CHAPTER 6

DIGITAL FILTERS II: RECURSIVE FILTERS AND SIMULATING LINEAR SYSTEMS

6.1 Introduction

We now turn to another kind of digital filter: one that will allow us to use a computer to imitate what some physical system does. We might need this when, for example, we want to model a seismogram. The first step would be to have a way of computing the the ground motion input to the seismometer. The next step would be to simulate what the output of the seismometer (a physical system) is for this ground motion. The FIR filters of the previous chapter are not well-suited for this, but other designs are, and it is these we will now describe.

But, to discuss these filters and how to design them we need to first spend some time introducing additional mathematics for the linear systems we discussed in Chapter 2. We will then return to the problem of designing digital filters, introducing what is known as a **recursive filter**; finally, we will show how to make such a digital filter accurately simulate an analog system.

6.2 Lumped-Parameter Systems

We saw in Chapter 2 that a linear time-invariant system could be characterized in three different ways:

- By its frequency response $\tilde{g}(f)$; this expresses the ratio (a complex number) between output and input when the input is a pure sinusoid, $e^{2\pi i f t}$.

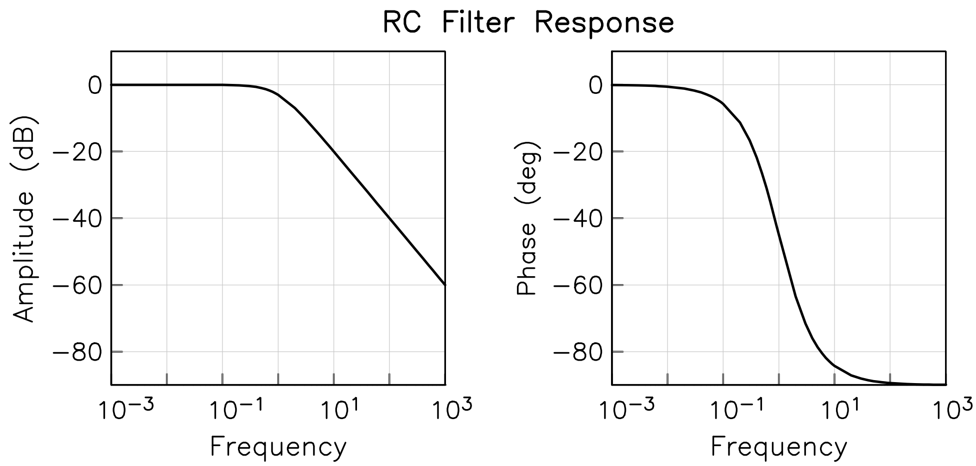


Figure 6.1: Frequency response of an RC filter with $\tau = 1$, shown as log amplitude, and phase, plotted against log frequency.

.

- By its impulse response $g(t)$: the time-domain behavior when the input is a delta-function. This is also the inverse Fourier transform of $\tilde{g}(f)$: $gfo(f) = \mathscr{F}[g(t)]$.

- By the step response $h(t)$, which is the integral of $g(t)$.

We now look at these characterizations for a special class of linear time-invariant systems – though one that includes many actual cases. This class is systems that are described by constant-coefficient linear differential equations; a common term, especially in electrical engineering, for these is **lumped-parameter systems**. For such systems the relationship between the input x and output y is

$$
\begin{aligned}
a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \ldots + a_1 \frac{dy}{dt} + a_0 y = \\
b_m \frac{d^m x}{dt^m} + b_{m-1} \frac{d^{m-1} x}{dt^{m-1}} + \ldots + b_1 \frac{dx}{dt} + b_0 x
\end{aligned}
\tag{6.1}
$$

The a's and b's are the parameters that describe the system.

There are standard procedures for solving such an equation to get explicit forms for $y(t)$, especially for certain classes of inputs $x(t)$; we shall instead look at the frequency response of such systems. But first we provide a few examples.

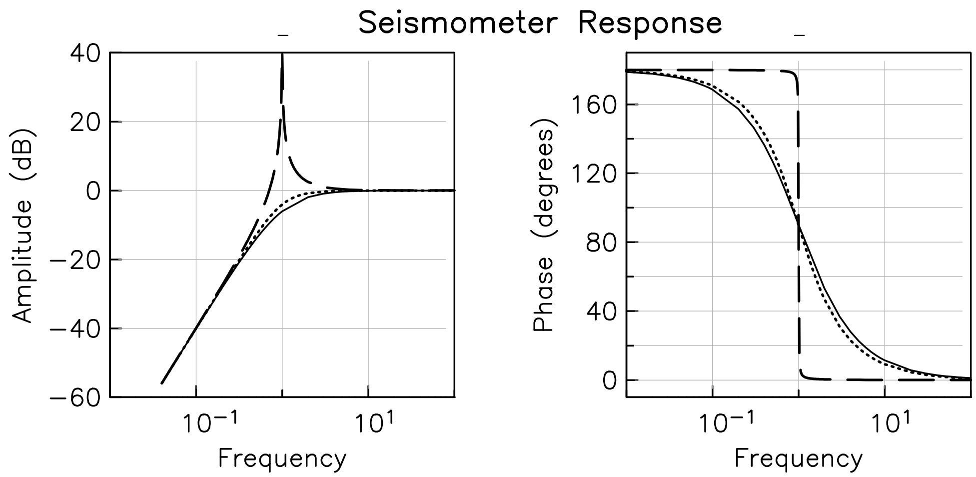
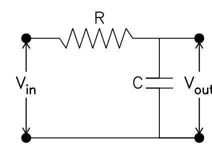


Figure 6.2: Frequency response of a seismometer to displacement. The natural frequency is $\omega_0 = 1$. The response is shown for $\lambda = 0$ (dashed), $\lambda = 0.8$ (dotted), and $\lambda = 1$ (solid).



