analyses of the step response, ramp response, and impulse response of the second-order systems are presented. Section 5–4 discusses the transient-response analysis of higher-order systems. Section 5–5 gives an introduction to the MATLAB approach to the solution of transient-response problems. Section 5–6 gives an example of a transient-response problem solved with MATLAB. Section 5–7 presents Routh's stability criterion. Section 5–8 discusses effects of integral and derivative control actions on system performance. Finally, Section 5–9 treats steady-state errors in unity-feedback control systems.

5–2 FIRST-ORDER SYSTEMS

Consider the first-order system shown in Figure 5–1(a). Physically, this system may represent an RC circuit, thermal system, or the like. A simplified block diagram is shown in Figure 5–1(b). The input-output relationship is given by

$$\frac{C(s)}{R(s)} = \frac{1}{Ts + 1} \tag{5–1}$$

In the following, we shall analyze the system responses to such inputs as the unit-step, unit-ramp, and unit-impulse functions. The initial conditions are assumed to be zero.

Note that all systems having the same transfer function will exhibit the same output in response to the same input. For any given physical system, the mathematical response can be given a physical interpretation.

Unit-Step Response of First-Order Systems. Since the Laplace transform of the unit-step function is $1/s$, substituting $R(s) = 1/s$ into Equation (5–1), we obtain

$$C(s) = \frac{1}{Ts + 1}\frac{1}{s}$$

Expanding $C(s)$ into partial fractions gives

$$C(s) = \frac{1}{s} - \frac{T}{Ts + 1} = \frac{1}{s} - \frac{1}{s + (1/T)} \tag{5–2}$$

Taking the inverse Laplace transform of Equation (5–2), we obtain

$$c(t) = 1 - e^{-t/T}, \qquad \text{for } t \geq 0 \tag{5–3}$$

Equation (5–3) states that initially the output $c(t)$ is zero and finally it becomes unity. One important characteristic of such an exponential response curve $c(t)$ is that at $t = T$ the value of $c(t)$ is 0.632, or the response $c(t)$ has reached 63.2% of its total change. This may be easily seen by substituting $t = T$ in $c(t)$. That is,

$$c(T) = 1 - e^{-1} = 0.632$$

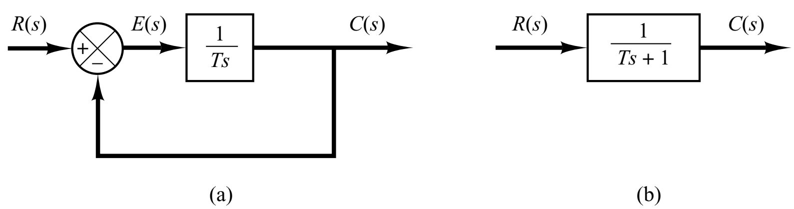

Figure 5–1
(a) Block diagram of a first-order system; (b) simplified block diagram.

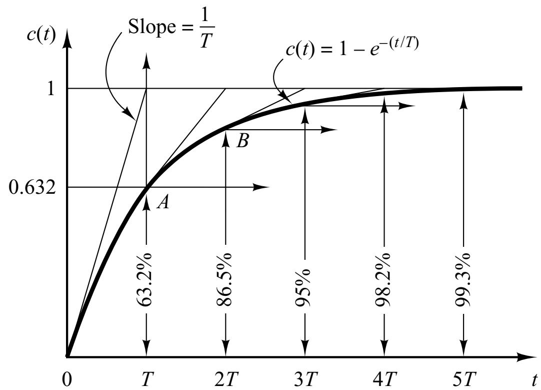

Figure 5–2 Exponential response curve.

Note that the smaller the time constant T, the faster the system response. Another important characteristic of the exponential response curve is that the slope of the tangent line at $t = 0$ is $1/T$, since

$$\left.\frac{dc}{dt}\right|_{t=0} = \left.\frac{1}{T}e^{-t/T}\right|_{t=0} = \frac{1}{T} \tag{5–4}$$

The output would reach the final value at $t = T$ if it maintained its initial speed of response. From Equation (5–4) we see that the slope of the response curve $c(t)$ decreases monotonically from $1/T$ at $t = 0$ to zero at $t = \infty$.

The exponential response curve $c(t)$ given by Equation (5–3) is shown in Figure 5–2. In one time constant, the exponential response curve has gone from 0 to 63.2% of the final value. In two time constants, the response reaches 86.5% of the final value. At $t = 3T, 4T$, and $5T$, the response reaches 95%, 98.2%, and 99.3%, respectively, of the final value. Thus, for $t \geq 4T$, the response remains within 2% of the final value. As seen from Equation (5–3), the steady state is reached mathematically only after an infinite time. In practice, however, a reasonable estimate of the response time is the length of time the response curve needs to reach and stay within the 2% line of the final value, or four time constants.

Unit-Ramp Response of First-Order Systems. Since the Laplace transform of the unit-ramp function is $1/s^2$, we obtain the output of the system of Figure 5–1(a) as

$$C(s) = \frac{1}{Ts + 1}\frac{1}{s^2}$$

Expanding $C(s)$ into partial fractions gives

$$C(s) = \frac{1}{s^2} - \frac{T}{s} + \frac{T^2}{Ts + 1} \tag{5–5}$$

Taking the inverse Laplace transform of Equation (5–5), we obtain

$$c(t) = t - T + Te^{-t/T}, \qquad \text{for } t \geq 0 \tag{5–6}$$

The error signal $e(t)$ is then

$$\begin{aligned} e(t) &= r(t) - c(t) \\ &= T\left(1 - e^{-t/T}\right) \end{aligned}$$

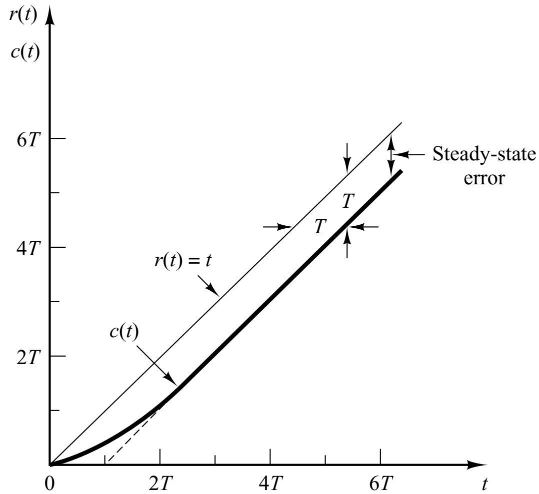

Figure 5–3 Unit-ramp response of the system shown in Figure 5–1(a).

