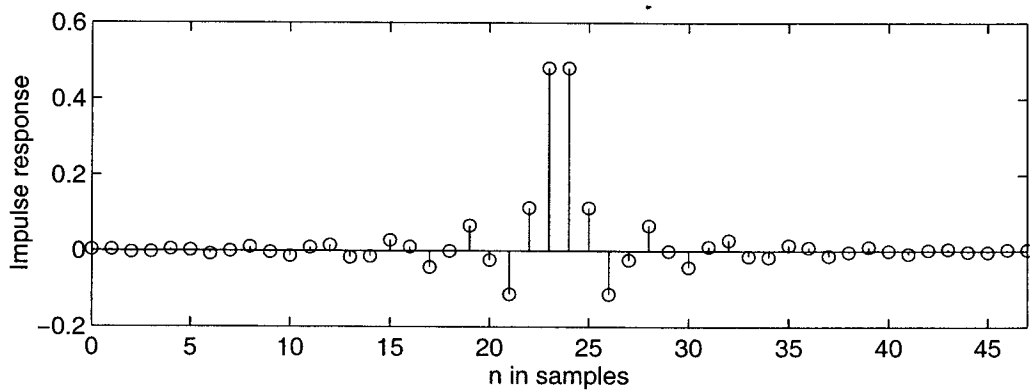
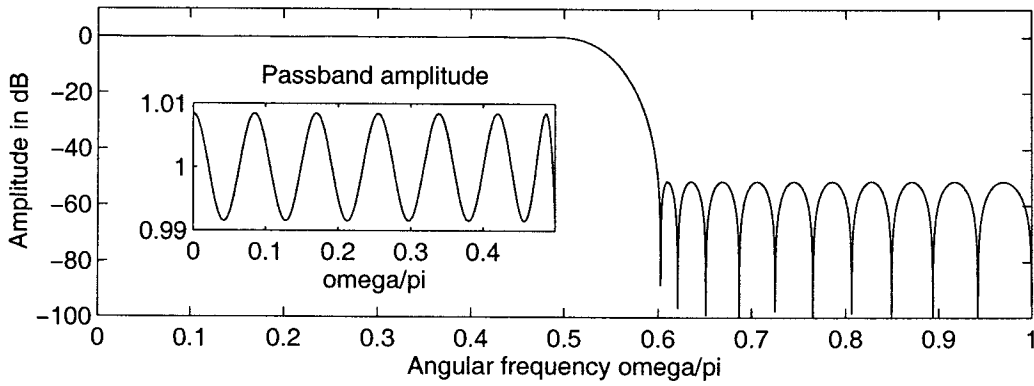


Example 1: Matlab code

67

```
%Design of an FIR filter of order 46
%(Type I) using the Remez algorithm
%with passband edge at 0.5pi, stopband
%edge at 0.6pi, passband ripple=0.01,
%stopband ripple =sqrt(.00001)
%(50 dB attenuation)
%Tapio Saramaki 28.10.1995
%This can be found in SUN's
%~ts/matlab/dsp/luefir1.m
m=[1 0];
dev=[.01,sqrt(.00001)];
f=[.5,.6];
[n,f0,m0,w]=remezord(f,m,dev,2);
n=46
h=remez(n,f0,m0,w);
[H,f]=freqz(h,1,4*2048,2);
figure(1)
subplot(2,1,1)
plot(f,20*log10(abs(H)));
axis([0 1 -100 10]);
ylabel('Amplitude in dB');
xlabel('Angular frequency omega/pi');
hold on;
axes('position',[.21 .68 .3 .14]);
plot(f,(abs(H)));axis([0 .5 .99 1.01]);
title('Passband amplitude');xlabel('omega/pi');
hold off;
subplot(2,1,2)
impz(h);xlabel('n in samples');
ylabel('Impulse response');
figure(2)
zplane(h);title('Zero-plot')
figure(3);
[H1,w1]=zeroam(h,.0,.5,500);
[H2,w2]=zeroam(h,.6,1.,500);
%zeroam is a routine evaluating
%the zero-phase frequency response;
%programmed by Tapio Saramaki
%This can be found in SUN's
%~ts/matlab/dsp/zeroam.m
subplot(2,1,1)
plot(w1/pi,H1-1,w2/pi,H2);grid;
xlabel('Angular frequency omega/pi');
ylabel('Zero-phase response');
title('Error H(w)-D(w)')
subplot(2,1,2)
plot(w1/pi,H1-1,w2/pi,H2*sqrt(10));grid;
xlabel('Angular frequency omega/pi');
ylabel('Zero-phase response');
title('Weighted error W(w)[H(w)-D(w)]')
```


EXAMPLE 2: Type II lowpass filter, $N = 47$

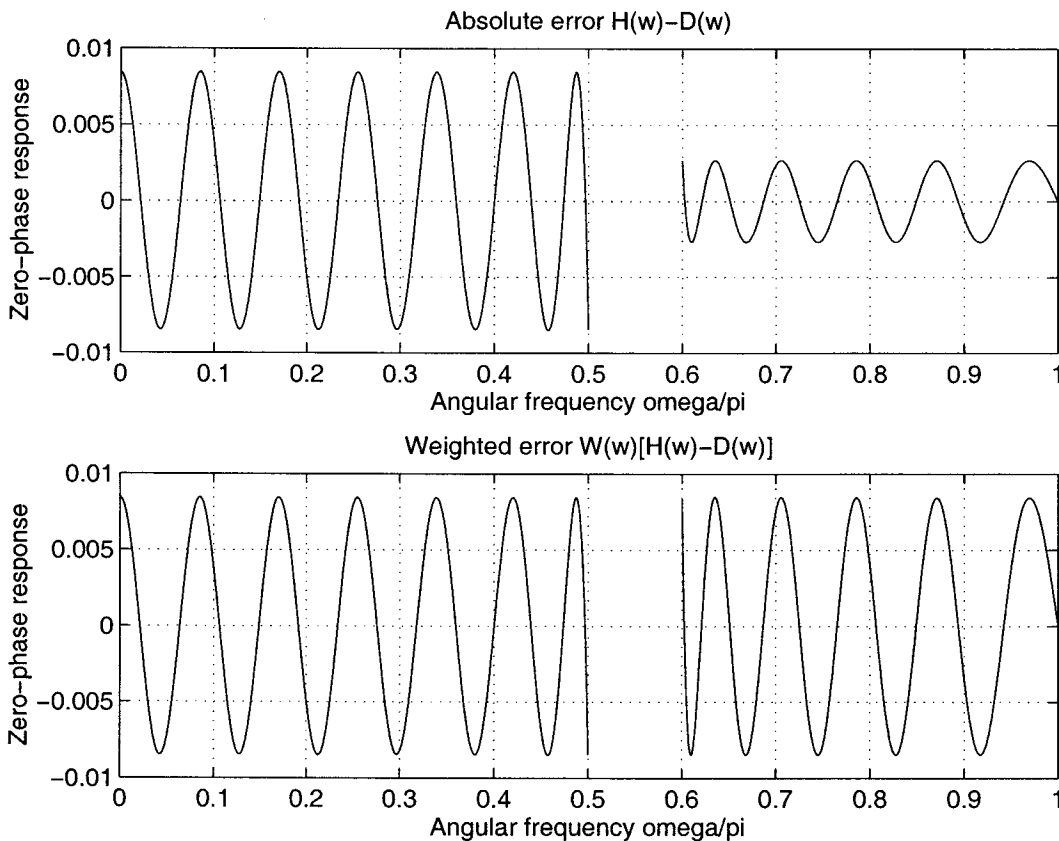


EXAMPLE 2: Error function

- The filter coefficients have been determined to minimize on $[0, 0.5\pi] \cup [0.6\pi, \pi]$ the peak absolute value of

$$E(\omega) = W(\omega)[H(\omega) - D(\omega)],$$

where $D(\omega) = 1$ and $W(\omega) = 1$ on $[0, 0.5\pi]$ and $D(\omega) = 0$ and $W(\omega) = \sqrt{10}$ on $[0.6\pi, \pi]$.

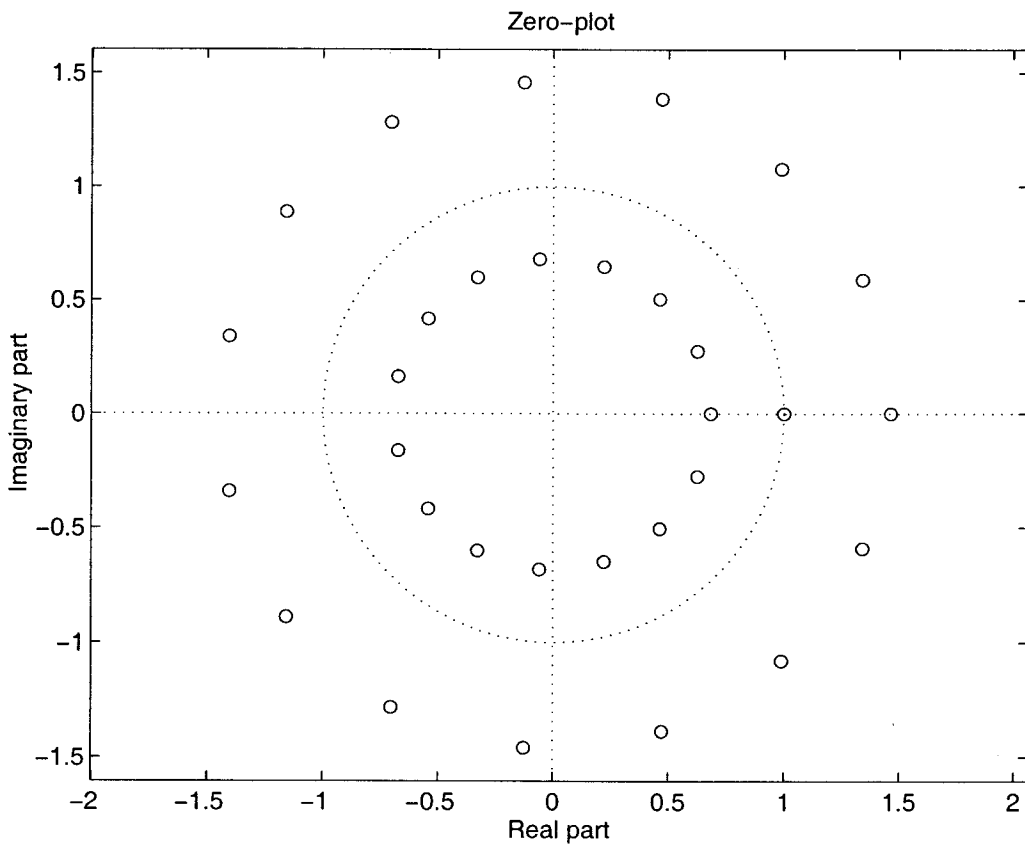
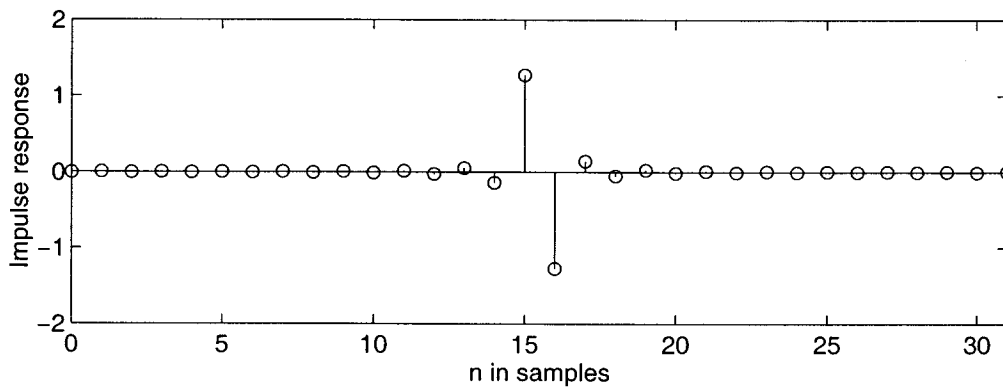
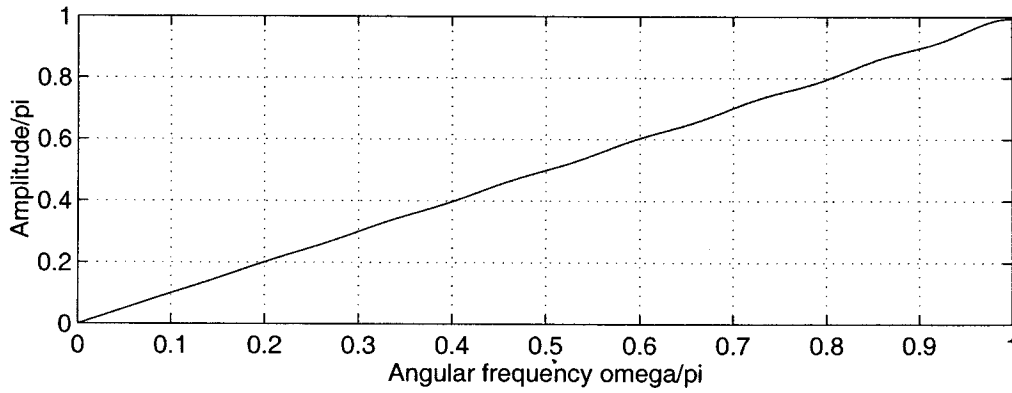


Example 2: Matlab code

70

```
%Design of an FIR filter of order 47
%(Type II) using the Remez algorithm
%with passband edge at 0.5pi, stopband
%edge at 0.6pi, passband ripple=0.01,
%stopband ripple =sqrt(.00001)
%(50 dB attenuation)
%Tapio Saramaki 28.10.1995
%This can be found in SUN's
%~ts/matlab/dsp/luefir2.m
m=[1 0];
dev=[.01,sqrt(.00001)];
f=[.5,.6];
[n,f0,m0,w]=remezord(f,m,dev,2);
n=47
h=remez(n,f0,m0,w);
[H,f]=freqz(h,1,4*2048,2);
figure(1)
subplot(2,1,1)
plot(f,20*log10(abs(H))),grid;
axis([0 1 -100 10]);grid;
ylabel('Amplitude in dB');
xlabel('Angular frequency omega/pi');
hold on;
axes('position',[.21 .68 .3 .14]);
plot(f,(abs(H)));axis([0 .5 .99 1.01]);
title('Passband amplitude');xlabel('omega/pi');
hold off;
subplot(2,1,2)
impz(h);xlabel('n in samples');
ylabel('Impulse response');
figure(2)
zplane(h);title('Zero-plot')
figure(3);
[H1,w1]=zeroam(h,.0,.5,500);
[H2,w2]=zeroam(h,.6,1.,500);
%zeroam is a routine evaluating
%the zero-phase frequency response;
%programmed by Tapio Saramaki
%This can be found in SUN's
%~ts/matlab/dsp/zeroam.m
subplot(2,1,1)
plot(w1/pi,H1-1,w2/pi,H2);grid;
xlabel('Angular frequency omega/pi');
ylabel('Zero-phase response');
title('Error H(w)-D(w)')
subplot(2,1,2)
plot(w1/pi,H1-1,w2/pi,H2*sqrt(10));grid;
xlabel('Angular frequency omega/pi');
ylabel('Zero-phase response');
title('Weighted error W(w)[H(w)-D(w)]')
```

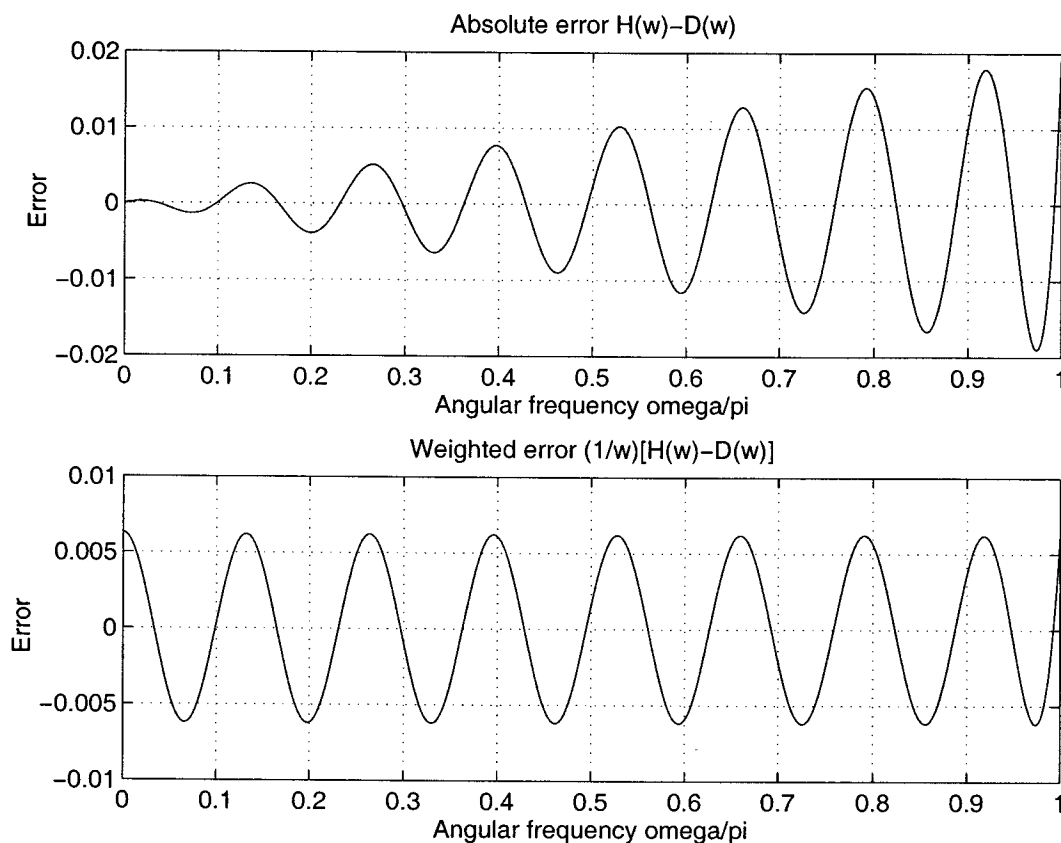
EXAMPLE 3: Type IV differentiator, $N = 31$



EXAMPLE 3: Error function

- The filter coefficients have been determined to minimize on $[0, \pi]$ the peak absolute value of the error function

$$E(\omega) = (1/\omega)[H(\omega) - \omega].$$

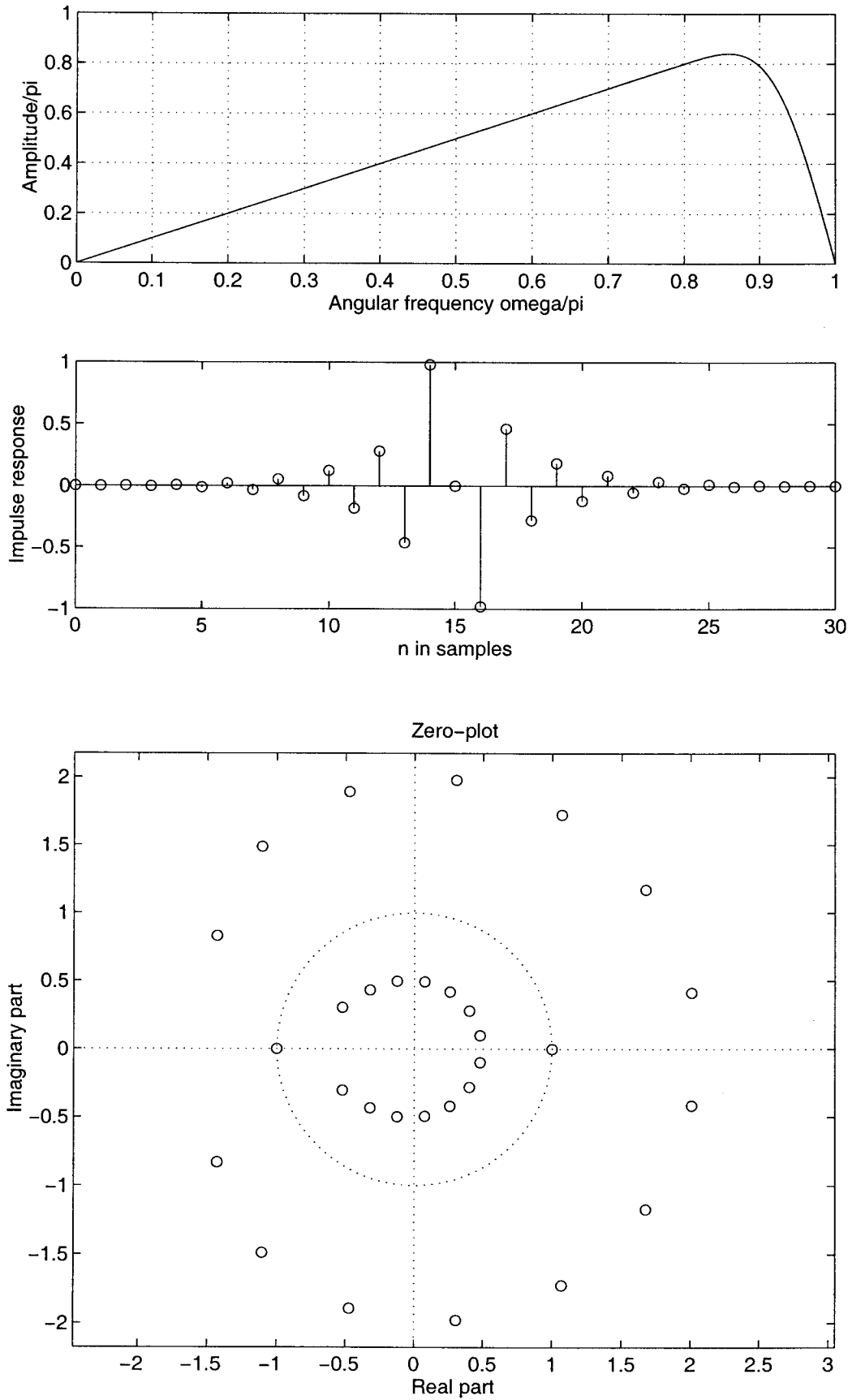


Example 3: Matlab code

73

```
%Design of an FIR differentiator of
%order 31 (Type IV) using
%the Remez algorithm with edges at 0
%and pi. The zero-phase response is
%desired to follow the desired function
%D(w)=w, that is, at pi the desired
%value is pi.
%Tapio Saramaki 28.10.1995
%This can be found in SUN's
%~ts/matlab/dsp/luefir3.m
f=[0 1]; m=[0 pi];
n=31;
h=remez(n,f,m,'differentiator');
h=-h;
%the algorithm gives a wrong sign
[H,w]=zeroam(h,0,1,4000);
%zeroam is a routine evaluating
%the zero-phase frequency response;
%programmed by Tapio Saramaki
%This can be found in SUN's
%~ts/matlab/dsp/zeroam.m
figure(1)
subplot(2,1,1)
plot(w/pi,(abs(H))/pi),grid;
ylabel('Amplitude/pi');
xlabel('Angular frequency omega/pi');
subplot(2,1,2)
impz(h);xlabel('n in samples');
ylabel('Impulse response');
figure(2)
zplane(h);title('Zero-plot')
figure(3);
subplot(2,1,1)
H1=H-w;
plot(w/pi,H1);grid;
xlabel('Angular frequency omega/pi');
ylabel('Error');
title('Error H(w)-D(w)')
subplot(2,1,2)
plot(w/pi,H1./w);grid;
xlabel('Angular frequency omega/pi');
ylabel('Error');
title('Weighted error (1/w)[H(w)-D(w)]')
```

EXAMPLE 4: Type III differentiator, $N = 30$

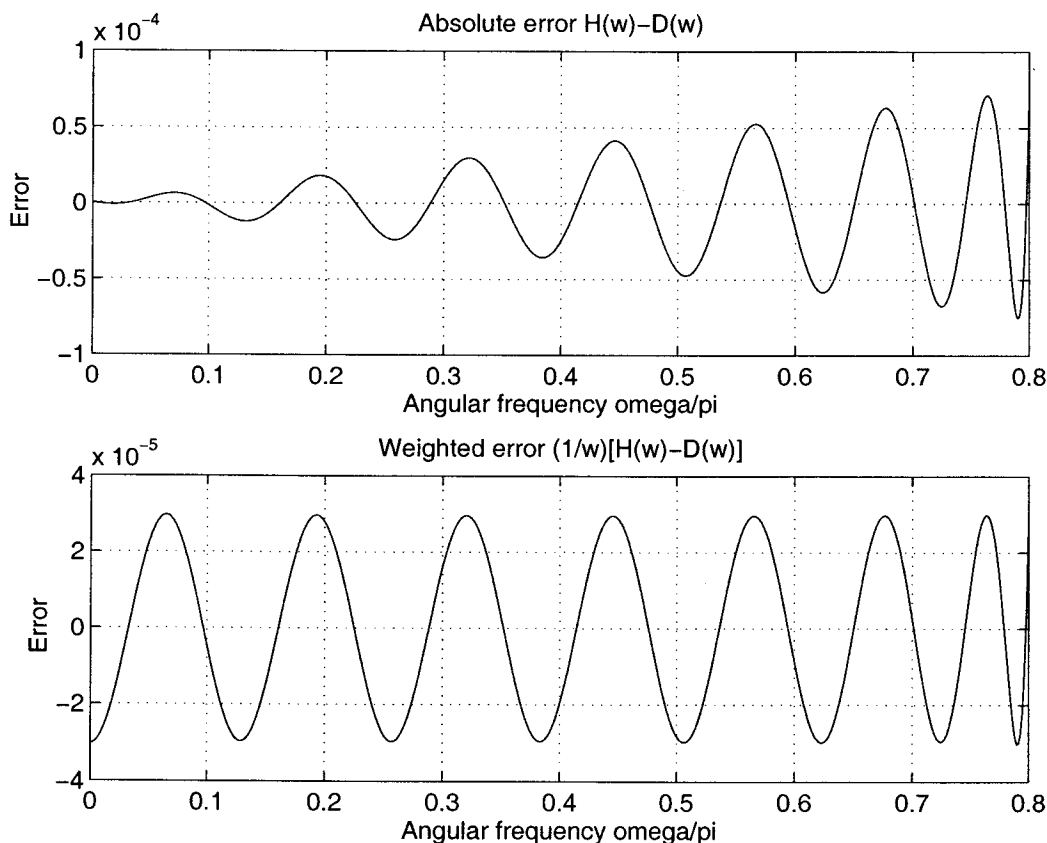


EXAMPLE 4: Error function

- The filter coefficients have been determined to minimize on $[0, 0.8\pi]$ the peak absolute value of the error function

$$E(\omega) = (1/\omega)[H(\omega) - \omega].$$

- Because of a fixed zero at $\omega = \pi$, the approximation interval is selected to be $[0, 0.8\pi]$ instead of $[0, \pi]$.

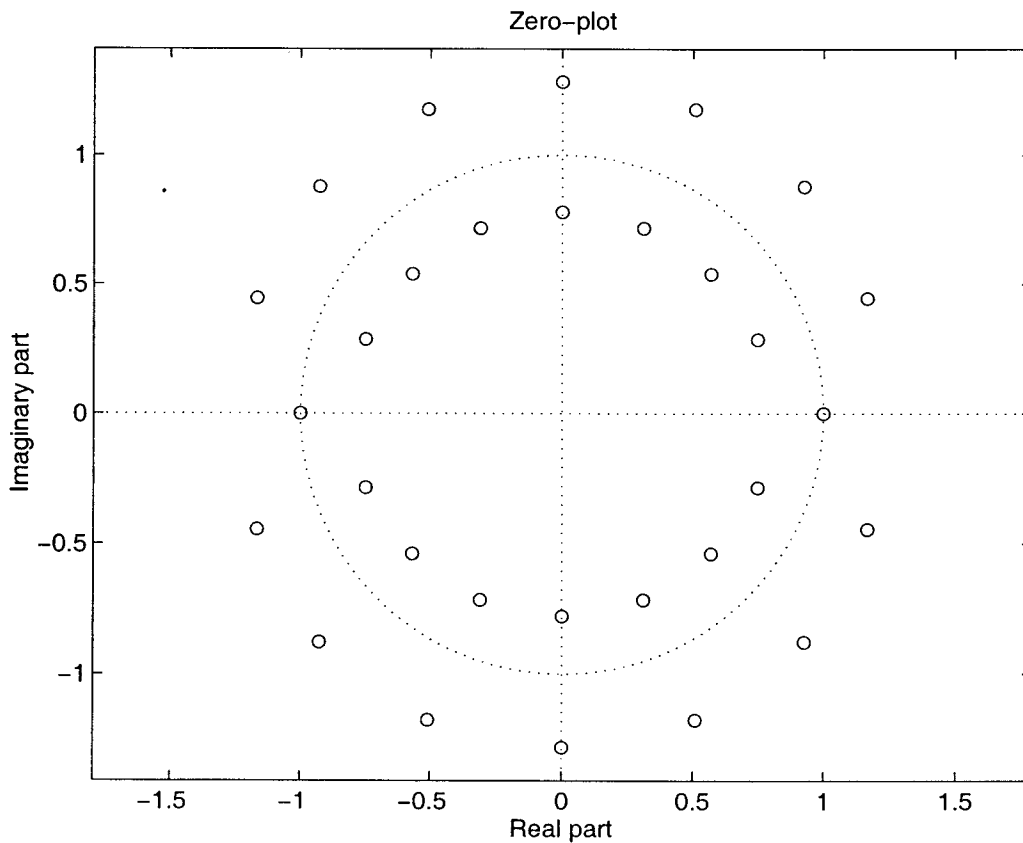
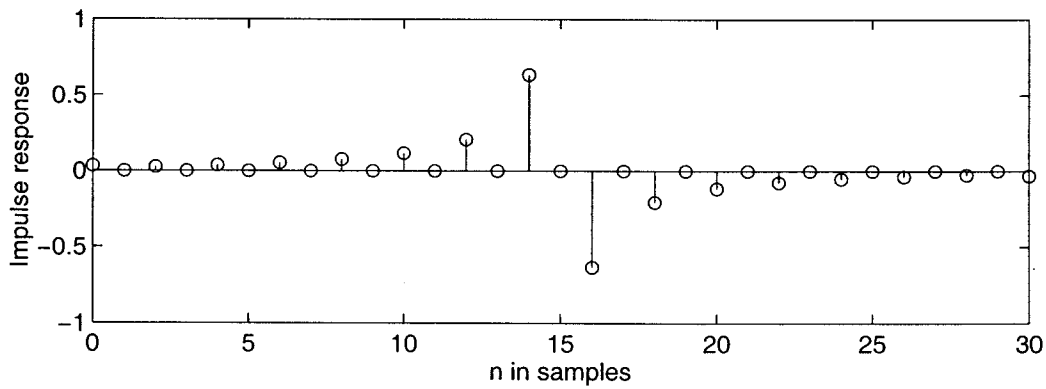
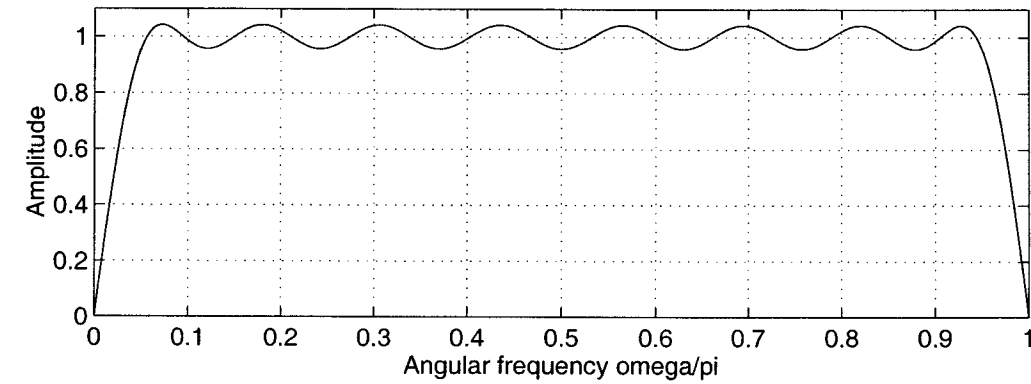


Example 4: Matlab code

76

```
%Design of an FIR differentiator of
%order 30 (Type III) using
%the Remez algorithm with edges at 0
%and 0.8*pi. The zero-phase response is
%desired to follow the desired function
%D(w)=w, that is, at pi the desired
%value is pi. However, since this filter
%is forced to have a zero at z=-1, the
%upper edge must be less than pi.
%Tapio Saramaki 28.10.1995
%This can be found in SUN's
%~ts/matlab/dsp/luefir4.m
f=[0 .8]; m=[0 0.8*pi];
n=30;
h=remez(n,f,m,'differentiator');
h=-h;
%the algorithm gives a wrong sign
[H,w]=zeroam(h,.0,1.,4000);
%zeroam is a routine evaluating
%the zero-phase frequency response;
%programmed by Tapio Saramaki
%This can be found in SUN's
%~ts/matlab/dsp/zeroam.m
figure(1)
subplot(2,1,1)
plot(w/pi,(abs(H))/pi),grid;
ylabel('Amplitude/pi');
xlabel('Angular frequency omega/pi');
subplot(2,1,2)
impz(h);xlabel('n in samples');
ylabel('Impulse response');
figure(2)
zplane(h);title('Zero-plot')
[H,w]=zeroam(h,.0,.8,4000);
figure(3);
subplot(2,1,1)
H1=H-w;
plot(w/pi,H1);grid;
xlabel('Angular frequency omega/pi');
ylabel('Error');
title('Error H(w)-D(w)')
subplot(2,1,2)
plot(w/pi,H1./w);grid;
xlabel('Angular frequency omega/pi');
ylabel('Error');
title('Weighted error (1/w)[H(w)-D(w)]')
```

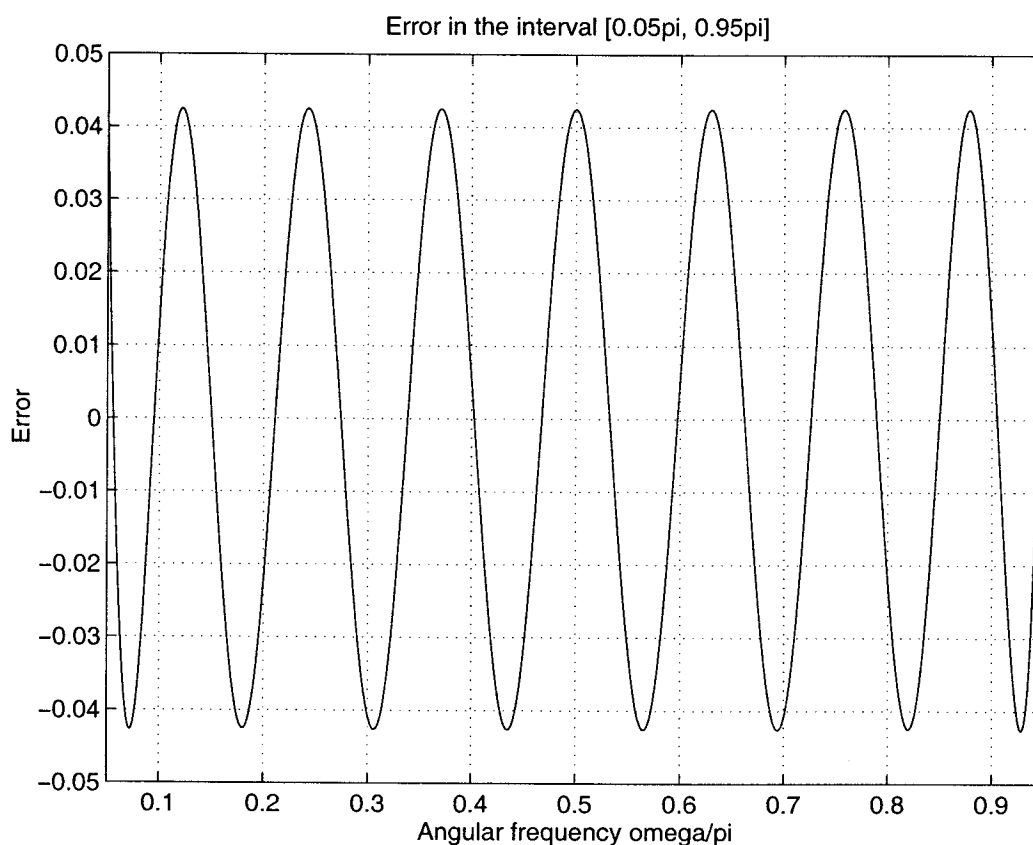
EXAMPLE 5: Type III Hilbert transformer, $N = 30$



EXAMPLE 5: Error function

- The filter coefficients have been determined to minimize on $[0.05\pi, 0.95\pi]$ the peak absolute value of the error function

$$E(\omega) = [H(\omega) - 1].$$

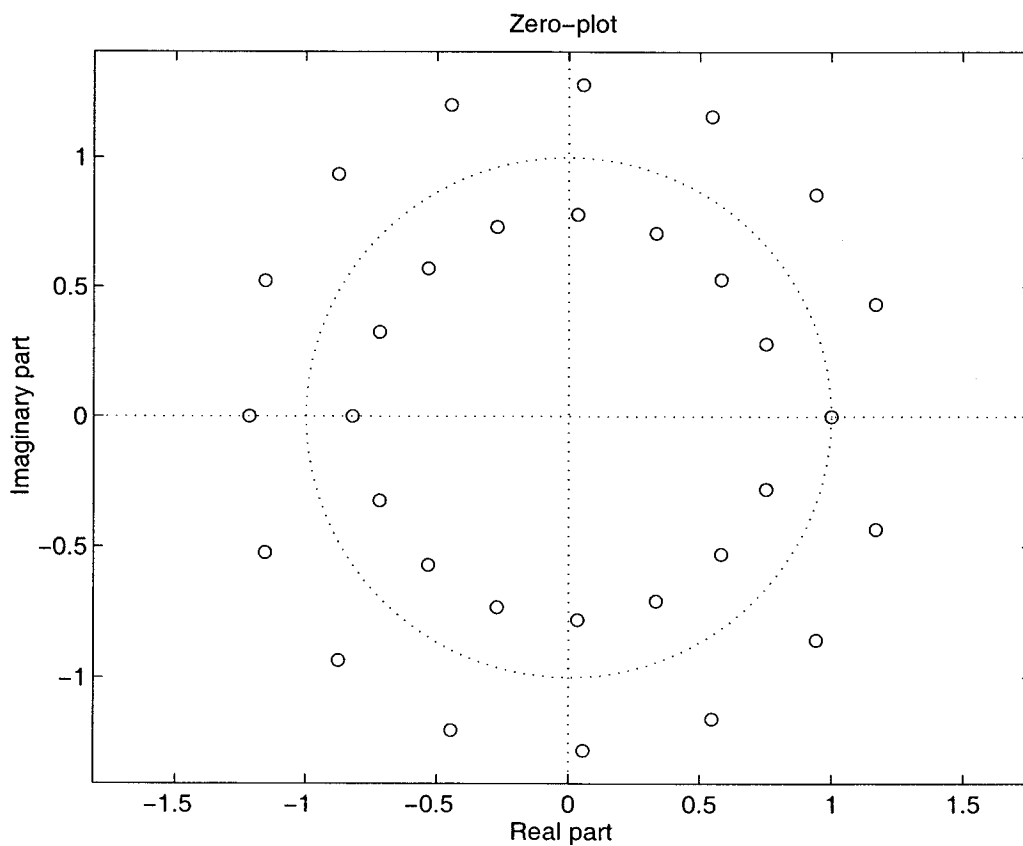
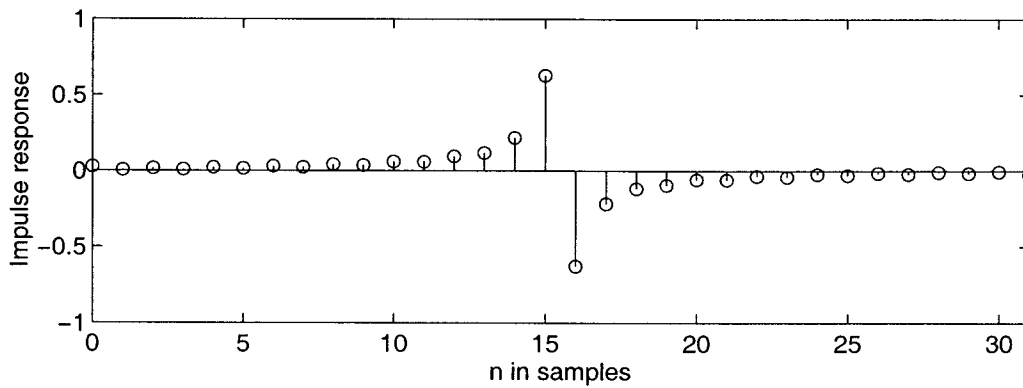
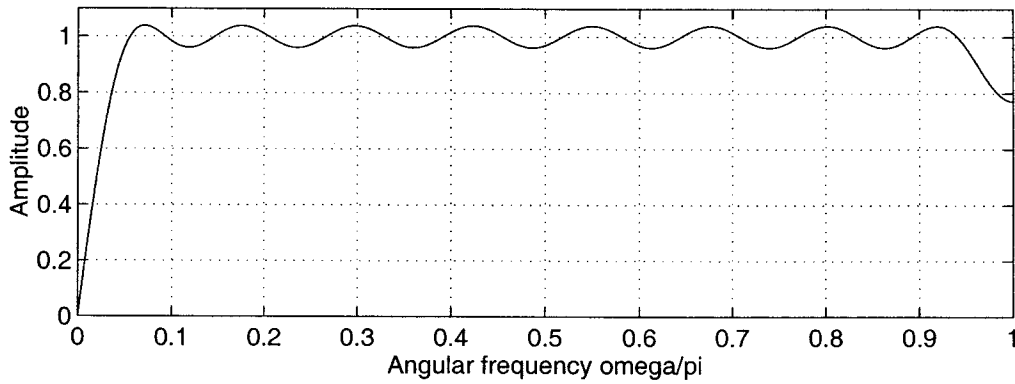


Example 5: Matlab code

79

```
%Design of an FIR Hilbert transformer
%of order 30 (Type III) using
%the Remez algorithm with edges at 0.05
%and 0.95*pi. The zero-phase response is
%desired approximate unity in this
%band.
%Tapio Saramaki 28.10.1995
%This can be found in SUN's
%~ts/matlab/dsp/luefir5.m
f=[.05 .95]; m=[1 1];
n=30;
h=remez(n,f,m,'Hilbert');
h=-h;
%the algorithm gives a wrong sign
[H,w]=zeroam(h,.0,1.,4000);
%zeroam is a routine evaluating
%the zero-phase frequency response;
%programmed by Tapio Saramaki
%This can be found in SUN's
%~ts/matlab/dsp/zeroam.m
figure(1)
subplot(2,1,1)
plot(w/pi,(abs(H))/pi,grid);
axis([0 1 0 1.1]);
ylabel('Amplitude');
xlabel('Angular frequency omega/pi');
subplot(2,1,2)
impz(h);xlabel('n in samples');
ylabel('Impulse response');
figure(2)
zplane(h);title('Zero-plot')
[H,w]=zeroam(h,.05,.95,4000);
figure(3);
plot(w/pi,H-1);grid;axis([.05 .95 -.05 .05]);
title('Error in the interval [0.05pi, 0.95pi]')
xlabel('Angular frequency omega/pi');
ylabel('Error');
```

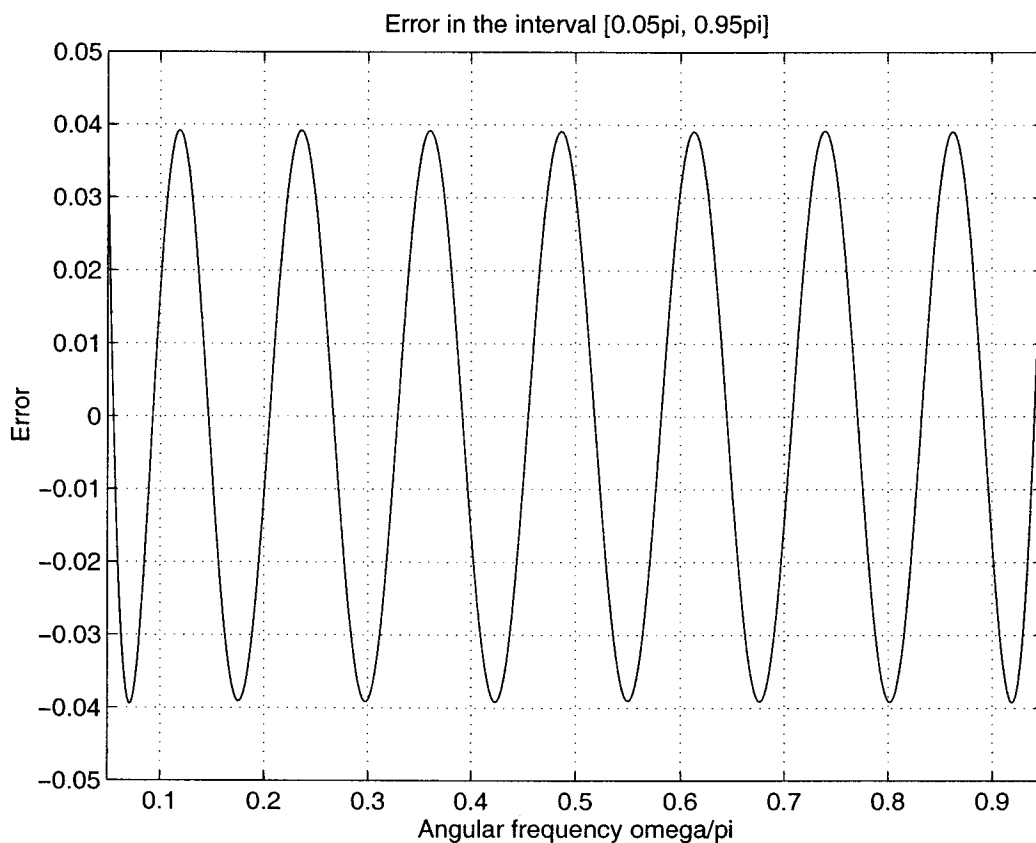
EXAMPLE 6: Type IV Hilbert transformer, $N = 31$



EXAMPLE 6: Error function

- The filter coefficients have been determined to minimize on $[0.05\pi, 0.95\pi]$ the peak absolute value of the error function

$$E(\omega) = [H(\omega) - 1].$$



Example 6: Matlab code

82

```
%Design of an FIR Hilbert transformer
%of order 31 (Type IV) using
%the Remez algorithm with edges at 0.05
%and 0.95*pi. The zero-phase response is
%desired approximate unity in this
%band.
%Tapio Saramaki 28.10.1995
%This can be found in SUN's
%~ts/matlab/dsp/luefir6.m
f=[.05 .95]; m=[1 1];
n=31;
h=remez(n,f,m,'Hilbert');
h=-h;
%the algorithm gives a wrong sign
[H,w]=zeroam(h,.0,1.,4000);
%zeroam is a routine evaluating
%the zero-phase frequency response;
%programmed by Tapio Saramaki
%This can be found in SUN's
%~ts/matlab/dsp/zeroam.m
figure(1)
subplot(2,1,1)
plot(w/pi,(abs(H))),grid;
axis([0 1 0 1.1])
ylabel('Amplitude');
xlabel('Angular frequency omega/pi');
subplot(2,1,2)
impz(h);xlabel('n in samples');
ylabel('Impulse response');
figure(2)
zplane(h);title('Zero-plot')
[H,w]=zeroam(h,.05,.95,4000);
figure(3);
plot(w/pi,H-1);grid;axis([.05 .95 -.05 .05]);
title('Error in the interval [0.05pi, 0.95pi]')
xlabel('Angular frequency omega/pi');
ylabel('Error');
```

SECTION 3: DESIGN OF LINEAR-PHASE FIR FILTERS BY WINDOWING

- The most straightforward approach to designing FIR filters is to determine the infinite-duration impulse response by expanding the frequency response of an ideal filter in a Fourier series and then to truncate and smooth this response using a window function.
- **The main advantage** of this design technique is that the impulse-response coefficients can be obtained in closed form and can be determined very fast even using a calculator.
- **The main drawback** is that the passband and stopband ripples of the resulting filter are restricted to be approximately equal.

