5.3. Impulse-Invariant Method (Impulse Invariant Transformation)

In the impulse invariance method, our objective is to design an IIR filter having a unit impulse sample *h(n)* that is the sampled version of the impulse response of the analog filter $h_A(t)$. That is,

$$h(n) = h(nT) = h_A(nT) = h_A(t) \quad n = 0, 1, 2, ...$$

where T is the sampling interval. In consequence of this result, the frequency response of the digital filter is an aliased version of the frequency response of the corresponding analog filter.

Let the transfer function of an analog filter be given

$$H_A(p) = \frac{L[y(t)]}{L[x(t)]} = \frac{B(p)}{A(p)} = \frac{\sum_{k=0}^{M} b_k p^k}{\sum_{k=0}^{N} a_k p^k}$$

To demonstrate how an analog filter is digitized using the impulse invariant transformation, we rewritten this equation in its partial expansion, as

$$H_A(p) = \sum_{i=1}^{N} \frac{c_i}{p + d_i}$$

where

$$c_i = H_A(p)(p + d_i)\big|_{p=-d_i}$$

and d_i is the location of the *i*-the pole. In writing this expression, we have assumed that the order of the numerator M is less that the order of the denominator N and that all the poles of $H_A(p)$ are simple. The assumption that $M<N$ must be valid for the system to be digitized – otherwise the aliasing in the digital system would be intolerable. If the poles of $H_A(p)$ are not simple, the discussion in this section can be appropriately modified.

The impulse response of the analog filter $h_A(t)$ with transfer function $H_A(p)$ is of the form

$$h_A(t) = L^{-1}[H_A(p)] = L^{-1}\left[\sum_{i=1}^{N} \frac{c_i}{p + d_i}\right] = \sum_{i=1}^{N} L^{-1}\left[\frac{c_i}{p + d_i}\right] = \sum_{i=1}^{N} c_i e^{-d_i t}$$

The corresponding digital impulse response *h(n)* can be written as the sampled version of $h_A(t)$; i.e.

$$h(n) = \sum_{i=1}^{N} c_i e^{-d_i nT}$$

The *z*-transform of *h(n)* is given by

$$H(z) = Z[h(n)] = \sum_{n=0}^{\infty} h(n) z^{-n} = \sum_{n=0}^{\infty} z^{-n} \sum_{i=1}^{N} c_i e^{-d_i nT}$$

Interchanging orders of summation and summing over n gives

$$H(z)=\sum_{i=1}^{N} c_i \sum_{n=0}^{\infty}\left(e^{-d_i T} z^{-1}\right)^n$$

The inner sum in the previous expressions converges because $p_k < 0$ and yields

$$\sum_{n=0}^{\infty}\left(e^{-d_i T} z^{-1}\right)^n = \frac{1}{1-e^{-p_k T} z^{-1}}$$

Therefore, the transfer (system) function of the digital filter is

$$H(z)=\sum_{i=1}^{N} \frac{c_i}{1-e^{-d_i T} z^{-1}}$$

By comparing *H(z)* and $H_A(p)$ it can be seen that *H(z)* is obtained from $H_A(p)$ by using the mapping relation

$$\frac{c_i}{p+d_i} \rightarrow \frac{c_i}{1-e^{-d_i T} z^{-1}}$$

for simple poles. When d_i is complex, then c_i is also complex. Since $h_A(t)$ is real, this implies that there is also a pole at $\overline{d_i}$ with $\overline{c_i}$. By combining these terms, we obtain the terms of the forms

$$\frac{c_i}{p+d_i}+\frac{\overline{c_i}}{p+\overline{d_i}}=\frac{2g_i p+2\left(\sigma_i g_i+\omega_i h_i\right)}{p^2+2\sigma_i p+\left(\sigma_i^2+\omega_i^2\right)}$$

$$\frac{c_i}{1-z^{-1}e^{-d_i T}}+\frac{\overline{c_i}}{1-z^{-1}e^{-\overline{d_i}T}}=\frac{2g_i - z^{-1}e^{-\sigma_i T}\left[2g_i\cos(\omega_i T)-2h_i\sin(\omega_i T)\right]}{1-2z^{-1}e^{-\sigma_i T}\cos(\omega_i T)+z^{-2}e^{-2\sigma_i T}}$$

where

$$d_i=\sigma_i+j\omega_i \quad and \quad c_i=g_i+jh_i .$$

With the above given forms for the transfer (system) function *H(z)*, the IIR filter is easily realized as a parallel bank of single-pole filters. If some of poles are complex valued, they may be paired together and combined to form two-pole filter sections. In addition, two factors containing real-valued poles may be combined to form two-pole filter. Consequently, the resulting filter may be realized as a parallel bank of two-pole filter.