- The first example is the simplest analog electronic filter: the RC lowpass filter, which consists of a resistor and capacitor in series, with input and output voltages as shown. The input voltage $x(t) = V_{in}$ is just the output voltage $y(t) = V_{out}$ plus the voltage drop across the resistor, which is given by $RI(t)$, where $I(t)$ is the current flowing through the capacitor. This current is given by $I(t) = C\dot{y}(t)$, making the differential equation

$$y(t) + RC\frac{dy}{dt}(t) = x(t) \tag{6.2}$$

To get the frequency response, we assume, as usual, a sinusoidal input $x(t) = e^{2\pi ift}$; by definition of $\tilde{g}$, $y(t) = \tilde{g}(f)e^{2\pi ift}$. Substituting these expressions into (6.2), we get

$$\tilde{g}(f) = \frac{1}{1 + 2\pi if\tau} \tag{6.3}$$

where $\tau = RC$ is the time constant of the filter. The two plots in Figure 6.1 show the amplitude and phase of $\tilde{g}(f)$ for τ equal to 1, plotted against the logarithm of the frequency (and with the amplitude plotted in dB); these are called **Bode plots**. Clearly this is a lowpass filter: high frequencies are attenuated.

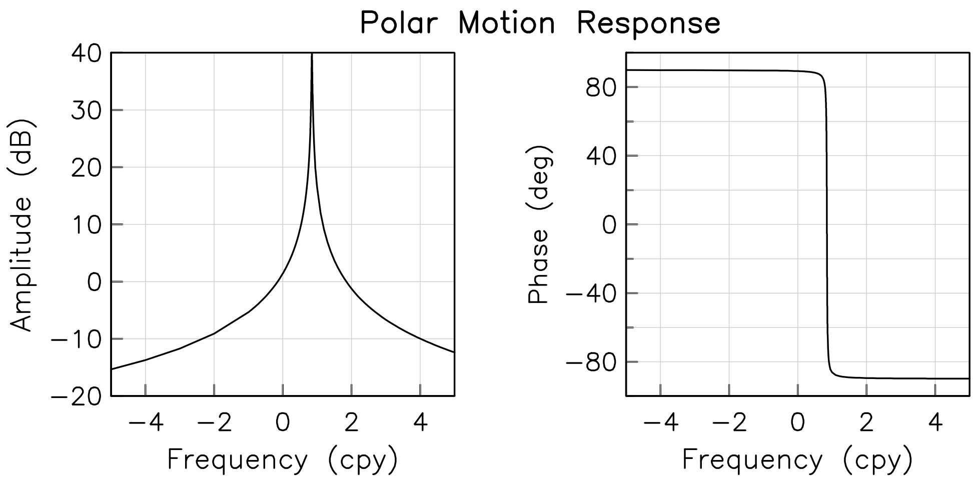


Figure 6.3: Response of the Earth's polar motion to excitation; note that since we represent the input and output (both 2-vectors) as complex numbers. we need to show both positive and negative frequency. The response is given for a Q of 100; the Chandler resonance is at 0.849 cycles/year. The rapid change of phase at this frequency makes it difficult to compare polar motion with excitations.

- A seismometer (Figure 6.2). The simplest form (a mass on a spring, or a pendulum) obeys the well-known equation for a simple harmonic oscillator:

$$\ddot{y} + 2\lambda\omega_0\dot{y} + \omega_0^2 y = \ddot{x}$$

where y is the displacement of the mass with respect to the frame, and x the displacement of the frame with respect to inertial space. The system parameters are the natural frequency ω_0, and the damping λ. Again, we substitute $e^{2\pi ift}$ for $x(t)$ and $\tilde{g}(f)e^{2\pi ift}$ for $y(t)$, and find

$$\tilde{g}(f) = \frac{-4\pi^2 f^2}{-4\pi^2 f^2 + 4i\pi\lambda\omega_0 f + \omega_0^2}$$

whose phase and amplitude response are shown in Bode plots (Figure 6.2) for several values of λ: $\lambda = 0$ is undamped, $\lambda = 0.8$ gives the flattest response, and $\lambda = 1$ is critically damped.

- The Earth's polar motion. The position of the rotation pole of the solid earth changes because of variations in the angular momentum and mass distribution of the fluid parts. The pole position can be described by two coordinates, p_1 and p_2, giving the (dimensionless) angular displacement of (say) the North rotation pole in directions toward the Greenwich meridian and at 90° to it. If these changes are small, it can be shown that the relationship between the pole position and the motions of the fluid parts is

$$\frac{1}{\omega_c}\frac{dp_1(t)}{dt}+p_2(t)=\psi_2(t) \qquad \frac{1}{\omega_c}\frac{dp_2(t)}{dt}+p_1(t)=\psi_1(t) \tag{6.4}$$

where ψ_1 and ψ_2 are integrals over the mass distribution and velocity of the fluid parts of the earth (usually termed the "excitation").[1] The frequency ω_c is determined by the properties of the solid earth; for a rigid ellipsoidal body $\omega_c = [(C-A)/A]\Omega$, where C and A are the polar and equatorial moments of inertia, and Ω is the frequency of the Earth's spin (once per day). The values of C and A for the Earth give a value for ω_c that corresponds to a period of 305 days; various corrections, too complicated to discuss here, make the actual period equal to 430 days. We can write this pair of equations as a single equation by forming the complex variables $\mathbf{p} = p_1 + ip_2$ and $\psi = \psi_1 + i\psi_2$; then we can combine the two equations in (6.4) to get:

$$\frac{i}{\omega_c}\dot{\mathbf{p}}(\mathbf{t})+\mathbf{p}(\mathbf{t})=\psi(\mathbf{t}) \tag{6.5}$$

