Chapter 18. The bilinear transformation and IIR filter transformations

In chapter 17, we computed the Butterworth filter using the inverse Laplace transform, which was cumbersome. The bilinear transformation is the standard approach used to convert a Laplace transform transfer function into a Z transform transfer function.

18.1. The bilinear transformation

One of the first results of the previous chapter was that the Laplace transform of a discrete signal evaluates to the Z transform with $z = e^{sT}$. The bilinear transformation is an approximation of this result.

The bilinear transformation is the substitution

$$z = e^{s\,T} \approx \frac{2 + s\,T}{2 - s\,T} \tag{18.1}$$

or alternatively

$$s \approx \frac{2}{T}\frac{z-1}{z+1} \tag{18.2}$$

where T is the sampling time, or $1\ /\ f_s$ for the sampling frequency f_s. Usually, when the bilinear substitution is performed, T is assumed to be to 1.

The result above follows from the Taylor series expansion of the function $e^{sT/2}$.

$$z = e^{s\,T} = \frac{e^{\frac{s\,T}{2}}}{e^{-\frac{s\,T}{2}}} = \frac{1 + \frac{\left(\frac{s\,T}{2}\right)^1}{1!} + \frac{\left(\frac{s\,T}{2}\right)^2}{2!} + \frac{\left(\frac{s\,T}{2}\right)^3}{3!} + \cdots}{1 + \frac{\left(-\frac{s\,T}{2}\right)^1}{1!} + \frac{\left(-\frac{s\,T}{2}\right)^2}{2!} + \frac{\left(-\frac{s\,T}{2}\right)^3}{3!} + \cdots} \approx \frac{1 + \frac{s\,T}{2}}{1 - \frac{s\,T}{2}} \tag{18.3}$$

Setting $T = 1$ for simplicity, we note that if $\mathrm{Re}(s) = \sigma < 0$, then

$$|z| = \left|\frac{2+s}{2-s}\right| = \left|\frac{2+\sigma-j\omega}{2-\sigma+j\omega}\right| = \left|\frac{4-(\sigma^2+\omega^2)-4j\omega}{4+(\sigma^2+\omega^2)-4\sigma}\right| < \frac{4-(\sigma^2+\omega^2)}{4+(\sigma^2+\omega^2)-4\sigma} < 1 \qquad (18.4)$$

Thus, if we have the Laplace transform transfer function of a stable filter with roots of the denominator in the left part of the s- complex plane, the transfer function that we will obtain with the bilinear transformation would have roots that are inside the unit circle and the filter will still be stable. The bilinear transformation preserves stability.

18.2. Butterworth filters with the bilinear transformation

If we use the bilinear transformation on the Laplace transform transfer function of the second order Butterworth filter in equation 17.31, we will obtain the following.

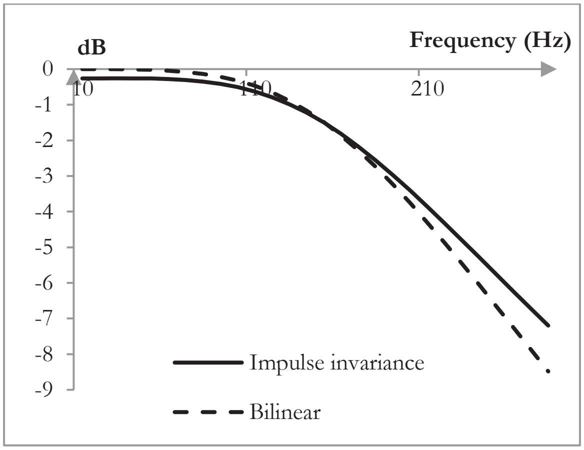
$$H(z) = (\omega_c{}^2 + 2\,\omega_c{}^2 z^{-1} + \omega_c{}^2 z^{-2}) \,/\, (4 - 4\,\omega_c\, cos\left(\frac{3\pi}{4}\right) + \omega_c{}^2 + (-8 + 2\omega_c{}^2)\, z^{-1} + \left(4 + 4\,\omega_c\, cos\left(\frac{3\pi}{4}\right) + \omega_c{}^2\right) z^{-2}) \qquad (18.5)$$

Suppose again that the cutoff frequency is $\omega_c = 0.6$. After scaling to obtain 1 at the beginning of the denominator, we get

$$H(z) = \frac{0.059435 + 0.118870\, z^{-1} + 0.059435\, z^{-2}}{1 - 1.201904\, z^{-1} + 0.439643\, z^{-2}} \qquad (18.6)$$

This filter is different than the *impulse invariant* filter described in the previous chapter, because the bilinear transformation is an approximation. Figure 112 compares the magnitude response of the two filters in the pass band and in the transition band.

