Filter Design Techniques

- Design of IIR Filters
- Design of FIR Filters
- Optimum Approximations of FIR Filters
- FIR Equiripple Approximation
Procedure of Filter Design

- Define the specifications of filter
- Selection of appropriate technique for filter's coefficient evaluation
- Selection of appropriate structure of filter
- Analysis of finite word-length effect
- Implementation
Define the Specification of Filter

- Signal characteristics (type of signal source and sink, I/O interface, data rate and width, maximum operating frequency)
- Filter characteristics (required amplitude and/or phase response, tolerance, operating speed and mode; real time or batch)
- Working environment
- Other criteria
Selection of Appropriate Technique for Filter's Coefficient Evaluation

- Impulse invariant: IIR
- Bilinear transformation: IIR
- Pole-zero placement: IIR
- Windowing: FIR
- Frequency sampling: FIR
- Optimal: FIR
Selection of Appropriate Structure of Filter

- Transversal (direct): FIR
- Frequency sampling: FIR
- Fast convolution: FIR
- Direct form: IIR
- Cascade: IIR
- Parallel: IIR
- Lattice: IIR or FIR
Analysis of Finite Word-Length Effect

- Input/output signal quantization effect
- Coefficient quantization effect → frequency response distortion in both FIR and IIR filters, and cause unstable effect in IIR filter.
- Arithmetic round-off errors (finite precision arithmetic operation) → required additional bit for representation of result value, and cause unstable effect in IIR filter as coefficient quantization effect.
- Overflow (finite word-length arithmetic operation) → error of result, and also cause unstable effect in IIR filter as coefficient quantization effect.
Implementation

- Hardware or software implementation

- Basic building block requirement
  - Memory unit (i.e. ROM) for storing filter's coefficients.
  - Memory unit (i.e. RAM) for storing input and output values.
  - Hardware or software multipliers.
  - Hardware or software adder and/or another arithmetic logic units.
Specification for effective frequency response of a continuous-time lowpass filter and its corresponding specifications for discrete-time system.

\[ \delta_p \text{ or } \delta_1 \text{ passband ripple} \]
\[ \delta_s \text{ or } \delta_2 \text{ stopband ripple} \]
\[ \Omega_p, \omega_p \text{ passband edge frequency} \]
\[ \Omega_s, \omega_s \text{ stopband edge frequency} \]
\[ \epsilon^2 \text{ passband ripple parameter} \]

\[ 1 - \delta_p = \frac{1}{\sqrt{1 + \epsilon^2}} \]

\[ \text{BW bandwidth} = \omega_u - \omega_l \]
\[ \omega_c \text{ 3-dB cutoff frequency} \]
\[ \omega_u, \omega_l \text{ upper and lower 3-dB cutoff frequencies} \]
\[ \Delta\omega \text{ transition band} = |\omega_p - \omega_s| \]
\[ A_p \text{ passband ripple in dB} \]
\[ = \pm 20\log_{10}(1 \pm \delta_p) \]
\[ A_s \text{ stopband attenuation in dB} \]
\[ = -20\log_{10}(\delta_s) \]
Example: Determining specifications for a discrete-time filter

- If a discrete-time filter will be used as a lowpass filter, we want the overall system have the following specifications when the sampling rate is 10^4 samples/s (T = 10^{-4} s):
  - The gain should be within ±0.01 of unity in the frequency band \(0 \leq \Omega \leq 2\pi(2000)\)
  - The gain should be no greater than 0.001 in the frequency band \(2\pi(3000) \leq \Omega\).
- From specifications above, the parameter should be:
  \(\delta_p = 0.01, \delta_s = 0.001, \Omega_p = 2\pi(2000), \text{ and } \Omega_s = 2\pi(3000)\).
- Because the sampling rate is 10^4 samples/s, thus
  \(\omega_p = \Omega_p T = 2\pi(2000) \times 10^{-4} = 0.4\pi \text{ radians}\)
  \(\omega_s = \Omega_s T = 2\pi(3000) \times 10^{-4} = 0.6\pi \text{ radians}\)
  maximum passband gain in dB = \(20\log_{10}(1 + 0.01) = 0.08643 \text{ dB}\)
  maximum stopband gain in dB = \(20\log_{10}(0.001) = -60 \text{ dB}\)
  stopband attenuation in dB, \(A_s = -\) maximum stopband gain in dB = 60 dB
  passband ripple in dB, \(A_p = -20\log_{10}(1 - 0.01) = 0.087296 \text{ dB} \neq \text{passband gain}\)
Design of Discrete-Time IIR Filters

- From Analog (Continuous-Time) Filters
  - Approximation of Derivatives
  - Impulse Invariance
  - the Bilinear Transformation
  - the Matched-z Transformation
- Frequency Transformation
- Based on Least-Squared Method
  - Pade Approximation Method
  - Least-Squares Design Method
Reasons of Design of Discrete-Time IIR Filters from Continuous-Time Filters

- The art of continuous-time IIR filter design is highly advanced and, since useful results can be achieved, it is advantageous to use the design procedures already developed for continuous-time filters.

- Many useful continuous-time IIR design methods have relatively simple closed-form design formulas. Therefore, discrete-time IIR filter design methods based on such standard continuous-time design formulas are rather simple to carry out.

- The standard approximation methods that work well for continuous-time IIR filters do not lead to simple closed-form design formulas when these methods are applied directly to the discrete-time IIR case.
Characteristics of Commonly Used Analog Filters

- Bessel Filter
- Butterworth Filter
- Chebyshev Filter
  - Chebyshev Type I
  - Chebyshev Type II of Inverse Chebyshev Filter
- Elliptic Filter
Bessel Filter

- Lowpass Bessel filter are characterized by
  \[ H(s) = \frac{1}{a_0 + a_1 s + a_2 s^2 + \ldots + a_{N-1} s^{N-1} + a_N s^N} = \frac{1}{B_N(s)} \]
  where \( B_N(s) \) is called the Bessel polynomial and can be derived via the recursion relation:
  \[ B_N(s) = (2N - 1) B_{N-1}(s) + s^2 B_{N-2}(s), \]
  which \( B_1(s) = s + 1 \), and \( B_2(s) = s^2 + 3s + 3 \).
  And the coefficients \( \{a_k\} \) are:
  \[ a_k = \frac{(2N - k)!}{2^{N-k} k!(N - k)!}, \quad k = 0, 1, \ldots, N. \]
Butterworth Filter

• Lowpass Butterworth filters are all-pole filters characterized by the magnitude-squared frequency response

\[ |H(\Omega)|^2 = 1/[1 + (\Omega/\Omega_c)^{2N}] = 1/[1 + \varepsilon^2(\Omega/\Omega_p)^{2N}] \]

where N is the order of the filter, \( \Omega_c \) is its -3-dB frequency (cutoff frequency), \( \Omega_p \) is the bandpass edge frequency, and \( 1/(1 + \varepsilon^2) \) is the band-edge value of \( |H(\Omega)|^2 \).

• At \( \Omega = \Omega_s \) (where \( \Omega_s \) is the stopband edge frequency) we have

\[ 1/[1 + \varepsilon^2(\Omega_s/\Omega_p)^{2N}] = \delta_2^2 \]

and

\[ N = (1/2)\log_{10}[(1/\delta_2^2) - 1]/\log_{10}(\Omega_s/\Omega_c) = \log_{10}(\delta/\varepsilon)/\log_{10}(\Omega_s/\Omega_p) \]

where \( \delta_2 = 1/\sqrt{1 + \delta_2^2} \).

• Thus the Butterworth filter is completely characterized by the parameters N, \( \delta_2 \), \( \varepsilon \), and the ratio \( \Omega_s/\Omega_p \).
Magnitude and phase of frequency response of lowpass Bessel and Butterworth filters.

Lowpass Bessel filter: \( H(s) = 1/B_N(s) \) where \( B_N(s) = \sum_{k=0}^{N} a_k s^k \) is the Nth-order Bessel polynomial.

Lowpass Butterworth filter:

\[
|H(\Omega)|^2 = 1/[1 + (\Omega/\Omega_c)^{2N}] = 1/[1 + e^{2}(\Omega/\Omega_p)^{2N}]
\]
Frequency response of lowpass Butterworth filters
Example

- Determine the order and the poles of a lowpass Butterworth filter that has a -3-dB bandwidth of 500 Hz and an attenuation of 40 dB at 1000 Hz.
- The critical frequencies are the -3-dB frequency $\Omega_c$ and the stopband frequency $\Omega_s$, which are
  $$\Omega_c = 1000\pi$$
  $$\Omega_s = 2000\pi$$
For an attenuation of 40 dB, $\delta_2 = 0.01$. We obtain
  $$N = \log_{10}(10^4 - 1)/2\log_{10}2 = 6.64$$
To meet the desired specifications, we select $N = 7$. The pole positions are
  $$s_k = 1000\pi e^{j[\pi/2 + (2k + 1)\pi/14]} \ , \ k = 0, 1, 2, \ldots, 6.$$
Chebyshev Filters

• The magnitude squared response of the analog lowpass Type I Chebyshev filter of Nth order is given by:
  \[ |H(\Omega)|^2 = \frac{1}{1 + \varepsilon^2 T_N^2(\Omega/\Omega_p)}. \]
where \( T_N(\Omega) \) is the Chebyshev polynomial of order N:
  \[ T_N(\Omega) = \begin{cases} \cos(N\cos^{-1}(\Omega)), & |\Omega| \leq 1, \\ \cosh(N\cosh^{-1}(\Omega)), & |\Omega| > 1. \end{cases} \]
The polynomial can be derived via a recurrence relation given by
  \[ T_r(\Omega) = 2\Omega T_{r-1}(\Omega) - T_{r-2}(\Omega), \quad r \geq 2, \]
with \( T_0(\Omega) = 1 \) and \( T_1(\Omega) = \Omega \).