We should recall that when a continuous time signal $h_A(t)$ with spectrum $H_A(F)$ is sampled at a rate $F_S=1/T$ samples per second, the spectrum of the sampled signal is the periodic repetition of the scaled spectrum $F_S H_A(F)$ with period F_S. Specifically, the relationship is

$$H(f)=F_S\sum_{k=-\infty}^{\infty} H_A\left[\left(f-k\right)F_S\right]$$

or

$$H(e^{j\omega}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(F - kF_S\right) = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(2\pi F - k2\pi F_S\right) = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(\Omega - 2\pi kF_S\right)$$

i.e.

$$H(e^{j\omega}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(\Omega - \frac{2\pi k}{T}\right) = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(\Omega - k\Omega_S\right)$$

where $f = F / F_S$ is normalized frequency. Aliasing occurs if the sampling rate F_S is less then twice the highest frequency contained in $H_A(F)$. The next figure depicts the frequency response of a low-pass analog filter and the frequency response of the corresponding digital filters.

It is clear that the digital filter with frequency response $H(\omega)$ will possess the frequency response characteristics of the corresponding analog filter if the sampling interval T is selected sufficiently small to avoid completely or minimize the effects of aliasing. It is also clear that the impulse invariance method is inappropriate for designing high-pass filters or stop-band filter due to spectrum aliasing that results from the sampling process.

To investigate the mapping between the *z*-plane and the *p*-plane implied by the sampling process, we rely on a generalization of the last expression which relates *z*-transform of *h(n)* to the Laplace transform of $h_A(t)$. This relationship is

$$H(z)\Big|_{z=e^{pT}} = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(p - j\frac{2\pi k}{T}\right)$$

where

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n}$$

$$H(z)\Big|_{z=e^{pT}} = \sum_{n=-\infty}^{\infty} h(n) e^{-pTn}$$

Note that when $p = j\Omega$,

$$H(z)\Big|_{z=e^{pT}} = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(p - j\frac{2\pi k}{T}\right)$$

reduces to

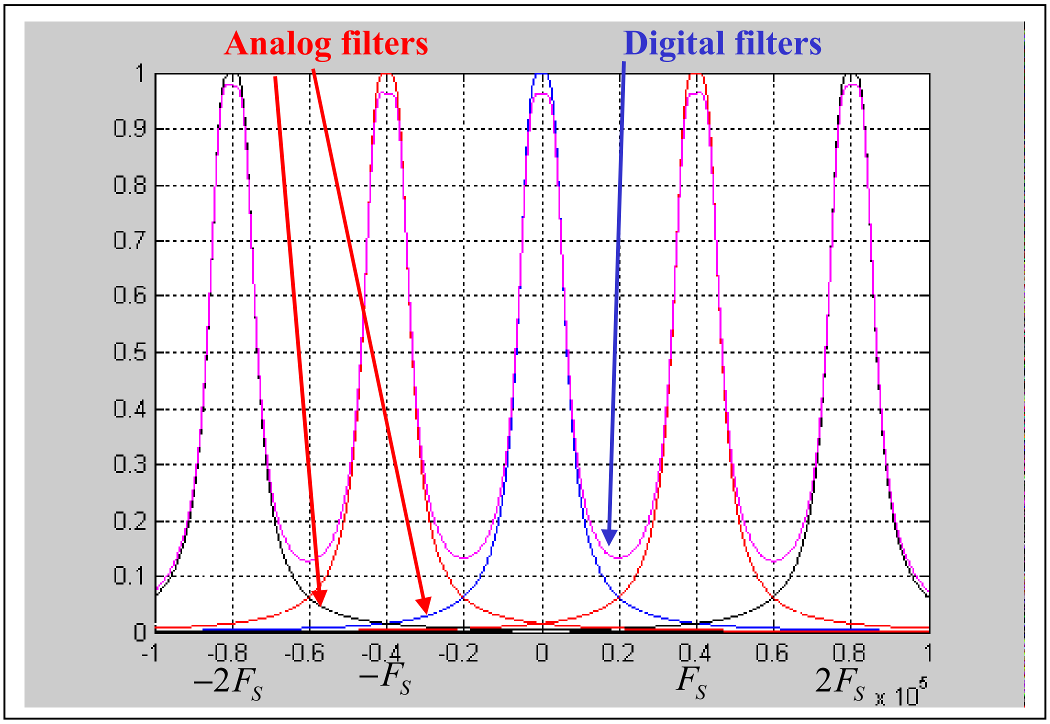
$$H(e^{j\omega}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} H_A\left(\Omega - k\Omega_S\right)$$

where the factor of *j* is suppressed in our notation.