If we make the input eft, and the output $\tilde{g}eft$, we find

$$\tilde{g}(f)=\frac{1}{1-(2\pi f/\omega_c)}$$

which has the response shown in Figure 6.3, in this case plotted against both positive and negative frequency, since the response is different for these: a common attribute of gyroscopic systems. (Remember that, with this trial function, negative frequency corresponds to clockwise [deasil] rotation, and positive to counterclockwise [widdershins].)

We will use these examples in the discussion below, both to illustrate concepts about systems, and also as examples for filter design.

[1] For a derivation of this equation, see Munk and McDonald (1960); the version given here has been revised to match the newer and more accurate result of Gross (1992).

6.3 The Laplace Transform: System Poles and Zeros

Yet another way of looking at the response of a system comes if we look at another integral transform: the **Laplace transform**. The Laplace transform of a function $x(t)$ is defined as

$$\tilde{x}(s) = \int_{-\infty}^{\infty} x(t)e^{-st}\,dt$$

for s a complex number: $\tilde{x}(s)$ is defined on the complex plane. The Fourier transform is $\tilde{x}(f)$ is just $\tilde{x}(s)$ evaluated along the imaginary axis of the complex plane; that is, $\tilde{x}(f) = \tilde{x}(s)$ for $s = 2\pi i f$. The Laplace transform can thus be thought of as a generalization of the Fourier transform; but, viewed as a transform, it is nowhere near as useful. And, mostly because $\tilde{x}(s)$ is a complex-valued function over the complex plane, the Laplace transform is much more difficult to visualize than the Fourier transform. The inverse Laplace transform is

$$x(t) = \frac{1}{2pi}\int_{c-i\infty}^{c+i\infty} \tilde{x}(s)e^{st}\,ds$$

where the integral is evaluated over a line parallel to the imaginary axis, the value of the real part depending on the nature of $\tilde{x}(s)$. The Laplace transform and its inverse are thus much less symmetrical than the Fourier transform pair.

The utility of the Laplace transform for us comes through applying it to the lumped-parameter systems described by equation (6.1). The value of the Laplace transform is that it enables us to get significant understanding of such a system without solving the differential equation at all.[2]

Taking the Fourier transform of both sides of equation (6.1) gives us the same result as we would get (and did get for the examples of the previous section) by substituting $e^{2\pi i f t}$ as the input. What we mean by "taking the Fourier transform" of the system is to find the equation that describes the connection between the Fourier transforms of the input and output: that is, between $\tilde{x}(f)$ and $\tilde{y}(f)$. We get this result by applying the theorem for the Fourier transform of the derivative of a function to equation (6.1); this gives us a polynomial in f:

$$[a_n(2\pi i f)^n + a_n - 1(2\pi i f)^n - 1 + \ldots + a_1(2\pi i f) + a_0]\tilde{y} =$$
$$[b_m(2\pi i f)^m + b_m - 1(2\pi i f)^m - 1 + \ldots + b_1(2\pi i f) + \ + b_0]\tilde{x}$$

[2] Deakin (1992) describes how this approach grew to dominance.

where $\tilde{y}$ and $\tilde{x}$ are the Fourier transforms of y and x. Therefore, the frequency response of the system is

$$\tilde{g}(f) = \frac{\tilde{y}(f)}{\tilde{x}(f)} = \frac{\sum_{j=0}^{m} b_j (2\pi i f)^j}{\sum_{k=0}^{n} a_k (2\pi i f)^k} \tag{6.6}$$

We can follow a similar route if we take the Laplace transform of both sides of (6.1). While this might seem like an unnecessary generalization (who needs the system response at a complex frequency?) it actually leads to some very useful insights into how the system will behave. We need the derivative theorem for the Laplace transform, with we simply state without proof: if $\tilde{x}(s)$ is the Laplace transform of $x(t)$, the Laplace transform of $\dot{x}$ is $s\tilde{x}(s)$.

Applying this rule to the differential equation shows that the ratio of the Laplace transforms of the input and output, which is called the **transfer function**, is

$$\frac{\tilde{y}(s)}{\tilde{x}(s)} = \frac{\sum_{j=0}^{m} b_j s^j}{\sum_{k=0}^{n} a_k s^k} \tag{6.7}$$

Comparing this with equation (6.6), we see that the frequency response is a special case of the transfer function: the frequency response $\tilde{g}(f)$ is just the transfer function $\tilde{g}(s)$ evaluated on the imaginary axis.

