4.1 Introduction

In Chapter 2 we saw how transfer functions can represent linear, time-invariant systems. In Chapter 3 systems were represented directly in the time domain via the state and output equations. After the engineer obtains a mathematical representation of a subsystem, the subsystem is analyzed for its transient and steady-state responses to see if these characteristics yield the desired behavior. This chapter is devoted to the analysis of system transient response.

It may appear more logical to continue with Chapter 5, which covers the modeling of closed-loop systems, rather than to break the modeling sequence with the analysis presented here in Chapter 4. However, the student should not continue too far into system representation without knowing the application for the effort expended. Thus, this chapter demonstrates applications of the system representation by evaluating the transient response from the system model. Logically, this approach is not far from reality, since the engineer may indeed want to evaluate the response of a subsystem prior to inserting it into the closed-loop system.

After describing a valuable analysis and design tool, poles and zeros, we begin analyzing our models to find the step response of first- and second-order systems. The order refers to the order of the equivalent differential equation representing the system—the order of the denominator of the transfer function after cancellation of common factors in the numerator or the number of simultaneous first-order equations required for the state-space representation.

4.2 Poles, Zeros, and System Response

The output response of a system is the sum of two responses: the *forced response* and the *natural response*.[1] Although many techniques, such as solving a differential equation or taking the inverse Laplace transform, enable us to evaluate this output response, these techniques are laborious and time-consuming. Productivity is aided by analysis and design techniques that yield results in a minimum of time. If the technique is so rapid that we feel we derive the desired result by inspection, we sometimes use the attribute *qualitative* to describe the method. The use of poles and zeros and their relationship to the time response of a system is such a technique. Learning this relationship gives us a qualitative "handle" on problems. The concept of poles and zeros, fundamental to the analysis and design of control systems, simplifies the evaluation of a system's response. The reader is encouraged to master the concepts of poles and zeros and their application to problems throughout this book. Let us begin with two definitions.

Poles of a Transfer Function

The *poles* of a transfer function are (1) the values of the Laplace transform variable, s, that cause the transfer function to become infinite or (2) any roots of the denominator of the transfer function that are common to roots of the numerator.

---

[1]The forced response is also called the *steady-state response* or *particular solution*. The natural response is also called the *homogeneous solution*.

Strictly speaking, the poles of a transfer function satisfy part (1) of the definition. For example, the roots of the characteristic polynomial in the denominator are values of s that make the transfer function infinite, so they are thus poles. However, if a factor of the denominator can be canceled by the same factor in the numerator, the root of this factor no longer causes the transfer function to become infinite. In control systems we often refer to the root of the canceled factor in the denominator as a pole even though the transfer function will not be infinite at this value. Hence, we include part (2) of the definition.

Zeros of a Transfer Function

The *zeros* of a transfer function are (1) the values of the Laplace transform variable, s, that cause the transfer function to become zero, or (2) any roots of the numerator of the transfer function that are common to roots of the denominator.

Strictly speaking, the zeros of a transfer function satisfy part (1) of this definition. For example, the roots of the numerator are values of s that make the transfer function zero and are thus zeros. However, if a factor of the numerator can be canceled by the same factor in the denominator, the root of this factor no longer causes the transfer function to become zero. In control systems we often refer to the root of the canceled factor in the numerator as a zero even though the transfer function will not be zero at this value. Hence, we include part (2) of the definition.

Poles and Zeros of a First-Order System: An Example

Given the transfer function $G(s)$ in Figure 4.1(a), a pole exists at $s = -5$, and a zero exists at -2. These values are plotted on the complex s-plane in Figure 4.1(b), using an $\times$ for the pole and a $\bigcirc$ for the zero. To show the properties of the poles and zeros, let us find the unit step response of the system. Multiplying the transfer function of Figure 4.1(a) by a step function yields

$$C(s) = \frac{(s+2)}{s(s+5)} = \frac{A}{s} + \frac{B}{s+5} = \frac{2/5}{s} + \frac{3/5}{s+5} \tag{4.1}$$

where

$$A = \left.\frac{(s+2)}{(s+5)}\right|_{s \to 0} = \frac{2}{5}$$

$$B = \left.\frac{(s+2)}{s}\right|_{s \to -5} = \frac{3}{5}$$

Thus,

$$c(t) = \frac{2}{5} + \frac{3}{5}e^{-5t} \tag{4.2}$$

From the development summarized in Figure 4.1(c), we draw the following conclusions:

1. A pole of the input function generates the form of the *forced response* (that is, the pole at the origin generated a step function at the output).

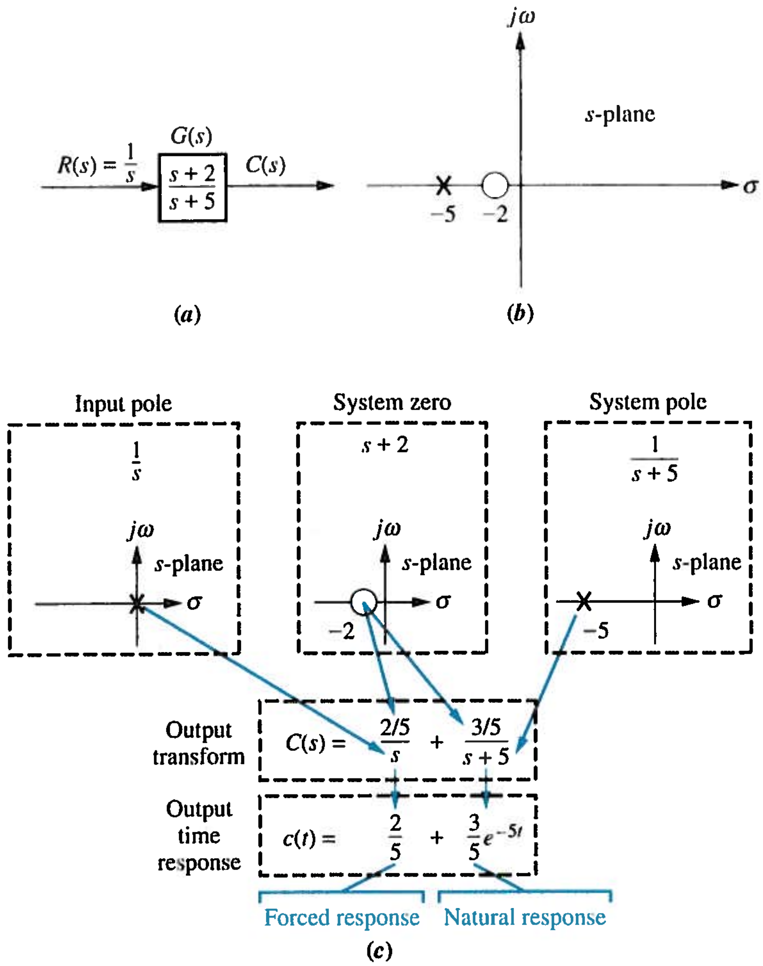


Figure 4.1 a. System showing input and output; **b.** pole-zero plot of the system; **c.** evolution of a system response. Follow blue arrows to see the evolution of the response component generated by the pole or zero.

2. A pole of the transfer function generates the form of the *natural response* (that is, the pole at -5 generated e^{-5t}).
3. A pole on the real axis generates an *exponential* response of the form $e^{-\alpha t}$, where $-\alpha$ is the pole location on the real axis. Thus, the farther to the left a pole is on the negative real axis, the faster the exponential transient response will decay to zero (again, the pole at -5 generated e^{-5t}; see Figure 4.2 for the general case).
4. The zeros and poles generate the *amplitudes* for both the forced and natural responses (this can be seen from the calculation of A and B in Eq. (4.1)).

Let us now look at an example that demonstrates the technique of using poles to obtain the form of the system response. We will learn to write the form of the response by inspection. Each pole of the system transfer function that is on the real axis generates an exponential response that is a component of the natural response. The input pole generates the forced response.

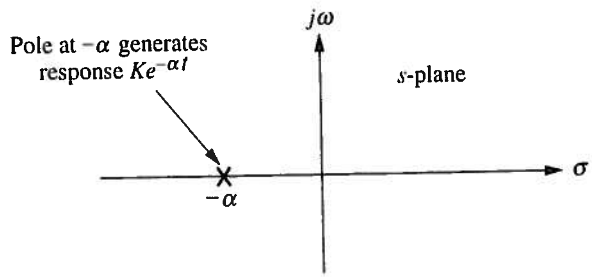


Figure 4.2 Effect of a real-axis pole upon transient response

Example 4.1

Evaluating response using poles

Problem Given the system of Figure 4.3, write the output, $c(t)$, in general terms. Specify the forced and natural parts of the solution.

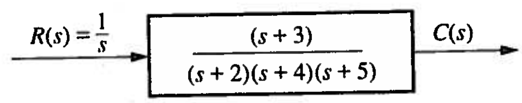


Figure 4.3 System for Example 4.1

Solution By inspection, each system pole generates an exponential as part of the natural response. The input's pole generates the forced response. Thus,

$$C(s) \equiv \underbrace{\frac{K_1}{s}}_{\text{Forced response}} + \underbrace{\frac{K_2}{(s+2)} + \frac{K_3}{(s+4)} + \frac{K_4}{(s+5)}}_{\text{Natural response}} \tag{4.3}$$

Taking the inverse Laplace transform, we get

$$c(t) \equiv \underbrace{K_1}_{\text{Forced response}} + \underbrace{K_2 e^{-2t} + K_3 e^{-4t} + K_4 e^{-5t}}_{\text{Natural response}} \tag{4.4}$$

Skill-Assessment Exercise 4.1

Problem A system has a transfer function, $G(s) = \dfrac{10(s+4)(s+6)}{(s+1)(s+7)(s+8)(s+10)}$.