ORGANIZATION OF THIS SECTION

- For simplicity, we start by designing Type I linear-phase filters and concentrate mainly on the lowpass case.
- For design purposes, first fixed window functions (the only adjustable variable is the window length) are used and then adjustable windows having a changeable variable are introduced.
- Finally, the windowing technique is generalized for designing also Type II, III, and IV filters.
- Also, four general-purpose matlab routines based on the use of the windowing technique are given in appendices in the end of this chapter.

BASIC DESIGN PROCESS

Step 1: Specify the ideal Type I zero-phase frequency response $H_{id}(\omega)$ that is even around $\omega = 0$ and has the periodicity of 2π .

- The ideal responses for the lowpass, highpass, bandpass, and bandstop cases are shown in the next transparency.
 - Note that there is no transition band(s) for the ideal responses. Only the cutoff frequency(ies) between the passband(s) and stopband(s) are given: ω_c for the lowpass and highpass cases and ω_{c1} and ω_{c2} for the bandpass and bandstop cases.

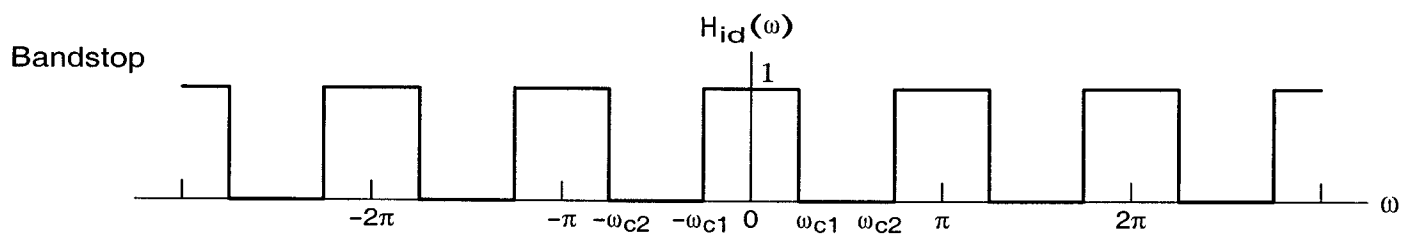
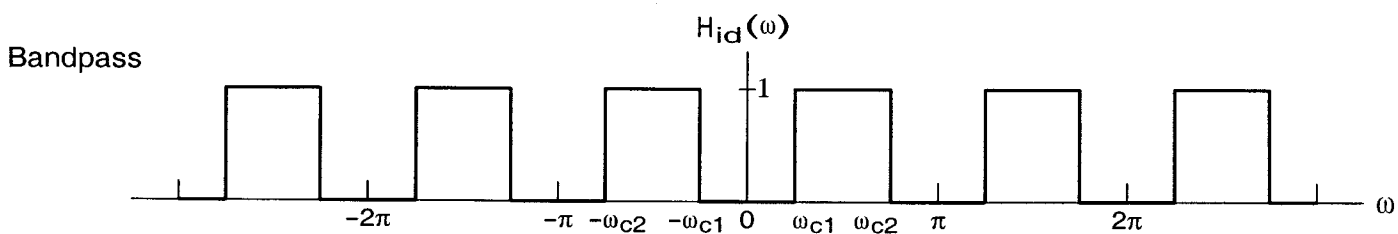
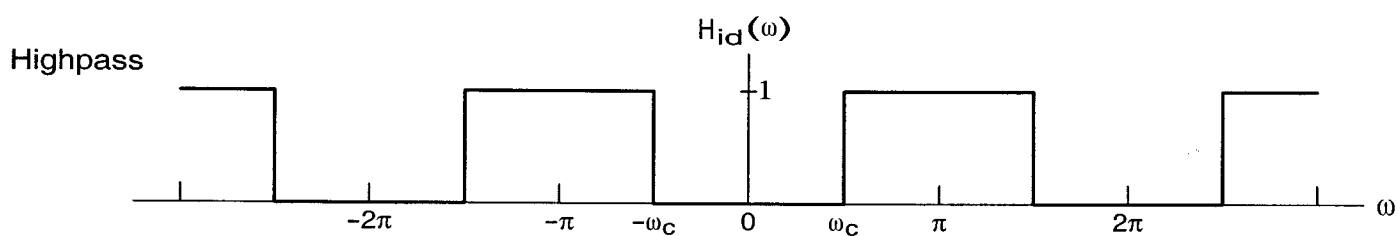
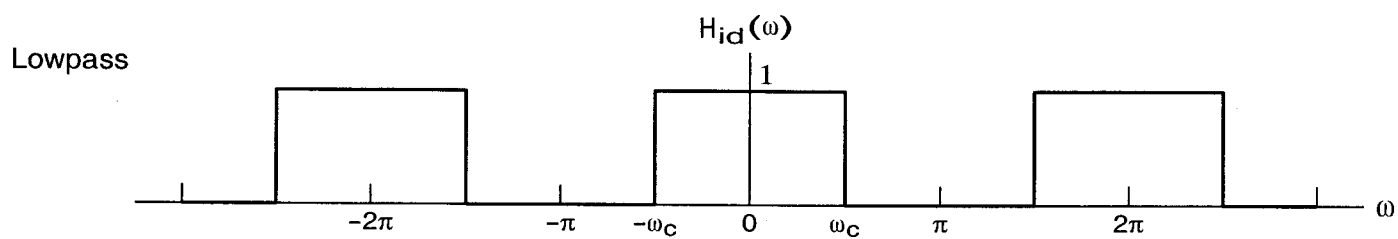
Step 2: Expand $H_{id}(\omega)$ in a Fourier series

$$H_{id}(\omega) = h_{id}^{(0)}[n] + 2 \sum_{n=1}^{\infty} h_{id}^{(0)}[n] \cos n\omega,$$

where

$$h_{id}^{(0)}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{id}(\omega) \cos(n\omega) d\omega, \quad 0 \leq n < \infty.$$

IDEAL TYPE I ZERO-PHASE FREQUENCY RESPONSES



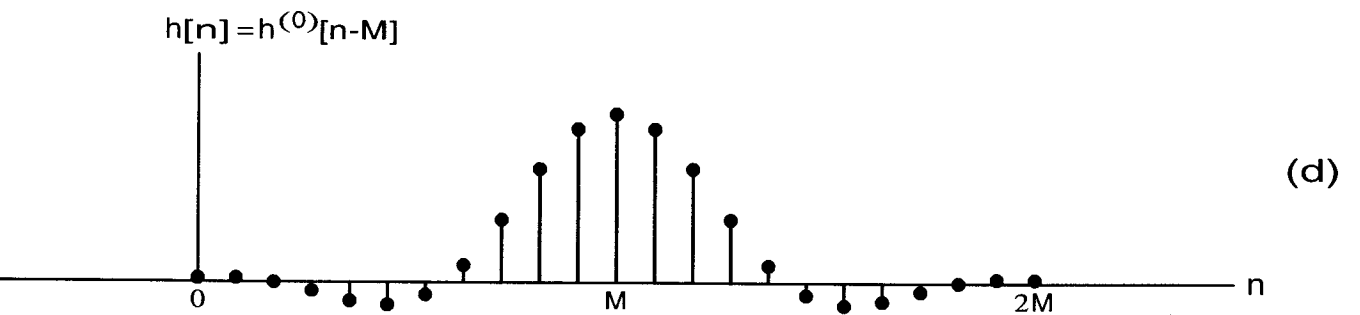
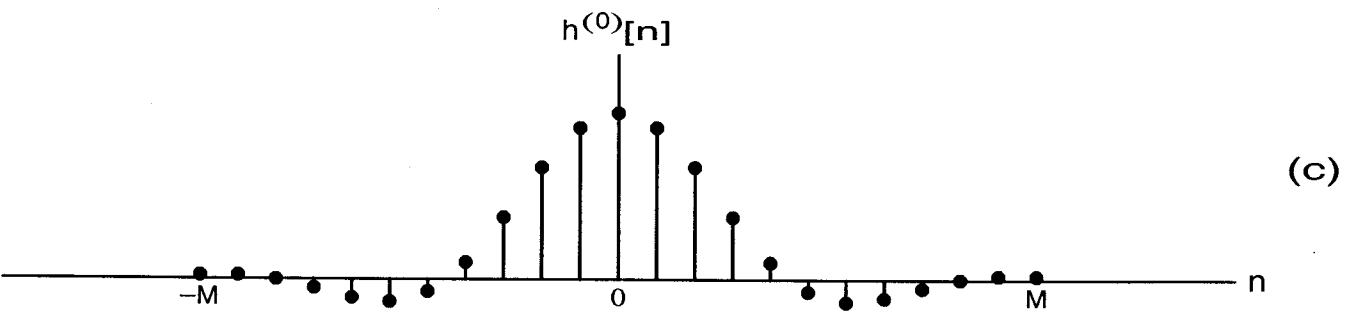
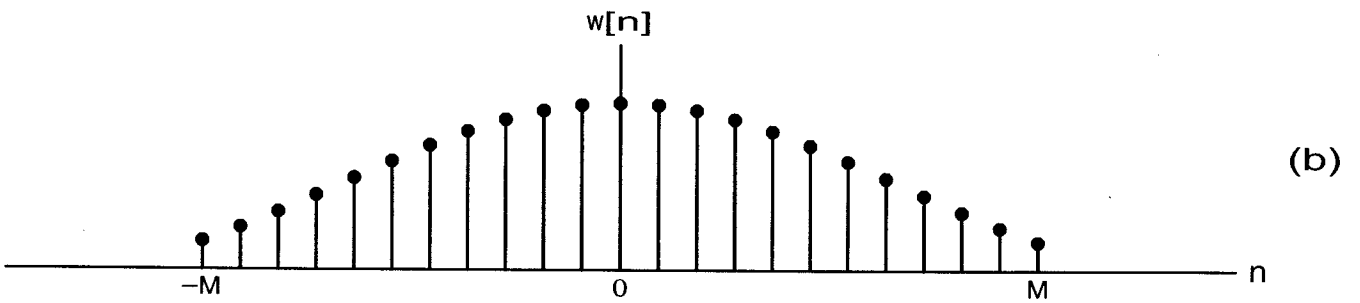
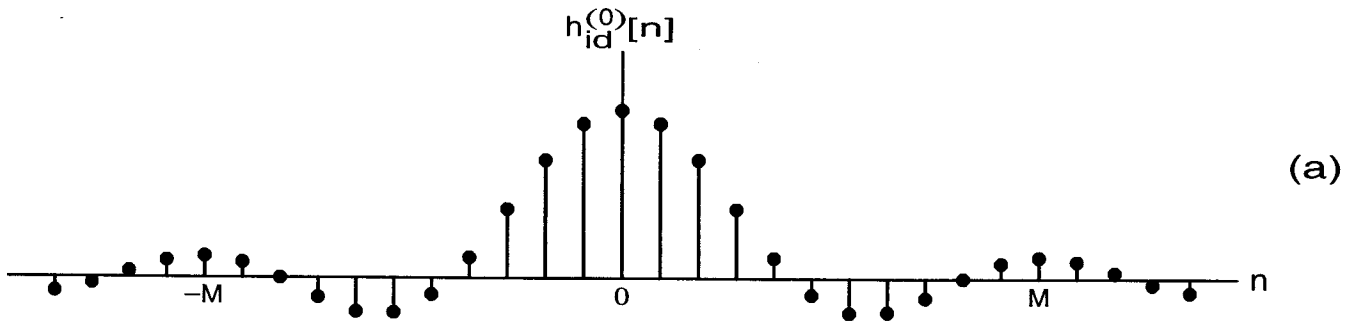
- The Fourier series contain only cosine terms since $H_{\text{id}}(\omega)$ has the periodicity of 2π and is even around $\omega = 0$.
- The above series for $H_{\text{id}}(\omega)$ can be interpreted as the frequency response of the ideal infinite-duration filter that has the impulse response values $h_{\text{id}}^{(0)}[n]$ for $n \geq 0$ and $h_{\text{id}}^{(0)}[n] = h_{\text{id}}^{(0)}[-n]$ for $n < 0$, as shown in the next transparency.
 - For this unrealizable uncausal filter, the infinite-duration impulse response is symmetric around $n = 0$.
 - The corresponding transfer function is given by

$$H_{\text{id}}^{(0)}(z) = h_{\text{id}}^{(0)}[0] + \sum_{n=1}^{\infty} h_{\text{id}}^{(0)}[n][z^{-n} + z^n],$$

from which the above Fourier series is obtained by using the substitution $z = e^{j\omega}$ and the identity $e^{jn\omega} + e^{-jn\omega} = 2 \cos n\omega$.

- The $h_{\text{id}}^{(0)}[n]$'s for the ideal lowpass, highpass, bandpass, and bandstop cases are given in transparency 89.

DESIGN PROCESS IN THE TIME DOMAIN



COEFFICIENT FOR IDEAL ZERO-PHASE TYPE I FILTERS

Lowpass filter with edge at $\omega = \omega_c$:

$$h_{\text{id}}^{(0)}[n] = \begin{cases} \omega_c/\pi, & n = 0 \\ \sin(\omega_c n)/(\pi n) & |n| > 0. \end{cases}$$

Highpass filter with edge at $\omega = \omega_c$:

$$h_{\text{id}}^{(0)}[n] = \begin{cases} 1 - \omega_c/\pi, & n = 0 \\ -\sin(\omega_c n)/(\pi n) & |n| > 0. \end{cases}$$

Bandpass filter with edges at $\omega = \omega_{c1}$ and
 $\omega = \omega_{c2}$:

$$h_{\text{id}}^{(0)}[n] = \begin{cases} (\omega_{c2} - \omega_{c1})/\pi, & n = 0 \\ [\sin(\omega_{c2}n) - \sin(\omega_{c1}n)]/(\pi n) & |n| > 0. \end{cases}$$

Bandstop filter with edges at $\omega = \omega_{c1}$ and
 $\omega = \omega_{c2}$:

$$h_{\text{id}}^{(0)}[n] = \begin{cases} 1 - (\omega_{c2} - \omega_{c1})/\pi, & n = 0 \\ [\sin(\omega_{c1}n) - \sin(\omega_{c2}n)]/(\pi n) & |n| > 0. \end{cases}$$

Step 3: Form the coefficients of the approximating finite-duration zero-phase (center of symmetry is at $n = 0$) filter according to

$$h^{(0)}[n] = w[n]h_{id}^{(0)}[n]$$

where $w[n]$ is a **window function** which is nonzero for $-M \leq n \leq M$ (see transparency 88).

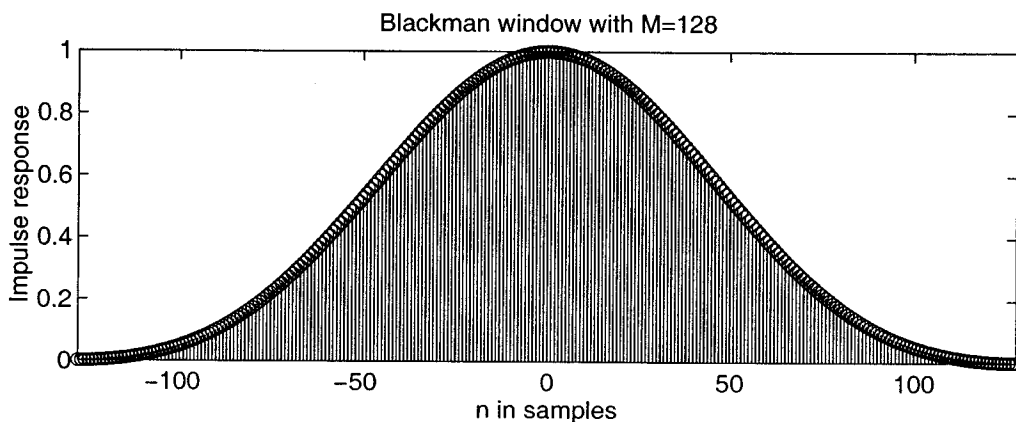
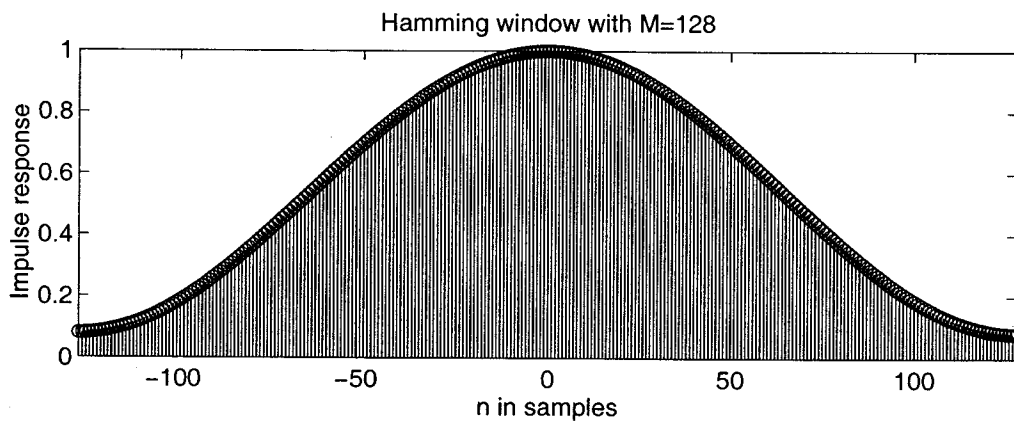
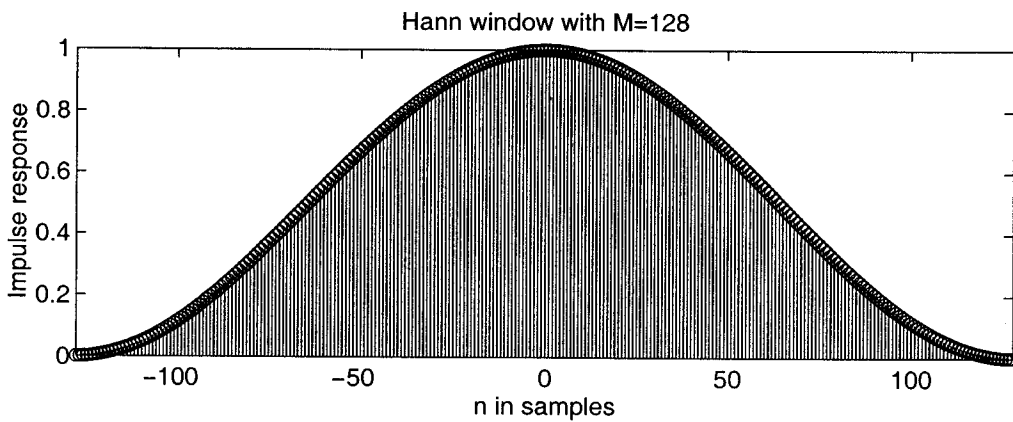
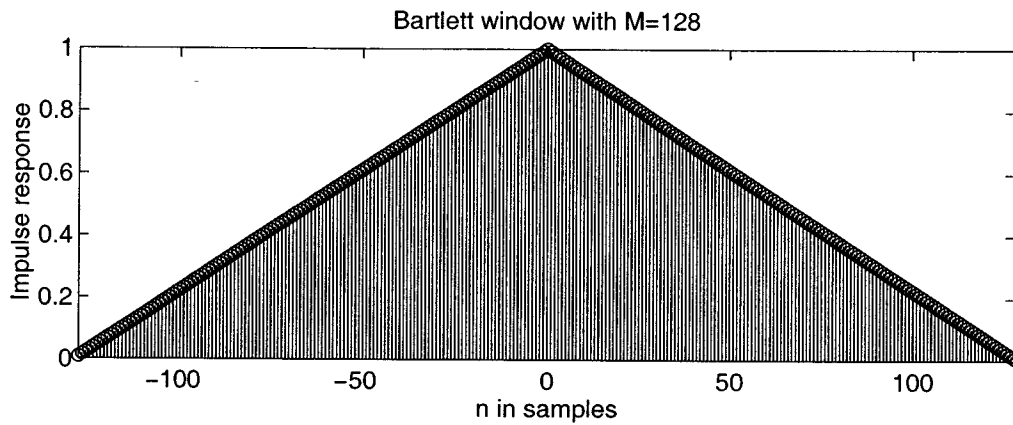
- Some commonly used fixed window functions (the only adjustable parameter is M , as we shall see later) are shown in transparency 91 for $M = 128$. The actual definitions are given later.

Step 4: The coefficients of the causal realizable filter are obtained by shifting the center of symmetry from $n = 0$ to $n = M$ (see transparency 88), yielding

$$h[n] = h^{(0)}[n - M].$$

- Note that the order of the resulting Type I filter is $N = 2M$.

SOME COMMONLY USED WINDOW FUNCTIONS FOR $M = 128$



DIRECT TRUNCATION

- When using the **rectangular window** defined by

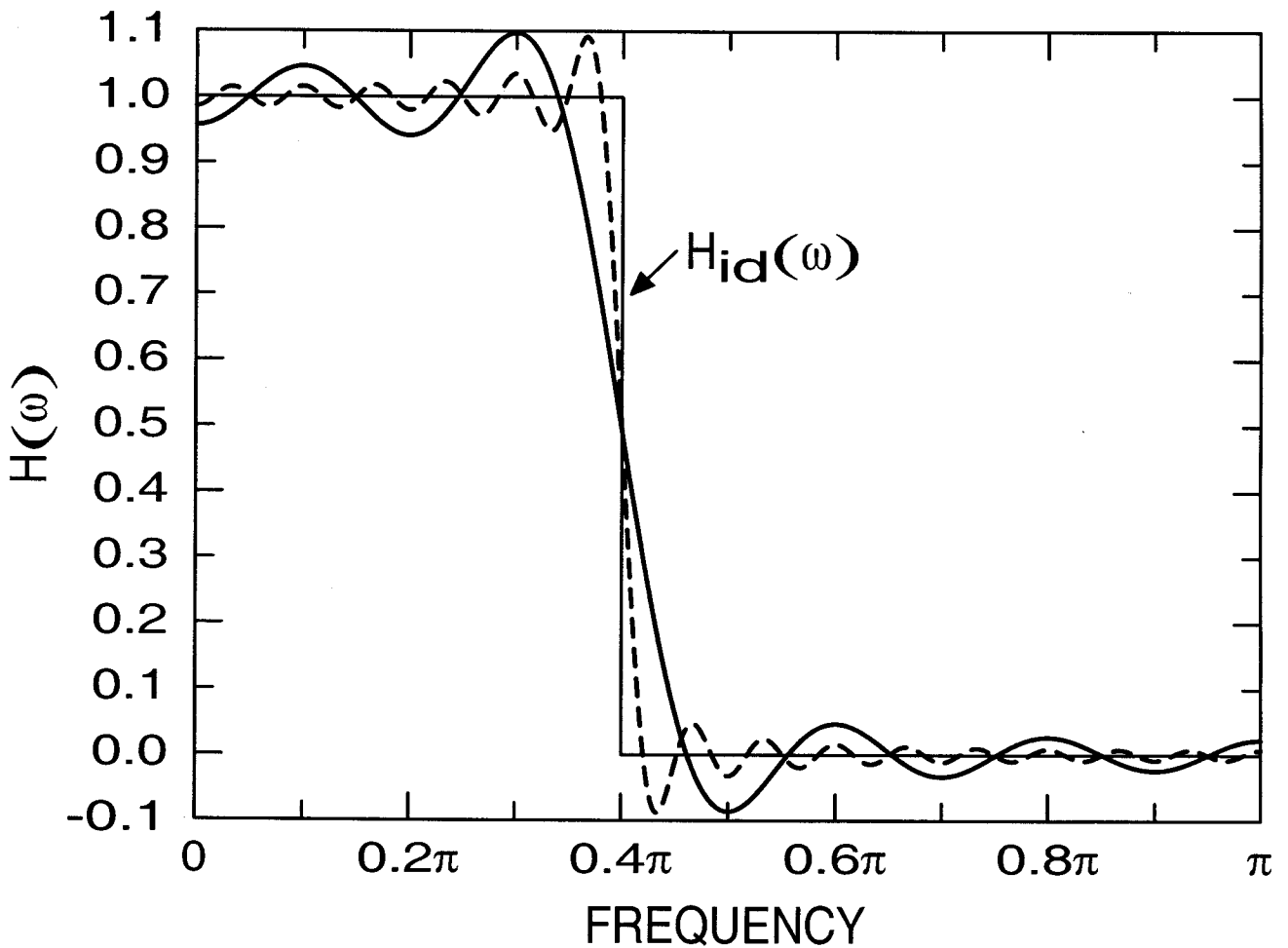
$$w[n] = \begin{cases} 1, & -M \leq n \leq M \\ 0, & \text{otherwise,} \end{cases}$$

we obtain

$$h^{(0)}[n] = \begin{cases} h_{id}^{(0)}[n], & -M \leq n \leq M \\ 0, & \text{otherwise.} \end{cases}$$

- This corresponds to the direct truncation of the ideal impulse response and leads to the well-known **Gibbs phenomenon**
- This phenomenon means that the response of the resulting filter exhibits large ripples before and after the discontinuity of the ideal response independent of the value of M .
- This phenomenon is illustrated in the next transparency that shows the resulting zero-phase frequency response $H(\omega)$ in the lowpass case with $\omega_c = 0.4\pi$ for $M = 10$ and $M = 30$.
- As M is increased, the transition bandwidth of $H(\omega)$ becomes narrower.
- However, for both cases the passband maximum and stopband minimum are approximately the same (1.09 and -0.09 , respectively).

Responses for Type I Lowpass Filters Designed Using the Rectangular Window. $\omega_c = 0.4\pi$. Solid and dashed lines are for $M = 10$ and $M = 30$, respectively.



Explanation of the Gibbs Phenomenon

- The Gibbs phenomenon can be explained by the fact that $H(\omega)$ is related to the ideal response $H_{\text{id}}(\omega)$ and the frequency response of the window function as given by

$$\Psi(\omega) = \sum_{n=-M}^M w[n]e^{-jn\omega}$$

through

$$H(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{\text{id}}(\theta) \Psi(\omega - \theta) d\theta. \quad (A)$$

- Since all the commonly used window functions satisfy $w[0] = 1$ and $w[-n] = w[n]$, the above $\Psi(\omega)$ is expressible as

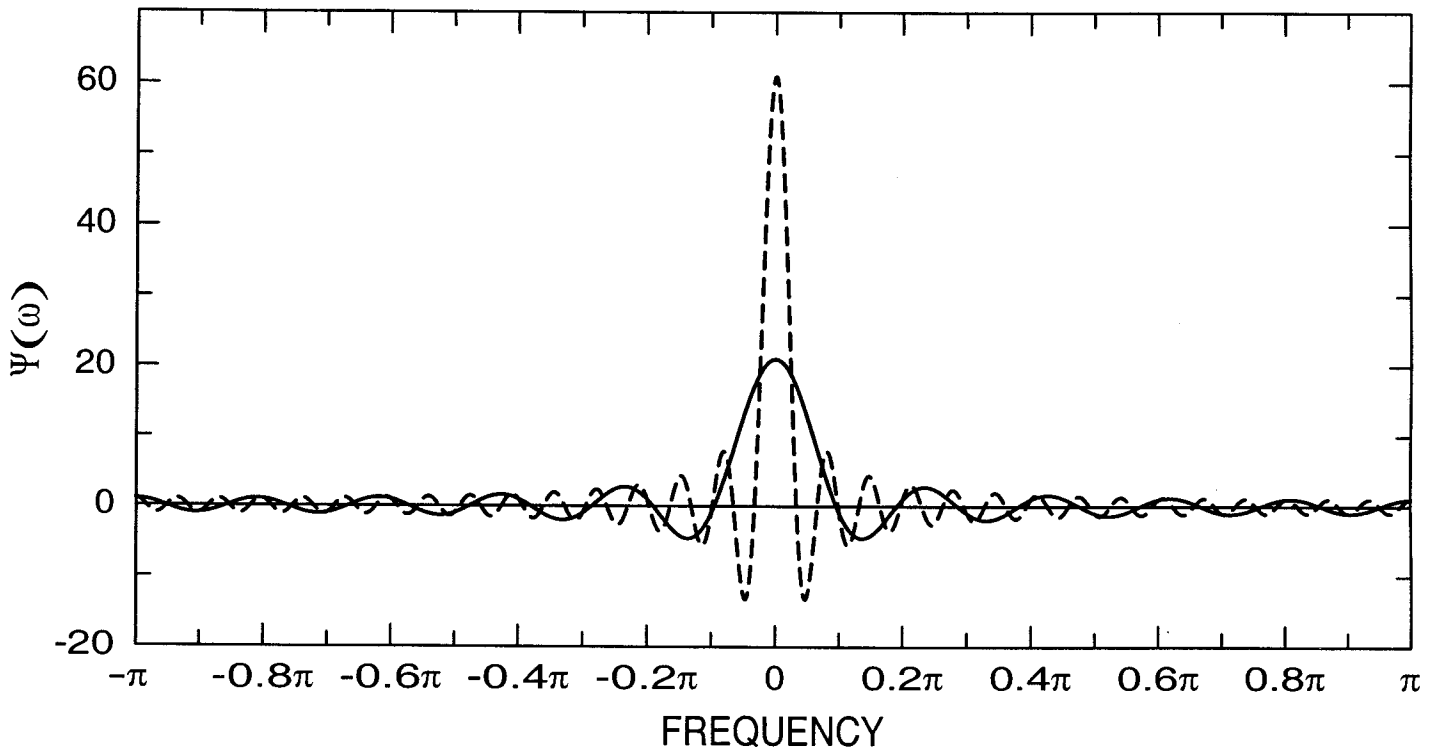
$$\Psi(\omega) = 1 + 2 \sum_{n=1}^M w[n] \cos n\omega.$$

- For the rectangular window,

$$\begin{aligned} \Psi(\omega) &= \sum_{n=-M}^M w[n]e^{-jn\omega} = \sum_{n=-M}^M e^{-jn\omega} \\ &= \frac{\sin[(2M+1)\omega/2]}{\sin(\omega/2)}. \end{aligned}$$

- The following transparency shows $\Psi(\omega)$ for $M = 10$ and $M = 30$.

Frequency responses for the rectangular window for $M = 10$ (solid line) and $M = 30$ (dashed line)

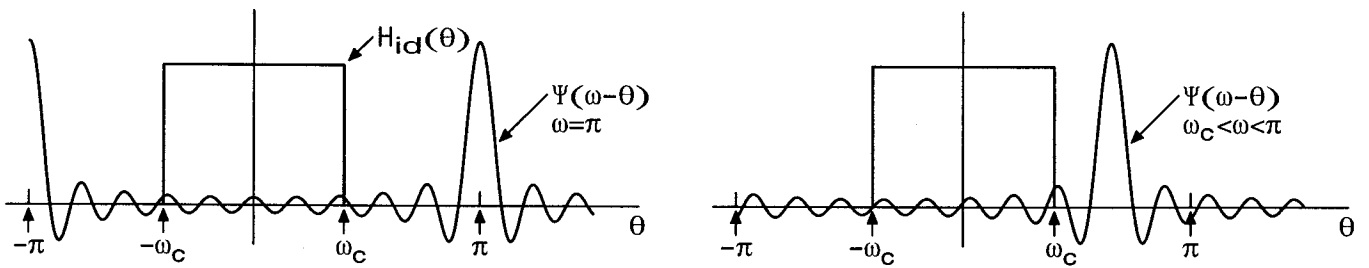


- As seen from the above figure, $\Psi(\omega)$ appears as a gradually decaying sinusoid.
- For later use, we give the following definitions:
- The **mainlobe** is the part of the frequency response situated around $\omega = 0$ between the points where $\Psi(\omega)$ crosses the value of zero.
- The **sidelobes** are the parts of the frequency response situated between two zero-crossings.
- For the rectangular window, the mainlobe width is twice the sidelobe widths that are the same.

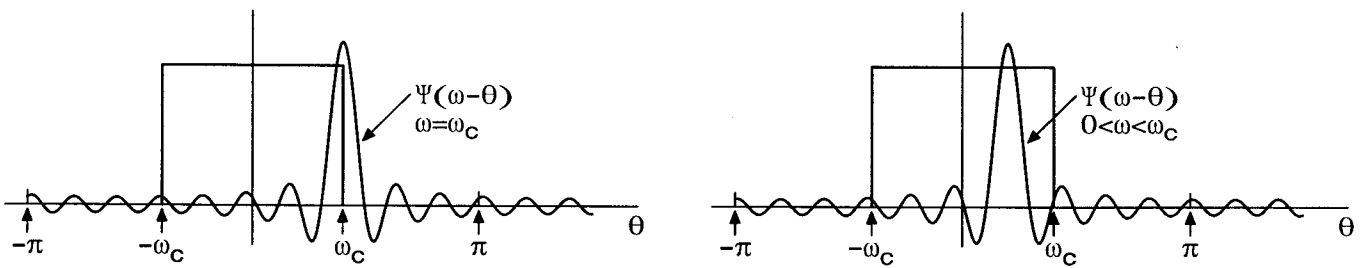
Explanation of the Gibbs phenomenon (continued)

- According to Eq. (A) of transparency 94, the value of $H(\omega)$ at any frequency point ω is obtained in the lowpass case with cutoff edge ω_c by integrating $\Psi(\omega - \theta)$ with respect to θ over the interval $[-\omega_c, \omega_c]$.
- This is the interval on $[-\pi, \pi]$ where $H_{\text{id}}(\omega) = 1$. Elsewhere it is zero.
- The next transparency illustrates the integration process.

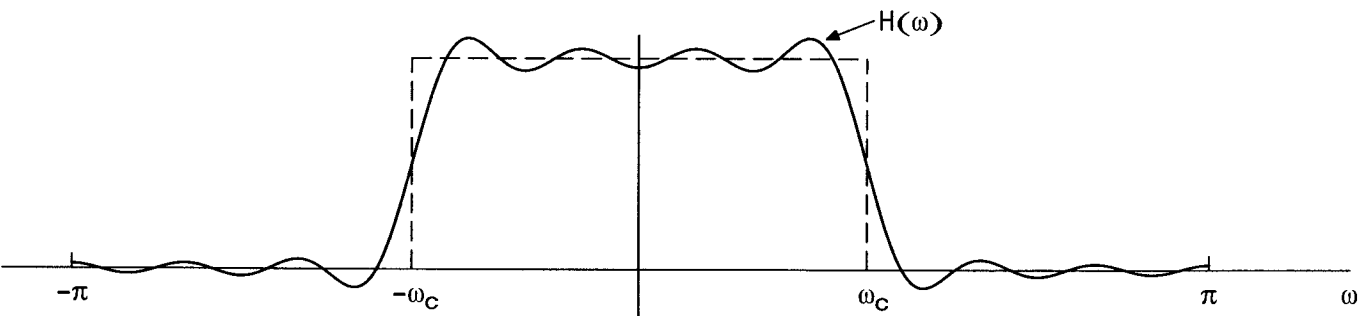
Explanation of the Gibbs phenomenon. (a) Convolution process. (b) Response for the resulting filter.



(a)



(b)



- As seen from the previous transparency, for $\omega = \pi$, only small ripples of $\Psi(\omega - \theta)$ are inside the interval $[-\omega_c, \omega_c]$, resulting in a small value of $H(\omega)$ at $\omega = \pi$.
- As ω is made smaller, larger ripples of $\Psi(\omega - \theta)$ are entering into the interval, resulting in larger values in $H(\omega)$ for $\omega < \pi$.
- The ripples are due to the fact that the area under every second sidelobe of $\Psi(\omega)$ is of opposite sign and the sidelobe heights are different (see transparency 95).
- For $\omega = \omega_c$, half the mainlobe is inside the interval $[-\omega_c, \omega_c]$. Since the integral of $\Psi(\omega)$ over the interval $[-\pi, \pi]$ is one and most of the energy is concentrated in the mainlobe, the value of $H(\omega)$ at $\omega = \omega_c$ is approximately $1/2$.
- When ω is further decreased, the whole mainlobe enters the interval and the area in this interval is approximately one, resulting in the passband response of $H(\omega)$.
- The ripples around one are due to fact that

the sidelobes of $\Psi(\omega - \theta)$, which are of different heights and of different signs, go inside the interval $[-\omega_c, \omega_c]$ and leave it as ω varies.

- As M is increased, the widths of the mainlobe and the sidelobes decrease.
- However, the area under each lobe remains the same since at the same time the heights of the lobes increase (see transparency 95).
- This means that as M is increased, the oscillations of the resulting filter response occur more rapidly but do not decrease.
- In summary:
 - As M is increased, the energy in the sidelobes remains the same.
 - As M is increased, the oscillations of $H(\omega)$ occur more rapidly but do not decrease (see transparency 93).

FIXED WINDOW FUNCTIONS

- The Gibbs phenomenon can be reduced by using a less abrupt truncation of the Fourier series.
- This is achieved by smoothing the coefficients of the ideal filter range $-M \leq n \leq M$ using a window that tapers smoothly to zero at both ends.
- Some of the well-known fixed window functions $w[n]$ are summarized in the first table of the next transparency along with their frequency responses $\Psi(\omega)$.
- These windows are called fixed since the only adjustable parameter is M , half the order ($N = 2M$) of the resulting filter.
- The second table of this transparency shows some characteristics of these windows as well as those of the resulting filters (to be considered later).
- The earlier transparency 91 showed the window functions for $M = 128$, whereas transparency 102 shows the frequency responses of these window functions for $M = 128$ as well as those of the resulting filters for $\omega_c = 0.4\pi$.