The additional insight to be gotten from looking at the transfer function comes if, instead of expressing the polynomials in (6.7) in terms of their coefficients, we instead write them as products of their roots:

$$\frac{\tilde{y}(s)}{\tilde{x}(s)} = C \frac{\prod_{j=1}^{m} (s - r_j)}{\prod_{k=1}^{n} (s - p_k)}$$

The roots of these polynomials in s come in two forms, the **zeros** r_j of the numerator, and the **poles** p_k of the denominator. At the zeros the transfer function is zero; and at the poles it is infinite. The scaling value, C, is needed to make the description complete. Looking at the locations of the poles and zeros, and particularly how close they are to the imaginary axis, will show where the frequency response will be large and where it will be small. Finding the locations for either set of roots has to be done numerically if m or n exceeds five, but our three examples can all be done analytically:

- The RC filter has a transfer function

$$\tilde{g}(s) = (1 + \tau s)^{-1} \tag{6.8}$$

which has a single pole on the negative real axis at $s = \tau^{-1}$. This is shown in the left-hand plot of Figure 6.4; as is conventional, we use a cross to show the location of a pole.

- The seismometer transfer function is

$$\tilde{g}(s) = \frac{s^2}{s^2 + \lambda\omega_0 s + \omega_0^2}$$

which has a pair of zeros at the origin of the s-plane. The poles lie at $\omega_0[-1 \pm \sqrt{\lambda^2 - 1}]$. For $\lambda = 0$, the poles are on the imaginary axis, at $\pm i\omega_0$. Since this axis corresponds to the frequency axis for the Fourier transform, these on-axis poles make $\tilde{g}(\pm\omega_0$ infinite. As λ increases from zero, the two poles leave the imaginary s-axis and follow a circular path about the origin: because the coefficients of the polynomial are real, its roots (these poles) must be complex conjugates. The case of $\lambda = 0.8$, which gives the flattest response, would be termed a two-pole Butterworth filter. At $\lambda = 1$ the poles meet; for larger values they lie on the negative real axis, one approaching and the other receding from the origin.

- The transfer function for polar motion is

$$\tilde{g}(s) = \frac{\mathbf{P(s)}}{\Psi\mathbf{(s)}} = \frac{\frac{1}{1+is}}{\omega_c} = \frac{\omega_c}{\omega_c + is}$$

which has a single pole, on the imaginary axis, at $s = i\omega_c$. Again, this gives an infinite response for $f = \omega_c$: in the Earth this resonance gives rise to the Chandler wobble, hence the subscript c on ω. In reality dissipation keeps the response from being infinite; we can add this effect to the equations by moving the pole of the transfer function away from the imaginary axis. We make this pole complex, rather than imaginary, by defining ω_c as

$$\omega_c = \frac{2\pi}{T_c}\left(1 + \frac{i}{2\pi Q_c}\right)$$

where T_c is now the period of the resonance, and Q_c a measure of the amount of dissipation.[3]

[3] It is certain that dissipation occurs in the Chandler wobble, but what causes it remains unclear. Smith and Dahlen (1981) provide an exhaustive discussion of the theoretical values for these two parameters for a given Earth model.

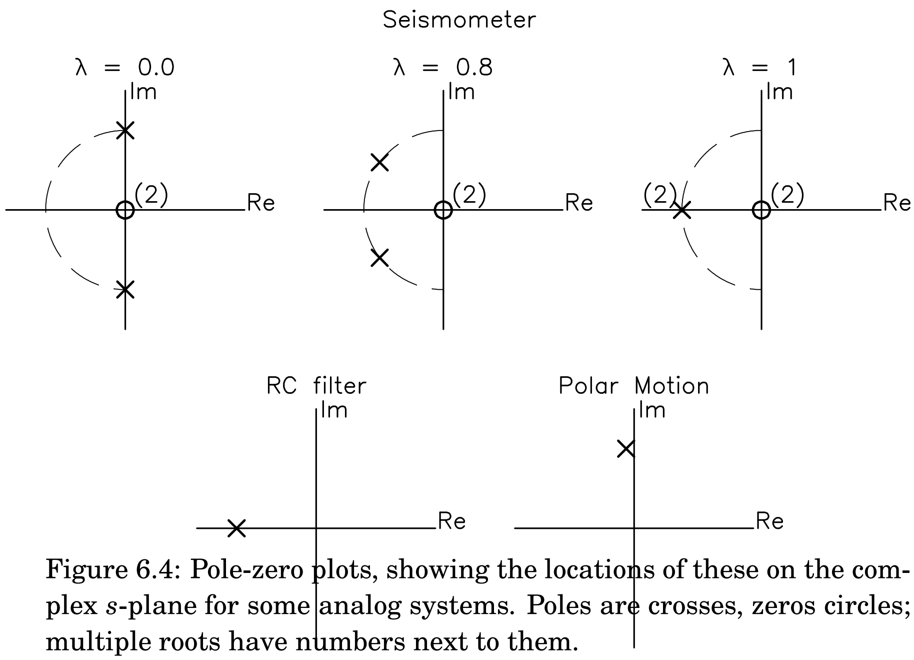


Figure 6.4: Pole-zero plots, showing the locations of these on the complex s-plane for some analog systems. Poles are crosses, zeros circles; multiple roots have numbers next to them.

For all our examples, including the last one with finite Q_c, the poles are left of the imaginary axis. This is good, because for any lumped-parameter system, stability requires that all the poles of the transfer function have negative real parts, lying in the left half-plane. And for an unstable system, any nonzero input leads to an infinite output. Whether or not a system is stable is not obvious from looking at the differential equation; but finding the poles shows this immediately. Likewise, looking at the poles we can easily see what at what frequencies the response is large: it will be those that are near poles that are close to the imaginary axis. Conversely, zeros close to the imaginary axis will produce dips in the amplitude response.