The general characteristic of the mapping

$$z = e^{pT}$$

can be obtained by substitution $p = \sigma + j\Omega$ and expressing the complex variable *z* in the polar form as $z = re^{j\omega}$.



With these substitutions we can expressed as

$$z = re^{j\omega} = e^{(\sigma + j\Omega)T} = e^{\sigma T} e^{j\Omega T}$$

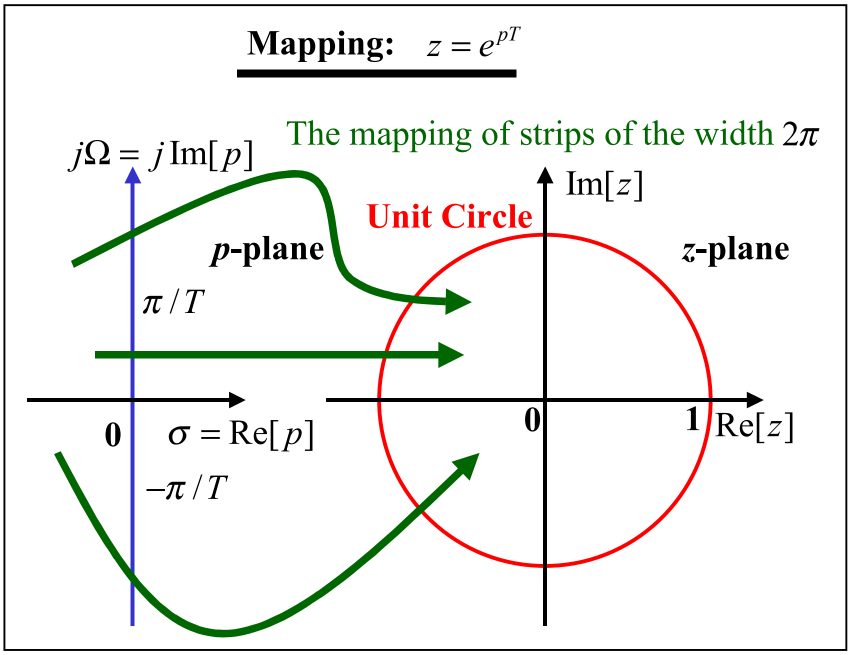
Clearly, we must have

$$r = e^{\sigma T} \quad \omega = \Omega T \quad \Omega = \omega / T$$

Consequently, $\sigma < 0$ implies that *0<r<1* and $\sigma > 0$ implies that *r>1*. When $\sigma = 0$, we have *r=1*. Therefore, the left-half of *p*-plane is mapped inside the unite circle in *z*-plane and right-half of *p*-plane is mapped into points that fall outside the unit circle in *z*-plane. This is one of the desirable properties of a good *p-z* mapping (transformation).

Also $j\Omega$-axis is mapped into the unit circle in *z*-plane as indicates above. However, the mapping of $j\Omega$-axis into the unit circle is not one-to-one. Since ω is unique over the range $< -\pi, \pi >$, the mapping $\omega = \Omega T$ implies that interval $-\pi / T \le \Omega \le \pi / T$ maps into the corresponding values of $-\pi \le \omega \le \pi$. Furthermore, the adjacent strips - frequency interval $\pi / T \le \Omega \le 3\pi / T$ also maps into the interval $-\pi \le \omega \le \pi$ and, in general, so does the interval $(2k-1)\pi / T \le \Omega \le (2k+1)\pi / T$, when *k* is integer. Thus, the mapping from the analog frequency Ω to the frequency variable ω in the digital domain is many-to-one, which simply reflects the effects of aliasing due to sampling. The next figures illustrate the mapping from *p*-plane to *z*-plane. Text: The mapping of $z = e^{pT}$ maps strips of the width $2\pi / T\,(for\,\sigma < 0)$ into the *p*-plane into points in the unit cirle.

From these figures it is clear that for the frequency responses of an analog filters and equivalent digital filter obtained by impulse invariant transformation to correspond, the analog filter must be bandlimited to the range $-\pi / T \le \Omega \le \pi / T$. This generally requires that the analog filter be suitably bandlimited prior to transformation.