• The magnitude squared response of the analog lowpass Type II or inverse Chebyshev filter of Nth order is given by:
  \[ |H(\Omega)|^2 = \frac{1}{1 + \varepsilon^2 \left\{ T_N(\Omega_s/\Omega_p)/ T_N(\Omega_s/\Omega) \right\}^2}. \]
Frequency response of lowpass Type I Chebyshev filter

\[ |H(\Omega)|^2 = \frac{1}{1 + \varepsilon^2 T_N^2(\Omega/\Omega_p)} \]

Frequency response of lowpass Type II Chebyshev filter

\[ |H(\Omega)|^2 = \frac{1}{1 + \varepsilon^2 \{T_N^2(\Omega_s/\Omega_p)/T_N^2(\Omega_s/\Omega)\}} \]
\[
N = \log_{10} \left[ \frac{\sqrt{1 - \delta^2} + \sqrt{1 - \delta^2(1 + \varepsilon^2))}/\varepsilon\delta}{\log_{10} \left[ (\Omega_s/\Omega_p) + \sqrt{(\Omega_s/\Omega_p)^2 - 1} \right]}
= \frac{\cosh^{-1}(\delta/\varepsilon)}{\cosh^{-1}(\Omega_s/\Omega_p)}
\]

for both Type I and II Chebyshev filters, and where

\[
\delta^2 = \frac{1}{\sqrt{1 + \delta^2}}.
\]

• The poles of a Type I Chebyshev filter lie on an ellipse in the s-plane with major axis \( r_1 = \Omega_p \{ (\beta^2 + 1)/2\beta \} \) and minor axis \( r_1 = \Omega_p \{ (\beta^2 - 1)/2\beta \} \) where \( \beta \) is related to \( \varepsilon \) according to

\[
\beta = \left[ \frac{\sqrt{1 + \varepsilon^2 + 1}}{\varepsilon} \right]^{1/N}
\]

• The zeros of a Type II Chebyshev filter are located on the imaginary axis.
Determination of the pole locations for a Chebyshev filter.

Type I: pole positions are
\[ x_k = r_2 \cos \phi_k \]
\[ y_k = r_1 \sin \phi_k \]
\[ \phi_k = \left[ \frac{\pi}{2} \right] + \left[ \frac{(2k + 1)\pi}{2N} \right] \]
\[ r_1 = \Omega_p [\beta^2 + 1]/2\beta \]
\[ r_2 = \Omega_p [\beta^2 - 1]/2\beta \]
\[ \beta = \left[ \frac{\sqrt{1 + \varepsilon^2} + 1}{\varepsilon} \right]^{1/N} \]

Type II: zero positions are
\[ s_k = j\Omega_s/\sin \phi_k \]
and pole positions are
\[ v_k = \Omega_s x_k/\sqrt{x_k^2 + y_k^2} \]
\[ w_k = \Omega_s y_k/\sqrt{x_k^2 + y_k^2} \]
\[ \beta = \left[ \frac{1 + \sqrt{1 - \delta_2^2}}{\delta_2} \right]^{1/N} \]
\[ k = 0, 1, ..., N-1. \]
Example

• Determine the order and the poles of a Type I lowpass Chebyshev filter that has a 1-dB ripple in the passband, a cutoff frequency $\Omega_p = 1000\pi$, a stopband frequency of $2000\pi$, and an attenuation of 40 dB or more for $\Omega \geq \Omega_s$.

• Determine the order of the filter: we have

$\quad 20\log_{10}(1 + \varepsilon^2) = 1 \quad \Rightarrow \varepsilon = 0.5088$.

Also,

$\quad 20\log_{10}\delta_2 = -40 \quad \Rightarrow \delta_2 = 0.01$.

So,

$\quad N = \frac{\log_{10}196.54}{\log_{10}(2 + \sqrt{3})} \approx 4.0$.

• Thus a Type I Chebyshev filter having four poles meets the specifications.
Elliptic Filter or Cauer Filter

- Elliptic filter has an equiripple passband and equiripple stopband magnitude response. Its magnitude-squared frequency response is:
  \[ |H(\omega)|^2 = \frac{1}{1 + \varepsilon^2 U_N(\omega_p/\omega_s)} \]
  where \( U_N(x) \) is the Jacobian elliptic function of order \( N \), which has been tabulated by Zverev [1967], and \( \varepsilon \) is a parameter related to the passband ripple.
- The zeros lie on the \( j\omega \)-axis.
- The filter order is given as:
  \[
  N = K(\omega_p/\omega_s)K(\sqrt{1 - (\varepsilon^2/\delta^2)})/K(\varepsilon/\delta)K(\sqrt{1 - (\omega_p/\omega_s)^2})
  \]
  where \( K(x) \) is the complete elliptic integral of the first kind:
  \[
  K(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x^2 \sin^2 \theta}} \, d\theta
  \]
  and \( \delta^2 = 1/\sqrt{1 + \delta^2} \). The approximate formula is:
  \[
  N \approx 2\log_{10}(4/k_1)/\log_{10}(1/\rho)
  \]
  where \( k_1 = \varepsilon \delta^2/\sqrt{1 - \delta^2} \), \( k' = \sqrt{1 - k^2} \), \( \rho_0 = [1 - \sqrt{k'}]/[2(1 + \sqrt{k'})] \),
  \[
  \rho = \rho_0 + 2(\rho_0)^5 + 15(\rho_0)^9 + 150(\rho_0)^{13}, \text{ and } k = \omega_p/\omega_s.\]
\[ |H(\Omega)|^2 = 1/[1 + \varepsilon^2 U_N(\Omega/\Omega_p)] \]

Magnitude-squared frequency response characteristics of lowpass elliptic filters.
Approximation of Derivative Method

- Approximation of derivative method is the simplest one for converting an analog filter into a digital filter by approximating the differential equation by an equivalent difference equation.
  - For the derivative \( \frac{dy(t)}{dt} \) at time \( t = nT \), we substitute the backward difference \( [y(nT) - y(nT - T)]/T \). Thus

\[
\left. \frac{dy(t)}{dt} \right|_{t=nT} = \frac{y(nT) - y(nT - T)}{T} \approx \frac{y[n] - y[n-1]}{T}
\]

where \( T \) represents the sampling period. Then, \( s = (1 - z^{-1})/T \)

- The second derivative \( \frac{d^2y(t)}{dt^2} \) is derived into second difference as follow:

\[
\left. \frac{dy(t)}{dt} \right|_{t=nT} \approx \frac{y[n] - 2y[n-1] + y[n-2]}{T^2}
\]

which \( s^2 = [(1 - z^{-1})/T]^2 \). So, for the kth derivative of \( y(t) \), \( s^k = [(1 - z^{-1})/T]^k \).
Hence, the system function for the digital IIR filter obtained as a result of the approximation of the derivatives by finite difference is

\[ H(z) = H_a(s) \big|_{s=(z-1)/Tz} \]

It is clear that points in the LHP of the s-plane are mapped into the corresponding points inside the unit circle in the z-plane and points in the RHP of the s-plane are mapped into points outside this circle.

- Consequently, a stable analog filter is transformed into a stable digital filter due to this mapping property.
Example: Approximation of derivative method

Convert the analog bandpass filter with system function

\[ H_a(s) = \frac{1}{[(s + 0.1)^2 + 9]} \]

Into a digital IIR filter by use of the backward difference for the derivative.

Substitute for \( s = (1 - z^{-1})/T \) into \( H_a(s) \) yields

\[ H(z) = \frac{1}{[((1 - z^{-1})/T) + 0.1)^2 + 9]} \]

\[ = \frac{T^2}{1 + 0.27 + 9.017^2} \cdot \frac{z^{-1}}{1 + 0.27 + 9.017^2} \cdot \frac{1}{z^{-2}} \]

\( T \) can be selected to satisfied specified requirements. For example, if \( T = 0.1 \), the poles are

\[ p_{1,2} = 0.91 \pm j0.27 = 0.949 \exp[\pm j16.5^o] \]
Pole-Zero Placement Method

- From filter’s specifications, place poles and zeros on z-plane.
- Calculate $H(z)$ form poles and zeros positions in step 1.
- Evaluate the difference equation from $H(z)$ in step 2.
- Obtain the coefficient from the difference equation in step 3.
Example: Pole-zero placement method

• A bandpass digital filter is required to meet the following specifications:
  - complete signal rejection at DC and 250 Hz
  - a narrow passband centered at 125 Hz
  - a 3-dB bandwidth of 10 Hz

Assuming a sampling frequency of 500 Hz, obtain the transfer function of the filter, by suitable placing z-plane poles and zeros, and its difference equation.

• Determine where to place the poles and zeros on the z-plane.
• Since a complete rejection is required at 0 and 250 Hz, we need to place zeros at corresponding points on the z-plane.
• These are at angle of 0° and 360° × 250/500 = 180° on the unit circle.
• To have the passband centered at 125 Hz requires us to place poles at ±360° × 125/500 = ±90°.
• To ensure that the coefficients are real, it is necessary to have a complex conjugate pole pair.
• The radius, $r$, of the poles is determined by the desired bandwidth. An approximate relationship between $r$, for $r > 0.9$, and bandwidth, $BW$, is given by

$$r \approx 1 - \frac{BW}{F_s}\pi.$$  

• For the problem, $BW = 10$ Hz and $F_s = 500$ Hz, giving $r = 0.937$. The pole-zero diagram is given in figure below. From the pole-zero diagram, the transfer function can be written down by inspection:

$$H(z) = \frac{(z - 1)(z + 1)}{(z - r e^{j\pi/2})(z - r e^{-j\pi/2})} = \frac{(1 - z^{-2})}{(1 + 0.877969z^{-2})}$$

• The difference equation is:

$$y[n] = -0.877969y[n - 2] + x[n] - x[n - 2]$$
Filter Design by Impulse Invariance Method

In the impulse invariance design procedure for transforming continuous-time filters into discrete-time, the impulse response of the discrete-time filter is chosen as equally spaced samples of the impulse response of the continuous-time filter:

\[ h[n] = T_d h_a(nT) \]

where \( T_d \) represents a sampling interval. Then, if

\[ H_a(s) = \sum_{k=1}^{N} \frac{C_k}{s - p_k} \quad \text{and} \quad h_a(t) = \sum_{k=1}^{N} c_k e^{p_k t}, \quad t \geq 0, \]

\[ H(z) = \sum_{k=1}^{N} \frac{C_k}{1 - e^{p_k T} z^{-1}} \]
Impulse Invariant Algorithm

- **Step 1:** define specifications of filter
  - Ripple in frequency bands
  - Critical frequencies: passband edge, stopband edge, and/or cutoff frequencies.
  - Filter band type: lowpass, highpass, bandpass, bandstop.
- **Step 2:** linear transform critical frequencies as follows
  \[ \Omega = \frac{\omega}{T_d} \]
- **Step 3:** select filter structure type and its order: Bessel, Butterworth, Chebyshev type I, Chebyshev type II or inverse Chebyshev, elliptic.
- **Step 4:** convert \( H_a(s) \) to \( H(z) \) using linear transform in step 2.
- **Step 5:** verify the result. If it does not meet requirement, return to step 3.
Example: Impulse invariant method

Convert the analog filter with system function

\[ H_a(s) = \frac{s + 0.1}{(s + 0.1)^2 + 9} \]

into a digital IIR filter by means of the impulse invariance method.