As t approaches infinity, $e^{-t/T}$ approaches zero, and thus the error signal $e(t)$ approaches T or

$$e(\infty) = T$$

The unit-ramp input and the system output are shown in Figure 5–3. The error in following the unit-ramp input is equal to T for sufficiently large t. The smaller the time constant T, the smaller the steady-state error in following the ramp input.

Unit-Impulse Response of First-Order Systems. For the unit-impulse input, $R(s) = 1$ and the output of the system of Figure 5–1(a) can be obtained as

$$C(s) = \frac{1}{Ts + 1} \tag{5–7}$$

The inverse Laplace transform of Equation (5–7) gives

$$c(t) = \frac{1}{T} e^{-t/T}, \qquad \text{for } t \geq 0 \tag{5–8}$$

The response curve given by Equation (5–8) is shown in Figure 5–4.

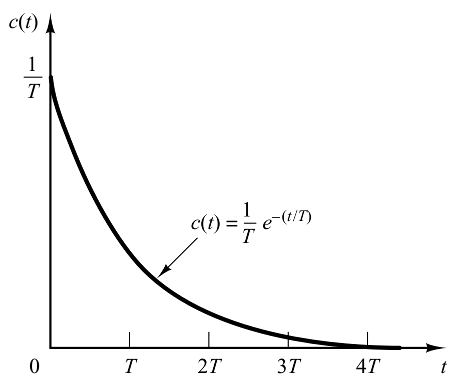

Figure 5–4 Unit-impulse response of the system shown in Figure 5–1(a).

An Important Property of Linear Time-Invariant Systems. In the analysis above, it has been shown that for the unit-ramp input the output $c(t)$ is

$$c(t) = t - T + Te^{-t/T}, \qquad \text{for } t \geq 0 \qquad \text{[See Equation (5–6).]}$$

For the unit-step input, which is the derivative of unit-ramp input, the output $c(t)$ is

$$c(t) = 1 - e^{-t/T}, \qquad \text{for } t \geq 0 \qquad \text{[See Equation (5–3).]}$$

Finally, for the unit-impulse input, which is the derivative of unit-step input, the output $c(t)$ is

$$c(t) = \frac{1}{T} e^{-t/T}, \qquad \text{for } t \geq 0 \qquad \text{[See Equation (5–8).]}$$

Comparing the system responses to these three inputs clearly indicates that the response to the derivative of an input signal can be obtained by differentiating the response of the system to the original signal. It can also be seen that the response to the integral of the original signal can be obtained by integrating the response of the system to the original signal and by determining the integration constant from the zero-output initial condition. This is a property of linear time-invariant systems. Linear time-varying systems and nonlinear systems do not possess this property.

5–3 SECOND-ORDER SYSTEMS

In this section, we shall obtain the response of a typical second-order control system to a step input, ramp input, and impulse input. Here we consider a servo system as an example of a second-order system.

Servo System. The servo system shown in Figure 5–5(a) consists of a proportional controller and load elements (inertia and viscous-friction elements). Suppose that we wish to control the output position c in accordance with the input position r.

The equation for the load elements is

$$J\ddot{c} + B\dot{c} = T$$

where T is the torque produced by the proportional controller whose gain is K. By taking Laplace transforms of both sides of this last equation, assuming the zero initial conditions, we obtain

$$Js^2C(s) + BsC(s) = T(s)$$

So the transfer function between $C(s)$ and $T(s)$ is

$$\frac{C(s)}{T(s)} = \frac{1}{s(Js + B)}$$

By using this transfer function, Figure 5–5(a) can be redrawn as in Figure 5–5(b), which can be modified to that shown in Figure 5–5(c). The closed-loop transfer function is then obtained as

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + Bs + K} = \frac{K/J}{s^2 + (B/J)s + (K/J)}$$

Such a system where the closed-loop transfer function possesses two poles is called a second-order system. (Some second-order systems may involve one or two zeros.)

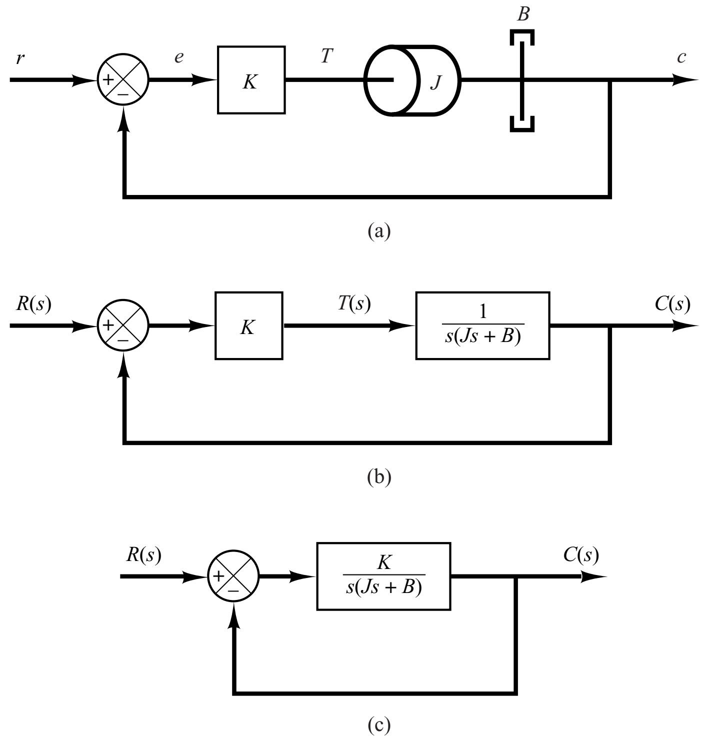

Figure 5–5
(a) Servo system;
(b) block diagram;
(c) simplified block diagram.

