The z-Transform and Flow Graphs

5

5.1 INTRODUCTION TO THE z-TRANSFORM

The z-transform, an important tool for discrete-time signal and system analysis, can be viewed as the discrete-time analog of the Laplace transform. It provides another domain in which signals and systems can be examined. The z-transform is very closely related to the discrete-time Fourier transform and finds its principal use in the analysis and manipulation of linear, constant coefficient difference equations.

The *two-sided* or *bilateral* z-transform is defined as the summation

at $z = e^{j\omega}$ you have DTFT

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n}. \tag{5.1}$$

Although other definitions are useful for some purposes, this is the most useful in digital signal processing and, consequently, is the only one that we will use in this text. The summation in equation (5.1) is a power series in the complex variable z^{-1}, the coefficients of which are the samples of the sequence $x[n]$. As with any complex variable, values of z can be associated with points in a plane. A point z, plotted in the z-plane as in Fig. 5.1, has as its ordinate the imaginary part of z and as its abscissa the real part of z. It is also convenient to express points in the z-plane in polar coordinates in which a value is parameterized by its distance from the origin and its angle measured relative to the real axis.

Two important properties of the z-transform are apparent from its definition in equation (5.1). First the transform is linear. If

$$x[n] \Longleftrightarrow X(z)$$

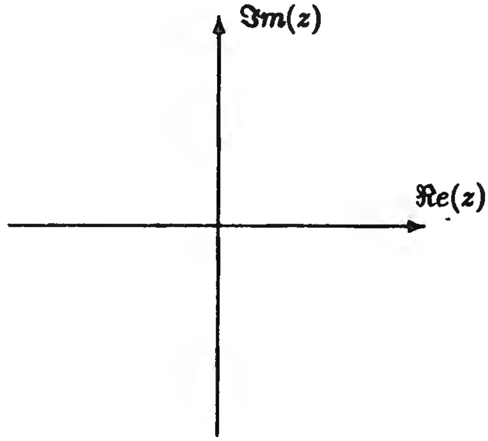

Figure 5.1. The z-plane for representing values of the complex variable z.

and

$$v[n] \Longleftrightarrow V(z),$$

then

$$ax[n] + bv[n] \Longleftrightarrow aX(z) + bV(z).$$

The second property allows us to deal with shifts in $x[n]$. Delaying $x[n]$ by one sample multiplies its z-transform by z^{-1}, i.e.,

$$x[n-1] \Longleftrightarrow z^{-1}X(z). \tag{5.2}$$

Thus in the z-domain multiplication by z^{-1} corresponds to a delay operation. Similarly, delaying a sequence by m samples multiplies its z-transform by z^{-m}. The shift, m, can be either positive or negative.

The transform pair

$$\delta[n] \Longleftrightarrow 1$$

follows immediately from the definition of the z-transform. Evaluating the z-transform of an arbitrary finite length sequence becomes trivial when this relation is used with the linearity and shift properties. For example, the z-transform of

$$x[n] = 2\delta[n+1] + \delta[n] + 2\delta[n-1] + 3\delta[n-2] + 8\delta[n-3]$$

is simply

$$X(z) = 2z + 1 + 2z^{-1} + 3z^{-2} + 8z^{-3}.$$

Sequences that are infinitely long are transformed the same way, but it is frequently possible to write these z-transforms in closed form. This is true, for example, for the impulse responses of linear, time-invariant systems that are described by linear, constant coefficient difference equations. As an example, consider the z-transform of the signal

$$h[n] = \alpha^n u[n].$$

$H(z)$ is found by substituting into the definition in equation (5.1) to obtain

$$H(z) = \sum_{n=-\infty}^{\infty} \alpha^n u[n] z^{-n} = \sum_{n=0}^{\infty} (\alpha z^{-1})^n. \tag{5.3}$$

The geometric series formulas that were discussed in Chapter 2, in particular,

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}, \tag{5.4}$$

can be used to express this summation in closed form provided that

$$|a| < 1. \tag{5.5}$$

If $|a|$ is not less than 1, the summation does not converge (i.e., its value is infinite).

Applying this formula to equation (5.3) in our example gives

$$H(z) = \frac{1}{1-\alpha z^{-1}}, \qquad |z| > |\alpha|.$$

For this to be valid, the convergence condition of equation (5.5) must hold. This implies that this z-transform is defined only for those values of z for which $|\alpha z^{-1}| < 1$ or $|z| > |\alpha|$. These values fall in the portion of the z-plane that is shaded in Fig. 5.2. This region is called the *region of convergence* (ROC). The specification of the ROC is part of the specification of the z-transform. This is made clear in the next example.

As a second example consider the z-transform of

$$x[n] = -\beta^n u[-n-1].$$

Inserting the sequence into the defining summation and manipulating it into the form of a geometric series produces the result

$$\begin{aligned} X(z) &= -\sum_{n=-\infty}^{-1} \beta^n z^{-n} = -\sum_{n=\infty}^{1} \beta^{-n} z^n \\ &= -\sum_{n=1}^{\infty} \beta^{-n} z^n = -\beta^{-1} z \sum_{n=0}^{\infty} (\beta^{-1} z)^n. \end{aligned} \tag{5.6}$$

In closed form this becomes

$$X(z) = -\frac{\beta^{-1} z}{1-\beta^{-1} z} = \frac{1}{1-\beta z^{-1}}.$$

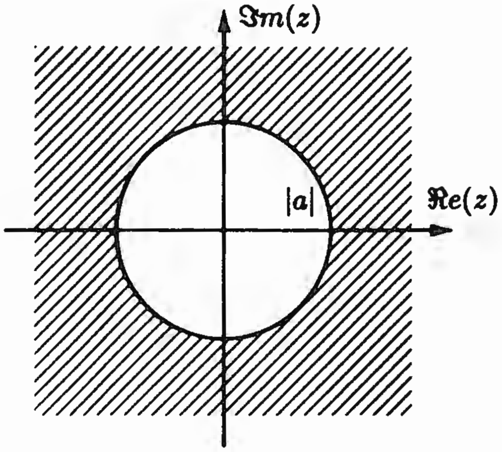

Figure 5.2. The region of convergence in the z-plane for the first example.

The region of convergence is seen to be $|z| < |\beta|$.

These two examples illustrate an important fact. The z-transform of the left-sided sequence $-\beta^n u[-n-1]$ is almost identical to the z-transform of the right-sided sequence $\beta^n u[n]$. Only their regions of convergence are different. For the left-sided sequence the ROC is $|z| < \beta$; for the right-sided sequence it is $|z| > \beta$.

From this discussion, three useful classes of signals can be identified:

1. ***Right-sided sequences*** of the form

$$x[n] = A_1\alpha_1^n u[n] + A_2\alpha_2^n u[n] + A_3\alpha_3^n u[n] + \cdots. \tag{5.7}$$

They have z-transforms of the form

$$X(z) = \frac{A_1}{1-\alpha_1 z^{-1}} + \frac{A_2}{1-\alpha_2 z^{-1}} + \frac{A_3}{1-\alpha_3 z^{-1}} + \cdots.$$

$A_1, A_2, A_3, \ldots$ are constants. The ROC consists of those values of z for which the z-transform of each of the exponential terms in equation (5.7) converges. This is the intersection of the individual ROCs for each term. Consequently, right-sided sequences will have regions of convergence that are the exteriors of circles in the z-plane, i.e., $|z| > |\alpha_m|$ where $|\alpha_m|=\max(|\alpha_k|, \quad k = 1, 2, \ldots)$ is the exponential of largest magnitude in equation (5.7).

2. ***Left-sided sequences*** of the form

$$x[n] = -B_1\beta_1^n u[-n-1] - B_2\beta_2^n u[-n-1] - B_3\beta_3^n u[-n-1] - \cdots. \tag{5.8}$$

They have z-transforms of the form

$$X(z) = \frac{B_1}{1-\beta_1 z^{-1}} + \frac{B_2}{1-\beta_2 z^{-1}} + \frac{B_3}{1-\beta_3 z^{-1}} + \cdots.$$

$B_1, B_2, B_3, \ldots$ are constants. Here the ROC is the interior of a circle, corresponding to the intersection of the regions of convergence of the individual terms. Convergence occurs for $|z| < |\beta_m|$ where $|\beta_m| = \min(|\beta_k|,\ k = 1, 2, \ldots)$ is the exponential of smallest magnitude in equation (5.8).

3. ***Finite duration sequences.*** These sequences have both a beginning and an end.

$$x[n] = c_0\delta[n-L] + c_1\delta[n-L-1] + \cdots \\ + c_{J-1}\delta[n-L-(J-1)] + c_J\delta[n-L-J] \tag{5.9}$$

They have z-transforms of the form

$$X(z) = c_0 z^{-L} + c_1 z^{-L-1} + \cdots + c_{J-1} z^{-L-(J-1)} + c_J z^{-L-J}.$$

In this case the ROC is the entire z-plane, since a finite sum of finite quantities is finite. The only points in the z-plane that may be excluded from the ROC are $z = 0$ and $z = \infty$. At these values some of the individual terms can become infinite.