5.3.1. Summary

1. Digital filter specification: ω_P, ω_S, δ_1 and δ_2.
2. Transformation of requirements to a digital filter to an analog filter: $\Omega = \omega / T$ ($\Omega_p \to \omega_P, \Omega_S \to \omega_S$, δ_1 and δ_2).
3. Analog filter design: $H_A(p) = \sum_{i=1}^{N} \frac{c_i}{p + d_i}$.
4. Analog filter conversion to a digital filter: $H(z) = T \sum_{i=1}^{N} \frac{c_i}{1 - e^{-d_i T} z^{-1}}$.

Comments on scaling factor T application:

The frequency response of the filter obtained by impulse invariant transformation is given by

$$H(e^{j\omega}) = \frac{1}{T} \sum_{k=-\infty}^{\infty} H_A(\Omega - k\Omega_S)$$

Under condition that

$$\left| H_A(\Omega - k\Omega_S) \right| \sim 0 \; for \left| k \right| > 0$$

we can obtain,

$$H(e^{j\omega}) \sim \frac{1}{T} H_A(\Omega)$$

If it is desired to get a digital filter with the same gain as the analog filter possesses, it is necessary to transform the expression for *H(z)* given by

$$H(z)=\sum_{i=1}^{N}\frac{c_i}{1-e^{-d_iT}z^{-1}}$$

in form

$$H(z)=T\sum_{i=1}^{N}\frac{c_i}{1-e^{-d_iT}z^{-1}}$$

5.4. The Matched Z-Transform

Another method for converting an analog filter into an digital filter is to map the poles and zeros of $H_A(p)$ directly into poles and zeros in the *z*-plane. Suppose that the system function of the analog filter is expressed in the factored form

$$H_A(p)=\frac{\sum_{k=0}^{M}b_kp^k}{\sum_{k=0}^{N}a_kp^k}=\frac{\prod_{k=1}^{M}(p-z_k)}{\prod_{k=1}^{N}(p-p_k)}$$

where z_k are the zeros and p_k are the poles of the filter. Then the transfer (system) function for digital filter is

$$H(z)=\frac{\prod_{k=1}^{M}\left(1-e^{z_kT}z^{-1}\right)}{\prod_{k=1}^{N}\left(1-e^{p_kT}z^{-1}\right)}$$

where *T* is sampling interval. Thus each factor *(p-a)* in $H_A(p)$ is mapped into the factor $\left(1-e^{aT}z^{-1}\right)$ i.e.

$$p-a\rightarrow 1-e^{aT}z^{-1}$$

This mapping is called the matched *z*-transformation.

We observe that the poles obtained from the matched *z*-transformation are identical to the poles obtained with the impulse invariance method. However, the two techniques result in different zero positions.

To preserve the frequency response characteristics of an analog filter, the sampling interval in the matched *z*-transformation must be selected properly to yield the pole and zero locations at the equivalent position in the *z*-plane. Thus aliasing must be avoided by selecting *T* sufficiently small.

Although the matched *z*-transformation is easy to apply, there are many cases when it is not a suitable mapping. E.g. if the analog system has zeros with center frequencies greater that half the sampling frequency, their *z*-plane positions will be greatly aliased. Another case where the matched *z*-transformation is unsuitable is where the continuous transfer function is an all-pole system. Then the digital transfer function is an all-pole system that, in any cases, does not adequately represent the desired continuous system. In general, use of impulse transformation is to be preferred over the matched *z*-transformation.

5.5. Bilinear Transformation Method

The IIR filter design techniques described in the previous sections have severe limitations in that they are appropriate only for low-pass filter and limited class of band-pass filters. This limitations is a result of the mapping that converts points in the *p*-plane to corresponding points in the *z*-plane.

In this section we describe a mapping from the *p*-plane to the *z*-plane, called the bilinear transformation, that overcomes the limitation of the other three design methods described previously. The bilinear transformation is a conformal mapping that transforms $j\Omega$-axis into the unit circle in the *z*-plane only once, thus avoiding aliasing of frequency components. Furthermore, all points of left-half *p*-plane are mapped inside the unit circle in the *z*-plane and all points in the right-half *p*-plane are mapped into corresponding points outside the unit circle in the *z*-plane.