Figure 112. Bilinear and impulse invariant Butterworth filters



Although in the discrete digital case the bilinear transformation filter performs better, in the analog case this same filter warps the frequency spectrum.

18.3. High pass and other Butterworth filters

We could attempt the filter transformations that we used to derive FIR high pass, band pass, and band stop filters when designing IIR filters as well. We would have to choose an easy an appropriate all pass filter and that could be difficult, considering that we would have to match the phase response of the all pass filter to the phase response of the low pass filter. It could also be fruitless, given that IIR filters do not behave as nicely in the stop band as do FIR filters.

The general prototype for an all pass filter was shown in equation 16.38. An all pass filter has the same, but inverted coefficients in the numerator and the denominator. If, for example, our low pass filter is the same as in the one in equation 18.6 then an easy all pass filter would be

$$H(z) = \frac{0.439643 - 1.201904\, z^{-1} + z^{-2}}{1 - 1.201904\, z^{-1} + 0.439643\, z^{-2}} \tag{18.7}$$

The high pass filter would be the difference between the two

$$H(z) = \frac{0.380208 - 1.320770\, z^{-1} + 0.940556\, z^{-2}}{1 - 1.201904\, z^{-1} + 0.439643\, z^{-2}} \tag{18.8}$$

The magnitude response of this high pass filter, however, is not as well behaved as we might have hoped. It has a normalized magnitude response that is greater than 1 in the pass band and that decreases slowly with higher frequencies, although that can be corrected to some extent by simply applying some gain to the coefficients in the numerator of the transfer function.

An easier filter transformation would be to realize that if the following is the magnitude response of a low pass filter

$$|H(j\omega)| = \frac{G}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}}} \tag{18.9}$$

then a good magnitude response of a high pass filter is

$$|H(j\omega)| = \frac{G}{\sqrt{1 + \left(\frac{\omega_c}{\omega}\right)^{2n}}} \tag{18.10}$$

As ω increases up to and past ω_c, this magnitude response would increase and settle at approximately 1.

> Given a low pass transfer function $H(s\ /\ \omega_c)$, we can typically create a high pass transfer function by substituting $s\ /\ \omega_c$ with $\omega_c\ /\ s$. Substituting $s\ /\ \omega_c$ with $(s^2 + \omega_c^2)\ /\ (B\ s)$ produces a band pass filter, where ω_c is the midpoint of the pass band and B is the width of the band. Substituting $s\ /\ \omega_c$ with $B\ s\ /\ (s^2 + \omega_c^2)$ produces a band stop filter.

We can then define the Butterworth filter again as the filter with the transfer function

$$H(s) = \frac{G}{B_n(S)} \tag{18.11}$$

where n is the order of the filter and $B_n(S)$ are the *normalized* Butterworth polynomials given by

$$\begin{aligned} B_n(S) &= \prod_{k=1}^{n/2} \left(S^2 - 2\,S\cos\left(\frac{2k+n-1}{2n}\pi\right) + 1\right), \quad n\ even \\ B_n(S) &= (S+1)\prod_{k=1}^{\frac{n-1}{2}} \left(S^2 - 2\,S\cos\left(\frac{2k+n-1}{2n}\pi\right) + 1\right), \quad n\ odd \end{aligned} \tag{18.12}$$

Substituting $S = s\ /\ \omega_c$ produces the Butterworth low pass filter, where ω_c is the cutoff frequency of the filter. Substituting $S = \omega_c\ /\ s$ produces the Butterworth high pass filter. Substituting $S = (s^2 + \omega_c^2)\ /\ (B\ s)$ produces the Butterworth band pass filter, where ω_c is the midpoint of the pass band and B is the width of the band. Substituting $S = B\,s\,/\,(s^2 + \omega_c^2)$ produces the Butterworth band stop filter. Since the transfer function and the Butterworth polynomials above are written irrespective of the cutoff frequency ω_c (i.e., they use the cutoff frequency $\omega_c = 1$) and the various forms of the Butterworth filter require substitutions, the transfer function and the polynomials are called *normalized*.