Write, by inspection, the output, $c(t)$, in general terms if the input is a unit step.

Answer $c(t) \equiv A + Be^{-t} + Ce^{-7t} + De^{-8t} + Ee^{-10t}$

In this section we learned that poles determine the nature of the time response: Poles of the input function determine the form of the forced response, and poles of the transfer function determine the form of the natural response. Zeros and poles of the input or transfer function contribute to the amplitudes of the component parts of the total response. Finally, poles on the real axis generate exponential responses.

4.3 First-Order Systems

We now discuss first-order systems without zeros to define a performance specification for such a system. A first-order system without zeros can be described by the transfer function shown in Figure 4.4(a). If the input is a unit step, where $R(s) = 1/s$, the Laplace transform of the step response is $C(s)$, where

$$C(s) = R(s)G(s) = \frac{a}{s(s+a)} \tag{4.5}$$

Taking the inverse transform, the step response is given by

$$c(t) = c_f(t) + c_n(t) = 1 - e^{-at} \tag{4.6}$$

where the input pole at the origin generated the forced response $c_f(t) = 1$, and the system pole at $-a$, as shown in Figure 4.4(b), generated the natural response $c_n(t) = -e^{-at}$. Equation (4.6) is plotted in Figure 4.5.

Let us examine the significance of parameter a, the only parameter needed to describe the transient response. When $t = 1/a$,

$$e^{-at}\big|_{t=1/a} = e^{-1} = 0.37 \tag{4.7}$$

or

$$c(t)\big|_{t=1/a} = 1 - e^{-at}\big|_{t=1/a} = 1 - 0.37 = 0.63 \tag{4.8}$$

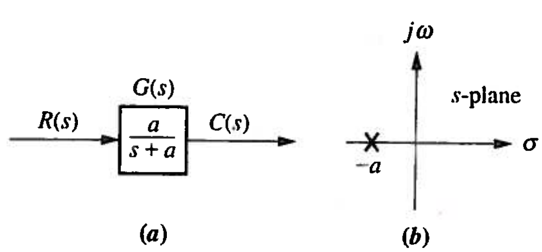
We now use Eqs. (4.6), (4.7), and (4.8) to define three transient response performance specifications.

Time Constant

We call $1/a$ the *time constant* of the response. From Eq. (4.7) the time constant can be described as the time for e^{-at} to decay to 37% of its initial value. Alternately, from Eq. (4.8) the time constant is the time it takes for the step response to rise to 63% of its final value (see Figure 4.5).

The reciprocal of the time constant has the units (1/seconds), or frequency. Thus, we can call the parameter a the *exponential frequency*. Since the derivative of e^{-at} is $-a$ when $t = 0$, a is the initial rate of change of the exponential at $t = 0$. Thus, the time constant can be considered a transient response specification for a

Figure 4.4
a. First-order system;
b. pole plot



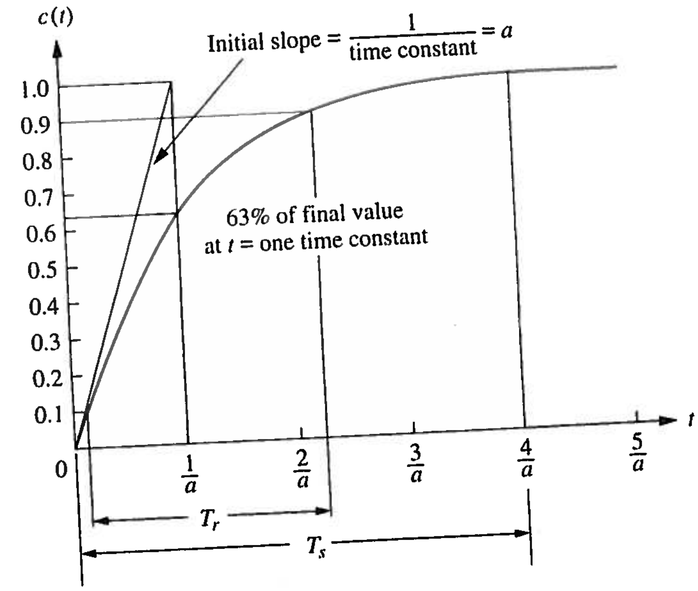


Figure 4.5 First-order system response to a unit step

first-order system, since it is related to the speed at which the system responds to a step input.

The time constant can also be evaluated from the pole plot (see Figure 4.4(*b*)). Since the pole of the transfer function is at $-a$, we can say the pole is located at the *reciprocal* of the time constant, and the farther the pole from the imaginary axis, the faster the transient response.

Let us look at other transient response specifications, such as rise time, T_r, and settling time, T_s, as shown in Figure 4.5.

Rise Time, T_r

Rise time is defined as the time for the waveform to go from 0.1 to 0.9 of its final value. Rise time is found by solving Eq. (4.6) for the difference in time at $c(t) = 0.9$ and $c(t) = 0.1$. Hence,

$$T_r = \frac{2.31}{a} - \frac{0.11}{a} = \frac{2.2}{a} \tag{4.9}$$

Settling Time, T_s

Settling time is defined as the time for the response to reach, and stay within, 2% of its final value.[2] Letting $c(t) = 0.98$ in Eq. (4.6) and solving for time, t, we find the settling time to be

$$T_s = \frac{4}{a} \tag{4.10}$$

---

[2]Strictly speaking, this is the definition of the *2% settling time*. Other percentages, for example 5%, also can be used. We will use *settling time* throughout the book to mean 2% settling time.

First-Order Transfer Functions via Testing

Often it is not possible or practical to obtain a system's transfer function analytically. Perhaps the system is closed, and the component parts are not easily identifiable. Since the transfer function is a representation of the system from input to output, the system's step response can lead to a representation even though the inner construction is not known. With a step input, we can measure the time constant and the steady-state value, from which the transfer function can be calculated.

Consider a simple first-order system, $G(s) = K/(s+a)$, whose step response is

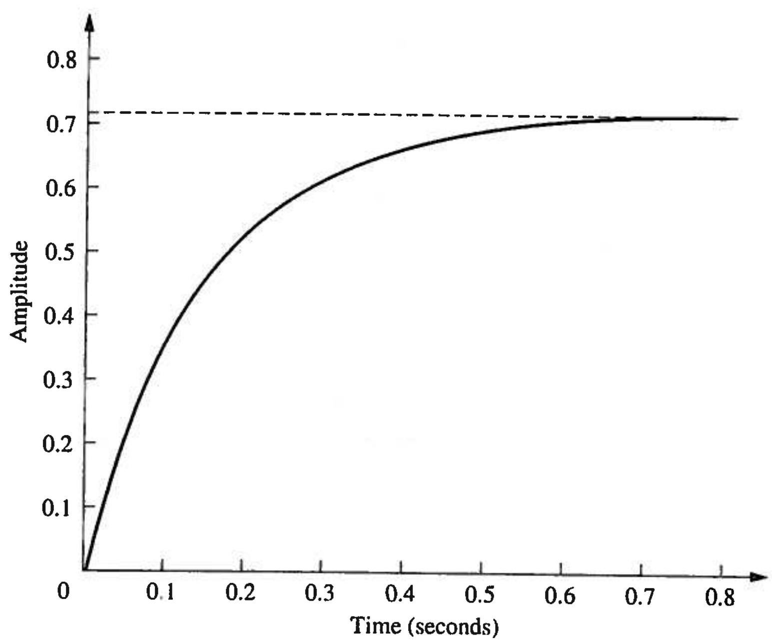
$$C(s) = \frac{K}{s(s+a)} = \frac{K/a}{s} - \frac{K/a}{(s+a)} \tag{4.11}$$

If we can identify K and a from laboratory testing, we can obtain the transfer function of the system.

For example, assume the unit step response given in Figure 4.6. We determine that it has the first-order characteristics we have seen thus far, such as no overshoot and nonzero initial slope. From the response, we measure the time constant, that is, the time for the amplitude to reach 63% of its final value. Since the final value is about 0.72, the time constant is evaluated where the curve reaches $0.63 \times 0.72 = 0.45$, or about 0.13 second. Hence, $a = 1/0.13 = 7.7$.

To find K, we realize from Eq. (4.11) that the forced response reaches a steady-state value of $K/a = 0.72$. Substituting the value of a, we find $K = 5.54$. Thus, the transfer function for the system is $G(s) = 5.54/(s+7.7)$. It is interesting to note

Figure 4.6
Laboratory results of a system step response test



that the response of Figure 4.6 was generated using the transfer function $G(s) = 5/(s+7)$.

Skill-Assessment Exercise 4.2

Problem A system has a transfer function, $G(s) = \dfrac{50}{s+50}$. Find the time constant, T_c, settling time, T_s, and rise time, T_r.

Answers $T_c = 0.02$ s, $T_s = 0.08$ s, and $T_r = 0.044$ s.

The complete solution is on the accompanying CD-ROM.