The analog filter has a zero at \( s = -0.1 \) and a pair of complex conjugate poles at \( p_k = -0.1 \pm j3 \). Thus,

\[ H_a(s) = \frac{1}{2} \left( s + 0.1 - j3 \right) + \frac{1}{2} \left( s + 0.1 + j3 \right) \]

Then

\[ H(z) = \frac{\frac{1}{2}}{1 - e^{-0.1T} e^{j3T} z^{-1}} + \frac{\frac{1}{2}}{1 - e^{-0.1T} e^{-j3T} z^{-1}} \]
Frequency response of digital filter.

Frequency response of analog filter.
Example: Impulse invariant with a Butterworth filter.

• By applying impulse invariant method to an appropriate Butterworth analog filter to design a lowpass digital filter with the specifications:
  \[0.89125 \leq |H(e^{j\omega})| \leq 1, \quad 0 \leq |\omega| \leq 0.2\pi,\]
  \[|H(e^{j\omega})| \leq 0.17783, \quad 0.3\pi \leq |\omega| \leq \pi.\]

• Since the parameter \(T\) cancels in the impulse invariant method, we can choose \(T = 1\), so that \(\omega = \Omega\). Hence,
  \[0.89125 \leq |H_a(j\Omega)| \leq 1, \quad 0 \leq |\Omega| \leq 0.2\pi,\]
  \[|H_a(j\Omega)| \leq 0.17783, \quad 0.3\pi \leq |\Omega| \leq \pi.\]

• Because the magnitude response of an analog Butterworth filter is monotonic function of frequency, the above specifications will be satisfied if
  \[|H_a(j0.2\pi)| \geq 0.89125 \quad \text{and} \quad |H_a(j0.3\pi)| \leq 0.17783.\]
  Then, \(1 + (0.2\pi/\Omega_c)^{2N} = (1/0.89125)^2\) and \(1 + (0.3\pi/\Omega_c)^{2N} = (1/0.17783)^2\).
  So, \(N = 5.8858\) and \(\Omega_c = 0.70474\).

• However, \(N\) must be rounded up to the next highest integer:
  \(N = 6\) and new value of \(\Omega_c = 0.7032\).

• It means that there are 12 poles of the magnitude-squared function:
\[ H_a(s)H_a(-s) = 1/[1 + (s/j \Omega_c)^{2N}] \]  
These poles are uniformly distributed in angle on a circle of radius \( \Omega_c = 0.7032 \).

- The poles of \( H_a(s) \) are the three pole pairs in the left half of the s-plane with the following coordinates:
  - Pole pair #1: \(-0.182 \ j0.679,\)
  - Pole pair #2: \(-0.497 \ j0.497,\)
  - Pole pair #3: \(-0.679 \ j0.182.\)

Therefore,

\[ H_a(s) = 0.12093/[(s^2 + 0.3640s + 0.4945)(s^2 + 0.9945s + 0.4945)(s^2 + 1.3585s + 0.4945)] \]

- Express \( H_a(s) \) as a partial fraction expansion, then perform the transformation, and combine complex-conjugate terms, the resulting system function of the digital filter is:

\[ H(z) = [(0.2871 - 0.4466z^{-1})/(1 - 1.2971z^{-1} + 0.6949z^{-2})] + [(-2.1428 + 1.1455z^{-1})/(1 - 1.0691z^{-1} + 0.3699z^{-2})] + [(1.8557 - 0.6303z^{-1})/(1 - 0.9972z^{-1} + 0.2570z^{-2})]. \]
s-plane locations for poles of $H_c(s)H_c(-s)$ and its frequency response for sixth-order Butterworth filter
Disadvantage of previous techniques: frequency warping $\rightarrow$ aliasing effect and error in specifications of designed filter (frequencies).
So, prewarping of frequency is considered.
Bilinear Transformation Method

- This method is a conformal mapping that transform the $j\Omega$-axis into the unit circle in the $z$-plane only once, thus avoiding aliasing of frequency components.
- All points in the LHP of $s$ are mapped inside the unit circle in the $z$-plane and all points in the RHP of $s$ are mapped into corresponding points outside the unit circle in the $z$-plane.
- The bilinear transformation is derived by applying the trapezoidal numerical integration approach to the differential equation representation of $H_a(s)$ that leads to the difference equation representation of $H(z)$. 
Mapping of the s-plane onto the z-plane using bilinear transformation
Mapping of the continuous-time frequency axis onto the discrete-time frequency axis by bilinear transformation.
Frequency warping inherent in the bilinear transformation of a continuous-time lowpass filter into a discrete-time lowpass filter. To achieve the desired discrete-time cutoff frequencies, the continuous-time cutoff frequencies must be prewarped.
Illustration of the effect of the bilinear transformation on a linear phase characteristics. (Dashed line is linear phase and solid line is phase resulting from bilinear transformation.)
Bilinear Transformation Algorithm

- **Step 1**: define specifications of filter
  - Ripple in frequency bands
  - Critical frequencies: passband edge, stopband edge, and/or cutoff frequencies.
  - Filter band type: lowpass, highpass, bandpass, bandstop.
- **Step 2**: select filter structure type and its order: Bessel, Butterworth, Chebyshev type I, Chebyshev type II or inverse Chebyshev, elliptic.
- **Step 3**: critical frequencies prewarping:
  \[ \Omega = \left[\frac{2}{T_d}\right] \tan(\omega/2) \]
- **Step 4**: bilinear transform by replace s variable in H(s) by
  \[ s = \left[\frac{2}{T_d}\right] \frac{(1-z^{-1})}{(1+z^{-1})} \]
  to obtain H(z).
- **Step 5**: verify the result. If it does not meet requirement, return to step 2.
## Poles-Zeros Mapping from Continuous-Time Lowpass Filter to Intermediate Continuous-Time Filter

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Mapping Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowpass</td>
<td>Lowpass</td>
<td>$s' = s / \omega_p'$</td>
</tr>
<tr>
<td>Lowpass</td>
<td>Highpass</td>
<td>$s' = \omega_p' / s$</td>
</tr>
<tr>
<td>Lowpass</td>
<td>Bandpass</td>
<td>$s' = (s^2 + \omega_o^2) / B$s</td>
</tr>
<tr>
<td>Lowpass</td>
<td>Bandstop</td>
<td>$s' = B s / (s^2 + \omega_o^2)$</td>
</tr>
</tbody>
</table>

Where $\omega_o^2 = \omega_1' \omega_2'$, $B =$ bandwidth $= \omega_2' - \omega_1'$, and $\omega_1'$ and $\omega_2'$ are desired lower and upper cutoff frequencies, respectively.
Example: Design of digital filter by selecting $T_d$

- Convert the analog filter with system function:
  $$H_a(s) = \frac{s + 0.1}{(s + 0.1)^2 + 16}$$
  into a digital IIR filter by means of the bilinear transformation. The digital filter is to have a resonant frequency of $\omega_r = \pi/2$.

- First, it is clear that the analog filter has a resonant frequency $\Omega_r = 4$. This frequency is to be mapped into $\omega_r = \pi/2$ by selecting the value of the parameter $T_d$. From the relationship $\omega = 2\tan^{-1}[\Omega T_d/2]$, we must select $T_d = \frac{1}{2}$ in order to have $\omega_r = \pi/2$. Thus the desired mapping is
  $$s = 4[1 - z^{-1}]/[1 + z^{-1}]$$
  The resulting digital filter has the system function
  $$H(z) = \frac{0.128 + 0.006z^{-1} - 0.122z^{-1}}{1 + 0.0006z^{-1} + 0.975z^{-2}}$$
  $$\approx \frac{0.128 + 0.006z^{-1} - 0.122z^{-1}}{1 + 0.975z^{-2}}$$
  The filter has poles at $p_{1,2} = 0.987\exp[\pm j\pi/2]$ and zeros at $z_{1,2} = -1, 0.95$. 
Example: Design of digital filter begins with specifications

- Design a single-pole lowpass digital filter with a 3-dB bandwidth of $0.2\pi$, using the bilinear transformation applied to the analog filter:
  \[ H_a(s) = \frac{\Omega_c}{s + \Omega_c} \]
  where $\Omega_c$ is the 3-dB bandwidth of the analog filter.
- The digital filter is specified to have its -3 dB gain at $\omega_c = 0.2\pi$. In the frequency domain of the analog filter $\omega_c = 0.2\pi$ corresponds to:
  \[ \Omega_c = \frac{2}{T_d}\tan(0.1\pi) = \frac{0.65}{T_d} \]
  Thus the analog filter has the system function:
  \[ H_a(s) = \frac{0.65/T_d}{s + 0.65/T_d} \]
  Now, we apply the bilinear transformation $s = \frac{2}{T_d}\frac{1 - z^{-1}}{1 + z^{-1}}$ to convert the analog filter into the desired digital filter. Thus we obtain:
  \[ H(z) = 0.245\frac{1 + z^{-1}}{1 + 0.509z^{-1}} \]
  where the parameter $T_d$ has been divided out. The frequency response of the digital filter is $H(\omega) = 0.245[1 + e^{-j\omega}]/[1 + 0.509e^{-j\omega}]$.