Step Response of Second-Order System. The closed-loop transfer function of the system shown in Figure 5–5(c) is

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + Bs + K} \tag{5–9}$$

which can be rewritten as

$$\frac{C(s)}{R(s)} = \frac{\frac{K}{J}}{\left[s + \frac{B}{2J} + \sqrt{\left(\frac{B}{2J}\right)^2 - \frac{K}{J}}\right]\left[s + \frac{B}{2J} - \sqrt{\left(\frac{B}{2J}\right)^2 - \frac{K}{J}}\right]}$$

The closed-loop poles are complex conjugates if $B^2 - 4JK < 0$ and they are real if $B^2 - 4JK \geq 0$. In the transient-response analysis, it is convenient to write

$$\frac{K}{J} = \omega_n^2, \qquad \frac{B}{J} = 2\zeta\omega_n = 2\sigma$$

where σ is called the *attenuation*; ω_n, the *undamped natural frequency*; and ζ, the *damping ratio* of the system. The damping ratio ζ is the ratio of the actual damping B to the critical damping $B_c = 2\sqrt{JK}$ or

$$\zeta = \frac{B}{B_c} = \frac{B}{2\sqrt{JK}}$$

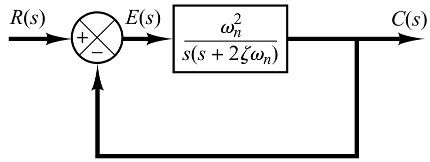

Figure 5–6
Second-order system.

In terms of ζ and ω_n, the system shown in Figure 5–5(c) can be modified to that shown in Figure 5–6, and the closed-loop transfer function $C(s)/R(s)$ given by Equation (5–9) can be written

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{5–10}$$

This form is called the *standard form* of the second-order system.

The dynamic behavior of the second-order system can then be described in terms of two parameters ζ and ω_n. If $0 < \zeta < 1$, the closed-loop poles are complex conjugates and lie in the left-half s plane. The system is then called underdamped, and the transient response is oscillatory. If $\zeta = 0$, the transient response does not die out. If $\zeta = 1$, the system is called critically damped. Overdamped systems correspond to $\zeta > 1$.

We shall now solve for the response of the system shown in Figure 5–6 to a unit-step input. We shall consider three different cases: the underdamped $(0 < \zeta < 1)$, critically damped $(\zeta = 1)$, and overdamped $(\zeta > 1)$ cases.

(1) *Underdamped case* $(0 < \zeta < 1)$: In this case, $C(s)/R(s)$ can be written

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{(s + \zeta\omega_n + j\omega_d)(s + \zeta\omega_n - j\omega_d)}$$

where $\omega_d = \omega_n\sqrt{1 - \zeta^2}$. The frequency ω_d is called the *damped natural frequency*. For a unit-step input, $C(s)$ can be written

$$C(s) = \frac{\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s} \tag{5–11}$$

The inverse Laplace transform of Equation (5–11) can be obtained easily if $C(s)$ is written in the following form:

$$\begin{aligned} C(s) &= \frac{1}{s} - \frac{s + 2\zeta\omega_n}{s^2 + 2\zeta\omega_n s + \omega_n^2} \\ &= \frac{1}{s} - \frac{s + \zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_d^2} - \frac{\zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_d^2} \end{aligned}$$

Referring to the Laplace transform table in Appendix A, it can be shown that

$$\mathcal{L}^{-1}\left[\frac{s + \zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_d^2}\right] = e^{-\zeta\omega_n t}\cos\omega_d t$$

$$\mathcal{L}^{-1}\left[\frac{\omega_d}{(s + \zeta\omega_n)^2 + \omega_d^2}\right] = e^{-\zeta\omega_n t}\sin\omega_d t$$

Hence the inverse Laplace transform of Equation (5–11) is obtained as

$$\begin{aligned}\mathscr{L}^{-1}[C(s)] &= c(t)\\ &= 1 - e^{-\zeta\omega_n t}\left(\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_d t\right)\\ &= 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\omega_d t + \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right), \qquad \text{for } t \geq 0 \end{aligned} \tag{5–12}$$

From Equation (5–12), it can be seen that the frequency of transient oscillation is the damped natural frequency ω_d and thus varies with the damping ratio ζ. The error signal for this system is the difference between the input and output and is

$$\begin{aligned} e(t) &= r(t) - c(t)\\ &= e^{-\zeta\omega_n t}\left(\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_d t\right), \qquad \text{for } t \geq 0\end{aligned}$$

This error signal exhibits a damped sinusoidal oscillation. At steady state, or at $t = \infty$, no error exists between the input and output.

If the damping ratio ζ is equal to zero, the response becomes undamped and oscillations continue indefinitely. The response $c(t)$ for the zero damping case may be obtained by substituting $\zeta = 0$ in Equation (5–12), yielding

$$c(t) = 1 - \cos\omega_n t, \qquad \text{for } t \geq 0 \tag{5–13}$$

Thus, from Equation (5–13), we see that ω_n represents the undamped natural frequency of the system. That is, ω_n is that frequency at which the system output would oscillate if the damping were decreased to zero. If the linear system has any amount of damping, the undamped natural frequency cannot be observed experimentally. The frequency that may be observed is the damped natural frequency ω_d, which is equal to $\omega_n\sqrt{1-\zeta^2}$. This frequency is always lower than the undamped natural frequency. An increase in ζ would reduce the damped natural frequency ω_d. If ζ is increased beyond unity, the response becomes overdamped and will not oscillate.

(2) *Critically damped case* ($\zeta = 1$): If the two poles of $C(s)/R(s)$ are equal, the system is said to be a critically damped one.

For a unit-step input, $R(s) = 1/s$ and $C(s)$ can be written

$$C(s) = \frac{\omega_n^2}{(s+\omega_n)^2 s} \tag{5–14}$$

The inverse Laplace transform of Equation (5–14) may be found as

$$c(t) = 1 - e^{-\omega_n t}(1 + \omega_n t), \qquad \text{for } t \geq 0 \tag{5–15}$$

This result can also be obtained by letting ζ approach unity in Equation (5–12) and by using the following limit:

$$\lim_{\zeta\to 1}\frac{\sin\omega_d t}{\sqrt{1-\zeta^2}} = \lim_{\zeta\to 1}\frac{\sin\omega_n\sqrt{1-\zeta^2}\,t}{\sqrt{1-\zeta^2}} = \omega_n t$$

(3) *Overdamped case* ($\zeta > 1$): In this case, the two poles of $C(s)/R(s)$ are negative real and unequal. For a unit-step input, $R(s) = 1/s$ and $C(s)$ can be written

$$C(s) = \frac{\omega_n^2}{\left(s + \zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}\right)\left(s + \zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}\right)s} \tag{5–16}$$

