Supplementary Notes for ELEN 4810 Lecture 17
Infinite Impulse Response (IIR) Filter Design

John Wright
Columbia University

Decmeber 11, 2015

Disclaimer: These notes are intended to be an accessible introduction to the subject, with no pretense at completeness. In general, you can find more thorough discussions in Oppenheim's book. Please let me know if you find any typos.

Reading suggestions: Oppenheim and Schafer 7.2-7.4, Appendix B,

1 IIR Filter Design

In filter design, the basic task is to go from design considerations (passband width, transition band width, stop band width, ect.) to a choice of $H(z)$, generally expressed in terms of its poles and zeros. Once we have $H(z)$, we can determine an appropriate block diagram for implementation. One approach to generating $H(z)$ is to design a continuous-time filter with the desired properties, and then transform it back to the discrete domain. This approach is popular mostly for historical reasons: by the time digital signal processing emerged in the 1960's and 70's, continuous-time filter design was an extensively studied area, with a vast literature and an array of practical design tools. Currently, the most widely-used discrete-time IIR designs for frequency-selective ("X-pass") filters are still based on transformations of continuous-time filters. Fortunately, these design tools are standard components of software packages such as Matlab's Signal Processing Toolbox. A relatively rudimentary understanding of the tradeoffs involved is enough to get started using them.[1]

2 A Reminder on the Laplace Transform

To borrow ideas from continuous-time filter design, we need to know at least a little about that area. Because our goal is to specify the discrete-time transfer function $H(z)$, we start with its continuous-time analogue, the *Laplace transform* $H_a(s)$ of the continuous-time impulse response $h_a(t)$. The Laplace transform of a continuous-time signal $f(t)$ is a function of a complex variable $s \in \mathbb{C}$:

$$F(s) = \int_{-\infty}^{\infty} f(t) e^{-st} dt. \tag{2.1}$$

[1] Of course, we still encourage you to study them with great depth and diligence!

Note that if $s = j\Omega$ is purely imaginary, $F(s)$ is simply the (continuous-time) Fourier transform, evaluated at Ω. So, just as the DTFT is the $\mathcal{Z}$-transform evaluated around the unit circle, the Fourier transform is the Laplace transform evaluated along the imaginary axis:

$$F(j\Omega) = \int_{-\infty}^{\infty} f(t)e^{-j\Omega t}dt = F(s)\Big|_{s=j\Omega}. \tag{2.2}$$

Just as the $\mathcal{Z}$-transform summation may only converge for certain $z \in \mathbb{C}$, the Laplace transform integral (2.1) may only be well-defined for a subset of $s \in \mathbb{C}$. Also, like the $\mathcal{Z}$-transform, it happens that often the Laplace transform $F(s)$ is a rational function of s. It is very useful to understand how the region of convergence of $F(s)$, as well as the poles and zeros, relate to the properties of the system whose impulse response is $f(t)$. If $h(t)$ is the (continuous-time) impulse response of a stable and causal system, then $h(t) = 0$ for $t < 0$, and $\int_{t=0}^{\infty} |h(t)|dt < +\infty$. This implies that for any s satisfying $\mathrm{Re}[s] \geq 0$,

$$\begin{aligned}
\int_{-\infty}^{\infty} |h(t)e^{-st}|dt &= \int_{0}^{\infty} |h(t)| \underbrace{|e^{-\mathrm{Re}[s]t}|}_{\leq 1, \text{ since } \mathrm{Re}[s]t \geq 0} \underbrace{|e^{-j\mathrm{Im}[s]t}|}_{=1} dt && (2.3)\\
&\leq \int_{0}^{\infty} |h(t)|dt && (2.4)\\
&< +\infty. && (2.5)
\end{aligned}$$

So, the *poles of a stable, causal continuous-time system lie in the left half plane* ($\mathrm{Re}[s] < 0$). There is a rich theory associated with the Laplace transform, and its utility in studying continuous time systems. You probably saw some of the highlights of this in your previous signals and systems course. For our purposes, we recap the two most important facts from the above discussion:

- The continuous-time Fourier transform is equal to the Laplace transform, evaluated along the imaginary axis $s = j\Omega$.
- For stable, causal systems, any poles of the Laplace transform lie in the left half plane ($\mathrm{Re}(s) < 0$).