SOME FIXED WINDOWS

Table 1: Some Fixed Windows

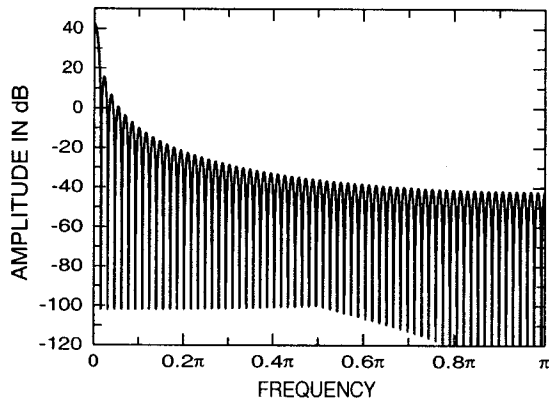
Window Type	Window Function, $w[n], -M \leq n \leq M$	Frequency Response, $\Psi(\omega)$
Rectangular	1	$\Psi_R(\omega) \equiv \sin[(2M+1)\omega/2]/\sin(\omega/2)$
Bartlett	$1 - \frac{ n }{M+1}$	$[\sin[(M+1)\omega/2]/\sin(\omega/2)]^2$
Hann	$\frac{1}{2} \left[1 + \cos\left[\frac{2\pi n}{2M+1}\right] \right]$	$0.5\Psi_R(\omega) + 0.25\Psi_R\left(\omega - \frac{2\pi}{2M+1}\right) + 0.25\Psi_R\left(\omega + \frac{2\pi}{2M+1}\right)$
Hamming	$0.54 + 0.46 \cos\left[\frac{2\pi n}{2M+1}\right]$	$0.54\Psi_R(\omega) + 0.23\Psi_R\left(\omega - \frac{2\pi}{2M+1}\right) + 0.23\Psi_R\left(\omega + \frac{2\pi}{2M+1}\right)$
Blackman	$0.42 + 0.5 \cos\left[\frac{2\pi n}{2M+1}\right] + 0.08 \cos\left[\frac{4\pi n}{2M+1}\right]$	$0.42\Psi_R(\omega) + 0.25\Psi_R\left(\omega - \frac{2\pi}{2M+1}\right) + 0.25\Psi_R\left(\omega + \frac{2\pi}{2M+1}\right) + 0.04\Psi_R\left(\omega - \frac{4\pi}{2M+1}\right) + 0.04\Psi_R\left(\omega + \frac{4\pi}{2M+1}\right)$

Table 2: Properties of Some Fixed Windows

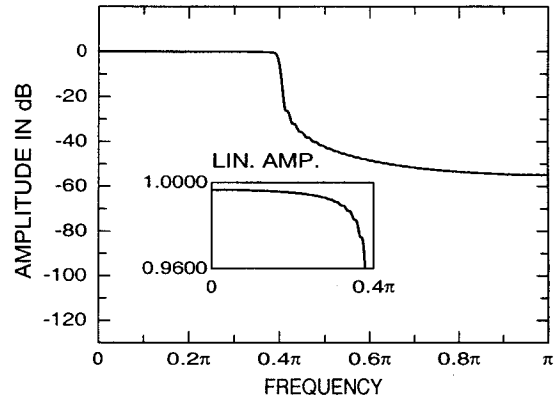
Window Type	Mainlobe Width Δ_M	Sidelobe Ripple	A_s	$\Delta\omega = \omega_s - \omega_p$
Rectangular	$\frac{4\pi}{2M+1}$	-13.3 dB	20.9 dB	$1.84\pi/(2M)$
Bartlett	$\frac{4\pi}{M+1}$	-26.5 dB	see text	see text
Hann	$\frac{8\pi}{2M+1}$	-31.5 dB	43.9 dB	$6.22\pi/(2M)$
Hamming	$\frac{8\pi}{2M+1}$	-44.0 dB	54.5 dB	$6.64\pi/(2M)$
Blackman	$\frac{12\pi}{2M+1}$	-58.1 dB	75.3 dB	$11.13\pi/(2M)$

Frequency Responses for Fixed Window Functions and the Resulting Filters for $M = 128$ and $\omega_c = 0.4\pi$.

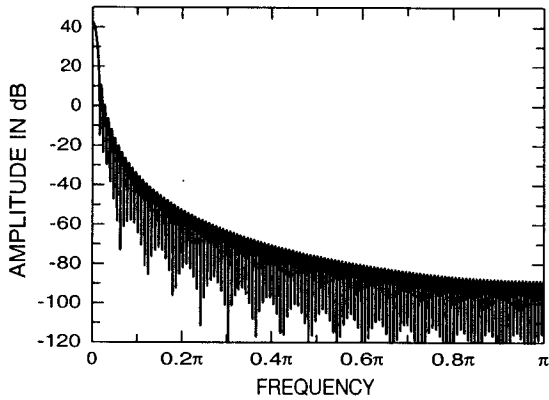
(a,b) Bartlett Window. (c,d) Hann Window. (e,f) Hamming Window. (g,h) Blackmann Window.



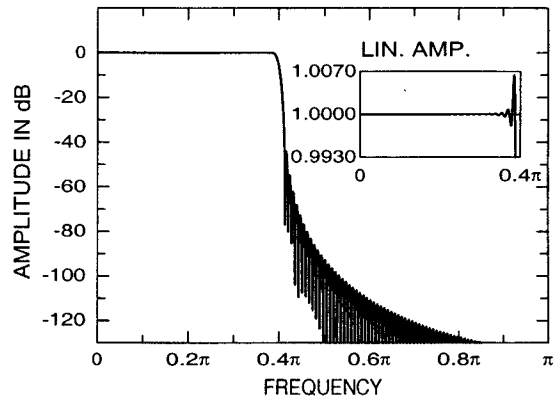
(a)



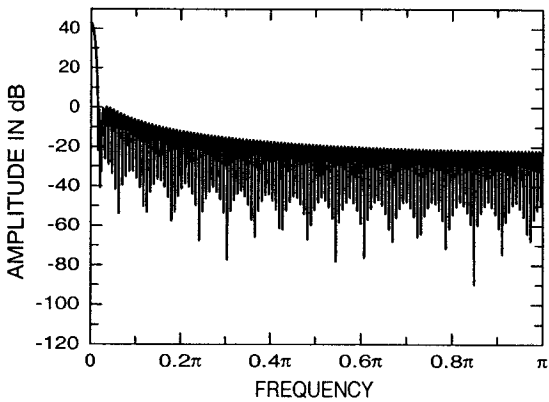
(b)



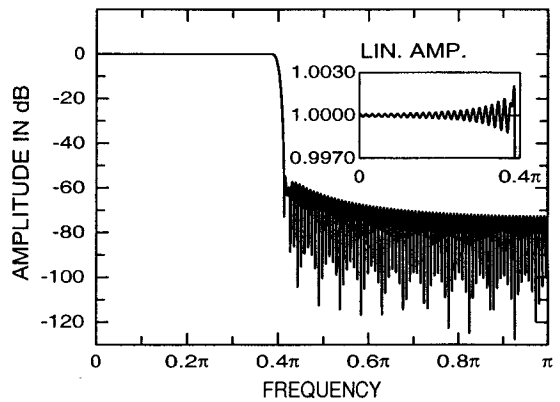
(c)



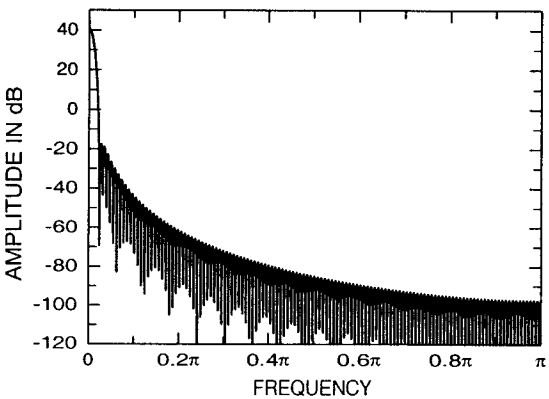
(d)



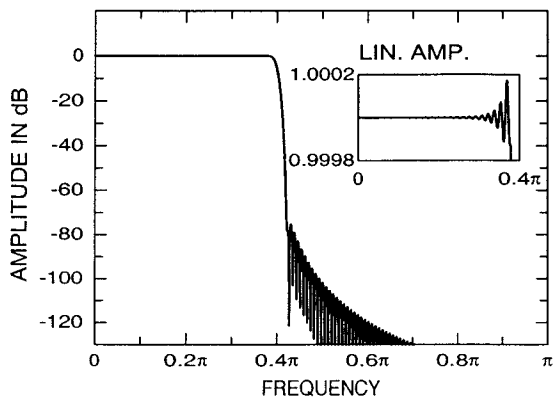
(e)



(f)



(g)

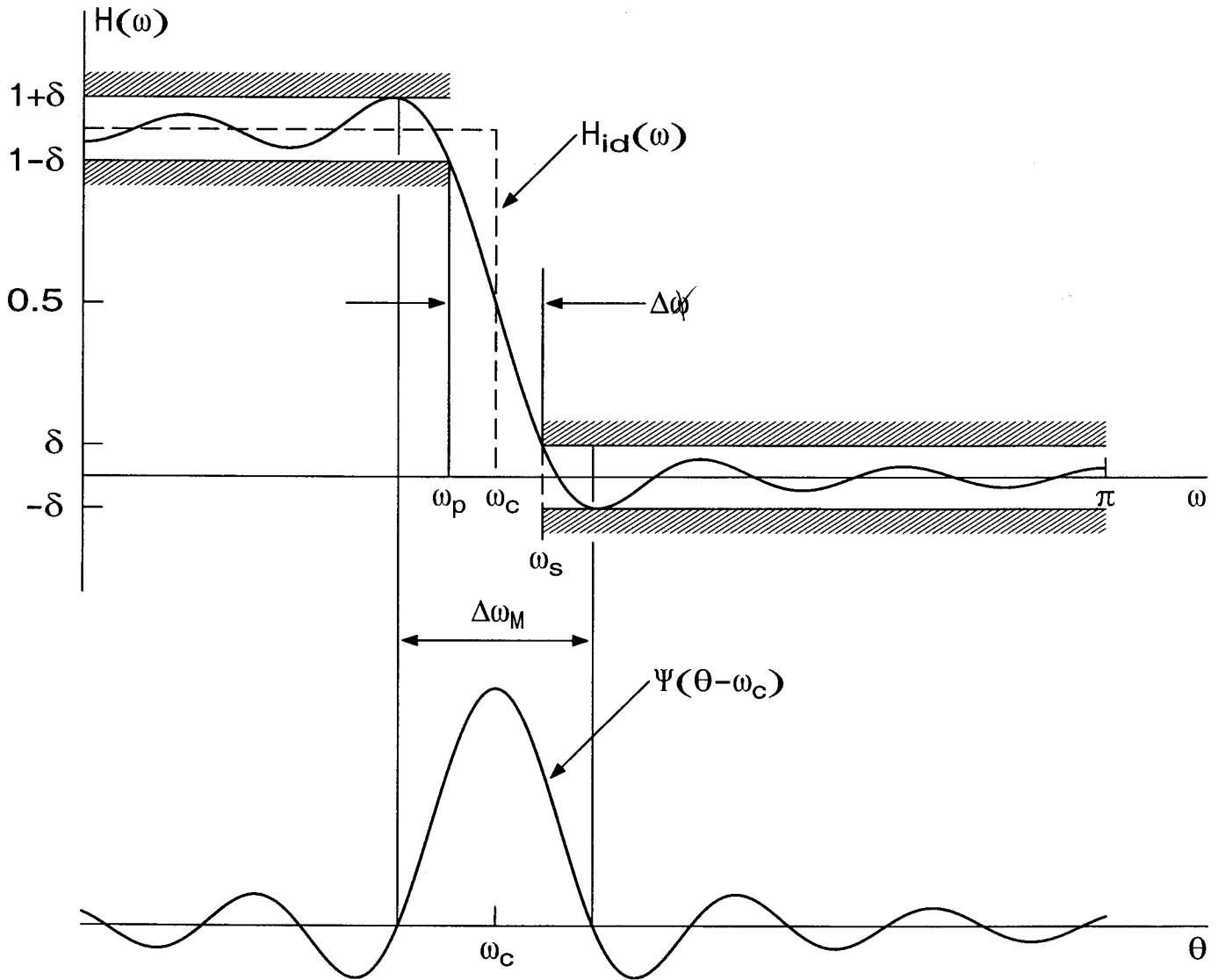


(h)

PROPERTIES OF FIXED WINDOWS

- The following transparency depicts, in the low-pass case, a typical relation between $H(\omega)$ and $\Psi(\omega)$, which is given in terms of $\theta - \omega_c$ in order to center the response at the cutoff edge.
- Notice a close similarity to the case where $\Psi(\omega)$, $H_{\text{id}}(\omega)$, and $H(\omega)$ correspond to the impulse response, the step excitation, and the response of a continuous-time filter, respectively.
- As seen from this transparency, $H(\omega)$ satisfies approximately $H(\omega_c + \omega) + H(\omega_c - \omega) = 1$ in the vicinity of the cutoff edge ω_c .
- This means that $H(\omega_c) \approx 1/2$.
- Furthermore, the maximum passband deviation from unity and the maximum stopband deviation from zero are about the same, and the peak passband overshoot $(1 + \delta)$ and the peak negative stopband undershoot $(-\delta)$ occur at the same distance from the discontinuity point ω_c .
- The distance between these two overshoot points is for most windows approximately equal to the mainlobe width Δ_M .

Typical Relations Between the Window Function and the Resulting Filter in the Lowpass Case With Cutoff Edge at ω_c



CRITERIA MET BY THE RESULTING FILTER

- The criteria met by $H(\omega)$ can be given by

$$1 - \delta \leq H(\omega) \leq 1 + \delta \quad \text{for } \omega \in [0, \omega_p]$$

$$-\delta \leq H(\omega) \leq \delta \quad \text{for } \omega \in [\omega_s, \pi].$$

- Here, ω_p (ω_s) is defined to be the highest frequency where $H(\omega) \geq 1 - \delta$ (the lowest frequency where $H(\omega) \leq \delta$) (see the previous transparency).
- The width of the transition band, $\Delta\omega = \omega_s - \omega_p$, is thus less than the mainlobe width Δ_M .
- This means that for a good window function, the mainlobe width has to be as narrow as possible.
- On the other hand, for a small ripple value δ , it is required that the area under the sidelobes of $\Psi(\omega)$ is as small as possible.
- These two requirements contradict each other and the fixed windows of transparency of 101 make proper compromises between these requirements.

PROPERTIES OF FIXED WINDOWS IN NUTSHELL

- The only adjustable parameter is M , half the order of the resulting filter.
- $H(\omega)$ satisfies

$$\begin{aligned} 1 - \delta_p &\leq H(\omega) \leq 1 + \delta_p && \text{for } \omega \in [0, \omega_p] \\ -\delta_s &\leq H(\omega) \leq \delta_s && \text{for } \omega \in [\omega_s, \pi]. \end{aligned}$$

- For each window,

$$\delta_p \approx \delta_s,$$

$$\omega_p \approx \omega_c - \Delta/2,$$

and

$$\omega_s \approx \omega_c + \Delta/2,$$

where

$$\Delta\omega = \omega_s - \omega_p \approx \gamma/(2M)$$

with δ_s and γ being (approximately) constants.

- The ripple values cannot be varied.
- The second table of transparency 101 summarizes the properties of the fixed windows under consideration.
- The relation $\omega_s - \omega_p \approx \gamma/(2M)$ can be seen from the last column of this table.

- The maximum sidelobe ripple for $\Psi(\omega)$ is given in decibels for the case where $\Psi(\omega)$ is normalized to achieve the value of unity at $\omega = 0$.
- $A_s = -20 \cdot \log_{10}(\delta)$ is the minimum stopband attenuation of the resulting filter.
- These values have been determined for the case $\omega_c = 0.4\pi$ and $M = 128$.
- For the Bartlett window γ and A_s are not given since it is very difficult to locate the stopband edge (see transparency 102).

DESIGN OF LOWPASS FILTERS USING FIXED WINDOWS

- If for one of the fixed windows the stopband attenuation, A_s given in the second table in transparency 101 is satisfactory, then the only adjustable parameters are M , half the filter order, and ω_c , the cutoff edge of the ideal filter.
- If the desired passband and stopband edges are at ω_p and ω_s , then these parameters are selected as follows:

Step I: Select ω_c to be the average of the passband and stopband edges, that is,

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

Step II: Find from the second table of transparency 101 the constant γ in the equation

$$\Delta\omega = \omega_s - \omega_p \approx \gamma/(2M).$$

Step III: Determine the smallest value of M satisfying

$$M \geq \frac{\omega_s - \omega_p}{\gamma/2}.$$

Step IV: Perform the design process described in the beginning of this section.

ADJUSTABLE WINDOWS

- These windows contain an additional parameter with which $\delta_p \approx \delta_s$ can be adjusted.

There exist four adjustable windows:

- Kaiser window
- Saramäki window
- Dolph-Chebyshev window
- Transitional window obtained from the Saramäki and Dolph-Chebyshev windows
- In the following, the basic characteristics of these windows are considered. If you are interested in more details, contact the lecturer.

THE KAISER WINDOW

- This window function is given by

$$w[n] = \begin{cases} I_0[\alpha \sqrt{1 - (\frac{n}{M})^2}] / I_0(\alpha), & -M \leq n \leq M \\ 0, & \text{otherwise.} \end{cases}$$

- M is half the filter order.
- α is the adjustable parameter.
- $I_0(x)$ is the modified zeroth-order Bessel function of the first kind:

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

- For most practical applications, about 20 terms in the above summation are sufficient to arrive at reasonably accurate values of $w[n]$.
- No analytic frequency-domain representation.

LOWPASS TYPE I FILTER DESIGN WITH THE KAISER WINDOW

Given: ω_p , ω_s , $\delta_p \approx \delta_s$, then α , M , and ω_c are determined as follows:

Step I:

$$\alpha = \begin{cases} 0.1102(A_s - 8.7), & A_s > 50 \\ 0.5842(A_s - 21)^{0.4} \\ +0.07886(A_s - 21), & 21 < A_s < 50 \\ 0, & A_s < 21 \end{cases}$$

where

$$A_s = -20 \cdot \log_{10} \delta_s$$

Step II:

$$M = \frac{A_s - 7.95}{14.36(\omega_s - \omega_p)/\pi}$$

Step III: Cutoff frequency of the ideal filter is

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

- The design process described in the beginning of this section yields a filter meeting approximately (very closely) the given criteria.

THE SARAMÄKI WINDOW

- The frequency response for the unscaled window ($\widehat{w}[0] \neq 1$) is given by

$$\begin{aligned}\widehat{\Psi}(\omega) &= \sum_{n=-M}^M \widehat{w}[n]e^{-jn\omega} = 1 + \sum_{k=1}^M 2T_k[\gamma \cos \omega + (\gamma - 1)] \\ &= \frac{\sin\left[\frac{2M+1}{2} \cos^{-1}\{\gamma \cos \omega + (\gamma - 1)\}\right]}{\sin\left[\frac{1}{2} \cos^{-1}\{\gamma \cos \omega + (\gamma - 1)\}\right]},\end{aligned}$$

where

$$\gamma = \frac{1 + \cos \frac{2\pi}{2M+1}}{1 + \cos \frac{2\beta\pi}{2M+1}}.$$

- β is the adjustable parameter.
- The mainlobe width is $4\beta\pi/(2M+1)$ that is β times that of the rectangular window.

SCALED RESPONSES ($\widehat{w}[0] = 1$)

- The scaled window and the corresponding frequency response are, respectively, given by

$$w[n] = \begin{cases} \widehat{w}[n]/\widehat{w}[0], & -M \leq n \leq M \\ 0, & \text{otherwise} \end{cases}$$

and

$$\Psi(\omega) = \widehat{\Psi}(\omega)/\widehat{w}(0).$$

- The $\widehat{w}(n)$'s can be expressed as

$$\widehat{w}[n] = v_0(n) + 2 \sum_{k=1}^M v_k[n],$$

where the $v_k[n]$'s can be calculated using the recursion relations:

$$v_0[n] = \begin{cases} 1, & n = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$v_1[n] = \begin{cases} \gamma - 1, & n = 0 \\ \gamma/2, & |n| = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$v_k[n] = \begin{cases} 2(\gamma - 1)v_{k-1}[n] - v_{k-2}[n] \\ +\gamma[v_{k-1}[n-1] + v_{k-1}[n+1]], & -k \leq n \leq k \\ 0, & \text{otherwise.} \end{cases}$$

FORMULAS FOR FILTER DESIGN

- Given ω_p , ω_s , and $\delta_p \approx \delta_s$ for the lowpass Type I filter, β , M , and ω_c are determined like for the Kaiser window.
- Empirical formulas for estimating β and M for the above Saramäki window are given in the table of the following transparency.
- This table gives the corresponding formulas also for the Dolph-Chebyshev and transitional windows to be considered later as well as for the Kaiser window.
- It should be pointed out that the estimation formulas for the Dolph-Chebyshev and transitional windows are not so accurate as they are for the Kaiser and Saramäki windows.

Estimation Formulas for M and the Adjustable Parameter for Adjustable Windows to Give the Desired Attenuation A_s and the Transition Bandwidth $\omega_s - \omega_p$

Window Type	Adjustable Parameter	M
Kaiser	$\alpha = \begin{cases} 0.1102(A_s - 8.7), & A_s > 50 \\ 0.5842(A_s - 21)^{0.4} \\ +0.07886(A_s - 21), & 21 < A_s < 50 \\ 0, & A_s < 21 \end{cases}$	$M = \frac{A_s - 7.95}{14.36(\omega_s - \omega_p)/\pi}$
Saramäki	$\beta = \begin{cases} 0.000121(A_s - 21)^2 \\ +0.0224(A_s - 21) \\ +1, & 21 \leq A_s \leq 65 \\ 0.033A_s + 0.062, & 65 < A_s \leq 110 \\ 0.0345A_s - 0.097, & A_s > 110 \end{cases}$	$M = \frac{A_s - 8.15}{14.36(\omega_s - \omega_p)/\pi}$
Dolph-Chebyshev	$\beta = \begin{cases} 0.0000769(A_s)^2 \\ +0.0248A_s + 0.330, & A_s \leq 60 \\ 0.0000104(A_s)^2 \\ +0.0328A_s + 0.079, & A_s > 60 \end{cases}$	$M = \frac{1.028A_s - 8.4}{14.36(\omega_s - \omega_p)/\pi}$
Transitional	$\beta = \begin{cases} 0.000154(A_s)^2 \\ +0.0153A_s + 0.465, & A_s \leq 60 \\ 0.0000204(A_s)^2 \\ +0.0303A_s + 0.032, & A_s > 60 \end{cases}$	$M = \frac{1}{14.36(\omega_s - \omega_p)/\pi} \times [0.00036(A_s)^2 + 0.951A_s - 9.4]$

THE DOLPH-CHEBYSHEV WINDOW

- For the mainlobe width of $4\beta\pi/(2M + 1)$ and for the order of $2M$, the frequency response for the unscaled window is given by

$$\widehat{\Psi}(\omega) = T_M[\gamma \cos \omega + (\gamma - 1)],$$

where

$$\gamma = (1 + \cos \frac{\pi}{2M}) / (1 + \cos \frac{2\beta\pi}{2M + 1}).$$

- The unscaled coefficients are

$$\widehat{w}[n] = v_M[n],$$

where the $v_M[n]$'s can be determined using the recursion relations of transparency 113.

- The scaled window function and the corresponding frequency response are, respectively, given by

$$w[n] = \begin{cases} \widehat{w}[n]/\widehat{w}[0], & -M \leq n \leq M \\ 0, & \text{otherwise} \end{cases}$$

and

$$\Psi(\omega) = \widehat{\Psi}(\omega)/\widehat{w}[0].$$

TRANSITIONAL WINDOW

- For the mainlobe width of $4\beta\pi/(2M + 1)$ and for the order of $N = 2M$, the frequency response for the unscaled window is given by

$$\widehat{\Psi}(\omega) = \sum_{n=-M}^M \widehat{w}(n) e^{-jn\omega} = \prod_{k=1}^M (\cos \omega - \cos \omega_k),$$

where

$$\omega_k = \rho \omega_k^{(1)} + (1 - \rho) \omega_k^{(2)}.$$

with

$$\omega_k^{(1)} = 2 \cos^{-1} \left[\frac{\cos[\beta\pi/(2M + 1)]}{\cos[\pi/(2M + 1)]} \cos\left[\frac{k\pi}{2M + 1}\right] \right]$$

and

$$\omega_k^{(2)} = 2 \cos^{-1} \left[\frac{\cos[\beta\pi/(2M + 1)]}{\cos[\pi/(4M)]} \cos\left[\frac{(2k - 1)\pi}{4M}\right] \right].$$

- Here, $\omega_k^{(1)}$ and $\omega_k^{(2)}$ for $k = 1, 2, \dots, M$ are the zero locations of the Saramäki and the Dolph-Chebyshev windows, respectively.
- For $\rho = 1$ and $\rho = 0$, $\widehat{\Psi}(\omega)$ is thus the unscaled frequency response for the Saramäki and the Dolph-Chebyshev window, respectively.
- For this transitional window, $0 < \rho < 1$ is an adjustable parameter in addition to β .

- In most cases,

$$\rho = \begin{cases} 0.4, & A_s \leq 50 \\ 0.5, & 50 < A_s \leq 75 \\ 0.6, & 75 < A_s \end{cases}$$

is a good selection.

- Accurate values for the unscaled window coefficients $\widehat{w}(n)$ are obtained from

$$\widehat{w}[n] = \frac{1}{2M+1} \left[\widehat{\Psi}(0) + 2 \sum_{k=1}^M \widehat{\Psi}\left(\frac{2\pi k}{2M+1}\right) \cos\left(\frac{2\pi nk}{2M+1}\right) \right].$$

- Alternatively, the coefficients can be determined by evaluating $\widehat{\Psi}(\omega)$ at 2^I ($> 2M+1$) equally-spaced frequencies and using the inverse fast Fourier transform (this is used in a matlab file generated by the lecturer for evaluating the transitional window; can be found in Appendix A).

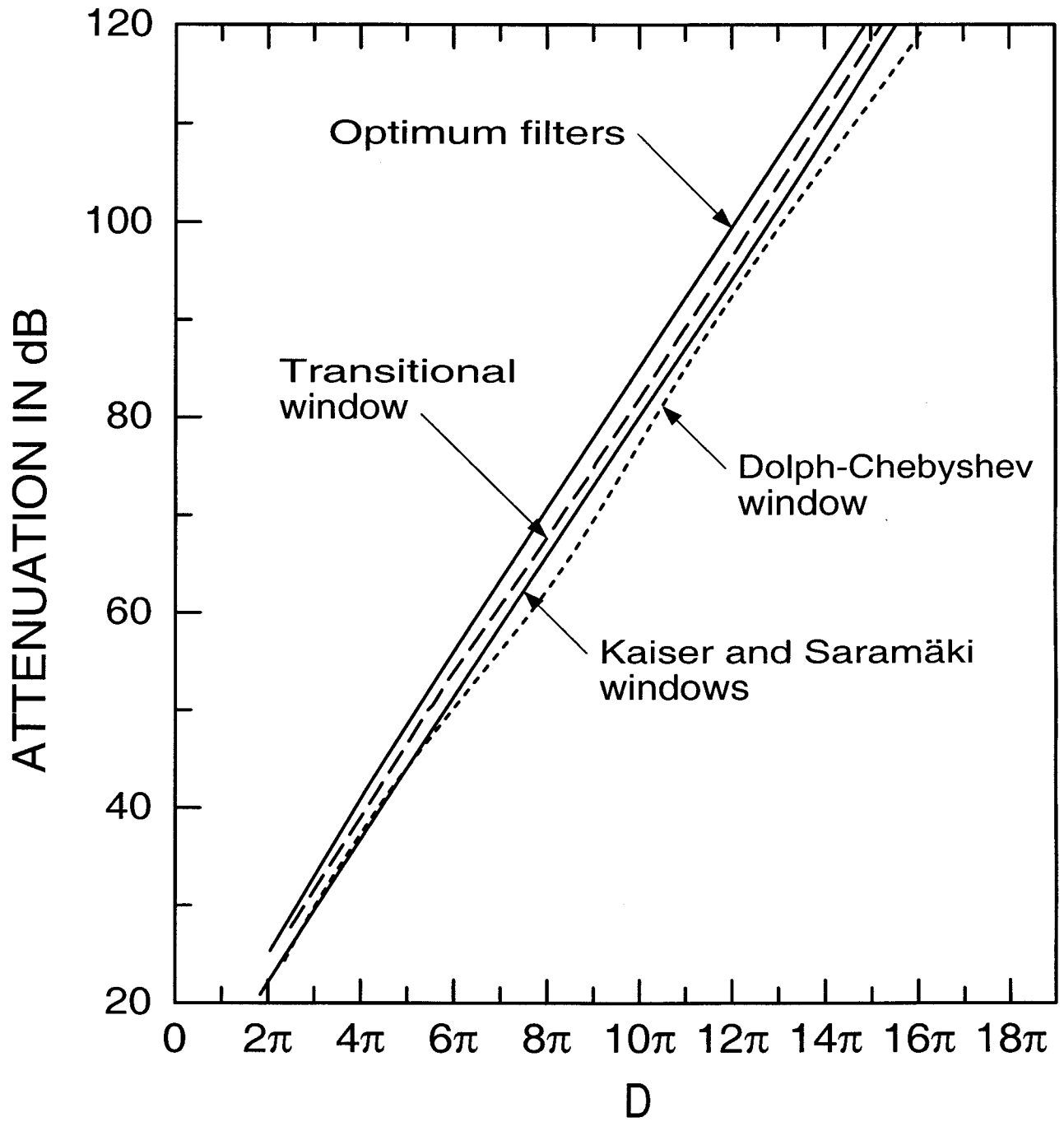
COMPARISON BETWEEN THE ADJUSTABLE WINDOWS

- An informative way to compare the performances of adjustable windows is to design several classes of filters with various values of the adjustable parameter for fixed values of M and ω_c .
- Based on the resulting filter frequency responses, a plot of the stopband attenuation as a function of the parameter $D = 2M(\omega_s - \omega_p)$ can be generated (D , instead of $\omega_s - \omega_p$, is used to make the plot almost independent of M).
- The figure of transparency 121 gives such plots for the above-mentioned adjustable windows for $\omega_c = 0.4\pi$ and $M = 128$.
- For the Kaiser window and the Saramäki window, the difference in the plots is very small.
- For comparison purposes, also a corresponding plot is included for filters for which the passband and stopband ripples $\delta_p = \delta_s$ are minimized in the minimax manner for the given value of D .
- This plot gives thus an upper limit for the stop-

band attenuation attainable using window functions.

- For the Kaiser window and the Saramäki window, the resulting attenuation is typically 6 dB less than this upper limit.
- The stopband attenuation obtained by the Dolph-Chebyshev window is 2 to 5 dB worse than that of the Kaiser or Saramäki window.
- For the transitional window, the improvement is typically 2 to 4 dB over the Kaiser and Saramäki windows and the resulting attenuation approaches the upper limit.

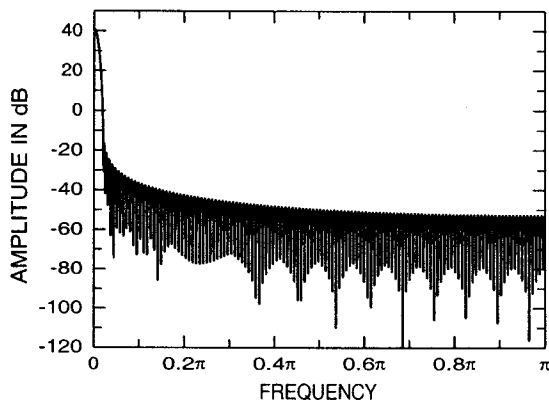
COMPARISON BETWEEN THE ADJUST- ABLE WINDOWS



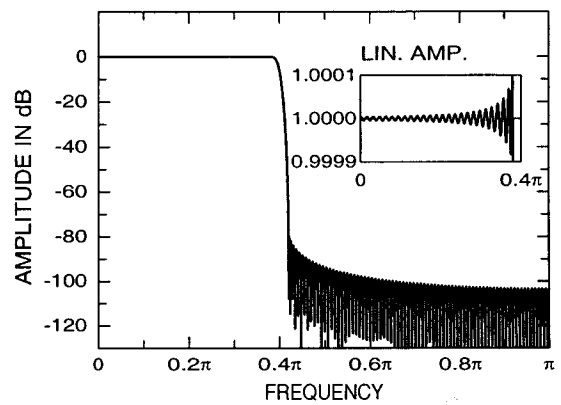
Example

- It is desired to design with each adjustable window considered above a filter with a 80-dB stop-band attenuation for $M = 128$ and $\omega_c = 0.4\pi$.
- Using the estimation formulas of transparency 115 above, the values for the adjustable parameters for the Kaiser, the Saramäki, the Dolph-Chebyshev, and the transitional windows are $\alpha = 7.857$, $\beta = 2.702$, $\beta = 2.770$, and $\beta = 2.587$, respectively.
- The resulting attenuations are 79.68, 80.17, 79.29, and 80.75 dB, respectively.
- The following two transparencies show the frequency responses of both the window function and the resulting filter in the case of an exactly 80-dB attenuation as well as the corresponding window in the time domain.

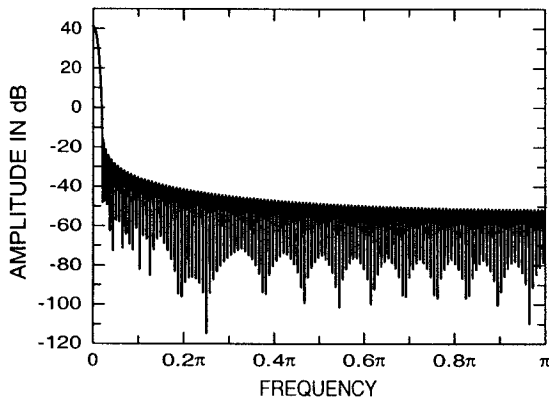
Responses for the adjustable windows and the resulting filters for an exactly 80-dB attenuation. (a,b) Kaiser. (c,d) Saramäki. (e,f) Dolph-Chebyshev. (g,h) Transitional.



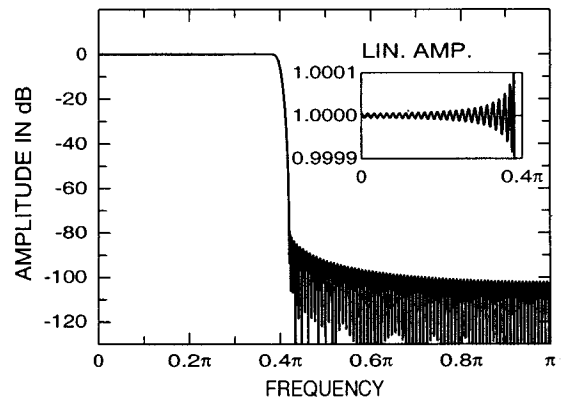
(a)



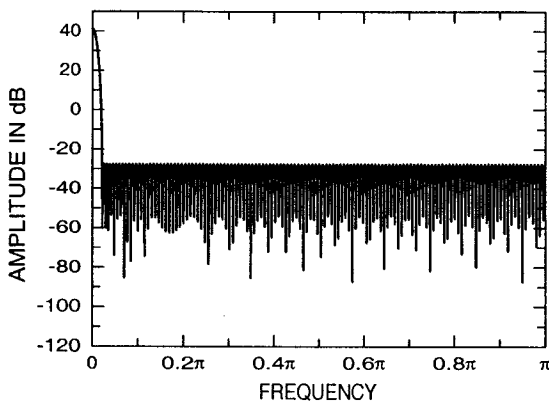
(b)



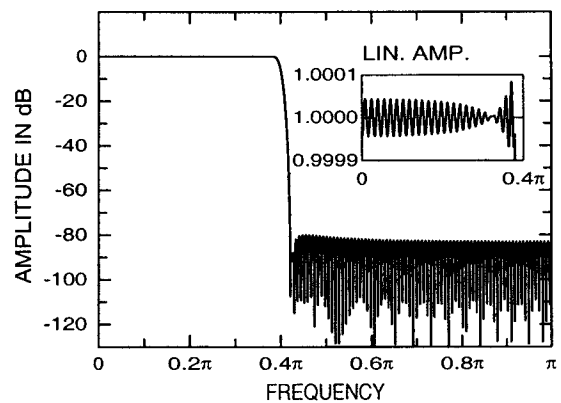
(c)



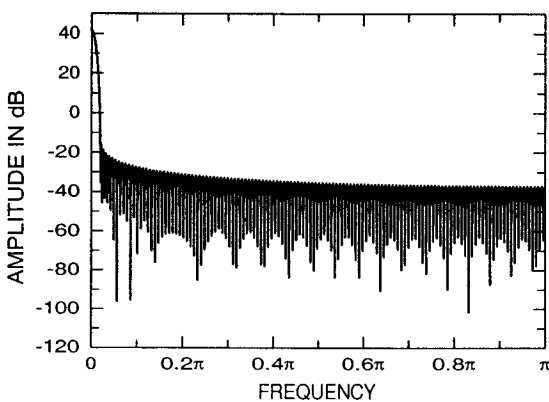
(d)



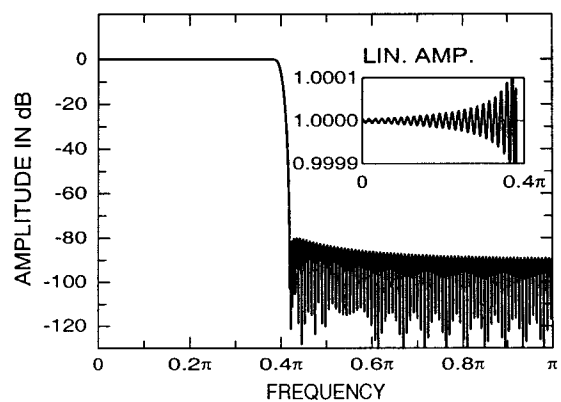
(e)



(f)

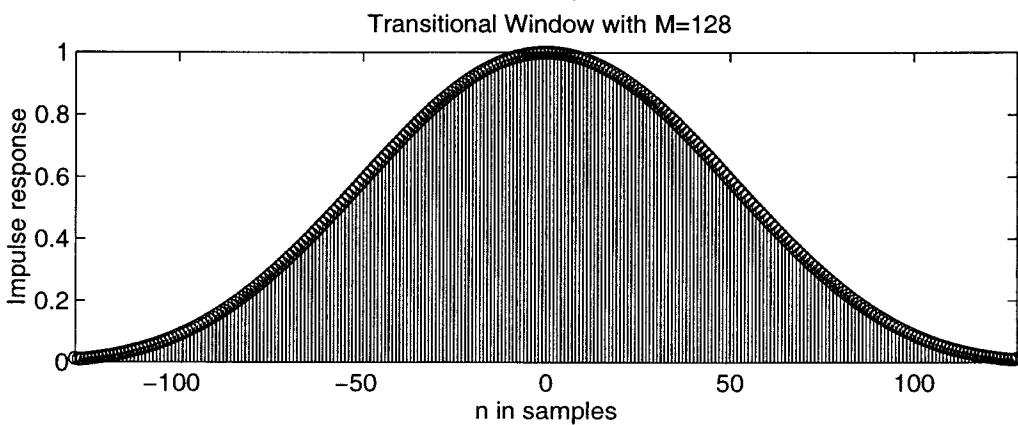
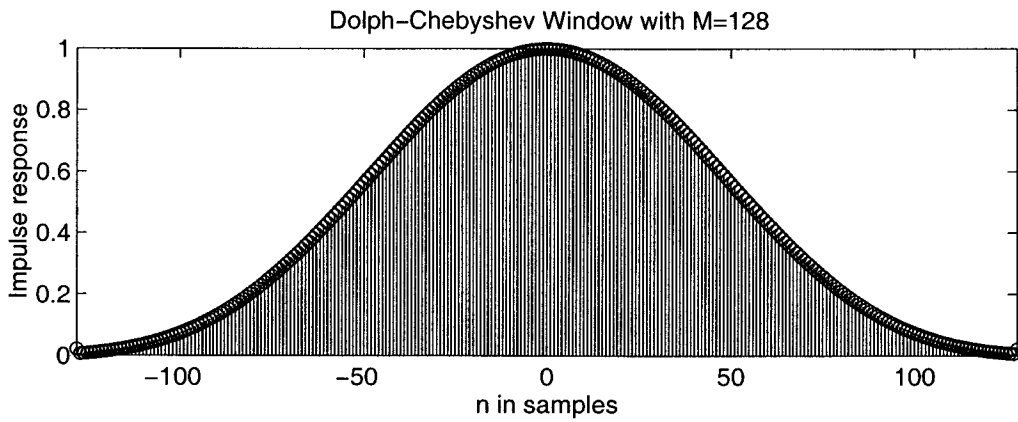
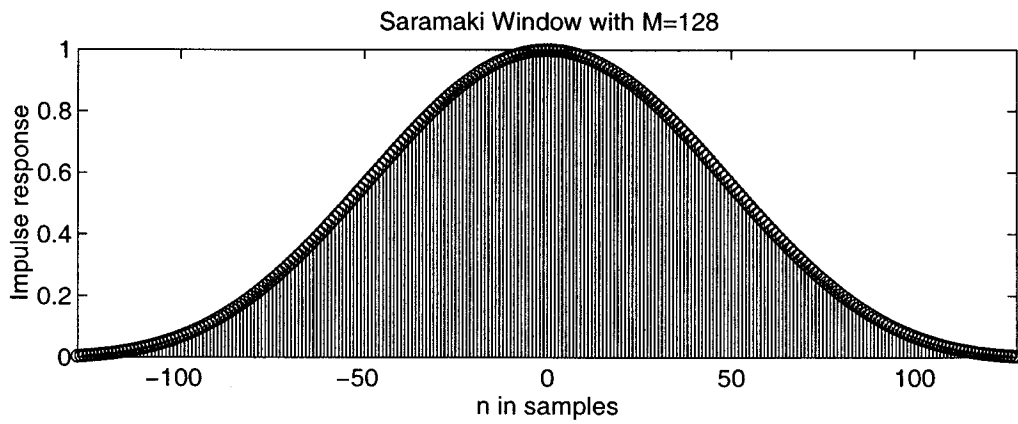
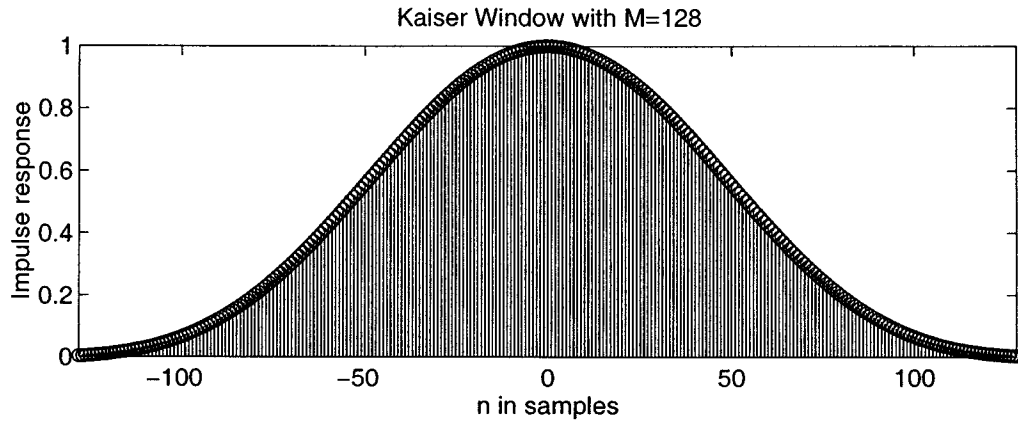


(g)



(h)

Adjustable windows giving exactly an 80-dB attenuation



Comparisons and Comments

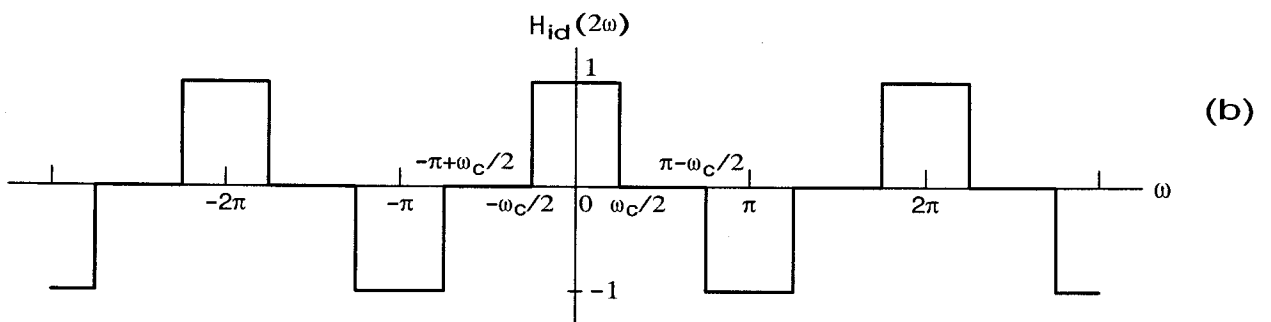
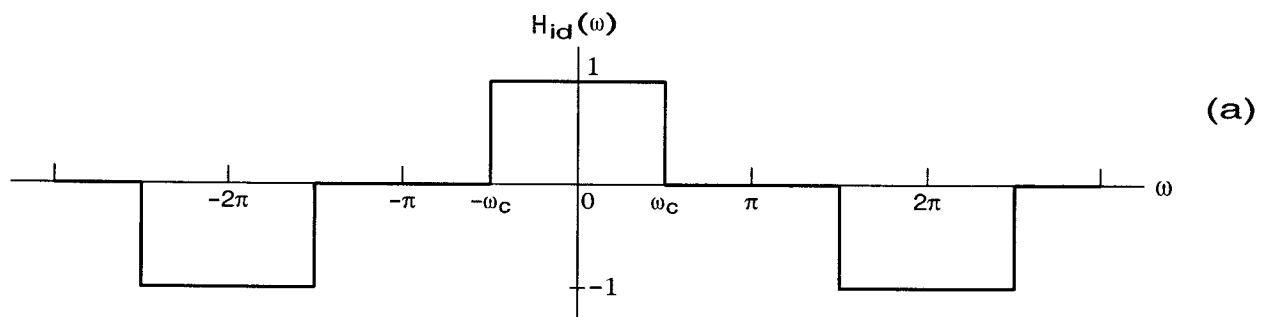
- The transition bandwidths for these filters are 0.0393π , 0.0390π , 0.0404π , and 0.0372π , respectively.
- To achieve the transition bandwidth resulting when using the Kaiser window, the Dolph-Chebyshev window requires $M = 132$, whereas for the transitional window M can be reduced to $M = 123$.
- It is interesting to observe from transparency 123, that for the Kaiser and Saramäki windows, the responses of both the window function and the filter have larger ripples near the stopband edge.
- For the Dolph-Chebyshev window, all the side-lobe levels are the same and there is a ‘hole’ in the response of the filter near the stopband edge.
- As expected, the behaviors of the transitional window and the resulting filter are between those of the Saramäki and Dolph-Chebyshev windows.

