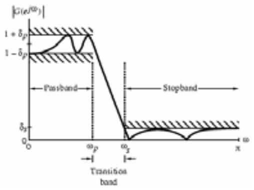
9 IIR Digital Filter Design

Digital Filter Design Steps

1) Derivation of a realizable transfer function $G(z)$
2) Realization of $G(z)$ using a suitable filter structure

   A choice between IIR and FIR digital filter has to be made

Digital Filter Specifications



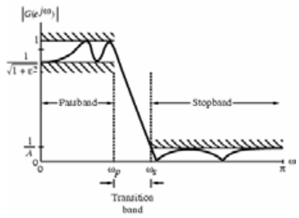
- Passband: $1-\delta_p \le |G(e^{j\omega})| \le 1+\delta_p \quad |\omega| \le \omega_p$
- Stopband: $|G(e^{j\omega})| \le \delta_s \quad \omega_p \le |\omega| \le \pi$
- Peak passband ripple: $\alpha_p = -20\log_{10}(1-\delta_p)$ dB
- Minimum stopband attenuation: $\alpha_s = -20\log_{10}(\delta_s)$ dB

Normalized Specifications

- Maximum value of the magnitude (gain) is assumed to be unity or the minimum value of the loss function is 0 dB

$$\alpha_{\max} = 20\log_{10}\left(\sqrt{1+\varepsilon^2}\right)$$

$$\alpha_{\max} \cong -20\log_{10}(1-2\delta_p)$$
$$\cong 2\alpha_p, \quad \delta_p << 1$$



Normalized Frequency Specifications

- Frequency specifications are normalized using the sampling rate:

$$\omega_p = \frac{\Omega_p}{F_T} = \frac{2\pi F_p}{F_T} = 2\pi F_p T$$

$$\omega_s = \frac{\Omega_s}{F_T} = \frac{2\pi F_s}{F_T} = 2\pi F_s T$$

- $\omega=\pi$ corresponds to the half of the sampling rate, $F_T/2$

Selection of the Filter Type

IIR filters:

- The transfer function is a real rational function of z^{-1}

$$H(z) = \frac{P(z)}{D(z)} = \frac{p_0 + p_1 z^{-1} + \ldots + p_{M-1}z^{M-1} + p_M z^{-M}}{d_0 + d_1 z^{-1} + \ldots + d_{N-1}z^{N-1} + d_N z^{-N}}$$

- $H(z)$ must be a stable transfer function
- For reduced computational complexity, it must be of lowest order

Selection of the Filter Type

FIR filters:

- The transfer function is a polynomial in z^{-1}

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

- For reduced computational complexity, the degree N of $H(z)$ must be as small as possible
- In addition, if linear phase is desired, then the FIR filter coefficients must satisfy the symmetry constraint

$$h[n]=\pm h[N-n]$$

Selection of the Filter Type

FIR filters:

\+ Linear phase response
\+ Stability with quantized coefficients
\- Higher order required than using IIR filters

IIR filters:

\+ Better attenuation properties
\+ Closed form approximation formulas
\- Nonlinear phase response
\- Instability (limit cycle oscillations) with finite wordlength computation

N_{FIR}/N_{IIR} is typically of the order of tens (or more)

IIR Filter Design

An analog filter transfer function $H_a(s)$ is transformed into the desired digital filter transfer function $G(z)$

- Analog approximation techniques are highly advanced
- They usually yield closed-form solutions
- Extensive tables are available for analog filter design or the methods are easy to program
- Digital filters often replace (or simulate) analog filters

$$H_a(s)=\frac{P_a(s)}{D_a(s)} \quad \Rightarrow \quad G(z)=\frac{P(z)}{D(z)}$$

Requirements of the Mapping

The basic idea behind the conversion of an analog prototype transfer function $H_a(s)$ to a digital filter transfer function $G(z)$ is to apply a mapping from the s-domain to the z-domain so that the essential properties of the analog frequency response are preserved

Requirements for the mapping are:

- The imaginary axis ($j\Omega$) of the the s-plane is mapped onto the unit circle in the z-plane
- Stable $H_a(s)$ must be transformed into a stable $G(z)$

Most widely used transformation: Bilinear Transformation

Impulse Invariance Method

A straight-forward approach for obtaining the digital filter:

- The impulse response of the digital filter is made identical to the impulse response of an analog prototype filter at sampling instants
- Analog transfer function: $H_a(s)$

$$h_a(t)=\mathcal{L}^{-1}\{H_a(s)\}$$

- The impulse response of the digital filter is obtained by sampling:

$$g[n]=h_a(nT), \quad n=0,1,2,...$$

Impulse Invariance Method

- The digital filter transfer function $G(z)$ is:

$$G(z)=\mathcal{Z}\{g[n]\}=\mathcal{Z}\{h_a[nT]\}$$
$$=\frac{1}{T}\sum_{k=-\infty}^{\infty} H_a\left(s+j\frac{2\pi k}{T}\right)\Bigg|_{s=\frac{1}{T}\ln z}$$

- The frequency responses are obtained by substituting $z=e^{j\omega}$ and $s=j\Omega$:

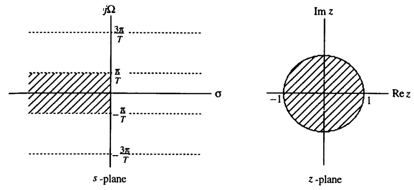
$$G(e^{j\omega})=\frac{1}{T}\sum_{k=-\infty}^{\infty} H_a\left(j\Omega+j\frac{2\pi k}{T}\right)$$

- According to the sampling theorem $G(e^{j\omega})$ is a periodic version of $H_a(j\Omega)$

Impulse Invariance Method

- Transformation from s-plane to z-plane: $z = e^{sT}$
- For $s=\sigma_0+j\Omega_0$: $z = re^{j\omega} = e^{\sigma_0 T} e^{j\Omega_0 T}$
  $|z| = r = e^{\sigma_0 T}$
- A point on the $j\Omega$-axis in the s-plane is characterized by $\sigma_0=0$
- A point on the frequency axis is mapped to a point on the unit circle in the z-plane
- A point on the left-half s-plane with $\sigma_0<0$ is mapped to z-plane with $|z|<1$, i.e., the left-half s-plane is mapped inside the unit circle

Impulse Invariance Method

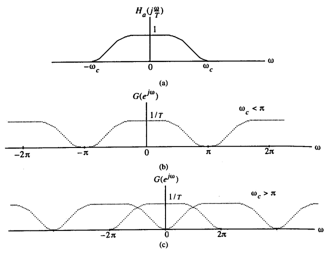


Thus, the impulse invariance mapping has the desired properties:

1) Frequency axis $j\Omega$ corresponds to unit circle
2) Stability is preserved

Impulse Invariance Method

- Due to sampling the mapping is many-to-one



- The strips of length $2\pi/T$ are all mapped onto the unit circle

Aliasing!

Impulse Invariance Method

- Derivation of $G(z)$ using impulse invariance method:

$$H_a(s) = \frac{A}{s+\alpha} \quad \leftrightarrow \quad h_a(t) = Ae^{-\alpha t}\mu(t)$$

- By sampling $h_a(t)$

$$g[n] = h_a(nT) = Ae^{-\alpha nT}\mu[n]$$

$$G(z) = A\sum_{n=0}^{\infty} e^{-\alpha nT} z^{-n} = \frac{A}{1-e^{-\alpha T}z^{-1}}$$

- $G(z)$ converges if $|e^{-\alpha T}|<1$ or $\alpha>0$, indicating that $H_a(s)$ is stable