The inverse Laplace transform of Equation (5–16) is

$$\begin{aligned} c(t) &= 1 + \frac{1}{2\sqrt{\zeta^2 - 1}\left(\zeta + \sqrt{\zeta^2 - 1}\right)} e^{-\left(\zeta + \sqrt{\zeta^2 - 1}\right)\omega_n t} \\ &\quad - \frac{1}{2\sqrt{\zeta^2 - 1}\left(\zeta - \sqrt{\zeta^2 - 1}\right)} e^{-\left(\zeta - \sqrt{\zeta^2 - 1}\right)\omega_n t} \\ &= 1 + \frac{\omega_n}{2\sqrt{\zeta^2 - 1}}\left(\frac{e^{-s_1 t}}{s_1} - \frac{e^{-s_2 t}}{s_2}\right), \qquad \text{for } t \geq 0 \end{aligned} \tag{5–17}$$

where $s_1 = \left(\zeta + \sqrt{\zeta^2 - 1}\right)\omega_n$ and $s_2 = \left(\zeta - \sqrt{\zeta^2 - 1}\right)\omega_n$. Thus, the response $c(t)$ includes two decaying exponential terms.

When ζ is appreciably greater than unity, one of the two decaying exponentials decreases much faster than the other, so the faster-decaying exponential term (which corresponds to a smaller time constant) may be neglected. That is, if $-s_2$ is located very much closer to the $j\omega$ axis than $-s_1$ (which means $|s_2| \ll |s_1|$), then for an approximate solution we may neglect $-s_1$. This is permissible because the effect of $-s_1$ on the response is much smaller than that of $-s_2$, since the term involving s_1 in Equation (5–17) decays much faster than the term involving s_2. Once the faster-decaying exponential term has disappeared, the response is similar to that of a first-order system, and $C(s)/R(s)$ may be approximated by

$$\frac{C(s)}{R(s)} = \frac{\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}}{s + \zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}} = \frac{s_2}{s + s_2}$$

This approximate form is a direct consequence of the fact that the initial values and final values of both the original $C(s)/R(s)$ and the approximate one agree with each other.

With the approximate transfer function $C(s)/R(s)$, the unit-step response can be obtained as

$$C(s) = \frac{\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}}{\left(s + \zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}\right)s}$$

The time response $c(t)$ is then

$$c(t) = 1 - e^{-\left(\zeta - \sqrt{\zeta^2 - 1}\right)\omega_n t}, \qquad \text{for } t \geq 0$$

This gives an approximate unit-step response when one of the poles of $C(s)/R(s)$ can be neglected.

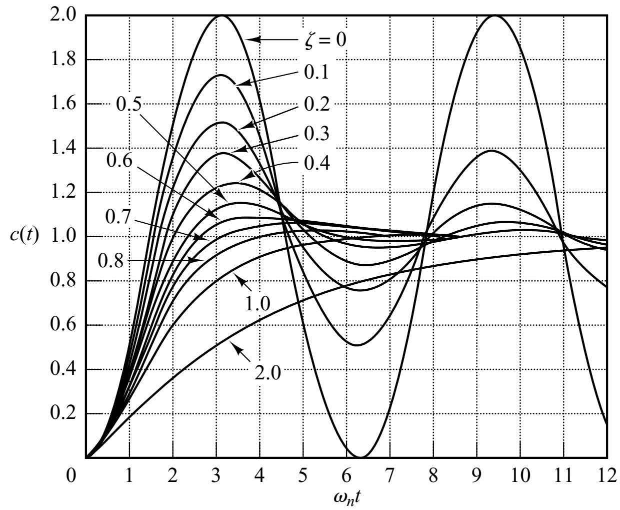

Figure 5–7
Unit-step response curves of the system shown in Figure 5–6.

A family of unit-step response curves $c(t)$ with various values of ζ is shown in Figure 5–7, where the abscissa is the dimensionless variable $\omega_n t$. The curves are functions only of ζ. These curves are obtained from Equations (5–12), (5–15), and (5–17). The system described by these equations was initially at rest.

Note that two second-order systems having the same ζ but different ω_n will exhibit the same overshoot and the same oscillatory pattern. Such systems are said to have the same relative stability.

From Figure 5–7, we see that an underdamped system with ζ between 0.5 and 0.8 gets close to the final value more rapidly than a critically damped or overdamped system. Among the systems responding without oscillation, a critically damped system exhibits the fastest response. An overdamped system is always sluggish in responding to any inputs.

It is important to note that, for second-order systems whose closed-loop transfer functions are different from that given by Equation (5–10), the step-response curves may look quite different from those shown in Figure 5–7.

Definitions of Transient-Response Specifications. Frequently, the performance characteristics of a control system are specified in terms of the transient response to a unit-step input, since it is easy to generate and is sufficiently drastic. (If the response to a step input is known, it is mathematically possible to compute the response to any input.)

The transient response of a system to a unit-step input depends on the initial conditions. For convenience in comparing transient responses of various systems, it is a common practice to use the standard initial condition that the system is at rest initially with the output and all time derivatives thereof zero. Then the response characteristics of many systems can be easily compared.

The transient response of a practical control system often exhibits damped oscillations before reaching steady state. In specifying the transient-response characteristics of a control system to a unit-step input, it is common to specify the following:

1. Delay time, t_d
2. Rise time, t_r

3. Peak time, t_p
4. Maximum overshoot, M_p
5. Settling time, t_s

These specifications are defined in what follows and are shown graphically in Figure 5–8.

1. Delay time, t_d: The delay time is the time required for the response to reach half the final value the very first time.
2. Rise time, t_r: The rise time is the time required for the response to rise from 10% to 90%, 5% to 95%, or 0% to 100% of its final value. For underdamped second-order systems, the 0% to 100% rise time is normally used. For overdamped systems, the 10% to 90% rise time is commonly used.
3. Peak time, t_p: The peak time is the time required for the response to reach the first peak of the overshoot.
4. Maximum (percent) overshoot, M_p: The maximum overshoot is the maximum peak value of the response curve measured from unity. If the final steady-state value of the response differs from unity, then it is common to use the maximum percent overshoot. It is defined by

$$\text{Maximum percent overshoot} = \frac{c(t_p) - c(\infty)}{c(\infty)} \times 100\%$$

   The amount of the maximum (percent) overshoot directly indicates the relative stability of the system.
5. Settling time, t_s: The settling time is the time required for the response curve to reach and stay within a range about the final value of size specified by absolute percentage of the final value (usually 2% or 5%). The settling time is related to the largest time constant of the control system. Which percentage error criterion to use may be determined from the objectives of the system design in question.

