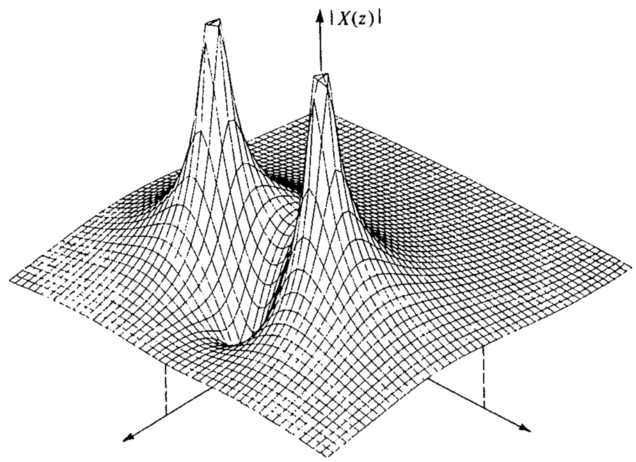

Figure 3.3.4 Graph of $|X(z)|$ for the z-transform in (3.3.3).

3.3.2 Pole Location and Time-Domain Behavior for Causal Signals

In this subsection we consider the relation between the z-plane location of a pole pair and the form (shape) of the corresponding signal in the time domain. The discussion is based generally on the collection of z-transform pairs given in Table 3.3 and the results in the preceding subsection. We deal exclusively with real, causal signals. In particular, we see that the characteristic behavior of causal signals depends on whether the poles of the transform are contained in the region $|z| < 1$, or in the region $|z| > 1$, or on the circle $|z| = 1$. Since the circle $|z| = 1$ has a radius of 1, it is called the *unit circle*.

If a real signal has a z-transform with one pole, this pole has to be real. The only such signal is the real exponential

$$x(n) = a^n u(n) \overset{z}{\longleftrightarrow} X(z) = \frac{1}{1 - az^{-1}}, \qquad \text{ROC: } |z| > |a|$$

having one zero at $z_1 = 0$ and one pole at $p_1 = a$ on the real axis. Figure 3.3.5 illustrates the behavior of the signal with respect to the location of the pole relative to the unit circle. The signal is decaying if the pole is inside the unit circle, fixed if the pole is on the unit circle, and growing if the pole is outside the unit circle. In addition, a negative pole results in a signal that alternates in sign. Obviously, causal signals with poles outside the unit circle become unbounded, cause overflow in digital systems, and in general, should be avoided.

A causal real signal with a double real pole has the form

$$x(n) = na^n u(n)$$

(see Table 3.3) and its behavior is illustrated in Fig. 3.3.6. Note that in contrast to the single-pole signal, a double real pole on the unit circle results in an unbounded signal.

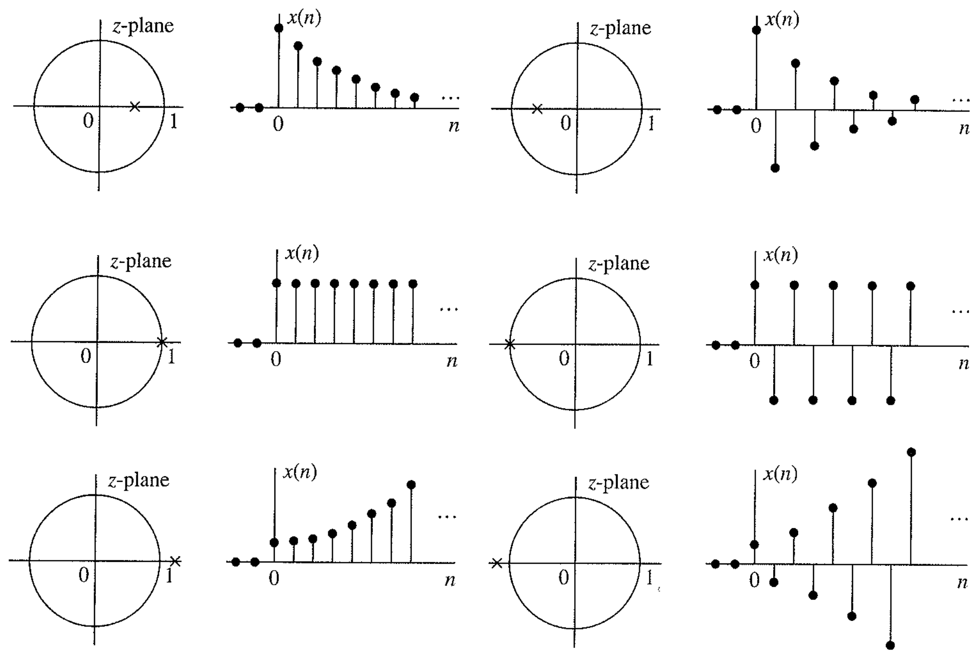

Figure 3.3.5 Time-domain behavior of a single-real-pole causal signal as a function of the location of the pole with respect to the unit circle.

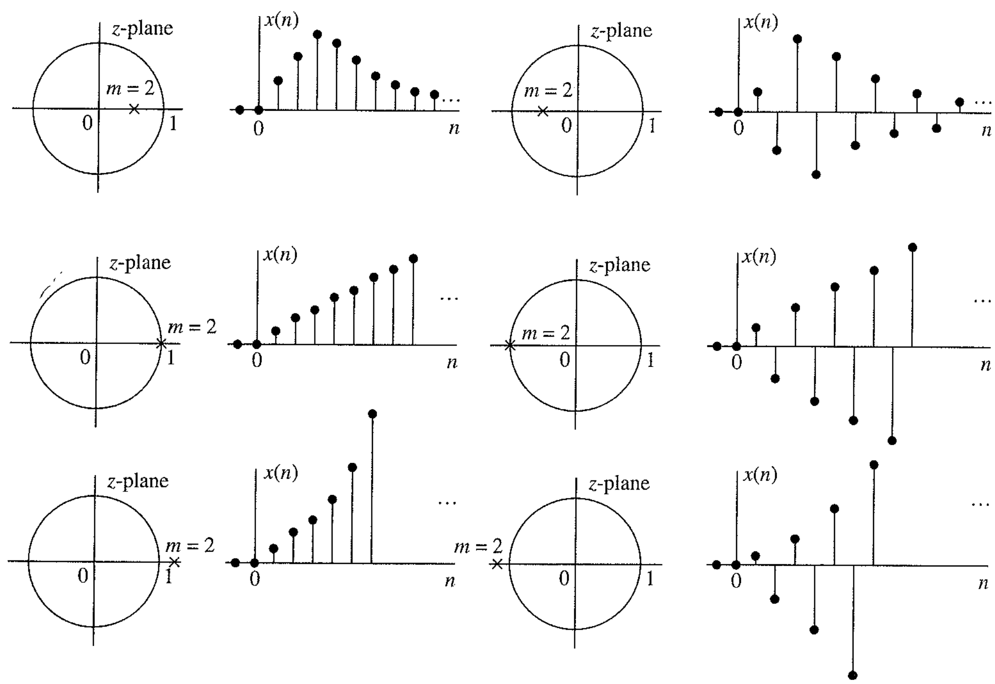

Figure 3.3.6 Time-domain behavior of causal signals corresponding to a double ($m = 2$) real pole, as a function of the pole location.

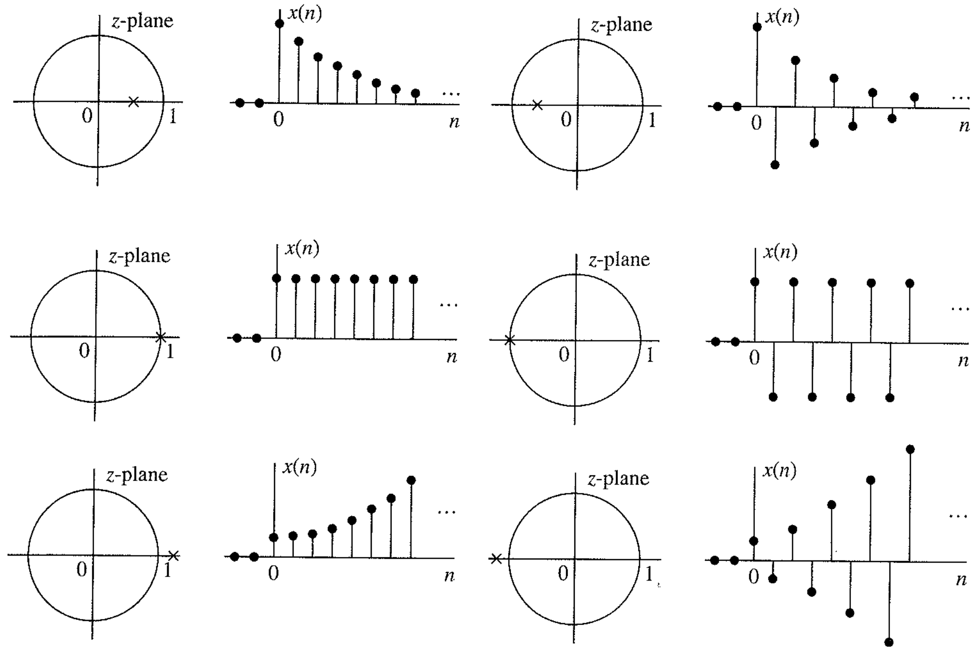

Figure 3.3.5 Time-domain behavior of a single-real-pole causal signal as a function of the location of the pole with respect to the unit circle.

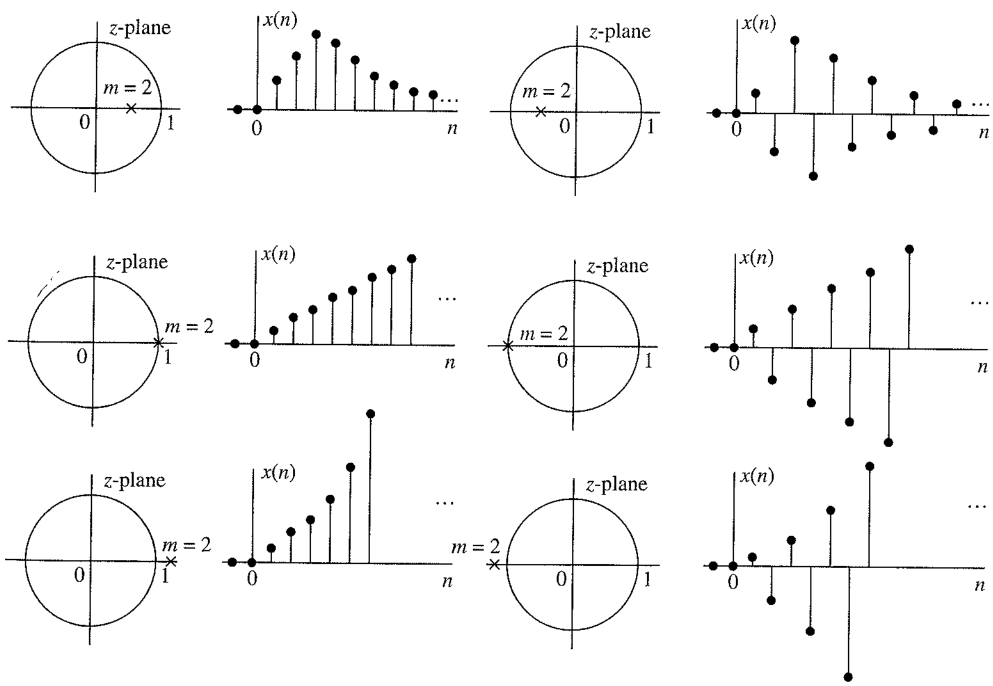

Figure 3.3.6 Time-domain behavior of causal signals corresponding to a double ($m = 2$) real pole, as a function of the pole location.