6.4 The z-transform for Digital Filters

In Chapter 3 we examined the Fourier transform for sequences. We have just seen that the Laplace transform can be regarded as the generalization of the Fourier transform; the equivalent generalization for sequences is called the **z-transform**. An infinite sequence x_n, has a z-transform $\zeta[x_n]$

that is a function of the complex variable z:

$$\zeta[x(n)] = \tilde{x}(z) = \sum_{n=-\infty}^{\infty} x_n z^{-n}$$

It is important to note that this is *not* a definition that is universally followed. In particular, the geophysical exploration industry (which does a lot of signal processing) defines the z-transform as a sum over $x_n z^n$, Our usage is that found in electrical engineering; if we call our z-variable z_{EE}, and the exploration one z_{oil}, $z_{\mathrm{EE}} = z_{\mathrm{oil}}^{-1}$. Since z is a complex variable, this usage is equivalent to an inversion of the complex plane on the unit circle: the outside of the circle becomes the inside, and vice-versa, with the origin and infinity mapping into each other. Either convention is consistent, but in looking at results (and software) you need to know which one is being used.

We have seen that the Fourier transform can be viewed as the Laplace transform evaluated on the imaginary axis of the s-plane. In just the same way, the Fourier transform of a sequence is a special case of the z-transform. Remember that the Fourier transform of a sequence is

$$\tilde{x}(f) = \sum_{n=-\infty}^{\infty} x_n e^{-2\pi i f n} \tag{6.9}$$

This is equivalent to the z-transform for $z = e^{2\pi i f}$; in words,

> The Fourier transform of a sequence is the z-transform evaluated on the unit circle in the complex z-plane.

There are other ways of looking at the z-transform. We can, for example, regard z^{-1} as a unit delay. This may sound peculiar, but is easily derived. Suppose y is the same series as x, but delayed by m terms: $y_n = x_{n-m}$; then the z-transform of y is just

$$\tilde{y}(z) = \sum_{n=-\infty}^{\infty} x_{n-m} z^{-n} = \sum_{l=-\infty}^{\infty} x_l z^{-(l+m)} = \tilde{x}(z) z^{-m} = \tilde{x}(z)(z^{-1})^m \tag{6.10}$$

so that the effect of delaying x is to multiply the z-transform by z^{-1} a total of m times: z^{-1} "represents" a delay. (That is, z_{EE}^{-1} does; in the usage of the exploration industry, z_{oil} does.)

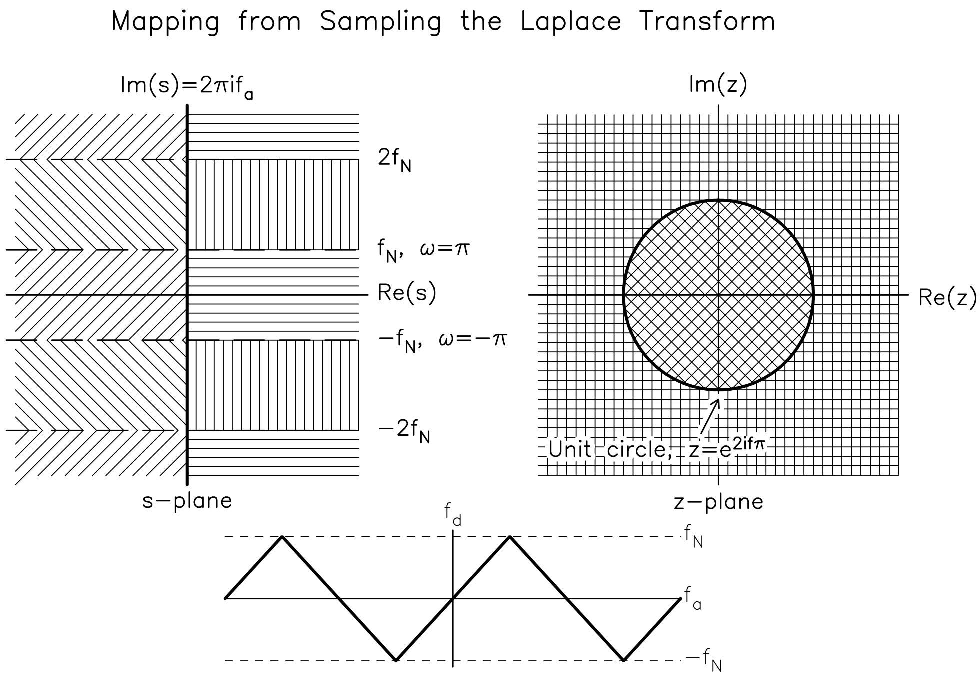


Figure 6.5: Mapping between the s-plane and the z-plane if we get our sequence in discrete time by sampling in continuous time. The right-hand side of the s-plane maps outside the unit circle in the z-plane, and the left-hand side to the inside of the unit circle, but in both cases the mapping is not one-to-one. The bottom plot shows how the continuous-time frequency f_a maps into discrete-time frequency f_d.

