

IIR Filter Design

- 1) Based on Analog Prototype
 - a) Impulse invariant design
 - b) Bilinear transformation (✓) ~ widely used
- 2) Computer-Aided Optimization

Designs in category 1) proceed by first designing an analog filter having a frequency response with the desired shape, and then “transforming” it to a digital filter. To use these design methods, we must first learn just a bit about analog filter design.

Elements of Analog Filter Design

Notation:

$$H_L(s) = \int_{-\infty}^{\infty} h_a(t) e^{-st} dt \quad (\text{Laplace transform})$$

$$\Rightarrow H_a(\Omega) = H_L(j\Omega) \quad (\text{Fourier transform})$$

Consider only lowpass Butterworth, Chebyshev, and elliptic (Cauer) filters.

For each of these types of filters, $H_L(s)$ is found indirectly from a specified $|H_a(\Omega)|^2$.

We need $H_L(s)$ because, later, this is what will be transformed into $H(z)$.

For Butterworth, Chebyshev, and elliptic filters, $|H_a(\Omega)|^2$ has the form:

$$|H_a(\Omega)|^2 = M(\Omega^2) = \frac{1}{1 + F(\Omega^2)}$$

↑
rational with
real coeffs.

12.2

How do we get $H_L(s)$ from $|H_a(\Omega)|^2$? That is, how do we find $H_L(s)$ satisfying

$$|H_L(j\Omega)|^2 = |H_a(\Omega)|^2 = M(\Omega^2) \quad (*)$$

Answer:

- 1) First find poles and zeros of $M(-s^2)$ where s is a complex variable. Since M has real coefficients and is a function of s^2 , the poles and zeros will have symmetry around both the real and imaginary axes.
- 2) Take $H_L(s)$ to be the left-half-plane pole factors (for stability) and left-half-plane zero factors (for smallest delay, called “minimum phase”).

But, does this work?

Need to show

$$|H_L(j\Omega)|^2 = M(\Omega^2) \quad (1)$$

Have:

$$H_L(s) H_L(-s) = M(-s^2)$$

which implies

$$H_L(j\Omega) H_L(-j\Omega) = M(\Omega^2)$$

So, (1) will be true if

$$H_L(-j\Omega) = H_L^*(j\Omega) \quad (2)$$

This follows, though, because the poles and zeros of $H_L(s)$ are symmetric around the real axis, and therefore occur in complex-conjugate pairs. For any pole pair or zero pair $(s-p)$ $(s-p^*)$ in $H_L(s)$, we have

$$(s-p)(s-p^*) \Big|_{s=-j\Omega} = (-j\Omega-p)(-j\Omega-p^*) = [(j\Omega-p^*)(j\Omega-p)]^* = \left[(s-p)(s-p^*) \Big|_{s=j\Omega} \right]^*$$

which proves (2), and therefore (1).

Example

$$|H_a(\Omega)|^2 = M(\Omega^2) = \frac{1}{1 + \Omega^2}$$

Find $H_L(s)$.

We have

$$\begin{aligned} M(-s^2) &= \frac{1}{1 - s^2} \\ &= \frac{1}{\underbrace{(1 - s)(1 + s)}_{\text{LHP factor}}} \end{aligned}$$

$$\Rightarrow \left[H_L(s) = \frac{1}{s + 1} \right]$$

Let's check to see if $|H_a(\Omega)|^2 =$ above $M(\Omega^2)$:

$$\begin{aligned} |H_a(\Omega)|^2 &= |H_L(j\Omega)|^2 \\ &= \left| \frac{1}{j\Omega + 1} \right|^2 = \frac{1}{|j\Omega + 1|^2} \\ &= \frac{1}{\Omega^2 + 1} = M(\Omega^2) \quad \checkmark \end{aligned}$$

Example

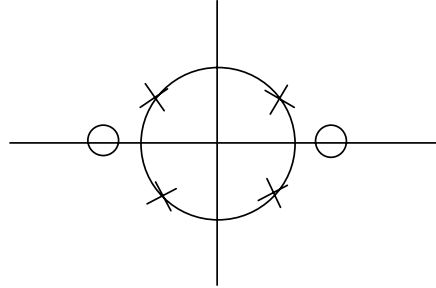
Suppose $M(\Omega^2) = \frac{2 + \Omega^2}{1 + \Omega^4}$ (not for a B, C, or E filter!)

Then

$$\begin{aligned} M(-s^2) &= \frac{2 - s^2}{1 + s^4} \\ &= \frac{(\sqrt{2} - s)(\sqrt{2} + s)}{\underset{\uparrow}{(s + \gamma)}(s + \gamma^*)(s - \gamma)(s - \gamma^*)} \end{aligned}$$

$$\gamma = e^{j\frac{\pi}{4}} = \frac{1+j}{\sqrt{2}}$$

Pole-zero diagram:



$$\Rightarrow \text{Take } H_L(s) = \frac{\sqrt{2} + s}{(s + \gamma)(s + \gamma^*)} \quad (\text{LHP factors})$$

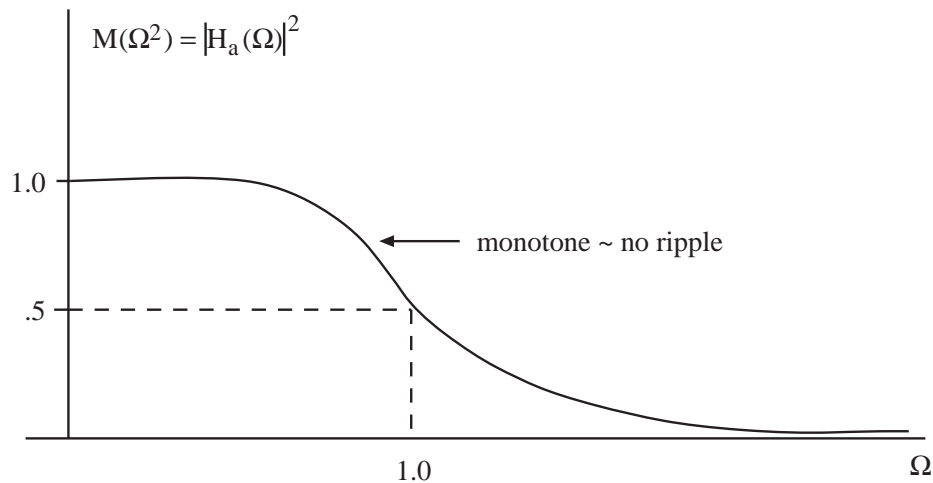
Can check that $|H_L(j\Omega)|^2 = M(\Omega^2)$.

$M(\Omega^2)$ for B, C, and E Filters

Butterworth

$$M(\Omega^2) = \frac{1}{1 + F(\Omega^2)} \text{ with } F(\Omega^2) = \Omega^{2n} \text{ for } n\text{-th-order filter}$$

Result:



$$M(\Omega^2) = .5 \text{ at } \Omega = \Omega_c = 1.0. \text{ For a general cutoff frequency } \Omega_c, \text{ use } F(\Omega^2) = \left(\frac{\Omega}{\Omega_c}\right)^{2n} .$$

Here, we are defining the cutoff frequency to be the value of Ω where $|H_a(\Omega)|^2$ is reduced to one-half its maximum height, or correspondingly, $|H_a(\Omega)|$ reaches $1/\sqrt{2}$ times its maximum value. This definition of cutoff frequency is common, particularly for smooth frequency responses that contain little or no ripple.