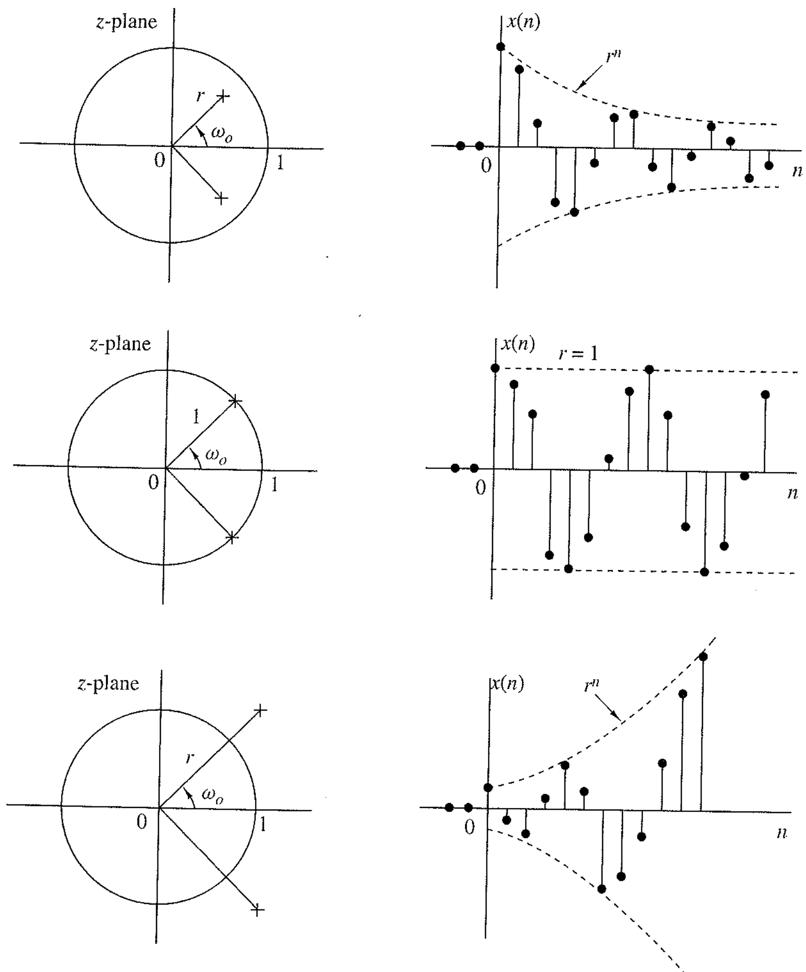

Figure 3.3.7 A pair of complex-conjugate poles corresponds to causal signals with oscillatory behavior.

Figure 3.3.7 illustrates the case of a pair of complex-conjugate poles. According to Table 3.3, this configuration of poles results in an exponentially weighted sinusoidal signal. The distance r of the poles from the origin determines the envelope of the sinusoidal signal and their angle with the real positive axis, its relative frequency. Note that the amplitude of the signal is growing if $r > 1$, constant if $r = 1$ (sinusoidal signals), and decaying if $r < 1$.

Finally, Fig. 3.3.8 shows the behavior of a causal signal with a double pair of poles on the unit circle. This reinforces the corresponding results in Fig. 3.3.6 and illustrates that multiple poles on the unit circle should be treated with great care.

To summarize, causal real signals with simple real poles or simple complex-conjugate pairs of poles, which are inside or on the unit circle, are always bounded in amplitude. Furthermore, a signal with a pole (or a complex-conjugate pair of poles)

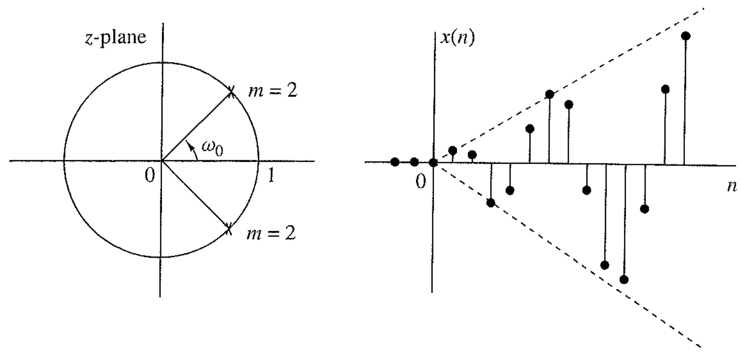

Figure 3.3.8 Causal signal corresponding to a double pair of complex-conjugate poles on the unit circle.

near the origin decays more rapidly than one associated with a pole near (but inside) the unit circle. Thus the time behavior of a signal depends strongly on the location of its poles relative to the unit circle. Zeros also affect the behavior of a signal but not as strongly as poles. For example, in the case of sinusoidal signals, the presence and location of zeros affects only their phase.

At this point, it should be stressed that everything we have said about causal signals applies as well to causal LTI systems, since their impulse response is a causal signal. Hence if a pole of a system is outside the unit circle, the impulse response of the system becomes unbounded and, consequently, the system is unstable.

3.3.3 The System Function of a Linear Time-Invariant System

In Chapter 2 we demonstrated that the output of a (relaxed) linear time-invariant system to an input sequence $x(n)$ can be obtained by computing the convolution of $x(n)$ with the unit sample response of the system. The convolution property, derived in Section 3.2, allows us to express this relationship in the z-domain as

$$Y(z) = H(z)X(z) \tag{3.3.4}$$

where $Y(z)$ is the z-transform of the output sequence $y(n)$, $X(z)$ is the z-transform of the input sequence $x(n)$ and $H(z)$ is the z-transform of the unit sample response $h(n)$.

If we know $h(n)$ and $x(n)$, we can determine their corresponding z-transforms $H(z)$ and $X(z)$, multiply them to obtain $Y(z)$, and therefore determine $y(n)$ by evaluating the inverse z-transform of $Y(z)$. Alternatively, if we know $x(n)$ and we observe the output $y(n)$ of the system, we can determine the unit sample response by first solving for $H(z)$ from the relation

$$H(z) = \frac{Y(z)}{X(z)} \tag{3.3.5}$$

and then evaluating the inverse z-transform of $H(z)$.

Since

$$H(z) = \sum_{n=-\infty}^{\infty} h(n) z^{-n} \tag{3.3.6}$$

it is clear that $H(z)$ represents the z-domain characterization of a system, whereas $h(n)$ is the corresponding time-domain characterization of the system. In other words, $H(z)$ and $h(n)$ are equivalent descriptions of a system in the two domains. The transform $H(z)$ is called the *system function*.

The relation in (3.3.5) is particularly useful in obtaining $H(z)$ when the system is described by a linear constant-coefficient difference equation of the form

$$y(n) = -\sum_{k=1}^{N} a_k y(n-k) + \sum_{k=0}^{M} b_k x(n-k) \tag{3.3.7}$$

In this case the system function can be determined directly from (3.3.7) by computing the z-transform of both sides of (3.3.7). Thus, by applying the time-shifting property, we obtain

$$\begin{aligned}
Y(z) &= -\sum_{k=1}^{N} a_k Y(z) z^{-k} + \sum_{k=0}^{M} b_k X(z) z^{-k} \\
Y(z)\left(1 + \sum_{k=1}^{N} a_k z^{-k}\right) &= X(z)\left(\sum_{k=0}^{M} b_k z^{-k}\right) \\
\frac{Y(z)}{X(z)} = H(z) &= \frac{\sum_{k=0}^{M} b_k z^{-k}}{1 + \sum_{k=1}^{N} a_k z^{-k}}
\end{aligned} \tag{3.3.8}$$

Therefore, a linear time-invariant system described by a constant-coefficient difference equation has a rational system function.

This is the general form for the system function of a system described by a linear constant-coefficient difference equation. From this general form we obtain two important special forms. First, if $a_k = 0$ for $1 \le k \le N$, (3.3.8) reduces to

$$H(z) = \sum_{k=0}^{M} b_k z^{-k} = \frac{1}{z^M} \sum_{k=0}^{M} b_k z^{M-k} \tag{3.3.9}$$

In this case, $H(z)$ contains M zeros, whose values are determined by the system parameters $\{b_k\}$, and an Mth-order pole at the origin $z = 0$. Since the system contains only trivial poles (at $z = 0$) and M nontrivial zeros, it is called an *all-zero system*. Clearly, such a system has a finite-duration impulse response (FIR), and it is called an FIR system or a moving average (MA) system.

On the other hand, if $b_k = 0$ for $1 \le k \le M$, the system function reduces to

$$H(z) = \frac{b_0}{1 + \sum_{k=1}^{N} a_k z^{-k}} = \frac{b_0 z^N}{\sum_{k=0}^{N} a_k z^{N-k}}, \qquad a_0 \equiv 1 \tag{3.3.10}$$

In this case $H(z)$ consists of N poles, whose values are determined by the system parameters $\{a_k\}$ and an Nth-order zero at the origin $z = 0$. We usually do not make reference to these trivial zeros. Consequently, the system function in (3.3.10) contains only nontrivial poles and the corresponding system is called an *all-pole system*. Due to the presence of poles, the impulse response of such a system is infinite in duration, and hence it is an IIR system.

The general form of the system function given by (3.3.8) contains both poles and zeros, and hence the corresponding system is called a *pole–zero system*, with N poles and M zeros. Poles and/or zeros at $z = 0$ and $z = \infty$ are implied but are not counted explicitly. Due to the presence of poles, a pole–zero system is an IIR system.

The following example illustrates the procedure for determining the system function and the unit sample response from the difference equation.