DESIGN OF TYPE II FILTERS

- The design of a Type II filter of odd order N can be carried out by using the following process:

Step I: Instead of the Type II ideal response $H_{id}(\omega)$ that is even around $\omega = 0$, odd around $\omega = \pi$, and has the periodicity of 4π , we first concentrate on the Type I ideal response $H_{id}(2\omega)$.

- As seen from the figure shown below for a low-pass case, the resulting $H_{id}(2\omega)$ is even around $\omega = \pi$ and the periodicity is 2π , as required by the Type I zero-phase frequency response.



Step II: Like in the process described in the beginning of this section, determine the coefficients of the infinite-duration zero-phase filter according to

$$h_{id}^{(0)}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{id}(2\omega) \cos(n\omega) d\omega, \quad -\infty < n < \infty.$$

- In the lowpass case of the previous transparency,

$$h_{id}^{(0)}[n] = \begin{cases} 0 & \text{for } n \text{ even} \\ \sin(\omega_c n/2)/(\pi n) & \text{for } n \text{ odd.} \end{cases}$$

- Also for other cases, $h_{id}^{(0)}[n] = 0$ for n even, since $H_{id}(2\omega)$ is odd around the points $\omega = \pm\pi/2$.

Step 3: Form the coefficients of the approximating finite-duration zero-phase (center of symmetry is at $n = 0$) filter according to

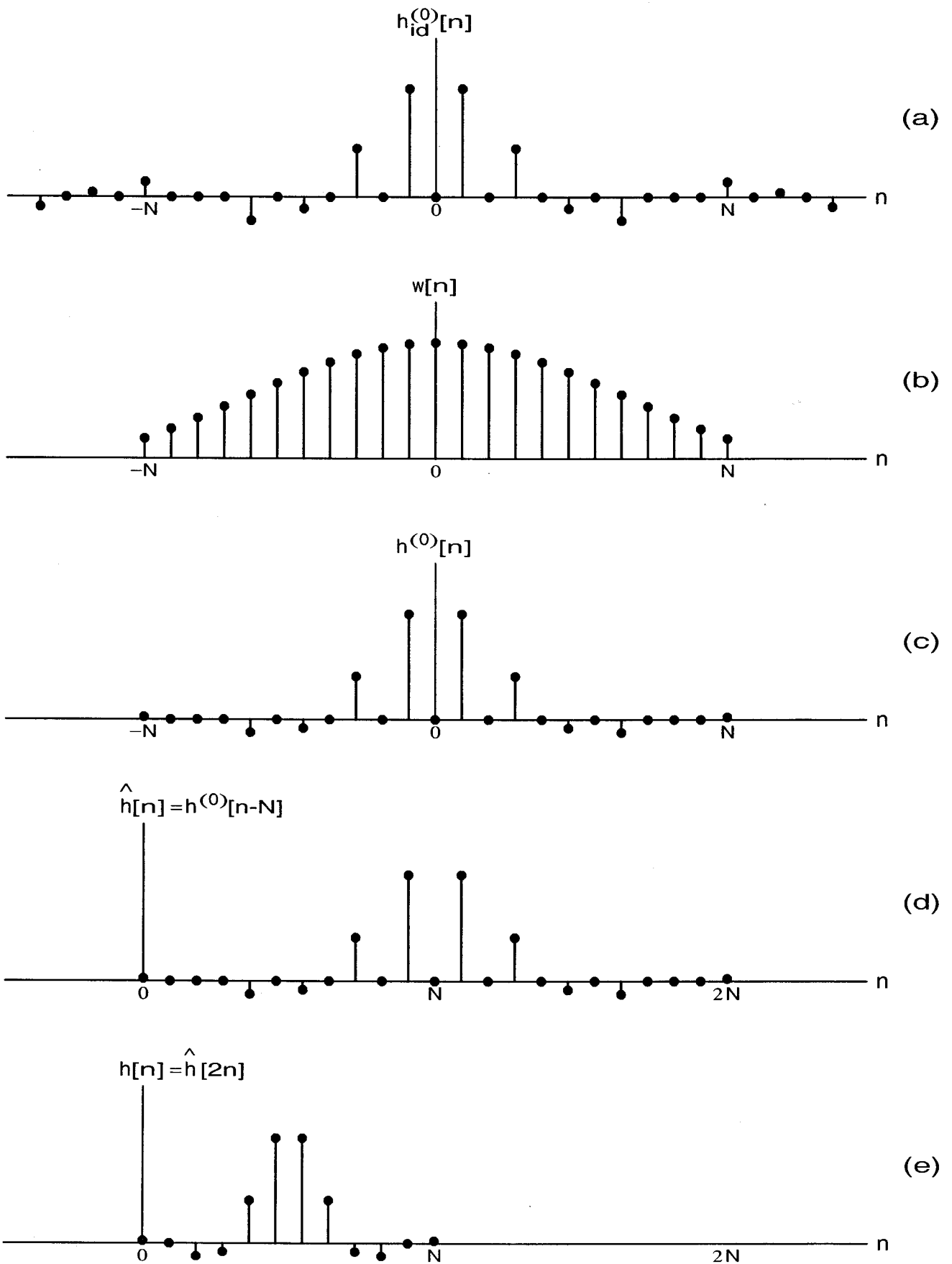
$$h^{(0)}[n] = w[n] h_{id}^{(0)}[n]$$

where $w[n]$ is a **window function** which is nonzero for $-N \leq n \leq N$ (see the next transparency).

Step 4: The coefficients of the causal realizable filter are obtained by shifting the center of symmetry from $n = 0$ to $n = N$ (see the next transparency), yielding

$$\hat{h}[n] = h^{(0)}[n - N].$$

DESIGN PROCESS FOR TYPE II FILTERS IN THE TIME DOMAIN



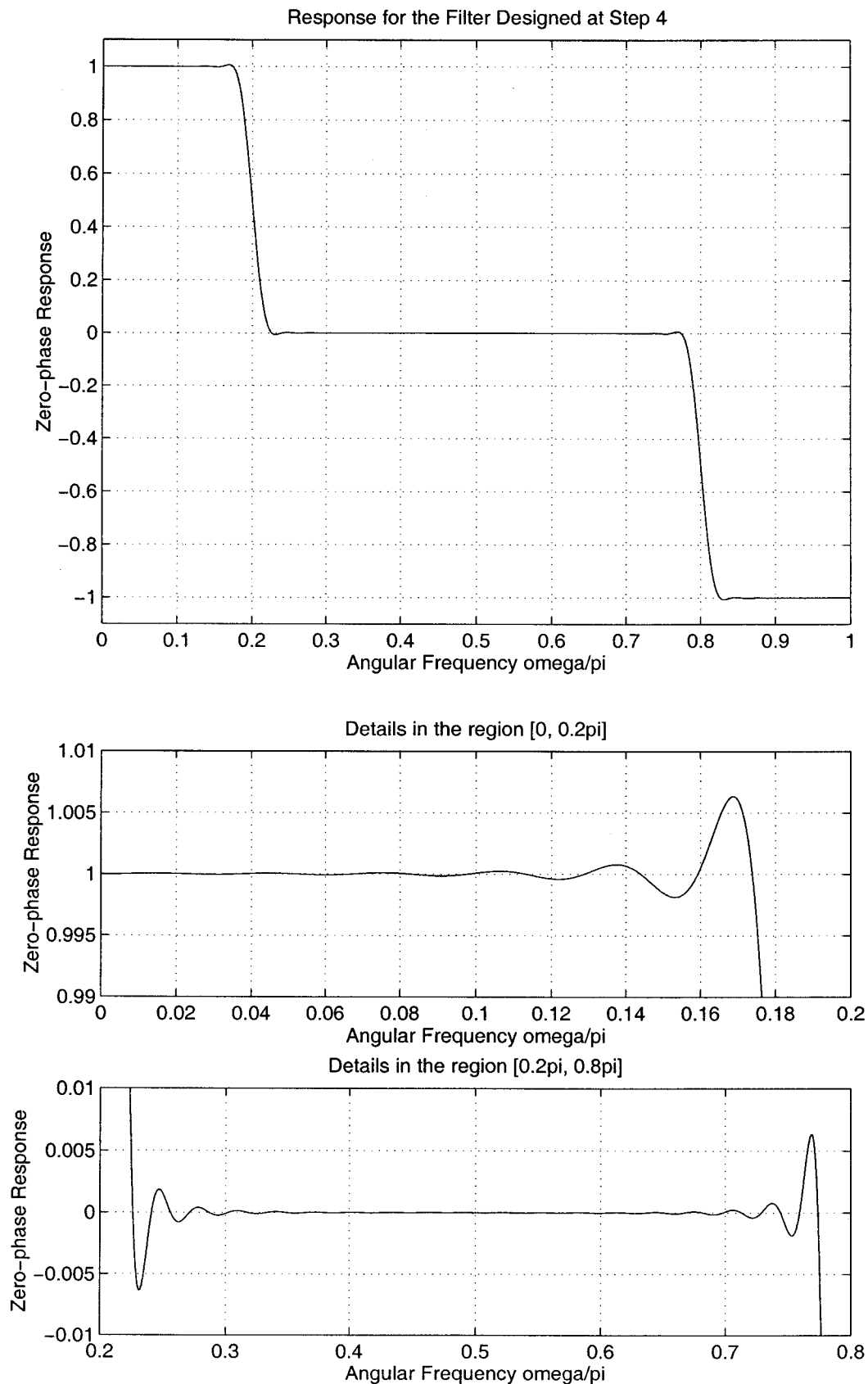
- The Type I filter having the impulse response coefficients $\hat{h}[n]$ obtained at the previous step is characterized by the facts (see both the previous and the next transparencies):
 - The filter order is $2N$, instead of the desired order N .
 - Its zero-phase frequency response approximates $H_{\text{id}}(2\omega)$, instead of the desired $H_{\text{id}}(\omega)$.
 - The impulse-response coefficients satisfy $\hat{h}[N \pm 2r] = 0$ for $r = 0, 1, \dots, (N - 1)/2$.

Step 5: The Type II filter with the zero-phase frequency approximating the original Type II ideal zero-phase frequency response is obtained by discarding the zero-valued impulse response samples of $\hat{h}[n]$, yielding (see the previous transparency and transparency 131)

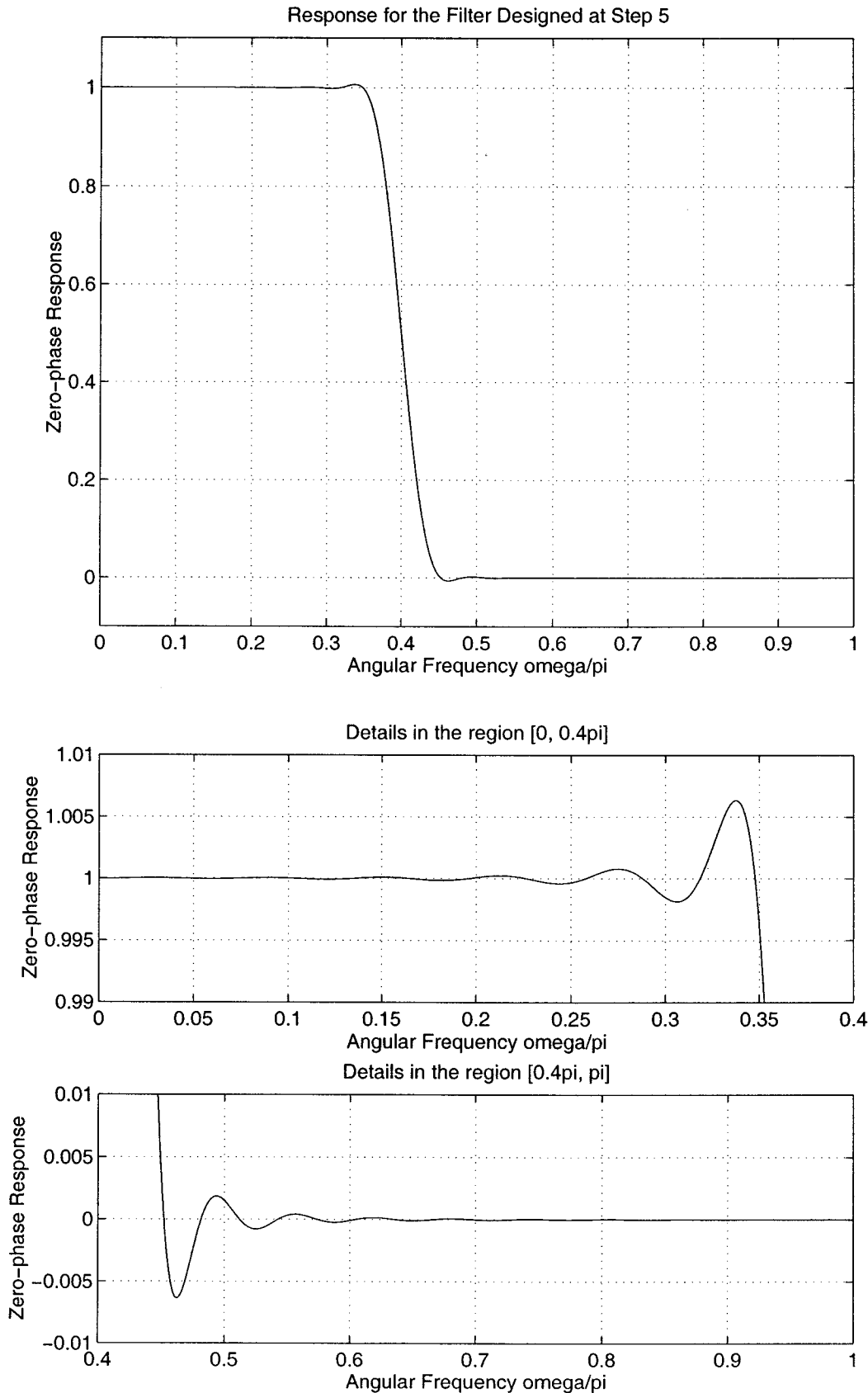
$$h[n] = \hat{h}[2n] \quad \text{for } n = 0, 1, \dots, N.$$

- The order of the resulting Type II filter is now N , as is desired.

Responses for the Type I Filter Obtained at Step 4: $N = 63$, $\omega_c = 0.4\pi$, and the Hann window has been used.



Responses for the Type II Filter Obtained at Step 5: $N = 63$, $\omega_c = 0.4\pi$, and the Hann window has been used.



COMBINED FORMULAS FOR THE CAUSAL TYPE I AND II FILTERS DESIGNED BY WINDOWING

- After some reasoning, the coefficients of the causal filters for both Type I (N even) and Type II (N odd) can be expressed in the lowpass, highpass, bandpass, and bandstop cases in the following common form:

$$h[n] = W[n]f[n] \quad \text{for } n = 0, 1, \dots, N.$$

- Here, $f[n]$ for $n = 0, 1, \dots, N$ is given by

$$f[n] = \begin{cases} \omega_c/\pi, & n = N/2 \text{ and } N \text{ even} \\ \frac{\sin[\omega_c(n - N/2)]}{\pi(n - N/2)}, & \text{otherwise} \end{cases}$$

for the lowpass case with cutoff edge at ω_c ,

$$f[n] = \begin{cases} 1 - \omega_c/\pi, & n = N/2 \text{ and } N \text{ even} \\ \frac{\sin[\omega_c(n - N/2)]}{\pi(n - N/2)}, & \text{otherwise} \end{cases}$$

for the highpass case with cutoff edge at ω_c ,

$$f[n] = \begin{cases} (\omega_{c2} - \omega_{c1})/\pi, & n = N/2 \text{ and } N \text{ even} \\ \frac{\sin[\omega_{c2}(n - N/2)] - \sin[\omega_{c1}(n - N/2)]}{\pi(n - N/2)}, & \text{otherwise} \end{cases}$$

for the bandpass case with cutoff edges at ω_{c1} and ω_{c2} , and

$$f[n] = \begin{cases} 1 - (\omega_{c2} - \omega_{c1})/\pi, & n = N/2 \text{ and } N \text{ even} \\ \frac{\sin[\omega_{c1}(n - N/2)] - \sin[\omega_{c2}(n - N/2)]}{\pi(n - N/2)}, & \text{otherwise} \end{cases}$$

for the bandpass case cutoff edges at ω_{c1} and ω_{c2} .

- Here, ω_c is the cutoff edge for the ideal lowpass and highpass cases, whereas ω_{c1} and ω_{c2} and the lower and upper cutoff edges for the ideal bandpass and bandstop cases.
- For highpass and bandstop filters, N is restricted to be even, since for N odd there exist a fixed zero at $z = -1$ ($\omega = \pi$). This is not allowed for these cases since the zero-phase frequency response should approximately unity in these cases at $\omega = \pi$.
- $W[n]$ for $n = 0, 1, \dots, N$ is given for the fixed windows by

$$W[n] \equiv 1$$

for the rectangular window,

$$W[n] = \begin{cases} 1 - \frac{|n - N/2|}{N/2+1}, & N \text{ even} \\ 1 - \frac{|n - N/2|}{N+1}, & N \text{ even} \end{cases}$$

for the Bartlett window,

$$W[n] = 0.5 + 0.5 \cos\left[\frac{2\pi(n - N/2)}{N + 1}\right]$$

for the Hann window,

$$W[n] = 0.54 + 0.46 \cos\left[\frac{2\pi(n - N/2)}{N + 1}\right]$$

for the Hamming window, and

$$W[n] = 0.42 + 0.5 \cos\left[\frac{2\pi(n - N/2)}{N + 1}\right] + 0.08 \cos\left[\frac{4\pi(n - N/2)}{N + 1}\right]$$

for the Blackman window.

- Among the adjustable windows,

$$W[n] = I_0\left[\alpha \sqrt{1 - \left(\frac{(n - N/2)}{N/2}\right)^2}\right] / I_0(\alpha)$$

for the Kaiser window.

- For the remaining three adjustable windows for N even the $w[n]$'s are determined according to the previous discussion for $-N/2 \leq n \leq N/2$ and $W[n] = w[n - N/2]$ for $n = 0, 1, \dots, N$.

- For N odd, the $w[n]$'s are determined according to the previous discussion for $-N \leq n \leq N$ and $W[n] = w[2n - N]$ for $n = 0, 1, \dots, N$.
- We are now ready to generalizing the design procedure described in the beginning of this section to include both Type I (N even) and Type II (N odd) filters.

DESIGN OF LOWPASS FILTERS BY WINDOWING FOR THE GIVEN VALUES OF ω_p , ω_s , δ_p , and δ_s

Step I: Determine $\omega_c = (\omega_p + \omega_s)/2$ and $A_s = -20 \log_{10} \min [\delta_p, \delta_s]$.

- If $\delta_p < \delta_s$, then A_s is determined according to the value of δ_p and δ_s is forced to be approximately equal to δ_s . However, in most cases, $\delta_p > \delta_s$ and A_s depends on the value of δ_s and $\delta_p \approx \delta_s$.

Step II: A fixed or an adjustable window function? This depends on whether one of the fixed window functions considered in transparency 101 provides a satisfactory value for A_s .

Step IIA: If a fixed window function is selected and both even and odd values of N are used, then $N/2$ takes the role of M and N is selected to be the smallest integer satisfying the condition given in the last column of the second table of transparency 101, that is,

$$\Delta\omega_p = (\omega_p - \omega_s) \leq \gamma/N$$

or

$$N \geq \gamma / (\omega_p - \omega_s),$$

where

$$\gamma = \begin{cases} 1.84\pi, & \text{for the rectangular window} \\ 6.22\pi, & \text{for the Hann window} \\ 6.644\pi, & \text{for the Hamming window} \\ 11.13\pi, & \text{for the Blackmann window.} \end{cases} \quad (A)$$

- What is left is to apply the formulas of transparencies of 132–135.

Step IIA: If an adjustable window function is selected, then $N/2$ takes again the role of M . α for the Kaiser window and β for the remaining three adjustable windows are selected according the second columns of the table in transparency 115. By replacing M by $N/2$, the last column of this table can rewritten as

$$N \geq \frac{2f(A_s)}{14.36(\omega_s - \omega_p)/\pi},$$

where

$$f(A_s) = \begin{cases} A_s - 7.95, & \text{Kaiser window} \\ A_s - 8.15, & \text{Saramäki window} \\ 1.028A_s - 0.84 & \text{Dolph-Chebyshev window} \\ 0.00036(A_s)^2 \\ +0.951A_s - 9.4, & \text{transitional window} \end{cases} \quad (A)$$

- Again, what is left is to apply the formulas of transparencies 132–135.

DESIGN OF HIGHPASS FILTERS BY WINDOWING FOR THE GIVEN VALUES OF ω_p , ω_s , δ_p , and δ_s

- Determine $A_s = -20 \log_{10} \min [\delta_p, \delta_s]$ and $\omega_c = (\omega_p + \omega_s)/2$, like for lowpass filters.
- The design of highpass filters is very similar to that of lowpass filters.
- The only differences are the following:
 - $\omega_p > \omega_s$.
 - N must be even in order to avoid a zero at $z = -1$ ($\omega = \pi$).

DESIGN OF BANDPASS FILTERS BY WINDOWING FOR THE GIVEN VALUES OF ω_{p1} , ω_{p2} , ω_{s1} , ω_{s2} , δ_p , and δ_s

Step I: Determine $\omega_{c1} = (\omega_{p1} + \omega_{s1})/2$, $\omega_{c2} = (\omega_{p2} + \omega_{s2})/2$, and $A_s = -20 \log_{10} \min [\delta_p, \delta_s]$. Determine also

$$\Delta\omega = \min [\omega_{p1} - \omega_{s1}, \omega_{s2} - \omega_{p2}].$$

Step II: A fixed or an adjustable window function?

Step IIA: If a fixed window function is selected, then N is selected to be the smallest integer satisfying

$$N \geq \gamma/(\Delta\omega),$$

where γ is given by Eq. (A) of transparency 137. Finally, the formulas of transparencies 132–135 are used.

Step IIA: If an adjustable window function is selected, then N is selected to be the smallest integer satisfying

$$N \geq \frac{2f(A_s)}{14.36\Delta\omega/\pi},$$

where $f(A_s)$ is given by Eq. (A) of transparency

138. Finally, the formulas described earlier are used.

- It should be pointed out that in the bandpass case, the resulting ripples $\delta_s \approx \delta_p$ are usually close to those of the lowpass filter.
- However, for filters having rather narrow bandpass region, one of these ripples is in the worst case twice those of the lowpass design (this occurs very rarely).

DESIGN OF BANDSTOP FILTERS BY WINDOWING FOR THE GIVEN VALUES OF ω_{p1} , ω_{p2} , ω_{s1} , ω_{s2} , δ_p , and δ_s

- Determine $\omega_{c1} = (\omega_{p1} + \omega_{s1})/2$, $\omega_{c2} = (\omega_{p2} + \omega_{s2})/2$, and $A_s = -20 \log_{10} \min [\delta_p, \delta_s]$. Determine also

$$\Delta\omega = \min [\omega_{s1} - \omega_{p1}, \omega_{p2} - \omega_{s2}].$$

- The design of bandstop filters is very similar to that of bandpass filters.
- The only differences are the following:
 - $\omega_{p1} < \omega_{s1}$ and $\omega_{p2} > \omega_{s2}$.
 - N must be even in order to avoid a zero at $z = -1$ ($\omega = \pi$).

DESIGN OF TYPE III FILTERS

- Before the actual design, consider an infinite-duration Type III filter whose impulse response values satisfies $h_{\text{id}}^{(0)}[0] = 0$ and $h_{\text{id}}^{(0)}[n] = -h_{\text{id}}^{(0)}[-n]$ for $n = 1, 2, \dots, \infty$.

- For this filter

$$H_{\text{zero}}(z) = \sum_{n=1}^{\infty} h_{\text{id}}^{(0)}[-n][z^n - z^{-n}]$$

and

$$\begin{aligned} H_{\text{zero}}(e^{j\omega}) &= j \sum_{n=1}^{\infty} h_{\text{id}}^{(0)}[-n][2 \sin n\omega] \\ &= e^{j\pi/2} H_{\text{id}}(\omega), \end{aligned}$$

where

$$H_{\text{id}}(\omega) = \sum_{n=1}^{\infty} h_{\text{id}}^{(0)}[-n][2 \sin n\omega].$$

- For this filter, the center of symmetry is at $n = 0$, instead at $n = N/2$ like for a causal Type III filter of order N .
- The above formulas show that the design of Type III filters can be carried out by windowing as will be described in the following transparencies.

BASIC DESIGN PROCESS FOR TYPE III FILTERS

Step 1: Specify the ideal Type III zero-phase frequency response $H_{\text{id}}(\omega)$ that is odd around $\omega = 0$ and has the periodicity of 2π .

- The ideal responses for the Hilbert transformer and the differentiator are shown in the next transparency. For the Hilbert transformer, the desired function is $+1$ for $0 \leq \omega \leq \pi$ and -1 for $-\pi \leq \omega \leq 0$. For the differentiator, the desired function is ω for $-\omega_c \leq \omega \leq \omega_c$ and zero elsewhere in the interval $[-\pi, \pi]$.

Step 2: Determine the coefficients of the ideal zero-phase filter as follows:

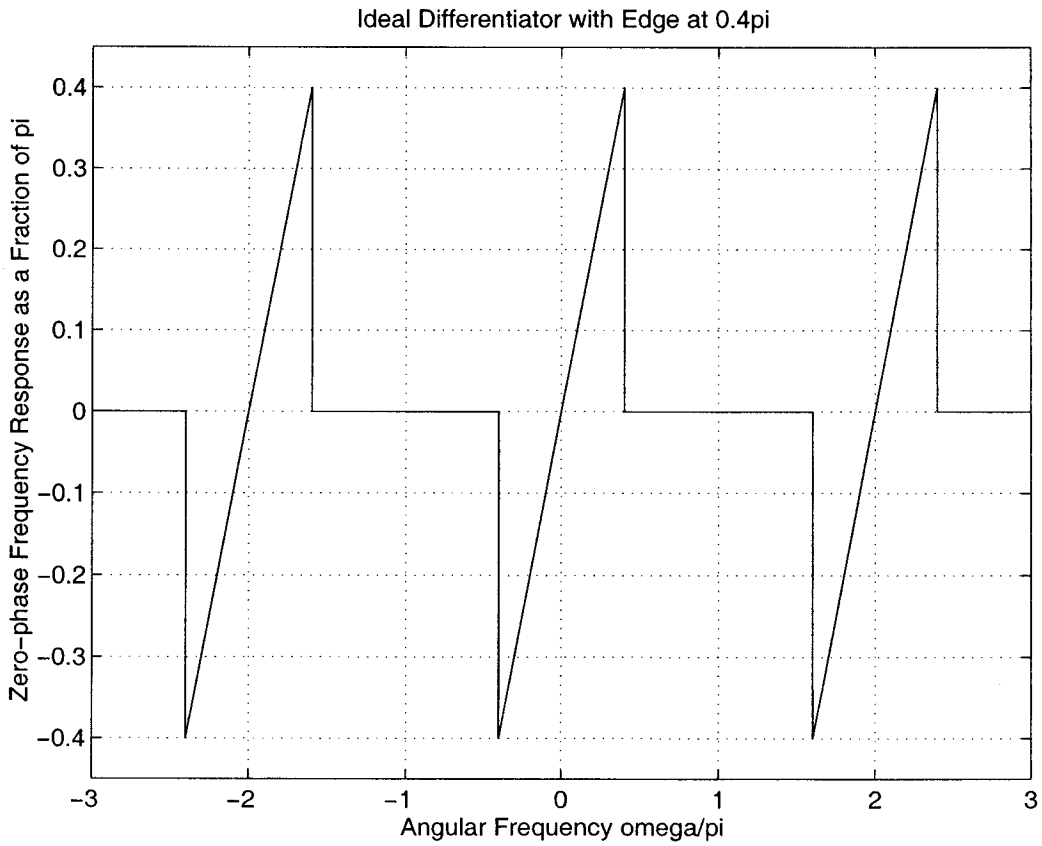
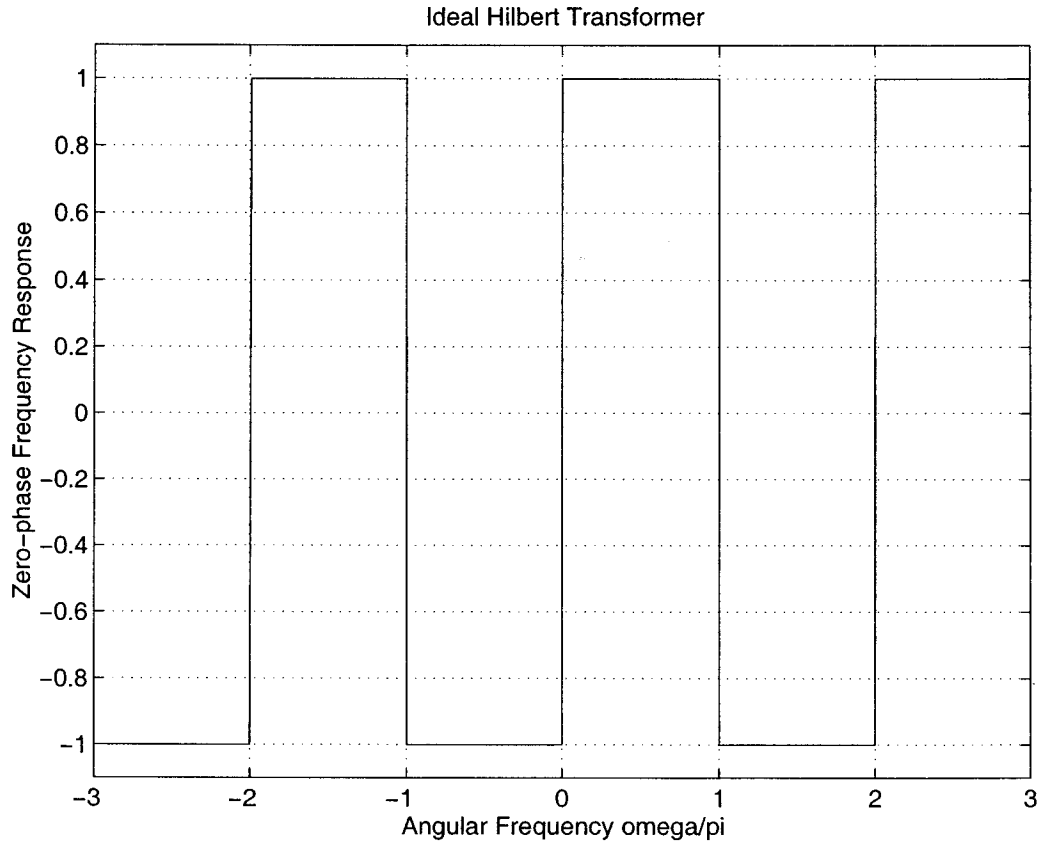
$$h_{\text{id}}^{(0)}[-n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_{\text{id}}(\omega) \sin(n\omega) d\omega, \quad 1 \leq n < \infty,$$

$$h_{\text{id}}^{(0)}[0] = 0,$$

and

$$h_{\text{id}}^{(0)}[n] = -h_{\text{id}}^{(0)}[-n], \quad 1 \leq n < \infty.$$

IDEAL TYPE III ZERO-PHASE FREQUENCY RESPONSES



- For the Hilbert transformer,

$$h_{\text{id}}^{(0)}[n] = \begin{cases} 0, & \text{for } n = 0 \\ -[1 - \cos(n/\pi)]/(n\pi) & \text{for } n \neq 0. \end{cases}$$

- For the differentiator with cutoff edge at ω_c ,

$$h_{\text{id}}^{(0)}[n] = \begin{cases} 0, & \text{for } n = 0 \\ -\frac{\sin(n\omega_c)}{\pi n^2} \\ +\frac{\omega_c \cos(n\omega_c)}{\pi n} & \text{for } n \neq 0. \end{cases}$$

Step 3: Form the coefficients of the approximating finite-duration zero-phase using a window function $w[n]$ being nonzero for $-M \leq n \leq M$ as

$$h^{(0)}[n] = w[n]h_{\text{id}}^{(0)}[n].$$

Step 4: The coefficients of the causal realizable filter are then

$$h[n] = h^{(0)}[n - M].$$

- Note that the order of the resulting Type III filter is $N = 2M$.

COMBINED FORMULAS FOR THE CAUSAL TYPE III AND IV FILTERS DESIGNED BY WINDOWING

- The design of Type IV Hilbert transformers and differentiators can be accomplished with the aid of Type III filters in a manner similar to the design of Type II filters with the aid of Type I filters.
- The coefficients of the causal filters for both Type III (N even) and Type IV (N odd) can be expressed for the Hilbert transformers and differentiators in the common form:

$$h[n] = W[n]f[n] \quad \text{for } n = 0, 1, \dots, N.$$

- For the Hilbert transformer, $f[n]$ for $n = 0, 1, \dots, N$ is given by

$$f[n] = \begin{cases} 0, & n = N/2 \text{ and } N \text{ even} \\ -\frac{1 - \cos[(n - N/2)\pi]}{(n - N/2)\pi}, & n \neq N/2 \text{ and } N \text{ even} \\ -\frac{1}{(n - N/2)\pi}, & N \text{ odd.} \end{cases}$$

- For the differentiator with cutoff ω_c , $f[n]$ for $n = 0, 1, \dots, N$ is given for $n = N/2$ and N even by

$$f[n] = 0,$$

whereas for N odd and $n \neq N/2$ and N even by

$$f[n] = - \frac{\sin[(n - N/2)\omega_c]}{\pi(n - N/2)^2} + \frac{\omega_c \cos[(n - N/2)\omega_c]}{\pi(n - N/2)}.$$

- The causal windows $W[n]$ are the same as those considered previously in connection with designing Type I and II filters (see transparencies 133–135).

DESIGN OF HILBERT TRANSFORMERS BY WINDOWING

- For Hilbert transformers, we give usually the maximum allowable deviation δ_h from unity on $[\omega_h, \pi - \omega_h]$ for N even and on $[\omega_h, \pi]$ for N odd.
- There are the following differences compared to the design of Type I and II lowpass filters.
- **If** a lowpass filter has been designed using a window function such that

1) The transition bandwidth is

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

2) The passband and stopband ripples are

$$\delta_p \approx \delta_s = \hat{\delta}.$$

- **Then** to the corresponding Hilbert transformer designed by the same window applies approximately:

1) The first stopband edge is at

$$\omega_h = \Delta\omega/2.$$

2) The maximum deviation from unity is

$$\delta_h = 2\hat{\delta}.$$

- Therefore, when designing Hilbert transformers,

$$\delta_s = \delta_h/2$$

and

$$\omega_s - \omega_p = 2\omega_h$$

can be used in the formulas described above.

DESIGN OF DIFFERENTIATORS BY WINDOWING

- For differentiators, we give usually the band $[0, \omega_{dp}]$ where it is desired to approximate ω with tolerance δ_{dp} and the band $[\omega_{ds}, \pi]$ where it is desired to approximate zero with tolerance δ_{ds} .
- Like for lowpass filters, the cutoff edge is selected to be

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

- **If** a lowpass filter has been designed using a window function such that
 - 1) The transition bandwidth is

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

- 2) The passband and stopband ripples are

$$\delta_p \approx \delta_s = \hat{\delta}.$$

- **Then** to the corresponding differentiator designed by the same window applies approximately:
 - 1) The transition bandwidth $\omega_{ds} - \omega_{dp}$ is the same as for the lowpass filter.

2) The passband and stopband ripples are

$$\delta_{dp} \approx \delta_{ds} = \omega_c \delta_s.$$

- Therefore, when designing differentiators

$$\delta_s = \widehat{\delta}_{ds} / \omega_c$$

and

$$\omega_s - \omega_p = \omega_{ds} - \omega_{dp}$$

can be used in the formulas described above.

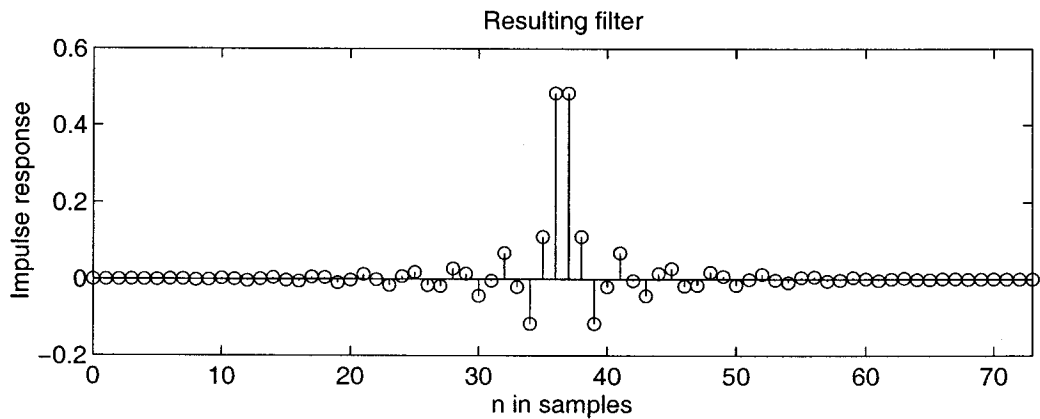
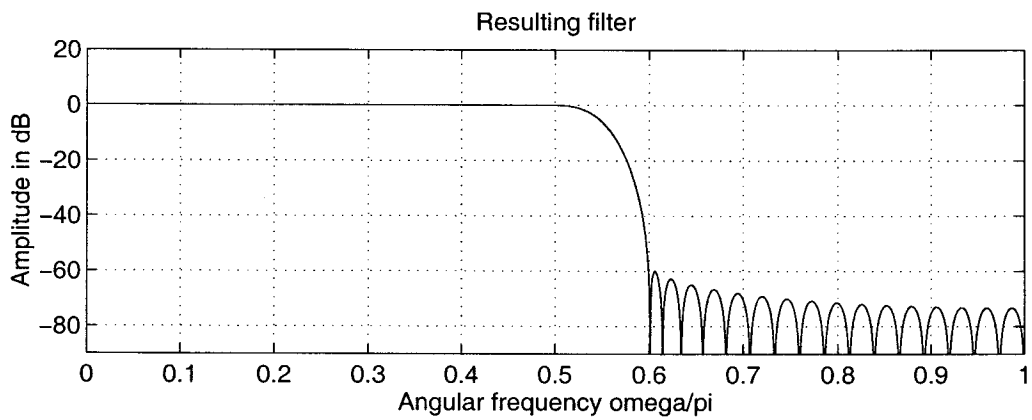
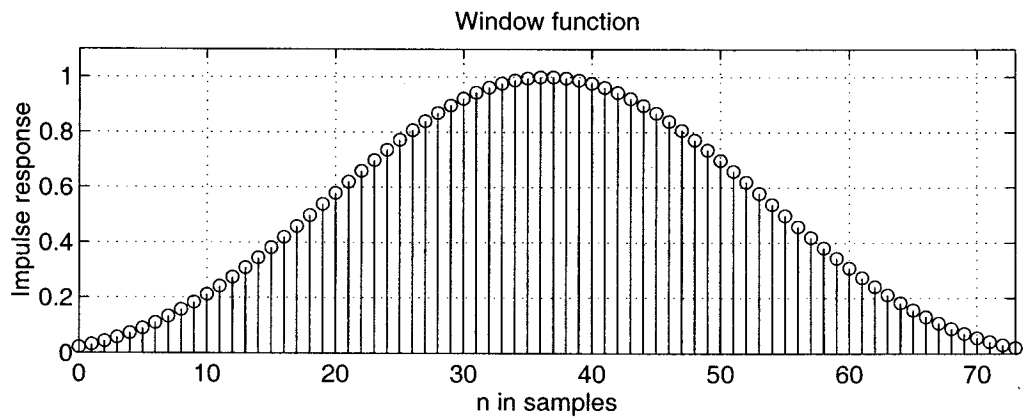
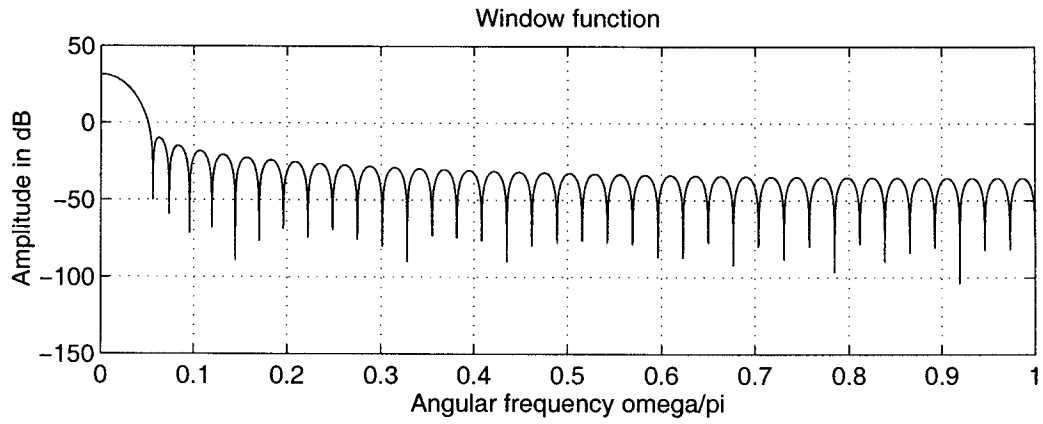
MATLAB ROUTINES

- The author of these lecture notes has generated four general-purpose routines:
- Given the criteria for a lowpass, highpass, bandpass, or bandstop filter (edges and ripples) as well as the adjustable window (Kaiser, Saramäki, Dolph-Chebyshev, or transitional window), `firwinad.m` automatically searches for the minimum filter length and finds the adjustable parameter to just meet the given criteria.
- This routine is given in Appendix A in the end of this chapter.
- `hilwinad.m` (Appendix B) and `difwinad.m` (Appendix C) do the same for the Hilbert transformers and differentiators, respectively.
- `firwifix.m` (Appendix D) is a routine for designing with the aid of fixed windows lowpass, highpass, bandpass, and bandstop filters as well as Hilbert transformers and differentiators.
- We consider next some examples.