3 Turning a Continuous-Time System into a Discrete-Time System

Suppose we are given $H_a(s)$, the Laplace transform of a continuous-time impulse response $h_a(t)$.[2] For the designs of interest here, $H_a(s)$ will be a *rational function* of s. We can generate another function $H(z)$ by making a substitution

$$s = \frac{z-1}{z+1}, \tag{3.1}$$

giving

$$H(z) = H_a\left(\frac{z-1}{z+1}\right). \tag{3.2}$$

[2]Note that typically, filter characteristics are specified in the Fourier domain. We usually start with the form of the magnitude response $|H_a(j\Omega)|^2$, and then determine a stable, causal $H_a(s)$ whose magnitude squared is $|H_a(j\Omega)|^2$. In designing the continuous-time system $H_a(s)$, we may never actually write down the impulse response $h_a(t)$.

In signal processing, this mapping is referred to as *bilinear transformation*.

The mapping $z \mapsto \psi(z) = \frac{z-1}{z+1}$ is well-defined for all $z \in \mathbb{C} \setminus \{-1\}$. Strictly speaking,

$$\psi : \mathbb{C} \setminus \{-1\} \to \mathbb{C} \setminus \{1\}. \tag{3.3}$$

Solving for z in the equation $s = \psi(z)$, we obtain an inverse relationship

$$z = \frac{s+1}{1-s}, \tag{3.4}$$

which is valid for all $s \in \mathbb{C} \setminus \{1\}$. The mappings ψ and ψ^{-1} are highly nonlinear transformations of the complex plane. From the definition of ψ, it is in no way obvious why this might be the *right* way to transform the complex plane to convert a continuous time transfer function $H_a(s)$ to a discrete time transfer function $H(z)$. We make three observations below, which justify the use of the transformation (3.1).

Bilinear transformation maps rational $H_a(s)$ to rational $H(z)$. Suppose that $H_a(s)$ is rational:

$$H_a(s) = \alpha \frac{\prod_{i=1}^{M}(s - c_i)}{\prod_{\ell=1}^{N}(s - d_\ell)}. \tag{3.5}$$

The following shows that $H(z)$ is also a rational function of z, and that its poles and zeros correspond to those of $H_a(s)$ in a natural way:

Proposition 3.1. *Suppose that $H_a(s)$ satisfies* (3.5)*, and that none of the c_i or d_ℓ are equal to* 1*. Then $H(z)$ is also a rational function of z, with zeros of the form $\psi^{-1}(c_i)$ and poles $\psi^{-1}(d_\ell)$.*

Proof. Write

$$\begin{aligned}
H(z) &= H_a\left(\frac{z-1}{z+1}\right) && (3.6)\\
&= \alpha \frac{\prod_{i=1}^{M}\left(\frac{z-1}{z+1} - c_i\right)}{\prod_{\ell=1}^{N}\left(\frac{z-1}{z+1} - d_\ell\right)} && (3.7)\\
&= \alpha(z+1)^{N-M} \frac{\prod_{i=1}^{M}\left(z - 1 - c_i(z+1)\right)}{\prod_{\ell=1}^{N}\left(z - 1 - d_\ell(z+1)\right)} && (3.8)\\
&= \alpha(z+1)^{N-M} \left(\frac{\prod_{i=1}^{M}(1-c_i)}{\prod_{\ell=1}^{N}(1-d_\ell)}\right) \frac{\prod_{i=1}^{M}\left(z - \frac{1+c_i}{1-c_i}\right)}{\prod_{\ell=1}^{N}\left(z - \frac{1+d_\ell}{1-d_\ell}\right)} && (3.9)\\
&= \alpha(z+1)^{N-M} \left(\frac{\prod_{i=1}^{M}(1-c_i)}{\prod_{\ell=1}^{N}(1-d_\ell)}\right) \frac{\prod_{i=1}^{M}\left(z - \psi^{-1}(c_i)\right)}{\prod_{\ell=1}^{N}\left(z - \psi^{-1}(d_\ell)\right)} && (3.10)
\end{aligned}$$

□

This is auspicious: practically all useful continuous-time filter design techniques produce rational designs. Applying bilinear transformation produces rational discrete-time filter, whose poles and zeros correspond to the poles and zeros of the continuous-time filter in a very natural way. As discussed briefly in the last two lectures, rational $H(z)$ are especially well-suited to implementation, and can be converted into block diagrams in a number of ways.