Optimality: This $M(\Omega^2)$ has maximum # of derivatives = 0 at the origin for its order. Thus, the response is very flat across lower frequencies.

Can show poles of $M(-s^2) = H_L(s) H_L(-s)$ are equally spaced on the unit circle. This fact helps in factoring $M(-s^2)$.

Chebyshev

$$F(\Omega^2) = \varepsilon^2 C_n^2(\Omega)$$

where ε is a real constant chosen by the designer and $C_n(\bullet)$ is the n th-order Chebyshev polynomial:

$$C_n(\Omega) = \begin{cases} \cos(n \cos^{-1}(\Omega)) & |\Omega| \leq 1 \\ \cosh(n \cosh^{-1}(\Omega)) & |\Omega| > 1 \end{cases}$$

with

$$\cosh t = \frac{e^t + e^{-t}}{2}$$

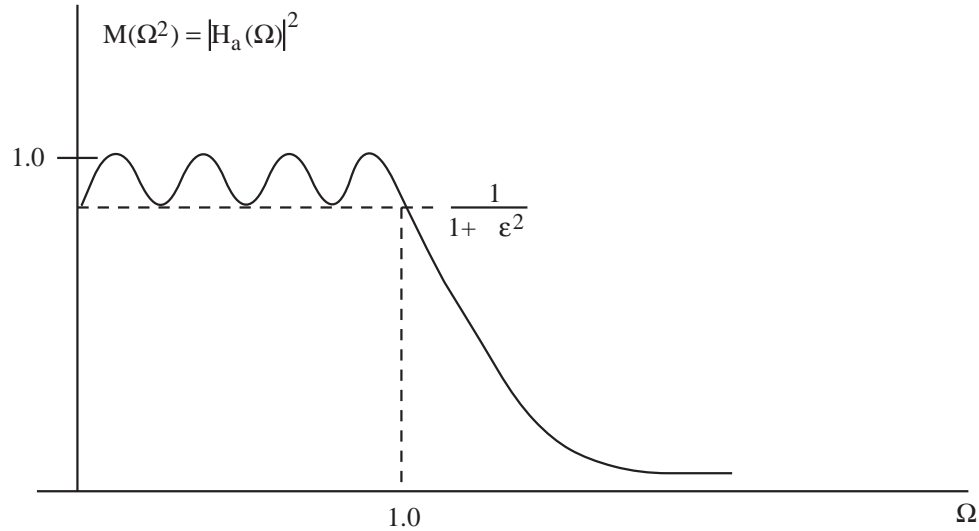
Can show:

$$C_0(\Omega) = 1, C_1(\Omega) = \Omega, C_2(\Omega) = 2\Omega^2 - 1,$$

and that there is a recursion relation:

$$C_{n+1}(\Omega) = 2\Omega C_n(\Omega) - C_{n-1}(\Omega)$$

Result:



For this type of filter, the cutoff frequency is defined to be the value of Ω where $|H_a(\Omega)|^2$ first drops below $1/(1 + \epsilon^2)$ or, correspondingly, $|H_a(\Omega)|$ first drops below $1/\sqrt{1 + \epsilon^2}$.

This is a “Type 1” Chebyshev filter. Its response is equiripple in the passband and monotone decreasing in the stopband. It has a narrower transition band than a Butterworth filter.

Tradeoff: Smaller ϵ gives smaller passband ripple but a wider transition band.

Poles of $M(-s^2) = H_L(s) H_L(-s)$ lie on an ellipse.

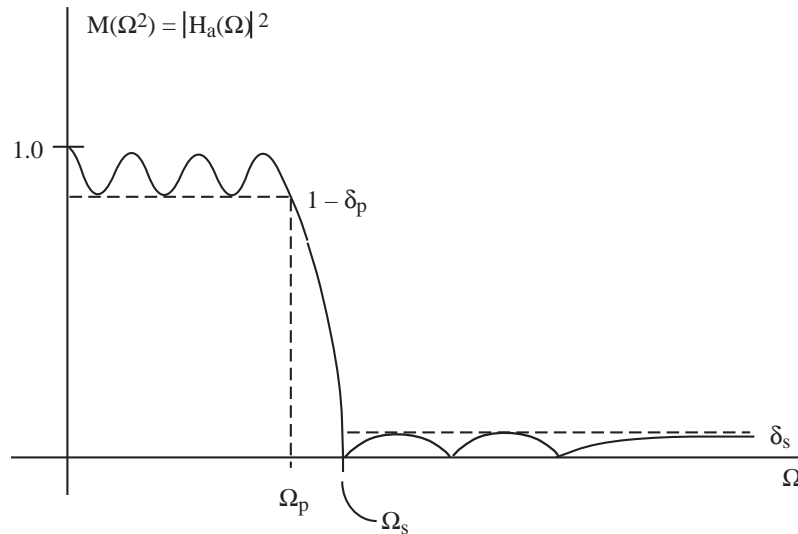
There is also a “Type 2” Chebyshev filter that has a monotone response in its passband and equiripple behavior in its stopband.

Elliptic

$$F(\Omega^2) = \epsilon^2 J_n^2(\Omega) \text{ where } J_n \text{ is the Jacobi elliptic function.}$$

The defining formulas for J_n are so cumbersome that they are not presented here.

Result:



The response is equiripple in both the passband and stopband.

Elliptic filters are optimal in the sense that for a given n , δ_p , δ_s , Ω_p , the transition bandwidth $\Omega_s - \Omega_p$ is the smallest possible.

$\angle H_a(\Omega)$ for B, C, and E filters is reasonably linear till you get near the edge of the passband, where it can be quite nonlinear.

The phase response is closest to linear for B, then C. Elliptic is worst.

All-pass filters are sometimes cascaded onto elliptic filters to compensate for the nonlinear phase of elliptic filters.

All-pass filters have $|H_a(\Omega)| = \text{constant}$ and the coefficients are chosen to shape $\angle H_a(\Omega)$ in a desired way.

Bilinear Transformation

Start with analog prototype $H_L(s)$.