\[ \square \]
- At $\omega = 0$, $H(0) = 1$, and at $\omega = 0.2\pi$, we have $|H(0.2\pi)| = 0.707$, which is the desired response.
Example: of Bilinear Transformation Design

- Design a second-order digital notch filter having a notch frequency at 60 Hz and a 3-dB notch bandwidth of 6 Hz. Tha sampling frequency employed is 400 Hz.
- The normalized angular notch frequency $\omega_0$ and the normalized angular 3-dB bandwidth BW are
  \[
  \omega_0 = 2\pi\left(\frac{60}{400}\right) = 0.3\pi, \quad \text{BW} = 2\pi\left(\frac{6}{400}\right) = 0.03\pi.
  \]
  Selecting a second-order analog notch filter with a transfer function as:
  \[
  H_a(s) = \frac{s^2 + \Omega^2}{s^2 + Bs + \Omega_0^2}.
  \]
  Using $T_d = 2$, then $s = (1-z^{-1})/(1+z^{-1})$ and
  \[
  H(z) = \frac{[(1+\Omega_0^2) - 2(1-\Omega_0^2) z^{-1} + (1+\Omega_0^2) z^{-2}]/[(1+\Omega_0^2 + B) - 2(1-\Omega_0^2) z^{-1} + (1+\Omega_0^2 - B) z^{-2}]
  \]
  where $B$ is equal to the 3-dB notch bandwidth $= \Omega_2 - \Omega_1$.
  Thus,
  \[
  H(z) = \frac{0.940809 - 1.105987z^{-1} + 0.940809z^{-2}}{1 - 1.105987z^{-1} + 0.881618z^{-2}}
  \]
Example: Bilinear transformation of a Butterworth filter

• By applying the bilinear transformation method to an appropriate Butterworth analog filter to design a lowpass digital filter with the specifications:
  
  \[
  0.89125 \leq |H(e^{j\omega})| \leq 1, \quad 0 \leq \omega \leq 0.2\pi, \\
  |H(e^{j\omega})| \leq 0.17783, \quad 0.3\pi \leq \omega \leq \pi.
  \]

• The critical frequencies of the digital filter must be prewarped to the corresponding analog frequencies using the relationship:
  
  \[\Omega = \left[\frac{2}{T_d}\right] \tan\left(\frac{\omega}{2}\right)\]

  given \(T_d = 1\):
  
  \[
  0.89125 \leq |H(e^{j\omega})| \leq 1, \quad 0 \leq \Omega \leq 2\tan(0.2\pi/2), \\
  |H(e^{j\omega})| \leq 0.17783, \quad 2\tan(0.3\pi/2) \leq \Omega \leq \infty.
  \]

• Since an analog Butterworth filter has a monotonic magnitude response, so
  
  \[
  |H_a(j2\tan(0.1\pi))| \geq 0.89125 \text{ and } |H_a(j2\tan(0.1\pi))| \leq 0.17783.
  \]

• Because the form of the magnitude-squared function for the Butterworth filter is:
  
  \[
  |H_a(\Omega)|^2 = 1/[1 + (\Omega/\Omega_c)^{2N}] .
  \]
Solving for $N$ and $\Omega_c$, we obtain

$$1 + \left( \frac{2\tan(0.1\pi)}{\Omega_c} \right)^2 N = \left( \frac{1}{0.89125} \right)^2$$

and

$$1 + \left( \frac{2\tan(0.15\pi)}{\Omega_c} \right)^2 N = \left( \frac{1}{0.17783} \right)^2 .$$

$N = 5.305$

Since $N$ must be an integer, we choose $N = 6$. Thus, we obtain $\Omega_c = 0.766$.

The obtained value of $\Omega_c$ exceeds the passband specifications (0.649839 ~ 1.019051) and the stopband specifications are met exactly.

In the s-plane, the 12 poles of the magnitude-squared function are uniformly distributed in angle on a circle of radius 0.766. The system function of the analog filter by selecting the LHP poles is:

$$H_a(s) = \frac{0.20238}{[(s^2 + 0.3996s + 0.5871)(s^2 + 1.0836s + 0.5871)(s^2 + 1.4802s + 0.5871)]}$$

The system function for the digital filter is obtained by applying the bilinear transformation with $T_d = 1$. The result is:

$$H(z) = \frac{0.0007378(1+z^{-1})^6}{[(1-1.2686z^{-1}+0.7051z^{-2})(1-1.0106z^{-1}+0.3583z^{-2})(1-0.9044z^{-1}+0.2155z^{-2})]}$$
s-plane locations for poles of $H_c(s)H_c(-s)$ and frequency response of sixth-order Butterworth lowpass filter transformed by bilinear transformation.
Example: Approximation methods for frequency-selective IIR analog filters → Butterworth, Chebyshev, and elliptic function approximation methods.

- Consider the lowpass digital filter specifications:
  \[0.99 \leq |H(e^{j\omega})| \leq 1.01, \quad |\omega| \leq 0.4\pi,\]
  \[|H(e^{j\omega})| \leq 0.001, \quad 0.6\pi \leq |\omega| .\]
Example: Butterworth approximation

Pole-zero plot and frequency response of 14th-order Butterworth lowpass filter.
Example: Chebyshev approximation

Pole-zero plot and frequency response of 8th-order Chebyshev type I lowpass filter.

Log Magnitude

Detailed plot of magnitude in passband

Group delay
Pole-zero plot and frequency response of 8th-order Chebyshev type II lowpass filter.

Log Magnitude

Detailed plot of magnitude in passband

Group delay
Example: Elliptic approximation

Pole-zero plot and frequency response of 6th-order elliptic lowpass filter.

Log Magnitude

Detailed plot of magnitude in passband

Group delay
Matched-z Transformation Method

- Converting an analog filter into an equivalent digital filter by mapping the poles and zeros of $H_a(s)$ directly into poles and zeros in the $z$-plane.

$$H_a(s) = \frac{\prod_{k=1}^{M} (s-Z_k)}{\prod_{k=1}^{N} (s-P_k)}$$

where $\{Z_k\}$ are the zeros and $\{P_k\}$ are the poles of $H_a(s)$, and $T$ is the sampling interval.

- Thus each factor of the form $(s - a)$ in $H_a(s)$ is mapped into the factor $(1 - e^{aT}z^{-1})$ in $H(z)$.
• The poles obtained from the matched-z transformation are identical to the poles obtained with the impulse invariant method. However, the two techniques result in different zero positions.

• To preserve the frequency response characteristics of the analog filter, the sampling interval in the matched-z transformation must be properly selected to yield the pole and zero locations at equivalent position in the z-plane.

• Thus aliasing must be avoid by selecting T sufficiently small.
Frequency Transformation Method

- A lowpass prototype filter will be selected and a frequency transformation will be performed to obtain a lowpass or a highpass or a bandpass or a bandstop filter.

- There are two alternative approaches:
  - To perform the frequency transformation in the analog domain and then to convert the analog filter into a corresponding digital filter by a mapping of the s-plane into the z-plane.
  - To convert the analog lowpass filter into a lowpass digital filter and then to transform the lowpass digital filter into the desired filter by a digital transformation.

- In general, these two approaches yield different results, except for the bilinear transformation, in which case the resulting filter designs are identical.
Frequency transformations for analog filters (prototype lowpass filter has band edge frequency $\Omega_p$)

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Transformation</th>
<th>Band edge frequency of new filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowpass</td>
<td>Lowpass</td>
<td>$s \rightarrow [\frac{\Omega_p}{\Omega'_p}]s$</td>
<td>$\Omega'_p$</td>
</tr>
<tr>
<td>Lowpass</td>
<td>Highpass</td>
<td>$s \rightarrow \frac{\Omega_p\Omega'_p}{s}$</td>
<td>$\Omega'_p$</td>
</tr>
<tr>
<td>Lowpass</td>
<td>Bandpass</td>
<td>$s \rightarrow \frac{\Omega_p(s^2 + \Omega_l\Omega_u)}{(\Omega_u - \Omega_l)s}$</td>
<td>$\Omega_l, \Omega_u$</td>
</tr>
<tr>
<td>Lowpass</td>
<td>Bandstop</td>
<td>$s \rightarrow \frac{\Omega_p(\Omega_u - \Omega_l)s}{(s^2 + \Omega_l\Omega_u)}$</td>
<td>$\Omega_l, \Omega_u$</td>
</tr>
</tbody>
</table>
Example: Frequency transformation for analog filter

• Transform the single-pole lowpass Butterworth filter with system function:
  \[ H(s) = \frac{\Omega_p}{s + \Omega_p} \]
  into a bandpass filter with upper and lower band edge frequencies \( \Omega_u \) and \( \Omega_l \), respectively.