Impulse Invariance Method

- Generalizing to higher order (N) analog transfer functions

$$H_a(s) = \sum_{k=1}^{N} \frac{A_k}{s+\alpha_k}$$

$$h_a(t) = \sum_{k=1}^{N} A_k e^{-\alpha_k t}\mu(t)$$

$$\Rightarrow \quad G(z) = \sum_{k=1}^{N} \frac{A_k}{1-e^{-\alpha_k T}z^{-1}}$$

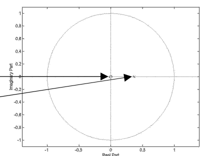
Example: First Order Butterworth Filter Designed Using the Impulse Invariant Method

$$H_a(s) = \frac{1}{s+1} = \frac{1/s}{1+(1/s)} \quad \leftrightarrow \quad h_a(t) = e^{-t}\mu(t)$$

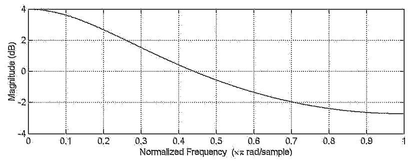
$$g[n] = h_a(nT) = e^{-nT}\mu[n] \quad \leftrightarrow \quad G(z) = \sum_{n=0}^{\infty} e^{-nT} z^{-n} = \frac{1}{1-e^{-T}z^{-1}}$$

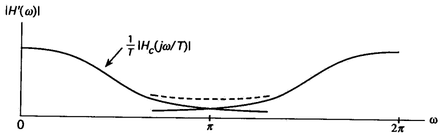
Let T=1: $G(z) = \frac{1}{1-e^{-1}z^{-1}}$

$\Rightarrow$ zero at z=0, pole at $z=1/e$



Example: First Order Butterworth Filter Designed Using the Impulse Invariant Method

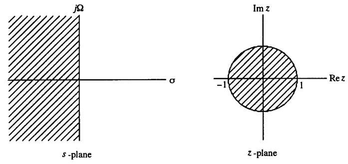




Bilinear Transform Method

To avoid aliasing, the mapping from s-plane to z-plane should be one-to-one, i.e., a single point in the s-plane should be mapped to a unique point in the z-plane and vice versa

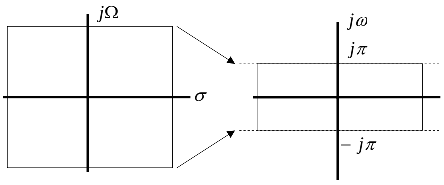
1) The entire $j\Omega$-axis should be mapped onto the unit circle
2) The entire left-half s-plane should be mapped inside the unit circle



Bilinear Transform Method

Derivation of the bilinear transform:

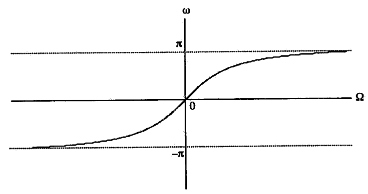
1) One-to-one mapping from s to s' which compresses the entire s-plane into the strip
   $-\pi/T \le \mathrm{Im}(s') \le \pi/T$
2) Employ impulse invariance method to s'-plane with $z=e^{s'T}$



s-plane s'-plane

Bilinear Transform Method

- The mapping: $s' = \frac{2}{T}\tanh^{-1}\left(\frac{sT}{2}\right)$
- Substituting: $s=j\Omega$ and $s'=j\Omega'$ $\quad \Omega' = \frac{2}{T}\tan^{-1}\left(\frac{\Omega T}{2}\right)$



Bilinear Transform Method

- The normalized frequency ω now corresponds to $\Omega' T$

$$\omega = 2\tan^{-1}\left(\frac{\Omega T}{2}\right)$$

- Thus, the entire $j\Omega$-axis is compressed to the interval $(-\omega, \omega)$ for ω in a one-to-one manner
- The mapping is highly nonlinear
- However, for small $\omega=\Omega T$ it is approximately linear

Bilinear Transform Method

- The desired transformation from s to z (via s')

$$s = \frac{2}{T}\tanh\left(\frac{s'T}{2}\right)$$

- Setting $z = e^{s'T}$ or $s' = \frac{1}{T}\ln z$

  yields $s = \frac{2}{T}\tanh\left(\frac{\ln z}{2}\right)$

- Writing $\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1-e^{-2x}}{1+e^{-2x}}$
- Now $x = \frac{\ln z}{2}$ or $2x = \ln z$ $\Rightarrow$ $e^{2x} = z$