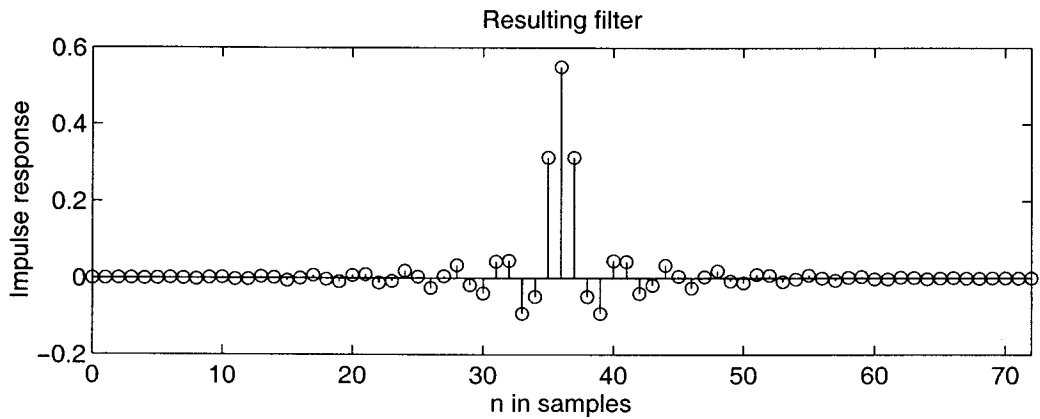
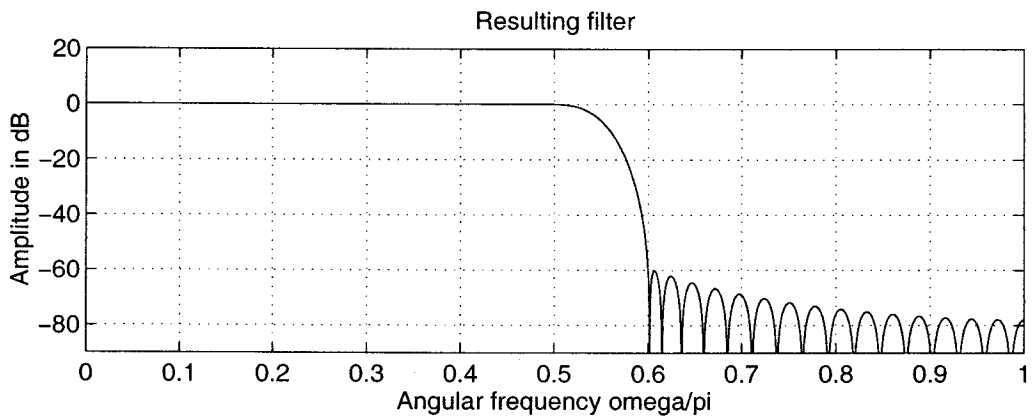
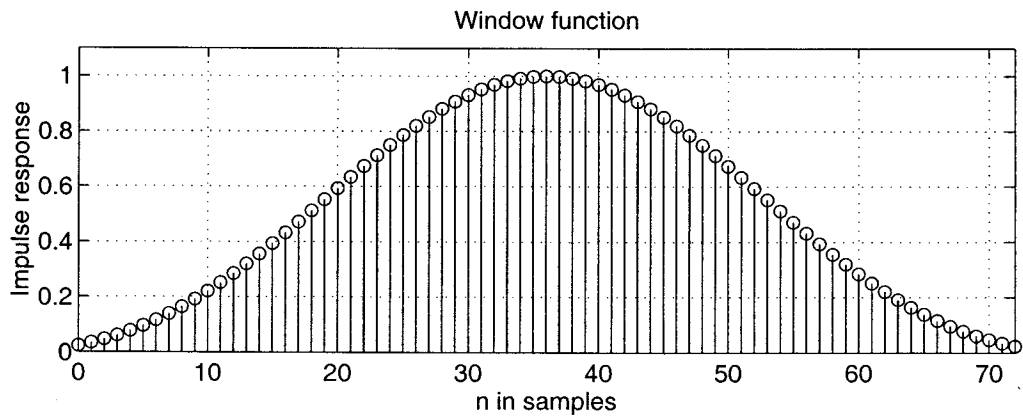
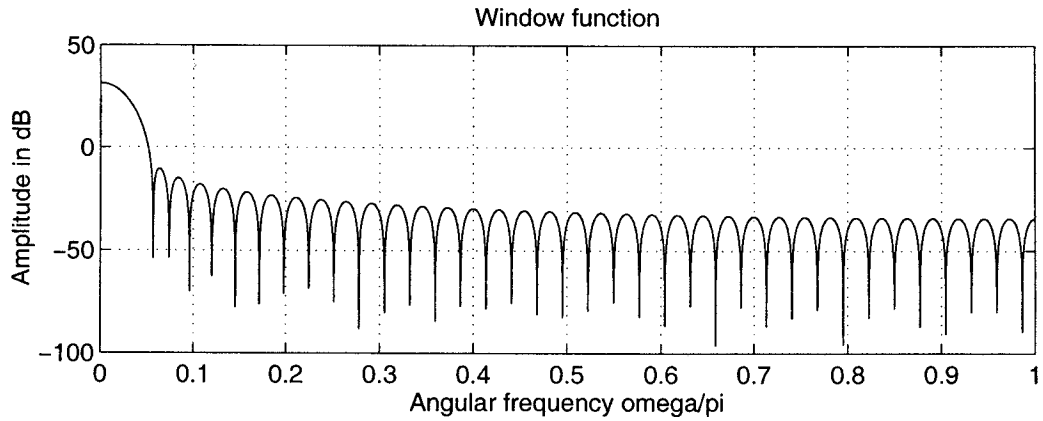
EXAMPLE 1

- It is desired to design with the aid of adjustable windows a linear-phase FIR filter (Type I or Type II) such that the maximum deviation from unity on $[0, 0.5\pi]$ is at most 0.002 and the maximum deviation from zero on $[0.6\pi, \pi]$ is at most 0.001 (60-dB attenuation).
- The results obtained using `firwinad.m` are given in the next four transparencies.
- The minimum orders for the Kaiser window, the Saramäki window, the Dolph-Chebyshev window, and the transitional window are 73, 72, 76, and 70, respectively.

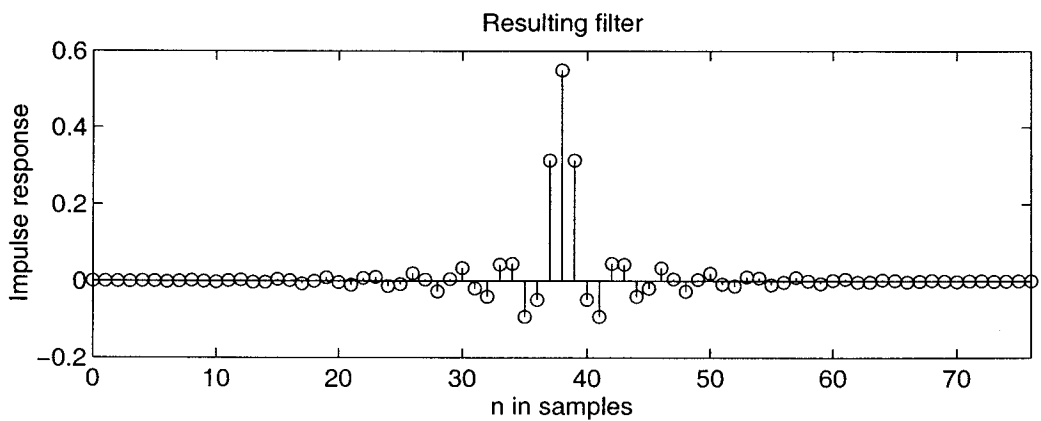
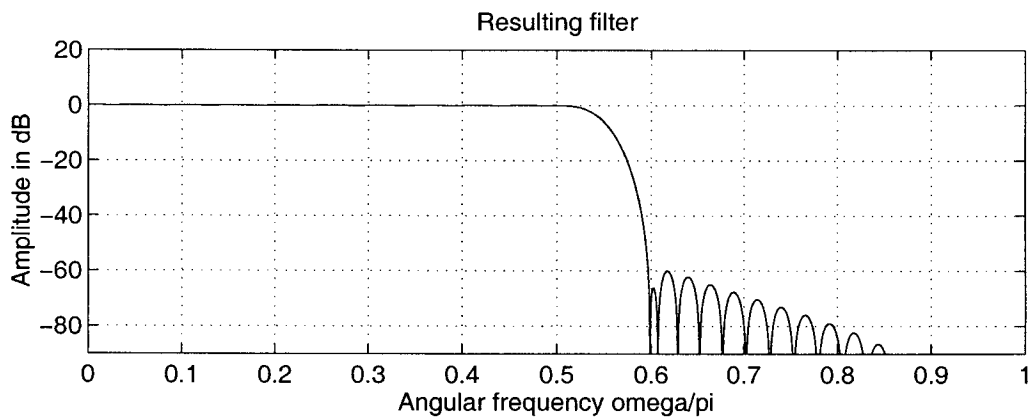
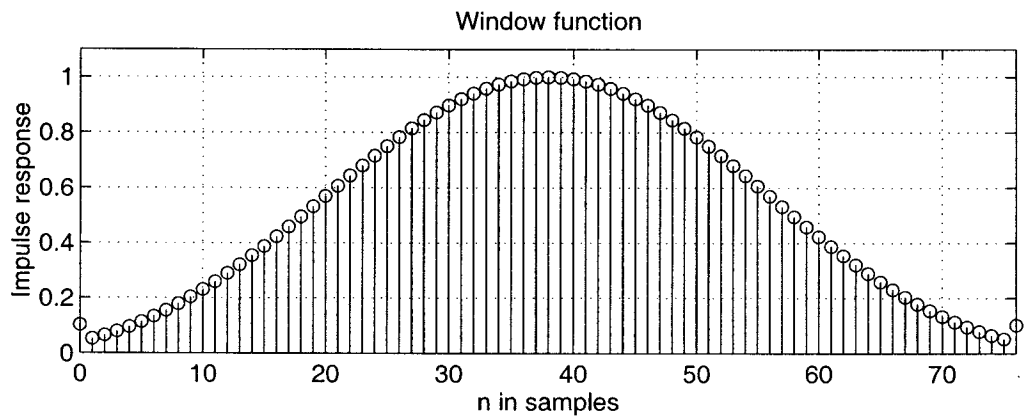
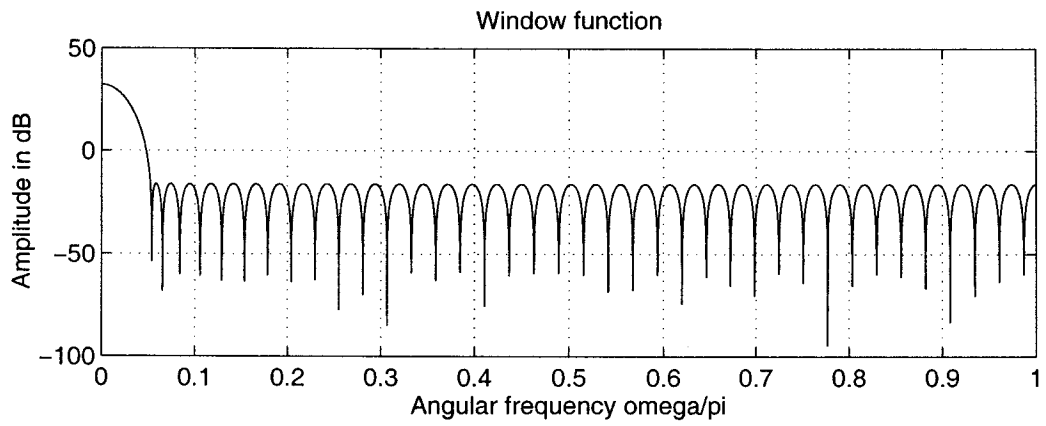
Kaiser window: $N = 73$



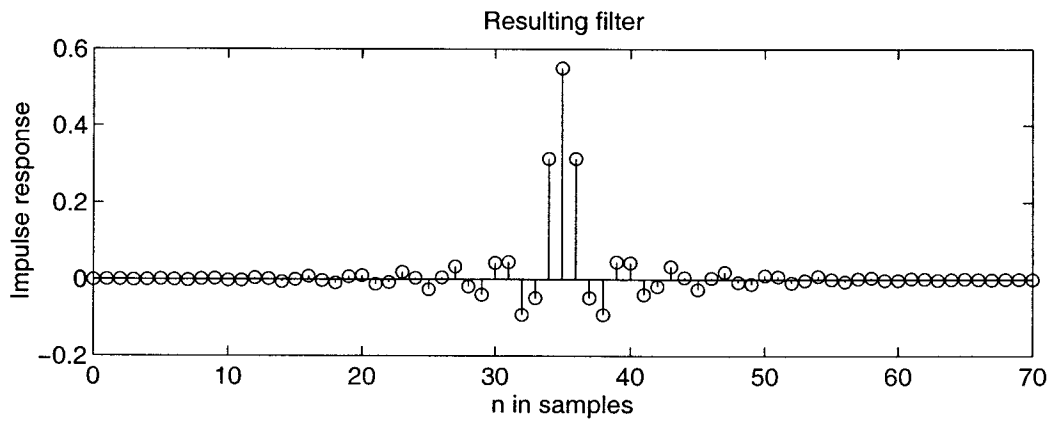
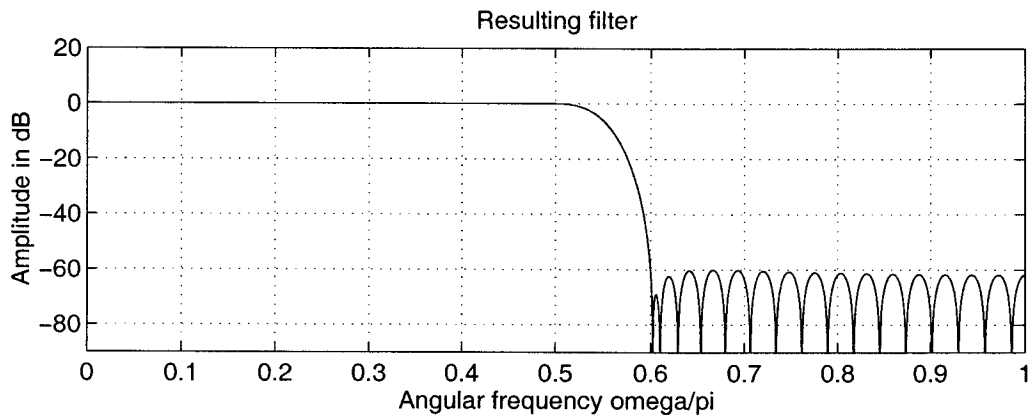
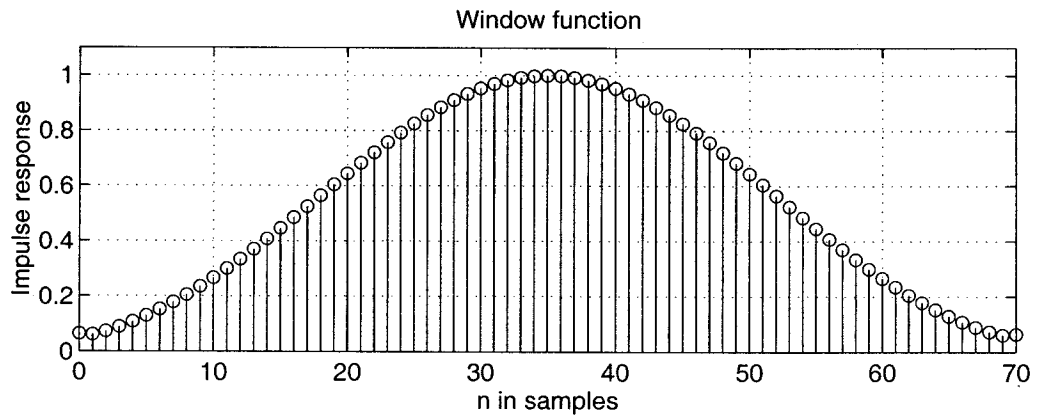
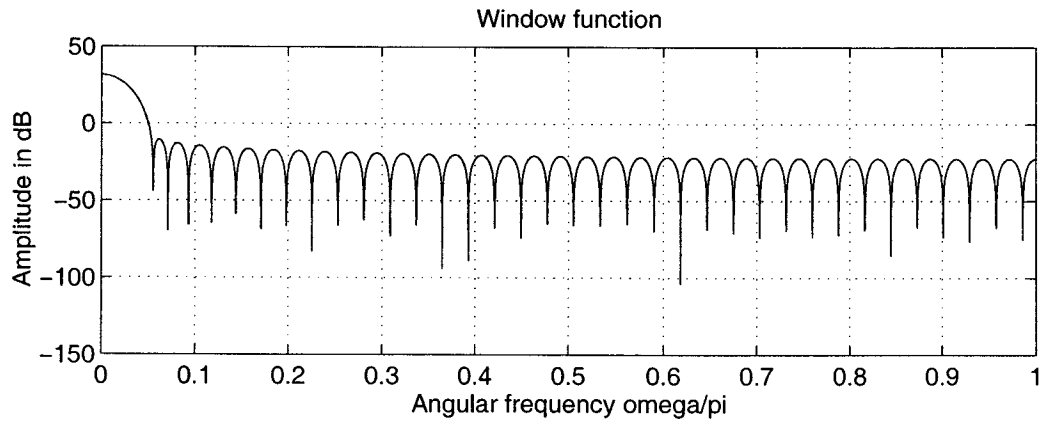
Saramäki window: $N = 72$



Dolph-Chebyshev window: $N = 76$

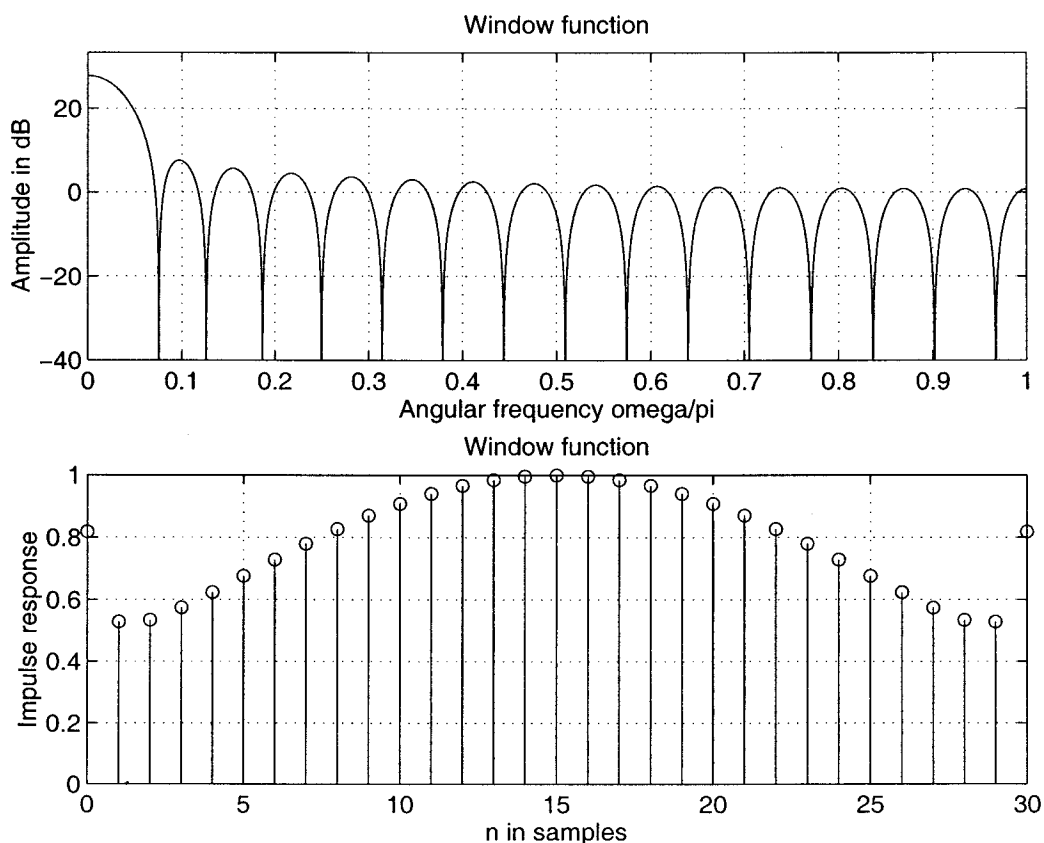


Transitional window: $N = 70$

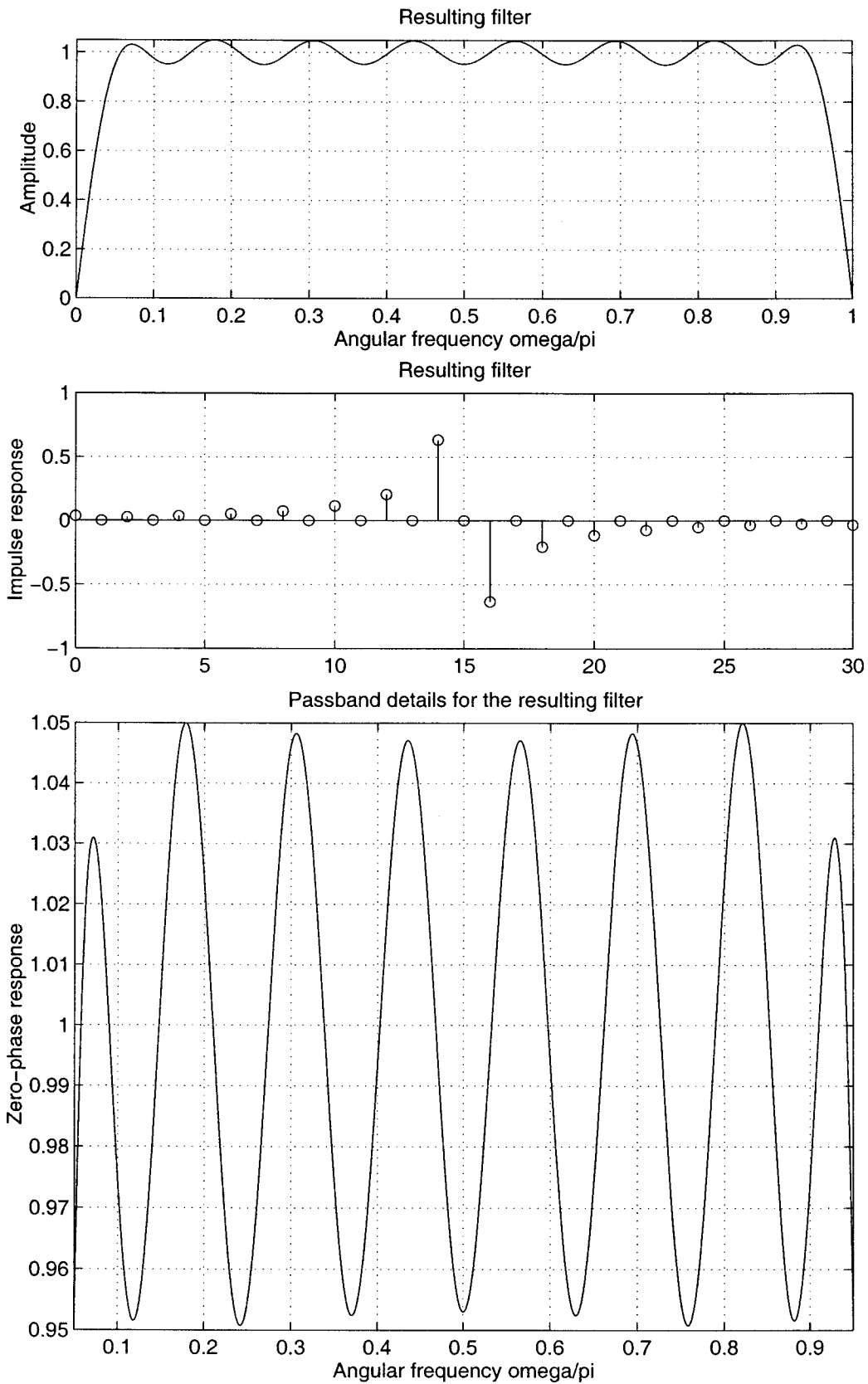


EXAMPLE 2

- It is desired to design with the aid of `hilwinad.m` a Hilbert Transformer such that for Type III (N even) the passband region is $[0.05\pi, 0.95\pi]$ and for Type IV $[0.05\pi, \pi]$. The maximum allowable deviation from unity is 0.05.
- The minimum orders for the Kaiser, Saramäki, Dolph-Chebyshev, transitional windows are 33, 33, 35, and 30, respectively.
- The figures below and in the next transparency show the characteristics of the transitional window and the resulting filter.

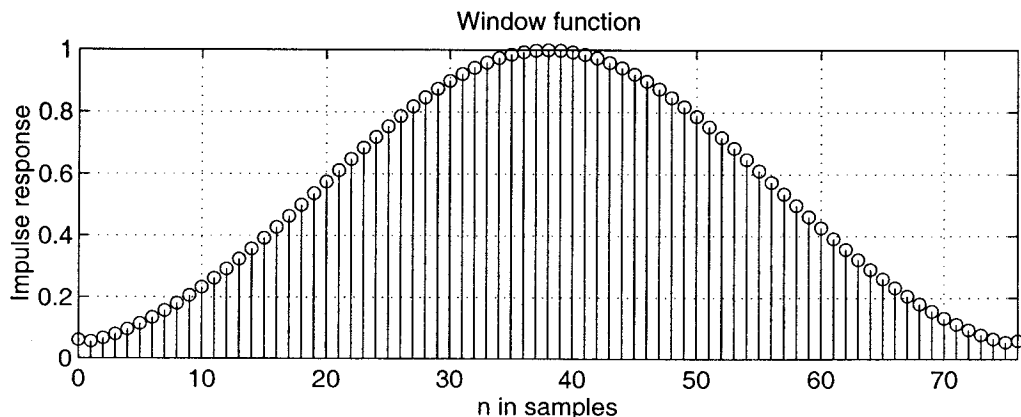
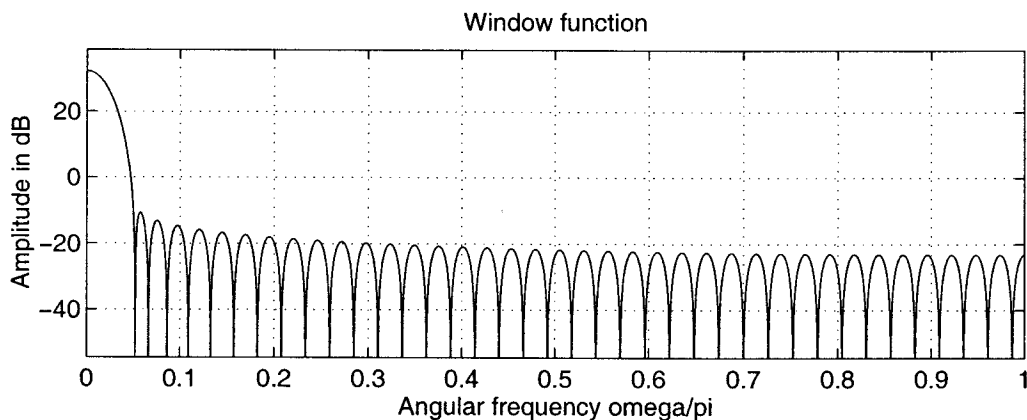


Hilbert Transformer Designed With the Aid of the Transitional Window

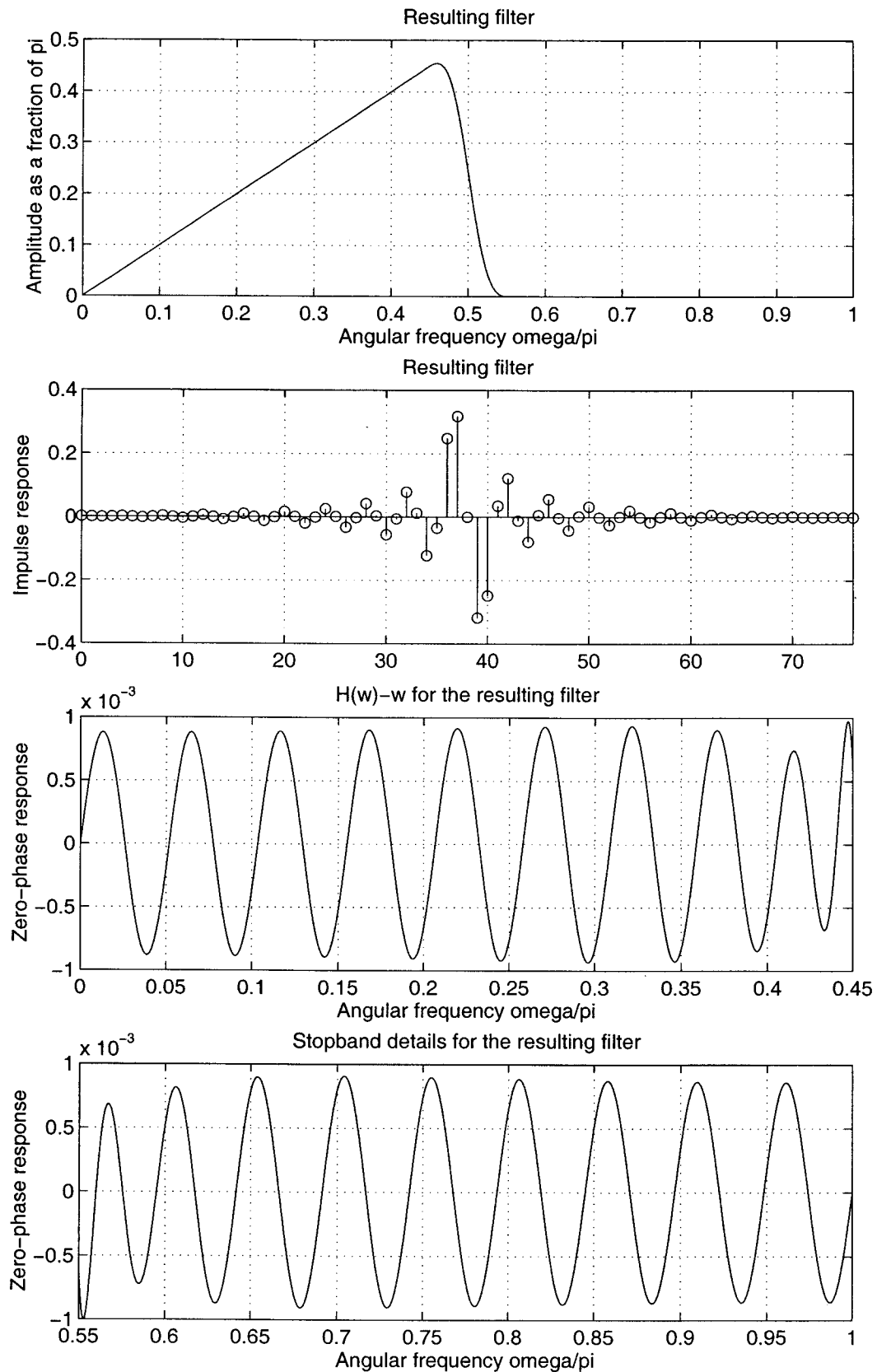


EXAMPLE 3

- It is desired to design with the aid of difwinad.m a differentiator such that its response approximates ω in the band $[0, 0.45\pi]$ with tolerance 0.001 and zero in the band $[0.55\pi, \pi]$ with the same tolerance.
- The minimum orders for the Kaiser, Saramäki, Dolph-Chebyshev, transitional windows are 78, 78, 82, and 76, respectively.
- The figures below and in the next transparency show the characteristics of the transitional window and the resulting filter.



Differentiator Designed With the Aid of the Transitional Window



SECTION 4: DESIGN OF LINEAR-PHASE FIR FILTERS IN THE LEAST-MEAN-SQUARE SENSE

- The second straightforward approach for designing FIR filters is based on the use of the least-squared approximation.
- The problem is to find the filter coefficients to minimize

$$E_2 = \int_X [W(\omega)[H(\omega) - D(\omega)]]^2 d\omega,$$

where X contains the passband and stopband regions, $D(\omega)$ is a desired response, and $W(\omega)$ is a positive weighting function.

- For the conventional lowpass case, $X = [0, \omega_p] \cup [\omega_s, \pi]$, $D(\omega)$ is unity in the passband and zero in the stopband.
- If $D(\omega)$ and $W(\omega)$ are sampled at a very dense grid of frequencies $\omega_1, \omega_2, \dots, \omega_K$ on X , then minimizing the above equation may be achieved by minimizing

$$E_2 = \sum_{k=1}^K [W(\omega_k)[H(\omega_k) - D(\omega_k)]]^2. \quad (1)$$

- According to above considerations, $H(\omega)$ can be expressed in the four different linear-phase cases in the form (see transparency 39)

$$H(\omega) = \sum_{n=0}^M b[n] \text{trig}(\omega, n)$$

- By substituting this for $H(\omega_k)$ in Eq. (1) and transferring $W(\omega_k)$ inside the parentheses yields

$$E_2 = \sum_{k=1}^K [W(\omega_k) \sum_{n=0}^M b[n] \text{trig}(\omega_k, n) - W(\omega_k) D(\omega_k)]^2.$$

This equation can be written in the following quadratic form

$$E = \mathbf{e}^T \mathbf{e},$$

where

$$\mathbf{e} = \mathbf{X}\mathbf{b} - \mathbf{d}$$

with

$$\mathbf{X} = \begin{pmatrix} W(\omega_1) \text{trig}(\omega_1, 0) & W(\omega_1) \text{trig}(\omega_1, 1) & \dots & W(\omega_1) \text{trig}(\omega_1, M) \\ W(\omega_2) \text{trig}(\omega_2, 0) & W(\omega_2) \text{trig}(\omega_2, 1) & \dots & W(\omega_2) \text{trig}(\omega_2, M) \\ \vdots & \vdots & \ddots & \vdots \\ W(\omega_K) \text{trig}(\omega_K, 0) & W(\omega_K) \text{trig}(\omega_K, 1) & \dots & W(\omega_K) \text{trig}(\omega_K, M) \end{pmatrix}$$

$$\mathbf{b} = [b[0], b[1], \dots, b[M]]^T$$

$$\mathbf{d} = [W(\omega_1)D(\omega_1), W(\omega_2)D(\omega_2), \dots, W(\omega_K)D(\omega_K)]^T$$

- Here, \mathbf{e} is a K length vector with the k -th element being $W(\omega_k)[H(\omega_k) - D(\omega_k)]$.

- The optimum solution of minimizing E_2 is given by

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{d} \quad (2)$$

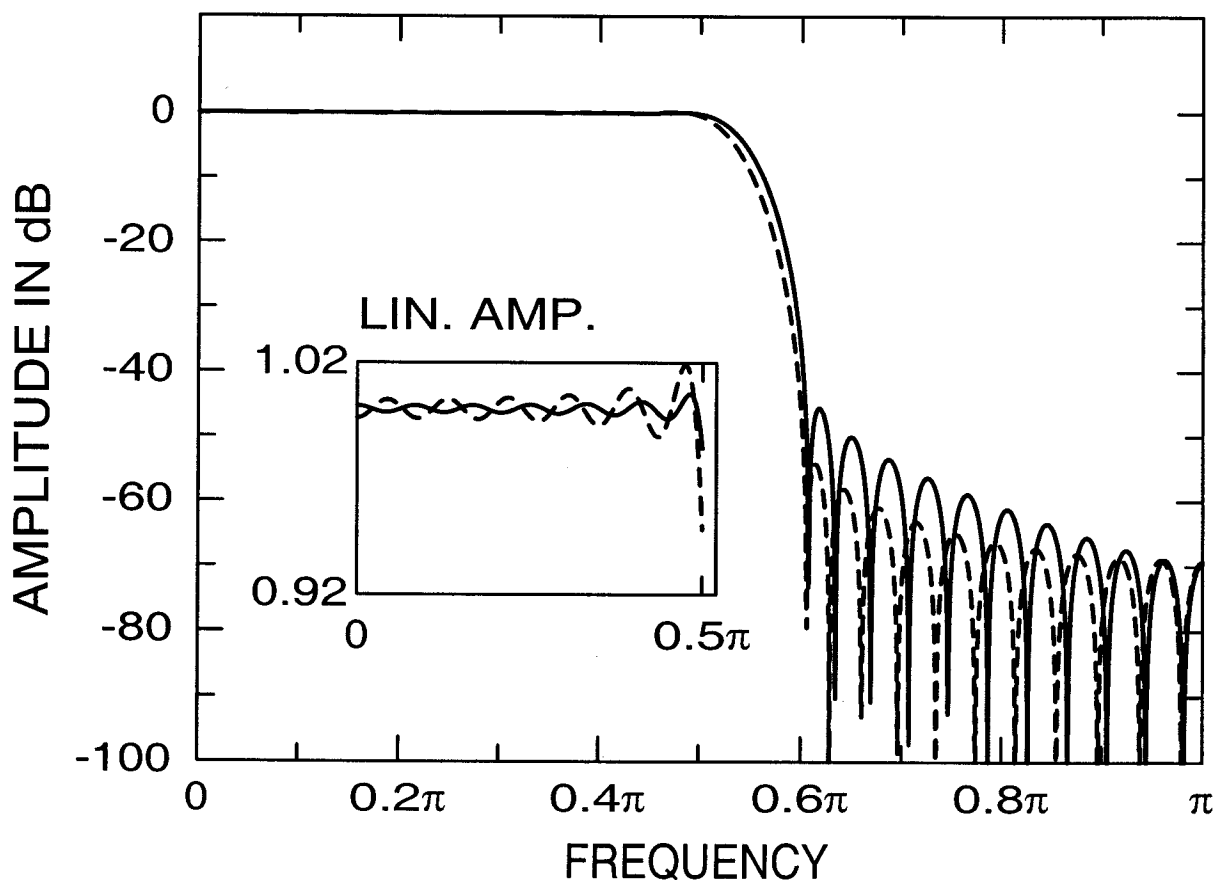
and it satisfies the "normal equations"

$$\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{X}^T \mathbf{d}.$$

- If K is much larger than M , then Eq. (2) should not be solved directly because it becomes ill conditioned.
- In this case, direct solution will probably have larger errors.
- Parks and Burrus recommend the use of the software package LINPACK, which has a special program for solving the above problem.
- In the case where both $W(\omega)$ and $D(\omega)$ are piecewise-constant functions, a significantly simpler procedure for finding the optimum solution can be generated. As a matter of fact, the matlab routine `firls.m` uses this alternative.
- Future work is devoted to generating a more general matlab routine.

Example

- Consider the design of a Type I filter of order $N = 46$ ($M = 23$) and having the passband and stopband edges of $\omega_p = 0.5\pi$ and $\omega_s = 0.6\pi$. $D(\omega) = 1$ on $[0, \omega_p]$ and $D(\omega) = 0$ on $[\omega_s, \pi]$.
- The following figure shows the resulting responses for two cases. In both cases, the $W(\omega) = 1$ on $[0, \omega_p]$, whereas $W(\omega) = W_s$ on $[\omega_s, \pi]$, where $W_s = 1$ in the first case and $W_s = 10$ in the second case.



Comments

- The effect of the stopband weighting is clearly seen from the figure of the previous transparency.
- It is also seen that the maximum deviations between the actual and the desired responses are much larger near the passband and stopband edges.
- This is characteristic of the least-squared-error designs.
- If the maximum deviations are desired to be minimized, then it is preferred to design the filter in the minimax sense.
- Compare the figure of the previous transparency to that of transparency 59, which gives a response for an FIR filter designed in the minimax sense.
- The filter orders in these two figures are the same.
- From the next transparency, you can find a matlab routine for designing the previous filters. Please try it!

MATLAB CODE GENERATING THE PREVIOUS FILTERS

```
%Design of FIR filters of order 46
%(Type I) in the least-mean-square sense
%with passband edge at 0.5pi, stopband
%edge at 0.6pi.
%Stopband weighting for the first filter
%is 1, and for the second one 100.
%Note that the weighting must be 10^2 in
%the case of the matlab routine to be used
%Tapio Saramaki 2.11.1995
%This can be found in SUN's
%~ts/matlab/dsp/luefir7.m
%First filter
h1=firls(46,[.0 .5 .6 1.], [1 1 0 0], [1 1]);
%Second filter
h2=firls(46,[.0 .5 .6 1.], [1 1 0 0], [1 100]);
%
[H,f]=freqz(h1,1,4*2048,2);
figure(1)
subplot(2,1,1)
plot(f,20*log10(abs(H))); axis([0 1 -100 10])
ylabel('Amplitude in dB');
xlabel('Angular frequency omega/pi');
title('Filter with stopband weighting of 1');
hold on;axes('position',[.21 .68 .3 .14]);
plot(f,(abs(H))); axis([0 .5 .98 1.02]);
xlabel('omega/pi');
title('Passband amplitude');
hold off;
subplot(2,1,2)
impz(h1);xlabel('n in samples');
ylabel('Impulse response');
figure(2)
zplane(h1);title('Zero-plot, stopband weighting is 1')
[H,f]=freqz(h2,1,4*2048,2);
figure(3);subplot(2,1,1)
plot(f,20*log10(abs(H))); axis([0 1 -100 10])
ylabel('Amplitude in dB');xlabel('Angular frequency omega/pi');
title('Filter with stopband weighting of 10');
hold on;axes('position',[.21 .68 .3 .14]);
plot(f,(abs(H))); axis([0 .5 .94 1.02]);xlabel('omega/pi');
title('Passband amplitude');hold off;
subplot(2,1,2);impz(h2);xlabel('n in samples');
ylabel('Impulse response');
figure(4):zplane(h2);title('Zero-plot, stopband weighting is 10')
```

MAXIMALLY-FLAT LINEAR-PHASE FIR FILTERS

- The third straightforward approach for designing FIR filters is to use filters with maximally flat response around $\omega = 0$ and $\omega = \pi$.
- The **advantages** of these filters are that the design is extremely simple and they are useful in applications where the signal is desired to be preserved with very small error near the zero frequency.
- In the case where the maximum deviation from the desired response is desired to be minimized, the **disadvantage** of these filters compared to the filters designed in the minimax sense is that the required filter order is much higher to meet the same criteria.

Conditions

- Consider a Type I filter with transfer function

$$H(z) = \sum_{n=0}^{2M} h[n]z^{-n}, \quad h[2M - n] = h[n].$$

- For maximally flat filters, it is advantageous to express $H(\omega)$ as an M th degree polynomial in $\cos \omega$ as follows (see transparency 61):

$$H(\omega) = \sum_{n=0}^M \alpha[n] \cos^n \omega.$$

- This $H(\omega)$ is determined in such a way that
 - $H(\omega)$ has $2K$ zeros at $\omega = \pi$.
 - $H(\omega) - 1$ has $2L = 2(M - K + 1)$ zeros at $\omega = 0$.
- M is thus related to L and K through $M = K + L - 1$.