• The desired transformation is given in previous table. Thus we have
  \[ H(s) = \frac{1}{[\{(s^2 + \Omega_l \Omega_u)/(\Omega_u - \Omega_l)s\} + 1]} = \frac{(\Omega_u - \Omega_l)s}{[s^2 + (\Omega_u - \Omega_l)s + \Omega_l \Omega_u]} \]
The resulting filter has a zero at \( s = 0 \) and poles at
  \[ s = \frac{[-(\Omega_u - \Omega_l) \pm \sqrt{\Omega_u^2 + \Omega_l^2 - 6 \Omega_u \Omega_l}]}{2} \]
Frequency Transformations in the Digital Domain

- The transformation involves replacing the variable $z^{-1}$ by a rational function $g(z^{-1})$, which must satisfy the following properties.
  - The mapping $z^{-1} \rightarrow g(z^{-1})$ must map points inside the unit circle in the $z$-plane into itself.
  - The unit circle must also be mapped into itself.
- Condition (2) implies that for $r = 1$,
  \[ e^{-j\omega} = g(e^{-j\omega}) \quad g(\omega) = |g(\omega)|e^{j\arg[g(\omega)]} \]
  It is clear that $|g(\omega)| = 1$ for all $\omega$. That is, the mapping must be all-pass. Hence it is of the form
  \[ g(z^{-1}) = \pm \prod_{k=1}^{n} \frac{z^{-1} - a_{k}}{1 - a_{k}z^{-1}} \]
  where $|a_{k}| < 1$ to ensure that a stable filter is transformed into another stable filter.
Frequency transformation for digital filters

<table>
<thead>
<tr>
<th>Transform to</th>
<th>Transformation Operator</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowpass</td>
<td>$z^{-1} \rightarrow \frac{(z^{-1} - a)}{(1 - az^{-1})}$</td>
<td>$a = \frac{\sin((\omega_p - \omega_{p'})/2)}{\sin((\omega_p + \omega_{p'})/2)}$</td>
</tr>
<tr>
<td>Highpass</td>
<td>$z^{-1} \rightarrow -\frac{(z^{-1} + b)}{(1 + bz^{-1})}$</td>
<td>$b = -\frac{\cos((\omega_p + \omega_{p'})/2)}{\cos((\omega_p - \omega_{p'})/2)}$</td>
</tr>
</tbody>
</table>
| Bandpass    | $z^{-1} \rightarrow -\frac{(z^{-2} - a_1z^{-1} + a_2)}{(a_2z^{-1} - a_1z^{-1} + 1)}$ | $a_1 = -2\alpha K/(K+1)$  
$\alpha = \frac{(K-1)}{(K+1)}$ |
| Bandstop    | $z^{-1} \rightarrow -\frac{(z^{-2} - b_1z^{-1} + b_2)}{(b_2z^{-1} - b_1z^{-1} + 1)}$ | $b_1 = -2\alpha/(M+1)$  
$b_2 = (1-M)/(1+M)$ |

Where $a = \frac{\cos((\omega_u + \omega_l)/2)}{\cos((\omega_u - \omega_l)/2)}$, $K = \cot[(\omega_u - \omega_l)/2] \tan(\omega_p/2)$, and $M = \tan [(\omega_u - \omega_l)/2] \tan(\omega_p/2)$.

$\omega_p$ is passband edge frequency new filter, $\omega_l$ and $\omega_u$ are lower and upper band edge frequencies.
Example: Frequency transformation for digital filter

• Convert the single-pole lowpass Butterworth filter with system function

\[ H(z) = \frac{0.245(1 + z^{-1})}{1 - 0.509z^{-1}} \]

into a bandpass filter with upper and lower cutoff frequencies \( \omega_u \) and \( \omega_l \), respectively. The lowpass filter has 3-dB bandwidth \( \omega_p = 0.2\pi \).

• The desired transformation is

\[ z^{-1} \rightarrow \frac{z^{-2} - a_1 z^{-1} + a_2}{a_2 z^{-2} - a_1 z^{-1} + 1} \]

where \( a_1 \) and \( a_2 \) are defined in previous table. Substitution into \( H(z) \) yields

\[
H(z) = 0.245\frac{1-[z^{-2}-a_1 z^{-1}+a_2]}/[a_2 z^{-2} - a_1 z^{-1} + 1]}{1+0.509[z^{-2}-a_1 z^{-1}+a_2]/[a_2 z^{-2} - a_1 z^{-1} + 1]}
\]

\[
= 0.245(1-a_2)(1-z^{-2})/[(1+0.509a_2)-1.509a_1z^{-1}+(a_2+0.509)z^{-2}]
\]

Note that the resulting filter has zeros at \( z = \pm 1 \) and a pair of poles that depend on the choice of \( \omega_u \) and \( \omega_l \).
• The impulse invariant method and the mapping of derivatives are inappropriate to use in designing highpass and many bandpass filters, due to the aliasing problem.
  - Consequently, one would not employ an analog frequency transformation followed by conversion of the result into the digital domain by use of these two mappings.
  - It is much better to perform the mapping from an analog lowpass filter into a digital lowpass filter by either of these mappings, and then to perform the frequency transformation in the digital domain to avoid the aliasing problem.
  - In the case of the bilinear transformation, where aliasing is not a problem, it does not matter whether the frequency transformation is performed in the analog domain or in the digital domain.
Directly Design of Digital Filters

• Methods for designing digital filters directly:
  - Pade Approximation Method (the specifications are given in the time domain and the design is also carried out in the time domain.)
  - Least-Squares Design Method (the specifications are given in the time domain and the design is also carried out in the time domain.)
  - Least-Squares Techniques (the design is carried out in the frequency domain.)
Pade Approximation Method

- Suppose that the desired impulse response $h_d[n]$ is specified for $n \geq 0$. The filter to be designed has the system function

$$H(z) = \frac{\sum_{k=0}^{M} b_k z^{-k}}{1 + \sum_{k=1}^{N} a_k z^{-k}} = \sum_{k=0}^{\infty} h[k] z^{-k}$$

where $h[k]$ are $L = M+N+1$ parameters, coefficients $\{a_k\}$ and $\{b_k\}$, which can be used to minimize some error criterion (the least-square error or the sum of the squared errors).
In general, $h[n]$ is a nonlinear function of the filter parameters and hence the minimization of errors involves the solution of a set of nonlinear equations.

However, if we select the upper limit as $U = L-1$, it is possible to match $h[n]$ perfectly to the desired response $h_d[n]$ for $0 \leq n \leq M+n$.

- The difference equation for the desired filter is

$$y[n] = - a_1 y[n-1] - a_2 y[n-2] - \ldots - a_N y[n-N] + b_0 x[n] + b_1 x[n-1] + \ldots + b_M x[n-M].$$

Suppose that the input to the filter is a unit sample ($x[n] = \delta[n]$). Then the response of the filter is $y[n] = h[n]$ and hence

$$h[n] = - a_1 h[n-1] - a_2 h[n-2] - \ldots - a_N h[n-N] + b_0 \delta[n] + b_1 \delta[n-1] + \ldots + b_M \delta[n-M].$$

Since $\delta[n-k] = 0$ except for $n = k$, above equation reduces to

$$h[n] = - a_1 h[n-1] - a_2 h[n-2] - \ldots - a_N h[n-N] + b_M, \quad 0 \leq n \leq M$$

For $n > M$, it becomes

$$h[n] = - a_1 h[n-1] - a_2 h[n-2] - \ldots - a_N h[n-N].$$

- The set of last two equations can be used to solve for filter parameters.
Example: The Pade approximation method

- Suppose that the desired unit sample response is
  \[ h_d[n] = 2(\frac{1}{2})^n u[n] \]
  Determine the parameters of the filter with system function
  \[ H(z) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1}} \]
  using the Pade approximation technique.

- With \( \delta[n] \) as the input to \( H(z) \), we obtain the output
  \[ h[n] = -a_1 h[n-1] + b_0 \delta[n] + b_1 \delta[n-1] . \]
  For \( n > 1 \), we have \( h[n] = -a_1 h[n-1] \)
  or, equivalently, \( h_d[n] = a_1 h_d[n-1] \).

- With the substitution for \( h_d[n] \), we obtain \( a_1 = -\frac{1}{2} \). Thus
  \[ h_d[n] = \frac{1}{2} h_d[n-1] + b_0 \delta[n] + b_1 \delta[n-1] . \]
  For \( n = 0 \) this equation yields \( b_0 = 2 \).
  For \( n = 1 \) we obtain the result \( b_1 = 0 \).
Example: The Pade approximation method

- A fourth-order Butterworth filter has the system function
  \[ H_d(z) = \frac{0.0048334(z + 1)^4}{(z^2 - 1.3205z + 0.6326)(z^2 - 1.0482z + 0.2959)} \]
  The unit impulse response corresponding to \( H_d(z) \) is illustrated in Figure below. Use the Pade approximation method to approximate \( H_d(z) \).

- The desired filter has \( M = 4 \) zeros and \( N = 4 \) poles. However, we try to determine the coefficients in the Pade approximation when the number of zeros and/or poles are different from the desired number of filter parameters.

- In the following Figure, the frequency response of the filters obtained by the Pade approximation method are plot in four cases: \( M=3,N=5; M=3,N=4; M=N=4; M=4,N=5 \).

- When \( M=3 \), the resulting frequency response is a relatively poor approximation to the desired response.

- However, an increase in the number of poles from \( N=4 \) to \( N=5 \) appears to compensate in part for the lack of the one zero.

- When \( M \) is increases from 3 to 4, we obtain a perfect match.
Magnitude response

Desired response (4-th order Butterworth)

$N = 4, M = 4$

$N = 5, M = 4$

$N = 5, M = 3$

$N = 4, M = 3$
Example: The Pade approximation method

- A three-pole and three-zero Type II lowpass Chebyshev digital filter has the system function
  
  \[ H_d(z) = \]
Since the design method matches $h_d[n]$ only up to the number of filter parameters, the more complex the filter, the better the approximation.

The major limitation with the Pade approximation method is the resulting filter must contain a large number of poles and zeros. Thus this method has found limited use in filter designs for practical applications.
Least Squares Design Methods

• The least-squares method in the following figure provides good estimates only for the pole parameters.
• Prony’s method: The parameter \( \{a_k\} \) that determine the poles are obtained by the method of least square while the parameters \( \{b_k\} \) that determine the zeros are obtained by the Pade approximation method.
Least-squares method for determining the poles and zeros of a filter following the Shank’s method

\[ \delta[n] \leftarrow \text{All-pole filter} \quad H_1(z) \leftarrow v[n] \quad \text{All-zero filter} \quad H_2(z) \rightarrow h_d[n] \]

- Shank's method: Both sets of parameters \( \{a_k\} \) and \( \{b_k\} \) are determined by application of the least-squares method (proposed by Shanks [1967]). Firstly, the parameter \( \{a_k\} \) are computed and then the parameters \( \{b_k\} \) can also be determined later.
The design is most easily carried out with the system function for the IIR filter expressed in the cascade form.

Then the filter gain and the filter coefficients are to be determined by iterative procedure in order to obtain the desired frequency response of the filter.

Instead of dealing with the phase of the filter, it is more convenient to deal with the group delay as a function of frequency.

The total weighted least-squares error over all frequencies will be minimized via the iterative process.