The time-domain specifications just given are quite important, since most control systems are time-domain systems; that is, they must exhibit acceptable time responses. (This means that, the control system must be modified until the transient response is satisfactory.)

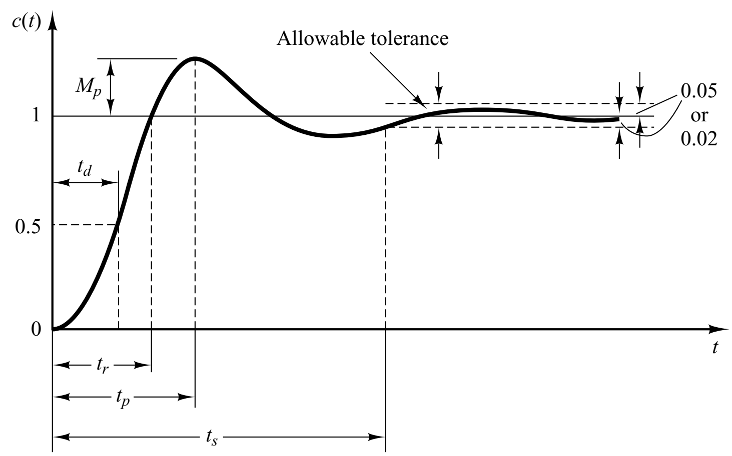

Figure 5–8
Unit-step response curve showing t_d, t_r, t_p, M_p, and t_s.

Note that not all these specifications necessarily apply to any given case. For example, for an overdamped system, the terms peak time and maximum overshoot do not apply. (For systems that yield steady-state errors for step inputs, this error must be kept within a specified percentage level. Detailed discussions of steady-state errors are postponed until Section 5–8.)

A Few Comments on Transient-Response Specifications. Except for certain applications where oscillations cannot be tolerated, it is desirable that the transient response be sufficiently fast and be sufficiently damped. Thus, for a desirable transient response of a second-order system, the damping ratio must be between 0.4 and 0.8. Small values of ζ (that is, $\zeta < 0.4$) yield excessive overshoot in the transient response, and a system with a large value of ζ (that is, $\zeta > 0.8$) responds sluggishly.

We shall see later that the maximum overshoot and the rise time conflict with each other. In other words, both the maximum overshoot and the rise time cannot be made smaller simultaneously. If one of them is made smaller, the other necessarily becomes larger.

Second-Order Systems and Transient-Response Specifications. In the following, we shall obtain the rise time, peak time, maximum overshoot, and settling time of the second-order system given by Equation (5–10). These values will be obtained in terms of ζ and ω_n. The system is assumed to be underdamped.

Rise time t_r: Referring to Equation (5–12), we obtain the rise time t_r by letting $c(t_r) = 1$.

$$c(t_r) = 1 = 1 - e^{-\zeta\omega_n t_r}\left(\cos\omega_d t_r + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_d t_r\right) \tag{5–18}$$

Since $e^{-\zeta\omega_n t_r} \neq 0$, we obtain from Equation (5–18) the following equation:

$$\cos\omega_d t_r + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_d t_r = 0$$

Since $\omega_n\sqrt{1-\zeta^2} = \omega_d$ and $\zeta\omega_n = \sigma$, we have

$$\tan\omega_d t_r = -\frac{\sqrt{1-\zeta^2}}{\zeta} = -\frac{\omega_d}{\sigma}$$

Thus, the rise time t_r is

$$t_r = \frac{1}{\omega_d}\tan^{-1}\left(\frac{\omega_d}{-\sigma}\right) = \frac{\pi-\beta}{\omega_d} \tag{5–19}$$

where angle β is defined in Figure 5–9. Clearly, for a small value of t_r, ω_d must be large.

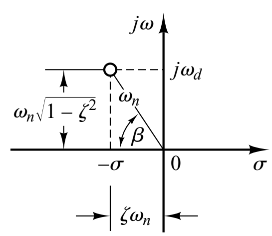

Figure 5–9 Definition of the angle β.

Peak time t_p: Referring to Equation (5–12), we may obtain the peak time by differentiating $c(t)$ with respect to time and letting this derivative equal zero. Since

$$\frac{dc}{dt} = \zeta\omega_n e^{-\zeta\omega_n t}\left(\cos\omega_d t + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\omega_d t\right)$$
$$+ e^{-\zeta\omega_n t}\left(\omega_d \sin\omega_d t - \frac{\zeta\omega_d}{\sqrt{1-\zeta^2}}\cos\omega_d t\right)$$

and the cosine terms in this last equation cancel each other, dc/dt, evaluated at $t = t_p$, can be simplified to

$$\left.\frac{dc}{dt}\right|_{t=t_p} = (\sin\omega_d t_p)\frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t_p} = 0$$

This last equation yields the following equation:

$$\sin\omega_d t_p = 0$$

or

$$\omega_d t_p = 0, \pi, 2\pi, 3\pi, \ldots$$

Since the peak time corresponds to the first peak overshoot, $\omega_d t_p = \pi$. Hence

$$t_p = \frac{\pi}{\omega_d} \tag{5–20}$$

The peak time t_p corresponds to one-half cycle of the frequency of damped oscillation.

Maximum overshoot M_p: The maximum overshoot occurs at the peak time or at $t = t_p = \pi/\omega_d$. Assuming that the final value of the output is unity, M_p is obtained from Equation (5–12) as

$$\begin{aligned} M_p &= c(t_p) - 1 \\ &= -e^{-\zeta\omega_n(\pi/\omega_d)}\left(\cos\pi + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\pi\right) \\ &= e^{-(\sigma/\omega_d)\pi} = e^{-(\zeta/\sqrt{1-\zeta^2})\pi} \end{aligned} \tag{5–21}$$

The maximum percent overshoot is $e^{-(\sigma/\omega_d)\pi} \times 100\%$.