A more general class of sequences is the class of two-sided sequences that can be expressed as composites of right-sided, left-sided, and finite duration sequences. The most useful z-transforms encountered in digital signal processing can be expressed as rational functions of the form

$$X(z) = \frac{B(z)}{A(z)} = z^{-L}\frac{b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}}{1 + a_1 z^{-1} + \cdots + a_N z^{-N}}$$

where $X(z)$ is a ratio of polynomials in z^{-1} and $N, M \geq 0$. The roots of the numerator polynomial (interpreted as a function of z, not z^{-1}) are called *zeros*, and the roots of the denominator polynomial are called *poles*. When a polynomial of the form

$$b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}$$

is written as

$$(1 - \beta_1 z^{-1})(1 - \beta_2 z^{-1}) \cdots (1 - \beta_M z^{-1}),$$

its roots are seen to occur at $z = \beta_1, z = \beta_2, \ldots, z = \beta_M$. The first example discussed earlier had a z-transform of the form

$$H(z) = \frac{1}{1 - \alpha z^{-1}}, \qquad |z| > |\alpha|.$$

This has a zero at $z = 0$ and a pole at $z = \alpha$. The second example, which has a z-transform with the same functional form but a different region of convergence, also has a zero at $z = 0$ and a pole at $z = \beta$.

It is convenient to plot the poles and zeros in the z-plane. This is called a *pole/zero* plot. The zeros and poles are typically plotted using $\circ$'s for the zeros and $\times$'s for the

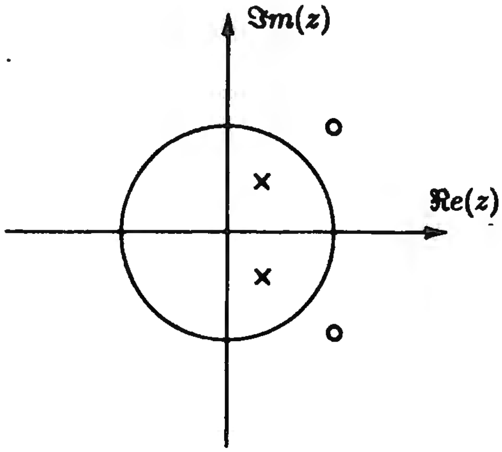

Figure 5.3. An example of a pole/zero plot.

poles. Figure 5.3 shows an example of a pole/zero plot for a signal with two complex poles and two complex zeros.

The locations of the poles influence the possible choices for the regions of convergence in the following ways:

1. The ROC of the z-transform has poles on its boundaries. A pole may never be included in the ROC, since at any value of z where the denominator polynomial is zero, $X(z)$ is not finite. The ROC may be either the interior of a circle, an annular ring bounded by two concentric circles, or the exterior of a circle.
2. If a signal is right-sided, its z-transform converges on the exterior of the circle that is defined by the pole with the largest magnitude. If a signal is left-sided, its z-transform converges on the interior of the circle that is defined by the pole with the smallest magnitude.

The z-transform of the impulse response of a linear time invariant (LTI) system is called its *system function*. The system function (with its ROC) is a unique representation for the system. Since the impulse response of a causal system is right-sided, its system function will converge on the exterior of the circle defined by the largest pole. The ROC of the system function of a stable system must include the unit circle. Inclusion of the unit circle implies that $|z|$ can equal unity in the z-transform summation. This, in turn, implies that the impulse response is absolutely summable. Thus we can determine whether an LTI system is stable by inspecting its ROC.

The z-transform has many well-known properties that make it a valuable and powerful analysis tool. Table 5.1 presents an abbreviated list of these properties. In this table the constants α and β are complex and the constant n_0 is an integer. Each of these properties can be derived from the z-transform definition. The regions of convergence of the transformed signals are not shown. They are related to the ROC of $X(z)$, but in most cases are different from it.

An abbreviated list of common z-transform pairs is given in Table 5.2.

Table 5.1. Properties of the z-Transform

Signal		z-Transform
$x[n]$	$\Longleftrightarrow$	$X(z)$
$x[n-n_0]$	$\Longleftrightarrow$	$z^{-n_0}X(z)$
$\alpha^n x[n]$	$\Longleftrightarrow$	$X(z/\alpha)$
$e^{-j\alpha n}x[n]$	$\Longleftrightarrow$	$X(e^{j\alpha}z)$
$x[-n]$	$\Longleftrightarrow$	$X(z^{-1})$
$nx[n]$	$\Longleftrightarrow$	$-z\frac{d}{dz}X(z)$
$x[n]*h[n]$	$\Longleftrightarrow$	$X(z)H(z)$

Table 5.2. Common z-Transform Pairs

Signal		z-Transform	
$x[n]$	$\Longleftrightarrow$	$X(z)$	
$\delta[n]$	$\Longleftrightarrow$	1	$0 \le \lvert z\rvert \le \infty$
$\delta[n-n_0]$	$\Longleftrightarrow$	z^{-n_0}	$\lvert z\rvert \ne 0$ or $\lvert z\rvert \ne \infty$
$u[n]$	$\Longleftrightarrow$	$\dfrac{1}{1-z^{-1}}$	$\lvert z\rvert > 1$
$\alpha^n u[n]$	$\Longleftrightarrow$	$\dfrac{1}{1-\alpha z^{-1}}$	$\lvert z\rvert > \lvert\alpha\rvert$
$-\alpha^n u[-n-1]$	$\Longleftrightarrow$	$\dfrac{1}{1-\alpha z^{-1}}$	$\lvert z\rvert < \lvert\alpha\rvert$
$n\alpha^n u[n]$	$\Longleftrightarrow$	$\dfrac{\alpha z^{-1}}{(1-\alpha z^{-1})^2}$	$\lvert z\rvert > \lvert\alpha\rvert$
$-n\alpha^n u[-n-1]$	$\Longleftrightarrow$	$\dfrac{\alpha z^{-1}}{(1-\alpha z^{-1})^2}$	$\lvert z\rvert < \lvert\alpha\rvert$
$\alpha^n \cos(\omega_0 n)\, u[n]$	$\Longleftrightarrow$	$\dfrac{1-(\alpha\cos\omega_0)z^{-1}}{1-(2\alpha\cos\omega_0)z^{-1}+\alpha^2 z^{-2}}$	$\lvert z\rvert > \lvert\alpha\rvert$
$\alpha^n \sin(\omega_0 n)\, u[n]$	$\Longleftrightarrow$	$\dfrac{(\alpha\sin\omega_0)z^{-1}}{1-(2\alpha\cos\omega_0)z^{-1}+\alpha^2 z^{-2}}$	$\lvert z\rvert > \lvert\alpha\rvert$

EXERCISE 5.1.1. **Evaluating the z-Transform Analytically**

This problem should be done without the aid of a computer.

(a) Determine the z-transform of each of the following sequences and specify its region of convergence.

(i) $h[n] = \delta[n] + 2\delta[n-1] + 3\delta[n-2]$;
(ii) $s[n] = 4\delta[n-7] + 3\delta[n-9] + 6\delta[n+3]$;
(iii) $g[n] = \left(\frac{1}{3}\right)^n u[n]$;
(iv) $r[n] = \left(\frac{1}{4}\right)^n u[n]$;
(v) $v[n] = g[n] + r[n]$;
(vi) $y[n] = g[n] + r[n] + h[n]$.

(b) For each of the signals above evaluate the z-transform at $z = 0$ and $z = \infty$ to determine whether these points are included in the ROC.

(c) Determine the poles for each sequence that contains poles in the finite z-plane.

EXERCISE 5.1.2. **Manipulating Rational z-Transforms**

Often a z-transform can be expressed as a rational function of the form $B(z)/A(z)$, which is a ratio of polynomials in z. This exercise provides an introduction to representing and processing rational functions or IIR filters.

Consider the system function $H(z)$ defined as

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + 2z^{-1} + 3z^{-2} + 4z^{-3} + 5z^{-4}}{1 + 2z^{-1} + 2z^{-2} + 1z^{-3}}$$

(a) This function can be represented as a file using the *create file* option in `x siggen`. Create $H(z)$, type the file to the screen, and record the results.

(b) Create two separate files representing the polynomials $B(z)$ and $A(z)$. In this case, however, use `x convert` to split $H(z)$ into $B(z)$ and $A(z)$. Next use `x rooter` to factor these polynomials. You can see the roots of $B(z)$ and $A(z)$ by typing the resulting files to the screen. In viewing the files, note that the roots appear as the sequence of complex numbers following the file header. These complex numbers are the real and imaginary parts of the roots. The function `x polar` can be used to convert the roots into polar form. Record these roots in terms of their magnitudes and phases.

(c) The roots of the numerator and denominator polynomials provide the essence of the pole/zero plot. Given these roots, the system function can be reconstructed to within a constant multiplier. Starting with the two files containing the poles and zeros of $H(z)$ (in polar form) use `x cartesian` to convert the roots back to their real and imaginary parts, `x rootmult` to expand these roots into polynomials, and `x revert` to merge the numerator and denominator polynomials into one system function file. Record your result. Compare it to $H(z)$ and verify that it is accurate to within a constant gain term.

EXERCISE 5.1.3. **Producing Pole/Zero Plots**

Pole/zero plots are an important visualization tool in digital signal processing. The plot is formed by factoring the numerator and denominator of a rational system function and plotting the zeros and poles in the z-plane. The function **x polezero** will make pole/zero plots. This exercise demonstrates the use of this function.

(a) Create a file for the system $H(z)$ defined by

$$H(z) = \frac{1 + 3z^{-1} + 3z^{-2} + z^{-3}}{1 + 0.5z^{-1} + 0.3z^{-2} + 0.1z^{-3}} = 1 + 2\,5z\,\frac{1\;1 + 1.08z^{-1} + 0\,36z^{-}}{1 + 0\,5z^{-1} + 0\,3z^{-2} + 0.1z}$$

using the *create file* option in **x siggen.** (improper)

The program **x polezero,** which is menu driven, will automatically find the roots of the numerator and denominator polynomials and make the plot. Execute the program by typing **x polezero**. Select option (1), *read a file*, in the main menu and enter the name of the file containing the coefficients for $H(z)$. Sketch the pole/zero plot that appears on the screen. A plot of the magnitude response in the range from $-\pi$ to π radians also appears in the lower right part of the screen.