Bilinear transformation maps the imaginary line to the unit circle. However, we have yet to say anything about how the frequency response of the resulting discrete time filter relates to that of the continuous-time filter we started with. Recall that the $H_a(s)$ restricted to $s = j\Omega$ gives the Fourier transform $H_a(j\Omega)$. In contrast, $H(z)$ evaluated around the unit circle gives the discrete-time Fourier transform $H(e^{j\omega})$. The bilinear transformation maps the imaginary axis to the unit circle:

Proposition 3.2. *The transformation $z = \frac{s+1}{1-s}$ maps the imaginary axis $\{s \mid s = j\Omega,\ \Omega \in\}$ onto the unit circle $\left\{z \mid z = e^{j\omega}\right\} \setminus \left\{e^{j\pi}\right\}$.*

Proof. For $s = j\Omega$,

$$\begin{aligned} z &= \frac{1+s}{1-s} \\ &= \frac{1+j\Omega}{1-j\Omega}. \end{aligned} \tag{3.11}$$

Because $|1 + j\Omega| = |1 - j\Omega|$, $|z| = 1$. □

So, under this transformation, the discrete-time Fourier transform $H(e^{j\omega})$ is simply the "compression" of the continuous-time Fourier transform $H_a(j\Omega)$ to the unit circle: if $z = e^{j\omega}$, the corresponding point s is $s = j\Omega$, with

$$\Omega = \tan\left(\omega/2\right), \tag{3.12}$$

and $\omega = 2\tan^{-1}(\Omega)$.[3] In particular, low-pass continuous-time filters transform into low-pass discrete-time filters; high-pass continuous-time filters transform into low-pass discrete-time filters.

Bilinear transformation maps poles in the left half plane to poles inside the unit circle. We mentioned above that stable, causal $H_a(s)$ have their poles in the left half-plane ($\mathrm{Re}[s] < 0$). Stable, causal $H(z)$ have poles inside the unit circle ($|z| < 1$). The previous proposition can be extended to show that for any s in the left half plane, the corresponding z is inside the unit circle. Indeed,

Proposition 3.3. *The transformation $z = \frac{s+1}{1-s}$ maps the left half plane $\{s \mid \mathrm{Re}[s] < 0\}$ onto the open disc $\{z \mid |z| < 1\}$.*

Proof. For $s = \alpha + j\Omega$, we simply notice that

$$z = \frac{1+s}{1-s} = \frac{1+\alpha+j\Omega}{1-\alpha-j\Omega}. \tag{3.16}$$

[3] To show these relationships, note that

$$\begin{aligned} s &= \frac{e^{j\omega}-1}{e^{j\omega}+1} = \frac{(e^{j\omega}-1)(e^{-j\omega}+1)}{|e^{j\omega}+1|^2} && (3.13)\\ &= \frac{2j\sin(\omega)}{(1+\cos\omega)^2+\sin^2(\omega)} = j\frac{2\sin(\omega)}{2+2\cos(omega)} && (3.14)\\ &= j\frac{\sin(\omega)}{1+\cos(\omega)} = j\tan(\omega/2). && (3.15) \end{aligned}$$

The magnitude of the numerator is $\sqrt{(1+\alpha)^2+\Omega^2}$, while the magnitude of the denominator is $\sqrt{(1-\alpha)^2+\Omega^2}$. If s is in the left half plane, α is negative, and the denominator has larger magnitude. Conversely, suppose that $|z|<1$. Then $z=re^{j\theta}$ with $|r|<1$. The corresponding s is

$$\begin{aligned} s &= \frac{z-1}{z+1} && (3.17)\\ &= \frac{re^{j\theta}-1}{re^{j\theta}+1} && (3.18)\\ &= \frac{(re^{j\theta}-1)(re^{-j\theta}+1)}{|re^{j\theta}+1|^2} && (3.19)\\ &= \frac{r^2+r(e^{j\theta}-e^{-j\theta})-1}{|re^{j\theta}+1|^2}. && (3.20) \end{aligned}$$

The denominator is real and positive. In the numerator, $r(e^{j\theta}-e^{-j\theta})$ is purely imaginary; the real part r^2-1 is necessarily negative. □

A few additional notes. The text describes the bilinear transform via the equivalent formula $s=\frac{1-z^{-1}}{1+z^{-1}}$. This leads to equivalent expressions for $H(z)$ phrased as rational functions of z^{-1}. Identical considerations and properties apply. The text also includes an additional multiplicative factor in the expression for s, writing $s=\frac{2}{T_d}\frac{z-1}{1+z}$. This multiplier has absolutely no consequence on the design process, and minimal consequences for the properties of the transformation, and so we omit it for simplicity.[4]

4 Four Designs for Continuous-Time Low-Pass Filters

We briefly describe four of the most popular types of continuous-time filters. In practical applications, each type can be constructed using existing computational tools – for example, Matlab's signal processing toolbox has a dedicated command for each. Using the Matlab command `bilinear`, any of these continuous-time filters can then be converted into a discrete time filter.[5] In the section, we briefly describe the properties of each of these types of filter. Figure **??** compares them on a concrete design example, and illustrates their characteristics.