Desired Solution

- The above conditions are satisfied if $H(\omega)$ can be written simultaneously in the forms

$$\begin{aligned} H(\omega) &= \left[\frac{1 + \cos \omega}{2} \right]^K \sum_{n=0}^{L-1} d[n] \left[\frac{1 - \cos \omega}{2} \right]^n \\ &= \cos^{2K}(\omega/2) \sum_{n=0}^{L-1} d[n] \sin^{2n}(\omega/2) \end{aligned}$$

and

$$\begin{aligned} H(\omega) &= 1 - \left[\frac{1 - \cos \omega}{2} \right]^L \sum_{n=0}^{K-1} \hat{d}[n] \left[\frac{1 + \cos \omega}{2} \right]^n \\ &= 1 - \sin^{2K}(\omega/2) \sum_{n=0}^{K-1} \hat{d}[n] \cos^{2n}(\omega/2). \end{aligned}$$

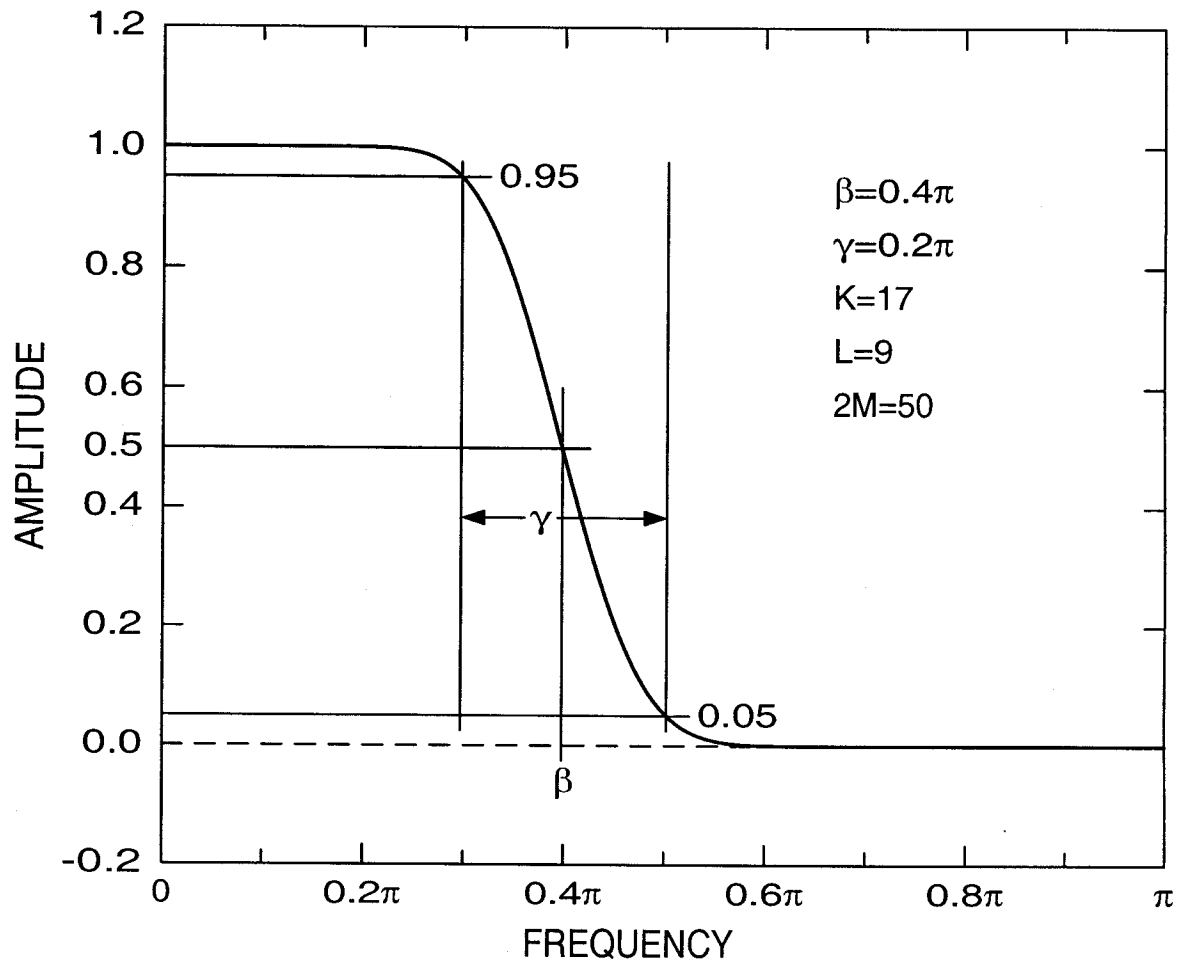
- The coefficients $d[n]$ and $\hat{d}[n]$ giving the desired solution are given by

$$d[n] = \frac{(K-1+n)!}{(K-1)!n!}, \quad \hat{d}[n] = \frac{(L-1+n)!}{(L-1)!n!}.$$

Characteristics of the Solution

- The resulting $H(\omega)$ is characterized by the following facts:
 - $H(\omega) = 1$ at $\omega = 0$ and its first $2L - 1$ derivatives are zero at this point.
 - $H(\omega) = 0$ at $\omega = \pi$ and its first $2K - 1$ derivatives are zero at this point.
- The primary unknowns of the above filters are K and L .
- Given the filter specifications, the problem is thus to determine these integers such that the given criteria are satisfied.

Specifications Stated by Kaiser



- Kaiser has stated the filter specifications as shown above.
- β is the center of the transition band.
- γ is the width of the transition band which is defined as the region where the response varies from 0.95 (passband edge angle) to 0.05 (stop-band edge angle).
- For meaningful specifications, γ has to satisfy $0 \leq \gamma \leq \min(2\beta, \pi - 2\beta)$.

Design Procedure Proposed by Kaiser

- The lower estimate for $M = K + L - 1$ (half the filter order) is given by

$$M_{\text{lower}} \approx (\pi/\gamma)^2.$$

- Then, ρ is determined by

$$\rho = (1 + \cos \beta)/2.$$

- The next step is to determine

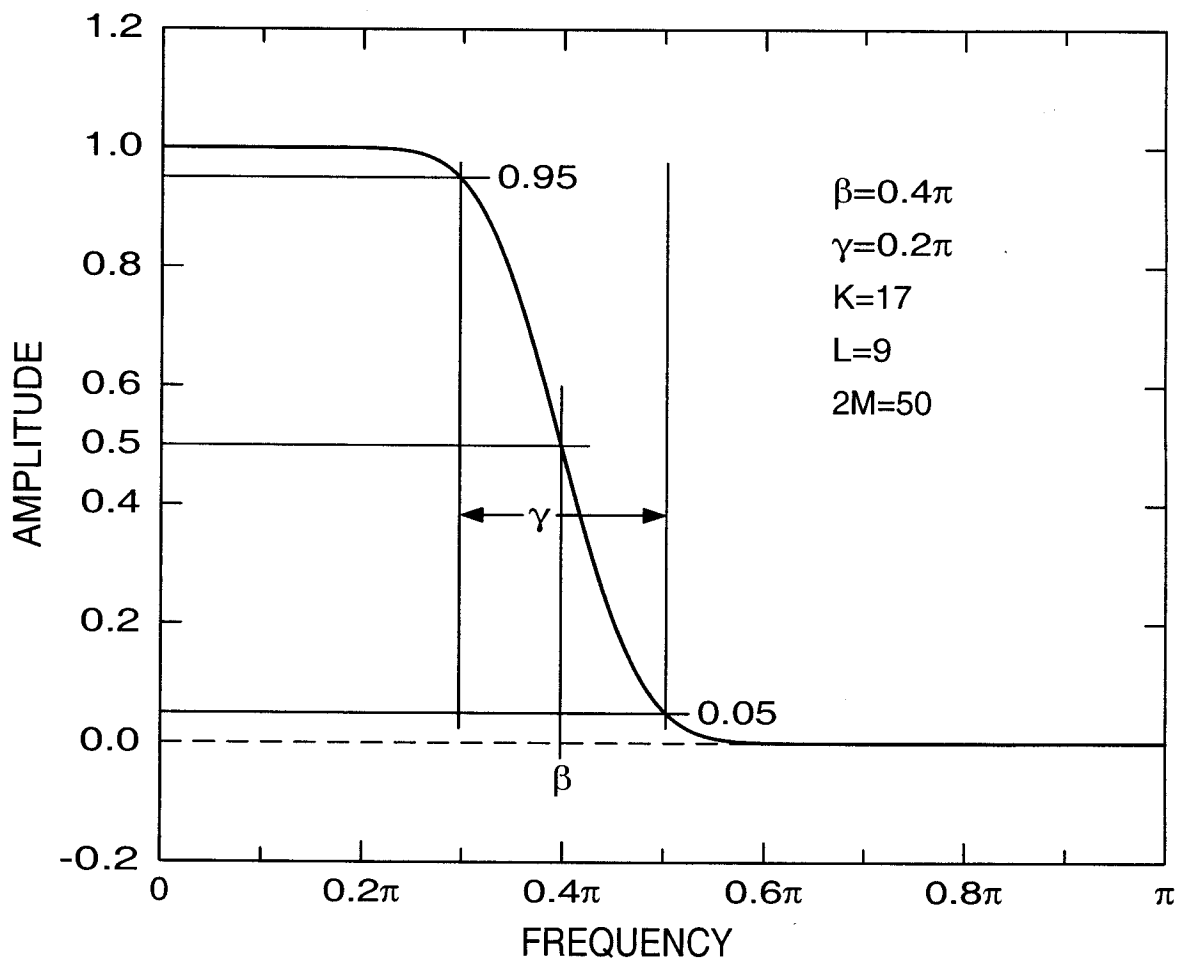
$$K_p = \langle \rho M_p \rangle,$$

where $\langle x \rangle$ stands for the nearest integer of x , for the values of M_p in the range $M_{\text{lower}} \leq M_p \leq 2M_{\text{lower}}$.

- Finally, the values of the integers K_p and M_p for which the ratio K_p/M_p is closest to ρ are selected.
- The desired values of K , L , M are then $K = K_p$, $L = M_p - K_p$, $M = M_p - 1$.
- With the above selections of K and L , the desired value of β can be achieved accurately.

Example

- Consider the specifications: $\beta = 0.4\pi$, $\gamma = 0.2\pi$.
- The above procedure results in $K = 17$ and $L = 9$. The order of the filter is thus $2(K + L - 1) = 50$.
- The amplitude response of this filter is depicted in the following figure.



Implementations

- The resulting $H(z)$ can be implemented using the conventional direct-form structure exploiting the coefficient symmetry (see transparencies 54 and 55).
- Alternatively, the transfer function can be written in the forms

$$H(z) = \left(\frac{1+z^{-1}}{2}\right)^{2K} \sum_{n=0}^{L-1} (-1)^n d[n] z^{-(L-1-n)} \left(\frac{1-z^{-1}}{2}\right)^{2n}$$

$$H(z) = z^{-M}$$

$$- (-1)^L \left(\frac{1-z^{-1}}{2}\right)^{2L} \sum_{n=0}^{K-1} \hat{d}[n] z^{-(K-1-n)} \left(\frac{1+z^{-1}}{2}\right)^{2n}.$$

- In the latter case, there are fewer multipliers, but the finite wordlength effects are worse.
- The next two transparency give a matlab routine to carry out the design procedure proposed by Kaiser and shows how the filter coefficients can be determined accurately.

Matlab routine for designing maximally-flat linear-phase FIR filters

```
% Matlab m-file (maxfir.m)
% This is a program for synthesizing maximally
% flat FIR filters according to the procedure
% suggested by Kaiser, see, e.g,
% T. Saramaki, "Finite impulse response Filter
% Design" in Handbook for Digital Signal Processing,
% S. K. Mitra and J. F. Kaiser, Eds, John Wiley &
% Sons, 1993, pages 193--195.
% Tapio Saramaki 1.2.96
% Can be found in SUN's: ~ts/matlab/dsp

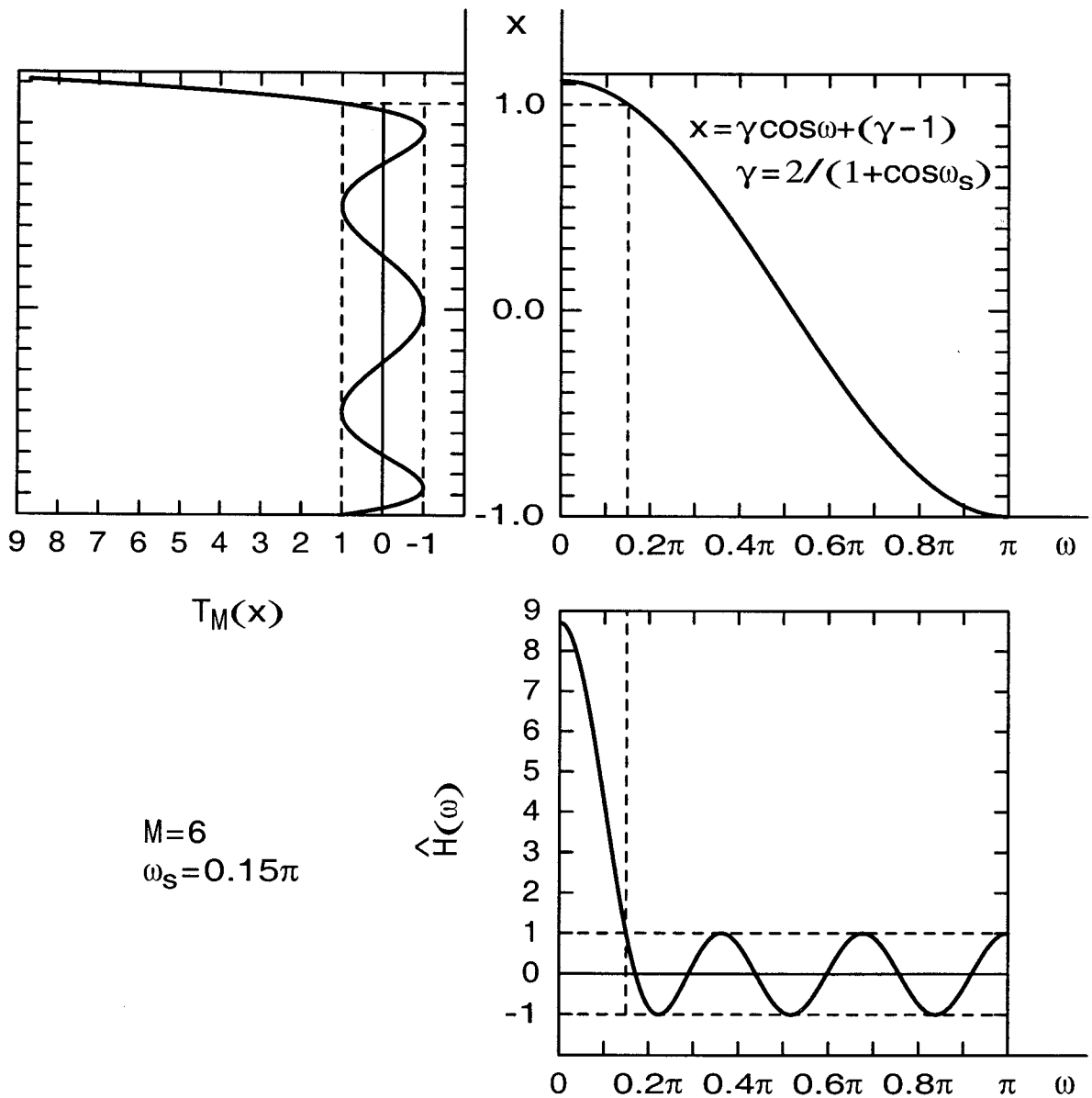
disp('This program designs maximally flat FIR filters');
disp('according to the procedure suggested by Kaiser. ');
disp('The parameters to be given are');
disp('beta, the center of the transition band ');
disp('gamma, the width of the transition band ');
disp('give both as a fraction of pi or as a fraction')
disp('of half the sampling rate')
disp('For meaningful specifications,');
disp('0<gamma<-min(2beta,2-2beta)')
beta=input('beta = ');
gamma=input('gamma = ');
beta=beta*pi;
gamma=gamma*pi;
Mlower=round((pi/gamma)^2);
rho=(1+cos(beta))/2;
help=200;
for k=Mlower:2*Mlower
Mp=k;
Kp=round(rho*Mp);
if abs(Kp/Mp-rho)< help
help=abs(Kp/Mp-rho);K=Kp;L=Mp-Kp;M=Mp-1;end
end
%
% Coefficients
l(1)=1
l(2)=1
for n=3:K-1+L l(n)=(n-1)*l(n-1);end
for n=0:L-1;d(n+1)=l(K-1+n+1)/(l(K-1+1)*l(n+1));end
%
% What is now remaining is to evaluate the impulse
```

```
% response coefficients of the resulting filter.
% To do this, we evaluate the zero-phase frequency
% response of our filter at  $2^l > 2M+1$  equally spaced
% frequencies and use the IFFT.
%
l=log2(2*M+1);l=floor(l)+1;k=2^l;
w=0:2*pi/k:2*(k-1)*pi/k;
A=((1+cos(w))/2).^K; B=zeros(size(w));
for n=0:L-1 B=B+d(n+1)*((1-cos(w))/2).^n;end
A=A.*B;
b=ifft(A);
b=real(b);
b=fftshift(b);
for k=1:2*M+1
h(k)=b(2^(l-1)+1-(M+1)+k);end
figure(1)
subplot(211);xlabel('n in samples');
impz(h);xlabel('n in samples');ylabel('Impulse response');
title(['Maximally-Flat FIR Filter with L = ', num2str(L), ' and K = ',
,num2str(K)])
[HH,w]=zeroam(h,.0,1.,1000);
subplot(212)
plot(w/pi,abs(HH));axis([0 1 -.01 1.01]);
xlabel('Frequency omega/pi');ylabel('Amplitude')
```

SECTION 6: SOME SIMPLE LINEAR-PHASE FIR FILTER DESIGNS

- There exist two special cases where the optimum solution in the minimax sense can be obtained analytically.
- The first analytically solvable case is the one where the zero-phase frequency response is monotonically decaying in the passband region and exhibits an equiripple behavior in the given stopband region $[\omega_s, \pi]$.
- An equiripple behavior on $[\omega_s, \pi]$ can be achieved by mapping the Chebyshev polynomial $T_M(x)$ to the ω -plane such that the region $[-1, 1]$, where $T_M(x)$ oscillates within ± 1 (see transparencies 62 and 63), is mapped to the region $[\omega_s, \pi]$.
- The next transparency shows how this mapping can be performed.

Generation of a zero-phase frequency response oscillating within the limits ± 1 in the stop-band $[\omega_s, \pi]$ based on mapping the M th degree Chebyshev polynomial $T_M(\omega)$ to the ω -plane



- As seen from the previous transparency, the desired transformation mapping $x = 1$ to $\omega = \omega_s$ and $x = -1$ to $\omega = \pi$ is given by

$$x = \gamma \cos \omega + (\gamma - 1), \quad \gamma = \frac{2}{1 + \cos \omega_s}.$$

- This result in the following zero-phase frequency response:

$$\hat{H}(\omega) = T_M[(2 \cos \omega + 1 - \cos \omega_s)/(1 + \cos \omega_s)].$$

- This is expressible as an M th order polynomial in $\cos \omega$.
- According to the discussion of transparency 61, the resulting $\hat{H}(\omega)$ is the zero-phase frequency response of a Type I filter of order $2M$.
- The response of the resulting filter shown in the previous transparency is not acceptable at all as it oscillates in the stopband region between ± 1 and achieves the value of almost 9 at $\omega = 0$.
- This problem can be solved by using a proper scaling, as will be described in the next transparency.

Scaled Response

- The response taking the value $1 + \delta_p$ at $\omega = 0$ is simply obtained by dividing $\hat{H}(\omega)$ by $\hat{H}(0)$ and multiplying by $1 + \delta_p$, giving

$$H(\omega) = (1 + \delta_p)\hat{H}(\omega)/\hat{H}(0)$$

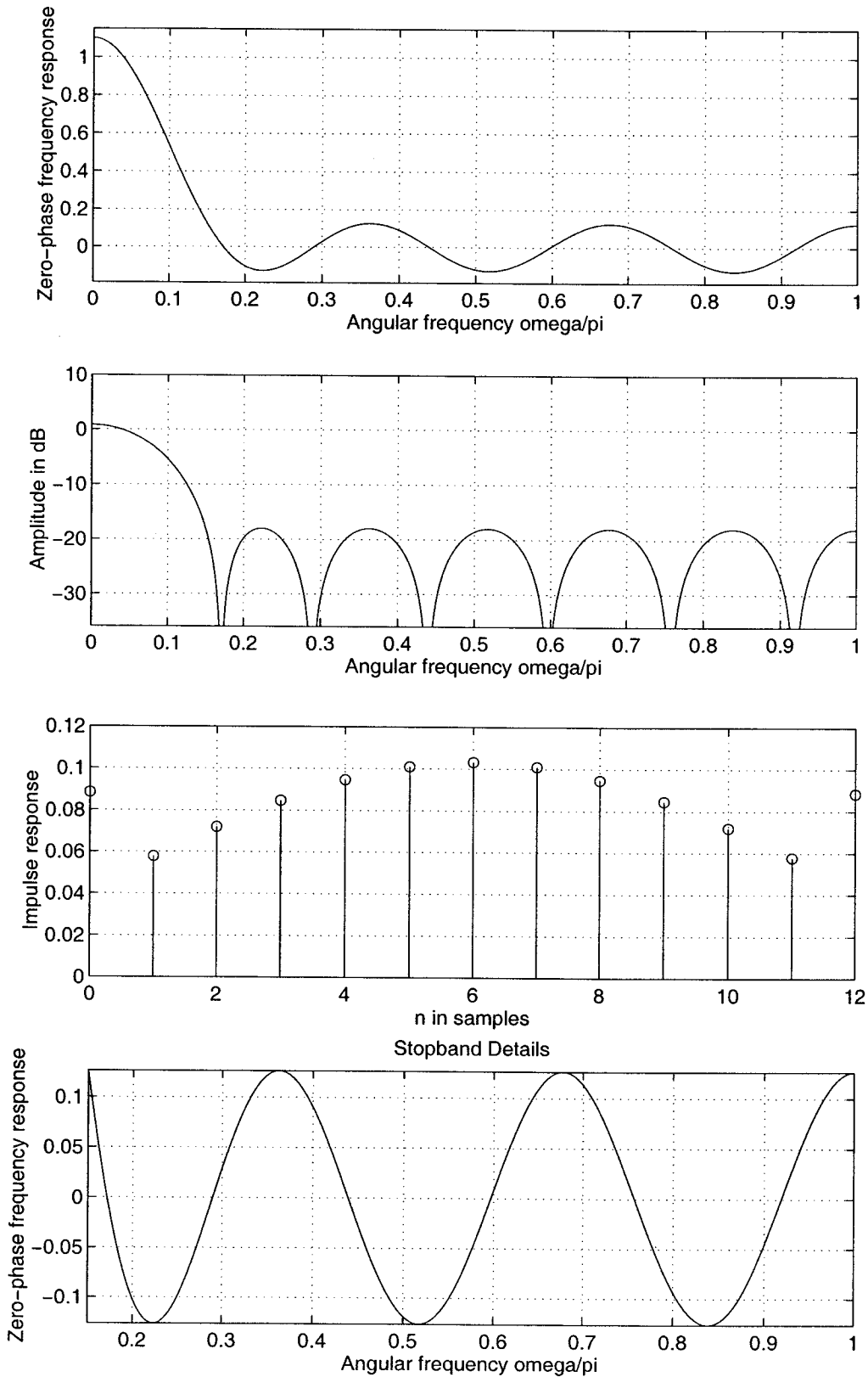
- This $H(\omega)$ oscillates on $[\omega_s, \pi]$ within the $\pm\delta_s$ where

$$\delta_s = (1 + \delta_p)/\hat{H}(0).$$

- The response of transparency 180 after scaling for $\delta_p = 0.1$ is shown in the next transparency.
- Based on the properties of Chebyshev polynomials, it can be shown that the minimum value of M (half the filter order) to give the specified ripples δ_s and δ_p is the smallest integer satisfying

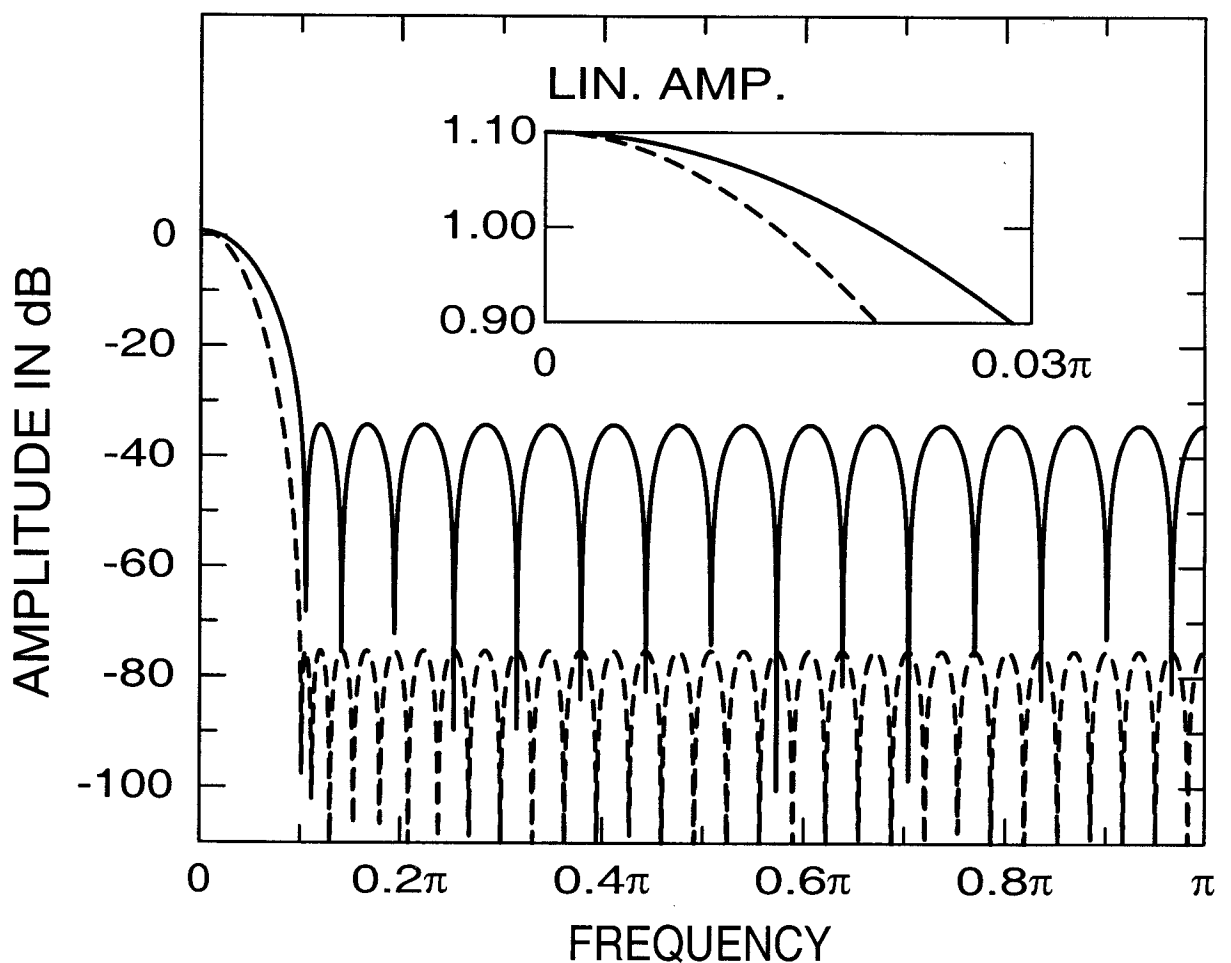
$$M \geq \frac{\cosh^{-1}[(1 + \delta_p)/\delta_s]}{\cosh^{-1}[(3 - \cos \omega_s)/(1 + \cos \omega_s)]}.$$

Response of the filter in transparency 180 after scaling it to take the value $1 + \delta_p$ with $\delta_p = 0.1$ at $\omega = 0$.



Example

- As an example, the following figure gives responses with $\omega_s = 0.1\pi$ and $\delta_p = 0.1$ for $M = 15$ and $M = 30$.
- The disadvantage of these designs is that all the zeros of the filter lie on the unit circle and the passband region where the response decays from $1 + \delta_p$ to $1 - \delta_p$ is narrow and cannot be controlled.



Matlab Routines for Designing Filters with Equiripple Stopband or Passband

- The next four transparencies give a matlab routine for generating filters having an equiripple stopband behavior.
- It can be used in the following two cases:
 - M , ω_s , and δ_p are specified and δ_s is determined according to the formula in transparency 182.
 - ω_s , δ_s , and δ_p are specified and M is determined according to the formula in transparency 182. Also δ_s is redetermined to be less than or equal to the specified value.
- The same routine can be used for designing linear-phase FIR filters with its zero-phase frequency response oscillating within $1 \pm \delta_p$ in the passband $[0, \omega_p]$ and being monotonically decaying outside this band achieving the value of $-\delta_s$ at $\omega = \pi$.

Matlab routine for designing linear-phase FIR filters with equiripple stopband or passband

% This program (simple.m) designs filters with equiripple
% stopband or passband region with the aid of Chebyshev
% polynomials

%-----

% Tapio Saramaki 14.11.1997

% This can be found in SUN's:

% ~ts/matlab/dsp/simple.m

%-----

```
disp('Hi there, I am a program of designing FIR filters')
disp('with the aid of Chebyshev polynomials')
disp('The order of the resulting filter is always twice')
disp('the degree of the Chebyshev polynomial')
disp('As input data, I need the following:')
disp('1 for filter with equiripple stopband');
itype1=input('2 for filter with equiripple passband: ');
disp('1 for specifying the order of the Chebyshev polynomial');
itype2=input('2 for specifying the filter criteria: ');
if itype2==1
    M=...
    input('half the filter order or the degree of the polynomial: ');
    if itype1==1
        oms=input('stopband egde as a fraction of pi: ');
        disp('the value at omega=0 is given as 1+delta_p');
        dp=input('delta_p: ');
    end
    if itype1==2
        omp=input('passband egde as a fraction of pi: ');
        disp('the value at omega=pi is given as -delta_s');
        ds=input('delta_s: ');
    end
end
end
if itype2==2
    if itype1==1
        oms=input('stopband egde as a fraction of pi: ');
        ds=input('stopband ripple delta_s: ');
        disp('the value at omega=0 is given as 1+delta_p');
        dp=input('delta_p: ');
        num=acosh((1+dp)/ds);
        den=acosh((3-cos(pi*oms))/(1+cos(pi*oms)));
        M=ceil(num/den);
    end
end
```

```
if itype1==2
    omp=input('passband egde as a fraction of pi: ');
    dp=input('passband ripple around unity delta_p');
    disp('the value at omega=pi is given as -delta_s');
    ds=input('delta_s: ');
    num=acosh((1+ds)/dp);
    den=acosh((3-cos(pi*(1-omp)))/(1+cos(pi*(1-omp))));
    M=ceil(num/den);
end
end
%-----
% Filter with equiripple stopband generated using the
% subroutine chepols(M,oms)
%-----
if itype1==1
    h=chepols(M,oms);
%-----
% Value at omega=0
%-----
[H,w]=zeroam(h,.0,.0,1);
h=(1+dp)*h/H(1);
ds=(1+dp)/H(1);
end
%-----
% Filter with equiripple passband generated using the
% subroutine chepols(M,oms)
%-----
if itype1==2
    oms=1-omp;
    h=chepols(M,oms);
    for k=1:2:M
        h(M+1-k)=-h(M+1-k);
        h(M+1+k)=-h(M+1+k);
    end
%-----
% Value at omega=pi
%-----
[H,w]=zeroam(h,1.,1.,1);
dp=(1+ds)/H(1);
h=-dp*h;
h(M+1)=h(M+1)+1;
end
%-----
% Plot the responses
%-----
```



```
[H,w]=zeroam(h,.0,1.,8000);
figure(1)
amax=1+1.5*max(H-1);
amin=1.5*min(H);
subplot(211)
plot(w/pi,H);axis([0 1 amin amax]); grid;
ylabel('Zero-phase frequency response');
xlabel('Angular frequency omega/pi');
subplot(212)
amax=40*log10(abs(ds));
HH=20*log10(abs(H));
plot(w/pi,HH); axis([0 1 amax 10]);grid;
ylabel('Amplitude in dB');
xlabel('Angular frequency omega/pi');
figure(2)
subplot(211)
impz(h);ylabel('Impulse response');grid;
xlabel('n in samples');
subplot(212)
if itype1==1
    plot(w/pi,H);grid;axis([0 1 -ds ds]);
    title('Stopband Details')
    ylabel('Zero-phase frequency response');
    xlabel('Angular frequency omega/pi');
end
if itype1==2
    plot(w/pi,H);grid;axis([0 1 -dp 1+dp]);
    title('Passband Details')
    ylabel('Zero-phase frequency response');
    xlabel('Angular frequency omega/pi');
end
end
*****
function [h]=chepols(M,omegas)
%-----
% [h]=chepols(N,omegas) evaluates the impulse response
% values of a Type I filter with zero-phase frequency
% response given by  $H(\omega)=T_M[\gamma\cos(\omega)+$ 
%  $(\gamma-1)]$  with  $\gamma=2/1+\cos(\omega)$ . The resulting
%  $H(\omega)$  oscillates within the +1 and -1 on
%  $[\omega, \pi]$ . The order of the resulting filter is
% 2M.
% omegas is given as a fraction of pi
%-----
```

```
%-----  
% Programmed by Tapio Saramaki, 14.11. 1997.  
% This can be found in SUN's  
% ~ts/matlab/dsp/chepols.m  
%-----  
gamma=2/(1+cos(pi*omegas));  
gamma1=gamma/2;  
gamma2=gamma-1;  
w1(1,2*M+1)=0; w2(1,M+2)=0; w3(1,M+2)=0;  
w1(1)=1; w2(1)=gamma2; w2(2)=gamma1;  
for i=2:M  
    w3(1)=2*gamma1*(w2(2)+w2(2))+2*gamma2*w2(1)-w1(1);  
    for k=2:i+1  
        w3(k)=2*gamma1*(w2(k-1)+w2(k+1))+2*gamma2*w2(k)-w1(k);  
    end  
    for k=1:M+1 w1(k)=w2(k); w2(k)=w3(k); end  
end  
for i=1:M+1 w1(i)=w3(M+2-i); end  
for i=2:M+1 w1(i+M)=w3(i); end  
h=w1;
```

Filters with Equiripple Passband

- The response which is equiripple in the passband $[0, \omega_p]$ oscillating within $1 \pm \delta_p$ and monotonically decaying in the region $[\omega_p, \pi]$ can be derived in the same manner.

- The desired response is

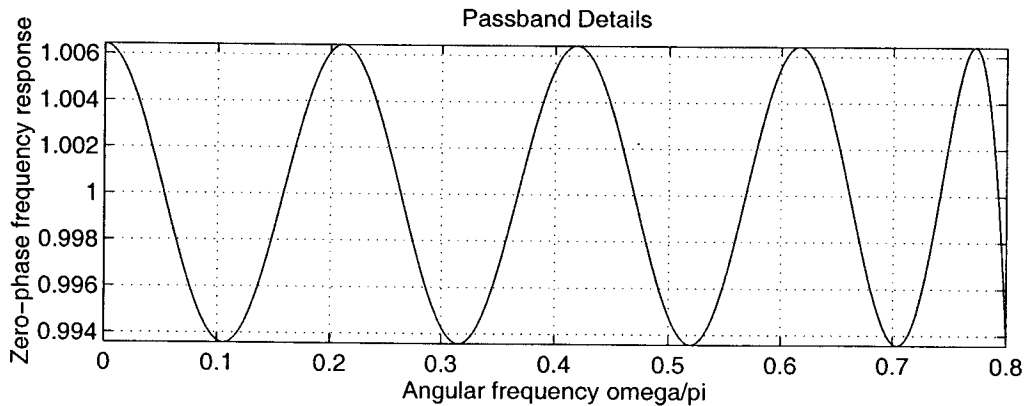
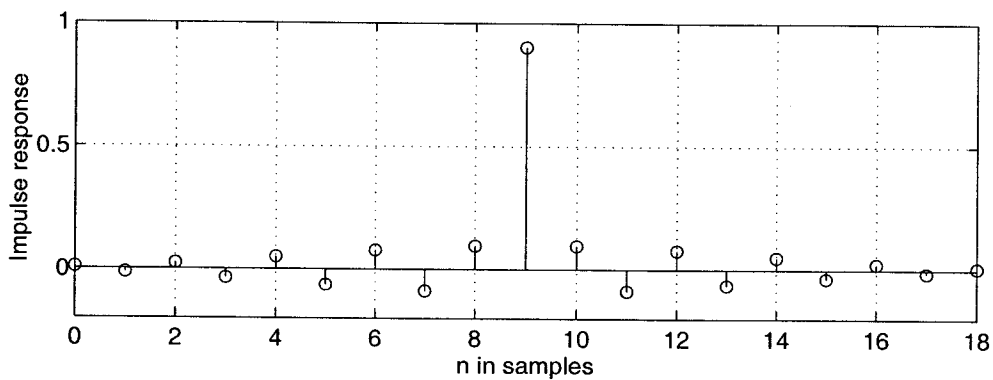
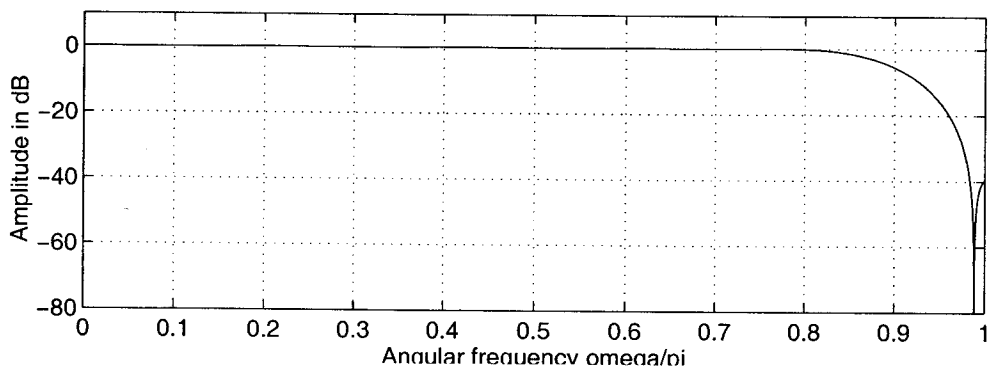
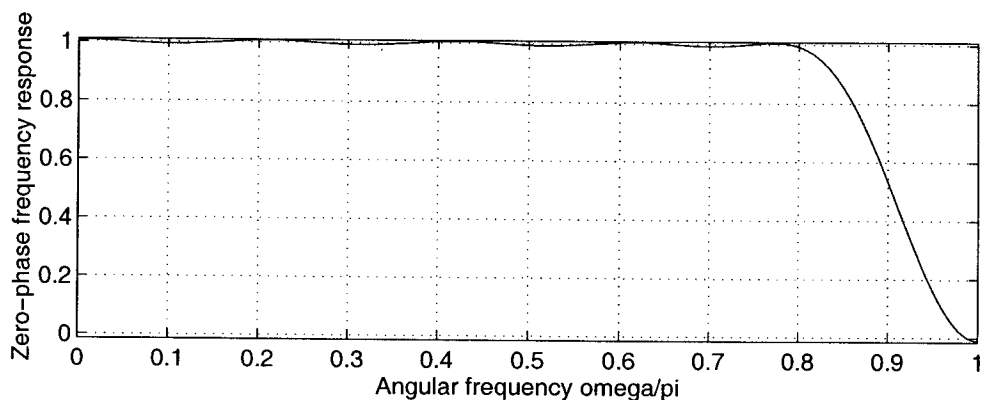
$$H(\omega) = 1 - \delta_p T_M[(-2 \cos \omega + 1 + \cos \omega_p)/(1 - \cos \omega_p)].$$

- If it is desired that $H(\pi) = -\delta_s$, then δ_p can be determined from

$$\delta_p = (1 + \delta_s)/T_M[(3 + \cos \omega_p)/(1 - \cos \omega_p)].$$

- The minimum value of M required to meet the given ripple requirements can be solved from the equation given previously in transparency 182 by interchanging δ_p and δ_s and by replacing ω_s by $\pi - \omega_p$.
- The next transparency shows the optimized filter for $\omega_p = 0.9\pi$ and $\delta_p = \delta_s = 0.01$.

Optimized Filter with Equiripple Behavior in the Passband $[0, 0.8\pi]$: $\delta_p = \delta_s = 0.01$



SECTION 7: DESIGN OF LINEAR-PHASE FIR FILTERS IN THE MINIMAX SENSE

- One of the main advantages of FIR filters over their IIR counterparts is that there exists an efficient algorithm for optimizing in the minimax sense **arbitrary-magnitude** FIR filters.
- For IIR filters, the design of arbitrary-magnitude filters is usually time-consuming and the convergence to the best solution is not always guaranteed.
- The most efficient method for designing optimum magnitude FIR filters with arbitrary specifications is the Remez multiple exchange algorithm.
- The most frequently used method for implementing this algorithm is the one originally advanced by Parks and McClellan.
- The actual program has been written by McClellan, Parks, and Rabiner.
- This is why this method is referred later to as the **MPR** algorithm.

Usefulness of the Algorithm

- This program is directly applicable to obtaining optimal designs for most types of FIR filters like lowpass, highpass, bandpass, and bandstop filters, Hilbert transformers, and digital differentiators.
- Also filters having several passbands and stopbands can be directly designed.
- Filters having some constraints in the time or frequency domains cannot be directly designed
- **Linear programming** can be used in most of these cases.
- We start this section with theory, but don't worry: The theory becomes, hopefully, clear with the aid of examples.

When Are We Able to Use the Remez Algorithm?

- The Remez multiple exchange algorithm is the most powerful algorithm for finding the coefficients $a[n]$ of the function

$$G(\omega) = \sum_{n=0}^M a[n] \cos n\omega$$

minimizing on a closed subset X of $[0, \pi]$ the peak absolute value of the following weighted error function

$$E(\omega) = \widehat{W}(\omega)[G(\omega) - \widehat{D}(\omega)],$$

that is, the quantity

$$\epsilon = \max_{\omega \in X} |E(\omega)|.$$

- All that is required is that $\widehat{D}(\omega)$ is continuous on X and $\widehat{W}(\omega) > 0$.
- For designing FIR filters, X is simply a union of the passband and stopband regions, as will be seen later on.

How to Use This Algorithm for Designing Linear-Phase Type I, II, III, and IV FIR Filters

- According to the discussion of Section 2, in all the four linear-phase cases, the zero-phase frequency response $H(\omega)$ of a filter of order N can be expressed as (see transparencies 42 and 49)

$$H(\omega) = F(\omega)G(\omega),$$

where

$$G(\omega) = \sum_{n=0}^M a[n] \cos n\omega,$$

$$F(\omega) = \begin{cases} 1 & \text{for Type I} \\ \cos(\omega/2) & \text{for Type II} \\ \sin \omega & \text{for Type III} \\ \sin(\omega/2) & \text{for Type IV,} \end{cases}$$

and

$$M = \begin{cases} N/2 & \text{for Type I} \\ (N - 1)/2 & \text{for Type II} \\ (N - 2)/2 & \text{for Type III} \\ (N - 1)/2 & \text{for Type IV.} \end{cases}$$

What Do We Get?

- If the desired function for $H(\omega)$ on X is $D(\omega)$ and the weighting function is $W(\omega)$, then the error function can be written into the desired form as follows:

$$\begin{aligned} E(\omega) &= W(\omega)[H(\omega) - D(\omega)] \\ &= W(\omega)[F(\omega)G(\omega) - D(\omega)] \\ &= W(\omega)F(\omega)[G(\omega) - D(\omega)/F(\omega)] \\ &= \widehat{W}(\omega)[G(\omega) - \widehat{D}(\omega)], \end{aligned}$$

where

$$\widehat{W}(\omega) = F(\omega)W(\omega), \quad \widehat{D}(\omega) = D(\omega)/F(\omega).$$

- This error function is of the form to which we can directly apply the Remez algorithm!!