(b) The program will also allow you to list the numerical values of the poles and zeros in terms of their real and imaginary parts. Using this option record the poles and zeros of $H(z)$.

EXERCISE 5.1.4. **Evaluating z-Transforms**

Find the system function of the LTI systems with the following impulse responses by manipulating the z-transform summation into the form of a geometric series. It is helpful to recognize the following useful identities involving infinite summations:

$$\sum_{n=n_0}^{\infty} a^n = a^{n_0} \sum_{n=0}^{\infty} a^n$$

$$\sum_{n=-\infty}^{-n_0} a^n = \sum_{n=-n_0}^{-\infty} a^n = \sum_{n=n_0}^{\infty} a^{-n}$$

where n_0 is an arbitrary integer.

(a) $h_a[n] = \left(\frac{1}{8}\right)^n u[3-n]$.

(b) $h_b[n] = n\left(\frac{1}{2}\right)^n u[n-2]$.

(c) $h_c[n] = \left(\frac{1}{2}\right)^n u[n-2] + 3^n u[n]$.

Tell whether each of the systems is stable. Note that this exercise does not involve the aid of a computer.

EXERCISE 5.1.5. **Calculating System Functions**

The systems below represent stable LTI systems that are defined by their impulse responses, denoted $h[n]$.

(i) $h[n] = \left(\frac{1}{5}\right)^n u[n]$;

(ii) $h[n] = \left(\frac{1}{5}\right)^n u[n-5]$;

(iii) $h[n] = \left(\frac{1}{5}\right)^{n-5} u[n-5]$;

(iv) $h[n] = 5^n u[-n-1]$;

(v) $h[n] = 2\delta[n] + \left(\frac{1}{5}\right)^n u[n-1]$.

(a) Determine (analytically) the system function $H(z)$ for each system.

(b) Sketch the pole/zero plot for each and shade the region of convergence. Do this without the aid of the computer.

(c) For each, construct a data file that represents the system using the *create file* option of `x siggen`. Use `x polezero` to display the poles and zeros of each system function. These plots should agree with your sketches in part (b). Note: `x polezero` will not display poles and zeros at the origin.

EXERCISE 5.1.6. **System Functions and Difference Equations**

Causal LTI systems are often defined in terms of difference equations relating the input sequence $x[n]$ to the output sequence $y[n]$. The system function can be obtained by evaluating the z-transform of each term in the difference equation using the second property given in Table 5.1, which is called the *shift property.* For example,

$$\begin{array}{rcl} x[n] & \leftrightarrow & X(z) \\ x[n-1] & \leftrightarrow & z^{-1}X(z) \\ x[n-2] & \leftrightarrow & z^{-2}X(z) \\ \vdots & \vdots & \vdots \\ y[n] & \leftrightarrow & Y(z) \\ y[n-1] & \leftrightarrow & z^{-1}Y(z) \\ \vdots & \vdots & \vdots \end{array}$$

The system function $H(z)$ is given by the ratio

$$H(z) = Y(z)/X(z). \tag{5.10}$$

This follows from the last property in Table 5.1, which is known as the *convolution theorem*.

(a) Determine the system functions for the following causal LTI systems and sketch their pole/zero plots.

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

(ii) $y[n] = \frac{1}{6}y[n-1] + \frac{1}{8}y[n-2] + x[n] + x[n-1]$;

(iii) $y[n] = \sum_{k=-3}^{3} x[n-k]$.

This may be done analytically or by using the *create file* option in **x siggen** and **x polezero** to display the plot.

(b) Now consider the causal LTI system with system function

$$H(z) = \frac{1+z^{-1}}{1-.5z^{-1}-.1z^{-2}}$$

Determine the linear constant coefficient difference equation (LCCDE) that describes this system.

(c) Using the *create file* option in **x siggen** construct a file containing the coefficients for $H(z)$ in part (b). Using **x filter** and an impulse for the input signal, generate 10 samples of the impulse response of $H(z)$. Remember that the input sequence is the first argument on the command line; $H(z)$ is the second. Now use **x lccde** to generate 10 samples of the impulse response. Compare and record these results.

The system function and the difference equation are two closely related ways to represent an LTI system.

EXERCISE 5.1.7. **The First Backward Difference and the Running Sum**

In Chapter 2 the first backward difference and the running sum operations were defined as

$$Y_1(z) = 1 - z^{-1} \Longleftrightarrow y[n] = x[n] - x[n-1]$$

$$Y_2(z) = \frac{X(z)}{1-z^{-1}} \Longleftrightarrow y[n] = \sum_{m=-\infty}^{n} x[m] = \sum_{m=-\infty}^{\infty} x[m]\cdot u[n-m] = x[n] * u[n]$$

where $x[n]$ is the input to the operator and $y[n]$ is its output. These operations are linear and time invariant and, therefore, can be uniquely characterized by their system functions. This problem examines the system functions of the first backward difference operation and the running sum.

(a) Determine the system function of the first backward difference. Using the *create file* option in **x siggen**, create a file that represents this system function.

(b) Repeat part (a) for the running sum operation. (*Hint:* The running sum can also be expressed as the convolution of $x[n]$ with the step function $u[n]$.)

(c) Use the *sine wave* option in **x siggen** to create 128 points of a cosine wave, $x[n]$. Let the frequency (alpha) equal 0.1, the phase (phi) equal $\pi/2$, and the starting point equal 0. Compute and sketch $v[n]$, the output of the first backward difference system with $x[n]$ as the input. Next compute $y[n]$, the output of the running sum system with input $v[n]$. Use x filter to perform the filtering. Remember that the sequence appears as the first argument on the command line and the IIR (running sum) filter appears as the second. Sketch the signals $x[n]$ and $y[n]$ and observe that these operations are inverses of each other.

—*EXERCISE 5.1.8.* **Finding System Functions from Poles and Zeros**

A system may often be described using its poles and zeros. This representation is accurate to within a constant scale factor. If the scale factor and region of convergence are known, then the impulse response is also uniquely specified.

This exercise will make use of the following $H(z)$:

$$H(z) = \frac{(1+0.2z^{-1})(1+0.3z^{-1})(1+0.4z^{-1})(1+0.8z^{-1})}{(1-0.5z^{-1})(1-0.4z^{-1})(1-0.3z^{-1})(1-0.2z^{-1})}$$

(a) Make a sketch of the pole/zero plot for $H(z)$.

(b) Determine the system function in the form

$$H(z) = \frac{B(z)}{A(z)}$$

and create the corresponding signal file. To do this, first put the roots for the numerator and denominator in separate files using the *create file* option in x siggen. Specify the coefficients as complex; the imaginary part will be zero. Use x rootmult to multiply these roots together to form $B(z)$ and $A(z)$. Then use the function x revert to construct $H(z)$. Record the coefficients for $B(z)$ and $A(z)$.

(c) Use x filter to evaluate the impulse response. This can be done by generating an impulse using x siggen and using it and $H(z)$ as the inputs to the x filter function. Sketch the first 10 points of the result.

—*EXERCISE 5.1.9.* **The DTFT of an IIR System**

The frequency response, $H(e^{j\omega})$, and the system function, $H(z)$, of a stable LTI system are identical when $z = e^{j\omega}$. Previously, we evaluated the frequency response of a filter whose impulse response was of finite length by calculating the discrete-time Fourier transform (DTFT) of the impulse response. The DTFT of a system with an infinite duration impulse response can also be computed in practice if $H(z)$ is a rational function. This can be done by computing the DTFT of the numerator and denominator separately and then dividing. To illustrate this, consider the system function

$$H(z) = \frac{B(z)}{A(z)}$$

where

$$B(z) = 1 + z^{-1} + z^{-2} + z^{-3} + z^{-4} + z^{-5}$$

and

$$A(z) = 1 - 0.9z^{-1}.$$

(a) Create separate files for $B(z)$ and $A(z)$ using the *create file* option of `x siggen`. Display and sketch the DTFT magnitude of $B(z)$ using `x dtft`. The function `x dtft` automatically creates an intermediate DSP file called `_dtft_` that contains the frequency-response values that are plotted to the screen. Copy this file to another file called `_dtft1_`. Now use `x dtft` to display and sketch the frequency response of the denominator polynomial $A(z)$. Since $H(e^{j\omega}) = B(e^{j\omega})/A(e^{j\omega})$ the magnitude response of the system can be obtained simply by dividing the frequency response of the numerator by the frequency response of the denominator. Perform this operation by typing

```
x divide    _dtft1_    _dtft_    outfile
```

Use `x look` to display and sketch the magnitude of `outfile` as a continuous waveform.

(b) Now use `x revert` to produce a file for $H(z)$. The function `x dtft` will automatically evaluate the DTFT of a rational function. Display and sketch the DTFT magnitude using `x dtft` directly.

(c) Generate an impulse using `x siggen` and compute 128 points of the impulse response of $H(z)$ using `x filter`. Remember that the first argument of `x filter` should be the impulse and the second should be the IIR filter. Display the impulse response $h[n]$ and observe that the samples of $h[n]$ decay rapidly as n increases. Consequently, the first 128 samples of the impulse response are a good approximation to the true infinite length impulse response. Evaluate the DTFT of the truncated impulse response and sketch the magnitude response.