Another way to look at the z-transform comes from taking the Laplace transform of the function which is equivalent to the sequence x_n, namely $\sum_{n=-\infty}^{\infty} x(n)\delta(t-n)$, which we used in Chapter 4 to discuss sampling of a function given in continuous time. The Laplace transform of this infinite sum of delta-functions is

$$\int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x_n \delta(t-n) e^{-st}\, dt = \sum_{n=-\infty}^{\infty} x_n \int_{-\infty}^{\infty} e^{-st} \delta(t-n)\, dt = \sum_{n=-\infty}^{\infty} x_n e^{-sn}$$

which becomes equivalent to the z-transform if we set

$$z = e^s \tag{6.11}$$

which for any s gives a value z. This derivation of the z-transform thus introduces a **mapping** from the s-plane to the z-plane; as we will see below, such mappings are important in devising digital (discrete-time) models of continuous-time systems. The particular mapping given by equation (6.11) has an important feature, namely that it is nonunique: different points in the s-plane all map to the same point in the z-plane.

If we consider the strip for which $|\Im(s)| \le \pi$, we can see that the region of the strip with negative real part maps inside the unit circle, the region with positive real part maps to outside the unit circle, and the segment of the imaginary axis that lies in the strip maps onto the unit circle.[4] But exactly the same mapping occurs for each strip defined by $m - \pi \le \Im(s) \le m + \pi$, for any integer m. This is yet another example of aliasing: when we form a sequence by sampling a function, one frequency in the sequence can come from many different frequencies in the function; with the mapping (6.11) we have generalized this to complex frequencies.

It is possible to use complex-variable theory to find the inverse of the z-transform, and also to derive extensions to the convolution theorems in order to show, for example, how to convolve two z-transforms to get the z-transform of the product of two sequences. Since these results are not needed to design filters for system simulation, we do not discuss them further; see Oppenheim and Schafer (1989) for the details.

6.5 Recursive Digital Filters

In a FIR filter, the output is just a weighted combination of input values. This kind of filter is inadequate for simulating an analog system, for which purpose we introduce a more general form: the **recursive filter**, where the output value depends on previous values of the output, as well as on the input values. The equation for such a digital filter is

$$\sum_{k=0}^{N-1} a_k y_{n-k} = \sum_{l=0}^{L-1} b_l x_{n-l} \tag{6.12}$$

where x_n is the input sequence and y_n is the output. Because the sums are one-sided, we can compute the "current" value of y, y_n, without needing any values of y except for the "past" ones that we have already computed. Using only past values of x is not necessary except in real-time processing; but is conventional and not restrictive in practice.

[4] In geophysical-exploration usage, the mappings to the inside and outside of the circle would be reversed.

Filters of this type, like the nonrecursive filters of the previous chapter, go by several different names: electrical engineers call them **Infinite Impulse Response** (IIR) filters, and statisticians call them **auto-regressive moving-average** (ARMA) systems. If $L = 1$, so that only the current value of the input is used, the statisticians' term is that the system is **autoregressive** (AR).

Just as we used the Laplace transform to find the transfer function of a differential equation, we may use the z-transform to find a transfer function for a discrete-time filter. If we take the z-transforms of both sides of equation (6.12), equation, and use the result (6.10) that the z-transform of a delayed sequence is z^{-m} times the transform of the original sequence, the two sides of (6.12) become polynomials in z. The transfer function becomes the ratio of these polynomials:

$$\frac{\tilde{y}(z)}{\tilde{x}(z)} = \frac{\sum_{l=0}^{L-1} b_l z^l}{\sum_{k=0}^{N-1} a_k z^k} \tag{6.13}$$

This is actually a general equation for any type of digital filter, recursive or nonrecursive; note that for zero phase shift FIR filters the range of summation will include negative as well as positive values of l.

The parallel between (6.12) and (6.13) on the one hand, and the differential equation (6.1) and its transfer function (6.7) is quite intentional: this parallelism shows how to reach our goal of simulating a system with a digital filter. The simulation problem now becomes how to find the coefficients of the filter (6.12), with z-transform (6.13), that will best imitate the behavior described by the differential equation (6.1) with transfer function (6.7).

We start with our simplest example, the lowpass RC filter, and begin with the step response. If x is constant and equal to x_0, the solution to the differential equation (6.2) is $y = x_0$. If at time $t = 0$, x then becomes zero, the solution for y for $t > 0$ is $y = x_0 e^{-t/\tau}$. An FIR filter cannot reproduce the infinitely long response to a step that this system shows; even approximating it would require a large number of weights. But a very simple recursive filter can give a similar response. This is

$$y_n - a y_{n-1} = x_n(1-a)$$

For x constant this gives a constant y equal to x. If x_n is 0 for $n \le 0$, and 1 for $n \ge 1$, it is easy to compute y:

$$y_0 = 1 \quad y_1 = a \quad y_2 = a^2 \qquad y_m = a^m = e^{m \ln a}$$

giving the same exponential falloff as in the analog filter. The transfer function of this digital filter is

$$\tilde{y}(z)(1-az^{-1}) = \tilde{x}(z)(1-a) \quad \text{or} \quad \frac{\tilde{y}(z)}{\tilde{x}(z)} = \frac{1-a}{1-az^{-1}}$$

which has the frequency response (using $z = e^{2\pi i f}$)

$$\frac{1-a}{1-ae^{-2\pi i f}} \tag{6.14}$$

where f is now the digital frequency, running over the range from $-\frac{1}{2}$ to $\frac{1}{2}$ (the range between the Nyquist frequencies). Clearly the analog response (6.3) and the digital response (6.14), though similar, are not the same; the problem we will now address is to devise a recursive filter that will be closer.