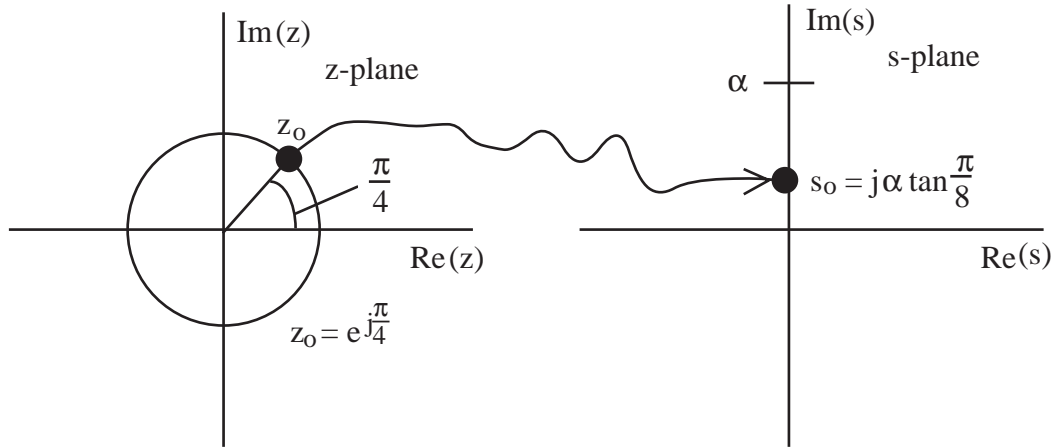
Take

$$H(z) = H_L(s) \Big|_{s=\alpha \frac{1-z^{-1}}{1+z^{-1}}}$$

α is a real positive constant that we will be able to choose to control the cutoff frequency of the digital filter.

$s = \alpha \frac{1 - z^{-1}}{1 + z^{-1}}$ is a bilinear transformation (BLT) of the z-plane to the s-plane. For

example, the point z_0 maps to the point s_0 as shown below.



To see this, note:

$$s_0 = \alpha \frac{1 - z_0^{-1}}{1 + z_0^{-1}} = \alpha \frac{1 - e^{-j\pi/4}}{1 + e^{-j\pi/4}} = \alpha \frac{e^{-j\pi/8} - e^{j\pi/8}}{e^{-j\pi/8} + e^{j\pi/8}}$$

$$= \alpha \frac{2j \sin \frac{\pi}{8}}{2 \cos \frac{\pi}{8}} = j \alpha \tan \frac{\pi}{8}$$

So, if we design a digital filter using the BLT, then $H(z_0) = H_d\left(\frac{\pi}{4}\right)$ will have the same value as $H_L(s_0) = H_a\left(\alpha \tan \frac{\pi}{8}\right)$.

We must have a much broader understanding than this, however. Questions:

- i) Stable $H_L(s) \Rightarrow$ stable $H(z)$?
- ii) How is $H_d(\omega) = H(e^{j\omega})$ related to $H_L(s)$?

Answer i) by considering a point $s = s_0$ and determining what z it gets mapped to.

$$\text{BLT mapping is } s = \alpha \frac{1 - z^{-1}}{1 + z^{-1}}$$

$$\Rightarrow s(1 + z^{-1}) = \alpha(1 - z^{-1})$$

$$\Rightarrow z^{-1}(s + \alpha) = \alpha - s$$

$$\Rightarrow z = \frac{\alpha + s}{\alpha - s}$$

So, a point $s_0 = \sigma_0 + j \Omega_0$.

gets mapped to:

$$z_0 = \frac{\alpha + \sigma_0 + j\Omega_0}{\alpha - \sigma_0 - j\Omega_0} \quad (*)$$

i.e., $H(z_0) = H_L(s_0)$.

From (*)

$$|z_0| = \left[\frac{(\alpha + \sigma_0)^2 + \Omega_0^2}{(\alpha - \sigma_0)^2 + \Omega_0^2} \right]^{1/2}$$

So:

$$|z_0| \begin{cases} < 1 & \sigma_0 < 0 \\ = 1 & \sigma_0 = 0 \\ > 1 & \sigma_0 > 0 \end{cases}$$

giving:

- Left-half s-plane is mapped inside the unit circle in the z-plane.
- Right-half s-plane is mapped outside the unit circle in the z-plane.
- Imaginary axis in s-plane is mapped onto the unit circle in the z-plane.

Note:

[a] \Rightarrow stable $H_L(s)$ results in stable $H(z)$

[c] $\Rightarrow H_d(\omega) = H(e^{j\omega})$ depends only on $H_L(j\Omega) = H_a(\Omega)$

What is the relationship between $H_d(\omega)$ and $H_a(\Omega)$? $H_d(\omega)$ is given by

12.10

$$H_d(\omega) = H(e^{j\omega}) = H_L(s) \Big|_{s=\alpha \frac{1-e^{-j\omega}}{1+e^{-j\omega}}}$$

Note:

$$\begin{aligned} \alpha \frac{1-e^{-j\omega}}{1+e^{-j\omega}} &= \alpha \frac{e^{-j\omega/2}}{e^{-j\omega/2}} \frac{e^{j\omega/2} - e^{-j\omega/2}}{e^{j\omega/2} + e^{-j\omega/2}} \\ &= \alpha \frac{2j \sin \frac{\omega}{2}}{2 \cos \frac{\omega}{2}} = j \alpha \tan \frac{\omega}{2} \end{aligned}$$

So,

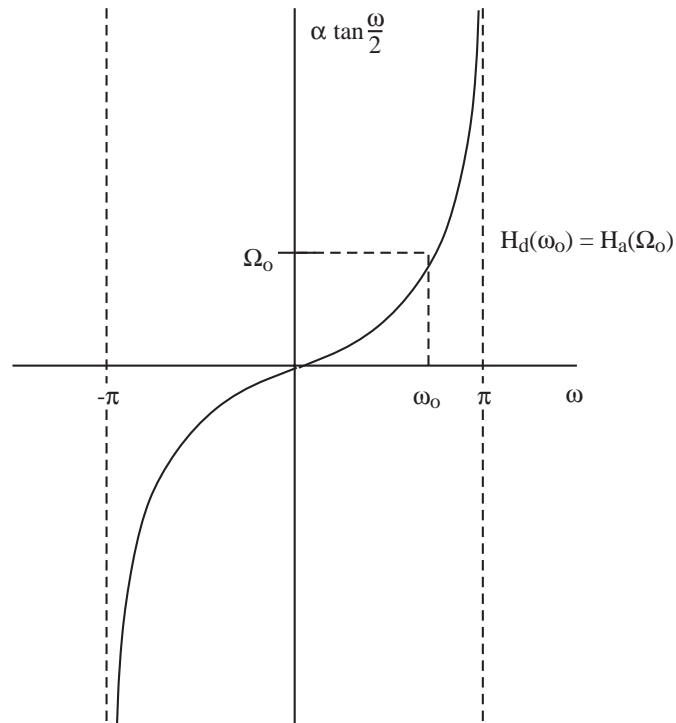
$$H_d(\omega) = H_L(s) \Big|_{s=j\alpha \tan \frac{\omega}{2}}$$

Since $H_a(\Omega) = H_L(j\Omega)$ we have

$$\boxed{H_d(\omega) = H_a\left(\alpha \tan \frac{\omega}{2}\right)} \quad (\square)$$

This equation tells us exactly how, when using the bilinear transformation design method, the frequency response of the designed digital filter will depend on the frequency response of the analog prototype.