— *EXERCISE 5.1.10.* **Stability**

If a filter is stable, the ROC of its system function must include the unit circle. If the system is also causal, the poles of the system function must lie inside the unit circle. To illustrate this point, generate files for the following polynomials using the *create file* option of `x siggen`:

(i) $A(z) = 1 + 2z^{-1} + 0.5z^{-2} + 0.2z^{-3}$;

(ii) $B(z) = 1 + 0.5z^{-1} + 0.2z^{-2} + 0.1z^{-3}$;

(iii) $C(z) = 1 + 4z^{-1} + 4z^{-2} + 2z^{-3}$.

Using the function `x revert` construct files for each of the systems given below.

1. $H_1(z) = A(z)/B(z)$
2. $H_2(z) = A(z)/C(z)$
3. $H_3(z) = B(z)/C(z)$
4. $H_4(z) = C(z)/B(z)$

(a) Use x **polezero** to display the pole/zero plots. If each of the systems is causal, determine if each is also stable.

(b) Assume that the system impulse responses are left-sided sequences. In other words, we are assuming that $H_1(z), \ldots, H_4(z)$ are noncausal systems. Tell whether each of the systems is stable.

(c) Now assume that each of the systems has an impulse response that is two-sided, i.e., it extends to infinity in both directions. Which of the systems, if any, is stable? For which system(s) is the answer indeterminate?

—*EXERCISE 5.1.11.* **An Allpole System**

A system with a system function of the form

$$H(z) = 1/A(z)$$

where $A(z)$ is a polynomial in z, is said to be *allpole*. Let $A(z) = 1 + 0.8z^{-1} + 0.4z^{-2} + 0.2z^{-3}$ and assume that the system is causal. Using the *create file* option in x **siggen**, create a file representing $H(z)$ and one representing $A(z)$.

(a) Using x **polezero**, display and sketch the pole/zero and magnitude response plots for $H(z)$ and $A(z)$.

(b) Determine a difference equation with input $x[n]$ and output $v[n]$ that will represent the system, $H(z)$. v(n) x(n) - 0 (n-1) - 4v(n-2) - 0.2v(n-3

(c) Determine another difference equation with input $v[n]$ and output $y[n]$ such that $x[n] = y[n]$. Generate 128 points of a chirp signal using x **siggen** with *alpha1* = 0.05, *alpha2* = 0.001, and *phi* = 1. Use this signal for the input $x[n]$. Apply the function x **filter** to first generate 128 samples of the signal $v[n]$ as defined in part (b). Then use x **filter** to filter $v[n]$ with $A(z)$ to produce 128 samples of $y[n]$. Display $x[n]$ and $y[n]$ one above the other using x **view2** and sketch these signals. y(n) = ( ) + 1) + 0

5.2 THE INVERSE z-TRANSFORM

Until now calculating the impulse response of an LTI system has involved exciting the system with a discrete impulse and observing its response. However, the system impulse response can also be found by taking the inverse z-transform of the system function. This section introduces a method for performing this calculation based on the use of partial fractions. This is a relatively easy way to calculate the inverse z-transform.

The discussion that follows will limit itself to the set of rational system functions with distinct poles. These system functions have the form

$$H(z) = \frac{B(z)}{A(z)} = z^{-L}\frac{b_0 + b_1z^{-1} + b_2z^{-2} + \cdots + b_Mz^{-M}}{1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_Nz^{-N}} \tag{5.11}$$

where L, M, and N are integers and M and N are assumed to be positive. Note that M and N are the orders of the numerator and denominator polynomials in equation (5.11). $H(z)$, in turn, can also be expressed as

$$H(z) = Cz^{-L}\frac{(1-\beta_1 z^{-1})(1-\beta_2 z^{-1})\cdots(1-\beta_M z^{-1})}{(1-\alpha_1 z^{-1})(1-\alpha_2 z^{-1})\cdots(1-\alpha_N z^{-1})} \tag{5.12}$$

where the poles $\alpha_1, \alpha_2, \ldots, \alpha_N$ are all assumed to be different. The constant C is a gain and z^{-L} represents an arbitrary delay. System functions of this form can always be expressed as a sum of right-sided components, left-sided components, and finite length components, i.e.,

$$H(z) = z^{-L}\left[\underbrace{\sum_{k=1}^{K-1}\frac{C_k}{1-\alpha_k z^{-1}}}_{\text{rt-sided terms}} + \underbrace{\sum_{\ell=K}^{N}\frac{C_\ell}{1-\alpha_\ell z^{-1}}}_{\text{lf-sided terms}} + \underbrace{\sum_{m=0}^{M-N} c_m z^{-m}}_{\text{finite length terms}}\right] \tag{5.13}$$

where $1 \leq K \leq N$. Once in this form, the inverse can be found by inspection using equations (5.7–5.9). The challenge is to express $H(z)$ in this form. The method of partial fractions, which is summarized below, accomplishes this.

1. Determine if $H(z)$ is a proper rational function. This means that the condition $M < N$ must be satisfied.[1] When $H(z)$ is proper, there is no finite length component. For example, the function

$$H(z) = \frac{1+2z^{-1}+3z^{-2}}{1+3z^{-1}+4z^{-2}}$$

   is an improper function because both the numerator and denominator have the same order which, in this case, is 2.

2. A partial fraction expansion can only be performed on a proper rational function. If $H(z)$ is not proper, the finite length terms must first be extracted by polynomial division (or *synthetic division* as it is often called). The procedure for dividing polynomials is illustrated later in Example 5.2.

3. Factor the denominator polynomial $A(z)$ into the form

$$A(z) = (1-\alpha_1 z^{-1})(1-\alpha_2 z^{-1})\cdots(1-\alpha_N z^{-1}).$$

4. $H(z)$ can now be written in partial fraction form as

$$H(z) = \frac{C_1}{(1-\alpha_1 z^{-1})} + \frac{C_2}{(1-\alpha_2 z^{-1})} + \cdots + \frac{C_N}{(1-\alpha_N z^{-1})}.$$

[1] In terms of equation (5.13) this means that all c_m terms in the last summation are zero. The finite length terms in this equation do not come into play when $M - N < 0$.

The coefficients $\{C_i\}$ can be found by using the formula

$$C_i = \lim_{z \to \alpha_i} H(z)(1 - \alpha_i z^{-1}). \tag{5.14}$$

The coefficients are determined by removing the pole at α_i from the function and evaluating the remaining expression at $z = \alpha_i$. This formula is only valid when the poles are distinct (i.e., all different), although generalizations of it exist for the more general case.

EXAMPLE 5.1.

As an example, let

$$H(z) = \frac{1 + 2z^{-1}}{1 - (1/4)z^{-2}}$$

be the system function for a causal LTI system. Since the order of the numerator is 1 and the order of the denominator is 2, the function is proper and therefore expressible in partial fraction form without any preliminary synthetic division, i.e.,

$$H(z) = \frac{1 + 2z^{-1}}{(1 - \frac{1}{2}z^{-1})(1 + \frac{1}{2}z^{-1})} = \frac{C_1}{(1 - \frac{1}{2}z^{-1})} + \frac{C_2}{(1 + \frac{1}{2}z^{-1})}.$$

The coefficients, C_1 and C_2, can be evaluated using equation (5.14)

$$C_1 = \lim_{z \to 1/2} H(z)\left(1 - \frac{1}{2}z^{-1}\right) = \left.\frac{1 + 2z^{-1}}{1 + (1/2)z^{-1}}\right|_{z=1/2} = \frac{5}{2}.$$

Similarly,

$$C_2 = \left.\frac{1 + 2z^{-1}}{1 - (1/2)z^{-1}}\right|_{z=-1/2} = -\frac{3}{2}.$$

Now that $H(z)$ is in this form, its inverse can be found by inspection. Causality implies that the impulse response is right sided. Therefore, equation (5.7) can be used to give

$$h[n] = \frac{5}{2}\left(\frac{1}{2}\right)^n u[n] - \frac{3}{2}\left(-\frac{1}{2}\right)^n u[n]. \qquad \blacksquare$$

When the order of the numerator is equal to or greater than that of the denominator, a direct application of the partial fraction expansion will not work. The procedure will work, however, if the system function is first transformed into a proper rational function by long division. To illustrate this procedure consider the following example.

EXAMPLE 5.2.

This example will compute the impulse response corresponding to the system function

$$H(z) = \frac{1 + .2z^{-1} + .5z^{-2} + .2z^{-3}}{1 + .5z^{-1}},$$

which again we assume to be causal. The procedure is to remove the polynomial part of the improper function leaving a fractional part that is proper. The division is performed by arranging the numerator and denominator polynomials in descending powers of z^{-1}.

$$
\begin{array}{r|l}
 & 0.4z^{-2} + 0.2z^{-1} + 3.6 \\
\cline{2-2}
0.5z^{-1} + 1 & 0.2z^{-3} + 0.5z^{-2} + 2.0z^{-1} + 1.0 \\
 & \underline{0.2z^{-3} + 0.4z^{-2}} \\
 & \qquad\quad 0.1z^{-2} + 2.0z^{-1} + 1.0 \\
 & \qquad\quad \underline{0.1z^{-2} + 0.2z^{-1}} \\
 & \qquad\qquad\qquad 1.8z^{-1} + 1.0 \\
 & \qquad\qquad\qquad \underline{1.8z^{-1} + 3.6} \\
 & \qquad\qquad\qquad\qquad\; - 2.6
\end{array}
$$

The division terminates when the remainder is of lower degree than the divisor. In this case, we see that the function can be expressed in the equivalent form

$$H(z) = 3.6 + 0.2z^{-1} + 0.4z^{-2} + \frac{-2.6}{1 + 0.5z^{-1}}$$

from which the inverse can be found by inspection.

$$h[n] = 3.6\delta[n] + 0.2\delta[n-1] + 0.4\delta[n-2] - 2.6(-0.5)^n u[n] \qquad \blacksquare$$

This inverse transform procedure is explored further in the exercises that follow. Some variations on this approach are introduced along with some alternative methods for finding the inverse z-transform.