4.4 Second-Order Systems: Introduction

Let us now extend the concepts of poles and zeros and transient response to second-order systems. Compared to the simplicity of a first-order system, a second-order system exhibits a wide range of responses that must be analyzed and described. Whereas varying a first-order system's parameter simply changes the speed of the response, changes in the parameters of a second-order system can change the *form* of the response. For example, a second-order system can display characteristics much like a first-order system or, depending on component values, display damped or pure oscillations for its transient response.

To become familiar with the wide range of responses before formalizing our discussion in the next section, we take a look at numerical examples of the second-order system responses shown in Figure 4.7. All examples are derived from Figure 4.7(*a*), the general case, which has two finite poles and no zeros. The term in the numerator is simply a scale or input multiplying factor that can take on any value without affecting the form of the derived results. By assigning appropriate values to parameters a and b, we can show all possible second-order transient responses. The unit step response then can be found using $C(s) = R(s)G(s)$, where $R(s) = 1/s$, followed by a partial-fraction expansion and the inverse Laplace transform. Details are left as an end-of-chapter problem, for which you may want to review Section 2.2.

We now explain each response and show how we can use the poles to determine the nature of the response without going through the procedure of a partial-fraction expansion followed by the inverse Laplace transform.

Overdamped Response, Figure 4.7(*b*)

For this response,

$$C(s) = \frac{9}{s(s^2+9s+9)} = \frac{9}{s(s+7.854)(s+1.146)} \tag{4.12}$$

This function has a pole at the origin that comes from the unit step input and two real poles that come from the system. The input pole at the origin generates the constant forced response; each of the two system poles on the real axis generates an exponential natural response whose exponential frequency is equal to the pole location. Hence, the output initially could have been written as $c(t) = K_1 + K_2e^{-7.854t} + K_3e^{-1.146t}$. This response, shown in Figure 4.7(*b*), is

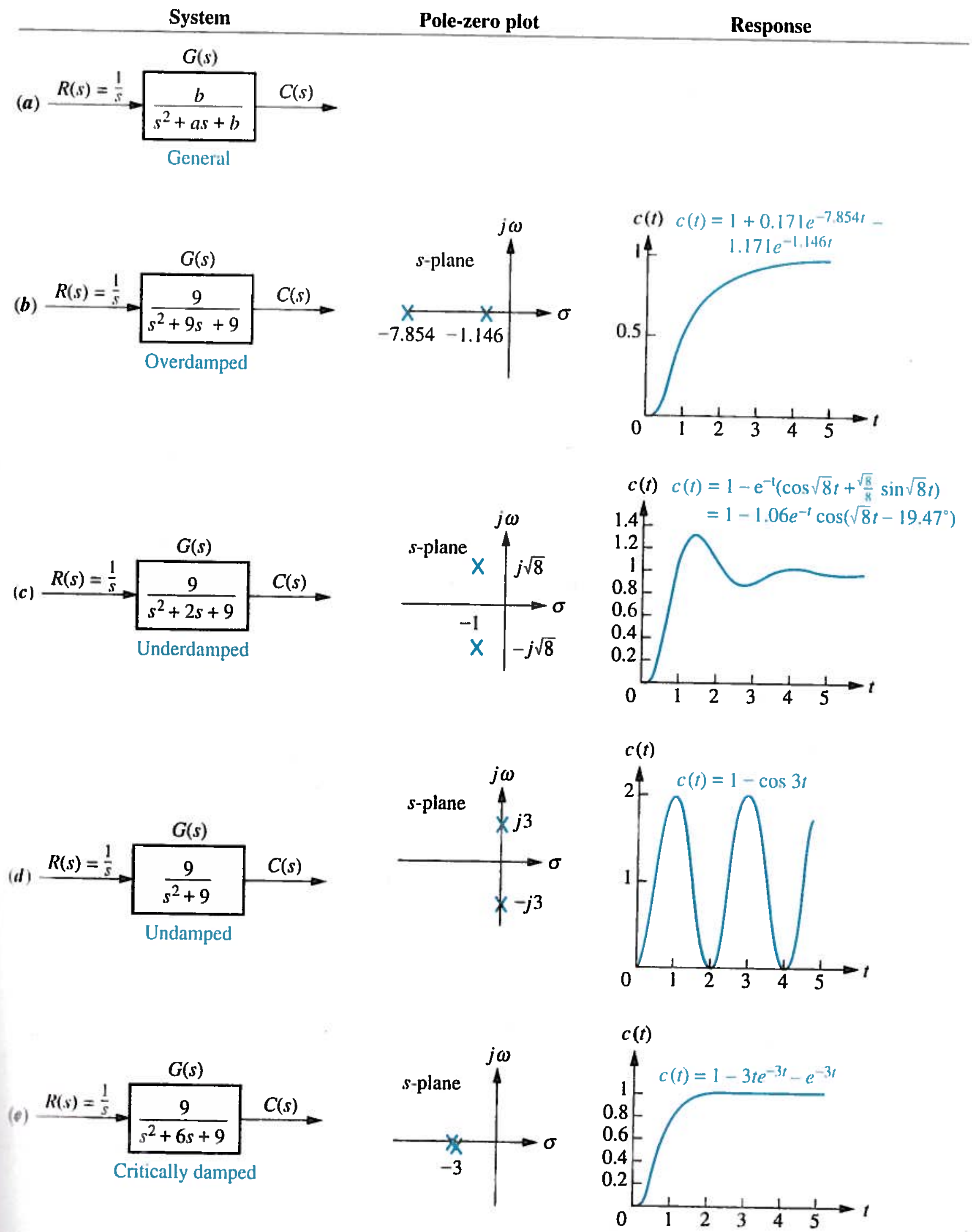


Figure 4.7

Second-order systems, pole plots, and step responses called *overdamped.*[3] We see that the poles tell us the form of the response without the tedious calculation of the inverse Laplace transform.

Underdamped Response, Figure 4.7(c)

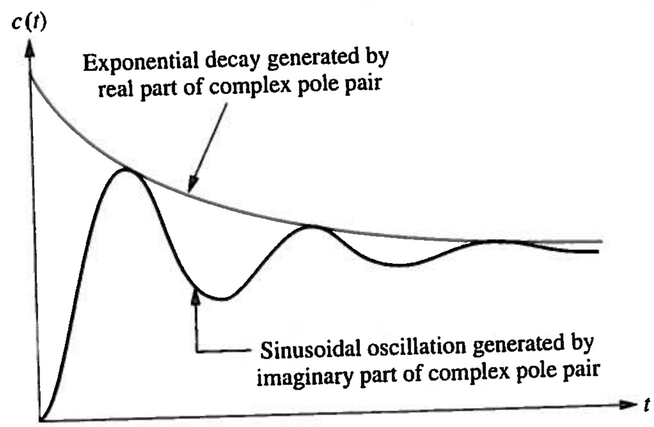
For this response,

$$C(s) = \frac{9}{s(s^2 + 2s + 9)} \tag{4.13}$$

This function has a pole at the origin that comes from the unit step input and two complex poles that come from the system. We now compare the response of the second-order system to the poles that generated it. First we will compare the pole location to the time function, and then we will compare the pole location to the plot. From Figure 4.7(*c*), the poles that generate the natural response are at $s = -1 \pm j\sqrt{8}$. Comparing these values to $c(t)$ in the same figure, we see that the real part of the pole matches the exponential decay frequency of the sinusoid's amplitude, while the imaginary part of the pole matches the frequency of the sinusoidal oscillation.

Let us now compare the pole location to the plot. Figure 4.8 shows a general, damped sinusoidal response for a second-order system. The transient response consists of an exponentially decaying amplitude generated by the real part of the system pole times a sinusoidal waveform generated by the imaginary part of the system pole. The time constant of the exponential decay is equal to the reciprocal of the real part of the system pole. The value of the imaginary part is the actual frequency of the sinusoid, as depicted in Figure 4.8. This sinusoidal frequency is given the name *damped frequency of oscillation*, ω_d. Finally, the steady-state response (unit step) was generated by the input pole located at the origin. We call the type of response shown in Figure 4.8 an *underdamped response,* one which approaches a steady-state value via a transient response that is a damped oscillation.

Figure 4.8 Second-order step response components generated by complex poles



[3]So named because *overdamped* refers to a large amount of energy absorption in the system, which inhibits the transient response from overshooting and oscillating about the steady-state value for a step input. As the energy absorption is reduced, an overdamped system will become underdamped and exhibit overshoot.

The following example demonstrates how a knowledge of the relationship between the pole location and the transient response can lead rapidly to the response form without calculating the inverse Laplace transform.

Example 4.2

Form of underdamped response using poles

Problem By inspection, write the form of the step response of the system in Figure 4.9.

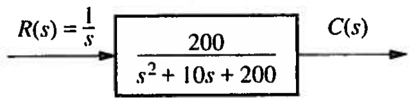


Figure 4.9 System for Example 4.2

Solution First we determine that the form of the forced response is a step. Next we find the form of the natural response. Factoring the denominator of the transfer function in Figure 4.9, we find the poles to be $s = -5 \pm j13.23$. The real part, -5, is the exponential frequency for the damping. It is also the reciprocal of the time constant of the decay of the oscillations. The imaginary part, 13.23, is the radian frequency for the sinusoidal oscillations. Using our previous discussion and Figure 4.7(*c*) as a guide, we obtain $c(t) = K_1 + e^{-5t}(K_2 \cos 13.23t + K_3 \sin 13.23t) = K_1 + K_4 e^{-5t}(\cos 13.23t - \phi)$, where $\phi = \tan^{-1} K_3/K_2$, $K_4 = \sqrt{K_2^2 + K_3^2}$, and $c(t)$ is a constant plus an exponentially damped sinusoid.