Butterworth filters. A (lowpass) Butterworth filter with cutoff Ω_c and order n has squared magnitude response

$$|H_c(j\Omega)|^2=\frac{1}{1+(\Omega/\Omega_c)^{2n}} \tag{4.1}$$

This function is monotone decreasing. For any choice of $n\geq 1$, $|H_c(j0)|=1$, $|H_c(j\Omega_c)|=1/\sqrt{2}$, and $\lim_{\Omega\to\infty}|H_c(j\Omega)|=0$. However, as n increases, the transition from 1 to 0 becomes increasingly abrupt.

[4] Roughly speaking, the parameter $\frac{2}{T_d}$ arises if we try to view the discrete-time filter as an approximation to the continuous-time filter, derived via numerical integration with the trapezoidal rule. We do not emphasize this viewpoint here.

[5] Alternatively, they can be specified directly in the discrete domain using the command `designfilter`.

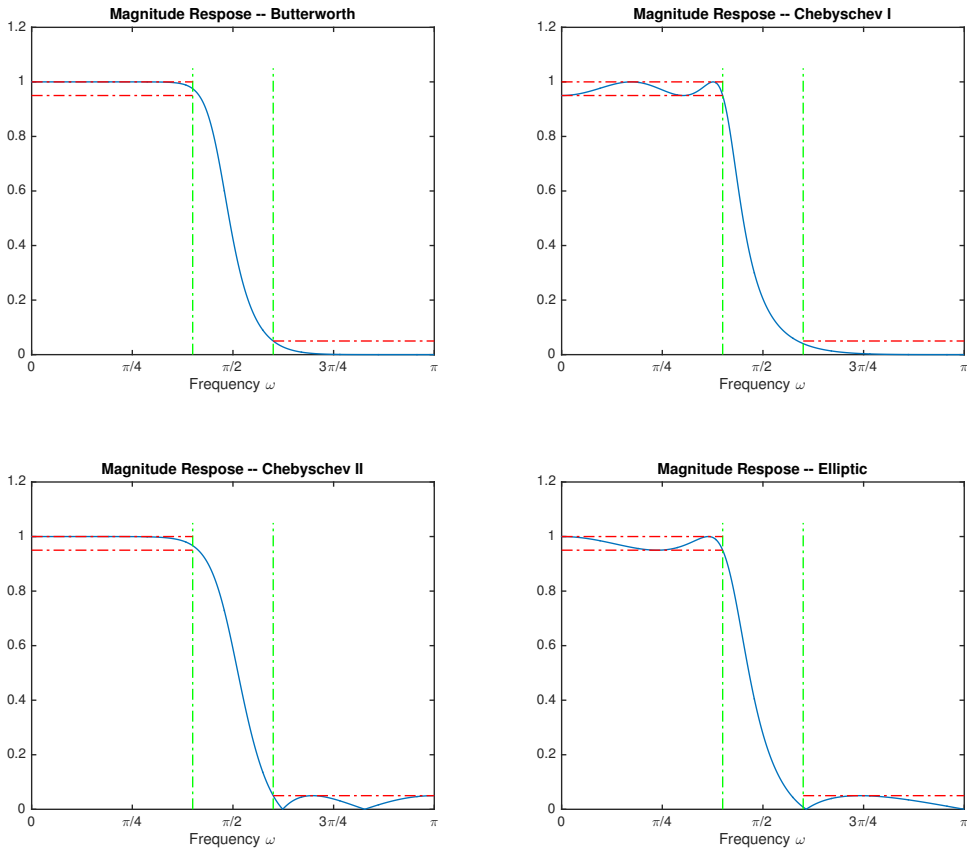

Figure 1: **Four IIR Designs.** We design IIR lowpass filters with the following characteristics: $\omega_p = 0.4 \times \pi$, $\omega_s = 0.6 \times \pi$, $\delta_p = \delta_s = 0.05$. Top left: Butterworth filter. An order $n = 7$ filter is required to meet the specification. Notice the monotone decrease in both the passband and stopband. Top right: Chebyschev I filter. An order $n = 4$ filter is required to meet the specification. Notice the presence of ripple in the passband, and monotone decrease in the stopband. Bottom left: Chebyshev II filter. Again an order $n = 4$ filter is required. Notice that now the magnitude response is monotone in the passband, but exhibits ripple in the stopband. Bottom right: an elliptic filter. Here an order $n = 3$ filter is sufficient. Notice that this is the lowest order of any of the four filter types considered here. Notice the ripple in both the passband and stopband.