Take the transfer function for the second order Butterworth high pass filter with $G = 1$.

$$H(s) = \frac{s^2}{\omega_c^2 - 2\,s\,\omega_c\cos\left(\frac{3\pi}{4}\right) + s^2} \tag{18.13}$$

Use the bilinear transformation $s = 2\,(z - 1)\,/\,(z + 1)$ to rewrite this transfer function as follows.

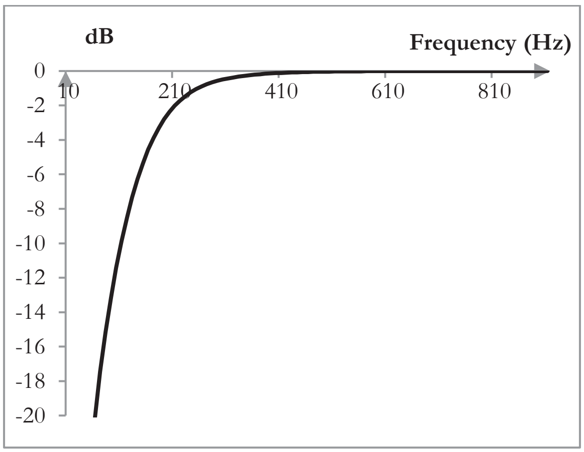
$$\begin{aligned} H(z) = {} & (4 - 8\,z^{-1} + 4\,z^{-2}) / \left(\left(\omega_c^2 - 2\sqrt{2}\omega_c + 4\right) + \left(2\omega_c^2 - 8\right) z^{-1} \right. \\ & \left. + \left(\omega_c^2 + 2\sqrt{2}\omega_c + 4\right) z^{-2}\right) \end{aligned} \tag{18.14}$$

When $\omega_c = 0.6$ for example, then the transfer function is

$$H(z) = \frac{0.6603 - 1.3208\,z^{-1} + 0.6604\,z^{-2}}{1 - 1.2019\,z^{-1} + 0.4396\,z^{-2}} \tag{18.15}$$

Take the sampling frequency of $f_s = 2000$ Hz. The cutoff frequency translates to $\omega_c = 191$ Hz and the transfer function above produces a filter with the magnitude response shown on figure 113.

Figure 113. Magnitude response of an example high pass Butterworth filter



Magnitude response of an example high pass second order Butterworth filter with $\omega_c = 0.6$.

The following is the formula for the second order Butterworth band pass filter after the bilinear transformation.

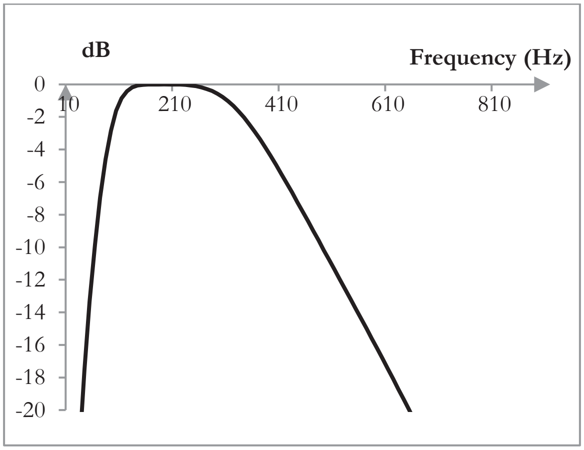
$$
\begin{gathered}
H(z) = \frac{a_0 + a_2 z^{-2} + a_4 z^{-4}}{b_0 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} + b_4 z^{-4}} \\
a_0 = 4\,B^2 \\
a_2 = -8\,B^2 \\
a_4 = 4\,B^2 \\
b_0 = 16 + 8\sqrt{2}B + 4(2{\omega_c}^2 + B^2) + 2\sqrt{2}B{\omega_c}^2 + {\omega_c}^4 \\
b_1 = -64 - 16\sqrt{2}B + 4\sqrt{2}B{\omega_c}^2 + 4{\omega_c}^4 \\
b_2 = 96 - 8(2{\omega_c}^2 + B^2) + 6{\omega_c}^4 \\
b_3 = -64 + 16\sqrt{2}B - 4\sqrt{2}B{\omega_c}^2 + 4{\omega_c}^4 \\
b_4 = 16 - 8\sqrt{2}B + 4(2{\omega_c}^2 + B^2) - 2\sqrt{2}B{\omega_c}^2 + {\omega_c}^4
\end{gathered}
\tag{18.16}
$$