We will revisit the second-order underdamped response in Sections 4.5 and 4.6, where we generalize the discussion and derive some results that relate the pole position to other parameters of the response.

Undamped Response, Figure 4.7(*d*)

For this response,

$$C(s) = \frac{9}{s(s^2 + 9)} \tag{4.14}$$

This function has a pole at the origin that comes from the unit step input and two imaginary poles that come from the system. The input pole at the origin generates the constant forced response, and the two system poles on the imaginary axis at $\pm j3$ generate a sinusoidal natural response whose frequency is equal to the location of the imaginary poles. Hence, the output can be estimated as $c(t) = K_1 + K_4 \cos(3t - \phi)$. This type of response, shown in Figure 4.7(*d*), is called *undamped*. Note that the absence of a real part in the pole pair corresponds to an exponential that does not decay. Mathematically, the exponential is $e^{-0t} = 1$.

Critically Damped Response, Figure 4.7(*e*)

For this response,

$$C(s) = \frac{9}{s(s^2 + 6s + 9)} = \frac{9}{s(s + 3)^2} \tag{4.15}$$

This function has a pole at the origin that comes from the unit step input and two multiple real poles that come from the system. The input pole at the origin generates the constant forced response, and the two poles on the real axis at -3 generate a natural response consisting of an exponential and an exponential multiplied by time, where the exponential frequency is equal to the location of the real poles. Hence, the output can be estimated as $c(t) = K_1 + K_2e^{-3t} + K_3te^{-3t}$. This type of response, shown in Figure 4.7(*e*), is called *critically damped*. Critically damped responses are the fastest possible without the overshoot that is characteristic of the underdamped response.

We now summarize our observations. In this section we defined the following natural responses and found their characteristics:

1. *Overdamped responses*

   Poles: Two real at $-\sigma_1$, $-\sigma_2$

   Natural response: Two exponentials with time constants equal to the reciprocal of the pole locations, or

$$c(t) = K_1e^{-\sigma_1 t} + K_2e^{-\sigma_2 t}$$

2. *Underdamped responses*

   Poles: Two complex at $-\sigma_d \pm j\omega_d$

   Natural response: Damped sinusoid with an exponential envelope whose time constant is equal to the reciprocal of the pole's real part. The radian frequency of the sinusoid, the damped frequency of oscillation, is equal to the imaginary part of the poles, or

$$c(t) = Ae^{-\sigma_d t}\cos(\omega_d t - \phi)$$

3. *Undamped responses*

   Poles: Two imaginary at $\pm j\omega_1$

   Natural response: Undamped sinusoid with radian frequency equal to the imaginary part of the poles, or

$$c(t) = A\cos(\omega_1 t - \phi)$$

4. *Critically damped responses*

   Poles: Two real at $-\sigma_1$

   Natural response: One term is an exponential whose time constant is equal to the reciprocal of the pole location. Another term is the product of time, t, and an exponential with time constant equal to the reciprocal of the pole location, or

$$c(t) = K_1e^{-\sigma_1 t} + K_2te^{-\sigma_1 t}$$

The step responses for the four cases of damping discussed in this section are superimposed in Figure 4.10. Notice that the critically damped case is the division between the overdamped cases and the underdamped cases and is the fastest response without overshoot.

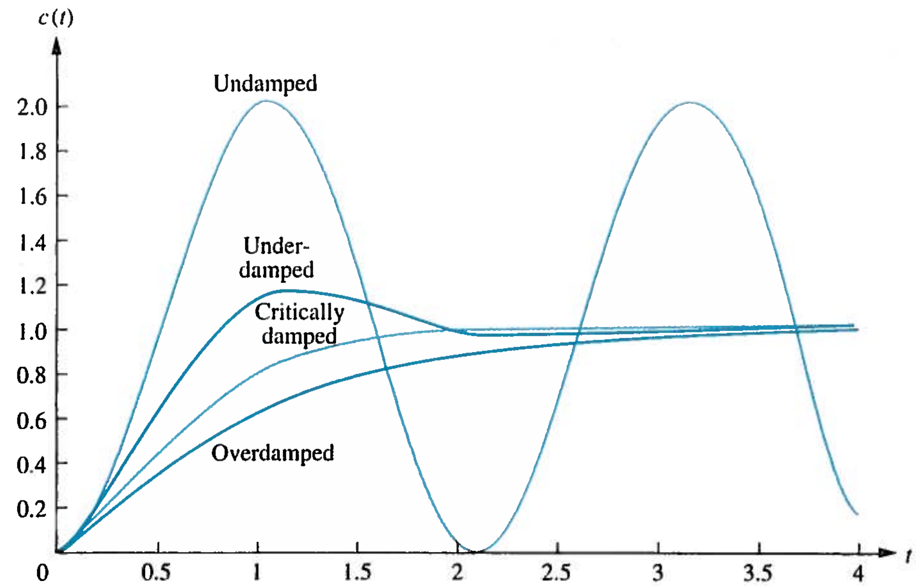


Figure 4.10 Step responses for second-order system damping cases

Skill-Assessment Exercise 4.3

Control Solution

Problem For each of the following transfer functions, write, by inspection, the general form of the step response:

a. $G(s) = \dfrac{400}{s^2 + 12s + 400}$

b. $G(s) = \dfrac{900}{s^2 + 90s + 900}$

c. $G(s) = \dfrac{225}{s^2 + 30s + 225}$

d. $G(s) = \dfrac{625}{s^2 + 625}$

Answers

a. $c(t) = A + Be^{-6t}\cos(19.08t + \phi)$

b. $c(t) = A + Be^{-78.54t} + Ce^{-11.46t}$

c. $c(t) = A + Be^{-15t} + Cte^{-15t}$

d. $c(t) = A + B\cos(25t + \phi)$

The complete solution is on the accompanying CD-ROM.

In the next section we will formalize and generalize our discussion of second-order responses and define two specifications used for the analysis and design of second-order systems. In Section 4.6 we will focus on the *underdamped* case and derive some specifications unique to this response that we will use later for analysis and design.

4.5 The General Second-Order System

Now that we have become familiar with second-order systems and their responses, we generalize the discussion and establish quantitative specifications defined in such a way that the response of a second-order system can be described to a designer without the need for sketching the response. In this section we define two physically meaningful specifications for second-order systems. These quantities can be used to describe the characteristics of the second-order transient response just as time constants describe the first-order system response. The two quantities are called natural frequency and damping ratio. Let us formally define them.

Natural Frequency, ω_n

The *natural frequency* of a second-order system is the frequency of oscillation of the system without damping. For example, the frequency of oscillation of a series *RLC* circuit with the resistance shorted would be the natural frequency.

Damping Ratio, ζ

Before we state our next definition, some explanation is in order. We have already seen that a second-order system's underdamped step response is characterized by damped oscillations. Our definition is derived from the need to quantitatively describe this damped oscillation regardless of the time scale. Thus, a system whose transient response goes through three cycles in a millisecond before reaching the steady state would have the same measure as a system that went through three cycles in a millennium before reaching the steady state. For example, the underdamped curve in Figure 4.10 has an associated measure that defines its shape. This measure remains the same even if we change the time base from seconds to microseconds or to millennia.

A viable definition for this quantity is one that compares the exponential decay frequency of the envelope to the natural frequency. This ratio is constant regardless of the time scale of the response. Also, the reciprocal, which is proportional to the ratio of the natural period to the exponential time constant, remains the same regardless of the time base.

We define the *damping ratio*, ζ, to be

$$\zeta = \frac{\text{Exponential decay frequency}}{\text{Natural frequency (rad/second)}} = \frac{1}{2\pi}\frac{\text{Natural period (seconds)}}{\text{Exponential time constant}}$$

Let us now revise our description of the second-order system to reflect the new definitions. The general second-order system shown in Figure 4.7(a) can be transformed to show the quantities ζ and ω_n. Consider the general system

$$G(s) = \frac{b}{s^2 + as + b} \tag{4.16}$$

Without damping, the poles would be on the $j\omega$ axis, and the response would be an undamped sinusoid. For the poles to be purely imaginary, $a = 0$. Hence,

$$G(s) = \frac{b}{s^2 + b} \tag{4.17}$$

By definition, the natural frequency, ω_n, is the frequency of oscillation of this system. Since the poles of this system are on the $j\omega$ axis at $\pm j\sqrt{b}$,

$$\omega_n = \sqrt{b} \tag{4.18}$$

Hence,

$$b = \omega_n^2 \tag{4.19}$$

Now what is the term a in Eq. (4.16)? Assuming an underdamped system, the complex poles have a real part, σ, equal to $-a/2$. The magnitude of this value is then the exponential decay frequency described in Section 4.4. Hence,

$$\zeta = \frac{\text{Exponential decay frequency}}{\text{Natural frequency (rad/second)}} = \frac{|\sigma|}{\omega_n} = \frac{a/2}{\omega_n} \tag{4.20}$$

from which

$$a = 2\zeta\omega_n \tag{4.21}$$

Our general second-order transfer function finally looks like this:

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{4.22}$$

In the following example we find numerical values for ζ and ω_n by matching the transfer function to Eq. (4.22).