EXAMPLE 3.3.4

Determine the system function and the unit sample response of the system described by the difference equation

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

Solution. By computing the z-transform of the difference equation, we obtain

$$Y(z) = \frac{1}{2}z^{-1}Y(z) + 2X(z)$$

Hence the system function is

$$H(z) = \frac{Y(z)}{X(z)} = \frac{2}{1 - \frac{1}{2}z^{-1}}$$

This system has a pole at $z = \frac{1}{2}$ and a zero at the origin. Using Table 3.3 we obtain the inverse transform

$$h(n) = 2(\frac{1}{2})^n u(n)$$

This is the unit sample response of the system.

We have now demonstrated that rational z-transforms are encountered in commonly used systems and in the characterization of linear time-invariant systems. In Section 3.4 we describe several methods for determining the inverse z-transform of rational functions.

3.4 Inversion of the z-Transform

As we saw in Section 3.1.2, the inverse z-transform is formally given by

$$x(n) = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz \tag{3.4.1}$$

where the integral is a contour integral over a closed path C that encloses the origin and lies within the region of convergence of $X(z)$. For simplicity, C can be taken as a circle in the ROC of $X(z)$ in the z-plane.

There are three methods that are often used for the evaluation of the inverse z-transform in practice:

1. Direct evaluation of (3.4.1), by contour integration.
2. Expansion into a series of terms, in the variables z, and z^{-1}.
3. Partial-fraction expansion and table lookup.

3.4.1 The Inverse z-Transform by Contour Integration

In this section we demonstrate the use of the Cauchy's integral theorem to determine the inverse z-transform directly from the contour integral.

Cauchy's integral theorem. Let $f(z)$ be a function of the complex variable z and C be a closed path in the z-plane. If the derivative $df(z)/dz$ exists on and inside the contour C and if $f(z)$ has no poles at $z = z_0$, then

$$\frac{1}{2\pi j} \oint_C \frac{f(z)}{z - z_0} dz = \begin{cases} f(z_0), & \text{if } z_0 \text{ is inside } C \\ 0, & \text{if } z_0 \text{ is outside } C \end{cases} \tag{3.4.2}$$

More generally, if the $(k+1)$-order derivative of $f(z)$ exists and $f(z)$ has no poles at $z = z_0$, then

$$\frac{1}{2\pi j} \oint_C \frac{f(z)}{(z - z_0)^k} dz = \begin{cases} \dfrac{1}{(k-1)!} \left.\dfrac{d^{k-1} f(z)}{dz^{k-1}}\right|_{z=z_0}, & \text{if } z_0 \text{ is inside } C \\ 0, & \text{if } z_0 \text{ is outside } C \end{cases} \tag{3.4.3}$$

The values on the right-hand side of (3.4.2) and (3.4.3) are called the residues of the pole at $z = z_0$. The results in (3.4.2) and (3.4.3) are two forms of the *Cauchy's integral theorem.*

We can apply (3.4.2) and (3.4.3) to obtain the values of more general contour integrals. To be specific, suppose that the integrand of the contour integral is a proper fraction $f(z)/g(z)$, where $f(z)$ has no poles inside the contour C and $g(z)$ is a polynomial with distinct (simple) roots $z_1, z_2, \ldots, z_n$ inside C. Then

$$\begin{aligned}\frac{1}{2\pi j}\oint_C \frac{f(z)}{g(z)}\,dz &= \frac{1}{2\pi j}\oint_C \left[\sum_{i=1}^{n} \frac{A_i}{z-z_i}\right] dz \\ &= \sum_{i=1}^{n} \frac{1}{2\pi j}\oint_C \frac{A_i}{z-z_i}\,dz \\ &= \sum_{i=1}^{n} A_i\end{aligned} \tag{3.4.4}$$

where

$$A_i = (z-z_i)\left.\frac{f(z)}{g(z)}\right|_{z=z_i} \tag{3.4.5}$$

The values $\{A_i\}$ are residues of the corresponding poles at $z = z_i$, $i = 1, 2, \ldots, n$. Hence the value of the contour integral is equal to the sum of the residues of all the poles inside the contour C.

We observe that (3.4.4) was obtained by performing a partial-fraction expansion of the integrand and applying (3.4.2). When $g(z)$ has multiple-order roots as well as simple roots inside the contour, the partial-fraction expansion, with appropriate modifications, and (3.4.3) can be used to evaluate the residues at the corresponding poles.

In the case of the inverse z-transform, we have

$$\begin{aligned}x(n) &= \frac{1}{2\pi j}\oint_C X(z)z^{n-1}\,dz \\ &= \sum_{\text{all poles } \{z_i\} \text{ inside } C} [\text{residue of } X(z)z^{n-1} \text{ at } z = z_i] \\ &= \sum_i (z-z_i)X(z)z^{n-1}|_{z=z_i}\end{aligned} \tag{3.4.6}$$

provided that the poles $\{z_i\}$ are simple. If $X(z)z^{n-1}$ has no poles inside the contour C for one or more values of n, then $x(n) = 0$ for these values.

The following example illustrates the evaluation of the inverse z-transform by use of the Cauchy's integral theorem.

EXAMPLE 3.4.1

Evaluate the inverse z-transform of

$$X(z) = \frac{1}{1-az^{-1}}, \qquad |z| > |a|$$

using the complex inversion integral.

Solution. We have

$$x(n) = \frac{1}{2\pi j} \oint_C \frac{z^{n-1}}{1 - az^{-1}}\, dz = \frac{1}{2\pi j} \oint_C \frac{z^n\, dz}{z - a}$$

where C is a circle at radius greater than $|a|$. We shall evaluate this integral using (3.4.2) with $f(z) = z^n$. We distinguish two cases.

1. If $n \geq 0$, $f(z)$ has only zeros and hence no poles inside C. The only pole inside C is $z = a$. Hence

$$x(n) = f(z_0) = a^n, \qquad n \geq 0$$

2. If $n < 0$, $f(z) = z^n$ has an nth-order pole at $z = 0$, which is also inside C. Thus there are contributions from both poles. For $n = -1$ we have

$$x(-1) = \frac{1}{2\pi j} \oint_C \frac{1}{z(z-a)}\, dz = \left.\frac{1}{z-a}\right|_{z=0} + \left.\frac{1}{z}\right|_{z=a} = 0$$

If $n = -2$, we have

$$x(-2) = \frac{1}{2\pi j} \oint_C \frac{1}{z^2(z-a)}\, dz = \left.\frac{d}{dz}\left(\frac{1}{z-a}\right)\right|_{z=0} + \left.\frac{1}{z^2}\right|_{z=a} = 0$$

By continuing in the same way we can show that $x(n) = 0$ for $n < 0$. Thus

$$x(n) = a^n u(n)$$

3.4.2 The Inverse z-Transform by Power Series Expansion

The basic idea in this method is the following: Given a z-transform $X(z)$ with its corresponding ROC, we can expand $X(z)$ into a power series of the form

$$X(z) = \sum_{n=-\infty}^{\infty} c_n z^{-n} \tag{3.4.7}$$

which converges in the given ROC. Then, by the uniqueness of the z-transform, $x(n) = c_n$ for all n. When $X(z)$ is rational, the expansion can be performed by long division.

To illustrate this technique, we will invert some z-transforms involving the same expression for $X(z)$, but different ROC. This will also serve to emphasize again the importance of the ROC in dealing with z-transforms.

EXAMPLE 3.4.2

Determine the inverse z-transform of

$$X(z) = \frac{1}{1 - 1.5z^{-1} + 0.5z^{-2}}$$

when

(a) ROC: $|z| > 1$

(b) ROC: $|z| < 0.5$

Solution.

(a) Since the ROC is the exterior of a circle, we expect $x(n)$ to be a causal signal. Thus we seek a power series expansion in negative powers of z. By dividing the numerator of $X(z)$ by its denominator, we obtain the power series

$$X(z) = \frac{1}{1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2}} = 1 + \frac{3}{2}z^{-1} + \frac{7}{4}z^{-2} + \frac{15}{8}z^{-3} + \frac{31}{16}z^{-4} + \cdots$$

By comparing this relation with (3.1.1), we conclude that

$$x(n) = \{\underset{\uparrow}{1}, \frac{3}{2}, \frac{7}{4}, \frac{15}{8}, \frac{31}{16}, \ldots\}$$

Note that in each step of the long-division process, we eliminate the lowest-power term of z^{-1}.

(b) In this case the ROC is the interior of a circle. Consequently, the signal $x(n)$ is anticausal. To obtain a power series expansion in positive powers of z, we perform the long division in the following way:

$$\begin{array}{r|l}
 & 2z^2 + 6z^3 + 14z^4 + 30z^5 + 62z^6 + \cdots \\
\hline
\frac{1}{2}z^{-2} - \frac{3}{2}z^{-1} + 1 & 1 \\
 & \underline{1 - 3z + 2z^2} \\
 & \quad 3z - 2z^2 \\
 & \quad \underline{3z - 9z^2 + 6z^3} \\
 & \qquad 7z^2 - 6z^3 \\
 & \qquad \underline{7z^2 - 21z^3 + 14z^4} \\
 & \qquad\quad 15z^3 - 14z^4 \\
 & \qquad\quad \underline{15z^3 - 45z^4 + 30z^5} \\
 & \qquad\qquad 31z^4 - 30z^5
\end{array}$$

Thus

$$X(z) = \frac{1}{1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2}} = 2z^2 + 6z^3 + 14z^4 + 30z^5 + 62z^6 + \cdots$$

In this case $x(n) = 0$ for $n \geq 0$. By comparing this result to (3.1.1), we conclude that

$$x(n) = \{\cdots 62, 30, 14, 6, 2, 0, \underset{\uparrow}{0}\}$$

We observe that in each step of the long-division process, the lowest-power term of z is eliminated. We emphasize that in the case of anticausal signals we simply carry out the long division by writing down the two polynomials in "reverse" order (i.e., starting with the most negative term on the left).

From this example we note that, in general, the method of long division will not provide answers for $x(n)$ when n is large because the long division becomes tedious. Although the method provides a direct evaluation of $x(n)$, a closed-form solution is not possible, except if the resulting pattern is simple enough to infer the general term $x(n)$. Hence this method is used only if one wishes to determine the values of the first few samples of the signal.

EXAMPLE 3.4.3

Determine the inverse z-transform of

$$X(z) = \log(1 + az^{-1}), \qquad |z| > |a|$$

Solution. Using the power series expansion for $\log(1 + x)$, with $|x| < 1$, we have

$$X(z) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} a^n z^{-n}}{n}$$

Thus

$$x(n) = \begin{cases} (-1)^{n+1}\frac{a^n}{n}, & n \geq 1 \\ 0, & n \leq 0 \end{cases}$$

Expansion of irrational functions into power series can be obtained from tables.

