HST.582J / 6.555J / 16.456J Biomedical Signal and Image Processing
Spring 2007

Chapter 6 - Z-TRANSFORMS

Introduction

Chapters 2 and 3 considered the design and analysis of digital filters in the time and frequency domains and illustrated the use of the discrete-time Fourier transform to determine the frequency response of linear, time-invariant systems. The limitation of that approach is that the DTFT only exists if the signal/unit-sample response is absolutely summable or contains finite energy. The Z-transform is a generalization of the DTFT and applies to signals/unit-sample responses that do not meet either of these criteria. Therefore, the Z-transform is a useful tool for investigating issues related to stability, including analysis of feedback systems.

6.1 Definition and properties

6.1.1 Definition

The Z-transform can be used to characterize the response of linear, time-invariant filters to complex exponential signals. Specifically, consider the response of a filter with unit-sample response $h[n]$ to the complex exponential z^n, where z is an arbitrary complex number:

$$y[n] = h[n] * z^n = \sum_{m=-\infty}^{\infty} h[m] z^{n-m}$$

This can be written as:

$$y[n] = z^n \sum_{m=-\infty}^{\infty} h[m] z^{-m} = x[n]H(z)$$

with

$$H(z) \triangleq \sum_{n=-\infty}^{\infty} h[n] z^{-n} \tag{6.1}$$

Thus, the output is equal to the input multiplied by the complex constant $H(z)$. The complex exponential signal z^n is said to be an *eigenfunction* of the linear, time-invariant system $h[n]$. When considered as a function of z, the *eigenvalue* $H(z)$ is the *Z-transform* of $h[n]$. Its argument is the complex variable z which defines a plane, with the real part and the imaginary part as orthogonal coordinates.

For LTI systems described by a unit-sample response $h[n]$, $H(z)$ is also referred to as the *system function*. The system function is a generalization of the frequency response, since the DTFT is a special case of the Z-transform, that is,

$$H(f) = H(z)|_{z=e^{j2\pi f}}$$

6.1.2 Z-transforms of filters defined by a difference equation

Consider the class of digital filters described by linear, constant-coefficient difference equations (LCCDEs) of the form

$$y[n] = \sum_{k=1}^{K} a_k y[n-k] + \sum_{m=0}^{M} b_m x[n-m] \tag{6.2}$$

The Z-transform of digital filters defined by this equation follows directly from the definition of $H(z)$ as the ratio of the output to the input for complex exponential inputs. Specifically, if $x[n] = z^n$, then $y[n] = H(z)z^n$. Substituting these expressions for $x[n]$ and $y[n]$ in Eq. 6.2 produces

$$y[n] = H(z)z^n = \sum_{k=1}^{K} a_k H(z) z^{n-k} + \sum_{m=0}^{M} b_m z^{n-m}$$

Dividing both sides of this equation by z^n, and rearranging terms, we obtain the system function:

$$H(z) = \frac{\sum_{m=0}^{M} b_m z^{-m}}{1 - \sum_{k=1}^{K} a_k z^{-k}} \tag{6.3}$$

Equation 6.3 shows that the Z-transform of a digital filter defined by a finite difference equation is a rational function in z. Conversely, for any rational function in z, we can write a corresponding difference equation in the form of Eq. 6.2 defining a digital filter. Thus, digital filters defined by finite difference equations correspond to LTI systems with rational system functions. Because any reasonably well-behaved function of z can be approximated by a rational function, any LTI filter can be approximated by a finite difference equation that can be implemented on a digital computer.

The K complex roots of the denominator of $H(z)$ are called the *poles* of the filter, while the M roots of the numerator are called *zeroes.* For FIR filters, the denominator of Eq. 6.3 is unity, so that the system function has only zeroes. On the other hand, for purely recursive filters, the numerator is the constant b_0, so that the system function has only poles. This explains the terms *all-zero* and *all-pole* to designate FIR filters and purely-recursive filters, respectively.

Examples

The Z-transform of simple digital filters can be computed either by direct application of Eq. 6.1 if the unit-sample response $h[n]$ is known or from Eq. 6.3 if the difference equation is known.

We will use the notation

$$h[n] \longleftrightarrow H(z)$$

to denote the relation between a signal $h[n]$ and its Z-transform $H(z)$.