The major difficulty with any iterative procedure that searches for the parameter values that minimize a nonlinear function is that the process may converge to a local minimum instead of a global minimum.
Design of Discrete-Time FIR Filters

- Windowing Method
- Frequency-Sampling Method
- Optimum Approximations
- Least-Squares Inverse (Wiener) Filter
Windowing Method

- FIR filters are almost entirely restricted to discrete-time implementations.
- The design techniques for FIR filters are based on directly approximating the desired frequency response of the discrete-time system.
- Most techniques for approximating the magnitude response of an FIR system assume a linear phase constraint, thereby avoiding the problem of spectrum factorization that complicates the direct design of IIR filters.
- The windowing technique is the simplest method of FIR filter design.
- This method generally begins with an ideal desired frequency response, \( H_d(e^{j\omega}) \), and evaluates its corresponding impulse response, \( h_d[n] \). Then, the desired impulse response, \( h[n] \), will be obtained by truncating \( h_d[n] \) with selected window function, \( w[n] \).
\[ H_d(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h_d[n] e^{-j\omega n} \]
\[ h_d[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega}) e^{j\omega n} d\omega \]
\[ h[n] = h_d[n] w[n] = h_d[n], \quad 0 \leq n \leq M, \]
\[ 0, \quad \text{otherwise}. \]
\[ H(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega}) W(e^{j(\omega-\theta)}) d\theta. \]

• That is \( H(e^{j\omega}) \) is the periodic convolution of the desired ideal frequency response with the Fourier transform of the window function.
• Thus, the frequency response \( H(e^{j\omega}) \) will be a “smeared” version of the desired response \( H_d(e^{j\omega}) \).
• In the case of the rectangular window:
\[ W(e^{j\omega}) = \sum_{n=0}^{M} e^{-j\omega n} = e^{-j\omega M/2} \frac{\sin(\omega [M+1]/2)}{\sin(\omega/2)}. \]
As \( M \) increases, the width of the main lobe decreases.
• The main lobe is the region between the first zero-crossings on either side of the origin.
• Gibbs Phenomenon
Convolution process implied by truncation of the ideal impulse response

Typical approximation resulting from windowing the ideal impulse response
Magnitude of the Fourier transform of a rectangular window, $M = 7$. 

\[ \left| \frac{\sin \left( \omega \frac{(M + 1)/2}{2} \right)}{\sin \left( \omega/2 \right)} \right| \]

$\omega$ range:
- $\frac{-2\pi}{(M + 1)}$ to $\frac{2\pi}{(M + 1)}$
- Peak sidelobe
- $\Delta \omega_m$ (Mainlobe width)

$(M = 7)$
## Window Functions for FIR Filter Design

<table>
<thead>
<tr>
<th>Window Type</th>
<th>Time-Domain Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectangular</td>
<td>[ w[n] = \begin{cases} 1, &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases} ]</td>
</tr>
<tr>
<td>Bartlett</td>
<td>[ w[n] = \begin{cases} 2n/M, &amp; 0 \leq n \leq M/2 \ 2-2n/M, &amp; M/2 &lt; n \leq M \ 0, &amp; \text{otherwise} \end{cases} ]</td>
</tr>
<tr>
<td>(Triangular)</td>
<td></td>
</tr>
<tr>
<td>Hanning</td>
<td>[ w[n] = \begin{cases} 0.5 - 0.5 \cos(2\pi n/M), &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases} ]</td>
</tr>
<tr>
<td>Hamming</td>
<td>[ w[n] = \begin{cases} 0.54 - 0.46 \cos(2\pi n/M), &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases} ]</td>
</tr>
<tr>
<td>Blackman</td>
<td>[ w[n] = \begin{cases} 0.42 - 0.5 \cos(2\pi n/M) + 0.08 \cos(4\pi n/M), &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases} ]</td>
</tr>
<tr>
<td>Kaiser</td>
<td>[ w[n] = \begin{cases} I_0[\beta(1 - {(n - \alpha)/\alpha}^2)^{1/2}]/I_0(\beta), &amp; 0 \leq n \leq M, \alpha = M/2 \ 0, &amp; \text{otherwise} \end{cases} ]</td>
</tr>
</tbody>
</table>

\( I_0(.) \) is zero order modified Bessel function of the first kind, \( \beta \) is window shape parameter.
Shape of commonly used window functions.
Log magnitude of Fourier transform of window functions

Rectangular
$M = 50$

Hamming
$M = 50$

Bartlett
$M = 50$

Blackman
$M = 50$

Hanning
$M = 50$

Kaiser
$M = 20$
• The width of the main lobe and the relative side-lobe amplitude depend primarily on the window length \( L \) and the shape (amount of tapering) of the window.

• Through the choice of the shape and duration of the window, we can control the properties of the resulting FIR filter:
  - The windows with the smaller side lobes yield better approximations of the ideal response at a discontinuity.
  - The smaller width of the main lobe which can be achieved by increasing \( M \) yield the narrower transition regions.

• Kaiser [1974] has developed a simple formalization of the window method using Kaiser window.

• Kaiser window overcomes the disadvantage occurred in using other window because we must try different windows and adjust their length by trial and error method.

• Filters designed by the window method inherently have \( \delta_1 = \delta_2 \), so must use the smaller value of ripple in the design procedure.
## Frequency-Domain Characteristics of Window Functions

<table>
<thead>
<tr>
<th>Type of window</th>
<th>Peak side lobe amplitude, dB</th>
<th>Approximation width of main lobe</th>
<th>Peak approximation error, 20log₁₀δ, dB</th>
<th>Equivalent Kaiser window, β</th>
<th>Transition width of equivalent Kaiser window, Δω</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectangular</td>
<td>-13</td>
<td>4π/(M+1)</td>
<td>-21</td>
<td>0</td>
<td>1.81π/M</td>
</tr>
<tr>
<td>Bartlett</td>
<td>-25</td>
<td>8π/M</td>
<td>-25</td>
<td>1.33</td>
<td>2.37π/M</td>
</tr>
<tr>
<td>Hanning</td>
<td>-31</td>
<td>8π/M</td>
<td>-44</td>
<td>3.86</td>
<td>5.01π/M</td>
</tr>
<tr>
<td>Hamming</td>
<td>-41</td>
<td>8π/M</td>
<td>-53</td>
<td>4.86</td>
<td>6.27π/M</td>
</tr>
<tr>
<td>Blackman</td>
<td>-57</td>
<td>12π/M</td>
<td>-74</td>
<td>7.04</td>
<td>9.19π/M</td>
</tr>
<tr>
<td>Kaiser</td>
<td>-50</td>
<td></td>
<td>4.54</td>
<td>4.86π/M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-70</td>
<td></td>
<td>6.76</td>
<td>8.64π/M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-90</td>
<td></td>
<td>8.96</td>
<td>11.42π/M</td>
<td></td>
</tr>
</tbody>
</table>

Δω = ωₛ − ωₚ
Illustration of type of approximation obtained at a discontinuity of the ideal frequency response.
Kaiser Window Function

- Window shape parameter, $b$

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

  $$
  A = -20\log_{10}\delta
  $$

  $$
  M = (A - 8)/2.285\Delta\omega
  $$

  Where $\delta$ is the peak approximation error, and $\Delta\omega = \omega_s - \omega_p$ is the transition width.
Kaiser window shapes and their frequency characteristics

M = 20

$\beta = 6$

$M = 10, 20, 40.$
Windowing Algorithm

- **Step 1**: define specifications of desired filter.
- **Step 2**: evaluate the system function $H_d(e^{j\omega})$ from step 1.
- **Step 3**: evaluate the impulse response sequence $h_d[n]$ as
  \[ h_d[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega})e^{j\omega n}d\omega \]
  where $H_d(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h_d[n]e^{-j\omega n}$
- **Step 4**: obtain finite duration sequence $h[n]$ from $h_d[n]$ as
  \[ h[n] = \begin{cases} h_d[n]w[n], & 0 \leq n \leq M \\ 0, & \text{otherwise} \end{cases} \]
  where $w[n]$ is a selective window function to meet the attenuation requirement, so
  \[ H(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\theta})W(e^{j(\omega-\theta)})d\theta \]
- **Step 5**: verify the result. If it does not meet requirement, return to step 4 by reselection of window width (M) and/or type ($w[n]$).
<table>
<thead>
<tr>
<th>Filter Type</th>
<th>Ideal Impulse Response, $h_d[n]$</th>
<th>$h_d[0]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowpass</td>
<td>$2f_c \left[ \sin \left( (n-M/2)\omega_c \right) / (n-M/2)\omega_c \right]$</td>
<td>$2f_c$</td>
</tr>
<tr>
<td>Highpass</td>
<td>$-2f_c \left[ \sin \left( (n-M/2)\omega_c \right) / (n-M/2)\omega_c \right]$</td>
<td>$1-2f_c$</td>
</tr>
<tr>
<td>Bandpass</td>
<td>$2f_2 \left[ \sin \left( (n-M/2)\omega_2 \right) / (n-M/2)\omega_2 \right]$</td>
<td>$2(f_2 - f_1)$</td>
</tr>
<tr>
<td></td>
<td>$-2f_1 \left[ \sin \left( (n-M/2)\omega_1 \right) / (n-M/2)\omega_1 \right]$</td>
<td></td>
</tr>
<tr>
<td>Bandstop</td>
<td>$2f_1 \left[ \sin \left( (n-M/2)\omega_1 \right) / (n-M/2)\omega_1 \right]$</td>
<td>$1-2(f_2 - f_1)$</td>
</tr>
<tr>
<td></td>
<td>$-2f_2 \left[ \sin \left( (n-M/2)\omega_2 \right) / (n-M/2)\omega_2 \right]$</td>
<td></td>
</tr>
</tbody>
</table>

Where $f_c$, $f_1$ and $f_2$ are passband or stopband edge frequencies, and $M$ is the filter length.
Example: Kaiser window design of a lowpass filter

- Consider the lowpass digital filter specifications:
  \[0.99 \leq |H(e^{j\omega})| \leq 1.01, \quad |\omega| \leq 0.4\pi,
  \]
  \[|H(e^{j\omega})| \leq 0.001, \quad 0.6\pi \leq |\omega|.
  \]

Using the design formulas for the Kaiser window to design an FIR lowpass filter to meet prescribed specifications.
- First, we set \(\delta = 0.001\).
- Next, the cutoff frequency of the ideal lowpass filter is
  \[\omega_c = (\omega_p + \omega_s)/2 = 0.5\pi.
  \]
- To determine the parameters of the Kaiser window, we first compute
  \[\Delta\omega = \omega_s - \omega_p = 0.2\pi, \quad A = -20\log_{10}\delta = 60.
  \]
  Then,
  \[\beta = 5.653, \quad M = 37.
  \]
- The impulse response of the filter is:
  \[h[n] = \frac{\sin(\omega_c[n-\alpha])}{\pi(n-\alpha)}\left\{I_0[\beta(1-[(n-\alpha)/\alpha]^2)^{1/2}] / I_0(\beta)\right\}, \quad 0 \leq n \leq M,
  \]
  \[0, \quad \text{otherwise.}
  \]

where \(\alpha = M/2 = 18.5\)
The Response Functions of Lowpass Filter for Kaiser Windows of $\beta = 5.653$ and $M = 37$. 