3.4.3 The Inverse z-Transform by Partial-Fraction Expansion

In the table lookup method, we attempt to express the function $X(z)$ as a linear combination

$$X(z) = \alpha_1 X_1(z) + \alpha_2 X_2(z) + \cdots + \alpha_K X_K(z) \tag{3.4.8}$$

where $X_1(z), \ldots, X_K(z)$ are expressions with inverse transforms $x_1(n), \ldots, x_K(n)$ available in a table of z-transform pairs. If such a decomposition is possible, then $x(n)$, the inverse z-transform of $X(z)$, can easily be found using the linearity property as

$$x(n) = \alpha_1 x_1(n) + \alpha_2 x_2(n) + \cdots + \alpha_K x_K(n) \tag{3.4.9}$$

This approach is particularly useful if $X(z)$ is a rational function, as in (3.3.1). Without loss of generality, we assume that $a_0 = 1$, so that (3.3.1) can be expressed as

$$X(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}}{1 + a_1 z^{-1} + \cdots + a_N z^{-N}} \tag{3.4.10}$$

Note that if $a_0 \neq 1$, we can obtain (3.4.10) from (3.3.1) by dividing both numerator and denominator by a_0.

A rational function of the form (3.4.10) is called *proper* if $a_N \neq 0$ and $M < N$. From (3.3.2) it follows that this is equivalent to saying that the number of finite zeros is less than the number of finite poles.

An improper rational function ($M \geq N$) can always be written as the sum of a polynomial and a proper rational function. This procedure is illustrated by the following example.

EXAMPLE 3.4.4

Express the improper rational transform

$$X(z) = \frac{1 + 3z^{-1} + \frac{11}{6}z^{-2} + \frac{1}{3}z^{-3}}{1 + \frac{5}{6}z^{-1} + \frac{1}{6}z^{-2}}$$

in terms of a polynomial and a proper function.

Solution. First, we note that we should reduce the numerator so that the terms z^{-2} and z^{-3} are eliminated. Thus we should carry out the long division with these two polynomials written in *reverse* order. We stop the division when the order of the remainder becomes z^{-1}. Then we obtain

$$X(z) = 1 + 2z^{-1} + \frac{\frac{1}{6}z^{-1}}{1 + \frac{5}{6}z^{-1} + \frac{1}{6}z^{-2}}$$

In general, any improper rational function ($M \geq N$) can be expressed as

$$X(z) = \frac{B(z)}{A(z)} = c_0 + c_1 z^{-1} + \cdots + c_{M-N} z^{-(M-N)} + \frac{B_1(z)}{A(z)} \tag{3.4.11}$$

The inverse z-transform of the polynomial can easily be found by inspection. We focus our attention on the inversion of proper rational transforms, since any improper function can be transformed into a proper function by using (3.4.11). We carry out the development in two steps. First, we perform a partial fraction expansion of the proper rational function and then we invert each of the terms.

Let $X(z)$ be a proper rational function, that is,

$$X(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}}{1 + a_1 z^{-1} + \cdots + a_N z^{-N}} \tag{3.4.12}$$

where

$$a_N \neq 0 \qquad \text{and} \qquad M < N$$

To simplify our discussion we eliminate negative powers of z by multiplying both the numerator and denominator of (3.4.12) by z^N. This results in

$$X(z) = \frac{b_0 z^N + b_1 z^{N-1} + \cdots + b_M z^{N-M}}{z^N + a_1 z^{N-1} + \cdots + a_N} \tag{3.4.13}$$

which contains only positive powers of z. Since $N > M$, the function

$$\frac{X(z)}{z} = \frac{b_0 z^{N-1} + b_1 z^{N-2} + \cdots + b_M z^{N-M-1}}{z^N + a_1 z^{N-1} + \cdots + a_N} \tag{3.4.14}$$

is also always proper.

Our task in performing a partial-fraction expansion is to express (3.4.14) or, equivalently, (3.4.12) as a sum of simple fractions. For this purpose we first factor the denominator polynomial in (3.4.14) into factors that contain the poles $p_1, p_2, \ldots, p_N$ of $X(z)$. We distinguish two cases.

Distinct poles. Suppose that the poles $p_1, p_2, \ldots, p_N$ are all different (distinct). Then we seek an expansion of the form

$$\frac{X(z)}{z} = \frac{A_1}{z - p_1} + \frac{A_2}{z - p_2} + \cdots + \frac{A_N}{z - p_N} \tag{3.4.15}$$

The problem is to determine the coefficients $A_1, A_2, \ldots, A_N$. There are two ways to solve this problem, as illustrated in the following example.

EXAMPLE 3.4.5

Determine the partial-fraction expansion of the proper function

$$X(z) = \frac{1}{1 - 1.5z^{-1} + 0.5z^{-2}} \tag{3.4.16}$$

Solution. First we eliminate the negative powers, by multiplying both numerator and denominator by z^2. Thus

$$X(z) = \frac{z^2}{z^2 - 1.5z + 0.5}$$

The poles of $X(z)$ are $p_1 = 1$ and $p_2 = 0.5$. Consequently, the expansion of the form (3.4.15) is

$$\frac{X(z)}{z} = \frac{z}{(z-1)(z-0.5)} = \frac{A_1}{z-1} + \frac{A_2}{z-0.5} \tag{3.4.17}$$

A very simple method to determine A_1 and A_2 is to multiply the equation by the denominator term $(z-1)(z-0.5)$. Thus we obtain

$$z = (z - 0.5)A_1 + (z - 1)A_2 \tag{3.4.18}$$

Now if we set $z = p_1 = 1$ in (3.4.18), we eliminate the term involving A_2. Hence

$$1 = (1 - 0.5)A_1$$

Thus we obtain the result $A_1 = 2$. Next we return to (3.4.18) and set $z = p_2 = 0.5$, thus eliminating the term involving A_1, so we have

$$0.5 = (0.5 - 1)A_2$$

and hence $A_2 = -1$. Therefore, the result of the partial-fraction expansion is

$$\frac{X(z)}{z} = \frac{2}{z-1} - \frac{1}{z-0.5} \tag{3.4.19}$$

The example given above suggests that we can determine the coefficients A_1, $A_2, \ldots, A_N$, by multiplying both sides of (3.4.15) by each of the terms $(z - p_k)$, $k = 1, 2, \ldots, N$, and evaluating the resulting expressions at the corresponding pole positions, $p_1, p_2, \ldots, p_N$. Thus we have, in general,

$$\frac{(z-p_k)X(z)}{z} = \frac{(z-p_k)A_1}{z-p_1} + \cdots + A_k + \cdots + \frac{(z-p_k)A_N}{z-p_N} \tag{3.4.20}$$

Consequently, with $z = p_k$, (3.4.20) yields the kth coefficient as

$$A_k = \left.\frac{(z-p_k)X(z)}{z}\right|_{z=p_k}, \qquad k = 1, 2, \ldots, N \tag{3.4.21}$$

EXAMPLE 3.4.6

Determine the partial-fraction expansion of

$$X(z) = \frac{1 + z^{-1}}{1 - z^{-1} + 0.5z^{-2}} \tag{3.4.22}$$

Solution. To eliminate negative powers of z in (3.4.22), we multiply both numerator and denominator by z^2. Thus

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

The poles of $X(z)$ are complex conjugates

$$p_1 = \frac{1}{2} + j\frac{1}{2}$$

and

$$p_2 = \frac{1}{2} - j\frac{1}{2}$$

Since $p_1 \neq p_2$, we seek an expansion of the form (3.4.15). Thus

$$\frac{X(z)}{z} = \frac{z+1}{(z-p_1)(z-p_2)} = \frac{A_1}{z-p_1} + \frac{A_2}{z-p_2}$$

To obtain A_1 and A_2, we use the formula (3.4.21). Thus we obtain

$$A_1 = \left.\frac{(z-p_1)X(z)}{z}\right|_{z=p_1} = \left.\frac{z+1}{z-p_2}\right|_{z=p_1} = \frac{\frac{1}{2} + j\frac{1}{2} + 1}{\frac{1}{2} + j\frac{1}{2} - \frac{1}{2} + j\frac{1}{2}} = \frac{1}{2} - j\frac{3}{2}$$

$$A_2 = \left.\frac{(z-p_2)X(z)}{z}\right|_{z=p_2} = \left.\frac{z+1}{z-p_1}\right|_{z=p_2} = \frac{\frac{1}{2} - j\frac{1}{2} + 1}{\frac{1}{2} - j\frac{1}{2} - \frac{1}{2} - j\frac{1}{2}} = \frac{1}{2} + j\frac{3}{2}$$

The expansion (3.4.15) and the formula (3.4.21) hold for both real and complex poles. The only constraint is that all poles be distinct. We also note that $A_2 = A_1^*$. It can be easily seen that this is a consequence of the fact that $p_2 = p_1^*$. In other words, *complex-conjugate poles result in complex-conjugate coefficients in the partial-fraction expansion.* This simple result will prove very useful later in our discussion.

Multiple-order poles. If $X(z)$ has a pole of multiplicity l, that is, it contains in its denominator the factor $(z - p_k)^l$, then the expansion (3.4.15) is no longer true. In this case a different expansion is needed. First, we investigate the case of a double pole (i.e., $l = 2$).

EXAMPLE 3.4.7

Determine the partial-fraction expansion of

$$X(z) = \frac{1}{(1+z^{-1})(1-z^{-1})^2} \tag{3.4.23}$$

Solution. First, we express (3.4.23) in terms of positive powers of z, in the form

$$\frac{X(z)}{z} = \frac{z^2}{(z+1)(z-1)^2}$$

$X(z)$ has a simple pole at $p_1 = -1$ and a double pole $p_2 = p_3 = 1$. In such a case the appropriate partial-fraction expansion is

$$\frac{X(z)}{z} = \frac{z^2}{(z+1)(z-1)^2} = \frac{A_1}{z+1} + \frac{A_2}{z-1} + \frac{A_3}{(z-1)^2} \tag{3.4.24}$$

The problem is to determine the coefficients A_1, A_2, and A_3.