1. Gain:

$$G\delta[n] \longleftrightarrow G$$

   In particular, the Z-transform of the unit sample $\delta[n]$ is the constant 1.

2. Delay by n_0 samples:
$$\delta[n-n_0] \longleftrightarrow z^{-n_0} \tag{6.4}$$
This explains why the notation z^{-1} is often used to designate a unit delay.

3. Rectangular ("boxcar") filter of length N:
$$R_N[n] \stackrel{\triangle}{=} u[n]-u[n-N] \longleftrightarrow \sum_{n=0}^{N-1} z^{-n} = \frac{1-z^{-N}}{1-z^{-1}}$$
This filter has $N-1$ zeroes equally spaced on the unit circle, except for $z=1$ where the zero is cancelled by a pole.

4. First-order recursive lowpass filter $y[n] = ay[n-1]+x[n]$:
$$a^n u[n] \longleftrightarrow \frac{1}{1-az^{-1}} \quad \text{for} \quad |z|>|a| \tag{6.5}$$
This filter has one pole at $z=a$. As a special case, the Z-transform of the unit step $u[n]$ is $\frac{1}{1-z^{-1}}$ for $|z|>1$.

6.1.3 Properties

The Z-transform is a *linear* operation in the sense that
$$c_1h_1[n]+c_2h_2[n] \longleftrightarrow c_1H_1(z)+c_2H_2(z) \tag{6.6}$$
for c_1 and c_2 arbitrary constants. This implies that the Z-transform of a parallel combination of filters is the sum of the transforms for each of the filters.

The most important property of Z-transforms is the *convolution theorem,* which states that
$$h[n]*x[n] \longleftrightarrow H(z)X(z) \tag{6.7}$$
Because it replaces the complicated convolution operation by a simpler multiplication, this theorem is useful for computing the responses of digital filters to signals given by analytic expressions, and for finding the unit-sample response of cascade combinations of filters. To prove Eq. 6.7, we write the definition of the Z-transform of $y[n]=x[n]*h[n]$:
$$Y(z) = \sum_{n=-\infty}^{\infty} y[n]z^{-n} = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} h[m]x[n-m]z^{-n}$$
Interchanging the order of summations over n and m, and making use of $z^{-n}=z^{-m}z^{-(n-m)}$, we obtain
$$Y(z) = \sum_{m=-\infty}^{\infty} h[m]z^{-m} \sum_{n=-\infty}^{\infty} x[n-m]z^{-(n-m)}$$
Making the change of variable $l=n-m$, we recognize the product of $H(z)$ and $X(z)$, proving Eq. 6.7.

An important special case of the convolution theorem is the *delay theorem*

$$h[n-n_0] = \delta[n-n_0] * h[n] \longleftrightarrow z^{-n_0}H(z) \tag{6.8}$$

where we have made use of Eq. 6.4.

Another useful property is the *Z-derivative,* or *ramp-multiplication* theorem:

$$nh[n] \longleftrightarrow -z\frac{dH(z)}{dz} \tag{6.9}$$

This is easily shown by taking the derivative of the definition of the Z-transform:

$$\frac{dH(z)}{dz} = \frac{d}{dz}\sum_{n=-\infty}^{\infty} h[n]z^{-n} = \sum_{n=-\infty}^{\infty} -nh[n]z^{-n-1} = -z^{-1}\sum_{n=-\infty}^{\infty}(nh[n])z^{-n}$$

Applying Eq. 6.9 to Eq. 6.5, one obtains

$$na^n u[n] \longleftrightarrow \frac{az^{-1}}{(1-az^{-1})^2} \quad \text{for} \quad |z| > |a| \tag{6.10}$$

This transform has a double pole at $z = a$.

6.2 Region of Convergence

For signals of finite duration, the Z-transform is well-defined for all values of the complex variable z (with the possible exception of the origin) because it is a finite sum of finite terms. Thus the Z-transforms of FIR filters alway converge. On the other hand, the Z-transforms of IIR filters are only defined for certain values of z. The *region of convergence* is the portion of the complex z-plane for which the summation of Eq. 6.1 converges.

As seen in Sec. 6.1.2, the Z-transform of the first-order, recursive lowpass filter $h[n] = a^n u[n]$ is $H(z) = \frac{1}{1-az^{-1}}$. The corresponding region of convergence is $|z| > |a|$.