Bilinear Transform Method

- The bilinear transform is:

$$s = \frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right) = \frac{2}{T}\left(\frac{z-1}{z+1}\right)$$

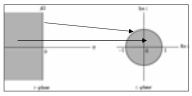
- The s-plane transfer function $H_a(s)$ gives a z-plane transfer function

$$G(z) = H_a(s)\Big|_{s=\frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right)}$$

Bilinear Transform Method

- Solving z and substituting $s=\sigma_0+j\Omega_0$ gives:

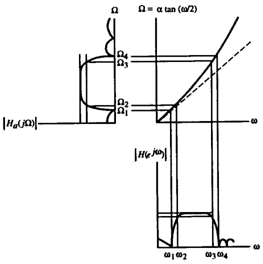
$$z = \frac{1+\frac{2}{T}s}{1-\frac{2}{T}s} = \frac{1+\frac{2}{T}(\sigma_0+j\Omega_0)}{1-\frac{2}{T}(\sigma_0+j\Omega_0)} = \frac{\left(1+\frac{2}{T}\sigma_0\right)+j\frac{2}{T}\Omega_0}{\left(1-\frac{2}{T}\sigma_0\right)-j\frac{2}{T}\Omega_0}$$



1) $j\Omega$-axis, $\mathrm{Re}(s)=0$; this gives $|z|=1$
   The frequency axis from s-plane is mapped onto the unit circle
2) Left-half s-plane, Re(s)<0; $|1+(T/2)s| < |1-(T/2)s|$ or $|z|<1$
   Left-half s-plane is mapped inside the unit circle
3) Right-half s-plane, Re(s)>0; $|1+(T/2)s| > |1-(T/2)s|$ or $|z|>1$
   The right-half s-plane is mapped outside the unit circle

Frequency Warping

- Distortion due to nonlinearity of the mapping:



$$j\Omega = \frac{1-e^{-j\omega}}{1+e^{-j\omega}} = j\tanh\left(\frac{\omega}{2}\right)$$

$$\Omega = \tan\left(\frac{\omega}{2}\right)$$

Frequency Warping

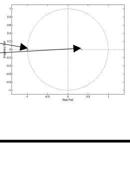
- To design a digital filter meeting the desired (digital) specifications we have to
  1) Prewarp the critical band-edge frequencies (ω_p and ω_s) to analog frequencies (Ω_p and Ω_s)
  2) Design an analog prototype filter $H_a(s)$ using the prewarped critical frequencies
  3) Transform $H_a(s)$ to $G(z)$ using the bilinear transformation

Example: First Order Butterworth Filter Designed by the Bilinear Transformation

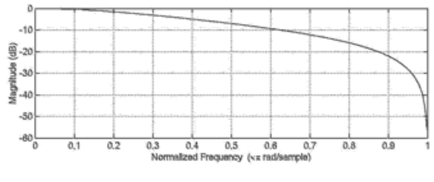
$$H_a(s) = \frac{1}{s+1} \rightarrow H(z) = \frac{1}{s+1}\Bigg|_{s=\frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right)} = \frac{1}{\frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right)+1}$$

$$\Rightarrow \quad H(z) = \frac{1+z^{-1}}{\frac{2}{T}\left(1-z^{-1}\right)+1+z^{-1}} = \frac{1+z^{-1}}{\left(1-\frac{2}{T}\right)z^{-1}+1+\frac{2}{T}}\Bigg|_{T=1} = \frac{1+z^{-1}}{3-z^{-1}}$$

⇨ zero at z=-1; pole at z=1/3



Example: First Order Butterworth Filter Designed by the Bilinear Transformation



- The entire frequency axis from the s-plane is mapped onto the unit circle in the z-plane one-to-one

=> NO ALIASING !