If the final value $c(\infty)$ of the output is not unity, then we need to use the following equation:

$$M_p = \frac{c(t_p) - c(\infty)}{c(\infty)}$$

Settling time t_s: For an underdamped second-order system, the transient response is obtained from Equation (5–12) as

$$c(t) = 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\omega_d t + \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right), \qquad \text{for } t \geq 0$$

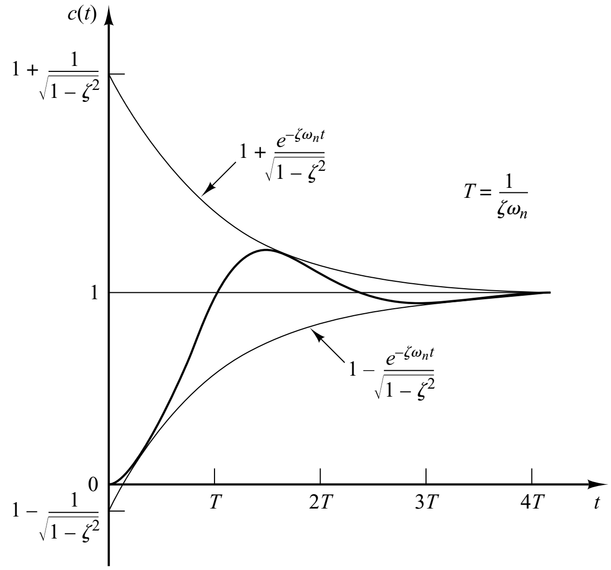

Figure 5–10 Pair of envelope curves for the unit-step response curve of the system shown in Figure 5–6.

The curves $1 \pm \left(e^{-\zeta\omega_n t}/\sqrt{1-\zeta^2}\right)$ are the envelope curves of the transient response to a unit-step input. The response curve $c(t)$ always remains within a pair of the envelope curves, as shown in Figure 5–10. The time constant of these envelope curves is $1/\zeta\omega_n$.

The speed of decay of the transient response depends on the value of the time constant $1/\zeta\omega_n$. For a given ω_n, the settling time t_s is a function of the damping ratio ζ. From Figure 5–7, we see that for the same ω_n and for a range of ζ between 0 and 1 the settling time t_s for a very lightly damped system is larger than that for a properly damped system. For an overdamped system, the settling time t_s becomes large because of the sluggish response.

The settling time corresponding to a $\pm 2\%$ or $\pm 5\%$ tolerance band may be measured in terms of the time constant $T = 1/\zeta\omega_n$ from the curves of Figure 5–7 for different values of ζ. The results are shown in Figure 5–11. For $0 < \zeta < 0.9$, if the 2% criterion is used, t_s is approximately four times the time constant of the system. If the 5% criterion is used, then t_s is approximately three times the time constant. Note that the settling time reaches a minimum value around $\zeta = 0.76$ (for the 2% criterion) or $\zeta = 0.68$ (for the 5% criterion) and then increases almost linearly for large values of ζ. The discontinuities in the curves of Figure 5–11 arise because an infinitesimal change in the value of ζ can cause a finite change in the settling time.

For convenience in comparing the responses of systems, we commonly define the settling time t_s to be

$$t_s = 4T = \frac{4}{\sigma} = \frac{4}{\zeta\omega_n} \qquad \text{(2\% criterion)} \tag{5–22}$$

or

$$t_s = 3T = \frac{3}{\sigma} = \frac{3}{\zeta\omega_n} \qquad \text{(5\% criterion)} \tag{5–23}$$

Note that the settling time is inversely proportional to the product of the damping ratio and the undamped natural frequency of the system. Since the value of ζ is usually determined from the requirement of permissible maximum overshoot, the settling time

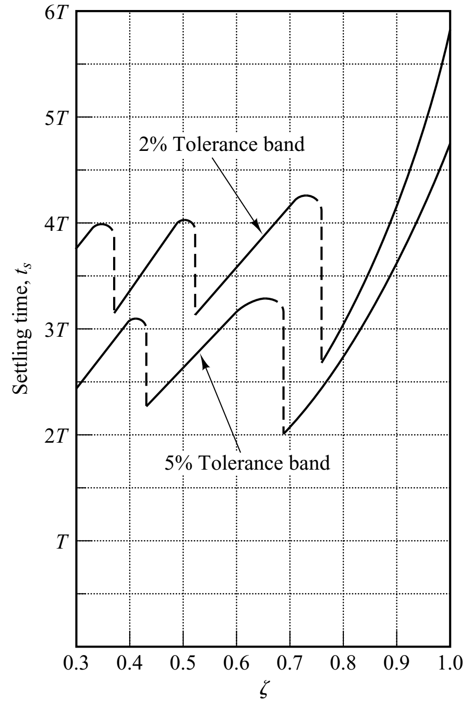

Figure 5–11
Settling time t_s versus ζ curves.

is determined primarily by the undamped natural frequency ω_n. This means that the duration of the transient period may be varied, without changing the maximum overshoot, by adjusting the undamped natural frequency ω_n.

From the preceding analysis, it is evident that for rapid response ω_n must be large. To limit the maximum overshoot M_p and to make the settling time small, the damping ratio ζ should not be too small. The relationship between the maximum percent overshoot M_p and the damping ratio ζ is presented in Figure 5–12. Note that if the damping ratio is between 0.4 and 0.7, then the maximum percent overshoot for step response is between 25% and 4%.

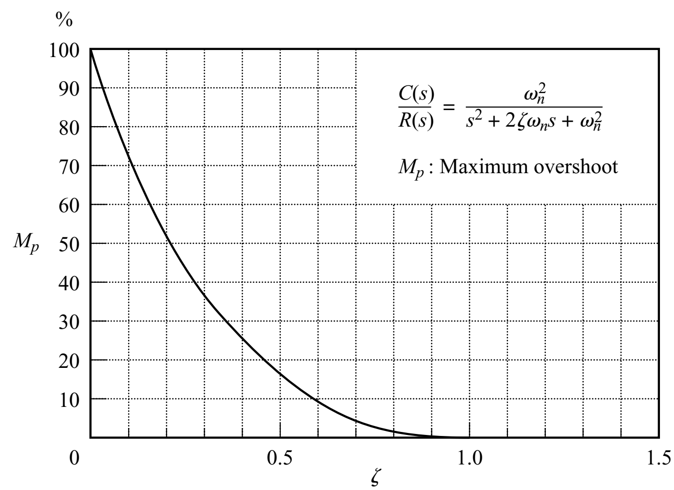

Figure 5–12
M_p versus ζ curve.