— *EXERCISE 5.2.1.* **Finding the Inverse z-Transform**

Using the preceding examples as illustrations, determine the impulse responses of the causal systems with the following system functions:

(a)

$$H(z) = \frac{1}{1 - \frac{1}{4}z^{-2}}.$$

(b)

$$H(z) = \frac{1 + 2z^{-2}}{1 - \frac{1}{4}z^{-4}}.$$

This exercise is intended to be worked analytically without the aid of the computer. However, you should feel free to use the computer to verify your answer by checking the first few terms of your impulse responses.

— *EXERCISE 5.2.2.* **More Inverse z-Transforms**

This exercise is intended to be worked analytically.

(a) Determine the inverse z-transform of the function

$$H(z) = \frac{1 + 2z^{-1} + 2z^{-2} + 2z^{-3}}{1 + 0.5z^{-1}}$$

if the system is causal.

(b) Determine the impulse response of the causal system with system function

$$H(z) = \frac{1 + 2z^{-1} + 0.5z^{-2} + 0.4z^{-3}}{1 + 0.5z^{-1} + 0.1z^{-2}}.$$

The previous discussion suggested long division as a method for converting an improper system function into a proper one plus a polynomial so that the inverse z-transform could be computed. This is not the only possible approach.

- Consider the system function

$$H(z) = \frac{1 + 2z^{-2}}{1 - (1/4)z^{-2}}$$

which is an improper rational function that represents a causal system. This can be rewritten in the form

$$H(z) = \frac{1}{1 - (1/4)z^{-2}} + \frac{2z^{-2}}{1 - (1/4)z^{-2}}.$$

The numerator of the second term, $2z^{-2}$, represents a weighted two-sample delay. Thus, if $v[n]$ is the inverse z-transform of the first term, it follows that

$$h[n] = v[n] + 2v[n-2].$$

Note that $v[n]$ can be found without doing any synthetic division. Clearly,

$$\begin{aligned} V(z) &= \frac{1}{(1 - (1/2)z^{-1})(1 + (1/2)z^{-1})} \\ &= \frac{C_1}{1 - (1/2)z^{-1}} + \frac{C_2}{1 + (1/2)z^{-1}} \end{aligned}$$

from which we see

$$v[n] = C_1 \left(\frac{1}{2}\right)^n u[n] + C_2 \left(-\frac{1}{2}\right)^n u[n].$$

EXERCISE 5.2.3. **Completing the Example**

(a) For the example above what are the numerical values of C_1 and C_2?

(b) Determine $h[n]$.

(c) Redo part (b) of Exercise 5.2.2 using this approach and compute the impulse response. Note that the form of the impulse response looks different. Evaluate and record the first three samples of $h[n]$ using both expressions and show that they are in fact the same.

— *EXERCISE 5.2.4.* **Computing the Inverse Transform**

Consider the system function of the causal LTI system, $H(z)$, where

$p_{1,2}$ $\pm 1/2$, p_3 -   4 -

$$H(z) = \frac{1}{1 - 5z^{-1} + 5.75z^{-2} + 1.25z^{-3} - 1.5z^{-4}}.$$

(a) Determine the impulse response using partial fractions. Use x siggen to first create a file for the denominator polynomial and then x rooter to factor the polynomial.

(b) Now assume that this system function corresponds to an anticausal system, for which the impulse response is a left-sided sequence. Sketch the pole/zero plot and indicate the region of convergence. Determine the impulse response.

— *EXERCISE 5.2.5.* **Effect of the ROC**

The z-transform is only unique when the ROC is specified.

(a) Consider the system function

$z_{1,2} = \pm\sqrt{2}$

$$H(z) = \frac{1 - 2z^{-2}}{1 + (1/6)z^{-1} - (1/6)z^{-2}}.$$

$p_{1,2} = -\frac{1}{2}, \frac{1}{3}$

By considering all possible regions of convergence, determine all possible impulse responses for the system.

(b) Now consider the system function

$H(z) = \dfrac{z^8}{(z+1)^2(z-1)^4(z-j)(z+j)}$

$$H(z) = \frac{1}{1 - 2z^{-1} + 2z^{-3} - 2z^{-4} + 2z^{-5} - 2z^{-7} + z^{-8}}.$$

Put $H(z)$ into a file using x siggen. Use x polezero to display the pole/zero plot. Determine the number of impulse response sequences that share this functional form for their z-transform.

EXERCISE 5.2.6. **Inversion by Long Division**

The approach discussed in this chapter for determining an inverse z-transform involves decomposing the transform into simple terms, each of which can be inverse transformed by inspection. This exercise considers a different method in which the inverse is computed term-by-term using long division.

(a) Consider the system function

$$H(z) = \frac{1}{(1 - (1/2)z^{-1})^2}.$$

Expand the denominator and arrange its terms in an order in which the powers of z^{-1} are ascending. Perform long division and generate five or six terms of the quotient. These are the first few terms of the impulse response. Does this impulse response correspond to a causal system?

(b) Repeat part (a) but reverse the order of the terms in the denominator polynomial so that the powers of z^{-1} are arranged in descending order. Is this system causal?

(c) There are difficulties with this method in terms of finding closed-form expressions for the inverses. Discuss how you would find inverses of transforms with annular ROCs.

The remaining exercises in this section illustrate some other properties of the z-transform.

—*EXERCISE 5.2.7.* **Time Reversal**

This exercise examines the effects of time reversal on the pole/zero plot and the region of convergence.

(a) Create a 6-point chirp sequence, $h[n]$, with *alpha1* = 0.5, *alpha2* = 0.8, *phi* = 1, and starting point zero. Use x **polezero** to display and sketch the pole/zero plot. Now form the sequence $h[-n]$ and sketch the pole/zero plot of the reversed signal. What has happened to the zeros? Show analytically why this relationship is always true.

(b) Consider the sequence $g[n] = (1/3)^n u[n]$, which has a system function with one pole. Based on your observations and analysis in part (a), describe what happens to the pole when $g[n]$ is time reversed. What is the ROC of the system function of the time-reversed signal?

—*EXERCISE 5.2.8.* **Modulation**

This exercise explores the effects of modulating a sequence. Create the sequence

$$x[n] = \begin{cases} (\frac{5}{4})^n, & 0 \leq n \leq 5 \\ 0, & \text{otherwise} \end{cases}$$

using the exponential option of x **siggen** and plot the zeros using x **polezero**.

(a) Use x **cexp** to generate the following signals:

(i) $x_i[n] = e^{j\frac{\pi}{2}n}x[n]$,

(ii) $x_{ii}[n] = e^{j\pi n}x[n]$,

(iii) $x_{iii}[n] = e^{-j\frac{\pi}{2}n}x[n]$.

Examine their pole/zero plots using **x polezero** and compare these to the pole/zero plot of $x[n]$. Describe the relationships and justify your observations analytically.

(b) Now determine the pole/zero plots of the signals:

(i) $y_i[n] = \left(\frac{1}{2}\right)^n x[n]$,

(ii) $y_{ii}[n] = \left(\frac{1}{2}e^{j\pi}\right)^n x[n]$,

(iii) $y_{iii}[n] = 2^n x[n]$

using **x cexp** and **x polezero**. Note that the modulation variable, *omega*, can be complex, which allows modulation to be performed by damped exponentials. For example, multiplication by $(1/2)^n$ can be performed when $\omega = j\ln(2) = j(0.69315)$. This can be done by specifying 0 0.69315 for *omega* in the **x cexp** function. Compare the resulting pole/zero plots to that of $x[n]$. Examine the roots carefully in polar coordinates and describe their relationship to the roots of $x[n]$. How does the ROC in each case change? Justify your observations with a short proof.

EXERCISE 5.2.9. **Differentiation of the z-Transform**

One of the well-known properties of the z-transform is the derivative property

$$n\,x[n] \Longleftrightarrow -z\,\frac{d}{dz}X(z),$$

which can be derived from the definition. To illustrate this property consider the simple example

$$X(z) = 1 + 2z^{-1} + 3z^{-2} + 4z^{-3}$$

corresponding to

$$x[n] = \delta[n] + 2\delta[n-1] + 3\delta[n-2] + 4\delta[n-3].$$

This exercise does not require the computer.

(a) Analytically compute the expression $-z(d/dz)X(z)$ and take its inverse z-transform. Evaluate the expression $nx[n]$ and compare this with your result.

(b) Using this property, determine the z-transform of $n\alpha^n u[n]$ where $|\alpha| < 1$.

(c) Determine the z-transform of $n^2\alpha^n u[n]$ for $|\alpha| < 1$.

—*EXERCISE 5.2.10.* **Convolution**

This exercise tests your intuition about discrete-time convolution.