6.2.1 Properties of the region of convergence

The properties of the region of convergence are stated here; justification for these properties can be found in *Oppenheim, Willsky and Nawab* or *Oppenheim and Schafer.*

1. The region of convergence is always a contiguous region, bounded by circles centered at the origin in the z-plane.

2. The region of convergence may contain zeros, never contains poles, and is always bounded by poles. (The fact that the region of convergence never contains poles is understood by realizing that a pole is a value of z for which $H(z)$ is infinite, and the region of convergence includes all values of z for which Eq. 6.1 converges.

3. If $h[n]$ is the unit-sample response of an FIR filter, the region of convergence of the system function $H(z)$ is the entire z-plane, with the possible exception of $z = 0$ and $z = \infty$.

4. If $h[n]$ is the unit-sample response of a right-sided IIR filter (that is, one that is finite in the negative-time direction but extends infinitely in the positive-time direction), then the region of convergence of $H(z)$ is the exterior of a circle.

5. If $h[n]$ is the unit-sample response of a left-sided IIR filter (that is, one that is finite in the positive-time direction but extends infinitely in the negative-time direction), then the region of convergence of $H(z)$ is the interior of a circle.

6. If $h[n]$ is the unit-sample response of an IIR filter that extends infinitely in both positive and negative time, then the region of convergence of $H(z)$ is a 'donut' in the complex z-plane. This is generally of little practical importance, since such IIR filters are not recursively computable, and we are not interested in filters that can not readily be implemented on a digital computer. However, such non-computable filters might be of use in modeling certain physical systems.

6.2.2 Causality and the region of convergence

Since causal filters are a subset of right-sided unit-sample responses, the region of convergence for a causal filter is the exterior of a circle centered at the origin of the z-plane (Property 4). Recalling that the form of the LCCDE specified in Eq. 6.2 implicitly defines a causal filter and using Property 2, we conclude that *for a causal filter, the region of convergence is the exterior of the circle defined by the pole(s) with the largest magnitude.*

6.2.3 Stability and the region of convergence

We saw in Chapter 2 that a system is stable if its unit-sample response is absolutely summable. In this case, the DTFT exists ($z = e^{j2\pi f}$) and the Z-transform converges for for $|z| = 1$, that is, for values of z that are on the unit circle. Thus, *the region of convergence of the Z-transform of a stable system always includes the unit circle.* Since the region of convergence of a stable filter must include the unit circle and since the region of convergence must be a contiguous region and not include any poles, we conclude that the the Z-transform of a stable filter has all poles either inside or outside the unit circle.

If we further restrict our consideration to causal filters, which requires that the region of convergence be the exterior of a circle, then we conclude that *the region of convergence of a causal stable filter is the outside of a circle whose radius is less than unity and the poles of a causal, stable filter are all inside the unit circle.* For example, the first-order lowpass filter has a pole at $z = a$. This filter will be stable if and only if this pole is inside the unit circle, that is, if $|a| < 1$. This result fits with the observation that the unit sample response $a^n u[n]$ is absolutely summable if and only if $|a| < 1$.

To consider another example, the trapezoidal integration formula is not stable because it has a pole on the unit circle for $z = 1$. Indeed, the response of the integrator to a unit step is not bounded. When, as for the integrator, the poles with the largest magnitude are on the unit circle, the response to a bounded input can only grow with time as a polynomial. This form of unstability is less severe than when there are poles outside the unit circle, in which case the output grows exponentially with time. A system for which the largest pole is on the unit circle can be considered to be marginally stable. Many practically important filters such as integrators and lowpass and bandpass filters are marginally stable.

Unlike poles, there is no restriction on the location of zeros for causal, stable filters. In the special case when a filter has all its poles and all its zeros inside the unit circle, both $H(z)$ and its inverse $\frac{1}{H(z)}$ are causal and stable. Such filters are called *minimum phase.* The term "minimum phase" is used because such filters have the lowest delay among all possible causal filters with the same magnitude $|H(z)|$ on the unit circle.