It is important to note that the equations for obtaining the rise time, peak time, maximum overshoot, and settling time are valid only for the standard second-order system defined by Equation (5–10). If the second-order system involves a zero or two zeros, the shape of the unit-step response curve will be quite different from those shown in Figure 5–7.

EXAMPLE 5–1 Consider the system shown in Figure 5–6, where $\zeta = 0.6$ and $\omega_n = 5$ rad/sec. Let us obtain the rise time t_r, peak time t_p, maximum overshoot M_p, and settling time t_s when the system is subjected to a unit-step input.

From the given values of ζ and ω_n, we obtain $\omega_d = \omega_n\sqrt{1-\zeta^2} = 4$ and $\sigma = \zeta\omega_n = 3$.

Rise time t_r: The rise time is

$$t_r = \frac{\pi - \beta}{\omega_d} = \frac{3.14 - \beta}{4}$$

where β is given by

$$\beta = \tan^{-1}\frac{\omega_d}{\sigma} = \tan^{-1}\frac{4}{3} = 0.93 \text{ rad}$$

The rise time t_r is thus

$$t_r = \frac{3.14 - 0.93}{4} = 0.55 \text{ sec}$$

Peak time t_p: The peak time is

$$t_p = \frac{\pi}{\omega_d} = \frac{3.14}{4} = 0.785 \text{ sec}$$

Maximum overshoot M_p: The maximum overshoot is

$$M_p = e^{-(\sigma/\omega_d)\pi} = e^{-(3/4)\times 3.14} = 0.095$$

The maximum percent overshoot is thus 9.5%.

Settling time t_s: For the 2% criterion, the settling time is

$$t_s = \frac{4}{\sigma} = \frac{4}{3} = 1.33 \text{ sec}$$

For the 5% criterion,

$$t_s = \frac{3}{\sigma} = \frac{3}{3} = 1 \text{ sec}$$

Servo System with Velocity Feedback. The derivative of the output signal can be used to improve system performance. In obtaining the derivative of the output position signal, it is desirable to use a tachometer instead of physically differentiating the output signal. (Note that the differentiation amplifies noise effects. In fact, if discontinuous noises are present, differentiation amplifies the discontinuous noises more than the useful signal. For example, the output of a potentiometer is a discontinuous voltage signal because, as the potentiometer brush is moving on the windings, voltages are induced in the switchover turns and thus generate transients. The output of the potentiometer therefore should not be followed by a differentiating element.)

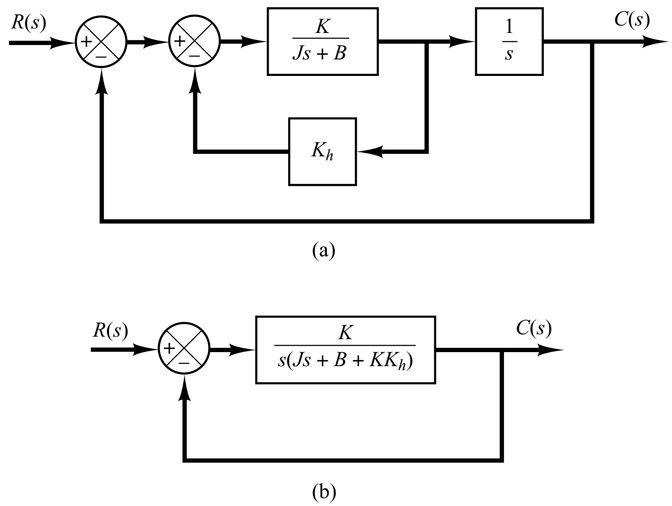

Figure 5–13
(a) Block diagram of a servo system;
(b) simplified block diagram.

The tachometer, a special dc generator, is frequently used to measure velocity without differentiation process. The output of a tachometer is proportional to the angular velocity of the motor.

Consider the servo system shown in Figure 5–13(a). In this device, the velocity signal, together with the positional signal, is fed back to the input to produce the actuating error signal. In any servo system, such a velocity signal can be easily generated by a tachometer. The block diagram shown in Figure 5–13(a) can be simplified, as shown in Figure 5–13(b), giving

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + (B + KK_h)s + K} \tag{5–24}$$

Comparing Equation (5–24) with Equation (5–9), notice that the velocity feedback has the effect of increasing damping. The damping ratio ζ becomes

$$\zeta = \frac{B + KK_h}{2\sqrt{KJ}} \tag{5–25}$$

The undamped natural frequency $\omega_n = \sqrt{K/J}$ is not affected by velocity feedback. Noting that the maximum overshoot for a unit-step input can be controlled by controlling the value of the damping ratio ζ, we can reduce the maximum overshoot by adjusting the velocity-feedback constant K_h so that ζ is between 0.4 and 0.7.

It is important to remember that velocity feedback has the effect of increasing the damping ratio without affecting the undamped natural frequency of the system.

EXAMPLE 5–2 For the system shown in Figure 5–13(a), determine the values of gain K and velocity-feedback constant K_h so that the maximum overshoot in the unit-step response is 0.2 and the peak time is 1 sec. With these values of K and K_h, obtain the rise time and settling time. Assume that $J = 1$ kg-m^2 and $B = 1$ N-m/rad/sec.

Determination of the values of K and K_h: The maximum overshoot M_p is given by Equation (5–21) as

$$M_p = e^{-(\zeta/\sqrt{1-\zeta^2})\pi}$$

This value must be 0.2. Thus,

$$e^{-(\zeta/\sqrt{1-\zeta^2})\pi} = 0.2$$

or

$$\frac{\zeta\pi}{\sqrt{1 - \zeta^2}} = 1.61$$

which yields

$$\zeta = 0.456$$

The peak time t_p is specified as 1 sec; therefore, from Equation (5–20),

$$t_p = \frac{\pi}{\omega_d} = 1$$

or

$$\omega_d = 3.14$$

Since ζ is 0.456, ω_n is

$$\omega_n = \frac{\omega_d}{\sqrt{1 - \zeta^2}} = 3.53$$

Since the natural frequency ω_n is equal to $\sqrt{K/J}$,

$$K = J\omega_n^2 = \omega_n^2 = 12.5 \text{ N-m}$$

Then K_h is, from Equation (5–25),

$$K_h = \frac{2\sqrt{KJ}\zeta - B}{K} = \frac{2\sqrt{K}\zeta - 1}{K} = 0.178 \text{ sec}$$