(a) Consider the sequences

$$\begin{aligned} x[n] = 0.75\delta[n] + 0.75\delta[n-1] - 1.25\delta[n-2] \\ -1.25\delta[n-3] - 0.75\delta[n-4] - 0.75\delta[n-5] \\ +0.25\delta[n-6] + 0.25\delta[n-7] \end{aligned}$$

and

$$h[n] = \left(\frac{9}{10}\right)^n u[n].$$

Sketch $x[n]$ and $h[n]$ on paper. Using the method of graphical convolution, produce a rough sketch of $x[n] * h[n]$. Verify your answer by creating a file for $x[n]$ using x siggen and a 50-sample approximation of $h[n]$ using the *exponential* option in x siggen. Convolve these sequences and sketch the result.

(b) Now repeat part (a) for

$$h[n] = \left(\frac{1}{2}\right)^n u[n].$$

Note that $h[n]$, although infinite in duration, can be expressed as the rational function

$$H(z) = \frac{1}{1-(1/2)z^{-1}}.$$

Create a file for $H(z)$ and use x filter to convolve the signals. Generating 20 samples of the output is sufficient. Why is the output not infinite in length?

5.3 FLOW GRAPHS

We have already seen that a discrete-time LTI system is often specified by a difference equation or by a series of difference equations. The system function of such a system is unique, but the converse is not true. For a given rational system function, it is possible to identify several sets of difference equations that will realize the system. Some of these realizations are better than others. For example, consider the causal LTI system defined by the system function

$$H(z) = 1 + b_1z^{-1} + b_2z^{-2} + b_3z^{-3} + b_4z^{-4}.$$

One way to implement this system is to use a single difference equation

$$y[n] = x[n] + b_1x[n-1] + b_2x[n-2] + b_3x[n-3] + b_4x[n-4]$$

where $x[n]$ is the input and $y[n]$ is the output. Another approach is to factor $H(z)$ into the form

$$H(z) = (1 + \alpha_1z^{-1} + \alpha_2z^{-2})(1 + \beta_1z^{-1} + \beta_2z^{-2})$$

and to express the system function as a pair of difference equations corresponding to a *cascade realization*

$$v[n] = x[n] + \alpha_1 x[n-1] + \alpha_2 x[n-2]$$
$$y[n] = v[n] + \beta_1 v[n-1] + \beta_2 v[n-2].$$

A common realization for a high-order IIR system is as a cascade of second-order sections. The system $H(z)$ is factored into subsystems $H_1(z)$, $H_2(z)$, ..., $H_{N/2}(z)$, each of which has two poles and at most two zeros. The input to the overall system is the input to $H_1(z)$. The system output is the output of $H_{N/2}(z)$, as shown in Fig. 5.4.

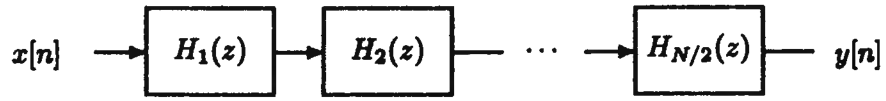

Figure 5.4. A cascade connection of $N/2$ systems.

An alternative is a *parallel implementation*, based on a parallel connection of subsystems or sections as shown in Fig. 5.5 where $H(z) = G_1(z) + G_2(z) + \cdots + G_{N/2}(z)$. These subsystems are typically second-order sections obtained from a partial fraction expansion. Quite often, one of these sections is a constant (i.e., a zero-order subsystem).

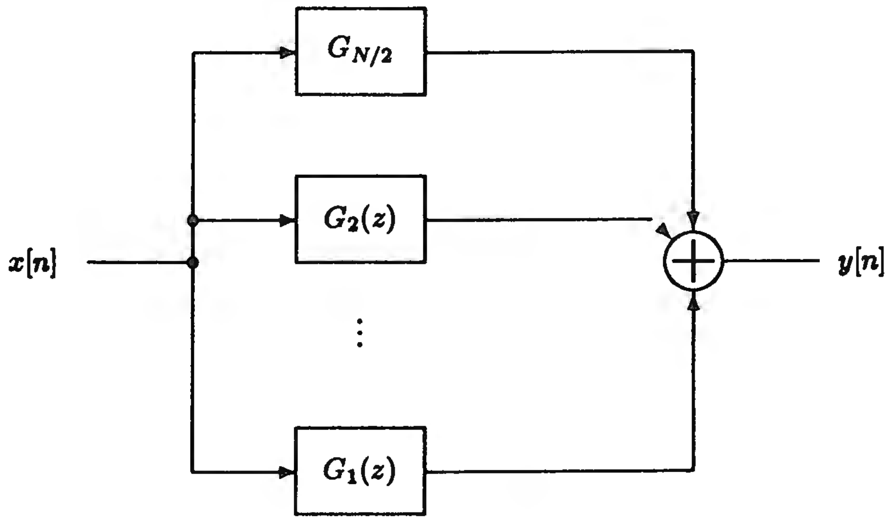

Figure 5.5. A parallel connection of $N/2$ systems.

Although these different implementations all have the same system function, the performance of the system implemented in these different ways will usually be different due to effects such as coefficient quantization and round-off errors introduced by the limited precision of the numerical processor. In addition, some of these implementations may be faster than others (i.e., they may require fewer arithmetic operations per output sample), while other forms may require less memory.

Digital flow graphs provide a means for defining and manipulating different structures for implementing systems. They can be used to describe a wide variety of systems including nonlinear and time-varying ones, but this text will limit itself to the flow graph elements necessary to represent LTI systems. These basic elements are shown in Fig. 5.6.

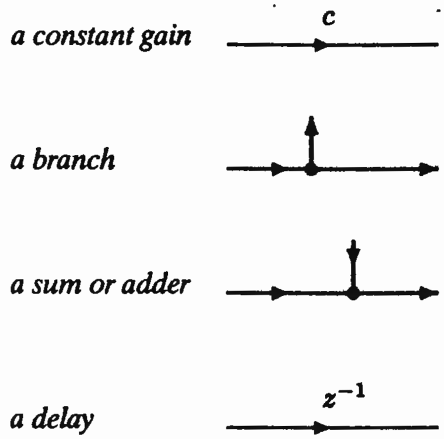

Figure 5.6. The basic elements of a flow graph.

As an example, consider the causal second-order LTI system function

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}.$$

A difference equation that will realize it is

$$y[n] = -a_1 y[n-1] - a_2 y[n-2] + b_0 x[n] + b_1 x[n-1] + b_2 x[n-2]. \tag{5.15}$$

The flow graph can be formed by placing the input node $x[n]$ on the left, the output node $y[n]$ on the right, and then using cascades of delay branches to generate the samples $x[n-1]$, $x[n-2]$, $y[n-1]$, and $y[n-2]$. The nodes corresponding to these sample values are then connected together using gains and summers, as shown in Fig. 5.7. The key to drawing flow graphs is to identify the signals represented at the various nodes in the figure and their interrelationships in the difference equation(s). For example, node C in Fig. 5.7 corresponds to $y[n]$. Going through the delay, we see that node F corresponds to $y[n-1]$, i.e., it is $y[n]$ delayed by one sample. Node E is the sum of two branches, one with the signal $-a_1 y[n-1]$, the other with $-a_2 y[n-2]$. Therefore, the signal at node E is $-a_1 y[n-1] - a_2 y[n-2]$. The signal at node A can be found in a similar way. The contributions from the top, middle, and bottom branches are $b_0 x[n]$, $b_1 x[n-1]$, and $b_2 x[n-2]$, respectively. Thus, at node A, the signal is $b_0 x[n] + b_1 x[n-1] + b_2 x[n-2]$. Following this analysis, it is straightforward to verify that the flow graph describes equation (5.15). This kind of analysis allows one to either draw the flow graphs given an equation or determine the difference equation from the flow graph. Since a linear constant coefficient difference equation (LCCDE) can also be expressed as a system function, it is possible to convert any one of these representations to the others.

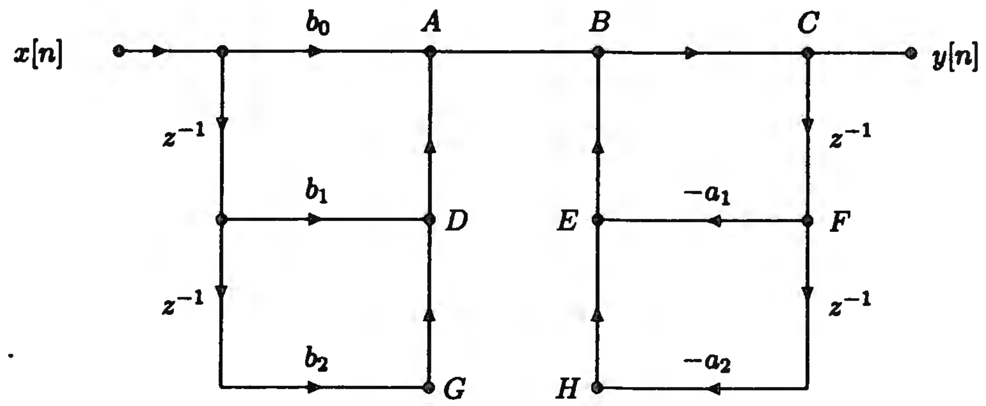

Figure 5.7. Flowgraph of a second-order IIR system.

For LTI systems, an equivalent system can be formed by taking the transpose of the flowgraph. The transpose is formed by the following three-step process:

1. Interchange the input and output nodes. This corresponds to interchanging $x[n]$ and $y[n]$ in Fig. 5.7, for example.
2. Reverse the directions of the arrows along each branch in the flow graph.
3. Change summing nodes into branch points and branch points into summing nodes.