We proceed as in the case of distinct poles. To determine A_1, we multiply both sides of (3.4.24) by $(z+1)$ and evaluate the result at $z = -1$. Thus (3.4.24) becomes

$$\frac{(z+1)X(z)}{z} = A_1 + \frac{z+1}{z-1}A_2 + \frac{z+1}{(z-1)^2}A_3$$

which, when evaluated at $z = -1$, yields

$$A_1 = \left.\frac{(z+1)X(z)}{z}\right|_{z=-1} = \frac{1}{4}$$

Next, if we multiply both sides of (3.4.24) by $(z-1)^2$, we obtain

$$\frac{(z-1)^2X(z)}{z} = \frac{(z-1)^2}{z+1}A_1 + (z-1)A_2 + A_3 \tag{3.4.25}$$

Now, if we evaluate (3.4.25) at $z = 1$, we obtain A_3. Thus

$$A_3 = \left.\frac{(z-1)2X(z)}{z}\right|_{z=1} = \frac{1}{2}$$

The remaining coefficient A_2 can be obtained by differentiating both sides of (3.4.25) with respect to z and evaluating the result at $z = 1$. Note that it is not necessary formally to carry out the differentiation of the right-hand side of (3.4.25), since all terms except A_2 vanish when we set $z = 1$. Thus

$$A_2 = \frac{d}{dz}\left[\frac{(z-1)^2X(z)}{z}\right]_{z=1} = \frac{3}{4} \tag{3.4.26}$$

The generalization of the procedure in the example above to the case of an mth-order pole $(z-p_k)^m$ is straightforward. The partial-fraction expansion must contain the terms

$$\frac{A_{1k}}{z-p_k} + \frac{A_{2k}}{(z-p_k)^2} + \cdots + \frac{A_{mk}}{(z-p_k)^m}$$

The coefficients $\{A_{ik}\}$ can be evaluated through differentiation as illustrated in Example 3.4.7 for $m = 2$.

Now that we have performed the partial-fraction expansion, we are ready to take the final step in the inversion of $X(z)$. First, let us consider the case in which $X(z)$ contains distinct poles. From the partial-fraction expansion (3.4.15), it easily follows that

$$X(z) = A_1 \frac{1}{1 - p_1 z^{-1}} + A_2 \frac{1}{1 - p_2 z^{-1}} + \cdots + A_N \frac{1}{1 - p_N z^{-1}} \tag{3.4.27}$$

The inverse z-transform, $x(n) = Z^{-1}\{X(z)\}$, can be obtained by inverting each term in (3.4.27) and taking the corresponding linear combination. From Table 3.3 it follows that these terms can be inverted using the formula

$$Z^{-1}\left\{\frac{1}{1 - p_k z^{-1}}\right\} = \begin{cases} (p_k)^n u(n), & \text{if ROC: } |z| > |p_k| \\ & \text{(causal signals)} \\ -(p_k)^n u(-n-1), & \text{if ROC: } |z| < |p_k| \\ & \text{(anticausal signals)} \end{cases} \tag{3.4.28}$$

If the signal $x(n)$ is causal, the ROC is $|z| > p_{\max}$, where $p_{\max} = \max\{|p_1|, |p_2|, \ldots, |p_N|\}$. In this case all terms in (3.4.27) result in causal signal components and the signal $x(n)$ is given by

$$x(n) = (A_1 p_1^n + A_2 p_2^n + \cdots + A_N p_N^n) u(n) \tag{3.4.29}$$

If all poles are real, (3.4.29) is the desired expression for the signal $x(n)$. Thus a causal signal, having a z-transform that contains real and distinct poles, is a linear combination of real exponential signals.

Suppose now that all poles are distinct but some of them are complex. In this case some of the terms in (3.4.27) result in complex exponential components. However, if the signal $x(n)$ is real, we should be able to reduce these terms into real components. If $x(n)$ is real, the polynomials appearing in $X(z)$ have real coefficients. In this case, as we have seen in Section 3.3, if p_j is a pole, its complex conjugate p_j^* is also a pole. As was demonstrated in Example 3.4.6, the corresponding coefficients in the partial-fraction expansion are also complex conjugates. Thus the contribution of two complex-conjugate poles is of the form

$$x_k(n) = [A_k (p_k)^n + A_k^* (p_k^*)^n] u(n) \tag{3.4.30}$$

These two terms can be combined to form a real signal component. First, we express A_j and p_j in polar form (i.e., amplitude and phase) as

$$A_k = |A_k| e^{j\alpha_k} \tag{3.4.31}$$

$$p_k = r_k e^{j\beta_k} \tag{3.4.32}$$

where α_k and β_k are the phase components of A_k and p_k. Substitution of these relations into (3.4.30) gives

$$x_k(n) = |A_k| r_k^n [e^{j(\beta_k n + \alpha_k)} + e^{-j(\beta_k n + \alpha_k)}] u(n)$$

or, equivalently,

$$x_k(n) = 2|A_k|r_k^n \cos(\beta_k n + \alpha_k)u(n) \tag{3.4.33}$$

Thus we conclude that

$$Z^{-1}\left(\frac{A_k}{1-p_k z^{-1}} + \frac{A_k^*}{1-p_k^* z^{-1}}\right) = 2|A_k|r_k^n \cos(\beta_k n + \alpha_k)u(n) \tag{3.4.34}$$

if the ROC is $|z| > |p_k| = r_k$.

From (3.4.34) we observe that each pair of complex-conjugate poles in the z-domain results in a causal sinusoidal signal component with an exponential envelope. The distance r_k of the pole from the origin determines the exponential weighting (growing if $r_k > 1$, decaying if $r_k < 1$, constant if $r_k = 1$). The angle of the poles with respect to the positive real axis provides the frequency of the sinusoidal signal. The zeros, or equivalently the numerator of the rational transform, affect only indirectly the amplitude and the phase of $x_k(n)$ through A_k.

In the case of *multiple* poles, either real or complex, the inverse transform of terms of the form $A/(z - p_k)^n$ is required. In the case of a double pole the following transform pair (see Table 3.3) is quite useful:

$$Z^{-1}\left\{\frac{pz^{-1}}{(1-pz^{-1})^2}\right\} = np^n u(n) \tag{3.4.35}$$

provided that the ROC is $|z| > |p|$. The generalization to the case of poles with higher multiplicity is obtained by using multiple differentiation.

EXAMPLE 3.4.8

Determine the inverse z-transform of

$$X(z) = \frac{1}{1 - 1.5z^{-1} + 0.5z^{-2}}$$

if

(a) ROC: $|z| > 1$

(b) ROC: $|z| < 0.5$

(c) ROC: $0.5 < |z| < 1$

Solution. This is the same problem that we treated in Example 3.4.2. The partial-fraction expansion for $X(z)$ was determined in Example 3.4.5. The partial-fraction expansion of $X(z)$ yields

$$X(z) = \frac{2}{1-z^{-1}} - \frac{1}{1-0.5z^{-1}} \tag{3.4.36}$$

To invert $X(z)$ we should apply (3.4.28) for $p_1 = 1$ and $p_2 = 0.5$. However, this requires the specification of the corresponding ROC.

(a) In the case when the ROC is $|z| > 1$, the signal $x(n)$ is causal and both terms in (3.4.36) are causal terms. According to (3.4.28), we obtain

$$x(n) = 2(1)^n u(n) - (0.5)^n u(n) = (2 - 0.5^n)u(n) \tag{3.4.37}$$

which agrees with the result in Example 3.4.2(a).

(b) When the ROC is $|z| < 0.5$, the signal $x(n)$ is anticausal. Thus both terms in (3.4.36) result in anticausal components. From (3.4.28) we obtain

$$x(n) = [-2 + (0.5)^n]u(-n-1) \tag{3.4.38}$$

(c) In this case the ROC $0.5 < |z| < 1$ is a ring, which implies that the signal $x(n)$ is two-sided. Thus one of the terms corresponds to a causal signal and the other to an anticausal signal. Obviously, the given ROC is the overlapping of the regions $|z| > 0.5$ and $|z| < 1$. Hence the pole $p_2 = 0.5$ provides the causal part and the pole $p_1 = 1$ the anticausal. Thus

$$x(n) = -2(1)^n u(-n-1) - (0.5)^n u(n) \tag{3.4.39}$$

EXAMPLE 3.4.9

Determine the causal signal $x(n)$ whose z-transform is given by

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

Solution. In Example 3.4.6 we have obtained the partial-fraction expansion as

$$X(z) = \frac{A_1}{1 - p_1 z^{-1}} + \frac{A_2}{1 - p_2 z^{-1}}$$

where

$$A_1 = A_2^* = \frac{1}{2} - j\frac{3}{2}$$

and

$$p_1 = p_2^* = \frac{1}{2} + j\frac{1}{2}$$

Since we have a pair of complex-conjugate poles, we should use (3.4.34). The polar forms of A_1 and p_1 are

$$A_1 = \frac{\sqrt{10}}{2} e^{-j71.565}$$

$$p_1 = \frac{1}{\sqrt{2}} e^{j\pi/4}$$

Hence

$$x(n) = \sqrt{10}\left(\frac{1}{\sqrt{2}}\right)^n \cos\left(\frac{\pi n}{4} - 71.565^\circ\right) u(n)$$

EXAMPLE 3.4.10

Determine the causal signal $x(n)$ having the z-transform

$$X(z) = \frac{1}{(1 + z^{-1})(1 - z^{-1})^2}$$

Solution. From Example 3.4.7 we have

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

By applying the inverse transform relations in (3.4.28) and (3.4.35), we obtain

$$x(n) = \frac{1}{4}(-1)^n u(n) + \frac{3}{4}u(n) + \frac{1}{2}nu(n) = \left[\frac{1}{4}(-1)^n + \frac{3}{4} + \frac{n}{2}\right]u(n)$$

3.4.4 Decomposition of Rational z-Transforms

At this point it is appropriate to discuss some additional issues concerning the decomposition of rational z-transforms, which will prove very useful in the implementation of discrete-time systems.