However, before doing so, we need to discuss the stability criterion for digital filters: what makes them stable or not. If we look at the output of the discrete-time filter for the case given above (equation (6.13), we see that for a less than 1, the output decays when x goes to zero (from one); but if a were greater than 1, the output would increase exponentially. The value of a thus determines if the filter is stable or unstable, the boundary being $a = 1$, for which the filter is metastable. In the z-plane, the one filter pole is at $z = a$, so the condition for stability is that the pole is inside the unit circle. This can be shown to be true in general. Just as a continuous-time system, to be stable, must have all poles on the negative real half of the s-plane, for a discrete-time filter, all its poles must fall within the unit circle. These criteria matter because we want any mapping from the s-plane to the z-plane to preserve stability: the "stable" part of each plane should map only into the stable part of the other. The mapping given by equation (6.11), corresponding to sampling, fulfills this criterion; some others do not. For example, creating discrete-time systems by replacing derivatives by forward differences does not; if we perform this substitution on a stable differential equation, the resulting discrete-time filter may not be stable.

A better mapping than (6.11) for system simulation is called the **bilinear transformation**. To motivate it, we again construct a discrete-time system from the differential equation, but instead of using numerical differences to approximate derivatives, we use numerical approximations to integrals.

For example, we can re-express the differential equation for an RC low-pass filter in integral form as:

$$y(t) = \int_{t_0}^{t} \dot{y}(u)du + y(t_0)$$

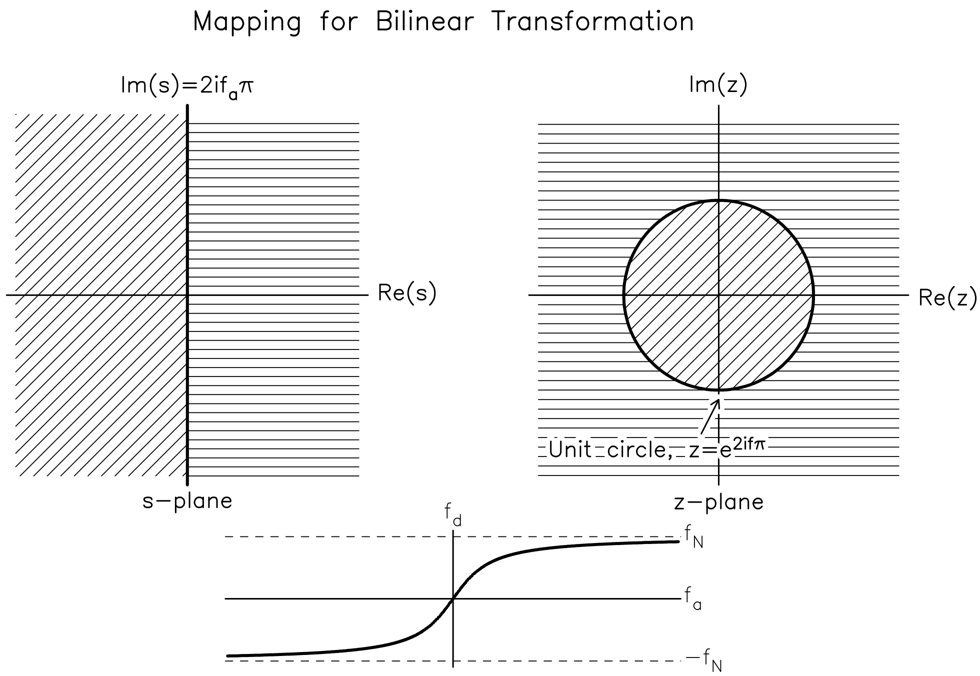


Figure 6.6: Mapping between the s-plane and the z-plane if the bilinear transformation is used. The right-hand side of the s-plane maps outside the unit circle in the z-plane, and the left-hand side to the inside of the unit circle; in both cases the mapping is one-to-one. The bottom plot shows how the continuous-time frequency f_a maps into discrete-time frequency f_d: the mapping is one-to-one, but nonlinear, though approximately linear for $f_a \ll f_n$.

For equispaced intervals, which we set equal to 1, this becomes

$$y(n) = \int_{n-1}^{n} \dot{y}(u)du + y(n-1)$$

We now use the trapezoidal rule for this integral to write

$$y(n) = \tfrac{1}{2}[\dot{y}(n) + \dot{y}(n-1)] + y(n-1)$$

The differential equation itself gives an expression for $\dot{y}$: $\dot{y}(t) = \tau^{-1}[x(t) - y(t)]$. We may then write the above expression for $y(n)$ as

$$y(n) = \frac{1}{2\tau}[x(n) + x(n-1) - y(n) - y(n-1)] + y(n-1)$$

which we may, finally, write in the form (6.12) of a recursive filter; using subscripts rather than arguments since we are now working with a sequence, we get

$$y_n\left(1+\frac{1}{2\tau}\right)-y_{n-1}\left(1+\frac{1}{2\tau}\right)=\frac{1}{2\tau}(x_n+x_{n-1})$$

The transfer function in the z-plane of this discrete-time filter is found in the usual way; we take z-transforms to get

$$\tilde{y}(z)\left[\left(1+\frac{1}{2\tau}\right)-z^{-1}\right]=\tilde{x}(z)[1+z^{-1}]$$

whence the transfer function is

$$\frac{1}{1+2\tau}\left(\frac{1-z^{-1}}{1+z^{-1}}\right)$$

We can equate this with the transfer function for the continuous-time case, equation (6.8), if we make the relationship between s and z

$$s=2\left(\frac{1-z^{-1}}{1+z^{-1}}\right) \tag{6.15}$$

which is the bilinear transformation.