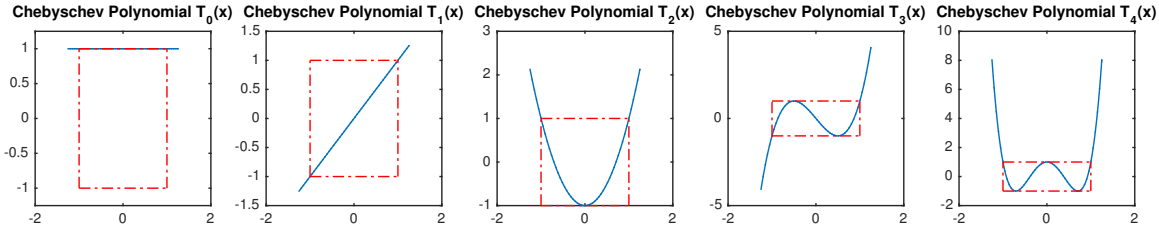

Figure 2: **Chebyshev polynomials of order** $0, \ldots 4$**.** Each figure plots $T_n(x)$ for $x \in [-1.5, 1.5]$ (blue). Also plotted is the unit box $[-1, 1] \times [-1, 1]$. Notice that for $x \in [-1, 1]$, $T_n(x)$ oscillates between -1 and 1. Moreover, for $n > 0$, $|T_n(x)|$ increases montonically when $x > 1$.

The main advantage of the Butterworth filter is that it has no ripples – the magnitude response is monotone in both the passband and the stopband. In fact, the filter is characterized by the fact that its magnitude response is "maximally flat" at $\Omega = 0$:

$$\frac{d^k}{d\omega^k}|H_c(j\Omega)|^2\big|_{\Omega=0} = 0 \tag{4.2}$$

for $k = 1, 2, \ldots, 2n-1$. This is the largest number of derivatives that can be zero for any order n filter.

The price of this relatively flat magnitude response is a wider transition region, compared to designs that allow ripples in the passband and/or stopband. Compared to designs that are not monotone, the magnitude response of the Butterworth "rolls off" much more slowly, and hence has a wider transition region.

Chebyschev Types I and II. Chebyschev filters are built out of Chebyschev polynomials, $T_n(x)$. For each n, T_n is a polynomial of degree n. The T_n can be inductively, via the relationships

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{4.3}$$

With a bit of trigonometry (which we did in the previous lecture), this can be shown to imply that

$$T_n(x) = \cos(n \cos^{-1}(x)), \tag{4.4}$$

for $x \in [-1, 1]$. In particular, for $x \in [-1, 1]$, $|T_n(x)| \leq 1$. This curious family of polynomials is interesting because it minimizes (over all possible degree-n polynomials) the maximum value over $[-1, 1]$:

Fact 4.1. *Let $T_n(x)$ denote the n-th order Chebyschev polynomial, and let $f(x) = \frac{1}{2^{n-1}}T_n(x)$. Then $f(x)$ minimizes $\max_{x\in[-1,1]} |g(x)|$ over all degree-n polynomials g with leading coefficient 1.*

Figure 2 plots the first five Chebyschev polynomials $T_0, \ldots, T_4$. In each plot, the unit box $[-1, 1] \times [-1, 1]$ is outlined in red. Notice that for x between -1 and 1, $T_n(x)$ indeed oscillates between -1 and 1. As $|x|$ increases away from 1, $|T_n(x)|$ increases to $+\infty$. Moreover, the larger n is, the more rapidly $|T_n(x)|$ increases outside of $[-1, 1]$.