Suppose that we have a rational z-transform $X(z)$ expressed as

$$X(z) = \frac{\sum_{k=0}^{M} b_k z^{-k}}{1 + \sum_{k=1}^{N} a_k z^{-k}} = b_0 \frac{\prod_{k=1}^{M}(1 - z_k z^{-1})}{\prod_{k=1}^{N}(1 - p_k z^{-1})} \tag{3.4.40}$$

where, for simplicity, we have assumed that $a_0 \equiv 1$. If $M \geq N$ [i.e., $X(z)$ is improper], we convert $X(z)$ to a sum of a polynomial and a proper function

$$X(z) = \sum_{k=0}^{M-N} c_k z^{-k} + X_{\text{pr}}(z) \tag{3.4.41}$$

If the poles of $X_{\text{pr}}(z)$ are distinct, it can be expanded in partial fractions as

$$X_{\text{pr}}(z) = A_1 \frac{1}{1 - p_1 z^{-1}} + A_2 \frac{1}{1 - p_2 z^{-1}} + \cdots + A_N \frac{1}{1 - p_N z^{-1}} \tag{3.4.42}$$

As we have already observed, there may be some complex-conjugate pairs of poles in (3.4.42). Since we usually deal with real signals, we should avoid complex coefficients in our decomposition. This can be achieved by grouping and combining terms containing complex-conjugate poles, in the following way:

$$\begin{aligned} \frac{A}{1 - pz^{-1}} + \frac{A^*}{1 - p^* z^{-1}} &= \frac{A - Ap^* z^{-1} + A^* - A^* p z^{-1}}{1 - pz^{-1} - p^* z^{-1} + pp^* z^{-2}} \\ &= \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} \end{aligned} \tag{3.4.43}$$

where

$$\begin{aligned} b_0 &= 2\,\text{Re}(A), & a_1 &= -2\,\text{Re}(p) \\ b_1 &= 2\,\text{Re}(Ap*), & a_2 &= |p|^2 \end{aligned} \tag{3.4.44}$$

are the desired coefficients. Obviously, any rational transform of the form (3.4.43) with coefficients given by (3.4.44), which is the case when $a_1^2 - 4a_2 < 0$, can be inverted using (3.4.34). By combining (3.4.41), (3.4.42), and (3.4.43) we obtain a partial-fraction expansion for the z-transform with *distinct* poles that contains real coefficients. The general result is

$$X(z) = \sum_{k=0}^{M-N} c_k z^{-k} + \sum_{k=1}^{K_1} \frac{b_k}{1 + a_k z^{-1}} + \sum_{k=1}^{K_2} \frac{b_{0k} + b_{1k} z^{-1}}{1 + a_{1k} z^{-1} + a_{2k} z^{-2}} \tag{3.4.45}$$

where $K_1 + 2K_2 = N$. Obviously, if $M = N$, the first term is just a constant, and when $M < N$, this term vanishes. When there are also multiple poles, some additional higher-order terms should be included in (3.4.45).

An alternative form is obtained by expressing $X(z)$ as a product of simple terms as in (3.4.40). However, the complex-conjugate poles and zeros should be combined to avoid complex coefficients in the decomposition. Such combinations result in second-order rational terms of the following form:

$$\frac{(1 - z_k z^{-1})(1 - z_k^* z^{-1})}{(1 - p_k z^{-1})(1 - p_k^* z^{-1})} = \frac{1 + b_{1k} z^{-1} + b_{2k} z^{-2}}{1 + a_{1k} z^{-1} + a_{2k} z^{-2}} \tag{3.4.46}$$

where

$$\begin{aligned} b_{1k} &= -2\,\mathrm{Re}(z_k), \qquad & a_{1k} &= -2\,\mathrm{Re}(p_k) \\ b_{2k} &= |z_k|^2, \qquad & a_{2k} &= |p_k|^2 \end{aligned} \tag{3.4.47}$$

Assuming for simplicity that $M = N$, we see that $X(z)$ can be decomposed in the following way:

$$X(z) = b_0 \prod_{k=1}^{K_1} \frac{1 + b_k z^{-1}}{1 + a_k z^{-1}} \prod_{k=1}^{K_2} \frac{1 + b_{1k} z^{-1} + b_{2k} z^{-2}}{1 + a_{1k} z^{-1} + a_{2k} z^{-2}} \tag{3.4.48}$$

where $N = K_1 + 2K_2$. We will return to these important forms in Chapters 9 and 10.

3.5 Analysis of Linear Time-Invariant Systems in the z-Domain

In Section 3.3.3 we introduced the system function of a linear time-invariant system and related it to the unit sample response and to the difference equation description of systems. In this section we describe the use of the system function in the determination of the response of the system to some excitation signal. In Section 3.6.3, we extend this method of analysis to nonrelaxed systems. Our attention is focused on the important class of pole–zero systems represented by linear constant-coefficient difference equations with arbitrary initial conditions.

We also consider the topic of stability of linear time-invariant systems and describe a test for determining the stability of a system based on the coefficients of the denominator polynomial in the system function. Finally, we provide a detailed analysis of second-order systems, which form the basic building blocks in the realization of higher-order systems.

3.5.1 Response of Systems with Rational System Functions

Let us consider a pole–zero system described by the general linear constant-coefficient difference equation in (3.3.7) and the corresponding system function in (3.3.8). We represent $H(z)$ as a ratio of two polynomials $B(z)/A(z)$, where $B(z)$ is the numerator polynomial that contains the zeros of $H(z)$, and $A(z)$ is the denominator polynomial that determines the poles of $H(z)$. Furthermore, let us assume that the input signal $x(n)$ has a rational z-transform $X(z)$ of the form

$$X(z) = \frac{N(z)}{Q(z)} \tag{3.5.1}$$

This assumption is not overly restrictive, since, as indicated previously, most signals of practical interest have rational z-transforms.

If the system is initially relaxed, that is, the initial conditions for the difference equation are zero, $y(-1) = y(-2) = \cdots = y(-N) = 0$, the z-transform of the output of the system has the form

$$Y(z) = H(z)X(z) = \frac{B(z)N(z)}{A(z)Q(z)} \tag{3.5.2}$$

Now suppose that the system contains simple poles $p_1, p_2, \ldots, p_N$ and the z-transform of the input signal contains poles $q_1, q_2, \ldots, q_L$, where $p_k \neq q_m$ for all $k = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, L$. In addition, we assume that the zeros of the numerator polynomials $B(z)$ and $N(z)$ do not coincide with the poles $\{p_k\}$ and $\{q_k\}$, so that there is no pole–zero cancellation. Then a partial-fraction expansion of $Y(z)$ yields

$$Y(z) = \sum_{k=1}^{N} \frac{A_k}{1 - p_k z^{-1}} + \sum_{k=1}^{L} \frac{Q_k}{1 - q_k z^{-1}} \tag{3.5.3}$$

The inverse transform of $Y(z)$ yields the output signal from the system in the form

$$y(n) = \sum_{k=1}^{N} A_k (p_k)^n u(n) + \sum_{k=1}^{L} Q_k (q_k)^n u(n) \tag{3.5.4}$$

We observe that the output sequence $y(n)$ can be subdivided into two parts. The first part is a function of the poles $\{p_k\}$ of the system and is called the *natural response* of the system. The influence of the input signal on this part of the response is through the scale factors $\{A_k\}$. The second part of the response is a function of the poles $\{q_k\}$ of the input signal and is called the *forced response* of the system. The influence of the system on this response is exerted through the scale factors $\{Q_k\}$.

We should emphasize that the scale factors $\{A_k\}$ and $\{Q_k\}$ are functions of both sets of poles $\{p_k\}$ and $\{q_k\}$. For example, if $X(z) = 0$ so that the input is zero, then $Y(z) = 0$, and consequently, the output is zero. Clearly, then, the natural response of the system is zero. This implies that the natural response of the system is different from the zero-input response.

When $X(z)$ and $H(z)$ have one or more poles in common or when $X(z)$ and/or $H(z)$ contain multiple-order poles, then $Y(z)$ will have multiple-order poles. Consequently, the partial-fraction expansion of $Y(z)$ will contain factors of the form $1/(1-p_l z^{-1})^k$, $k = 1, 2, \ldots, m$, where m is the pole order. The inversion of these factors will produce terms of the form $n^{k-1}p_l^n$ in the output $y(n)$ of the system, as indicated in Section 3.4.3.

3.5.2 Transient and Steady-State Responses

As we have seen from our previous discussion, the zero-state response of a system to a given input can be separated into two components, the natural response and the forced response. The natural response of a causal system has the form

$$y_{\mathrm{nr}}(n) = \sum_{k=1}^{N} A_k (p_k)^n u(n) \tag{3.5.5}$$

where $\{p_k\}$, $k = 1, 2, \ldots, N$ are the poles of the system and $\{A_k\}$ are scale factors that depend on the initial conditions and on the characteristics of the input sequence.

If $|p_k| < 1$ for all k, then, $y_{\mathrm{nr}}(n)$ decays to zero as n approaches infinity. In such a case we refer to the natural response of the system as the *transient response*. The rate at which $y_{\mathrm{nr}}(n)$ decays toward zero depends on the magnitude of the pole positions. If all the poles have small magnitudes, the decay is very rapid. On the other hand, if one or more poles are located near the unit circle, the corresponding terms in $y_{\mathrm{nr}}(n)$ will decay slowly toward zero and the transient will persist for a relatively long time.

The forced response of the system has the form

$$y_{\mathrm{fr}}(n) = \sum_{k=1}^{L} Q_k (q_k)^n u(n) \tag{3.5.6}$$

where $\{q_k\}$, $k = 1, 2, \ldots, L$ are the poles in the forcing function and $\{Q_k\}$ are scale factors that depend on the input sequence and on the characteristics of the system. If all the poles of the input signal fall inside the unit circle, $y_{\mathrm{fr}}(n)$ will decay toward zero as n approaches infinity, just as in the case of the natural response. This should not be surprising since the input signal is also a transient signal. On the other hand, when the causal input signal is a sinusoid, the poles fall on the unit circle and consequently, the forced response is also a sinusoid that persists for all $n \geq 0$. In this case, the forced response is called the *steady-state response* of the system. Thus, for the system to sustain a steady-state output for $n \geq 0$, the input signal must persist for all $n \geq 0$.

The following example illustrates the presence of the steady-state response.

EXAMPLE 3.5.1

Determine the transient and steady-state responses of the system characterized by the difference equation

$$y(n) = 0.5y(n-1) + x(n)$$

when the input signal is $x(n) = 10\cos(\pi n/4)u(n)$. The system is initially at rest (i.e., it is relaxed).

Solution. The system function for this system is

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

and therefore the system has a pole at $z = 0.5$. The z-transform of the input signal is (from Table 3.3)

$$X(z) = \frac{10(1 - (1/\sqrt{2})z^{-1})}{1 - \sqrt{2}z^{-1} + z^{-2}}$$

Consequently,

$$\begin{aligned}
Y(z) &= H(z)X(z) \\
&= \frac{10(1 - (1/\sqrt{2})z^{-1})}{(1 - 0.5z^{-1})(1 - e^{j\pi/4}z^{-1})(1 - e^{-j\pi/4}z^{-1})} \\
&= \frac{6.3}{1 - 0.5z^{-1}} + \frac{6.78e^{-j28.7^\circ}}{1 - e^{j\pi/4}z^{-1}} + \frac{6.78e^{j28.7^\circ}}{1 - e^{-j\pi/4}z^{-1}}
\end{aligned}$$

The natural or transient response is

$$y_{\text{nr}}(n) = 6.3(0.5)^n u(n)$$

and the forced or steady-state response is

$$\begin{aligned}
y_{\text{fr}}(n) &= [6.78e^{-j28.7}(e^{j\pi n/4}) + 6.78e^{j28.7}e^{-j\pi n/4}]u(n) \\
&= 13.56 \cos\left(\frac{\pi}{4}n - 28.7^\circ\right) u(n)
\end{aligned}$$

Thus we see that the steady-state response persists for all $n \geq 0$, just as the input signal persists for all $n \geq 0$.