The bilinear transformation can be linked to the trapezoidal formula for numerical integration. For example, let us consider an analog linear filter with transfer function

$$H_A(p) = \frac{b}{p+a}$$

This system is also characterized by the differential equation

$$\frac{dy}{dt} + ay(t) = bx(t)$$

Instead of substituting a finite difference for the derivative, suppose that we integrate the derivative and approximate the integral by the trapezoidal formula:

$$\int_{x_1}^{x_2} f(x)dx = \frac{1}{2}(x_2 - x_1)\left[f(x_2) + f(x_1)\right]$$

Thus

$$y(t) = \int_{t_0}^{t} y'(\tau)d\tau + y(t_0) = \int_{nT-T}^{nT} y'(\tau)d\tau + y(nT-T)$$

where $y'(t)$ denotes the derivative of $y(t)$. The approximation of the previous integral by trapezoidal formula is

$$y(nT) = \frac{T}{2}\left[y'(nT) + y'(nT-T)\right] + y(nT-T)$$

Now the differential equation evaluated in $t = nT$

$$y'(nT) = -ay(nT) + bx(nT)$$

We use this expression to substitute for derivative and thus obtain a difference equation for the equivalent discrete-time system. Then, we obtain the following results:

$$y(nT) = \frac{T}{2}\left[-ay(nT) + b(x(nT) - ay(nT-T) + bx(nT-T)\right] + y(nT-T)$$

and

$$\left(1+\frac{aT}{2}\right)y(nT)-\left(1-\frac{aT}{2}\right)y(nT-T)=\frac{bT}{2}\left[x(nT)-x(nT-T)\right]$$

The z-transform of this equation is

$$\left(1+\frac{aT}{2}\right)Y(z)-\left(1-\frac{aT}{2}\right)z^{-1}Y(z)=\frac{bT}{2}\left[X(z)+z^{-1}X(z)\right]$$

Consequently, the system function of the equivalent digital filter is

$$H(z)=\frac{Y(z)}{X(z)}=\frac{\frac{bT}{2}\left(1+z^{-1}\right)}{1+\frac{aT}{2}-\left(1-\frac{aT}{2}\right)z^{-1}}=\frac{\frac{bT}{2}\left(1+z^{-1}\right)}{\left(1-z^{-1}\right)+\frac{aT}{2}\left(1+z^{-1}\right)}=\frac{\frac{bT}{2}}{\frac{\left(1-z^{-1}\right)}{\left(1+z^{-1}\right)}+\frac{aT}{2}}=\frac{b}{a+\frac{2}{T}\frac{\left(1-z^{-1}\right)}{\left(1+z^{-1}\right)}}$$

i.e.

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

Clearly, the mapping from p-plane to the z-plane is

$$p=\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}$$

This is called the bilinear transformation. Similarly, the mapping from the z-plane is the mapping to p-plane is

$$z=\frac{\frac{2}{T}+p}{\frac{2}{T}-p}$$

Although our derivation of the bilinear transformation was performed for a first-order differential equation, it holds in general, for N-th order differential equation.

To investigate the characteristics of the bilinear transformation, let

$$z=re^{j\omega}$$

$$p=\sigma+j\Omega$$

When $p=j\Omega$

$$z = re^{j\omega} = \frac{\frac{2}{T} + j\Omega}{\frac{2}{T} - j\Omega} = \frac{\sqrt{\left(\frac{2}{T}\right)^2 + \Omega^2}}{\sqrt{\left(\frac{2}{T}\right)^2 + (-\Omega)^2}} \frac{e^{jarct\frac{\Omega T}{2}}}{e^{-jarct\frac{\Omega T}{2}}} = e^{j2arct\frac{\Omega T}{2}}$$

From this equation it can be found

$$|z| = r = 1$$

$$\omega = 2arct\frac{\Omega T}{2}$$

$$\Omega = \frac{2}{T}tg\frac{\omega}{2}$$

Then, we can see that the imaginary $j\Omega$-axis is mapped into unit circle. It is also interesting to note that the bilinear transformation maps the point $\Omega = \infty$ into the point $z = -1$ and the point $\Omega = 0$ into the point $z = 1$. Between these limits the angle of *z* varies monotonically from $0\, to\, \pi$. The relationship between frequency variables in the two domains is illustrated in the next figure. We observe that entire range in Ω is mapped only once into range $-\pi \le \omega \le \pi$, the aliasing errors inherent with impulse invariant transformations are eliminated. However, the mapping is highly nonlinear. We observe a frequency compression or frequency warping, as it is usually called, due to the nonlinearity of the arctangent function.