6.2.4 Feedback example

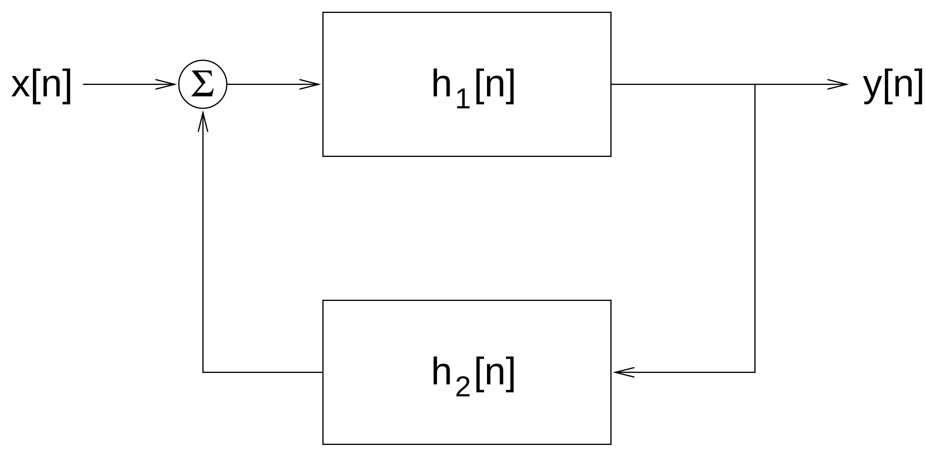

Figure 6.1: General feedback system

Consider feedback of the form shown in Fig. 6.1. The overall system function, $H(z)$ can be determined by writing an equation for the output $y[n]$, and then converting the equation to the Z-transform domain using the linearity property (Eq.6.6) and the convolution theorem:

$$y[n] = h_1[h] * (x[n] + h2[n] * y[n])$$

$$Y(z) = H_1(z)[X(z) + H_2(z)Y(z)] \tag{6.11}$$

Rearranging Eq. 6.11 gives

$$H(z) = \frac{Y(z)}{X(z)} = \frac{H_1(z)}{1 - H_1(z)H_2(z)} \tag{6.12}$$

From Eq. 6.12, we see that the addition of feedback, $H_2(z)$ has a substantial effect on the overall system function. Such feedback alters the locations of the poles of the system function and can be used to make a stable system unstable or to make an unstable system stable. The following example illustrates the former.

Consider a public address systems intended to pick up a signal spoken into a microphone and play an amplified version via a loudspeaker. The simple unit-sample response of this system is $h_1[n] = G\delta[n]$, where $G > 1$. In addition to the desired signal, the microphone also picks up the loudspeaker output with some delay and attenuation dependent on the distance from the loudspeaker to the microphone. We model this simple feedback path as $h_2[n] = a\delta[n - n_0]$, where $0 < a < 1$. Substituting $H_1(z) = G$ and $H_2(z) = az^{-n_0}$ into Eq. 6.12 gives

$$H(z) = \frac{G}{1 - Gaz^{-n_0}},$$

which has poles at the n_0^{th} roots of Ga. Since $H(z)$ is causal, in order for it also to be stable, these poles must lie within the unit circle, that is, have magnitude less than one. Since the n_0^{th} roots of Ga must have magnitude less than one, then $Ga < 1$, or $a < \frac{1}{G}$. Practically, this indicates that an unstable (that is, squealing) public address system can be made stable by reducing G (turning down the gain) or by reducing a (increasing the distance between the loudspeaker and the microphone, or shielding the microphone from the loudspeaker signal).

6.3 Inverse Z-transforms

Time signals can be exactly recovered from their Z-transforms, providing that the region of convergence is specified. Such *inverse Z-transforms* are useful for computing the response of digital filters to known signals by means of the convolution theorem, and for determining the unit-sample response of IIR filters. While it is possible to give a formal inverse Z-transform formula (see, for example, *Oppenheim and Schafer),* a strategy used in practice is to decompose the Z-transform into a sum of simple terms whose inverse transforms can be recognized by inspection. This can be seen as replacing a complex filter by a parallel combination of simpler filters. In the case of rational Z-transforms, a general inverse transform technique is to carry out a partial-fraction expansion. Specifically, let $A(z)$ and $B(z)$ be the denominator and numerator of Eq. 6.3, respectively. $A(z)$ and $B(z)$ are polynomials in z^{-1} of degrees K and M respectively. By polynomial division, $B(z)$ can be written uniquely in the form

$$B(z) = A(z)Q(z) + R(z)$$

where $Q(z)$ and $R(z)$ are polynomials in z^{-1}, and the degree of $R(z)$ is less than K, the degree of $A(z)$. In the special case when $M < K$ in Eq. 6.3, $Q(z)$ is zero, and $R(z) = B(z)$. On the other hand, if $M \geq K$, the degree of $Q(z)$ is $M - K$. Using this notation, Eq. 6.3 becomes