3.5.3 Causality and Stability

As defined previously, a causal linear time-invariant system is one whose unit sample response $h(n)$ satisfies the condition

$$h(n) = 0, \qquad n < 0$$

We have also shown that the ROC of the z-transform of a causal sequence is the exterior of a circle. Consequently, *a linear time-invariant system is causal if and only if the ROC of the system function is the exterior of a circle of radius $r < \infty$, including the point $z = \infty$.*

The stability of a linear time-invariant system can also be expressed in terms of the characteristics of the system function. As we recall from our previous discussion, a necessary and sufficient condition for a linear time-invariant system to be BIBO stable is

$$\sum_{n=-\infty}^{\infty} |h(n)| < \infty$$

In turn, this condition implies that $H(z)$ must contain the unit circle within its ROC.

Indeed, since

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

it follows that

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

When evaluated on the unit circle (i.e., $|z| = 1$),

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

Hence, if the system is BIBO stable, the unit circle is contained in the ROC of $H(z)$. The converse is also true. Therefore, a *linear time-invariant system* is BIBO stable if and only if the ROC of the system function includes the unit circle.

We should stress, however, that the conditions for causality and stability are different and that one does not imply the other. For example, a causal system may be stable or unstable, just as a noncausal system may be stable or unstable. Similarly, an unstable system may be either causal or noncausal, just as a stable system may be causal or noncausal.

For a causal system, however, the condition on stability can be narrowed to some extent. Indeed, a causal system is characterized by a system function $H(z)$ having as a ROC the exterior of some circle of radius r. For a stable system, the ROC must include the unit circle. Consequently, a causal and stable system must have a system function that converges for $|z| > r < 1$. Since the ROC cannot contain any poles of $H(z)$, it follows that a *causal linear time-invariant system* is BIBO stable if and only if all the poles of $H(z)$ are inside the unit circle.

EXAMPLE 3.5.2

A linear time-invariant system is characterized by the system function

$$H(z) = \frac{3 - 4z^{-1}}{1 - 3.5z^{-1} + 1.5z^{-2}}$$
$$= \frac{1}{1 - \frac{1}{2}z^{-1}} + \frac{2}{1 - 3z^{-1}}$$

Specify the ROC of $H(z)$ and determine $h(n)$ for the following conditions:

(a) The system is stable.

(b) The system is causal.

(c) The system is anticausal.

Solution. The system has poles at $z = \frac{1}{2}$ and $z = 3$.

(a) Since the system is stable, its ROC must include the unit circle and hence it is $\frac{1}{2} < |z| < 3$. Consequently, $h(n)$ is noncausal and is given as

$$h(n) = (\frac{1}{2})^n u(n) - 2(3)^n u(-n-1)$$

(b) Since the system is causal, its ROC is $|z| > 3$. In this case

$$h(n) = (\frac{1}{2})^n u(n) + 2(3)^n u(n)$$

This system is unstable.

(c) If the system is anticausal, its ROC is $|z| < 0.5$. Hence

$$h(n) = -[(\frac{1}{2})^n + 2(3)^n]u(-n-1)$$

In this case the system is unstable.

3.5.4 Pole–Zero Cancellations

When a z-transform has a pole that is at the same location as a zero, the pole is canceled by the zero and, consequently, the term containing that pole in the inverse z-transform vanishes. Such pole–zero cancellations are very important in the analysis of pole–zero systems.

Pole–zero cancellations can occur either in the system function itself or in the product of the system function with the z-transform of the input signal. In the first case we say that the order of the system is reduced by one. In the latter case we say that the pole of the system is suppressed by the zero in the input signal, or vice versa. Thus, by properly selecting the position of the zeros of the input signal, it is possible to suppress one or more system modes (pole factors) in the response of the system. Similarly, by proper selection of the zeros of the system function, it is possible to suppress one or more modes of the input signal from the response of the system.

When the zero is located very near the pole but not exactly at the same location, the term in the response has a very small amplitude. For example, nonexact pole–zero cancellations can occur in practice as a result of insufficiant numerical precision used in representing the coefficients of the system. Consequently, one should not attempt to stabilize an inherently unstable system by placing a zero in the input signal at the location of the pole.

EXAMPLE 3.5.3

Determine the unit sample response of the system characterized by the difference equation

$$y(n) = 2.5y(n-1) - y(n-2) + x(n) - 5x(n-1) + 6x(n-2)$$

Solution. The system function is

$$H(z) = \frac{1 - 5z^{-1} + 6z^{-2}}{1 - 2.5z^{-1} + z^{-2}}$$

$$= \frac{1 - 5z^{-1} + 6z^{-2}}{(1 - \frac{1}{2}z^{-1})(1 - 2z^{-1})}$$

This system has poles at $p_1 = 2$ and $p_1 = \frac{1}{2}$. Consequently, at first glance it appears that the unit sample response is

$$Y(z) = H(z)X(z) = \frac{1 - 5z^{-1} + 6z^{-2}}{(1 - \frac{1}{2}z^{-1})(1 - 2z^{-1})}$$

$$= z\left(\frac{A}{z - \frac{1}{2}} + \frac{B}{z - 2}\right)$$

By evaluating the constants at $z = \frac{1}{2}$ and $z = 2$, we find that

$$A = \frac{5}{2}, \qquad B = 0$$

The fact that $B = 0$ indicates that there exists a zero at $z = 2$ which cancels the pole at $z = 2$. In fact, the zeros occur at $z = 2$ and $z = 3$. Consequently, $H(z)$ reduces to

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

$$= 1 - \frac{2.5z^{-1}}{1 - \frac{1}{2}z^{-1}}$$

and therefore

$$h(n) = \delta(n) - 2.5\left(\frac{1}{2}\right)^{n-1}u(n-1)$$

The reduced-order system obtained by canceling the common pole and zero is characterized by the difference equation

$$y(n) = \frac{1}{2}y(n-1) + x(n) - 3x(n-1)$$

Although the original system is also BIBO stable due to the pole–zero cancellation, in a practical implementation of this second-order system, we may encounter an instability due to imperfect cancellation of the pole and the zero.

EXAMPLE 3.5.4

Determine the response of the system

$$y(n) = \frac{5}{6}y(n-1) - \frac{1}{6}y(n-2) + x(n)$$

to the input signal $x(n) = \delta(n) - \frac{1}{3}\delta(n-1)$.

Solution. The system function is

$$H(z) = \frac{1}{1 - \frac{5}{6}z^{-1} + \frac{1}{6}z^{-2}}$$

$$= \frac{1}{\left(1 - \frac{1}{2}z^{-1}\right)\left(1 - \frac{1}{3}z^{-1}\right)}$$

This system has two poles, one at $z = \frac{1}{2}$ and the other at $z = \frac{1}{3}$. The z-transform of the input signal is

$$X(z) = 1 - \frac{1}{3}z^{-1}$$

In this case the input signal contains a zero at $z = \frac{1}{3}$ which cancels the pole at $z = \frac{1}{3}$. Consequently,

$$Y(z) = H(z)X(z)$$

$$Y(z) = \frac{1}{1 - \frac{1}{2}z^{-1}}$$

and hence the response of the system is

$$y(n) = (\frac{1}{2})^n u(n)$$

Clearly, the mode $(\frac{1}{3})^n$ is suppressed from the output as a result of the pole–zero cancellation.

3.5.5 Multiple-Order Poles and Stability

As we have observed, a necessary and sufficient condition for a causal linear time-invariant system to be BIBO stable is that all its poles lie inside the unit circle. The input signal is bounded if its z-transform contains poles $\{q_k\}$, $k = 1, 2, \ldots, L$, which satisfy the condition $|q_k| \leq 1$ for all k. We note that the forced response of the system, given in (3.5.6), is also bounded, even when the input signal contains one or more distinct poles on the unit circle.

In view of the fact that a bounded input signal may have poles on the unit circle, it might appear that a stable system may also have poles on the unit circle. This is not the case, however, since such a system produces an unbounded response when excited by an input signal that also has a pole at the same position on the unit circle. The following example illustrates this point.

EXAMPLE 3.5.5

Determine the step response of the causal system described by the difference equation

$$y(n) = y(n-1) + x(n)$$

Solution. The system function for the system is

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

We note that the system contains a pole on the unit circle at $z = 1$. The z-transform of the input signal $x(n) = u(n)$ is

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

which also contains a pole at $z = 1$. Hence the output signal has the transform

$$\begin{aligned} Y(z) &= H(z)X(z) \\ &= \frac{1}{(1 - z^{-1})^2} \end{aligned}$$

which contains a double pole at $z = 1$.

The inverse z-transform of $Y(z)$ is

$$y(n) = (n + 1)u(n)$$

which is a ramp sequence. Thus $y(n)$ is unbounded, even when the input is bounded. Consequently, the system is unstable.

Example 3.5.5 demonstrates clearly that BIBO stability requires that the system poles be strictly inside the unit circle. If the system poles are all inside the unit circle and the excitation sequence $x(n)$ contains one or more poles that coincide with the poles of the system, the output $Y(z)$ will contain multiple-order poles. As indicated previously, such multiple-order poles result in an output sequence that contains terms of the form

$$A_k n^b (p_k)^n u(n)$$

where $0 \le b \le m - 1$ and m is the order of the pole. If $|p_k| < 1$, these terms decay to zero as n approaches infinity because the exponential factor $(p_k)^n$ dominates the term n^b. Consequently, no bounded input signal can produce an unbounded output signal if the system poles are all inside the unit circle.

Finally, we should state that the only useful systems which contain poles on the unit circle are the digital oscillators discussed in Chapter 5. We call such systems *marginally stable*.

3.5.6 Stability of Second-Order Systems

In this section we provide a detailed analysis of a system having two poles. As we shall see in Chapter 9, two-pole systems form the basic building blocks for the realization of higher-order systems.

Let us consider a causal two-pole system described by the second-order difference equation

$$y(n) = -a_1 y(n - 1) - a_2 y(n - 2) + b_0 x(n) \tag{3.5.7}$$

The system function is

$$\begin{aligned} H(z) = \frac{Y(z)}{X(z)} &= \frac{b_0}{1 + a_1 z^{-1} + a_2 z^{-1}} \\ &= \frac{b_0 z^2}{z^2 + a_1 z + a_2} \end{aligned} \tag{3.5.8}$$

This system has two zeros at the origin and poles at

$$p_1, p_2 = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2 - 4a_2}{4}} \tag{3.5.9}$$

The system is BIBO stable if the poles lie inside the unit circle, that is, if $|p_1| < 1$ and $|p_2| < 1$. These conditions can be related to the values of the coefficients a_1 and a_2. In particular, the roots of a quadratic equation satisfy the relations

$$a_1 = -(p_1 + p_2) \tag{3.5.10}$$

$$a_2 = p_1 p_2 \tag{3.5.11}$$

From (3.5.10) and (3.5.11) we easily obtain the conditions that a_1 and a_2 must satisfy for stability. First, a_2 must satisfy the condition

$$|a_2| = |p_1 p_2| = |p_1||p_2| < 1 \tag{3.5.12}$$

The condition for a_1 can be expressed as

$$|a_1| < 1 + a_2 \tag{3.5.13}$$

Therefore, a two-pole system is stable if and only if the coefficients a_1 and a_2 satisfy the conditions in (3.5.12) and (3.5.13).