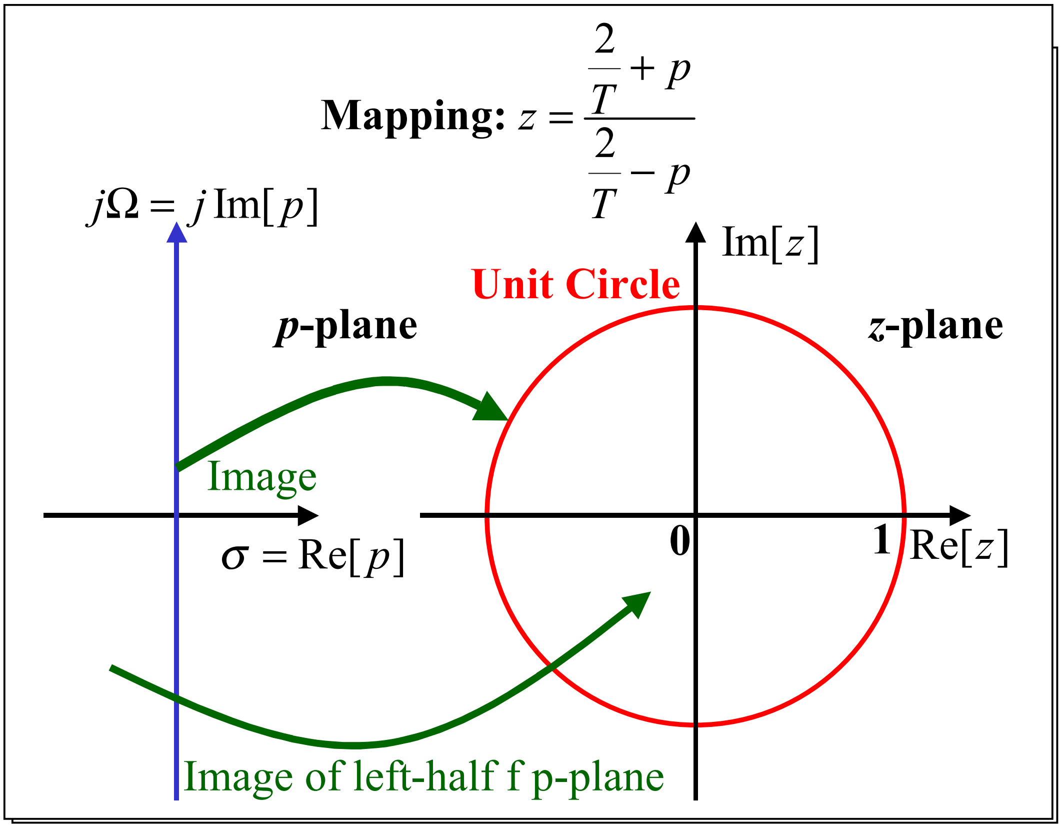
If $p = \sigma + j\Omega$ we obtain for *z*

$$z = \frac{\frac{2}{T} + \sigma + j\Omega}{\frac{2}{T} - \sigma - j\Omega} = \frac{\sqrt{\left(\frac{2}{T} + \sigma\right)^2 + \Omega^2}}{\sqrt{\left(\frac{2}{T} - \sigma\right) + \Omega^2}} \frac{e^{jarct\frac{\Omega}{2/T+\sigma}}}{e^{jarct\frac{\Omega}{2/T-\sigma}}}$$

i.e.

$$|z| = \frac{\sqrt{\left(\frac{2}{T} + \sigma\right)^2 + \Omega^2}}{\sqrt{\left(\frac{2}{T} - \sigma\right) + \Omega^2}}$$

Then, when $\sigma < 0$ (left-half *p*-plane), we find $|z| < 1$ (inside unit circle) and $\sigma > 0$ (right-half *p*-plane), we find $|z| > 1$ (outside unit circle). The next figures illustrate the mapping from *p*-plane to *z*-plane.



5.5.1. Summary

The overall bilinear transformation method of design is as follows:

1. For each of the desired critical digital frequencies (e.g. ω_P and ω_S), determine the corresponding analog frequencies, using

$$\Omega = \frac{2}{T} tg \frac{\omega}{2}$$

2. Given the prewarped analog frequencies, derived the appropriate analog prototype transfer function $H_A(p)$.
3. Apply the bilinear transformation

$$p = \frac{2}{T}\frac{1 - z^{-1}}{1 + z^{-1}}$$

to $H_A(p)$ to obtain the desired digital filter transfer function.