Spectral Transformations of IIR Filters

- Transformation of a given digital IIR lowpass transfer function $G_L(z)$ to another digital transfer function $G_D(z)$
- Prototype lowpass $G_L(z)$; variable z^{-1}
  Transformed filter $G_D(z')$; variable z'^{-1}
- Transformation from z-domain to z'-domain: $z=F(z')$
- Now, $G_L(z)$ is transformed to $G_D(z')$ through

$$G_D(z) = G_L\{F(z')\}$$

- To transform a rational $G_L(z)$ into a rational $G_D(z')$, $F(z')$ must be a rational function in z'
- The inside of the z-plane should be mapped into the inside of z'-plane

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 9 31

Spectral Transformations of IIR Filters

- In order to map a lowpass magnitude response to one of the four basic types of magnitude responses, points on the unit circle in z-plane should be mapped onto the unit circle in z'-plane
- The requirements

$$|F(z')| = \begin{cases} >1, & \text{if } |z|>1 \\ =1, & \text{if } |z|=1 \\ <1, & \text{if } |z|<1 \end{cases}$$

- $F^{-1}(z')$ must be a stable allpass function
- The most general form of $F^{-1}(z')$ with real coefficients is given by

$$F^{-1}(z') = \pm\prod_{l=1}^{L}\left(\frac{1-\alpha_l{}^* z'}{z'-\alpha_l}\right)$$

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 9 32

Lowpass-to-Lowpass Transformation

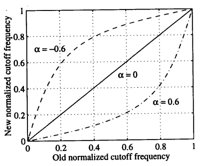
- $G_L(z)$ with cutoff frequency ω_c is transformed to another lowpass filter $G_D(z')$ with ω'_c

$$z^{-1} = F^{-1}(z') = \frac{1-\alpha z'}{z'-\alpha} = \frac{z'^{-1}-\alpha}{1-\alpha z'^{-1}}$$

with α real $\quad e^{-j\omega} = \dfrac{e^{-j\omega'}-\alpha_l}{1-\alpha_l e^{-j\omega'}}$

$$\tan\left(\frac{\omega}{2}\right) = \frac{1+\alpha}{1-\alpha}\tan\left(\frac{\omega'}{2}\right)$$

$$\alpha = \frac{\sin\left(\dfrac{\omega_c-\omega_c'}{2}\right)}{\sin\left(\dfrac{\omega_c+\omega_c'}{2}\right)}$$



© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 9 33

Spectral Transformations of IIR Filters

Filter type	Spectral transformation	Design parameters
Lowpass	$z^{-1} = \frac{\hat{z}^{-1}-\alpha}{1-\alpha\hat{z}^{-1}}$	$\alpha = \frac{\sin\left(\frac{\omega_c-\hat{\omega}_c}{2}\right)}{\sin\left(\frac{\omega_c+\hat{\omega}_c}{2}\right)}$ $\hat{\omega}_c$ = desired cutoff frequency
Highpass	$z^{-1} = -\frac{\hat{z}^{-1}+\alpha}{1+\alpha\hat{z}^{-1}}$	$\alpha = -\frac{\cos\left(\frac{\omega_c+\hat{\omega}_c}{2}\right)}{\cos\left(\frac{\omega_c-\hat{\omega}_c}{2}\right)}$ $\hat{\omega}_c$ = desired cutoff frequency
Bandpass	$z^{-1} = -\frac{\hat{z}^{-2}-\frac{2\alpha\beta}{\beta+1}\hat{z}^{-1}+\frac{\beta-1}{\beta+1}}{\frac{\beta-1}{\beta+1}\hat{z}^{-2}-\frac{2\alpha\beta}{\beta+1}\hat{z}^{-1}+1}$	$\alpha = \frac{\cos\left(\frac{\hat{\omega}_{c2}+\hat{\omega}_{c1}}{2}\right)}{\cos\left(\frac{\hat{\omega}_{c2}-\hat{\omega}_{c1}}{2}\right)}$ $\beta = \cot\left(\frac{\hat{\omega}_{c2}-\hat{\omega}_{c1}}{2}\right)\tan\left(\frac{\omega_c}{2}\right)$ $\hat{\omega}_{c2}, \hat{\omega}_{c1}$ = desired upper and lower cutoff frequencies
Bandstop	$z^{-1} = \frac{\hat{z}^{-2}-\frac{2\alpha}{1+\beta}\hat{z}^{-1}+\frac{1-\beta}{1+\beta}}{\frac{1-\beta}{1+\beta}\hat{z}^{-2}-\frac{2\alpha}{1+\beta}\hat{z}^{-1}+1}$	$\alpha = \frac{\cos\left(\frac{\hat{\omega}_{c2}+\hat{\omega}_{c1}}{2}\right)}{\cos\left(\frac{\hat{\omega}_{c2}-\hat{\omega}_{c1}}{2}\right)}$ $\beta = \tan\left(\frac{\hat{\omega}_{c2}-\hat{\omega}_{c1}}{2}\right)\tan\left(\frac{\omega_c}{2}\right)$ $\hat{\omega}_{c2}, \hat{\omega}_{c1}$ = desired upper and lower cutoff frequencies