For LTI systems, the system functions of a flowgraph and of its transpose are the same. This provides a convenient way to construct an alternate but equivalent LTI system. The exercises that follow provide some experience in working with flow graphs and implementing systems.

EXERCISE 5.3.1. **A Minimum Delay Implementation**

There are many different ways to implement an LTI system, as we have seen. The implementation in Fig. 5.7 is called a *direct form I* implementation. It is straightforward to derive, but it uses more than the minimum number of delays, since past samples of both the input and output need to be stored. This exercise will consider an alternative.

(a) Referring to Fig. 5.7, split this flow graph into two subsystems: the first with $x[n]$ as its input and the signal at node A as its output, and the second with the signal at node A as its input and $y[n]$ as its output. For convenience, call the first subsystem $H_1(z)$ and the second $H_2(z)$. Thus, the flowgraph is composed of $H_1(z)$ in cascade with $H_2(z)$. Now reverse the order of the systems in the cascade and sketch the resulting flow graph.

(b) Revise this flow graph to remove the redundant delays. The resulting implementation is called a *direct form II* implementation. How many delay elements are required to implement this system?

EXERCISE 5.3.2. **Sketching the Flow Graph from LCCDEs**

Listed below are several difference equations, each of which describes an LTI system. Sketch the flow graph for each of these systems. (This exercise does not involve the use of the computer.)

(a) $y[n] = x[n] + 3x[n-1] - 2x[n-2] - 0.4y[n-1]$.

(b) $y[n] + 0.2y[n-1] - 0.3y[n-2] + 0.5y[n-3] + 0.6y[n-4] = x[n]$.

(c) $y[n] = x[n] + 2x[n-1] - 3x[n-2] - 2x[n-3] + 0.5x[n-4] + 0.2x[n-5]$.

(d) $y[n] = x[n-4] + 2y[n-3] - 7x[n-7]$.

EXERCISE 5.3.3. **Sketching a Flow Graph from a System Function**

The system functions shown below correspond to causal LTI systems.

(i) $H(z) = \dfrac{1 + b_1z^{-1} + b_2z^{-2} + b_3z^{-3}}{1 + a_1z^{-1}};$

(ii) $H(z) = \dfrac{1 + b_1z^{-1} + b_2z^{-2} + b_3z^{-3}}{1 + a_3z^{-3}};$

(iii) $H(z) = c_0\dfrac{1 + b_2z^{-2}}{1 + a_1z^{-1} + a_2z^{-2}}.$

(a) Sketch the flow graphs for each.

(b) Sketch the transpose of each flow graph in part (a).

(This exercise does not involve the use of the computer.)

EXERCISE 5.3.4. **Determining the LCCDE from the Flow Graph**

Consider the causal LTI system with input $x[n]$ and output $y[n]$ described by the flow graph shown in Fig. 5.7 where $b_0 = 1$, $b_1 = \frac{3}{4}$, $b_2 = \frac{1}{8}$, $a_1 = \frac{2}{3}$, and $a_2 = \frac{1}{9}$.

(a) Determine the system function, $H(z)$, that corresponds to this system.

(b) Sketch the pole/zero plot associated with the system.

(c) Determine the impulse response of the system.

(d) Sketch the transposed flow graph corresponding to this system.

(This exercise does not involve the use of the computer.)

EXERCISE 5.3.5. **Cascading Second-Order Sections**

Chapter 2 discussed realizations of systems as cascades of second-order subsystems. Such systems are realized by sets of difference equations each of which describes a constituent subsystem.

Consider the causal fourth-order LTI system described by the difference equation

$$y[n] - y[n-1] + \frac{1}{4}y[n-2] + \frac{1}{4}y[n-3] - \frac{1}{8}y[n-4]$$
$$= x[n] + 2x[n-1] + x[n-2].$$

(a) Sketch the flow graph of this system as a direct form I structure.

(b) Find the system function and write it as the product of two second-order system functions. You may wish to use `x siggen` to create a file representing the denominator polynomial of the system function. This polynomial may be decomposed into second-order subsystems using the *delete root* and *write file* options in `x polezero`. Observe that some of the roots are complex. How should the poles within each section be arranged to avoid complex coefficients in the flow graph?

(c) Sketch the flow graph corresponding to the cascade structure that is consistent with your result in part (b).

EXERCISE 5.3.6. **A Parallel Form Implementation**

In Chapter 2, a brief discussion of the parallel form implementation was given. This implementation is based on the representation of a system function as a sum of low-order rational functions, i.e.,

$$H(z) = C_0 + H_1(z) + H_2(z) + \cdots + H_{N/2}(z).$$

Consider the causal fourth-order LTI system given in the previous exercise. Express this system as a sum of second-order subsystems plus a constant and sketch the flow graph corresponding to this realization.

5.4 PLOTTING THE FREQUENCY RESPONSE

Computer-generated plots of the frequency response are routinely produced for systems analysis. These plots can be generated by explicitly evaluating discrete samples of the DTFT using the FFT algorithm that is discussed in the next chapter. These samples can be displayed as a continuous function. While these representations are both accurate and important, there are occasions when a computer is not available and a rough sketch of the frequency response may be satisfactory. Such a rough sketch of the magnitude and phase responses can be obtained from the pole/zero plot using a set of rules based on a geometric interpretation of the DTFT. This method, although not very accurate, allows sketches to be made virtually by inspection. An even more important reason for considering this procedure, however, is the insight that it provides to the system designer.

To develop this method recall that the frequency response is equal to the system function evaluated on the unit circle, i.e.,

$$H(e^{j\omega}) = H(z)|_{z=e^{j\omega}}. \tag{5.16}$$

This allows the frequency response of a system to be related to its poles and zeros. Let $H(z)$ be an arbitrary rational function written in the form we have seen before:

$$H(z) = Cz^{-L}\frac{(1-\beta_1 z^{-1})(1-\beta_2 z^{-1})\cdots(1-\beta_M z^{-1})}{(1-\alpha_1 z^{-1})(1-\alpha_2 z^{-1})\cdots(1-\alpha_N z^{-1})}. \tag{5.17}$$

The magnitude of the frequency response is seen to be

$$\begin{aligned}|H(e^{j\omega})| &= |C||e^{-j\omega L}|\frac{|(1-\beta_1 e^{-j\omega})||(1-\beta_2 e^{-j\omega})|\cdots|(1-\beta_M e^{-j\omega})|}{|(1-\alpha_1 e^{-j\omega})||(1-\alpha_2 e^{-j\omega})|\cdots|(1-\alpha_N e^{-j\omega})|}\\ &= |C|\frac{|(e^{j\omega}-\beta_1)||(e^{j\omega}-\beta_2)|\cdots|(e^{j\omega}-\beta_M)|}{|(e^{j\omega}-\alpha_1)||(e^{j\omega}-\alpha_2)|\cdots|(e^{j\omega}-\alpha_N)|}.\end{aligned} \tag{5.18}$$

The last equality follows from the fact that $|e^{j\omega}| = 1$.

A geometrical interpretation can be found by examining the general term $|(e^{j\omega} - \alpha)|$. The point $e^{j\omega}$ lies on the unit circle at an angle ω as shown in Fig. 5.8. The location of the *zero*, which occurs at $z = \alpha = |\alpha|e^{j\theta}$ in the illustration, can also be represented by a vector in the z-plane. It has a magnitude equal to $|\alpha|$ and an angle $\theta = \angle\alpha$. The quantity of interest, however, is $|(e^{j\omega} - \alpha)|$, which is the magnitude of the difference of these two vectors. This is also shown in Fig. 5.8. We observe that as the angle ω corresponding to the point $e^{j\omega}$ on the unit circle is varied, the magnitude of this difference vector, $(|(e^{j\omega} - \alpha)|)$, changes. The implication of this geometric observation is that the frequency-response magnitude can be interpreted as the length of this vector for a given angle ω. Note that the angle ω is simply the frequency in the frequency-response plot and can assume values between 0 and 2π, $-\pi$ and π, or any other interval of 2π radians. For this example the magnitude of the difference vector at $\omega = 0$ is initially large. As ω increases, i.e., moves counterclockwise around the unit circle, the length of the difference vector decreases. It attains its smallest value at $\omega = \pi$ and then increases in length until $\omega = 2\pi$. By plotting this change in length as a function of ω we obtain a plot of the magnitude response for $0 \le \omega \le 2\pi$. Notice how the periodicity of the DTFT is illustrated by this geometric viewpoint. If one continues plotting the magnitude response for ever increasing values of ω, we see that this merely corresponds to continually circling the unit circle. Each revolution about the circle represents a period that, in turn, leads to the inherent periodicity property of the DTFT.