For example, using $\omega_c = 0.6$ and $B = 1$ and scaling to obtain $b_0 = 1$ produces the filter

$$
H(z) = \frac{0.1132 - 0.2264\,z^{-2} + 0.1132\,z^{-4}}{1 - 2.3789\,z^{-1} + 2.3490\,z^{-2} - 1.2136\,z^{-3} + 0.3021\,z^{-4}} \tag{18.17}
$$

Given a sampling frequency of, say, $f_s = 2000$ Hz, the midpoint frequency translates to $\omega_c = 191$ Hz and the band width translates to $B = 318$ Hz. The magnitude response of this band pass filter is shown on figure 114.

Figure 114. Magnitude response of an example band pass Butterworth filter



Magnitude response of a second order band pass Butterworth filter with $\omega_c = 0.6$ and $B = 1$.

18.4. Higher order Butterworth filters

When designing higher order Butterworth filters we have two choices. We can use the methods describe above. Those require many computations and may be tedious. Alternatively, we can design higher order filters by stacking lower order filters. Note, for example, that in a low pass Butterworth filter of order four, the roots of the denominator form two complex conjugate pairs. s_1 and s_4 are a pair and s_2 and s_3 are a pair.

$$\begin{aligned} s_1 &= \omega_c e^{j\frac{5\pi}{8}} \\ s_2 &= \omega_c e^{j\frac{7\pi}{8}} \\ s_3 &= \omega_c e^{j\frac{9\pi}{8}} = \omega_c e^{-j\frac{7\pi}{8}} \\ s_4 &= \omega_c e^{j\frac{11\pi}{8}} = \omega_c e^{-j\frac{5\pi}{8}} \end{aligned} \tag{18.18}$$

We can rewrite the transfer function of the fourth order low pass Butterworth filter as follows.

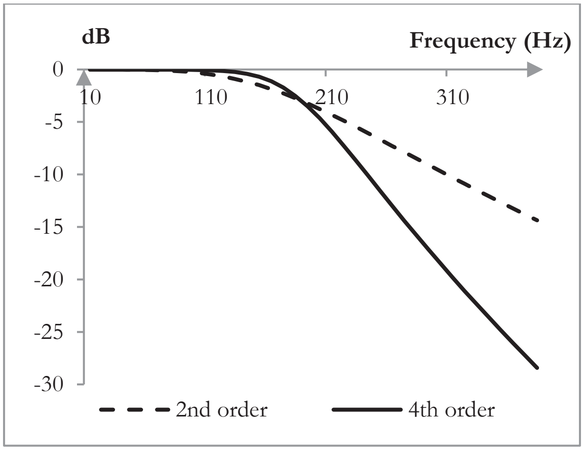
$$H(s) = \left(\frac{{\omega_c}^2}{(s - s_1)(s - s_4)}\right)\left(\frac{{\omega_c}^2}{(s - s_2)(s - s_3)}\right) \tag{18.19}$$

This is a multiplication of two transfer functions of two filters of order two. These two filters are not Butterworth filters – their roots are not the roots of the Butterworth filter of order two. Notwithstanding, we can then treat these two transfer functions as two separate filters.

The product of two Laplace transforms is the Laplace transform of the convolution of the two respective functions. In this case specifically, the multiplication of two transfer functions is the convolution of two impulse responses. This means that the filters are to be stacked one after the other. The output of the first filter is the input to the second filter.