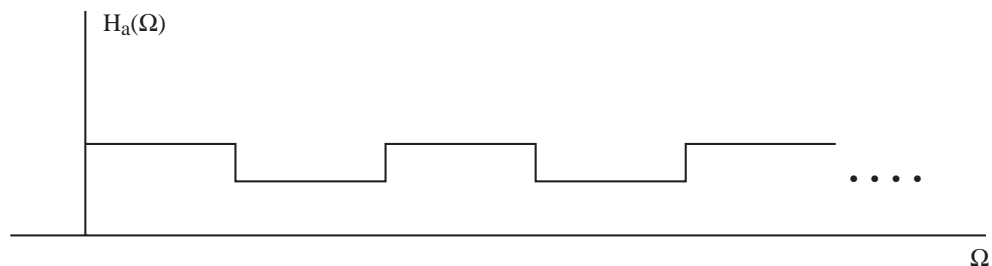
Pictorial interpretation of (□):



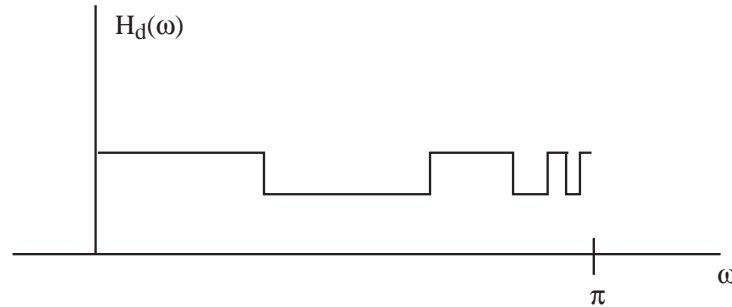
So, $H_d(\omega)$ takes on exactly the same set of values as $H_a(\Omega)$, but there is a squashing of the analog frequency axis, according to the above curve. This mapping is nonlinear, and it has to be, since the infinite-length analog frequency axis $-\infty < \Omega < \infty$ is mapped onto the finite-length interval $-\pi < \omega < \pi$. Because of this, $H_d(\omega)$ won't look like $H_a(\Omega)$ in general. It will be a warped version. This effect is sometimes called “frequency warping.”

Example

Applying the BLT to an analog prototype having frequency response:



results in:



In general, to design $H_d(\omega)$ having a desired shape, we would need to design $H_a(\Omega)$ so that after application of the BLT, the frequency warping produces the desired $H_d(\omega)$. Thus, we would need to “prewarp.”

Let the desired $H_d(\omega)$ be $D(\omega)$.

$$(\square) \Rightarrow \text{want } H_a\left(\alpha \tan \frac{\omega}{2}\right) = D(\omega)$$

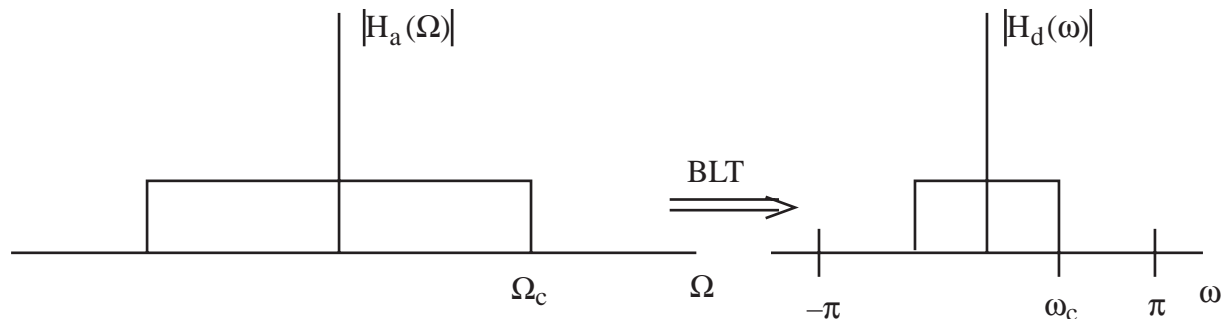
$$\Rightarrow H_a(\Omega) = D\left(2 \tan^{-1} \frac{\Omega}{\alpha}\right) \quad (\Delta)$$

If we *could* design $H_L(s)$ to satisfy (Δ) , the BLT, would then give

$$H_d(\omega) = D(\omega) .$$

This is problematic, however. The design of H_L with a general shape will require computer optimization. Thus, we may as well design H_d directly, using computer optimization.

Fortunately, for LPF's, BPF's and HPF's, there is no such problem. For these kinds of filters, frequency warping just affects the cutoff frequencies, e.g.,



From (\square) :

$$\Omega_c = \alpha \tan \frac{\omega_c}{2} \quad (\square\square)$$

So, we can simply pick Ω_c to give the desired ω_c . Equivalently, if $H_a(\Omega)$ is normalized to $\Omega_c = 1$, we can choose α in the BLT to give the desired ω_c .

This suggests two alternative, but equivalent, design procedures:

1. a) Choose α arbitrarily, say $\alpha = 1$.
 b) Then design the analog prototype to have cutoff Ω_c given by (□□).
 c) Apply the BLT.
2. a) Use an analog prototype with Ω_c .
 b) Choose α to give the desired ω_c .

From (□□),

$$\alpha = \Omega_c \cot\left(\frac{\omega_c}{2}\right)$$

so that the BLT method becomes

$$H(z) = H_L(s) \left|_{s = \Omega_c \cot \frac{\omega_c}{2} \frac{1-z^{-1}}{1+z^{-1}}}$$

Analog prototype filters are usually designed with normalized cutoff frequencies $\Omega_c = 1.0$. In this case, the bilinear transformation method of design reduces to

$$\boxed{H(z) = H_L(s) \left|_{s = \cot \frac{\omega_c}{2} \frac{1-z^{-1}}{1+z^{-1}}}} \quad \text{(BLT)}$$

In this course, we will use the second option, with Eq. (BLT), to perform the bilinear transformation method of design.

Example

12.14

Design a first-order Butterworth digital filter with $\omega_c = \frac{\pi}{4}$, using the BLT method

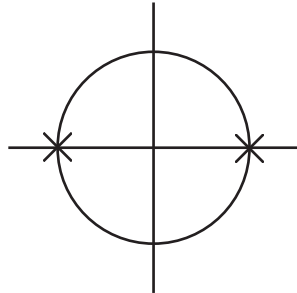
Solution:

First, find the analog prototype:

$$|H_a(\Omega)|^2 = B(\Omega^2) = \frac{1}{1 + \Omega^2}$$

$$B(-s^2) = \frac{1}{1 - s^2} = \frac{1}{1 - s} \cdot \frac{1}{1 + s}$$

Pole locations:



Take LHP factor for $H_L(s)$:

$$H_L(s) = \frac{1}{s + 1}$$

Apply (BLT):

$$H(z) = \frac{1}{s + 1} \Bigg|_{s = \cot \frac{\pi}{8} \frac{1 - z^{-1}}{1 + z^{-1}}}$$