In the more general case where the sampling interval is Δ, the same derivation may be used to give the same result, with the leading 2 replaced by $2/\Delta$. The inverse mapping from s to z has a similar form, namely

$$z=\frac{1+\frac{1}{2}s\Delta}{1-\frac{1}{2}s\Delta} \tag{6.16}$$

While our derivation used a particular differential equation, using multiple integrals for any constant-coefficient linear differential equation will produce the same mapping. This mapping takes the imaginary axis of the s-plane onto the unit circle in the z-plane; since it maps the entire left half of the s-plane into the inside of the unit circle on the z-plane, it maintains stability.

The bilinear transformation thus gives us a design procedure for getting discrete-time filters from continuous-time ones. The steps are:

1. Find the transfer function in the s-plane from the differential equation of the system; this will be a ratio of polynomials in s.

2. Perform the bilinear mapping (6.16) to produce a transfer function in the z-plane; reduce this to a ratio of polynomials in z.

3. Get the filter weights from the coefficients of the polynomials, as in the relation between (6.12) and (6.13).

There is often a step before step 1: we may have to modify the differential equation to minimize a distortion from the bilinear mapping. Let the frequency of the discrete-time system be f_d, so that (equation (6.9)) $z = e^{2\pi i f_d}$, and of the analog system be f_a, so that $s = 2\pi i f_a$. Then

$$\frac{1-z^{-1}}{1+z^{-1}} = i\tan(\pi f_d)$$

so that the bilinear transformation of the frequencies is

$$f_d = \frac{1}{\pi}\arctan(\pi f_a \Delta)$$

which maps the entire analog frequency axis, from $-\infty$ to ∞, onto the digital frequencies from $-\frac{1}{2}$ to $\frac{1}{2}$: that is, from the lower to the upper Nyquist frequencies. This warping of the frequency axis, though it avoids aliasing of the filter response, means that we must adjust the time constants of the differential equation so that after applying the bilinear transformation these have the correct frequency, an adjustment known as **prewarping**.

To show how to do this, we design a filter to simulate the polar-motion equation (6.4). If we apply the bilinear transformation (6.15) to the transfer function (6.5), and compute the frequency response of the resulting digital filter, we find that it is

$$\frac{\omega_c'}{\omega_c' - (2/\Delta)\tan(\pi f_d)}$$

where we have used ω_c' to denote that this parameter is not necessarily the same one as in the actual continuous-time system. For ω_c real, the frequency of the resonance in the continuous-time system is at $f_a = \omega_c/2\pi$; when the series is sampled at an interval Δ, this becomes the (nondimensional) frequency $\omega_c\Delta/2\pi$, providing it is not aliased. In the discrete-time system the resonance will be at

$$f_d = \frac{1}{\pi}\arctan\left(\frac{\omega_c'\Delta}{2}\right)$$

Equating these two digital frequencies gives us the relationship

$$\omega'_c = \frac{2}{\Delta}\tan\left(\frac{\Delta}{2}\omega_c\right)$$

for the time constant (in this case a frequency) used in the discrete-time filter as a function of the one actually present in the system. If $\omega_c\Delta \ll 1$, the two are nearly equal; but as this product approaches π more and more correction is needed. In this particular example, the usual sample interval Δ for $\mathbf{p}$ and ψ is 30 days; for $\omega_c = 0.01461$rad/day, $\omega'_c = 0.01485$, only a 2% change.

Having made this correction, we can compute the actual filter weights by the procedure just described. We find that the polynomial in z is

$$\frac{\omega'_c\Delta + \omega'_c\Delta z^{-1}}{(\omega'_c\Delta + 2i) + (\omega'_c\Delta - 2i)z^{-1}}$$

which gives a recursive filter

$$\mathbf{p_n}(\omega'_c\Delta + 2i) = -\mathbf{p_{n-1}}(\omega'_c\Delta - 2i) + \omega'_c\Delta(\psi_\mathbf{n} + \psi_\mathbf{n-1})$$

for producing simulated polar motion from a possible excitation.

CHAPTER 7

DIGITAL FILTERS III: DIFFERENTIATORS, LEAST-SQUARES, AND MINIMUM-PHASE FILTERS

7.1 FIR Filters: Other Applications

In this chapter we look at some other uses for FIR filters than removing bands of frequencies: differentiation, Hilbert transforms, and linear regression. We also describe the most important class of non-symmetric FIR filters for frequency selection, the minimum-phase filters; these are sometimes important for use in seismology.

7.2 Differentiators and Digital Hilbert Transformers

We have mentioned before (Section 2.8) that differentiation can be thought of as a linear system; this makes it appropriate for digital filtering. Equation (2.5)

$$\mathscr{F}[\dot{x}(t)] = 2\pi i f \mathscr{F}[x(t)]$$

shows that differentiation is a filter with frequency response $2\pi i f$: that is, a constant phase shift of $\pi/2$ (the i part of the coefficient) and an amplitude response that is proportional to frequency.

To imitate this digitally, we make the number of weights odd, with $N = 2M+1$, and also make them antisymmetric:

$$w_{-n} = -w_n \quad \text{for} \quad n = 0, \ldots M$$