Figure 115 compares the magnitude responses of the second and fourth order low pass Butterworth filters, where the fourth order filter is implemented as two second order filters stacked one after the other. The two filters themselves are computed after using bilinear transformations on the two separate multiples of the transfer function. Note that the two filters cross at around -3 dB, which, with Butterworth filters, happens at the cutoff frequency. The normalized cutoff frequency in this example is $\omega_c = 0.6$ radians per second, which is equivalent to approximately 191 Hz, given the sampling frequency of $f_s = 2000$ Hz.

Figure 115. Magnitude response of Butterworth filters of different order



Comparison of the magnitude response of a fourth order and a second order Butterworth filters with $\omega_c = 0.6$ and $f_s = 2000$ Hz.

18.5. Phase response of the Butterworth filter

Since we know the transfer function $H(z)$ of the Butterworth filter, we can compute its phase response with equation 16.13.

$$\Phi(\omega) = atan2\left(\frac{Re\big(H(z)\big)}{Im\big(H(z)\big)}\right) \tag{18.20}$$

For example, if the transfer function of some second order low pass Butterworth filter is

$$H(z) = \frac{a_0 - a_1\, z^{-1} + a_2\, z^{-2}}{1 - b_1\, z^{-1} + b_2\, z^{-2}} = \frac{a_0 z^2 - a_1\, z + a_2}{z^2 - b_1 z + b_2} \tag{18.21}$$

Then its phase response at $z = e^{j\,\omega} = \cos(\omega) - j\,\sin(\omega)$ is

$$\Phi(\omega) = atan2\left(\frac{\frac{A\,C + B\,D}{B^2 + D^2}}{\frac{B\,C - A\,D}{B^2 + D^2}}\right)$$

$$A = a_0\, cos(2\omega) + a_1\, cos(\omega) + a_2 \tag{18.22}$$

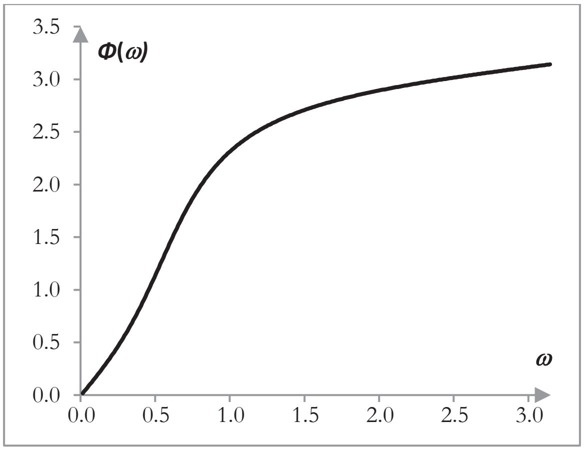
$$B = -a_0\, sin(2\omega) + a_1 sin(\omega)$$

$$C = cos(2\omega) + b_1\, cos(\omega) + b_2$$

$$D = -\,sin(2\omega) + b_1\, sin(\omega)$$

The phase response of our example second order low pass Butterworth filter is shown on figure 116 for the normalized ω between 0 and π. Unlike the phase response of symmetric FIR filters, this phase response is not linear.

Figure 116. Phase response of a Butterworth filter



Phase response of a second order Butterworth filter with $\omega_c = 0.6$.

18.6. Equivalent noise bandwidth of the Butterworth filter

While chapter 14 discusses the term equivalent noise bandwidth only as it applies to windows, the term is very appropriate for filters.

The simplest (first order) low pass Butterworth filter, for example, is defined by the transfer function

$$H(s) = \frac{1}{1 + \frac{s}{\omega_c}} = \frac{1}{1 + j\frac{\omega}{\omega_c}} \tag{18.23}$$

The magnitude response of this filter is

$$|H(j\omega)| = \frac{\omega_c}{\sqrt{{\omega_c}^2 + \omega^2}} \tag{18.24}$$

Since the magnitude response of the Butterworth filter is monotonically decreasing, $H_{max} = H(0) = 1$. The equivalent noise bandwidth is

$$B_N = \frac{1}{2\pi}\int_0^\infty \frac{{\omega_c}^2}{{\omega_c}^2 + \omega^2}\,d\omega = \frac{\pi}{2}\frac{\omega_c}{2\pi} \approx 1.57\,\frac{\omega_c}{2\pi} \tag{18.25}$$

Thus, the equivalent noise bandwidth is proportional to the normalized cutoff frequency. Note that the magnitude response at the cutoff frequency is

$$|H(j\omega)| = \frac{\omega_c}{\sqrt{{\omega_c}^2 + {\omega_c}^2}} = \frac{1}{\sqrt{2}} \approx -3\ dB \tag{18.26}$$

The ENBW relative to the -3 dB frequency is 1.57.