The stability conditions given in (3.5.12) and (3.5.13) define a region in the coefficient plane (a_1, a_2), which is in the form of a triangle, as shown in Fig. 3.5.1. The system is stable if and only if the point (a_1, a_2) lies inside the triangle, which we call the *stability triangle*.

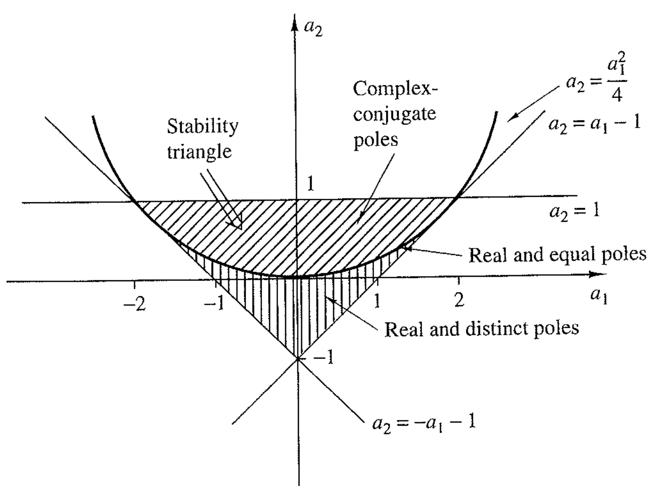

Figure 3.5.1 Region of stability (stability triangle) in the (a_1, a_2) coefficient plane for a second-order system.

The characteristics of the two-pole system depend on the location of the poles or, equivalently, on the location of the point (a_1, a_2) in the stability triangle. The poles of the system may be real or complex conjugate, depending on the value of the discriminant $\Delta = a_1^2 - 4a_2$. The parabola $a_2 = a_1^2/4$ splits the stability triangle into two regions, as illustrated in Fig. 3.5.1. The region below the parabola $(a_1^2 > 4a_2)$ corresponds to real and distinct poles. The points on the parabola $(a_1^2 = 4a_2)$ result in real and equal (double) poles. Finally, the points above the parabola correspond to complex-conjugate poles.

Additional insight into the behavior of the system can be obtained from the unit sample responses for these three cases.

Real and distinct poles $(a_1^2 > 4a_2)$. Since p_1, p_2 are real and $p_1 \neq p_2$, the system function can be expressed in the form

$$H(z) = \frac{A_1}{1 - p_1 z^{-1}} + \frac{A_2}{1 - p_2 z^{-1}} \tag{3.5.14}$$

where

$$A_1 = \frac{b_0 p_1}{p_1 - p_2}, \qquad A_2 = \frac{-b_0 p_2}{p_1 - p_2} \tag{3.5.15}$$

Consequently, the unit sample response is

$$h(n) = \frac{b_0}{p_1 - p_2}(p_1^{n+1} - p_2^{n+1})u(n) \tag{3.5.16}$$

Therefore, the unit sample response is the difference of two decaying exponential sequences. Figure 3.5.2 illustrates a typical graph for $h(n)$ when the poles are distinct.

Real and equal poles $(a_1^2 = 4a_2)$. In this case $p_1 = p_2 = p = -a_1/2$. The system function is

$$H(z) = \frac{b_0}{(1 - pz^{-1})^2} \tag{3.5.17}$$

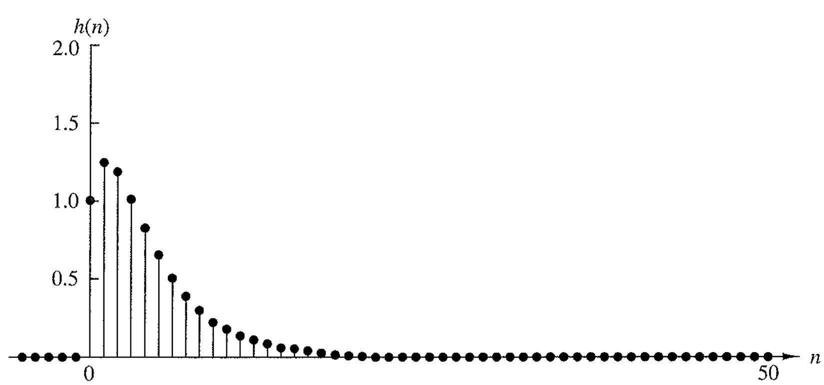

Figure 3.5.2 Plot of $h(n)$ given by (3.5.16) with $p_1 = 0.5$, $p_2 = 0.75$; $h(n) = [1/(p_1 - p_2)](p_1^{n+1} - p_2^{n+1})u(n)$.

and hence the unit sample response of the system is

$$h(n) = b_0(n+1)p^n u(n) \tag{3.5.18}$$

We observe that $h(n)$ is the product of a ramp sequence and a real decaying exponential sequence. The graph of $h(n)$ is shown in Fig. 3.5.3.

Complex-conjugate poles ($a_1^2 < 4a_2$). Since the poles are complex conjugate, the system function can be factored and expressed as

$$\begin{aligned} H(z) &= \frac{A}{1 - pz^{-1}} + \frac{A^*}{1 - p^*z^{-1}} \\ &= \frac{A}{1 - re^{j\omega_0}z^{-1}} + \frac{A^*}{1 - re^{-j\omega_0}z^{-1}} \end{aligned} \tag{3.5.19}$$

where $p = re^{j\omega}$ and $0 < \omega_0 < \pi$. Note that when the poles are complex conjugates, the parameters a_1 and a_2 are related to r and ω_0 according to

$$\begin{aligned} a_1 &= -2r\cos\omega_0 \\ a_2 &= r^2 \end{aligned} \tag{3.5.20}$$

The constant A in the partial-fraction expansion of $H(z)$ is easily shown to be

$$\begin{aligned} A &= \frac{b_0 p}{p - p^*} = \frac{b_0 r e^{j\omega_0}}{r(e^{j\omega_0} - e^{-j\omega_0})} \\ &= \frac{b_0 e^{j\omega_0}}{j2\sin\omega_0} \end{aligned} \tag{3.5.21}$$

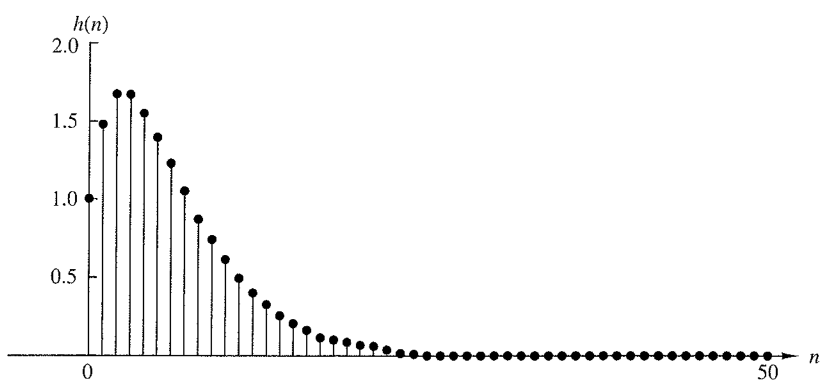

Figure 3.5.3 Plot of $h(n)$ given by (3.5.18) with $p = \frac{3}{4}$; $h(n) = (n+1)p^n u(n)$.

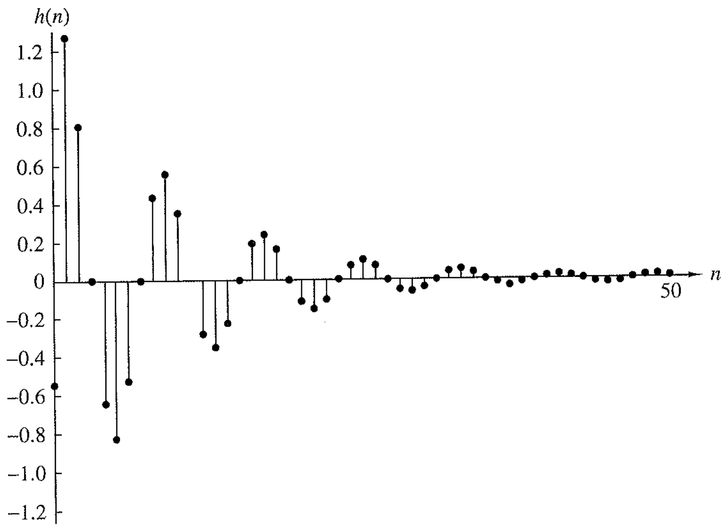

Figure 3.5.4 Plot of $h(n)$ given by (3.5.22) with $b_0 = 1$, $\omega_0 = \pi/4$, $r = 0.9$; $h(n) = [b_0 r^n/(\sin\omega_0)]\sin[(n+1)\omega_0]u(n)$.

Consequently, the unit sample response of a system with complex-conjugate poles is

$$
\begin{aligned}
h(n) &= \frac{b_0 r^n}{\sin\omega_0}\frac{e^{j(n+1)\omega_0} - e^{-j(n+1)\omega_0}}{2j}u(n) \\
&= \frac{b_0 r^n}{\sin\omega_0}\sin(n+1)\omega_0 u(n)
\end{aligned}
\tag{3.5.22}
$$

In this case $h(n)$ has an oscillatory behavior with an exponentially decaying envelope when $r < 1$. The angle ω_0 of the poles determines the frequency of oscillation and the distance r of the poles from the origin determines the rate of decay. When r is close to unity, the decay is slow. When r is close to the origin, the decay is fast. A typical graph of $h(n)$ is illustrated in Fig. 3.5.4.

3.6 The One-sided z-Transform

The two-sided z-transform requires that the corresponding signals be specified for the entire time range $-\infty < n < \infty$. This requirement prevents its use for a very useful family of practical problems, namely the evaluation of the output of nonrelaxed systems. As we recall, these systems are described by difference equations with nonzero initial conditions. Since the input is applied at a finite time, say n_0, both input and output signals are specified for $n \geq n_0$, but by no means are zero for $n < n_0$. Thus the two-sided z-transform cannot be used. In this section we develop the one-sided z-transform which can be used to solve difference equations with initial conditions.