The simple example illustrated in Fig. 5.8 can be generalized to represent arbitrary functions of the form of equation (5.18). The magnitude response of such an arbitrary

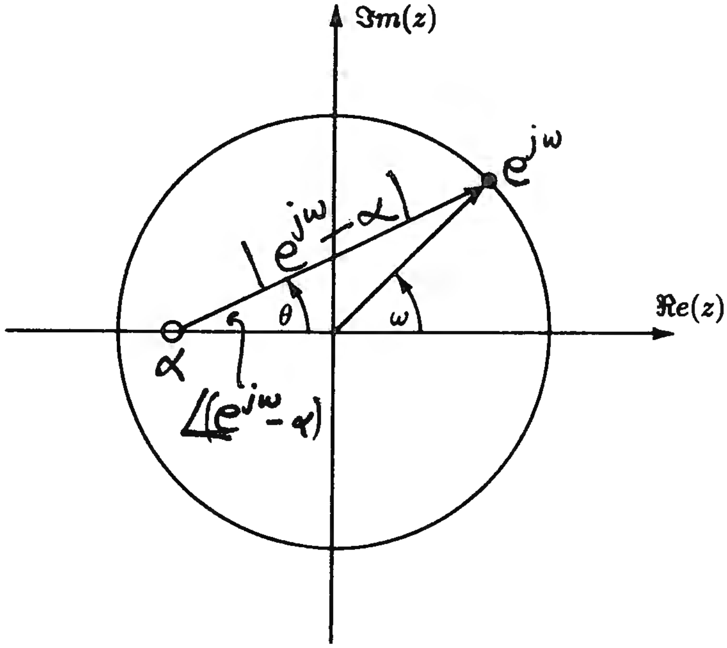

Figure 5.8. A geometric interpretation of the contribution to the magnitude and phase response due to a single zero.

function can be expressed as the product of individual difference vector magnitudes for the zeros divided by the product of vector magnitudes for the poles. Poles and zeros at the origin can be ignored, since the distance to these is unity for all ω.

Several useful observations arise out of this geometric interpretation. Zeros close to the unit circle cause a dip in the DTFT magnitude in the region near the zero. A zero that lies on the unit circle causes the DTFT magnitude to become zero at that frequency. Similarly, poles close to the unit circle result in peaks in the magnitude response at frequencies close to the angle of the pole. Thus a simple rule for sketching the DTFT magnitude can be formulated. As we move around the unit circle, the DTFT peaks when we pass close to a pole and dips when we pass close to a zero. The sharpness of the peak or dip is directly related to the closeness of the pole or zero to the unit circle. Roots that are far from the unit circle do not affect the magnitude response significantly.

A method for plotting the phase response can be derived similarly, but the procedure is a little more difficult. Consider expressing equation (5.17) in factored form in terms of the individual magnitude and phase terms, i.e.,

$$H(e^{j\omega}) = |A|e^{j\omega K} \frac{\prod_{n=0}^{N} |(e^{j\omega} - \beta_n)| e^{j\angle(e^{j\omega} - \beta_n)}}{\prod_{m=0}^{M} |(e^{j\omega} - \alpha_m)| e^{j\angle(e^{j\omega} - \alpha_m)}}. \tag{5.19}$$

The phase is given by

$$\angle H(e^{j\omega}) = \omega K + \sum_{n=0}^{N} \angle(e^{j\omega} - \beta_n) - \sum_{m=0}^{M} \angle(e^{j\omega} - \alpha_m) \tag{5.20}$$

The contribution of each of the individual phase terms corresponds to the direction (or angle) of the difference vector in Fig. 5.8 from its zero or pole. To compute the magnitude response, we plot the difference vector length as a function of ω. To compute the phase response we plot the angle θ as a function of ω. There is a phase plot associated with each individual pole and zero. To plot the phase response for an arbitrary function of the form of equation (5.19), add the phase plots associated with the zeros and subtract the phase plots associated with the poles. This is the geometric procedure implied by equation (5.20).

— *EXERCISE 5.4.1.* **Sketching the Magnitude Response**

(a) Consider the system with the system function

$$H(z) = 1 - 0.5z^{-1}.$$

Draw a freehand sketch of the pole/zero plot. Divide the unit circle into 16 evenly spaced segments and place a tic mark on the circle at each segment boundary. Measure the distance from the zero to each point on the unit circle. Plot these distances as a function of angle, proceeding counterclockwise around the unit circle. This is a crude plot of the magnitude of the DTFT in the range $0 \le \omega \le 2\pi$. Resketch the plot in the range $-\pi < \omega < \pi$.

• Create a file for $H(z)$ using the *create file* option in `x siggen`. Use `x polezero` to check the accuracy of your plot.

(b) Now consider the system with system function

$$H(z) = \frac{1}{1 - 0.5z^{-1}}.$$

Sketch the DTFT magnitude using the same procedure that you used in part (a). Here your plot is the reciprocal of the distances from the pole to points on the unit circle. Again check your answer using `x polezero`.

— *EXERCISE 5.4.2.* **Sketching Magnitude Responses from Pole/Zero Plots**

The geometric interpretation of the pole/zero plot provides a useful method for obtaining a quick sketch of the magnitude response. Consider the following set of system functions expressed in terms of their poles and zeros in polar coordinate form:

System	Poles (mag., angle)	Zeros (mag., angle)
$H_1(z)$	(0.5, 3.1416)	(1.0, 0.0000)
$H_2(z)$	(1.5, 3.1416)	(0.5, 1.5708) (0.5, −1.5708)
$H_3(z)$	(1.0, 0.3927)	(1.0, −0.3927)
$H_4(z)$	(0.6086, 0.96924) (0.6086, −0.96924)	(1.0, 2.4444) (1.0, −2.4444)
$H_5(z)$	(0.6086, 2.17235) (0.6086, −2.17235)	(1.0, 0.6972) (1.0, −0.6972)

Sketch the pole/zero plot for each of these systems. Using the geometric rules developed in the discussion, provide a rough sketch of the magnitude response for each system. Feel free to use the computer to check your results.

EXERCISE 5.4.3. **Sketching the Phase Response**

(a) Consider the system with system function

$$H(z) = 1 - 0.7z^{-1}.$$

Create a file for this function using the *create file* option in x siggen. Display the pole/zero plot using x polezero and sketch it. Now divide the unit circle into 16 evenly spaced segments and place a tic mark on the circle at each segment boundary. Measure the angle ϕ, which is shown in Fig. 5.8 as a function of ω at each tic mark. Plot these phase angles as a function of ω to obtain a rough phase plot for $0 \leq \omega < 2\pi$. To check your answer on the computer, use x dtft and display the phase response. *Note:* The phase plot in x dtft is in the range $-\pi \leq \omega \leq \pi$.

(b) Repeat part (a) for the single pole system function

$$H(z) = \frac{1}{1 - 0.7z^{-1}}.$$

Note that for a pole the phase is a plot of $-\phi$ as a function of ω.

EXERCISE 5.4.4. **Poles and Zeros at Infinity**

Poles and zeros at $z = 0$ and $z = \infty$ are often downplayed as being relatively unimportant. This is because these poles and zeros simply cause a shift in the sequence.

This exercise examines the impact of these poles and zeros using examples. It does not require the use of a computer.

Consider the sequence

$$h[n] = \left(\frac{1}{2}\right)^n u[n] + \left(\frac{1}{4}\right)^n u[n]$$

with the z-transform

$$H(z) = \frac{2 - (3/4)z^{-1}}{1 - (3/4)z^{-1} + (1/8)z^{-2}}.$$

(a) Sketch the pole/zero plot.

(b) How do the following modifications to the pole/zero plot affect $h[n]$?

(i) Introduction of a zero at $z = 0$.

(ii) Introduction of a pole at $z = 0$.

(iii) Introduction of a zero at $z = \infty$.

(iv) Introduction of a pole at $z = \infty$.

(c) What is the relationship between zeros at $z = 0$ and poles at $z = \infty$?

—*EXERCISE 5.4.5.* **Determining Poles and Zeros from the DTFT**

The impact of poles and zeros (located near the unit circle) on the magnitude response was demonstrated in the introductory discussion. This exercise looks at the reverse relationship by trying to determine the location of poles and zeros from the frequency response.

(a) Using the *create file* option in `x siggen`, create a file to represent the system function of the allpole system

$$H_1(z) = \frac{1}{1 - 0.5z^{-1} + 0.2z^{-2} - 0.1z^{-3} + 0.007z^{-4} + 0.14z^{-5} + 0.15z^{-6}}.$$

This system function has six poles.

(i) Use `x dtft` to display and sketch the DTFT magnitude of this allpole filter. Try to sketch the pole/zero plot of this function based on the DTFT magnitude.

(ii) Now use `x dtft` to display and sketch the phase response of the system. Based on this plot, attempt a sketch of the pole/zero plot. This one is somewhat more difficult.

(iii) Use `x polezero` to display the true pole/zero plot and assess the accuracy of your previous attempts.

(b) Consider the system function $H_2(z)$ where

$$H_2(z) = 1 + 1.9z^{-1} + 0.8z^{-2} - 0.8z^{-3} - 0.7z^{-4}.$$

This corresponds to a system with finite length impulse response. Create a file containing $H_2(z)$ using the *create file* option in **x siggen**.

(i) Examine and sketch the DTFT magnitude of this function using **x dtft**. Based on this sketch, attempt to draw the pole/zero plot.

(ii) Use **x polezero** to display the true pole/zero plot and compare.

(c) What visible cues in the magnitude and phase plots provide information about the pole and zero locations? How are these cues affected by poles and zeros located away from the unit circle?

— *EXERCISE 5.4.6.* **Interaction of Poles and Zeros**

The poles and zeros of a function have opposite effects on the magnitude response. In this exercise, the interaction of a pole and a zero in close proximity is examined using the following example:

$$H(z) = \frac{1.0 - \alpha z^{-1}}{1.0 - \beta z^{-1}}.$$

(a) Use **x siggen** to create the IIR filter $H(z)$ with $\beta = 0.5$ and $\alpha = 0.8$. Use **x polezero** to display and sketch the pole/zero plot of this function.

(b) Now assume that the pole and zero are variable but constrained to be on the real axis. Use the *change pole/zero* option in **x polezero** to vary the distance between the pole and zero. Pressing the "*d*" key will update the magnitude display. What happens when the pole and zero approach each other? Now move the pole and zero closer to the unit circle. What is the effect of these interactions on the magnitude response?