Example 4.3

Finding ζ and ω_n for a second-order system

Problem Given the transfer function of Eq. (4.23), find ζ and ω_n.

$$G(s) = \frac{36}{s^2 + 4.2s + 36} \tag{4.23}$$

Solution Comparing Eq. (4.23) to (4.22), $\omega_n^2 = 36$, from which $\omega_n = 6$. Also, $2\zeta\omega_n = 4.2$. Substituting the value of ω_n, $\zeta = 0.35$.

Now that we have defined ζ and ω_n, let us relate these quantities to the pole location. Solving for the poles of the transfer function in Eq. (4.22) yields

$$s_{1,2} = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2 - 1} \tag{4.24}$$

From Eq. (4.24) we see that the various cases of second-order response are a function of ζ; they are summarized in Figure 4.11.[4]

In the following example we find the numerical value of ζ and determine the nature of the transient response.

[4]The student should verify Figure 4.11 as an exercise.

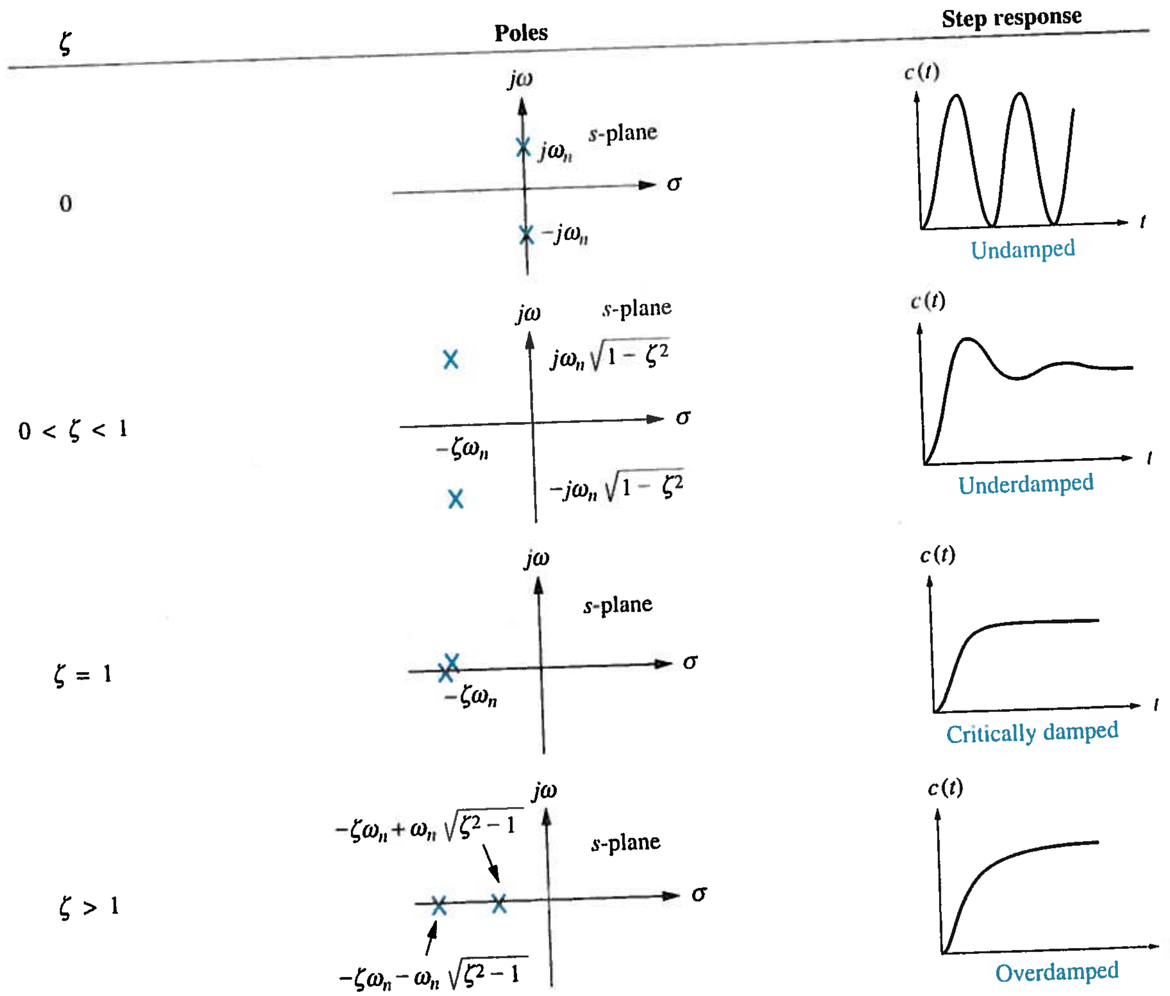


Figure 4.11 Second-order response as a function of damping ratio

Example 4.4

Characterizing response from the value of ζ

Problem For each of the systems shown in Figure 4.12, find the value of ζ and report the kind of response expected.

Solution First match the form of these systems to the forms shown in Eqs. (4.16) and (4.22). Since $a = 2\zeta\omega_n$ and $\omega_n = \sqrt{b}$,

$$\zeta = \frac{a}{2\sqrt{b}} \tag{4.25}$$

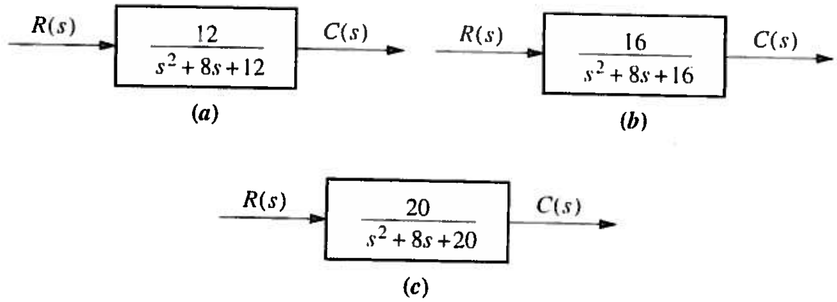


Figure 4.12 Systems for Example 4.4

Using the values of a and b from each of the systems of Figure 4.12, we find $\zeta = 1.155$ for system (a), which is thus overdamped, since $\zeta > 1$; $\zeta = 1$ for system (b), which is thus critically damped; and $\zeta = 0.894$ for system (c), which is thus underdamped, since $\zeta < 1$.

Skill-Assessment Exercise 4.4

Problem For each of the transfer functions in Skill-Assesment Exercise 4.3, do the following: (1) Find the values of ζ and ω_n; (2) characterize the nature of the response.

Answers

a. $\zeta = 0.3$, $\omega_n = 20$; system is underdamped

b. $\zeta = 1.5$, $\omega_n = 30$; system is overdamped

c. $\zeta = 1$, $\omega_n = 15$; system is critically damped

d. $\zeta = 0$, $\omega_n = 25$; system is undamped

The complete solution is on the accompanying CD-ROM.

This section defined two specifications, or parameters, of second-order systems: natural frequency, ω_n, and damping ratio, ζ. We saw that the nature of the response obtained was related to the value of ζ. Variations of damping ratio alone yield the complete range of overdamped, critically damped, underdamped, and undamped responses.

4.6 Underdamped Second-Order Systems

Now that we have generalized the second-order transfer function in terms of ζ and ω_n, let us analyze the step response of an *underdamped* second-order system. Not only will this response be found in terms of ζ and ω_n, but more specifications indigenous to the underdamped case will be defined. The underdamped second-order system, a common model for physical problems, displays unique behavior that must be itemized; a detailed description of the underdamped response is necessary for both analysis and design. Our first objective is to define transient specifications associated with underdamped responses. Next we relate these specifications to the pole location, drawing an association between pole location and the form of the underdamped second-order response. Finally, we tie the pole location to system parameters, thus closing the loop: Desired response generates required system components.

Let us begin by finding the step response for the general second-order system of Eq. (4.22). The transform of the response, $C(s)$, is the transform of the input times the transfer function, or

$$C(s) = \frac{\omega_n^2}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)} = \frac{K_1}{s} + \frac{K_2 s + K_3}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (4.26)$$

where it is assumed that $\zeta < 1$ (the underdamped case). Expanding by partial fractions, using the methods described in Section 2.2, Case 3, yields

$$C(s) = \frac{1}{s} - \frac{(s + \zeta\omega_n) + \dfrac{\zeta}{\sqrt{1-\zeta^2}}\omega_n\sqrt{1-\zeta^2}}{(s + \zeta\omega_n)^2 + \omega_n^2(1-\zeta^2)} \quad (4.27)$$

Taking the inverse Laplace transform, which is left as an exercise for the student, produces

$$\begin{aligned} c(t) &= 1 - e^{-\zeta\omega_n t}\left(\cos\omega_n\sqrt{1-\zeta^2}\,t + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_n\sqrt{1-\zeta^2}\,t\right) \\ &= 1 - \frac{1}{\sqrt{1-\zeta^2}}e^{-\zeta\omega_n t}\cos\left(\omega_n\sqrt{1-\zeta^2}\,t - \phi\right) \end{aligned} \quad (4.28)$$

where $\phi = \tan^{-1}(\zeta/\sqrt{1-\zeta^2})$.

A plot of this response appears in Figure 4.13 for various values of ζ, plotted along a time axis normalized to the natural frequency. We now see the relationship between the value of ζ and the type of response obtained: The lower the value of ζ, the more oscillatory the response. The natural frequency is a time-axis scale factor and does not affect the nature of the response other than to scale it in time.