Characterization Theorem:

Let $G(\omega)$ be

$$G(\omega) = \sum_{n=0}^M a[n] \cos n\omega.$$

Then $G(\omega)$ is the best unique solution minimizing

$$\epsilon = \max_{\omega \in X} |E(\omega)|,$$

where

$$E(\omega) = \widehat{W}(\omega)[G(\omega) - \widehat{D}(\omega)]$$

if and only if there exists at least $M + 2$ points $\omega_1, \omega_2, \dots, \omega_{M+2}$ in X such that

$$\omega_1 < \omega_2 < \dots < \omega_{M+1} < \omega_{M+2}$$

$$E(\omega_{i+1}) = -E(\omega_i), \quad i = 1, 2, \dots, M + 1$$

$$|E(\omega_i)| = \epsilon, \quad i = 1, 2, \dots, M + 2.$$

- In other words, the optimum solution is characterized by the fact that the weighted error function $E(\omega)$ alternately achieves the values $\pm\epsilon$, with ϵ being the peak absolute value of the weighted error, at least at $M + 2$ consecutive points in X .

What Does This Theorem Mean?

- To illustrate the meaning of the above characterization theorem, consider the figure of the next transparency showing the response of a typical optimum Type I lowpass filter of order $N = 12$ and the corresponding error function.
- For this Type I filter, $M = N/2 = 6$, $H(\omega) \equiv G(\omega)$, $\hat{D}(\omega) = D(\omega)$ and $\hat{W}(\omega) = W(\omega)$.
- The weighted error function is

$$E(\omega) = \hat{W}(\omega)[G(\omega) - \hat{D}(\omega)],$$

where

$$G(\omega) = \sum_{n=0}^6 a[n] \cos n\omega,$$

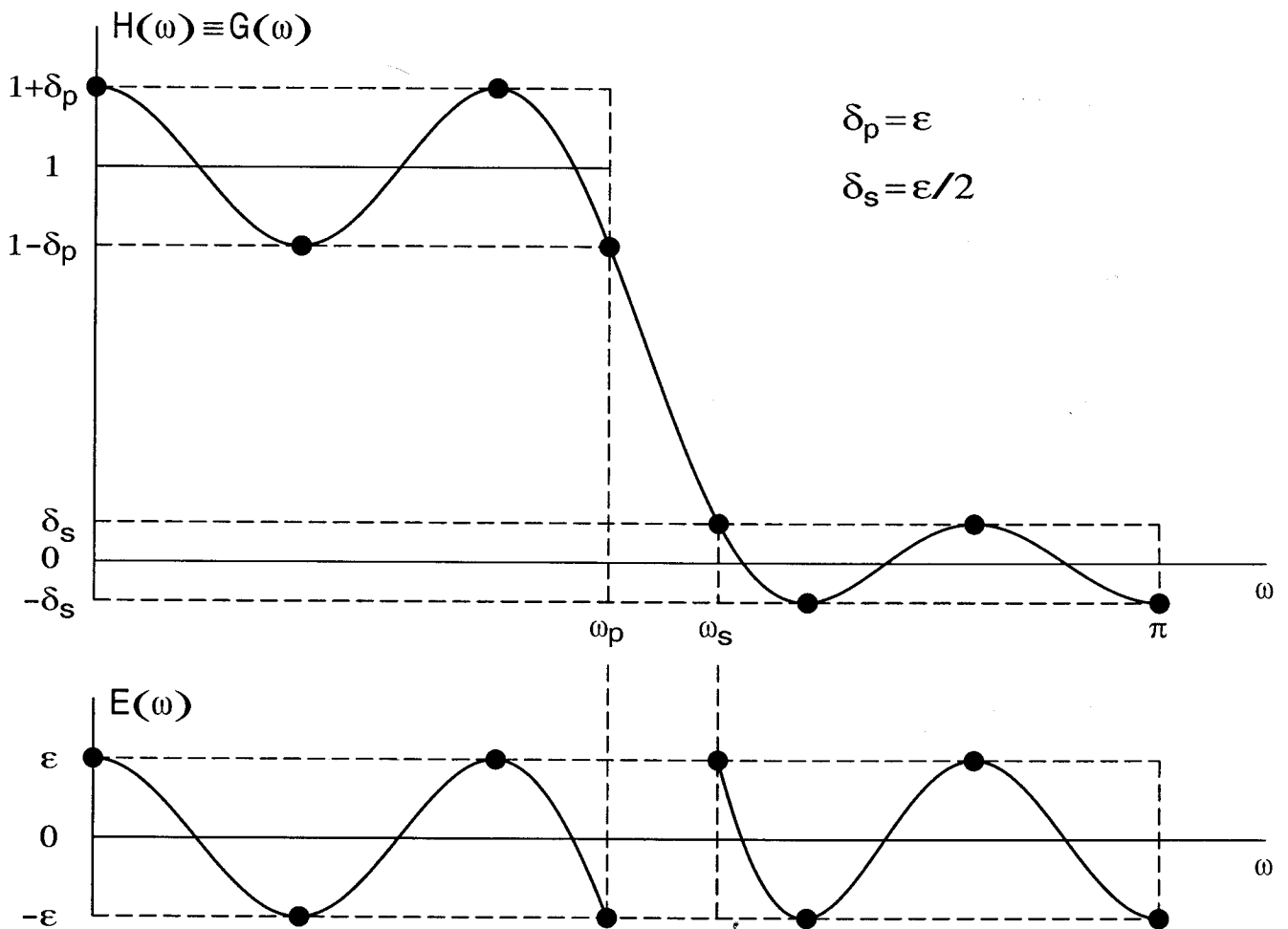
$$\hat{D}(\omega) = \begin{cases} 1, & \omega \in [0, \omega_p] \\ 0, & \omega \in [\omega_s, \pi], \end{cases}$$

$$\hat{W}(\omega) = \begin{cases} 1, & \omega \in [0, \omega_p] \\ 2, & \omega \in [\omega_s, \pi], \end{cases}$$

and

$$X = [0, \omega_p] \cup [\omega_s, \pi].$$

A Typical Optimum Type I Filter of Order $N = 12$: $M = 6$ and $H(\omega) \equiv G(\omega)$. $\widehat{D}(\omega) = D(\omega) = 1$ and $\widehat{W}(\omega) = W(\omega) = 1$ for $\omega \in [0, \omega_p]$, whereas $\widehat{D}(\omega) = D(\omega) = 0$ and $\widehat{W}(\omega) = W(\omega) = 2$ for $\omega \in [\omega_s, \pi]$.



- As seen from the figure of the previous transparency, for the optimum solution, there exist $M + 2 = 8$ extremal points ω_k for $k = 1, 2, \dots, 8$ (marked by dots) in X (four in the passband and four in the stopband) where $|E(\omega)|$ achieves the value of ϵ such that $E(\omega_{k+1}) = -E(\omega_k)$ for $k = 0, 1, \dots, 7$.
- $G(\omega)$ contains seven unknowns $a[0], a[1], \dots, a[6]$ so that there are one is **one more** extremal point than there are unknowns.
- Another very crucial fact to point out is:
- Since the stopband weighting is two and the passband weighting is two, the passband ripple $\delta_p = \epsilon$ and the stopband ripple is $\delta_s = \epsilon/2 = \delta_p/2$.

- In general, for the optimum solution there exists $M + 2$ points ω_k in X such that

$$E(\omega_k) = \widehat{W}(\omega_k)[G(\omega_k) - \widehat{D}(\omega_k)] = (-1)^k \epsilon.$$

- Here, ϵ is either positive or negative and the pear absolute value is $|\epsilon|$.
- For the unweighted “real” error $G(\omega) - \widehat{D}(\omega)$ this implies that

$$G(\omega_k) - \widehat{D}(\omega_k) = (-1)^k \epsilon / \widehat{W}(\omega_k).$$

- This shows that the deviation of $G(\omega)$ form $\widehat{D}(\omega)$ is forced to become smaller in those parts of X where $\widehat{W}(\omega)$ is larger.
- The larger is the relative difference in the weight function $\widehat{W}(\omega)$, the larger is the difference in the deviation.

How to Use the Theorem to Check the Optimality of a Lowpass Filter Solution?

- According to the theorem, it is thus easy to check whether a given lowpass filter solution is the optimum one.
- If the relative weighting between the passband and stopband errors is K and there exists a solution $H(\omega)$ which alternately goes through the values $1 \pm \epsilon$ in the passband and through the values $\pm\epsilon/K$ in the stopband and the overall number of these extrema is $M + 2$, then this solution is, according to the characterization theorem, the best unique solution.
- It should be noted that in the lowpass case, both ω_p and ω_s are always extremal points, and $H(\omega_p) = 1 - \epsilon$ and $H(\omega_s) = \epsilon/K$ so that $E(\omega_p) = -\epsilon$ and $E(\omega_s) = \epsilon$.

Starting Point for Constructing the McClellan-Parks-Rabiner (MPR) Algorithm

- Given a set of $M + 2$ points on X , denoted by $\Omega = \{\omega_1, \omega_2, \dots, \omega_{M+2}\}$, the unknowns coefficients $a[0], a[1], \dots, a[M]$ and ϵ can be determined such that $E(\omega)$ satisfies

$$E(\omega_k) = \widehat{W}(\omega_k)[G(\omega_k) - \widehat{D}(\omega_k)] = (-1)^k \epsilon,$$

$$k = 1, 2, \dots, M + 2.$$

- This can be achieved by solving for the unknowns the following system of $M + 2$ linear equations:

$$\sum_{n=0}^M a[n] \cos n\omega_k - (-1)^k \epsilon / \widehat{W}(\omega_k) = \widehat{D}(\omega_k),$$

$$k = 1, 2, \dots, M + 2.$$

- This $E(\omega)$ goes alternately through the values $\pm \epsilon$ at the points ω_k .
- If X consists of the above set of $M + 2$ points, i.e. $X = \Omega$, then $|\epsilon|$ is the peak absolute value of $E(\omega)$ on X and the conditions of the above characterization theorem are satisfied.

- The Remez exchange algorithm makes use of the above fact.
- The problem is simply to find a set Ω on X in such a way that the optimum solution on Ω is simultaneously the optimum solution on the overall set X .
- This is achieved if value of $|\epsilon|$ is simultaneously the peak absolute value of $E(\omega)$ on the overall set X .
- We start by giving the Remez algorithm in the mathematical form.
- After that, a simple filter design problem is considered to make this algorithm easier to understand.

The Remez Algorithm for Iteratively Finding the Desired Set of $M + 2$ Extremal Points

1. Set $l = 1$. Select an initial set of $M + 2$ extremal points $\Omega^{(l)} = \{\omega_1^{(l)}, \omega_2^{(l)}, \dots, \omega_{M+2}^{(l)}\}$ in X .
2. Solve the following system of $M + 2$ linear equations

$$\sum_{n=0}^M a^{(l)}[n] \cos n\omega_k^{(l)} - (-1)^k \epsilon^{(l)} / \widehat{W}(\omega_k^{(l)}) = \widehat{D}(\omega_k^{(l)}),$$

$$k = 1, 2, \dots, M + 2$$
 for the unknowns $a^{(l)}[0], \dots, a^{(l)}[M]$ and $\epsilon^{(l)}$.
3. Find on X , $M + 2$ extremal points of the resulting $E^{(l)}(\omega) = W(\omega)[G^{(l)}(\omega) - D(\omega)]$ where $|E^{(l)}(\omega)| \geq |\epsilon^{(l)}|$. If there are more than $M + 2$ extremal points, retain $M + 2$ extrema such that the largest absolute values are included with the condition that the sign of the error function $E^{(l)}(\omega)$ alternate at the selected points. Store the abscissae of the extrema into $\Omega^{(l+1)} = \{\omega_1^{(l)}, \omega_2^{(l+1)}, \dots, \omega_{M+2}^{(l+1)}\}$.
4. If $|\omega_k^{(l)} - \omega_k^{(l+1)}| \leq \alpha$ for $k = 1, 2, \dots, M + 2$ (α is a small number), then go to the next step. Otherwise, set $l = l + 1$ and go to Step 2.
5. Calculate the filter coefficients.

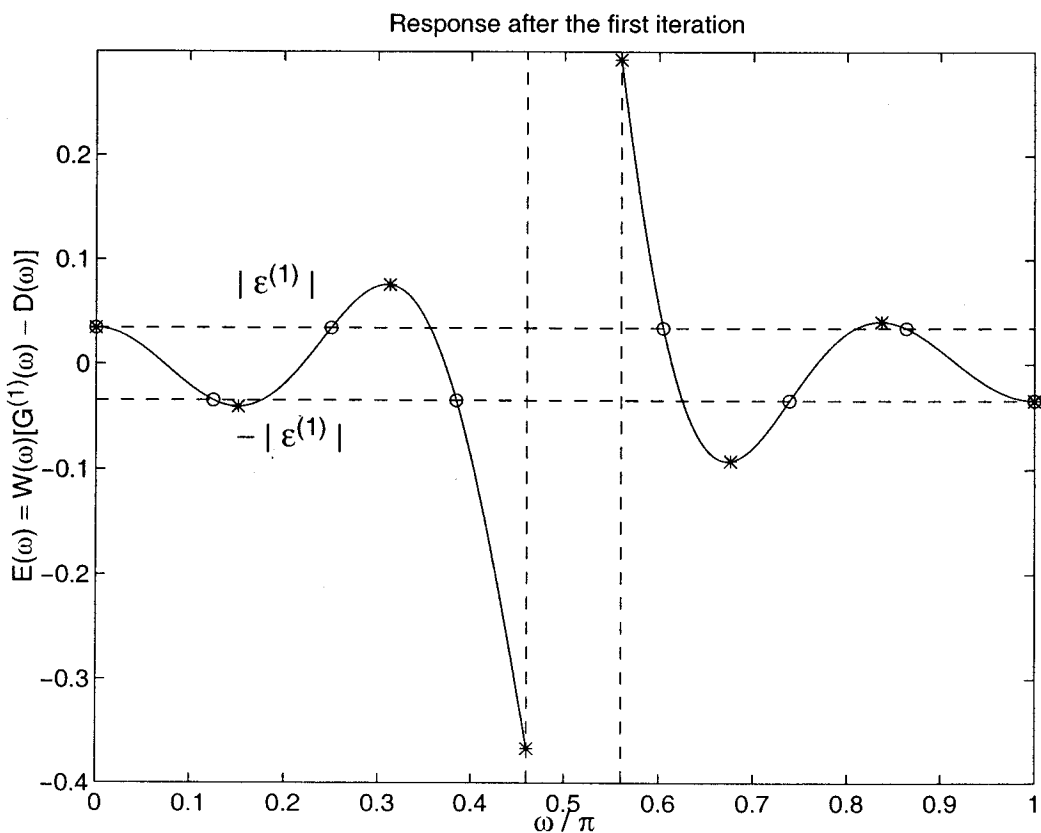
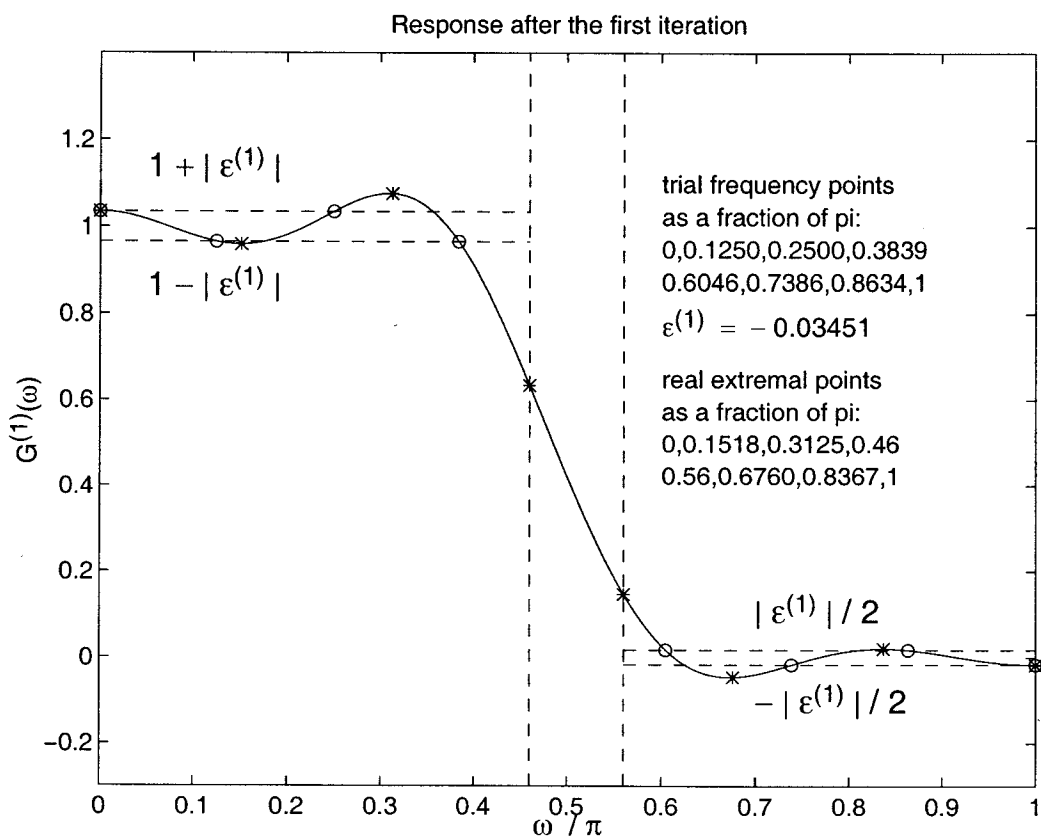
EXAMPLE ON HOW THE REMEZ ALGORITHM WORKS

- As an example, we consider the design of a low-pass filter of order 12 having the passband and stopband edges at 0.46π and 0.56π , respectively. The weighting in the passband is 1 and in the stopband 2.
 - These are the same criteria as those for the filter in transparency 199.
 - The three transparencies following the next one illustrate the responses after the first, second, and third iterations.
 - After the third iteration, we arrive at the desired optimum solution.
 - The trial extremal points $\omega_k^{(l)}$ are denoted by circles, whereas the true extremal points $\omega_k^{(l+1)}$ by asterisks.

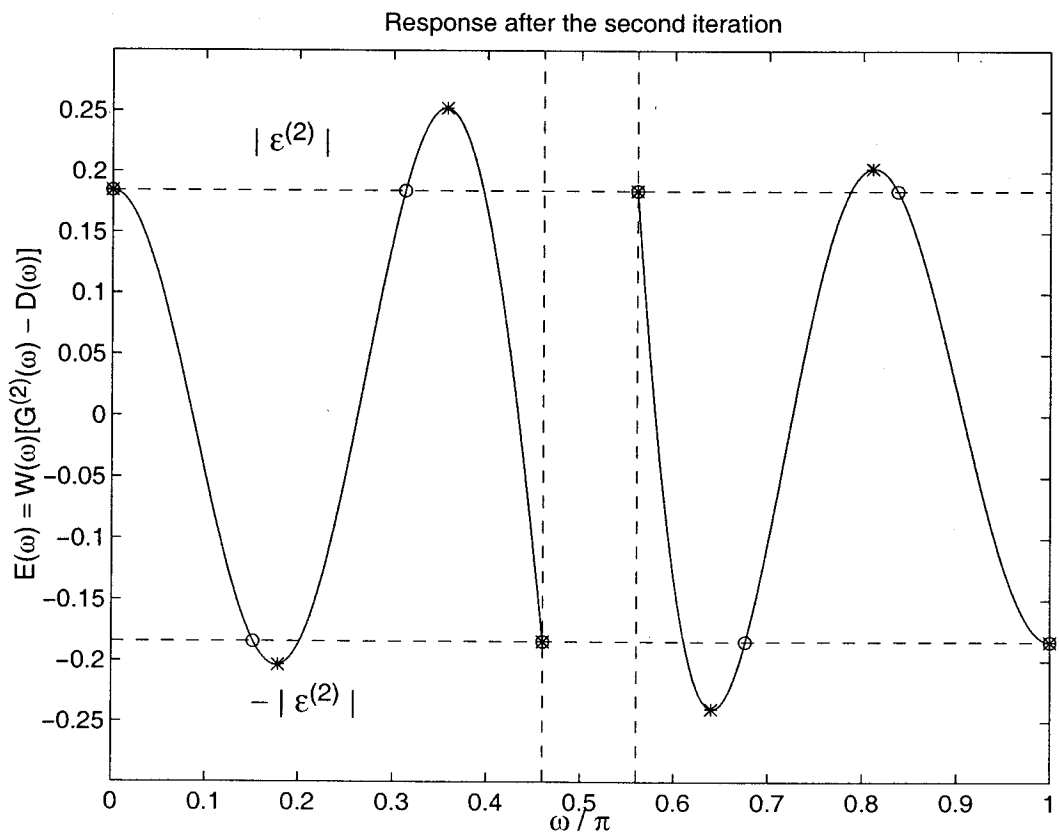
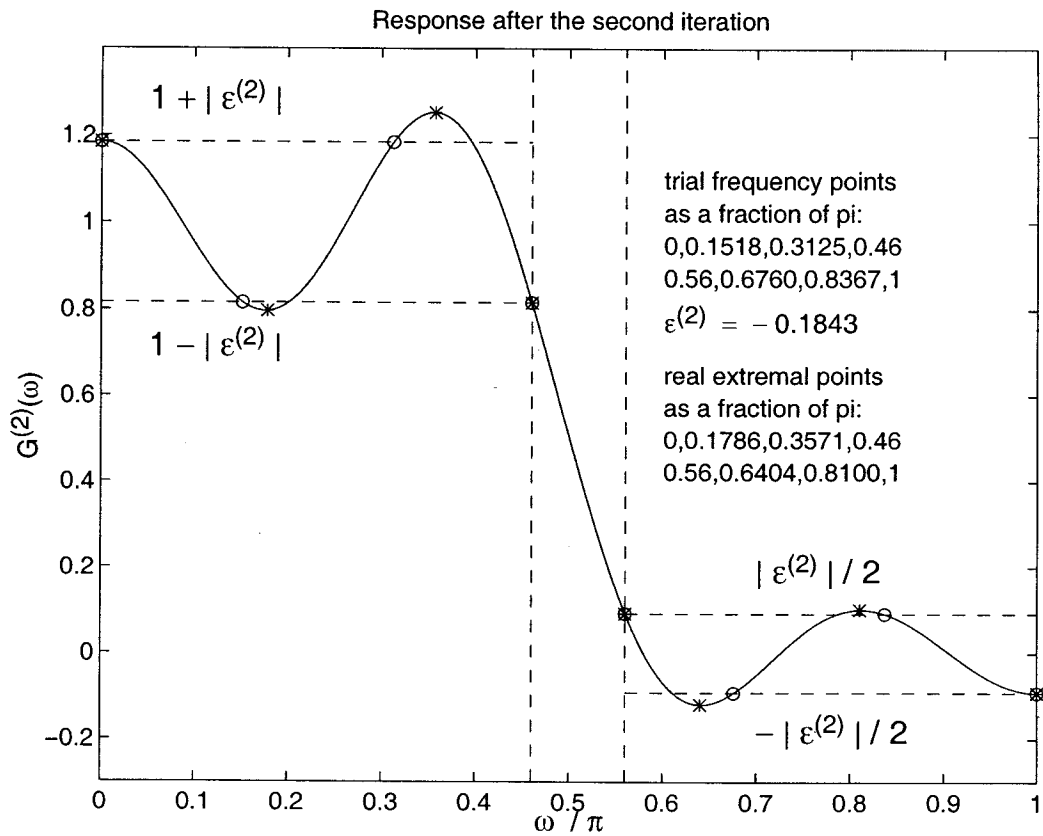
- As seen from the figures, the Remez multiple exchange algorithm works in a simplified form as follows:

- 1) Select $M + 2$ trial extremal points.
- 2) Determine the $M + 1$ unknowns of $G(\omega)$ as well as ϵ such that at all the trial extremal points the absolute value of $E(\omega)$ is $|\epsilon|$ and the sign alternates.
- 3) Find $M + 2$ true extremal points of $E(\omega)$.
- 4) If the trial and true extremal points are the same, the optimum solution has been found. Otherwise, use as new trial extremal points the true extremal points obtained at Step 3 and go to Step 2.

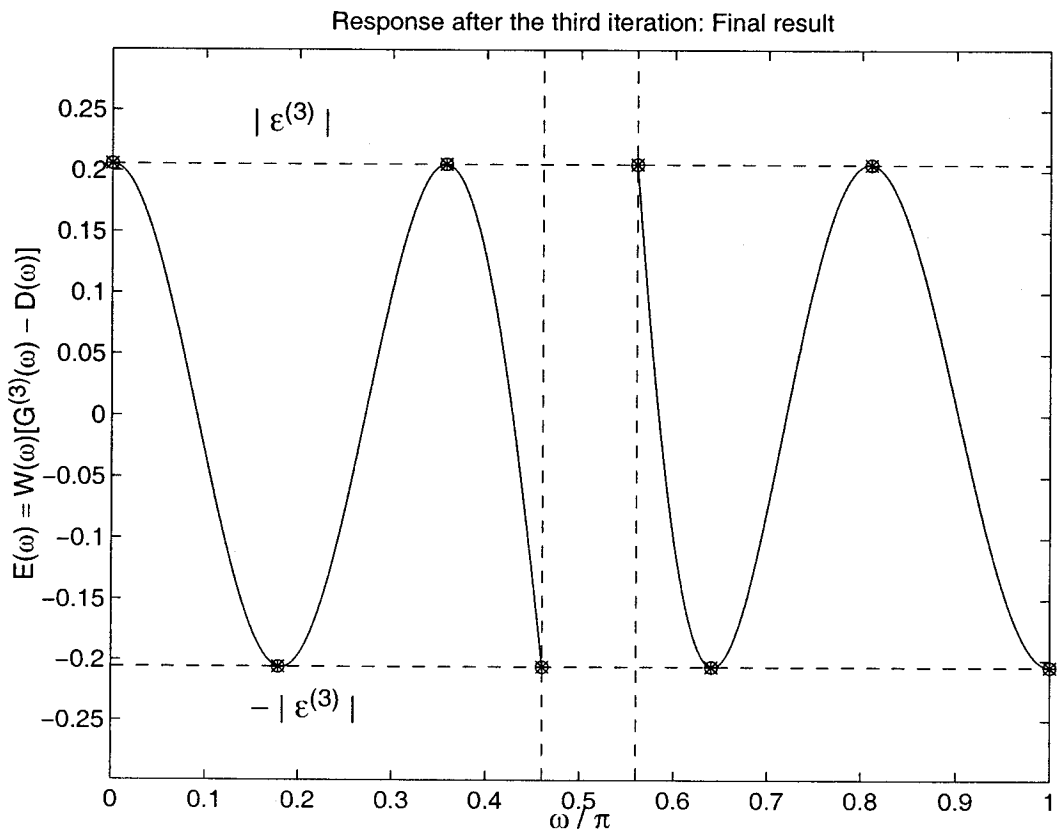
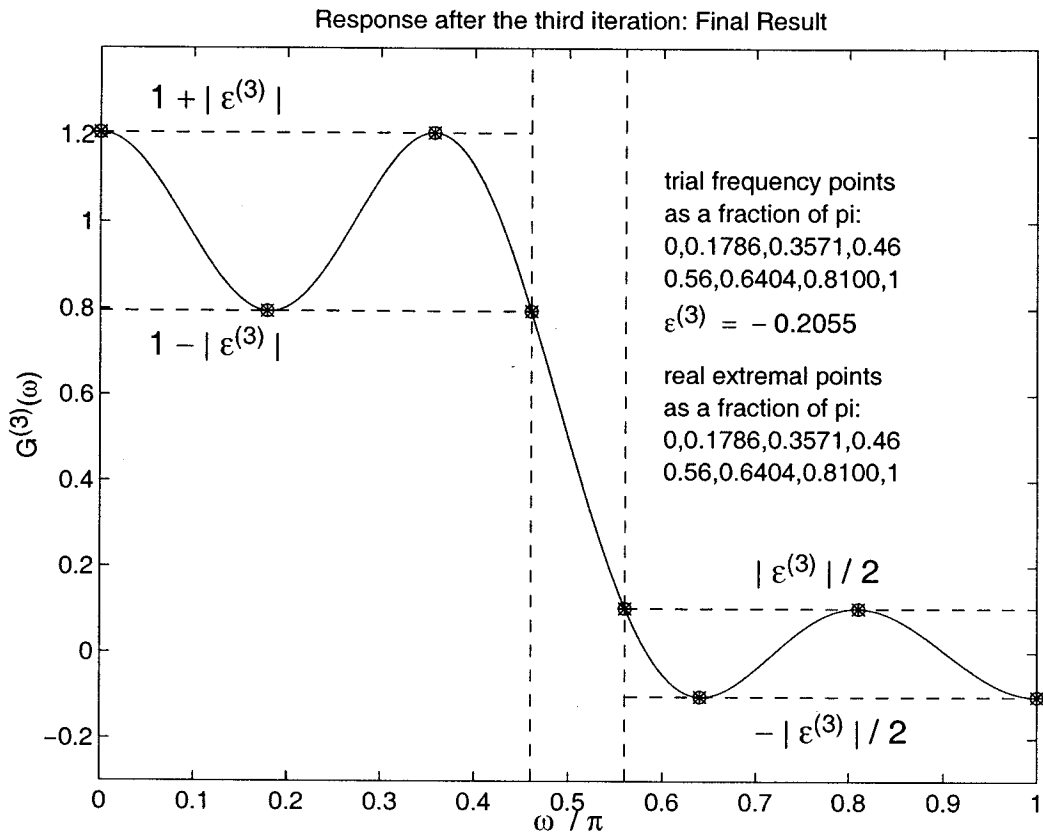
RESPONSES AFTER THE FIRST ITERATION



RESPONSES AFTER THE SECOND ITERATION



RESPONSES AFTER THE THIRD ITERATION: OPTIMUM SOLUTION



Some Details

- The set of linear equations at Step 2 of the Remez algorithm described in transparency 205 can be solved conveniently by first calculating $\epsilon^{(l)}$ analytically as

$$\epsilon = \frac{b_1 \widehat{D}(\omega_1^{(l)}) + b_2 \widehat{D}(\omega_2^{(l)}) + \dots + b_{M+2} \widehat{D}(\omega_{M+2}^{(l)})}{b_1 / \widehat{W}(\omega_1^{(l)}) - b_2 / \widehat{W}(\omega_2^{(l)}) + \dots + (-1)^{M+1} b_{M+2} / \widehat{W}(\omega_{M+2}^{(l)})},$$

where

$$b_k = \prod_{\substack{i=1 \\ i \neq k}}^{M+2} \frac{1}{(\cos \omega_k^{(l)} - \cos \omega_i)}.$$

- After calculating $\epsilon^{(l)}$, $G^{(l)}(\omega)$ achieves the value

$$C_k = \widehat{D}(\omega_k^{(l)}) - (-1)^k \epsilon^{(l)} / \widehat{W}(\omega_k^{(l)}),$$

at the k th extremal point. To get around the numerical sensitivity problems, the Lagrange interpolation formula in the barycentric form is used to express $G^{(l)}(\omega)$ as

$$G^{(l)}(\omega) = \sum_{k=1}^{M+1} \left(\frac{\beta_k}{\cos \omega - \cos \omega_k^{(l)}} \right) C_k / \sum_{k=1}^{M+1} \left(\frac{\beta_k}{\cos \omega - \cos \omega_k^{(l)}} \right),$$

where

$$\beta_k = \prod_{\substack{i=1 \\ i \neq k}}^{M+1} \frac{1}{(\cos \omega_k^{(l)} - \cos \omega_i^{(l)})}.$$

More details

- In the MPR algorithm, $G(\omega)$ is expressed in the above form. This is because the actual coefficient values $a[n]$ are not needed in intermediate calculations.
- After the convergence of the above algorithm, the $a[n]$'s are determined by evaluating $G(\omega)$ at 2^I equally-spaced frequency points ($2^I > 2M$) and then applying the inverse discrete Fourier transformation.
- From the $a[n]$'s the impulse-response coefficients $g[n]$ of $G(z)$ for $n = 0, 1, \dots, 2M$ can be determined as $g[M] = a[0]$, $g[M \pm n] = a[n]/2$ for $n = 1, 2, \dots, M$ (see transparency 49). From these coefficients, the impulse-response coefficients of $H(z)$ can then be obtained according to discussion of transparency 44 in all the linear-phase cases.
- In the practical implementation of the MPR algorithm, the extrema of $E(\omega)$ at Step 3 are located by evaluating $E(\omega)$ over a dense set of frequen-

cies spanning the approximation region X .

- As a rule of thumb, a good selection of the number of grid points is $16M$.
- The matlab routine `remez.m` uses this number of grid points. If a larger number of grid points is desired to be used, `remez.m` can be modified by changing the integer $lgrid = 16$.
- Typically, four to eight iterations of the above algorithm are required to arrive at the optimum solution in the lowpass cases.
- In designing filters having several passband and stopband regions, the number of iterations is typically two or three times that required for designing lowpass filters.

Properties of the Optimum Filters

- In the lowpass case, the filter design parameters are the passband edge ω_p , the stopband edge ω_s , the passband ripple δ_p , and the stopband ripple δ_s .
- The remaining parameter to be determined is the minimum filter order N to meet the given criteria.
- If N is prescribed, then the ripple ratio

$$k = \delta_p / \delta_s,$$

instead of δ_p and δ_s , is usually specified.

- In the latter case, the desired optimum result is obtained by using the following desired response and weighting function

$$D(\omega) = \begin{cases} 1 & \text{for } \omega \in [0, \omega_p] \\ 0 & \text{for } \omega \in [\omega_s, \pi] \end{cases}$$

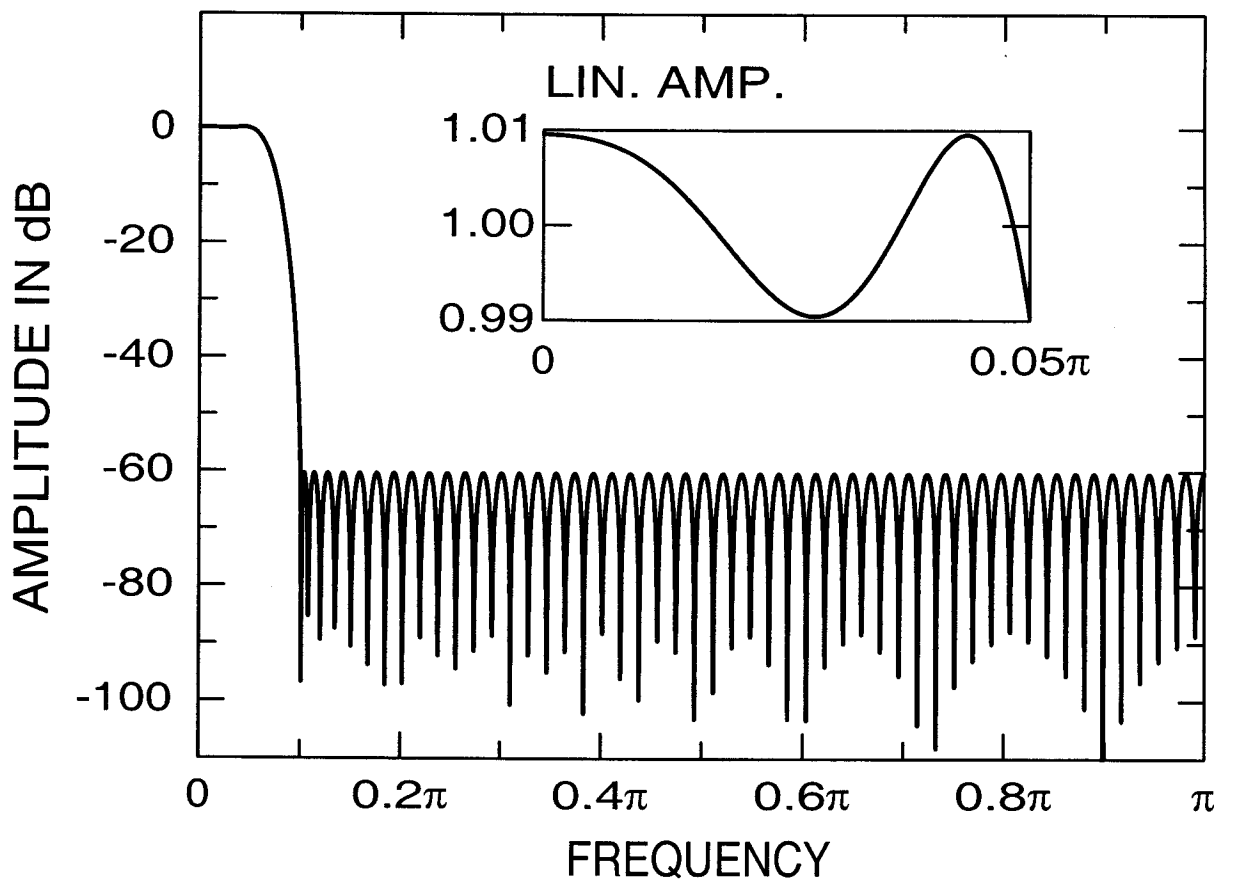
$$W(\omega) = \begin{cases} 1 & \text{for } \omega \in [0, \omega_p] \\ k & \text{for } \omega \in [\omega_s, \pi] \end{cases}$$

in the MPR algorithm.

- In this case, $X = [0, \omega_p] \cup [\omega_s, \pi]$.

Example

- As an example, the following figure gives an optimized response for $N = 108$ ($M = 54$), $\omega_p = 0.05\pi$, $\omega_s = 0.1\pi$, and $k = 10$.
- The resulting ripples are given by $\delta_p = 0.00955$ and $\delta_s = 0.000955$.
- $N = 108$ is the minimum filter order to meet the ripple requirements $\delta_p \leq 0.01$ and $\delta_s \leq 0.001$.



Order Estimation

Except for the case of the Chebyshev solutions there exists no analytic relations between the low-pass filter parameters N , ω_p , ω_s , δ_p , and δ_s .

- Rather accurate estimates, based on empirical data, have been reported.
- The simple formula due to Kaiser is given by

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

- This formula gives rather accurate values when both δ_p and δ_s are relatively small.
- For large values of δ_p and δ_s , the above formula is not so accurate; for very large values, it may even give negative values for N !!

- A somewhat more accurate formula due to Herrmann, Rabiner, and Chan is given by

$$N \approx \frac{D_\infty(\delta_p, \delta_s) - F(\delta_p, \delta_s)[(\omega_s - \omega_p)/(2\pi)]^2}{(\omega_s - \omega_p)/(2\pi)},$$

where

$$D_\infty(\delta_p, \delta_s) = [a_1(\log_{10} \delta_p)^2 + a_2 \log_{10} \delta_p + a_3] \log_{10} \delta_s \\ - [a_4(\log_{10} \delta_p)^2 + a_5 \log_{10} \delta_p + a_6]$$

$$F(\delta_p, \delta_s) = b_1 + b_2[\log_{10} \delta_p - \log_{10} \delta_s]$$

with

$$a_1 = 0.005309, \quad a_2 = 0.07114, \quad a_3 = -0.4761,$$

$$a_4 = 0.00266, \quad a_5 = 0.5941, \quad a_6 = 0.4278$$

$$b_1 = 11.01217, \quad b_2 = 0.51244.$$

- This formula has been developed for $\delta_s < \delta_p$. If $\delta_s > \delta_p$, then the estimate is obtained by interchanging δ_p and δ_s in the formula.
- For this formula, the estimation error is typically less than two percent except for rather narrow-band and wideband filters. The lowpass filter design considered above belongs to these exceptional cases.

- From the above formulas, it is seen that the required filter order is roughly inversely proportional to the transition bandwidth.

Formulas for Highpass, Bandpass, and Bandstop Cases

- For the highpass case, the above formulas apply by interchanging ω_p and ω_s . In order to avoid the zero at $z = -1$ that is in the passband of the highpass filter, use only an even value of N .
- In bandpass and bandstop cases, rather good estimates are obtained by replacing in the previous formulas $\omega_s - \omega_p$ with the narrower transition bandwidth. Like for highpass filters, use only even values of N for bandstop filters.

Example

- In the case of the specifications of the previous example ($\omega_p = 0.05\pi$, $\omega_s = 0.1\pi$, $\delta_p = 0.01$, and $\delta_s = 0.001$) both of the formulas give $N = 101$.
- For the optimized filter designed using the MPR algorithm, the ripples are given by $\delta_p = 0.0157$ and $\delta_s = 0.00157$, showing that the filter order has to be increased.
- When determining the actual minimum filter order, it must be taken into consideration that sometimes a filter of order $N - 1$ has lower ripple values than a filter of order N .
- For instance, for the case $\omega_p = 0.6856\pi$, $\omega_s = 0.83236\pi$, and $k = 1$, the Type I filter of order $N = 10$ achieves $\delta_p = \delta_s = 0.1282$, whereas the Type II filter of order $N = 9$ achieves $\delta_p = \delta_s = 0.1$.
- Based on this, it is advantageous to determine separately the minimum orders for both Type I filters (N is even) and Type II filters (N is odd), and, then, to select the lower order.

- For the above specifications, the minimum orders of Type I and Type II filters to meet the ripple requirements of $\delta_p = 0.01$ and $\delta_s = 0.001$ are 108 and 109, respectively, so that $N = 108$ is the minimum order.

Different Types of Optimum Solutions

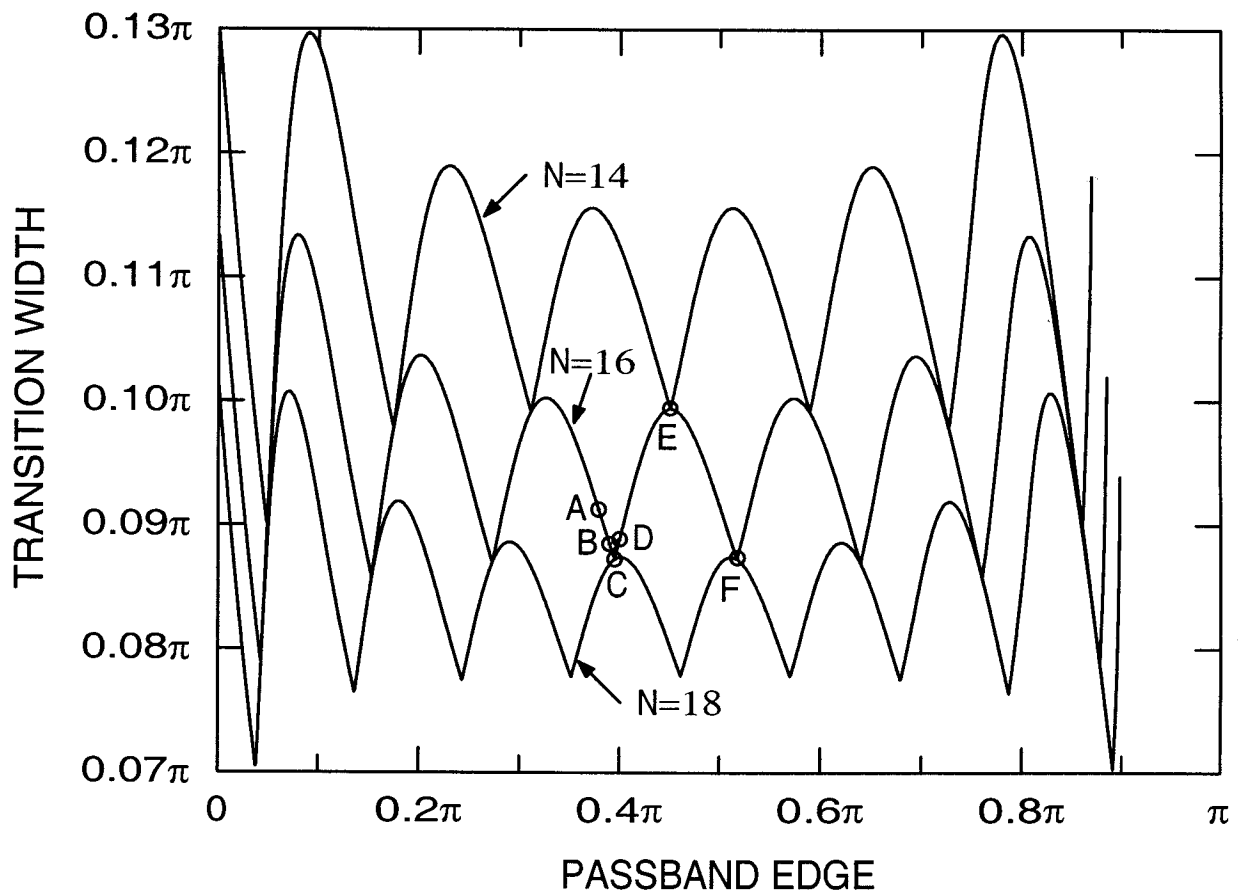
- An informative way to study the various types of optimum lowpass filter solutions is to plot the transition bandwidth

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

of the filter versus ω_p for fixed values of N , δ_p , and δ_s .