\uparrow
 2.4142

$$\begin{aligned} \Rightarrow H(z) &= \frac{1}{2.4142 \frac{1 - z^{-1}}{1 + z^{-1}} + 1} \\ &= \frac{1 + z^{-1}}{2.4142 (1 - z^{-1}) + 1 + z^{-1}} \\ &= \frac{1 + z^{-1}}{3.4142 - 1.4142 z^{-1}} \end{aligned}$$

$$= \boxed{\frac{.2929 + .2929 z^{-1}}{1 - .4142 z^{-1}}}$$

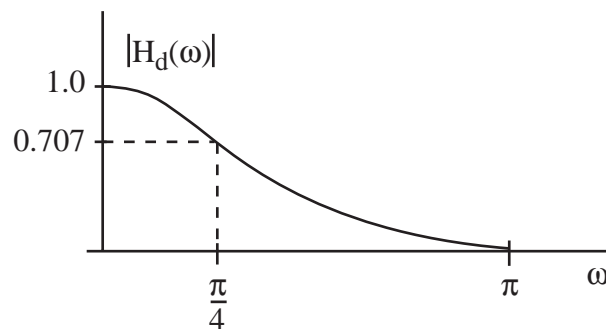
Now, since $H(z)$ resulted from a BLT design using an analog prototype filter with $|H_a(0)| = 1$ and $|H_d(\infty)| = 0$, we know

$$|H_d(0)| = |H_a(0)| = 1 \quad \text{and} \quad |H_a(\pi)| = |H_a(\infty)| = 0$$

Furthermore, since the digital cutoff is $\frac{\pi}{4}$, we should have

$$\left|H_d\left(\frac{\pi}{4}\right)\right| = \frac{1}{\sqrt{2}}$$

Thus, $|H_d(\omega)|$ should look like



The correctness of the values of $|H_d(\omega)|$ at $\omega = 0, \pi$ can be easily verified from the transfer function $H(z)$. Note that

$$H_d(0) = H(e^{j0}) = H(1) = \frac{.2929 + .2929}{1 - .4142} = 1.0$$

$$H_d(\pi) = H(e^{j\pi}) = H(-1) = \frac{.2929 - .2929}{1 - .4142} = 0$$

For $\omega \neq 0$ or π we can always evaluate the magnitude response via the lengthier calculation

$$\begin{aligned} |H_d(\omega)| &= \frac{|.2929 + .2929 e^{-j\omega}|}{|1 - .4142 e^{-j\omega}|} \\ &= \frac{\sqrt{(.2929 + .2929 \cos \omega)^2 + (.2929 \sin \omega)^2}}{\sqrt{(1 - .4142 \cos \omega)^2 + (.4142 \sin \omega)^2}} \end{aligned}$$

Finally, a plot of $\angle H_d(\omega)$ would show that the phase is nearly linear for $|\omega| < \frac{\pi}{4}$ and becomes more nonlinear for $|\omega| \approx \frac{\pi}{4}$ and larger.

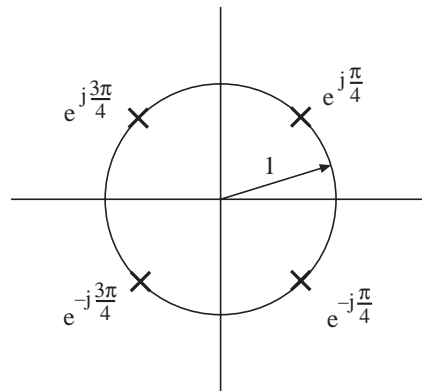
Example

Similar to previous example, but now design a second-order Butterworth digital filter with $\omega_c = \frac{\pi}{4}$.

To find analog prototype, note

$$|H_a(\Omega)|^2 = M(\Omega^2) = \frac{1}{1 + \Omega^4}$$

$$\Rightarrow B(-s^2) = \frac{1}{1 + s^4}$$



Left-half-plane poles are used for $H_L(s)$:

$$H_L(s) = \frac{1}{\left(s - e^{j\frac{3\pi}{4}}\right)\left(s - e^{-j\frac{3\pi}{4}}\right)} = \frac{1}{s^2 + \sqrt{2}s + 1}$$

$$\begin{aligned} H(z) &= H_L(s) \Big|_{s = \cot \frac{\pi}{8} \frac{1-z^{-1}}{1+z^{-1}}} \\ &= \frac{1}{(2.4142)^2 \frac{(1-z^{-1})^2}{(1+z^{-1})^2} + 2.4142 \sqrt{2} \frac{1-z^{-1}}{1+z^{-1}} + 1} \\ &= \frac{(1+z^{-1})^2}{(2.4142)^2 (1-2z^{-1}+z^{-2}) + 2.4142\sqrt{2}(1-z^{-2}) + (1+2z^{-1}+z^{-2})} \end{aligned}$$

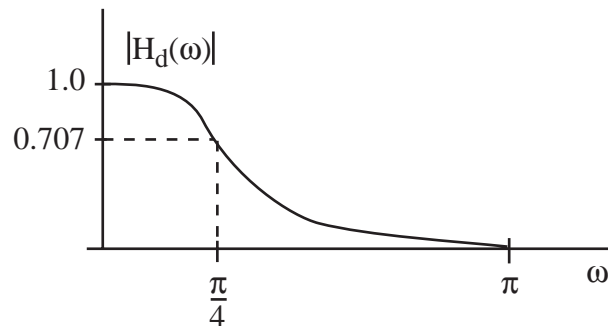
$$= \frac{1 + 2z^{-1} + z^{-2}}{10.23 - 9.66z^{-1} + 3.41z^{-2}}$$

$$= \frac{0.098 + 0.196z^{-1} + 0.098z^{-2}}{1 - 0.944z^{-1} + 0.333z^{-2}}$$

Now, once again, from the shape of the analog prototype $H_a(\Omega)$ we know

$$H_d(0) = 1, \quad \left| H_d\left(\frac{\pi}{4}\right) \right| = \frac{1}{\sqrt{2}}, \quad |H_d(\pi)| = 0$$

However, for the second-order filter, the frequency response makes a sharper transition around $\omega_c = \frac{\pi}{4}$ and looks like

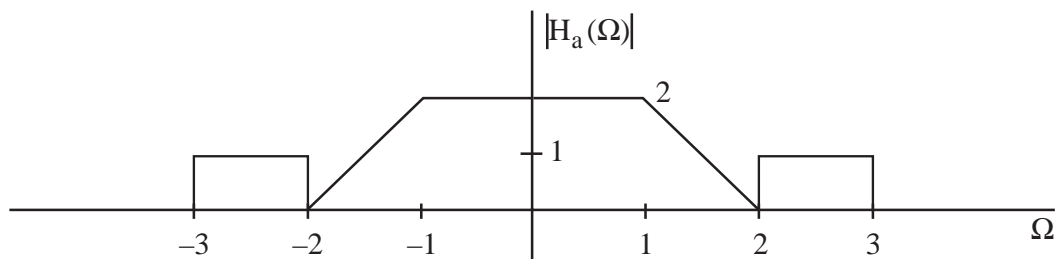


The phase $\angle H_d(\omega)$ will again be quite linear across the passband and more nonlinear across the stopband.

Let's try to get a better feel for the BLT mapping by considering one further example.

Example

Suppose that an analog prototype filter $H_L(s)$ has the frequency response



and that the bilinear transformation is used to produce a digital filter with $H(z) = H_L(s) \Big|_{s=\frac{1-z^{-1}}{1+z^{-1}}}$. Sketch $|H_d(\omega)|$. Label all critical frequencies and amplitudes.