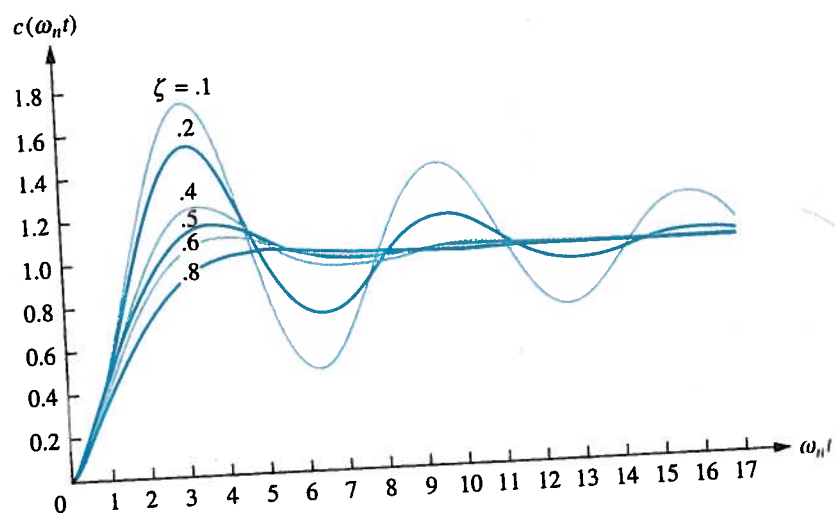


Figure 4.13 Second-order underdamped responses for damping ratio values

CUT THIS OUT AND INCLUDE IN IPYTHON

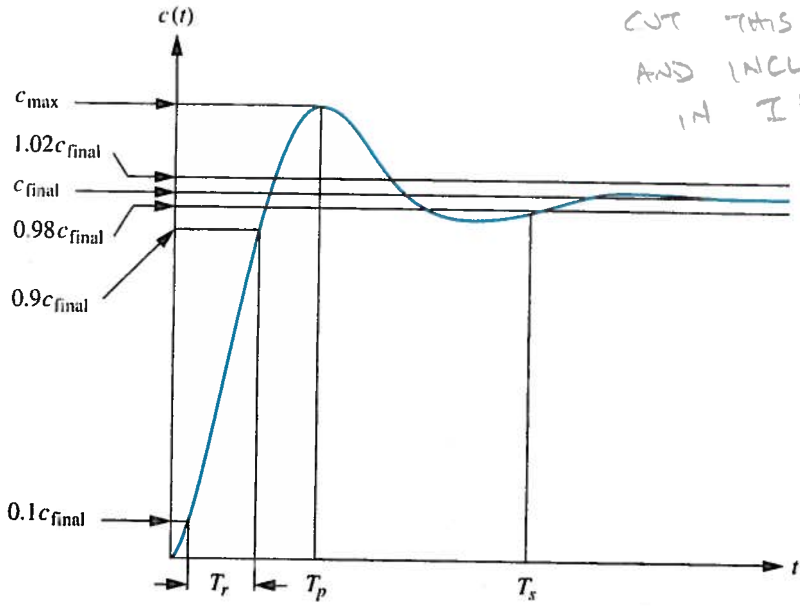


Figure 4.14 Second-order underdamped response specifications

We have defined two parameters associated with second-order systems, ζ and ω_n. Other parameters associated with the underdamped response are rise time, peak time, percent overshoot, and settling time. These specifications are defined as follows (see also Figure 4.14):

1. *Rise time,* T_r. The time required for the waveform to go from 0.1 of the final value to 0.9 of the final value.
2. *Peak time,* T_p. The time required to reach the first, or maximum, peak.
3. *Percent overshoot, %OS.* The amount that the waveform overshoots the steady-state, or final, value at the peak time, expressed as a percentage of the steady-state value.
4. *Settling time,* T_s. The time required for the transient's damped oscillations to reach and stay within $\pm 2\%$ of the steady-state value.

Notice that the definitions for settling time and rise time are basically the same as the definitions for the first-order response. All definitions are also valid for systems of order higher than 2, although analytical expressions for these parameters cannot be found unless the response of the higher-order system can be approximated as a second-order system, which we do in Sections 4.7 and 4.8.

Rise time, peak time, and settling time yield information about the speed of the transient response. This information can help a designer determine if the speed and the nature of the response do or do not degrade the performance of the system. For example, the speed of an entire computer system depends on the time it takes for a floppy disk drive head to reach steady state and read data; passenger comfort depends in part on the suspension system of a car and the number of oscillations it goes through after hitting a bump.

We now evaluate T_p, $\%OS$, and T_s as functions of ζ and ω_n. Later in this chapter we relate these specifications to the location of the system poles. A precise analytical expression for rise time cannot be obtained; thus, we present a plot and a table showing the relationship between ζ and rise time.

Evaluation of T_p

T_p is found by differentiating $c(t)$ in Eq. (4.28) and finding the first zero crossing after $t = 0$. This task is simplified by "differentiating" in the frequency domain by using Item 7 of Table 2.2. Assuming zero initial conditions and using Eq. (4.26), we get

$$\mathscr{L}[\dot{c}(t)] = sC(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{4.29}$$

Completing squares in the denominator, we have

$$\mathscr{L}[\dot{c}(t)] = \frac{\omega_n^2}{(s + \zeta\omega_n)^2 + \omega_n^2(1 - \zeta^2)} = \frac{\dfrac{\omega_n}{\sqrt{1-\zeta^2}}\omega_n\sqrt{1-\zeta^2}}{(s + \zeta\omega_n)^2 + \omega_n^2(1 - \zeta^2)} \tag{4.30}$$

Therefore,

$$\dot{c}(t) = \frac{\omega_n}{\sqrt{1-\zeta^2}}e^{-\zeta\omega_n t}\sin\omega_n\sqrt{1-\zeta^2}\,t \tag{4.31}$$

Setting the derivative equal to zero yields

$$\omega_n\sqrt{1-\zeta^2}\,t = n\pi \tag{4.32}$$

or

$$t = \frac{n\pi}{\omega_n\sqrt{1-\zeta^2}} \tag{4.33}$$

Each value of n yields the time for local maxima or minima. Letting $n = 0$ yields $t = 0$, the first point on the curve in Figure 4.14 that has zero slope. The first peak, which occurs at the peak time, T_p, is found by letting $n = 1$ in Eq. (4.33):

$$T_p = \frac{\pi}{\omega_n\sqrt{1-\zeta^2}} \tag{4.34}$$

Evaluation of $\%OS$

From Figure 4.14 the percent overshoot, $\%OS$, is given by

$$\%OS = \frac{c_{\max} - c_{\text{final}}}{c_{\text{final}}} \times 100 \tag{4.35}$$

The term $c_{\max}$ is found by evaluating $c(t)$ at the peak time, $c(T_p)$. Using Eq. (4.34) for T_p and substituting into Eq. (4.28) yields

$$\begin{aligned} c_{\max} = c(T_p) &= 1 - e^{-(\zeta\pi/\sqrt{1-\zeta^2})}\left(\cos\pi + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\pi\right) \\ &= 1 + e^{-(\zeta\pi/\sqrt{1-\zeta^2})} \end{aligned} \tag{4.36}$$

For the unit step used for Eq. (4.28),

$$c_{\text{final}} = 1 \tag{4.37}$$

Substituting Eqs. (4.36) and (4.37) into Eq. (4.35), we finally obtain

$$\%OS = e^{-(\zeta\pi/\sqrt{1-\zeta^2})} \times 100 \tag{4.38}$$

Notice that the percent overshoot is a function only of the damping ratio, ζ.

Whereas Eq. (4.38) allows one to find $\%OS$ given ζ, the inverse of the equation allows one to solve for ζ given $\%OS$. The inverse is given by

$$\zeta = \frac{-\ln(\%OS/100)}{\sqrt{\pi^2 + \ln^2(\%OS/100)}} \tag{4.39}$$

The derivation of Eq. (4.39) is left as an exercise for the student. Equation (4.38) (or, equivalently, (4.39)) is plotted in Figure 4.15.