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

By linearity, the inverse transform of this sum of expressions is the sum of the inverse transforms of the two expressions. The inverse transform $q[n]$ of $Q(z)$ is found by inspection from the definition of the Z-transform Eq. 6.1: It is a finite signal of length $M - K + 1$ such that, for each $0 \leq n \leq M - K$, $q[n]$ is the coefficient of z^{-n} in $Q(z)$.

The denominator $A(z)$ is a polynomial of degree K in z^{-1}, and has therefore K roots $z_i, 1 \leq z_i \leq K$ which are the poles of $H(z)$. The *residue theorem* states that, if the K roots are distinct,

$R(z)/A(z)$ can be expressed as a sum of K first-degree partial fractions:

$$\frac{R(z)}{A(z)} = \sum_{i=1}^{K} \frac{R_i}{1 - z_i z^{-1}}, \tag{6.13}$$

where the *residues* R_i are constants given by

$$R_i = \left[H(z)(1 - z_i z^{-1})\right]_{z=z_i}$$

The inverse Z-transform of each term is readily recognized by applying Eq. 6.5 for $a = z_i$:

$$R_i z_i^n u[n] \longleftrightarrow \frac{R_i}{1 - z_i z^{-1}}$$

Thus, the inverse Z-transform of $H(z)$ is a sum of K complex exponentials plus the finite signal $q[n]$:

$$h[n] = q[n] + u[n] \sum_{i=1}^{K} R_i z_i^n$$

The residue theorem also applies if $A(z)$ has multiple roots, but the expression for the residue is more complex. Specifically, if z_i is a root of multiplicity m, the partial fraction expansion includes a sum of m terms of the form:

$$\sum_{j=1}^{m} \frac{R_{ij}}{(1 - z_i z^{-1})^j}, \quad \text{with} \quad R_{ij} = \frac{(-z_i)^{-(m-j)}}{(m-j)!} \left[\frac{d^{m-j}}{dz^{-(m-j)}} H(z)(1 - z_i z^{-1})^m\right]_{z=z_i}$$

Making repeated use of the ramp-multiplication theorem, this means that the inverse transform $h[n]$ includes a sum of terms of the form $n^{j-1} z_i^n u[n]$ for $1 \leq j \leq m$.

6.3.1 Finding unit-sample responses of IIR filters defined by LCCDEs

The inverse Z-transform can be used to determine the unit-sample response of an IIR filter defined by a linear, constant-coefficient difference equation.

Example: Second-order filter

For example, consider the second-order filter

$$y[n] = 1.4y[n-1] - 0.48y[n-2] + 5x[n] - 6x[n-1] + 2.4x[n-2]$$

From Eq. 6.3, its Z-transform is

$$H(z) = \frac{5 - 6z^{-1} + 2.4z^{-2}}{1 - 1.4z^{-1} + 0.48z^{-2}} = \frac{5 - 6z^{-1} + 2.4z^{-2}}{(1 - 0.8z^{-1})(1 - 0.6z^{-1})}$$

There are two poles at $z = 0.8$ and $z = 0.6$.

The degree of the numerator is the same as that of the denominator. Therefore, the first step is to carry out a polynomial division:

$$H(z) = 5 + \frac{z^{-1}}{1 - 1.4z^{-1} + 0.48z^{-2}}$$

Using the residue theorem Eq. 6.13, the second term can be expanded into a sum of two first-order partial fractions:

$$H(z) = 5 + \frac{R_1}{1 - 0.8z^{-1}} + \frac{R_2}{1 - 0.6z^{-1}}$$

where the residues are given by:

$$R_1 = \Big[H(z)(1 - 0.8z^{-1})\Big]_{z=0.8} = \frac{\frac{1}{0.8}}{1 - \frac{0.6}{0.8}} = 5$$

$$R_2 = \Big[H(z)(1 - 0.6z^{-1})\Big]_{z=0.6} = \frac{\frac{1}{0.6}}{1 - \frac{0.8}{0.6}} = -5$$