Solution

$H_d(\omega)$ is a squashed version of $H_a(\Omega)$, described by

$$H_d(\omega) = H_a\left(\alpha \tan \frac{\omega}{2}\right) = H_a\left(\tan \frac{\omega}{2}\right)$$

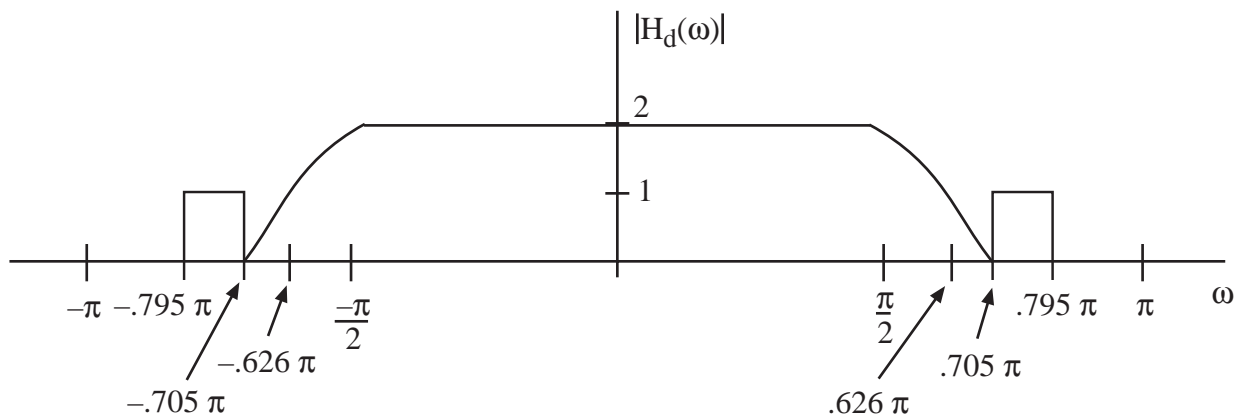
Thus, the value of $H_d(\omega)$ equals the value of $H_a(\Omega)$ at $\Omega = \tan \frac{\omega}{2}$. So a feature (e.g., jump) in $H_a(\Omega)$ that occurs at $\Omega = \Omega_0$ will occur in $H_d(\omega)$ at

$$\omega_0 = 2 \tan^{-1} \Omega_0$$

The interesting features in $H_a(\Omega)$ occur at $\Omega_0 = 1, 2,$ and 3 . In addition let's also consider $\Omega_0 = 1.5$. The corresponding ω_0 are given in the table below.

Ω_0	1	1.5	2	3
ω_0	0.5π	0.626π	0.705π	0.795π

Thus, $|H_d(\omega)|$ looks like



Notice that $H_d(\omega)$ takes on exactly the same set of values taken on by $H_a(\Omega)$. It just does so at different frequencies, which causes a change in shape.

2. Computer-Aided Optimization

Used for IIR designs with general magnitude and/or phase specifications (i.e., $H_d(\omega)$ not LPF, HPF, or BPF).

Digital Frequency Transformation

By making a “change of variable” in $H_L(s)$ or $H(z)$, we can transform an analog or digital LPF to a LPF having a different cutoff frequency, or to a HPF or BPF.

We will consider only digital transformations (transformations of $H(z)$).

Procedure:

Let $H(z)$ be the transfer function of a low-pass digital filter. Then simply substitute for z^{-1} in $H(z)$, using the expressions below, to produce filters having the described characteristics.

Lowpass \rightarrow Lowpass

$$z^{-1} \rightarrow \frac{z^{-1} - \beta}{1 - \beta z^{-1}}$$

where

ω'_c = cutoff of new filter

$$\beta = \frac{\sin [(\omega_c - \omega'_c)/2]}{\sin [(\omega_c + \omega'_c)/2]}$$

Lowpass \rightarrow Highpass

$$z^{-1} \rightarrow -\frac{z^{-1} - \beta}{1 - \beta z^{-1}}$$

ω'_c = cutoff of new filter

$$\beta = \frac{\cos [(\omega_c + \omega'_c)/2]}{\cos [(\omega_c - \omega'_c)/2]}$$

Lowpass \rightarrow Bandpass

$$z^{-1} \rightarrow -\frac{z^{-2} - \beta_1 z^{-1} + \beta_2}{\beta_2 z^{-2} - \beta_1 z^{-1} + 1}$$

ω_ℓ = lower cutoff of BPF

ω_u = upper cutoff of BPF

$$\beta_1 = 2\gamma K/(K+1)$$

$$\beta_2 = (K-1)/(K+1)$$

$$\gamma = \frac{\cos [(\omega_u + \omega_\ell)/2]}{\cos [(\omega_u - \omega_\ell)/2]}$$

$$K = \cot \frac{\omega_u - \omega_\ell}{2} \tan \frac{\omega_c}{2}$$

Relationship Between Pole and Zero Locations and Frequency Response

In digital filter design we choose the coefficients of

$$H(z) = \frac{a_0 + a_1 z^{-1} + \dots + a_N z^{-N}}{1 + b_1 z^{-1} + \dots + b_N z^{-N}}$$

to shape the frequency response $H_d(\omega) = H(e^{j\omega})$ in a desired way. Since $H(z)$ can also be written in terms of its poles and zeros as

$$H(z) = a_0 \prod_{i=1}^N \frac{z - z_i}{z - p_i}$$

this provides an alternative parameterization of $H(z)$. Choosing the pole and zero locations of the filter is basically equivalent to choosing the $\{a_i\}$ and $\{b_i\}$. In this connection, it is worth exploring how the pole and zero locations affect the shape of $H_d(\omega)$.

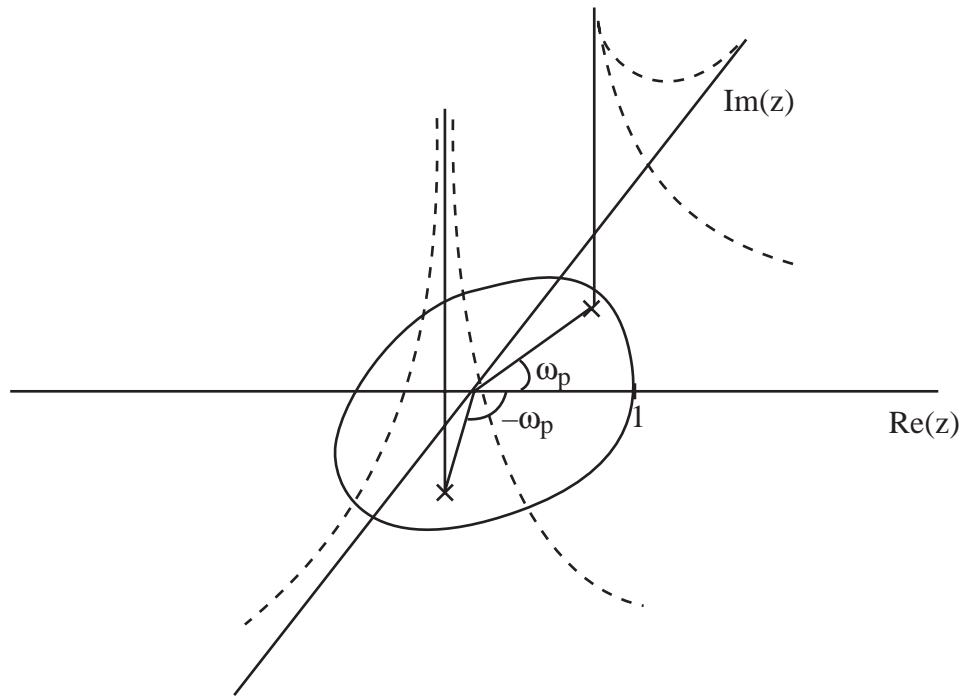
The general shape of $|H_d(\omega)|$ often can be visualized from knowledge of the pole and zero locations of $H(z)$. This is especially true for situations where poles are near the unit circle and zeros are either on or near the unit circle.