Evaluation of T_s

In order to find the settling time, we must find the time for which $c(t)$ in Eq. (4.28) reaches and stays within $\pm 2\%$ of the steady-state value, c_{final}. Using our definition, the settling time is the time it takes for the amplitude of the decaying sinusoid in Eq. (4.28) to reach 0.02, or

$$e^{-\zeta\omega_n t}\frac{1}{\sqrt{1-\zeta^2}} = 0.02 \tag{4.40}$$

This equation is a conservative estimate, since we are assuming that $\cos(\omega_n\sqrt{1-\zeta^2}t - \phi) = 1$ at the settling time. Solving Eq. (4.40) for t, the settling time is

$$T_s = \frac{-\ln\left(0.02\sqrt{1-\zeta^2}\right)}{\zeta\omega_n} \tag{4.41}$$

Figure 4.15 Percent overshoot versus damping ratio

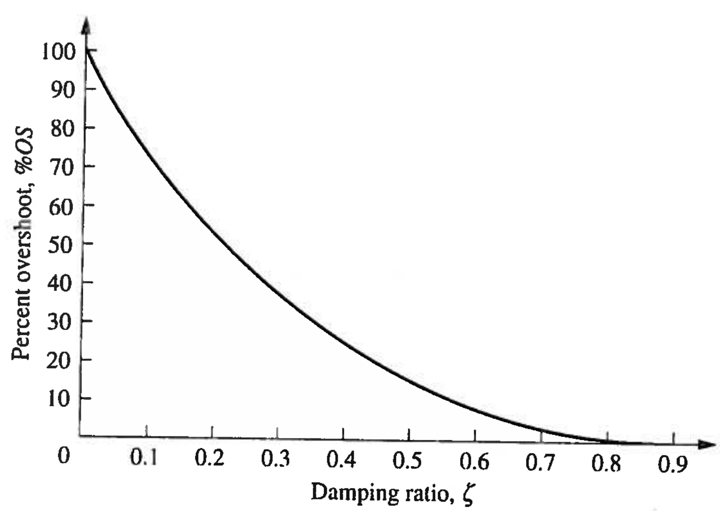
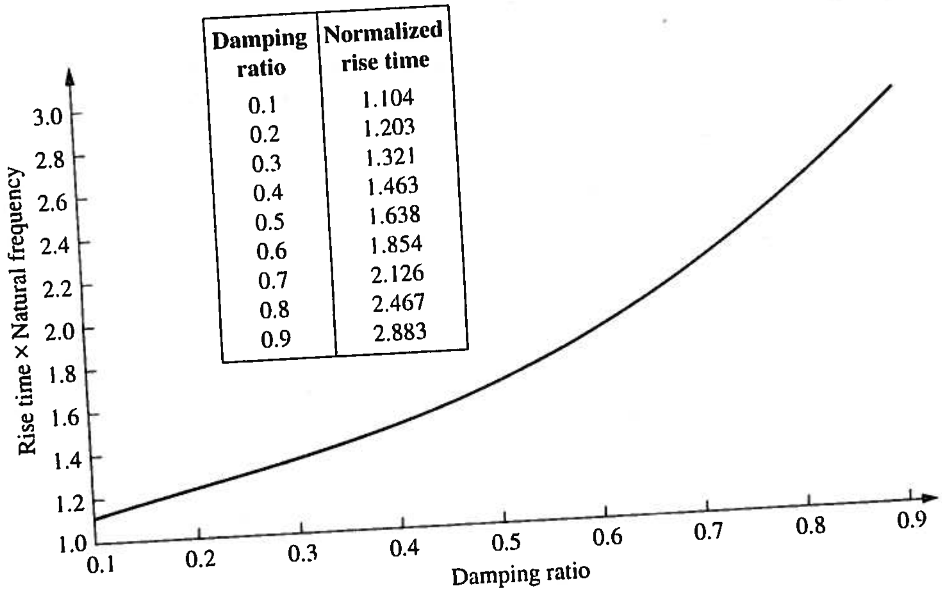


Figure 4.16 Normalized rise time versus damping ratio for a second-order underdamped response



Damping ratio	Normalized rise time
0.1	1.104
0.2	1.203
0.3	1.321
0.4	1.463
0.5	1.638
0.6	1.854
0.7	2.126
0.8	2.467
0.9	2.883

You can verify that the numerator of Eq. (4.41) varies from 3.91 to 4.74 as ζ varies from 0 to 0.9. Let us agree on an approximation for the settling time that will be used for all values of ζ; let it be

$$T_s = \frac{4}{\zeta\omega_n} \tag{4.42}$$

Evaluation of T_r

A precise analytical relationship between rise time and damping ratio, ζ, cannot be found. However, using a computer and Eq. (4.28), the rise time can be found. We first designate $\omega_n t$ as the normalized time variable and select a value for ζ. Using the computer, we solve for the values of $\omega_n t$ that yield $c(t) = 0.9$ and $c(t) = 0.1$. Subtracting the two values of $\omega_n t$ yields the normalized rise time, $\omega_n T_r$, for that value of ζ. Continuing in like fashion with other values of ζ, we obtain the results plotted in Figure 4.16.[5] Let us look at an example.

Example 4.5

Finding T_p, %OS, T_s, and T_r from a transfer function

Problem Given the transfer function

$$G(s) = \frac{100}{s^2 + 15s + 100} \tag{4.43}$$

find T_p, $\%OS$, T_s, and T_r.

[5]Figure 4.16 can be approximated by the following polynomials: $\omega_n T_r = 1.76\zeta^3 - 0.417\zeta^2 + 1.039\zeta + 1$ (maximum error less than $\frac{1}{2}\%$ for $0 < \zeta < 0.9$), and $\zeta = 0.115(\omega_n T_r)^3 - 0.883(\omega_n T_r)^2 + 2.504(\omega_n T_r) - 1.738$ (maximum error less than 5% for $0.1 < \zeta < 0.9$). The polynomials were obtained using MATLAB's **polyfit** function.

Solution ω_n and ζ are calculated as 10 and 0.75, respectively. Now substitute ζ and ω_n into Eqs. (4.34), (4.38), and (4.42) and find, respectively, that $T_p = 0.475$ second, $\%OS = 2.838$, and $T_s = 0.533$ second. Using the table in Figure 4.16, the normalized rise time is approximately 2.3 seconds. Dividing by ω_n yields $T_r = 0.23$ second. This problem demonstrates that we can find T_p, $\%OS$, T_s, and T_r without the tedious task of taking an inverse Laplace transform, plotting the output response, and taking measurements from the plot.

We now have expressions that relate peak time, percent overshoot, and settling time to the natural frequency and the damping ratio. Now let us relate these quantities to the location of the poles that generate these characteristics.

The pole plot for a general, underdamped second-order system, previously shown in Figure 4.11, is reproduced and expanded in Figure 4.17 for focus. We see from the Pythagorean theorem that the radial distance from the origin to the pole is the natural frequency, ω_n, and the $\cos\theta = \zeta$.

Now, comparing Eqs. (4.34) and (4.42) with the pole location, we evaluate peak time and settling time in terms of the pole location. Thus,

$$T_p = \frac{\pi}{\omega_n\sqrt{1-\zeta^2}} = \frac{\pi}{\omega_d} \tag{4.44}$$

$$T_s = \frac{4}{\zeta\omega_n} = \frac{4}{\sigma_d} \tag{4.45}$$

where ω_d is the imaginary part of the pole and is called the *damped frequency of oscillation*, and σ_d is the magnitude of the real part of the pole and is the *exponential damping frequency.*

Equation (4.44) shows that T_p is inversely proportional to the imaginary part of the pole. Since horizontal lines on the s-plane are lines of constant imaginary value, they are also lines of constant peak time. Similarly, Eq. (4.45) tells us that settling time is inversely proportional to the real part of the pole. Since vertical lines on the s-plane are lines of constant real value, they are also lines of constant settling time. Finally, since $\zeta = \cos\theta$, radial lines are lines of constant ζ. Since

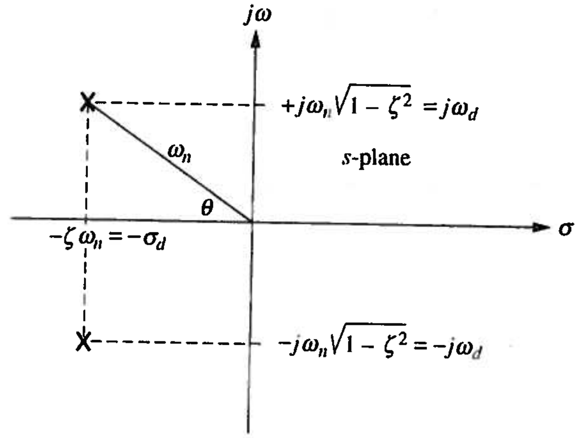


Figure 4.17 Pole plot for an underdamped second-order system

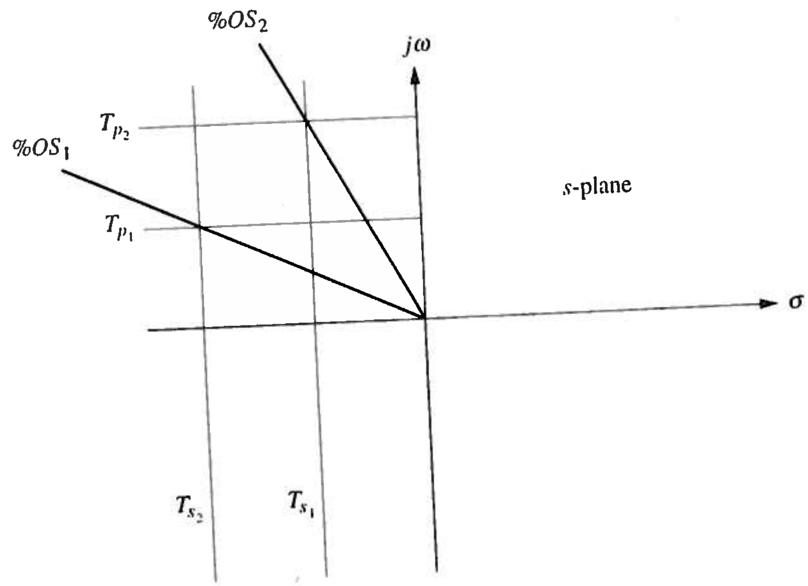


Figure 4.18 Lines of constant peak time, T_p, settling time, T_s, and percent overshoot, %OS. Note: $T_{s_2} < T_{s_1}$; $T_{p_2} < T_{p_1}$; $\%OS_1 < \%OS_2$.

percent overshoot is only a function of ζ, radial lines are thus lines of constant percent overshoot, $\%OS$. These concepts are depicted in Figure 4.18, where lines of constant T_p, T_s, and $\%OS$ are labeled on the s-plane.