- Impulse response
- Log magnitude
- Approximation error
• Increasing the order of $M$ may lead to more unsatisfactory result.
• Type II FIR linear-phase systems are generally not appropriate approximations for either highpass or bandstop filters.
The response function for type I and II FIR highpass filters.

**Type I:** $M = 24$

**Type II:** $M = 25$

- **Impulse response**
- **Log magnitude**
- **Approximation error**
## Window Functions for FIR Filter Design

<table>
<thead>
<tr>
<th>Window Type</th>
<th>Time-Domain Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectangular</td>
<td>$w[n] = \begin{cases} 1, &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td>Bartlett</td>
<td>$w[n] = \begin{cases} 2n/M, &amp; 0 \leq n \leq M/2 \ 2-2n/M, &amp; M/2 &lt; n \leq M \ 0, &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td>(Triangular)</td>
<td></td>
</tr>
<tr>
<td>Hanning</td>
<td>$w[n] = \begin{cases} 0.5 - 0.5\cos(2\pi n/M), &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td>Hamming</td>
<td>$w[n] = \begin{cases} 0.54 - 0.46\cos(2\pi n/M), &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td>Blackman</td>
<td>$w[n] = \begin{cases} 0.42 - 0.5\cos(2\pi n/M) + 0.08\cos(4\pi n/M), &amp; 0 \leq n \leq M \ 0, &amp; \text{otherwise} \end{cases}$</td>
</tr>
<tr>
<td>Kaiser</td>
<td>$w[n] = \begin{cases} I_0[\beta(1 - {(n - \alpha)/\alpha}^2)^{1/2}]/I_0(\beta), &amp; 0 \leq n \leq M, \alpha = M/2 \ 0, &amp; \text{otherwise} \end{cases}$</td>
</tr>
</tbody>
</table>

$I_0(.)$ is zero order modified Bessel function of the first kind, $\beta$ is window shape parameter.
Frequency-Domain Characteristics of Window Functions

<table>
<thead>
<tr>
<th>Type of window</th>
<th>Peak side lobe amplitude, dB</th>
<th>Approximation width of main lobe: $\Delta \omega$</th>
<th>Peak approximation error, $20\log_{10}\delta$, dB</th>
<th>Equivalent Kaiser window, $\beta$</th>
<th>Transition width of equivalent Kaiser window, $\Delta \omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectangular</td>
<td>-13</td>
<td>$4\pi/(M+1)$</td>
<td>-21</td>
<td>0</td>
<td>$1.81\pi/M$</td>
</tr>
<tr>
<td>Bartlett</td>
<td>-25</td>
<td>$8\pi/M$</td>
<td>-25</td>
<td>1.33</td>
<td>$2.37\pi/M$</td>
</tr>
<tr>
<td>Hanning</td>
<td>-31</td>
<td>$8\pi/M$</td>
<td>-44</td>
<td>3.86</td>
<td>$5.01\pi/M$</td>
</tr>
<tr>
<td>Hamming</td>
<td>-41</td>
<td>$8\pi/M$</td>
<td>-53</td>
<td>4.86</td>
<td>$6.27\pi/M$</td>
</tr>
<tr>
<td>Blackman</td>
<td>-57</td>
<td>$12\pi/M$</td>
<td>-74</td>
<td>7.04</td>
<td>$9.19\pi/M$</td>
</tr>
<tr>
<td>Kaiser</td>
<td></td>
<td></td>
<td></td>
<td>-50</td>
<td>$4.54\pi/M$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-70</td>
<td>$6.76\pi/M$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-90</td>
<td>$8.96\pi/M$</td>
</tr>
</tbody>
</table>

$\Delta \omega = \omega_s - \omega_p$  
Transition width = width of the main lobe
Example: Windowing Method → Hamming Window

- Obtain the coefficients of an FIR lowpass filter to meet the specifications given below using the window method.
  
  - passband edge frequency: 1.5 kHz
  - transition width: 0.5 kHz
  - stopband attenuation: > 50 dB
  - sampling frequency: 8 kHz

- We select $h_d[n]$ for lowpass filter which is given by:

\[ h_d[n] = 2f_c \frac{\sin((n-M/2)\omega_c)}{((n-M/2)\omega_c)} = \frac{\sin(n\omega_c)}{n\pi}, \quad n \neq M/2, \]
\[ 2f_c, \quad n = M/2 \]

- From characteristics table, it indicates that only the Hamming, Blackman or Kaiser ($b = 4.54$) windows will satisfy the stopband attenuation requirements. If we use the Hamming window for simplicity.

\[ \Delta f = 0.5k/8k = 0.0625 \rightarrow \frac{8\pi}{M} = \Delta\omega \rightarrow M = 64 \]

And $h[n] = h_d[n]w[n]$, 
\[ w[n] = \begin{cases} 
0.54 - 0.46\cos(2\pi n/M), & 0 \leq n \leq M, \\
0, & \text{otherwise}.
\end{cases} \]
\[ h[0] = h[64] = h_d[0]w[0] = 0 \times 0.08 = 0 \]
\[ h[1] = h[63] = h_d[1]w[1] = -0.01007 \times 0.08222 = -0.00083 \]
\[ h[3] = h[61] = h_d[3]w[3] = 0.00913 \times 0.09981 = 0.00091 \]
\[ h[4] = h[60] = h_d[4]w[4] = 0.00804 \times 0.11502 = 0.00093 \]
\[ h[5] = h[59] = h_d[5]w[5] = -0.00655 \times 0.13432 = -0.00088 \]
\[ h[6] = h[58] = h_d[6]w[6] = -0.01131 \times 0.15752 = -0.00178 \]
\[ \ldots \]
\[ h[30] = h[34] = h_d[30]w[30] = 0.06091 \times 0.99116 = 0.06037 \]
\[ h[31] = h[33] = h_d[31]w[31] = 0.31219 \times 0.99779 = 0.31150 \]

Where \( f_c \) will be chosen to the center of the transition band = \( f_c + \Delta f/2 = [1.5k + 0.5k/2]/8k = 0.21875 \)
Example: Windowing Method → Kaiser Window

- Design an FIR digital filter to meet the following specifications:
  - passband: 150~250 Hz
  - transition width: 50 Hz
  - passband ripple: 0.1 dB
  - stopband attenuation: 60 dB
  - sampling frequency: 1 kHz

Obtain the filter coefficients using the window method.

- Compare the ripples: $20\log_{10}(1 + \delta_p) = 0.1 \text{ dB} \rightarrow \delta_p = 0.0115$
  and $-20\log_{10}\delta_s = 60 \text{ dB} \rightarrow d_s = 0.001 < \delta_p$
  Thus $\delta = \min(\delta_p, \delta_s) = 0.001 \rightarrow A = -20\log_{10}\delta = 60 \text{ dB}$

- The attenuation requirements (60 dB) can only be met by the Kaiser or the Blackman window. If we select the Kaiser window, the number of coefficients is

  $M = (A - 8)/(2.285\Delta\omega) = 72.44 \rightarrow M = 73$, and $b = 0.1102(A - 8.7) = 5.65$.

  where $\Delta f = 50/1k = 0.05$. 
• When we select the Blackman window:
\[ \Delta \omega = \frac{12\pi}{M} \rightarrow M = 120 \]
• It is clearly that the complexity of the designed filter using the Blackman window is nearly 2 times greater than that using the Kaiser window.
Advantages and Disadvantages of the Window Method

- **Simplicity**
  - It is simple to apply and simple to understand. It involves a minimum amount of computational effort, even for the more complicated Kaiser window.

- **Lack of flexibility**
  - Both the peak passband and stopband ripples are approximately equal, so that the designer may end up with either too small a passband ripple or too large a stopband attenuation.

- **Unprecision**
  - Because of the effect of convolution of the spectrum of the window function and the desired response, the passband and stopband edge frequencies cannot be precisely specified.

- **Clumsy (trial and error technique)**
  - For a given window (except the Kaiser) the maximum ripple amplitude in the filter response is fixed regardless of how large we make N. Thus the stopband attenuation for a given window is fixed. Thus, for a given attenuation specification, the filter designer must find a suitable window.

- **Lack of capability**
  - In some applications, the expression for the desired filter response, $H_d(\omega)$, will be too complicated for $h_d[n]$ to be obtained analytically. In these cases $h_d[n]$ may be obtained via the frequency sampling method before the window function is applied.
Frequency Sampling Method

- This method allows us to design nonrecursive FIR filters for both standard frequency selective filters (lowpass, highpass, bandpass) and filters with arbitrary frequency response.
- It also allows recursive implementation of FIR filters, leading to computationally efficient filters.