Example

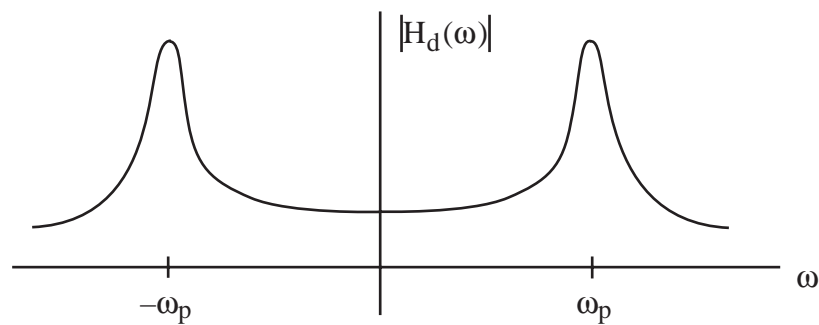
Consider a causal, stable all-pole filter with

$$H(z) = \frac{1}{z^2 - (2\alpha \cos \omega_p)z + \alpha^2} = \frac{1}{(z - \alpha e^{j\omega_p})(z - \alpha e^{-j\omega_p})}$$

and $0 < \alpha < 1$. $H(z)$ approaches ∞ at the pole locations and approaches zero for $|z|$ large. Thus, we expect $|H(z)|$ to look somewhat like a two-pole circus tent:



Here, the pole locations are marked by \times and are at a distance α from the origin. Since $H(z)$ is causal, its ROC is $\{z : |z| > \alpha\}$ and the above tent covers only this set of z . (The tent is undefined elsewhere.) Since $\alpha < 1$, ROC_H includes the unit circle. Thus, the frequency response $H_d(\omega) = H(e^{j\omega})$ is well defined and is a circular slice of the circus tent, around the unit circle. Now, since $H(z)$ is infinite at $z = \alpha e^{\pm j\omega_p}$, we expect that $H(z)$ will be large for z near the poles. If α is nearly one then the poles are close to the unit circle and $H_d(\omega)$ will be large for ω such that $e^{j\omega}$ is close to the poles. This suggests that $|H_d(\omega)|$ will look like

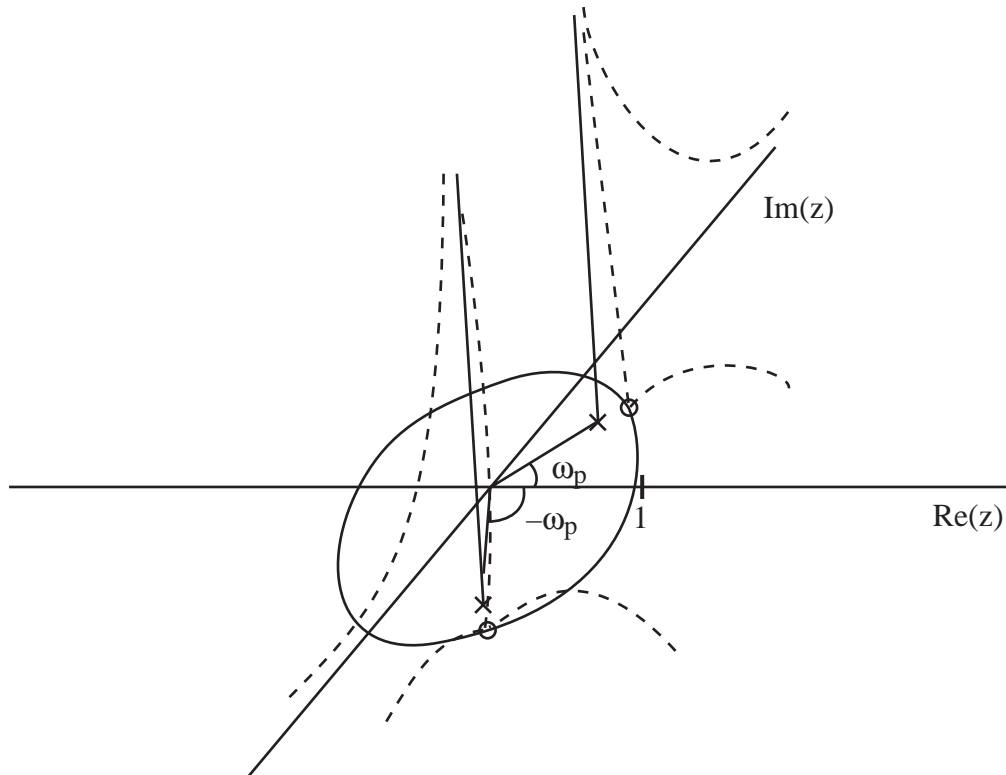


where the peaks occur near $\omega = \omega_p$.

Example

$$H(z) = \frac{z^2 - (2 \cos \omega_p)z + 1}{z^2 - (2\alpha \cos \omega_p)z + \alpha^2} = \frac{(z - e^{j\omega_p})(z - e^{-j\omega_p})}{(z - \alpha e^{j\omega_p})(z - \alpha e^{-j\omega_p})}$$

As before, $H(z)$ approaches ∞ at the pole locations $z = \alpha e^{\pm j\omega_p}$. $H(z) = 0$ at the zero locations $z = e^{\pm j\omega_p}$. Thus, $|H(z)|$ is similar to the previous two-pole circus tent except it “touches the ground” at the zero locations $z = e^{\pm j\omega_p}$ as shown below, where poles are indicated by \times and zeros are represented by \circ .