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 9 34

Computer-Aided Design of IIR Digital Filters

- Transformation of an analog prototype filter into a digital filter is used in application with frequency-selective magnitude response, i.e., lowpass, highpass, bandpass, and bandstop characteristics
- In applications requiring IIR digital filters with other types of frequency responses, filter design algorithms rely on some type of iterative optimization techniques that are used to minimize the error between the desired frequency response and the computer-generated filter

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 9 35

Computer-Aided Design of IIR Digital Filters

- Let $H(e^{j\omega})$ denote the frequency response of the digital transfer function $H(z)$ to be designed so that it approximates the desired frequency response $D(e^{j\omega})$, given by a piecewise linear function of ω
- Determine iteratively the coefficients of $H(z)$ so that the difference between $H(e^{j\omega})$ and $D(e^{j\omega})$ for all values of ω over closed subintervals of $0\leq\omega\leq\pi$ is minimized
- This difference usually specified as a weighted error function

$$\mathrm{E}(\omega) = W(e^{j\omega})[H(e^{j\omega})-D(e^{j\omega})]$$

© 2007 Olli Simula T-61.3010 Digital Signal Processing; Mitra 3rd Edition: Chapter 9 36

Computer-Aided Design of IIR Digital Filters

- **Chebyshev** or **minimax criterion** - Minimizes the peak absolute value of the weighted error:

$$\varepsilon = \max_{\omega \in R} |E(\omega)|$$

where R is the set of disjoint frequency bands in the range $0 \le \omega \le \pi$, on which the desired frequency response, $D(e^{j\omega})$, is defined

- R is composed of the passbands and stopbands of the filter to be designed

Computer-Aided Design of Digital Filters

- **Least-*p* Criterion** - Minimize

$$\varepsilon = \int_{\omega \in R} \left| W(e^{j\omega})[H(e^{j\omega}) - D(e^{j\omega})] \right|^p d\omega$$

over the specified frequency range R with p a positive integer

- $p = 2$ yields the ***least-squares criterion***
- As $p \to \infty$, the least p-th solution approaches the minimax solution

Computer-Aided Design of Digital Filters

- **Least-*p* Criterion -** In practice, the p-th power error measure is approximated as

$$\varepsilon = \sum_{i=1}^{K} \{W(e^{j\omega_i})[H(e^{j\omega_i}) - D(e^{j\omega_i})]\}^p$$

where ω_i, $1 \le i \le K$, is a suitably chosen dense grid of digital angular frequencies

- For linear-phase FIR filter design, $H(e^{j\omega})$ and are zero-phase frequency responses
- For IIR filter design, $H(e^{j\omega})$ and $D(e^{j\omega})$ are magnitude functions

Summary

- IIR digital filter design is usually carried out by transforming a prototype analog transfer function from the s-domain into the z-domain
- The widely used bilinear transform method is based on this approach
- The design programs, e.g., Matlab Signal Processing Toolbox, produce filters with Butterworth, Chebyshev, and elliptic magnitude responses
- Computer-aided iterative techniques are capable of designing filters with more general magnitude functions and responses