Thus, the unit-sample response is

$$h[n] = 5\delta[n] + [5(0.8)^n - 5(0.6)^n]\, u[n]$$

Example: Repeated poles

Consider the difference equation

$$y[n] = 1.8y[n-1] - 0.81y[n-2] + x[n] - x[n-1]$$

The Z-transform

$$H(z) = \frac{1 - z^{-1}}{1 - 1.8z^{-1} + 0.81z^{-2}} = \frac{1 - z^{-1}}{(1 - 0.9z^{-1})^2}$$

has a double pole for $z = 0.9$, so that it can be expanded into the partial fraction

$$H(z) = \frac{R_1}{1 - 0.9z^{-1}} + \frac{R_2}{(1 - 0.9z^{-1})^2}$$

with

$$R_1 = -0.9^{-1}\left[\frac{d}{dz^{-1}}H(z)(1 - 0.9z^{-1})^2\right]_z = 0.9 = \frac{10}{9}$$

$$R_2 = \Big[H(z)(1 - 0.9z^{-1})^2\Big]_{z=0.9} = 1 - \frac{1}{0.9} = -\frac{1}{9}$$

Using Eq. 6.10 and the delay theorem (Eq. 6.8), we conclude that the unit-sample response is:

$$h[n] = \left[\frac{10}{9}(0.9)^n - \frac{1}{9}(n+1)0.9^n\right] u[n] = (1 - \frac{n}{9})0.9^n u[n]$$

6.4 Laplace transform (optional)

The *Laplace transform* plays the same role for analog filters as the Z-transform does for digital filters. The Laplace transform of a signal $h(t)$ is a function of the complex variable s defined by the integral

$$H(s) \triangleq \int_{-\infty}^{\infty} h(t)e^{-st}dt$$

In general, this integral is only mathematically defined for certain values of s that constitute the region of convergence of the Laplace transform. For causal filters, the region of convergence is a right-sided half plane, i.e. the set of s such that the real part $\mathrm{Re(s)} > \mathrm{x}_0$. For stable filters, the region of convergence must include the imaginary axis, corresponsing to $\mathrm{Re(s)} = 0$.

Laplace transforms verify a convolution theorem similar to that for Z-transforms:

$$h(t) * x(t) \longleftrightarrow H(s)X(s)$$

The Laplace transforms of filters defined by linear, constant coefficient differential of the form

$$\sum_{k=0}^{K} a_k \frac{d^k y(t)}{dt^k} = \sum_{m=0}^{M} b_m \frac{d^m x(t)}{dt^m}$$

are rational functions of s:

$$H(s) = \frac{\sum_{m=0}^{M} b_m s^m}{\sum_{k=0}^{K} a_k s^k} \tag{6.14}$$

This expression has K poles and M zeroes. The inverse Laplace transform of rational functions can be determined by expanding Eq. 6.14 into a partial fraction, and noting that the transform of each first-order term is

$$e^{-\alpha t}u(t) \longleftrightarrow \frac{1}{s+\alpha}$$

Thus, the impulse responses of analog filters that have rational Laplace transforms are finite sums of complex exponential signals.

For causal filters, the region of convergence is the right-half plane bounded by the pole that has the largest real part. For causal, stable filters, this means that all the poles of the transform must have negative real parts.

Summary

The Z-transform simplifies the analysis and design of digital filters, particularly infinite-impulse-response (IIR) filters. The Z-transform of a discrete-time signal $h[n]$ is

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

The simplifications provided by the Z-transforms are largely due to the convolution theorem, which states that the transform of a convolution of signals is the product of the transforms.

Filters defined by linear, constant-coefficient difference equations have rational Z-transforms. Their impulse responses, which can be determined by expanding the transform into a sum of partial fractions, are finite sums of complex exponential signals. Causal filters defined by a linear, constant coefficient difference equation are stable if and only if all their poles are inside the unit circle. Unlike the DTFT, the Z-transform exists for unstable systems as well as stable systems. Therefore, it is a useful tool for analyzing potentially unstable systems and determining the conditions required to insure stability.

Further reading

- *Oppenheim and Schafer,* Chapter 3
- *Karu,* Chapter 7
- *Siebert,* Chapter 8, Chapter 10
- *Oppenheim, Willsky and Nawab,* Chapters 9 and 10