Since the pole and zero locations are close together, it may seem difficult to determine what $H_d(\omega) = \{H(z) \text{ for } z \text{ on the unit circle}\}$ will look like. A simple observation helps greatly, though. Suppose

$$H(z) = \frac{(z - z_1)(z - z_2)}{(z - p_1)(z - p_2)} \quad (\square)$$

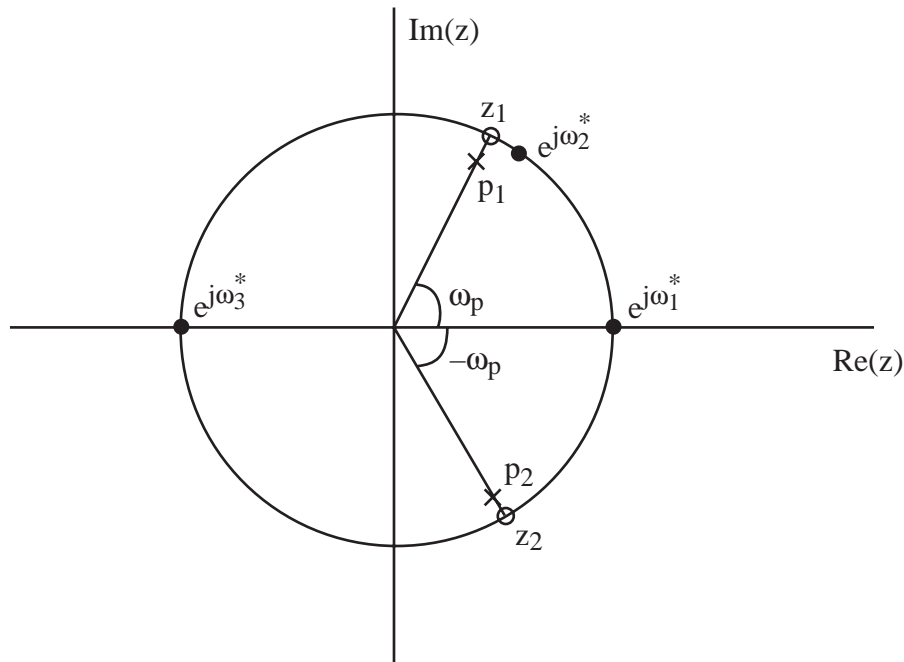
so that

$$|H_d(\omega)| = |H(e^{j\omega})| = \frac{|e^{j\omega} - z_1| |e^{j\omega} - z_2|}{|e^{j\omega} - p_1| |e^{j\omega} - p_2|}$$

Then, the value of $|H_d(\omega)|$ at some specific frequency ω^* is

$$\frac{|e^{j\omega^*} - z_1| |e^{j\omega^*} - z_2|}{|e^{j\omega^*} - p_1| |e^{j\omega^*} - p_2|}$$

The factor $|e^{j\omega^*} - z_i|$ is the distance in the complex plane between $e^{j\omega^*}$ and z_i . Likewise, the factor $|e^{j\omega^*} - p_i|$ is the distance in the complex plane between $e^{j\omega^*}$ and p_i . Thus, $|H_d(\omega^*)|$ is the product of the distances between $e^{j\omega^*}$ and z_1 , and between $e^{j\omega^*}$ and z_2 , divided by the product of the distances between $e^{j\omega^*}$ and p_1 , and between $e^{j\omega^*}$ and p_2 . In our example, these distances can be visualized for three different frequencies ω_1^* by examining the figure below.



We see that the distances from $e^{j\omega_1^*}$ to the zero z_1 at $e^{j\omega_p}$ and to the neighboring pole p_1 are nearly the same. Likewise, the distances from $e^{j\omega_1^*}$ to the zero z_2 at $e^{-j\omega_p}$ and its neighboring pole p_2 are nearly the same. Thus,

$$|H_d(\omega_1^*)| = \frac{|e^{j\omega_1^*} - z_1|}{|e^{j\omega_1^*} - p_1|} = \frac{|e^{j\omega_1^*} - z_2|}{|e^{j\omega_1^*} - p_2|} \approx (1)(1) = 1$$

The same result holds for any ω_1^* that is close to zero.

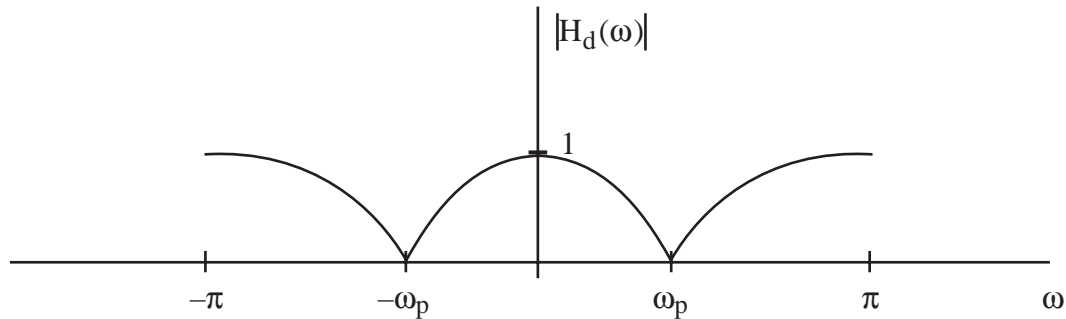
Similarly, in the case of $e^{j\omega_3^*}$ we see that distances to all poles and zeros are nearly equal. Thus,

$$|H_d(\omega_3^*)| \approx 1,$$

which holds for any ω_3^* roughly satisfying $\frac{\pi}{2} \leq |\omega_3^*| \leq \pi$. The distance between $e^{j\omega_2^*}$ and the zero at $e^{j\omega_p}$ approaches zero as $\omega_2^* \rightarrow \omega_p$. Thus

$$|H_d(\omega_2^*)| \approx 0$$

for ω_2^* close enough to ω_p . These considerations lead to the $|H_d(\omega)|$ sketched below.



Here, we can find the precise values of $|H_d(0)|$ and $|H_d(\pi)|$ by using

$$H_d(0) = H(1) = \frac{2 - 2\cos\omega_p}{1 + \alpha^2 - 2\alpha\cos\omega_p}$$

$$H_d(\pi) = H(-1) = \frac{2 + 2\cos\omega_p}{1 + \alpha^2 + 2\alpha\cos\omega_p}.$$

The above frequency response is that of a crude notch filter where signal components near $\omega = \omega_p$ are greatly attenuated and signal components at other frequencies are passed with nearly unit amplitude. The nulls (notches) in $H_d(\omega)$ at $\omega = \omega_p$ are caused by the zeros of $H(z)$ at $z = e^{\pm j\omega_p}$.

It is only slightly more difficult to gain a rough idea of what $\angle H_d(\omega)$ will look like. From (□) we have

$$\angle H_d(\omega) = \angle(e^{j\omega} - z_1) + \angle(e^{j\omega} - z_2) - \angle(e^{j\omega} - p_1) - \angle(e^{j\omega} - p_2)$$

Each term $(e^{j\omega} - z_i)$ or $(e^{j\omega} - p_i)$ is a vector in the complex plane. $\angle(e^{j\omega} - z_i)$ and $\angle(e^{j\omega} - p_i)$ are simply the angles of these vectors with respect to the positive real axis. This interpretation can be helpful when trying to visualize $\angle H_d(\omega)$ for low-order filters. In general, however, Matlab should be used to plot both $|H_d(\omega)|$ and $\angle H_d(\omega)$ for higher-order filters

Example

For an FIR lowpass filter, use Matlab to find the zero locations of $H(z)$. (All poles are at $z = 0$.) You will find that zeros in $|H_d(\omega)|$ within the stopband are caused by zeros of $H(z)$ on the unit circle. Other zeros of $H(z)$ are strategically placed off the unit circle to give a flat response in the passband.

Example

For a Butterworth lowpass filter you will find that some poles are located near the unit circle $e^{j\omega}$ for ω corresponding to the cutoff frequency.

Example

For an elliptic lowpass filter you will find poles near $e^{j\omega}$ for $\omega =$ cutoff frequency, and some zeros on the unit circle at locations corresponding to the stopband.