With filters of higher orders, the equivalent noise bandwidth decreases.

Order 1 ENBW ≈ 1.57

Order 2 ENBW ≈ 1.11

Order 3 ENBW ≈ 1.05

Order 4 ENBW ≈ 1.03

Order 5 ENBW ≈ 1.02

Order 6 ENBW ≈ 1.01

18.7. Warping of the frequency domain and the biquad transformation

The bilinear transformation filter performs better and is easier to derive, but warps the magnitude response of the filter. The properties of the resulting filter are evaluated at $s = -j\omega$ or $z = e^{-j\omega}$. However,

$$H(s) = H(-j\omega) \tag{18.27}$$

whereas

$$H\left(\frac{2\,(z-1)}{z+1}\right) = H\left(\frac{2\,(e^{-j\omega}-1)}{e^{-j\omega}+1}\right) = H\left(2\,\frac{e^{-\frac{j\omega}{2}}\left(e^{-\frac{j\omega}{2}} - e^{\frac{j\omega}{2}}\right)}{e^{-\frac{j\omega}{2}}\left(e^{-\frac{j\omega}{2}} + e^{\frac{j\omega}{2}}\right)}\right) = H\left(2\,\frac{-2\,j\,sin\left(\frac{\omega}{2}\right)}{2\,cos\left(\frac{\omega}{2}\right)}\right) \tag{18.28}$$
$$= H\left(-2\,j\,tan\left(\frac{\omega}{2}\right)\right)$$

The discrete filter $H(z)$, obtained after the bilinear transformation of $H(s)$, will behave at the discrete frequency ω_d the same way the continuous filter $H(s)$ behaves at the frequency $\omega_a = 2\tan(\omega_d/2)$ (alternatively, $\omega_d = 2\arctan(\omega_a/2)$).

For example, the squared magnitude response of the second order low pass Butterworth filter, computed from $H(s)$, is

$$|H(j\omega)H(-j\omega)| = \frac{{\omega_c}^4}{\omega^4 + {\omega_c}^4} \tag{18.29}$$

The squared magnitude response of the filter after bilinear transformation can be computed from

$$H(z) = \frac{{\omega_c}^2 z + 2{\omega_c}^2 + {\omega_c}^2 z^{-1}}{\left(4 + 2\sqrt{2}\omega_c + {\omega_c}^2\right)z + \left(-8 + 2{\omega_c}^2\right) + \left(4 - 2\sqrt{2}\omega_c + {\omega_c}^2\right)z^{-1}} \tag{18.30}$$

and is

$$\left|H\left(e^{j\omega}\right)H\left(e^{-j\omega}\right)\right| = \frac{{\omega_c}^4(\cos(\omega)+1)^2}{{\omega_c}^4(\cos(\omega)+1)^2 + 16(\cos(\omega)+1)^2 - 64\cos(\omega)}$$
$$= \frac{{\omega_c}^4}{{\omega_c}^4 + 16\frac{(\cos(\omega)-1)^2}{(\cos(\omega)+1)^2}} = \frac{{\omega_c}^4}{{\omega_c}^4 + \left(2\tan\left(\frac{\omega}{2}\right)\right)^4} \tag{18.31}$$

This warping of the frequency domain is small for small ω and increases as ω increases. At $\omega_a = 1$, for example, $\omega_d \approx 0.927$.

When designing digital filters, this frequency warping can be remedied in one of two ways. First, a discrete filter with the cutoff frequency ω_d can be designed with the bilinear transformation on the continuous filter with the cutoff frequency $\omega_a = 2\tan(\omega_d/2)$. Alternatively, instead of the bilinear transformation, one can use the biquad transformation

$$s = \frac{1}{K}\frac{z-1}{z+1}, \qquad K = \tan\left(\frac{\omega_c}{2}\right) \tag{18.32}$$

An example filter designed with the biquad transformation is presented in chapter 19.