At this point we can understand the significance of Figure 4.18 by examining the actual step response of comparative systems. Depicted in Figure 4.19(a) are the step responses as the poles are moved in a vertical direction, keeping the real part the same. As the poles move in a vertical direction, the frequency increases, but the envelope remains the same since the real part of the pole is not changing. The figure shows a constant exponential envelope, even though the sinusoidal response is changing frequency. Since all curves fit under the same exponential decay curve, the settling time is virtually the same for all waveforms. Note that as overshoot increases, the rise time decreases.

Let us move the poles to the right or left. Since the imaginary part is now constant, movement of the poles yields the responses of Figure 4.19(b). Here the frequency is constant over the range of variation of the real part. As the poles move to the left, the response damps out more rapidly, while the frequency remains the same. Notice that the peak time is the same for all waveforms because the imaginary part remains the same.

Moving the poles along a constant radial line yields the responses shown in Figure 4.19(c). Here the percent overshoot remains the same. Notice also that the responses look exactly alike, except for their speed. The farther the poles are from the origin, the more rapid the response.

We conclude this section with some examples that demonstrate the relationship between the pole location and the specifications of the second-order underdamped response. The first example covers analysis. The second example is a simple design problem consisting of a physical system whose component values we want to design to meet a transient response specification.

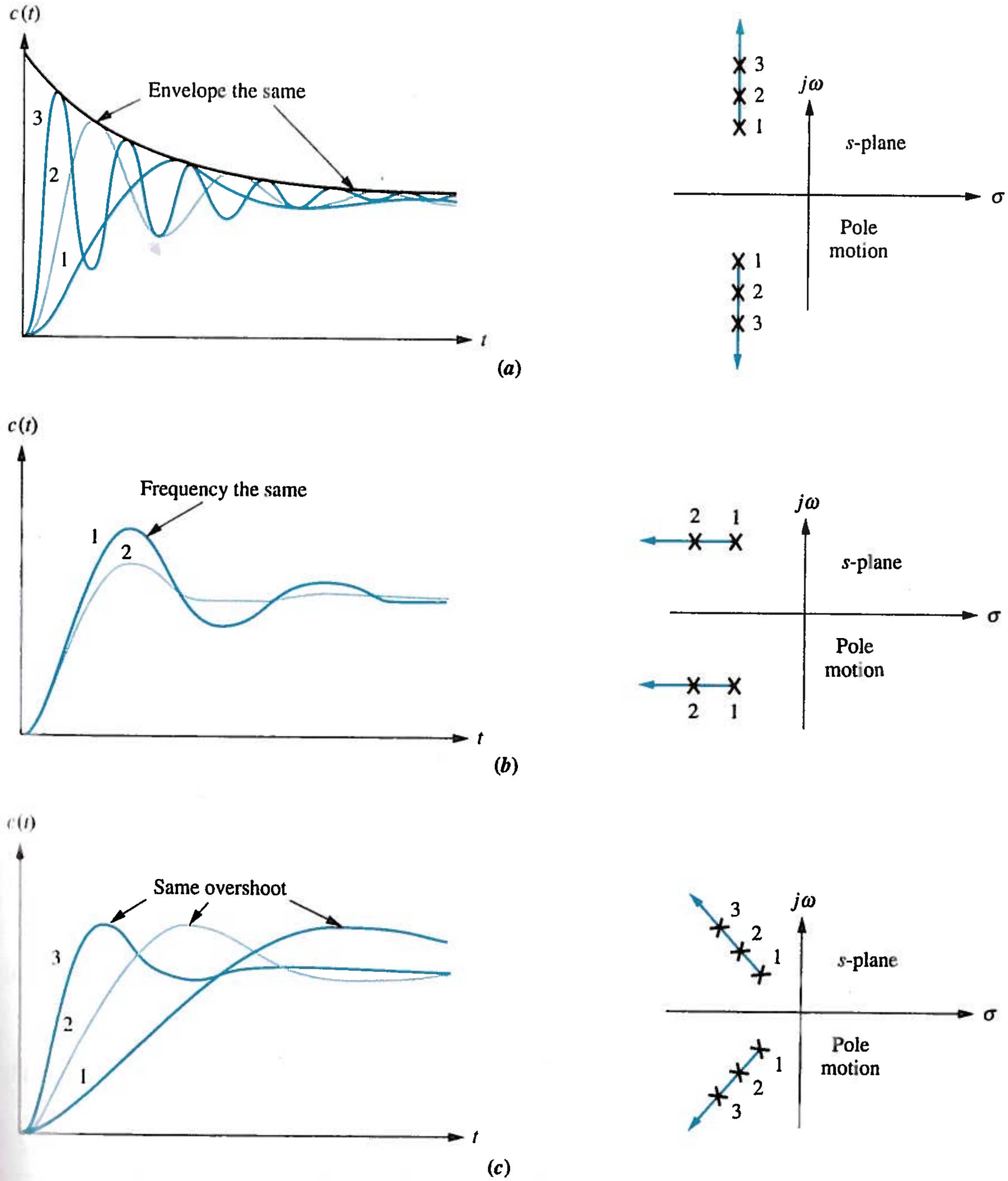


Figure 4.19
Step responses of second-order underdamped systems as poles move:
a. with constant real part;
b. with constant imaginary part;
c. with constant damping ratio

Example 4.6

Finding T_p, %OS, and T_s from pole location

Problem Given the pole plot shown in Figure 4.20, find ζ, ω_n, T_p, $\%OS$, and T_s.

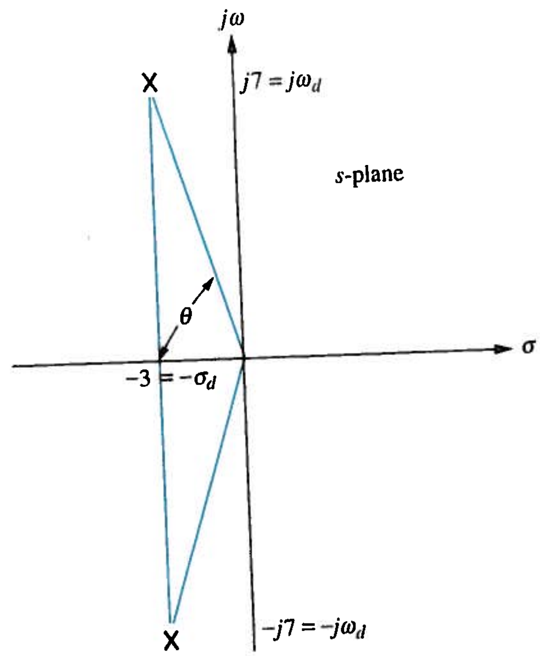


Figure 4.20 Pole plot for Example 4.6

Solution The damping ratio is given by $\zeta = \cos\theta = \cos[\arctan(7/3)] = 0.394$. The natural frequency, ω_n, is the radial distance from the origin to the pole, or $\omega_n = \sqrt{7^2 + 3^2} = 7.616$. The peak time is

$$T_p = \frac{\pi}{\omega_d} = \frac{\pi}{7} = 0.449 \text{ second} \tag{4.46}$$

The percent overshoot is

$$\%OS = e^{-(\zeta\pi/\sqrt{1-\zeta^2})} \times 100 = 26\% \tag{4.47}$$

The approximate settling time is

$$T_s = \frac{4}{\sigma_d} = \frac{4}{3} = 1.333 \text{ seconds} \tag{4.48}$$

MATLAB

Students who are using MATLAB should now run ch4p1 in Appendix B. You will learn how to generate a second-order polynomial from two complex poles as well as extract and use the coefficients of the polynomial to calculate T_p, %OS, and T_s. This exercise uses MATLAB to solve the problem in Example 4.6.

Example 4.7

Design

Transient response through component design

Problem Given the system shown in Figure 4.21, find J and D to yield 20% overshoot and a settling time of 2 seconds for a step input of torque $T(t)$.

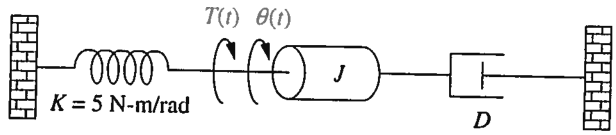


Figure 4.21 Rotational mechanical system for Example 4.7

Solution First, the transfer function for the system is

$$G(s) = \frac{1/J}{s^2 + \frac{D}{J}s + \frac{K}{J}} \tag{4.49}$$

From the transfer function,

$$\omega_n = \sqrt{\frac{K}{J}} \tag{4.50}$$

and

$$2\zeta\omega_n = \frac{D}{J} \tag{4.51}$$

But, from the problem statement,

$$T_s = 2 = \frac{4}{\zeta\omega_n} \tag{4.52}$$

or $\zeta\omega_n = 2$. Hence,

$$2\zeta\omega_n = 4 = \frac{D}{J} \tag{4.53}$$

Also, from Eqs. (4.50) and (4.52),

$$\zeta = \frac{4}{2\omega_n} = 2\sqrt{\frac{J}{K}} \tag{4.54}$$

From Eq. (4.39), a 20% overshoot implies $\zeta = 0.456$. Therefore, from Eq. (4.54),

$$\zeta = 2\sqrt{\frac{J}{K}} = 0.456 \tag{4.55}$$

Hence,

$$\frac{J}{K} = 0.052 \tag{4.56}$$