From the DFT, \( h[n] = \frac{1}{N} \sum_{k=0}^{N-1} H[k] e^{j2\pi nk/N} \), it can be shown that for linear phase filters, with positive symmetrical impulse response and for \( N \) even:

\[
h[n] = \frac{1}{N} \sum_{k=0}^{(N/2)-1} |H[k]| \cos \left(2\pi k(n - \frac{N-1}{2})/N \right) + H(0)
\]

where \( \alpha = (N - 1)/2 \). And for \( N \) odd, it becomes:

\[
h[n] = \frac{1}{N} \sum_{k=0}^{(N-1)/2} |H[k]| \cos \left(2\pi k(n - \alpha)/N \right) + H(0)
\]
Frequency Sampling Algorithm

- Step 1: define specifications of desired filter.
- Step 2: select frequency sample type
  - Type I: sampling frequency position is at $kF_s/N$.
  - Type II: sampling frequency position is at $(k + \frac{1}{2})F_s/N$.
- Step 3: calculate required total number of frequency sample, $N$, and evaluate the number of frequency sample in transition band, $M$, and their magnitudes, $T_i$; $i = 1, 2, ..., M$.
- Step 4: evaluate coefficient values of the filter using appropriate formula.
- Step 5: verify the result. If it does not meet requirement, return to step 3 to reselect $N$ and/or $M$ or step 2 to reselect frequency sample type.
Example: Frequency Sampling Method

- Consider a lowpass FIR filter with the following specifications:
  - passband: 0-5 kHz
  - sampling frequency: 18 kHz
  - filter length: 9

Obtain the filter coefficients using the frequency sampling method.

- The frequency samples are taken at intervals of $kF_s/N$, that is at intervals of $18/9 = 2$ kHz. Thus the frequency samples are given by
  $$|H[k]| = 1 \text{ at } k = 0, 1, 2$$
  $$= 0 \text{ at } k = 3, 4$$

- Because $N$ is even, then:
  $$h[0] = h[8] = 7.2522627 \times 10^{-2}$$
  $$h[1] = h[7] = -1.1111111 \times 10^{-1}$$
  $$h[4] = 5.5555556 \times 10^{-1}.$$
(a) Ideal frequency response showing sampling points.
(b) Frequency response of frequency sampling filter.

The 4 possible z-plane sampling grids for the 2 types of frequency sampling filters.
Example: Frequency sampling method

- Determine the coefficients of a linear-phase FIR filter of length $M = 15$ which has a symmetric unit sample response and a frequency response that satisfies the conditions:
  \[ H[2\pi k/15] = \begin{cases} 1 , & k = 0, 1, 2, 3 \\ 0.4 , & k = 4 \\ 0 , & k = 5, 6, 7 \end{cases} \]

- Since $h[n]$ is symmetric and the frequencies are selected to correspond to the case of Type I and because $N$ is even,

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.014112893</td>
<td>-0.001945309</td>
<td>0.04000004</td>
<td>0.01223454</td>
<td>-0.09138802</td>
<td>-0.01808986</td>
<td>0.3133176</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Optimizing the Amplitude Response


- For a lowpass filter, the stopband attenuation increases, approximately, by 20 dB for each transition band frequency sample [Rabiner et al., 1970], with a corresponding increase in the transition width:

  Approximate stopband attenuation \((25 + 20M)\) dB

  Approximate transition width \((M + 1)F_s/N\)

  where \(M\) is the number of transition band frequency samples and \(N\) is the filter length.

  - For one transition frequency sample: \(0.250 < T_1 < 0.450\)
  - For two transition frequency samples: \(0.250 < T_1 < 0.150\), \(0.450 < T_2 < 0.650\)
  - For three transition frequency samples: \(0.003 < T_1 < 0.035\), \(0.100 < T_2 < 0.300\), \(0.550 < T_3 < 0.750\)
Lowpass filter frequency samples including three transition band samples.
Optimum transition band frequency samples for type I lowpass frequency sampling filters for N = 15 [adapted from Rabiner et al., 1970]

<table>
<thead>
<tr>
<th>BW</th>
<th>Stopband attenuation (dB)</th>
<th>T_1</th>
<th>T_2</th>
<th>T_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>One transition band frequency sample, N = 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>42.309 322 83</td>
<td>0.433 782 96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>41.262 992 86</td>
<td>0.417 938 23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>41.253 337 86</td>
<td>0.410 473 63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>41.949 077 13</td>
<td>0.404 058 84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>44.371 245 38</td>
<td>0.392 681 89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>56.014 165 88</td>
<td>0.357 665 25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two transition band frequency samples, N = 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>70.605 405 85</td>
<td>0.095 001 22</td>
<td>0.589 954 18</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>69.261 681 56</td>
<td>0.103 198 24</td>
<td>0.593 571 18</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>69.919 734 95</td>
<td>0.100 836 18</td>
<td>0.589 432 70</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>75.511 722 56</td>
<td>0.084 074 93</td>
<td>0.557 153 12</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>103.460 783 00</td>
<td>0.051 802 06</td>
<td>0.499 174 24</td>
<td></td>
</tr>
<tr>
<td>Three transition band frequency samples, N = 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>94.611 661 91</td>
<td>0.014 550 78</td>
<td>0.184 578 82</td>
<td>0.668 976 13</td>
</tr>
<tr>
<td>2</td>
<td>104.998 130 80</td>
<td>0.010 009 77</td>
<td>0.173 607 13</td>
<td>0.659 515 26</td>
</tr>
<tr>
<td>3</td>
<td>114.907 193 18</td>
<td>0.008 734 13</td>
<td>0.163 973 10</td>
<td>0.647 112 64</td>
</tr>
<tr>
<td>4</td>
<td>157.292 575 84</td>
<td>0.003 787 99</td>
<td>0.123 939 63</td>
<td>0.601 811 54</td>
</tr>
</tbody>
</table>

BW refers to the number of frequency samples in the passband.
Frequency response of frequency sampling filter with (a) no transition band samples; (b) one transition band sample; (c) two transition band samples; (d) three transition Band samples.
The Optimal Method

- The optimal method is based on the concept of “equiripple” passband and stopband.
- The window method and the frequency sampling method have a major problem that is the lack of precise control of the critical frequencies such as $\omega_p$ and $\omega_s$.
- The filter design method selected to implement the optimal design is formulated as a Chebyshev approximation problem.
- It will be viewed as an optimum design criterion in the sense that the weighted approximation error between the desired frequency response and the actual frequency response is spread evenly across the passband and evenly across the stopband of the filter minimizing the maximum error.
- The resulting filter designs have ripples in both the passband and stopband.
Optimal Approximation of FIR Filters

- Alternation Theorem and Polynomials
  - It provides a necessary and sufficient condition for a polynomial to satisfy in order that it is the polynomial that minimizes the maximum weighted error for a given order.

- The Parks-McClellan Algorithm
  - Parks and McClellan [1972] applied the alternation theorem to the optimum approximation of FIR filter design problem.
Alternation Theorem

• In the Chebyshev approximation approach, the amplitude response of a type I linear phase lowpass N-tap FIR filter is
  \[ A(e^{j\omega}) = \sum_{k=0}^{r} a_k \cos(\omega k) \]
  where \( r = (N+1)/2 \).

• The response \( A(e^{j\omega}) \) will be unique, best-weighted Chebyshev approximation to the desired response \( D(e^{j\omega}) \) iff the error function \( \varepsilon = W(e^{j\omega})[D(e^{j\omega}) - A(e^{j\omega})] \) exhibits at least \( r + 1 \) extrema at frequencies in both passband and stopband.
  - \( W(e^{j\omega}) \) is the weighting function in each band.

• The frequencies at which extrema occur are called extremal frequencies.

• The maxima and minima alternate (hence, alternation theorem).
The Parks-McClellan Algorithm

• Parks and McClellan [1972] have written a computer program for designing linear phase FIR filters based on the Chebyshev approximation criterion and implemented with the Remez exchange algorithm.

• This program can be used to design lowpass, highpass or bandpass filters, differentiators, and Hilbert transformers.

• This algorithm requires a number of input parameters which determine the filter characteristics.

• If the length of the filter is increased, the attenuation in the stopband will be decreased.

• To increase the attenuation in the stopband by keeping the filter length fixed, it should decrease the weighting function in the passband.

• In general, there is no guarantee that the transition regions of a multiband filter will be monotonic, because the Parks-McClellan algorithm leaves these regions completely unconstrained.
Flowchart of Parks-McClellan Algorithm

1. Initial guess of (L+2) extremal frequencies
2. Calculate the optimum $\delta$ on extremal set
3. Interpolate through (L+1) points to obtain $A_\omega(e^{j\omega})$
4. Calculate error $E(\omega)$ and find local maxima where $|E(\omega)| \geq \delta$
5. More than (L+2) Extrema?
   - Yes → Retain (L+2) largest extrema
   - No → Check whether the extremal points changed
     - Changed → Retain (L+2) largest extrema
     - Unchanged → Best approximation
Least-Squares Inverse (Wiener) Filter Method

- The least-squares error criterion can be used to optimize the M+1 coefficients of the FIR filter.
- The error sequence $e[n]$ between the desired output sequence $d[n]$ and the actual output sequence is
  
  \[ e[n] = d[n] - \sum_{k=0}^{M} b_k h[n-k] \rightarrow \min \varepsilon = \min \sum_{n=0}^{\infty} e^2[n] \]

where the \{b_k\} are the filter coefficients.
Advantage and Disadvantage of FIR Design Methods

- **Windowing method**
  - Most simplify method, and simple understandably conceptual design.
  - Critical frequencies and/or ripples in frequency bands could not manipulated into the desired precision easily.
  - Equally ripple in each frequency band.

- **Frequency sampling method**
  - Technique may be selected as both recursive and non-recursive.
  - Applicable to both typical and general filter types.
  - Problem to manipulate band edge frequencies and passband ripple into the desired precision.

- **Optimum method**
  - All of parameters can be manipulated.
  - Coefficient calculation method is easy and efficient.
  - For the same value of M, the result in amplitude is the best.
  - For some filter, i.e. Hilbert transformer, differentiator, this technique is more suitable for in comparable to another method.
Selection of Using FIR and IIR Filter

- If linear phase is required → FIR
- If stability is required → FIR (since it can be non-recursive technique)
- Finite word-length effect to FIR less than to IIR.
- If sharp cut-off frequencies are required → FIR requires more coefficients, processing time, and memory size than IIR. (However, FFT algorithm or multirate technique may be used for FIR to compensate these disadvantages.)
- IIR can be used analog filter as prototype, but FIR can be synthesized more easily for any required frequency response. (However, in general, to synthesize FIR is required CAD because its algebraic design technique is very difficult.)