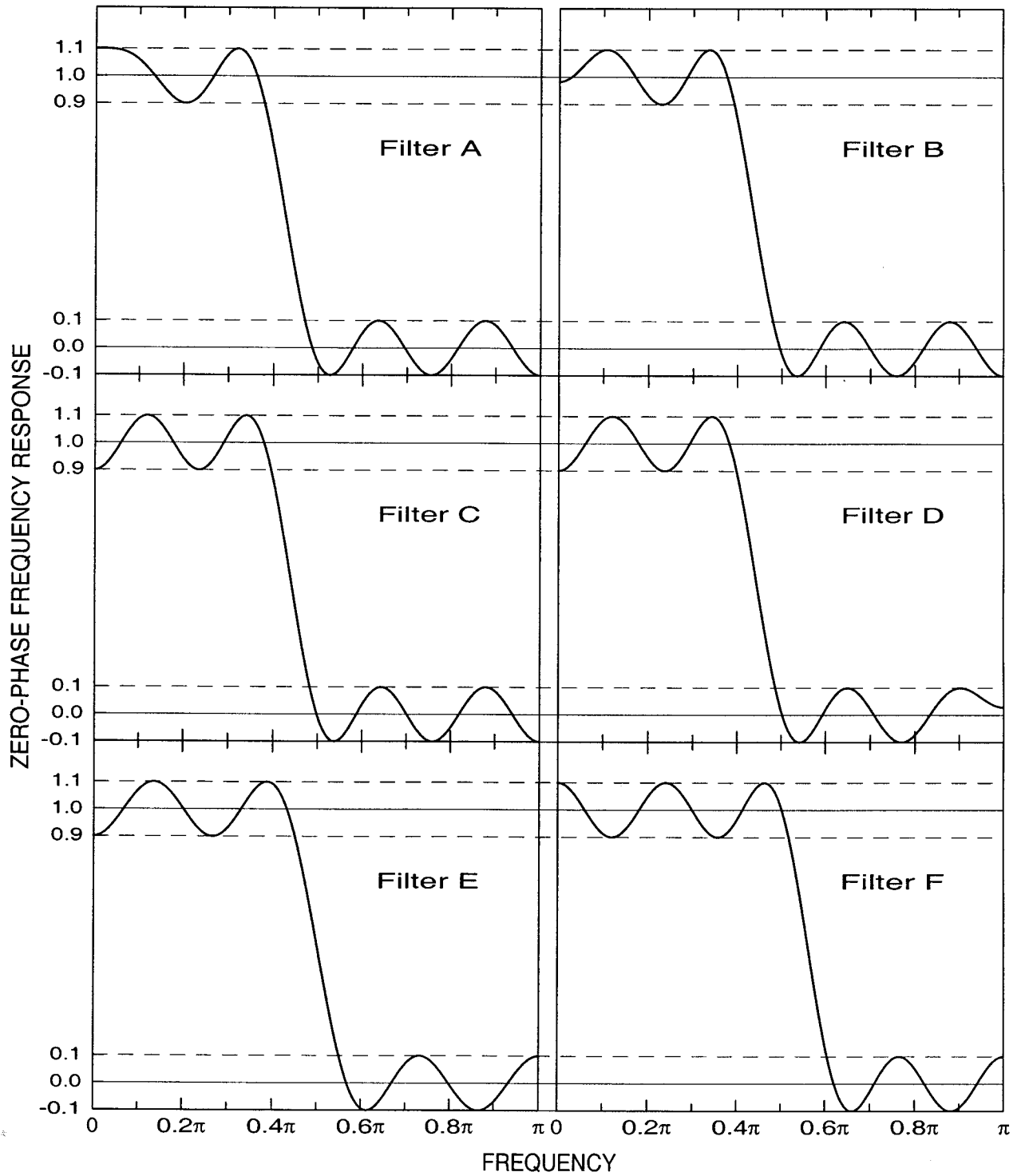
- The next transparency shows such plots for Type I optimum filters with $N = 14$ ($M = 7$), $N = 16$ ($M = 8$), and $N = 18$ ($M = 9$) for $\delta_p = \delta_s = 0.1$.
- As seen from this figure, all the three curves alternate between sharp minima and flat-topped maxima.

Transition Bandwidth ($\omega_s - \omega_p$) as a Function of the Passband Edge ω_p for $\delta_p = \delta_s = 0.1$ for Filters with $N = 14$ ($M = 7$), $N = 16$ ($M = 8$), and $N = 18$ ($M = 9$).



- We consider in greater detail the filters corresponding to the six points, denoted by the letters A, B, C, D, E, and F, in the curve for $N = 16$ ($M = 8$).
- The responses of these filters are given in the figure of the next transparency.
- Filters C and F correspond to the points where the local minimum of $\Delta\omega$ occurs with respect to ω_p .
- These are special **extraripple** or **maximal ripple** solutions whose error function exhibits $M + 3 = 11$ extrema with equal amplitude.
- This is one more than that required by the characterization theorem.
- Furthermore, it follows from this theorem that these extraripple solutions are also optimum solutions for $\widehat{M} = M + 1$, or equivalently for $N = 18$. This is because the number of extrema is $\widehat{M} + 2$ for this filter.

Responses for six of the filters of the figure of transparency 223



- The explanation to this is that the first and last impulse response coefficients $h[n]$ of the filter with higher order become exactly zero when the filter with lower order has the extraripple solution.
- If the transfer functions of the filters with higher and lower orders are denoted by $H(z)$ and $\hat{H}(z)$, then these are related through the equation

$$H(z) = z^{-1}\hat{H}(z),$$

that is the impulse response of $H(z)$ is obtained from that of $\hat{H}(z)$ by shifting the center of symmetry by one sample.

- When ω_p corresponding to the extraripple solution is made smaller, the resulting filter has $M + 2$ equal amplitude extrema, as well as one smaller amplitude extremum at $\omega = 0$ (Filter B).
- When ω_p is further increased, the extra extremum disappears (Filter A).
- On the other hand, if ω_p is made larger, then the resulting filter (Filter D) has one smaller ripple at $\omega = \pi$.

- Also this ripple disappears when ω_p is further increased.
- Filter E in the figure of transparency 225 corresponds to the case where the filter with $N = 14$ has the same solution (extraripple solution for the filter with $N = 14$).
- Hence, for Type I filters there are three kinds of optimum solutions:
 - Solutions having $M + 2$ equal amplitude extrema
 - Special solutions having $M + 3$ equal amplitude extrema
 - Solutions having, in addition to $M + 2$ equal amplitude extrema, one smaller extremum.
- For Type II filters, the properties are quite similar. The basic difference is that the Type II filters have an odd order ($N = 2M + 1$) and they have a fixed zero at $z = -1$ ($\omega = \pi$).

Some Useful Properties of Type I Filters

- Consider a Type I transfer function of the form

$$H(z) = \sum_{n=0}^{2M} h[n]z^{-n}, \quad h[2M - n] = h[n].$$

- The corresponding zero-phase frequency response $H(\omega)$ is given by

$$H(\omega) = h[M] + \sum_{n=1}^M 2h[M - n] \cos n\omega.$$

- On the basis of $H(z)$, the following three transfer functions can be generated:

$$G(z) = \begin{cases} z^{-M} - H(z) & \text{for Case A} \\ (-1)^M H(-z) & \text{for Case B} \\ z^{-M} - (-1)^M H(-z) & \text{for Case C.} \end{cases}$$

- The zero-phase frequency responses of these three filters can be written as

$$G(\omega) = \begin{cases} 1 - H(\omega) & \text{for Case A} \\ H(\pi - \omega) & \text{for Case B} \\ 1 - H(\pi - \omega) & \text{for Case C.} \end{cases}$$

- In Case A, the impulse response coefficients of $G(z)$ are related to the coefficients of $H(z)$ via $g[M] = 1 - h[M]$ and $g[n] = -h[n]$ for $n = 0, 1, \dots, M - 1$ and for $n = M + 1, M + 2, \dots, 2M$.

- By substituting these values into

$$G(\omega) = g[M] + \sum_{n=1}^M 2g[M - n] \cos n\omega$$

we end up with $G(\omega)$ as shown above.

- In Case B, the coefficients $g[n]$ are related to the $h[n]$'s via $g[M - n] = h[M - n]$ for n even and $g[M - n] = -h[M - n]$ for n odd.
- Using the facts that $\cos n\omega = \cos n(\pi - \omega)$ for n even and $-\cos n\omega = \cos(\pi - \omega)$ for n odd, we can write $G(\omega)$ in the above form.
- The fact that $G(\omega)$ is expressible in Case C as shown above follows directly from the properties of the Case A and Case B filters.

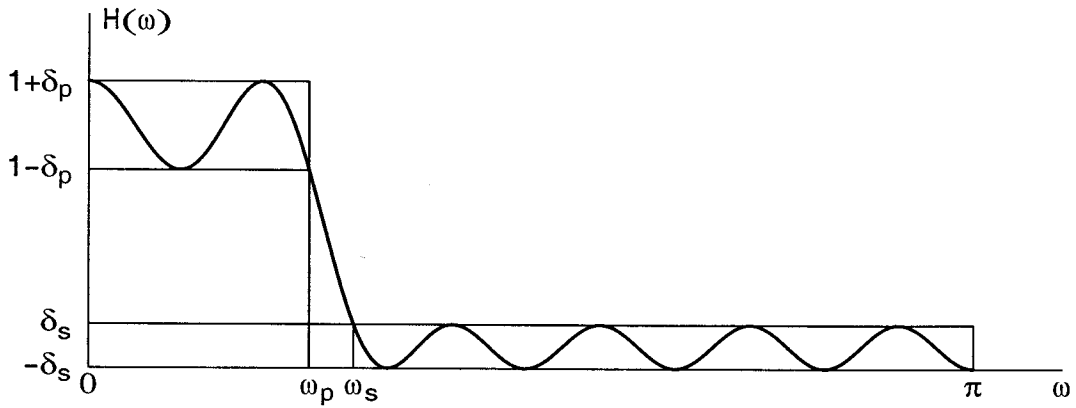
Properties of Case A, Case B, and Case C Transfer Functions

- In Case A, the filter pair $H(z)$ and $G(z)$ is called a **complementary filter pair** since the sum of their zero-phase frequency responses is unity, that is,

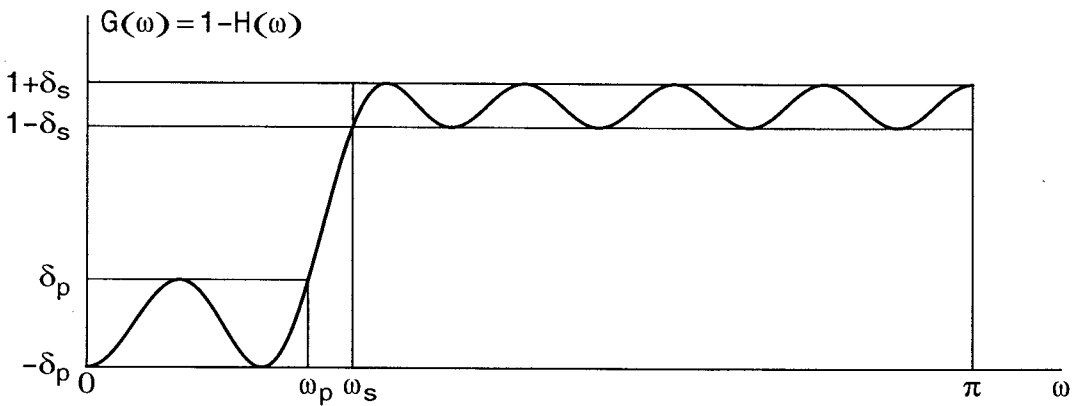
$$H(\omega) + G(\omega) = 1.$$

- This means that if $H(z)$ is a lowpass design with $H(\omega)$ oscillating within $1 \pm \delta_p$ on $[0, \omega_p]$ and within $\pm \delta_s$ on $[\omega_s, \pi]$, then $G(z)$ is a highpass filter with $G(\omega)$ oscillating within $\pm \delta_p$ on $[0, \omega_p]$ and within $1 \pm \delta_s$ on $[\omega_s, \pi]$ (compare figures (a) and (b) in the next transparency).
- An implementation of $G(z)$ is shown in transparency 232. The delay term z^{-M} can be shared with $H(z)$ in this implementation after proper arrangements.
- Hence, at the expense of one additional adder, a complementary filter pair can be implemented.

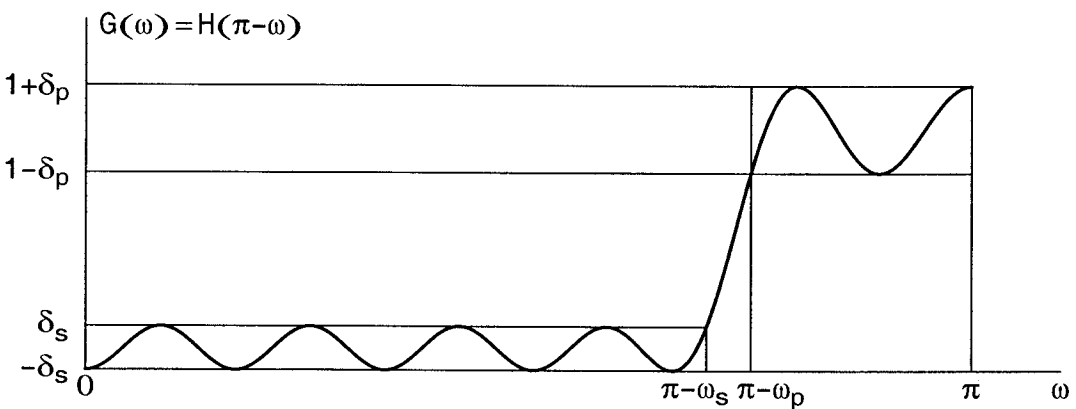
Responses for the Prototype Filter (a), the Case A Filter (b), the Case B Filter (c), and the Case C Filter (d)



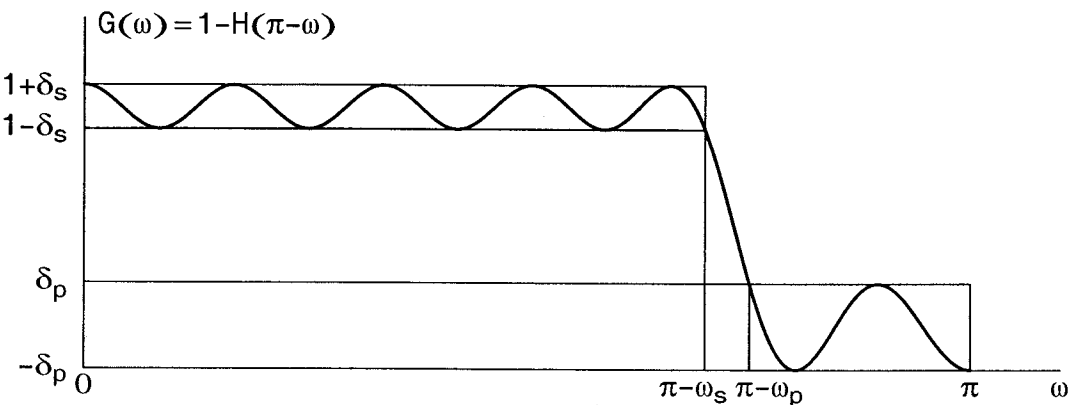
(a)



(b)



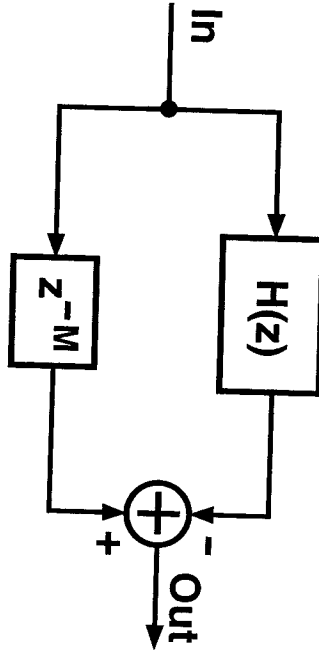
(c)



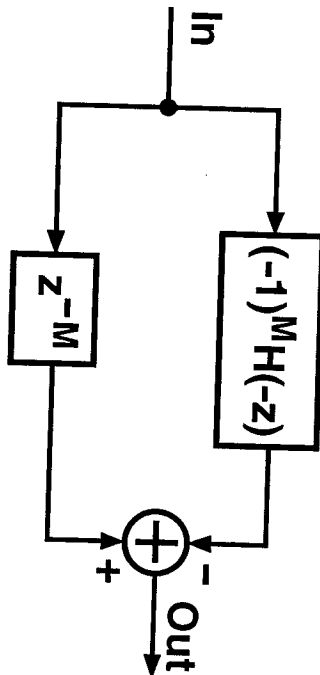
(d)

Implementations for Case A and C Filters

Case A



Case C



- If $H(\omega)$ is as shown in figure (a) in transparency 231, then the Case B filter is a highpass design with $G(\omega)$ oscillating within $\pm\delta_s$ on $[0, \pi - \omega_s]$ and within $1 \pm \delta_p$ on $[\pi - \omega_p, \pi]$ [see figure (c) in transparency 231].
- $G(\omega)$ for the Case C filter, in turn, varies within $1 \pm \delta_s$ on $[0, \pi - \omega_s]$ and within $\pm\delta_p$ on $[\pi - \omega_p, \pi]$ [see figure (d) in transparency 231].
- An implementation of the Case C filter is depicted in transparency 232.
- This implementation is very important in many cases as it allows us to implement a wideband filter $G(z)$ using a delay term and a transfer function which is obtained from a narrowband filter $H(z)$ by simply changing the sign of every second coefficient value.
- This is because there are computationally efficient implementations for narrowband filters, as will be seen Section 10 (not included in the basic course).

The Use of the MPR Algorithm

- There exists a matlab routine `remez.m` that implements the MPR algorithm.
- When designing conventional filters, the user first specifies the filter order N as well as K bands $[\omega_{l,k}, \omega_{u,k}]$ for $k = 1, 2, \dots, K$, where $\omega_{l,k} > \omega_{l,k}$ and $\omega_{l,k+1} > \omega_{u,k}$.
- The above edges are normally given as fractions of π (half the sampling rate).
- The desired function in each band is specified by giving two constants a_k and b_k . The desired function is then a line taking on the values a_k and b_k at the lower and upper edges, respectively. For a normal passband, $a_k = b_k = 1$ and for a normal stopband, $a_k = b_k = 0$.
- The weighting function in each band is specified by given one constants w_k that is then directly the weight in this band.
- If arbitrary desired and weighting functions are desired to be used, then `remez.m` must be modified, as we shall show later.

- On how to use `remez.m` for designing differentiators and Hilbert transformers see transparencies 71–82 considered earlier.

Example 1

- Design a lowpass filter with passband and stopband regions $[0, 0.3\pi]$ and $[0.4\pi, \pi]$, respectively.
- The passband ripple δ_p is restricted to be at most 0.002 on $[0, 0.15\pi]$ and at most 0.01 in the remaining region.
- The stopband ripple δ_s is at most 0.0001 (80 dB attenuation) on $[0.4\pi, 0.6\pi]$ and at most 0.001 (60 dB attenuation) on $[0.6\pi, \pi]$.
- Furthermore, it is desired to implement the overall filter in the form

$$H_{\text{ove}}(z) = H_{\text{fix}}(z)H(z)$$

where the fixed term $H_{\text{fix}}(z)$ has zero pairs on the unit circle at the frequencies $\omega_1 = 0.4\pi$, $\omega_2 = 0.45\pi$, $\omega_3 = 0.5\pi$, $\omega_4 = 0.55\pi$, $\omega_5 = 0.6\pi$, $\omega_6 = 0.65\pi$.

- $H_{\text{fix}}(z)$ can thus be expressed as

$$H_{\text{fix}}(z) = \prod_{k=1}^6 [1 + z^{-2} - 2 \cos(\omega_k)z^{-1}].$$

How to design?

- For the overall filter, the weighted error function is given by

$$E(\omega) = W(\omega)[H_{\text{ove}}(\omega) - D(\omega)],$$

where

$$D(\omega) = \begin{cases} 1, & \omega \in [0, 0.3\pi] \\ 0, & \omega \in [0.4\pi, \pi] \end{cases}$$

and

$$W(\omega) = \begin{cases} 5, & \omega \in [0, 0.15\pi] \\ 1, & \omega \in [0.15\pi, 0.3\pi] \\ 100, & \omega \in [0.4\pi, 0.6\pi] \\ 10, & \omega \in [0.6\pi, \pi]. \end{cases}$$

- The given criteria are met if the peak absolute value of the above error function on $[0, 0.3\pi] \cup [0.4\pi, \pi]$ becomes less than or equal to 0.01. This is the maximum allowable deviation on $[0.15\pi, 0.3\pi]$, where the weighting is unity.
- Using the substitution $H_{\text{ove}}(\omega) = H_{\text{fix}}(\omega)H(\omega)$, the above error function is expressible as

$$E(\omega) = W(\omega)H_{\text{fix}}(\omega)[H(\omega) - D(\omega)/H_{\text{fix}}(\omega)].$$

- Therefore, the desired overall filter can be obtained by designing $H(z)$ using the following de-

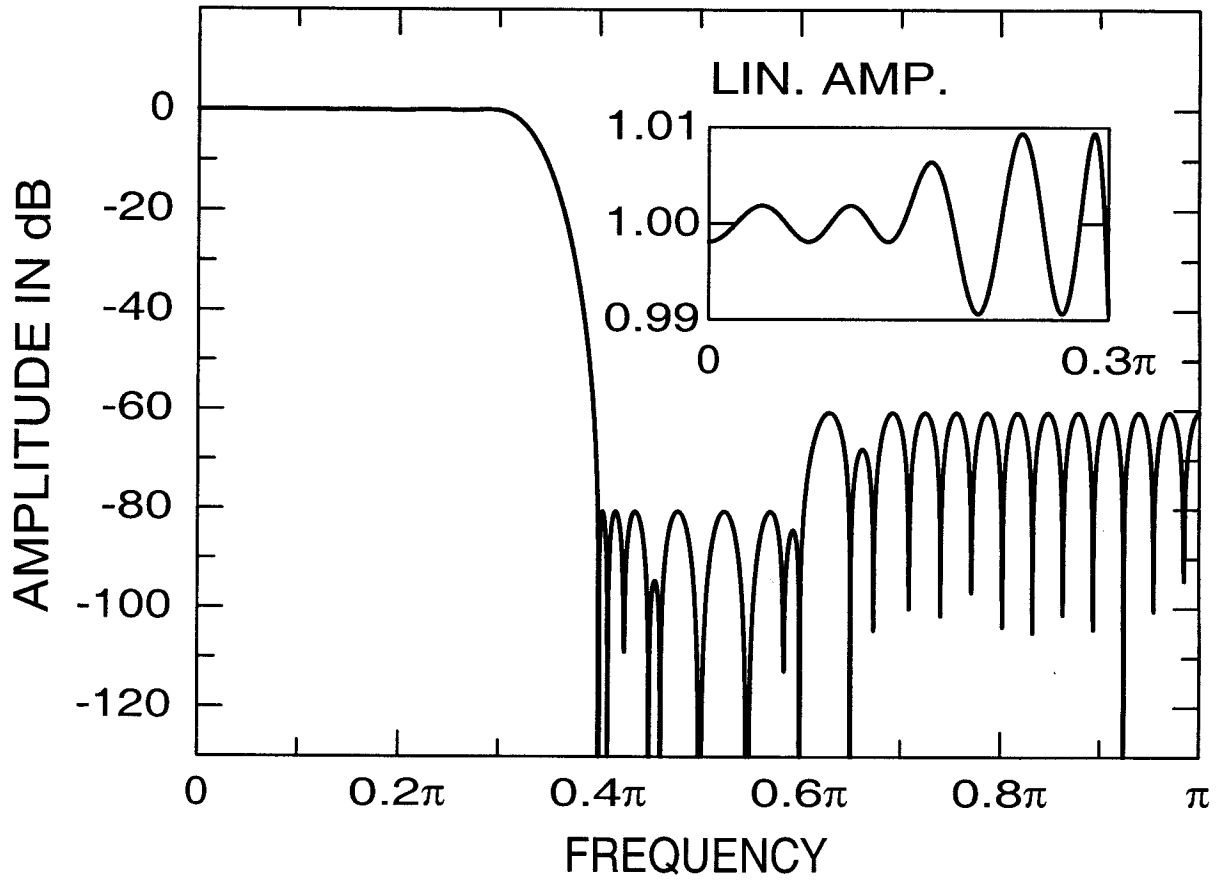
sired and weighting functions:

$$D(\omega) = \begin{cases} 1/|H_{\text{fix}}(\omega)|, & \omega \in [0, 0.3\pi] \\ 0, & \omega \in [0.4\pi, \pi] \end{cases}$$

$$W(\omega) = \begin{cases} 5|H_{\text{fix}}(\omega)|, & \omega \in [0, 0.15\pi] \\ |H_{\text{fix}}(\omega)|, & \omega \in [0.15\pi, 0.3\pi] \\ 100|H_{\text{fix}}(\omega)|, & \omega \in [0.4\pi, 0.6\pi] \\ 10|H_{\text{fix}}(\omega)|, & \omega \in [0.6\pi, \pi]. \end{cases}$$

- In the above, absolute values of $H_{\text{fix}}(\omega)$ are used in order to make the weighting function positive.
- Since the weighting function has to be positive, instead $|H_{\text{fix}}(\omega)|$ a small value must be used at the points where $|H_{\text{fix}}(\omega)|$ becomes zero.
- Appendix E in the end of this chapter shows how to actually perform the design of $H(z)$ by slightly modifying `remez.m`
- The given criteria are met when the peak absolute value of the resulting error function becomes smaller than or equal to 0.01.
- The minimum order of $H(z)$ to meet the criteria is 55.
- The amplitude response of the resulting overall filter is depicted in the next transparency.

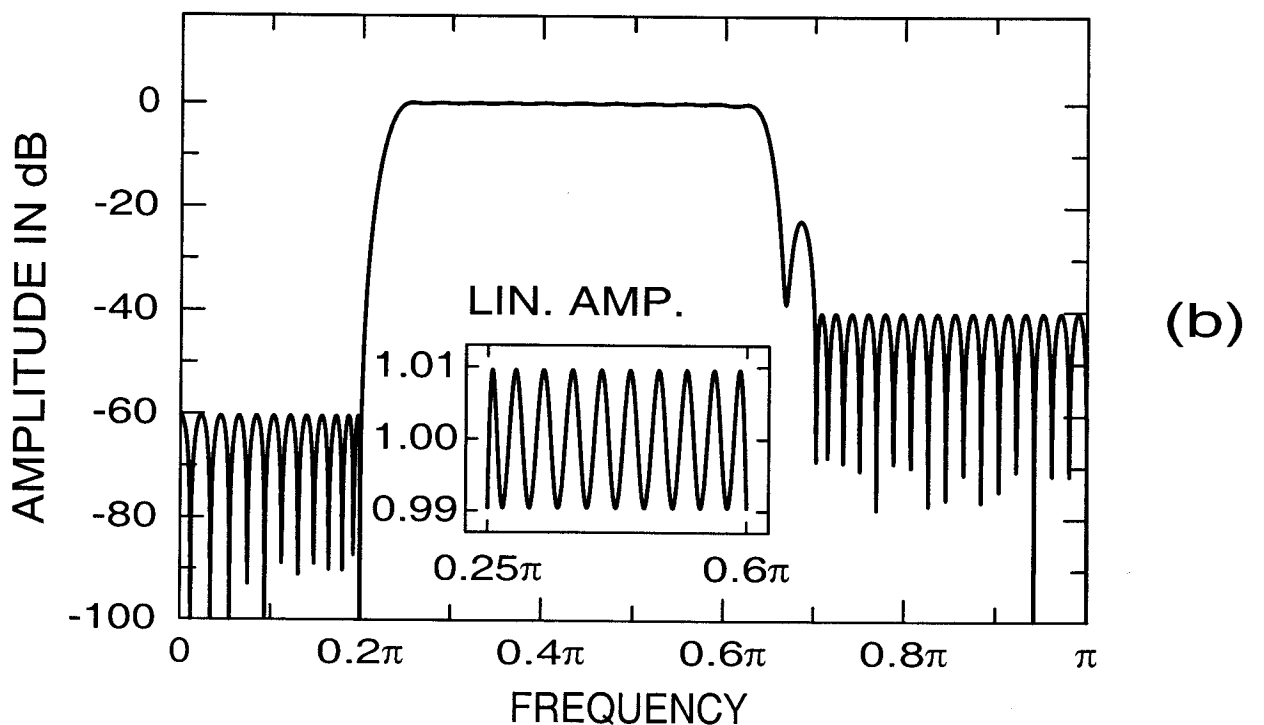
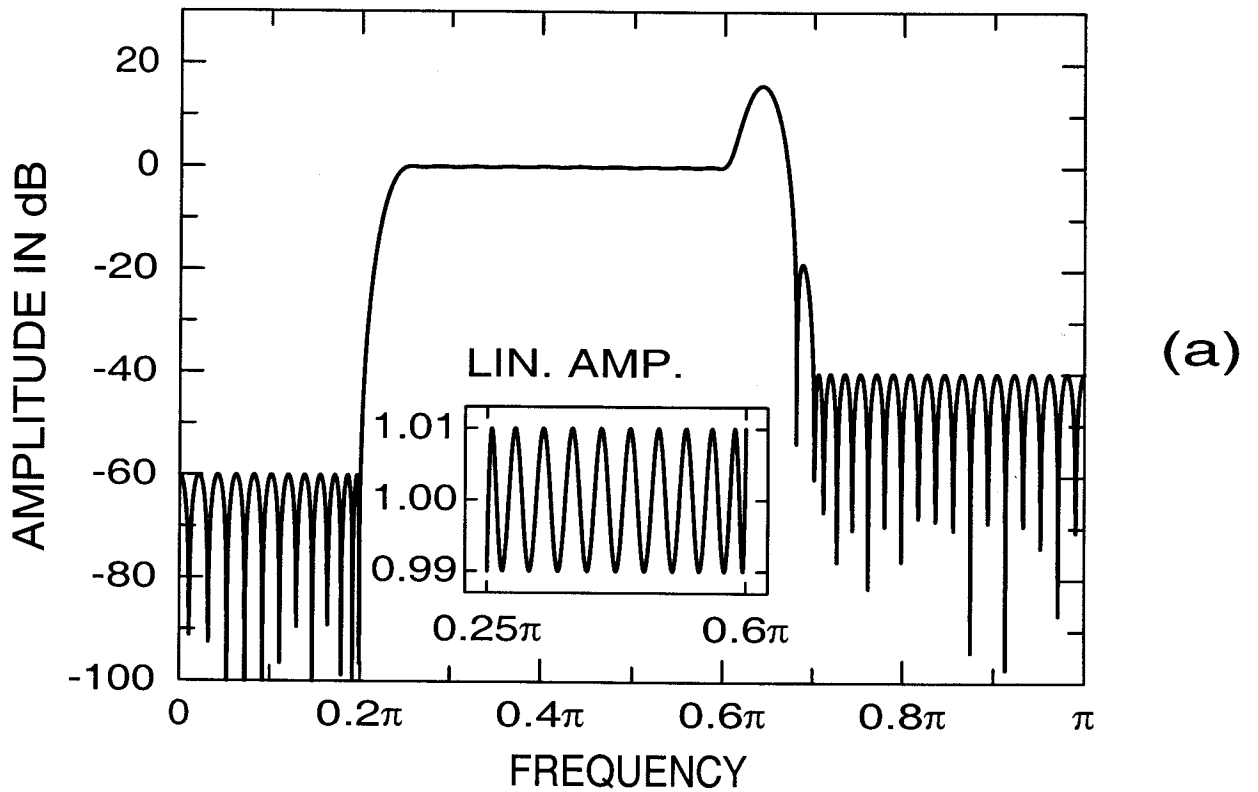
Optimized Filter of Example 1



Example 2

- Let the bandpass filter specifications be: $\omega_{s1} = 0.2\pi$, $\omega_{p1} = 0.25\pi$, $\omega_{p2} = 0.6\pi$, $\omega_{s2} = 0.7\pi$, $\delta_{s1} = 0.001$, $\delta_p = \delta_{s2} = 0.01$.
- The desired response is obtained when the weighting is 10 in the first stopband and 1 in both the passband and the second stopband.
- The minimum order to meet the above specifications is 102.
- Figure (a) in the next transparency shows the resulting response.

Amplitude Responses for Bandpass Filters.
(a) Filter Designed Without Transition Band Constraints. **(b) Filter Designed With Transition Band Constraints.**



Transition band ripples

- This response is optimal according to the characterization theorem even though it has an unacceptable transition band peak of 15 dB.
- This is possible because the approximation is restricted to the passband and stopband regions only and the transition bands are considered as don't-care bands.
- For designs with a single transition band there are no unacceptable transition band ripples.
- However, for filters having more than one transition band, this phenomenon of large transition band peaks occurs when the widths of the transition bands are different; the larger the difference, the greater the problem.

How to avoid the transition band ripple?

- The transition band peak can be easily attenuated by including the transition bands in the overall approximation interval and requiring that the response stays within the limits $-\delta_{s1}$ and $1 + \delta_p$ in the first transition band and within the limits $-\delta_{s2}$ and $1 + \delta_p$ in the second transition band.
- This can be done by selecting the desired function to be $(1/2)(1 + \delta_p - \delta_{s1})$ and $(1/2)(1 + \delta_p - \delta_{s2})$ in the first and second transition bands, respectively.
- If the weighting in the passband is unity, then the weighting functions in the transition bands are selected to be $\delta_p / [(1/2)(1 + \delta_p + \delta_{s1})]$ and $\delta_p / [(1/2)(1 + \delta_p + \delta_{s2})]$, respectively.

Overall Weighted Error Function

- To guarantee the continuity of the desired function, it is worth having very small transition bands between the intervals and to select the approximation interval to be

$$X = X_{s1} \cup X_{t1} \cup X_p \cup X_{t2} \cup X_{s2},$$

where

$$X_{s1} = [0, \omega_{s1}], \quad X_{t1} = [\omega_{s1} + \alpha, \omega_{p1} - \alpha],$$

$$X_p = [\omega_{p1}, \omega_{p2}], \quad X_{t2} = [\omega_{p1} + \alpha, \omega_{s2} - \alpha],$$

and

$$X_p = [\omega_{s2}, \pi].$$

- Here, α is a small positive number.
- The desired function and the weighting functions are, respectively, given by

$$D(\omega) = \begin{cases} 0, & \omega \in X_{s1} \\ (1 + \delta_p - \delta_{s1})/2, & \omega \in X_{t1} \\ 1, & \omega \in X_p \\ (1 + \delta_p - \delta_{s2})/2, & \omega \in X_{t2} \\ 0, & \omega \in X_{s2} \end{cases}$$

and

$$W(\omega) = \begin{cases} \delta_p/\delta_{s1}, & \omega \in X_{s1} \\ 2\delta_p/(1 + \delta_p + \delta_{s1}), & \omega \in X_{t1} \\ 1, & \omega \in X_p \\ 2\delta_p/(1 + \delta_p + \delta_{s2}), & \omega \in X_{t2} \\ \delta_p/\delta_{s2}, & \omega \in X_{s2}. \end{cases}$$

What do we achieve?

- These selections guarantee that if the passband ripple of the resulting filter is less than or equal to the specified δ_p , then the response stays within the desired limits in the transition bands.
- When including the transition bands in the approximation problem, the minimum filter order to meet the resulting specifications has be increased only by one (to 103).
- Figure (b) in transparency 241 shows this response.

Another alternative

- The desired function and the weighting functions can also be selected to be

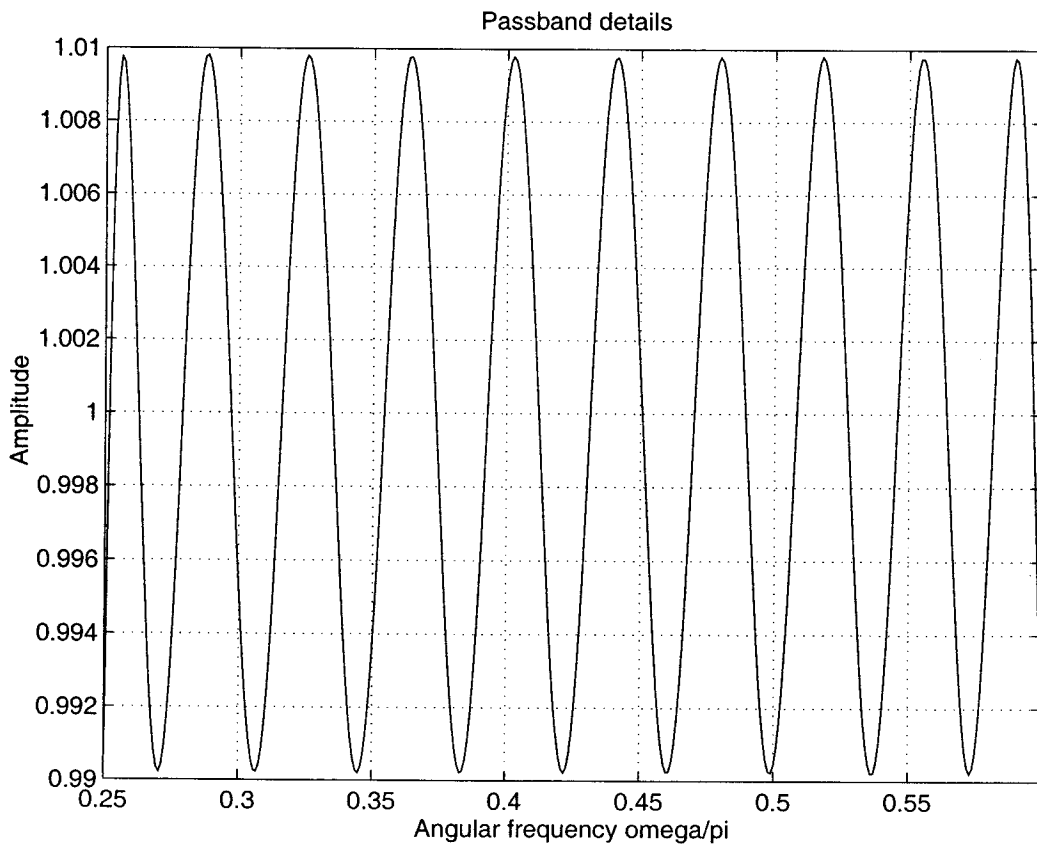
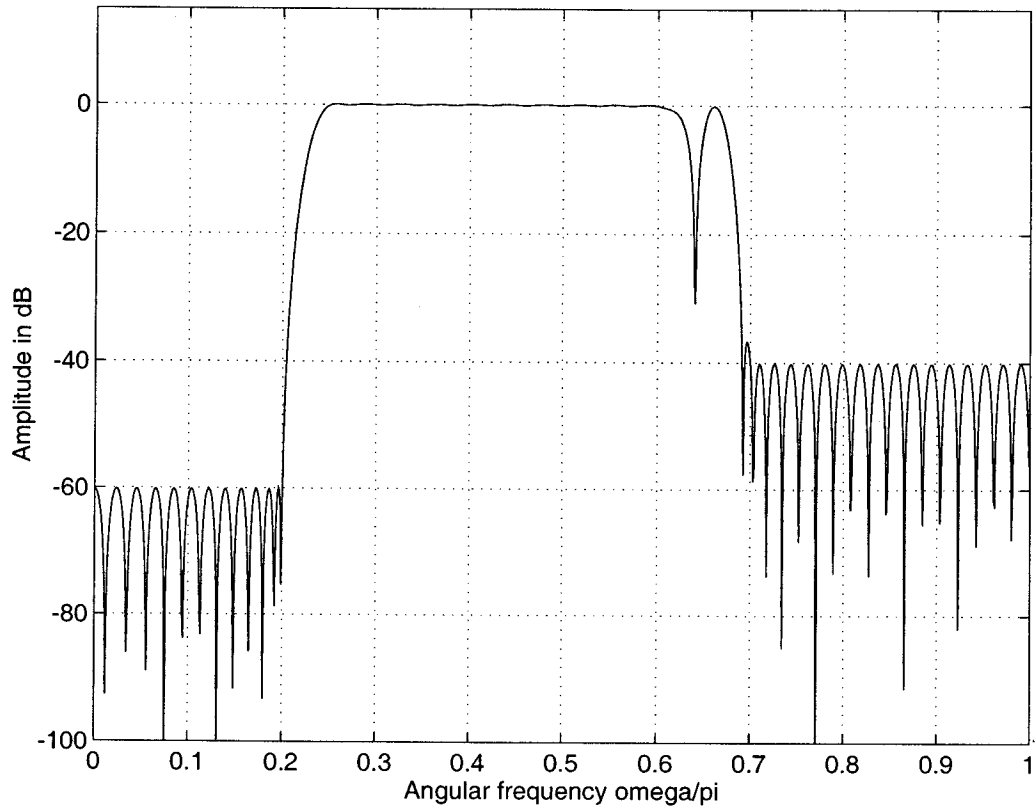
$$D(\omega) = \begin{cases} 0, & \omega \in X_{s1} \\ 0, & \omega \in X_{t1} \\ 1, & \omega \in X_p \\ 0, & \omega \in X_{t2} \\ 0, & \omega \in X_{s2} \end{cases}$$

and

$$W(\omega) = \begin{cases} \delta_p/\delta_{s1}, & \omega \in X_{s1} \\ \delta_p/(1 + \delta_p), & \omega \in X_{t1} \\ 1, & \omega \in X_p \\ \delta_p/(1 + \delta_p), & \omega \in X_{t2} \\ \delta_p/\delta_{s2}, & \omega \in X_{s2}. \end{cases}$$

- In this case, it is guaranteed that if the response meets the criteria in the passband and in the stopbands, then in the transition bands the response is between the limits $\pm(1 + \delta_p)$.
- The amplitude response of the resulting filter is shown in the next transparency.

Resulting Optimized Response



Other Examples

- Transparencies 71 to 82 show how to design differentiators and Hilbert transformers using the Remez algorithm.
- Appendix F in the end of this FIR filter design chapter gives a general-purpose matlab-program, called `firgen.m`, for designing multiband Type I and Type II filters.
- Given the band edges as well as the allowable ripples, it automatically designs a filter with the minimum order.
- If desired, it also takes care of the transition band ripples according to the two above-mentioned techniques.
- It should be pointed out that this program does not always give the optimum solution since the author of this FIR chapter noticed that there is an error in the Remez algorithm written by Parks and McClellan!!!
- People have been using this algorithm for twenty years and there is still an error.

- The next step is to locate this error in the routine.
- If `firgen.m` does not work, please try `firgenni.m`, which works better with a plenty of extra printings, since it is at present a test version.
- In the future, the corrected file will also be called `firgenni.m`.