Rise time t_r: From Equation (5–19), the rise time t_r is

$$t_r = \frac{\pi - \beta}{\omega_d}$$

where

$$\beta = \tan^{-1}\frac{\omega_d}{\sigma} = \tan^{-1} 1.95 = 1.10$$

Thus, t_r is

$$t_r = 0.65 \text{ sec}$$

Settling time t_s: For the 2% criterion,

$$t_s = \frac{4}{\sigma} = 2.48 \text{ sec}$$

For the 5% criterion,

$$t_s = \frac{3}{\sigma} = 1.86 \text{ sec}$$

Impulse Response of Second-Order Systems. For a unit-impulse input $r(t)$, the corresponding Laplace transform is unity, or $R(s) = 1$. The unit-impulse response $C(s)$ of the second-order system shown in Figure 5-6 is

$$C(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

The inverse Laplace transform of this equation yields the time solution for the response $c(t)$ as follows:

For $0 \le \zeta < 1$,

$$c(t) = \frac{\omega_n}{\sqrt{1 - \zeta^2}} e^{-\zeta\omega_n t} \sin \omega_n \sqrt{1 - \zeta^2}\, t, \qquad \text{for } t \ge 0 \tag{5–26}$$

For $\zeta = 1$,

$$c(t) = \omega_n^2 t e^{-\omega_n t}, \qquad \text{for } t \ge 0 \tag{5–27}$$

For $\zeta > 1$,

$$c(t) = \frac{\omega_n}{2\sqrt{\zeta^2 - 1}} e^{-(\zeta - \sqrt{\zeta^2 - 1})\omega_n t} - \frac{\omega_n}{2\sqrt{\zeta^2 - 1}} e^{-(\zeta + \sqrt{\zeta^2 - 1})\omega_n t}, \qquad \text{for } t \ge 0 \tag{5–28}$$

Note that without taking the inverse Laplace transform of $C(s)$ we can also obtain the time response $c(t)$ by differentiating the corresponding unit-step response, since the unit-impulse function is the time derivative of the unit-step function. A family of unit-impulse response curves given by Equations (5–26) and (5–27) with various values of ζ is shown in Figure 5–14. The curves $c(t)/\omega_n$ are plotted against the dimensionless variable $\omega_n t$, and thus they are functions only of ζ. For the critically damped and overdamped cases, the unit-impulse response is always positive or zero; that is, $c(t) \ge 0$. This can be seen from Equations (5–27) and (5–28). For the underdamped case, the unit-impulse response $c(t)$ oscillates about zero and takes both positive and negative values.

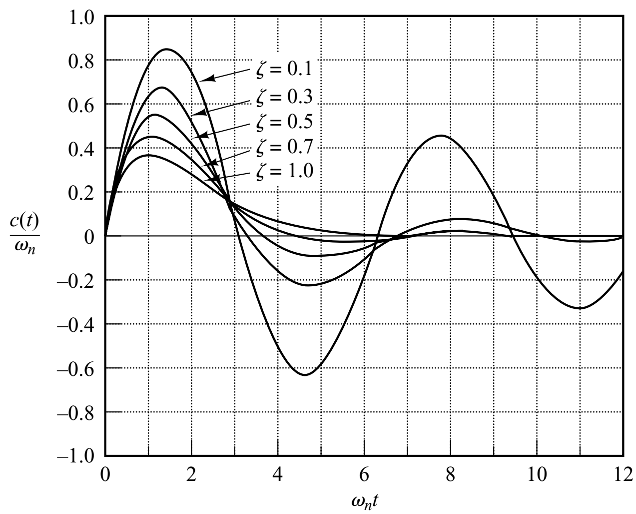

Figure 5–14 Unit-impulse response curves of the system shown in Figure 5–6.

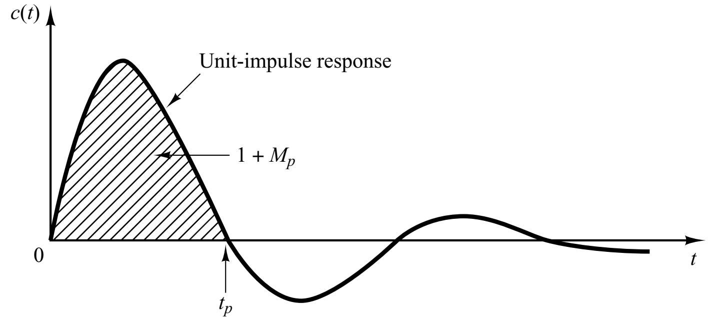

Figure 5–15 Unit-impulse response curve of the system shown in Figure 5–6.

From the foregoing analysis, we may conclude that if the impulse response $c(t)$ does not change sign, the system is either critically damped or overdamped, in which case the corresponding step response does not overshoot but increases or decreases monotonically and approaches a constant value.

The maximum overshoot for the unit-impulse response of the underdamped system occurs at

$$t = \frac{\tan^{-1} \dfrac{\sqrt{1-\zeta^2}}{\zeta}}{\omega_n \sqrt{1-\zeta^2}}, \qquad \text{where } 0 < \zeta < 1 \tag{5–29}$$

[Equation (5–29) can be obtained by equating dc/dt to zero and solving for t.] The maximum overshoot is

$$c(t)_{\max} = \omega_n \exp\left(-\frac{\zeta}{\sqrt{1-\zeta^2}} \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}\right), \qquad \text{where } 0 < \zeta < 1 \tag{5–30}$$

[Equation (5–30) can be obtained by substituting Equation (5–29) into Equation (5–26).]

Since the unit-impulse response function is the time derivative of the unit-step response function, the maximum overshoot M_p for the unit-step response can be found from the corresponding unit-impulse response. That is, the area under the unit-impulse response curve from $t = 0$ to the time of the first zero, as shown in Figure 5–15, is $1 + M_p$, where M_p is the maximum overshoot (for the unit-step response) given by Equation (5–21). The peak time t_p (for the unit-step response) given by Equation (5–20) corresponds to the time that the unit-impulse response first crosses the time axis.

5–4 HIGHER-ORDER SYSTEMS

In this section we shall present a transient-response analysis of higher-order systems in general terms. It will be seen that the response of a higher-order system is the sum of the responses of first-order and second-order systems.