The Chebyschev Type I filter takes advantage of this behavior to construct a filter which is equiripple within the passband, and falls off monotonically outside the passband. The magnitude-squared function of the Chebyschev Type I filter is

$$|H_c(j\Omega)|^2 = \frac{1}{1+\varepsilon^2 T_n^2(\Omega/\Omega_c)}. \tag{4.5}$$

Notice that for $0 \leq \Omega \leq \Omega_c$, $|H_c(j\Omega)|^2$ oscillates between $\frac{1}{1+\varepsilon^2}$ and 1. Moreover, when Ω exceeds Ω_c, $T_n^2(\Omega/\Omega_c)$ blows up to ∞, and so $|H_c(j\Omega)|^2 \to 0$.

The Type I Chebyschev filter exhibits ripples in the passband, but is monotone in the stopband. We can also achieve the opposite behavior (monotone in passband, but ripples in the stopband) by using the Chebyschev polynomials in a slightly different way. A Type II Chebyschev filter satisfies

$$|H_c(j\Omega)|^2 = \frac{1}{1+(\varepsilon^2 T_n^2(\Omega_c/\Omega))^{-1}}. \tag{4.6}$$

Notice that when $\Omega \to 0$, $\Omega_c/\Omega \to \infty$, and so $T_n^2(\Omega_c/\Omega) \to \infty$; this implies that as $\Omega \searrow 0$, $|H_c(j\Omega)|^2$ approaches 1. Moreover, $|H_c(j\Omega)|^2$ is monotone decreasing on $(0, \Omega_c)$. When $\Omega > \Omega_c$, $0 < \Omega_c/\Omega < 1$, and so $T_n^2(\Omega_c/\Omega)$ oscillates between zero and one in the stopband.

Elliptic Filters. Whereas Chebyshev filters either attempt to control the maximum error in the passband (Type I) or stop band (Type II), *elliptic filters* minimize the transition width for a given order N. The magnitude response has the form

$$|H_c(j\Omega)|^2 = \frac{1}{1+\varepsilon^2 U_n^2(\Omega)}, \tag{4.7}$$

where U_n is a Jacobi elliptic function. Important properties of elliptic filters are that: (i) they exhibit ripples in both the passband and the stopband and (ii) they exhibit the sharpest possible transition between the passband and stopband, for a given choice of N, ω_p, and the filter tolerances.

5 Frequency Transformations for High-Pass and Bandpass Filters

Using the designs described above and bilinear transformation, we can produce a variety of discrete-time low-pass filters. Moreover, by performing appropriate transformations of either the s-plane or the z-plane, we can convert these low-pass filters into bandpass or even high-pass filters. Such transformations can be applied to the continuous-time filter, before bilinear transformation, or to the discrete-time filter, after bilinear transformation.

Following Section 7.4 of the text, we describe this process in the $\mathcal{Z}$-domain. Suppose we are given a low-pass filter, with system function

$$H_{lp}(z^{-1}) \tag{5.1}$$

We generate a new filter $H(z^{-1}) = H_{lp}(Z^{-1})\big|_{Z^{-1}=\varphi(z^{-1})}$. Similar to bilinear transformation, we wish to guarantee to map rational H_{lp} to rational H. We also would like to map the unit circle to itself, so that the frequency response of H can be related directly to the frequency response of H_{lp}.

Finally, we wish to map stable, causal systems to stable, causal systems, and so the open disc should map to itself. It can be shown that if we wish to satisfy these requirements, we must set

$$Z^{-1} = \pm \prod_{k=1}^{N} \frac{z^{-1} - \alpha_k}{1 - \alpha_k z^{-1}}, \tag{5.2}$$

where the complex scalars α_k have magnitude strictly less than one.

For example, to map a low-pass filter to a high-pass filter, we can set

$$Z^{-1} = -\frac{z^{-1} + \alpha}{1 + \alpha z^{-1}}, \qquad \alpha = -\frac{\cos((\theta_c + \omega_c)/2)}{\cos((\theta_c - \omega_c)/2)}, \tag{5.3}$$

where θ_c is the cutoff frequency of H_{lp}, and ω_c is the target cutoff frequency. Because $\varphi(z^{-1})$ is all-pass, whenever $z = e^{j\omega}$, $Z = e^{j\theta}$, for some θ which depends on ω. For example, if $\alpha = 0$, we have $Z^{-1} = -z^{-1}$, and so we may take $\theta = \omega + \pi$. In this situations, low frequencies ω map to high frequencies θ, and vice-versa.

The text also provides formulas for converting low-pass filters to low-pass filters with different cutoff frequencies, as well as formulas for producing bandpass and bandstop filters.