Chapter Six

Transient and Steady State Responses

In control system analysis and design it is important to consider the complete system response and to design controllers such that a satisfactory response is obtained for all time instants $t \geq t_0$, where t_0 stands for the initial time. It is known that the system response has two components: transient response and steady state response, that is

$$y(t) = y_{tr}(t) + y_{ss}(t)$$

The transient response is present in the short period of time immediately after the system is turned on. If the system is asymptotically stable, the transient response disappears, which theoretically can be recorded as

$$\lim_{t\to\infty} y_{tr}(t) = 0$$

However, if the system is unstable, the transient response will increase very quickly (exponentially) in time, and in the most cases the system will be practically unusable or even destroyed during the unstable transient response (as can occur, for example, in some electrical networks).

Even if the system is asymptotically stable, the transient response should be carefully monitored since some undesired phenomena like high-frequency oscillations (e.g. in aircraft during landing and takeoff), rapid changes, and high magnitudes of the output may occur.

Assuming that the system is asymptotically stable, then the system response in the long run is determined by its steady state component only.

For control systems it is important that steady state response values are as close as possible to desired ones (specified ones) so that we have to study the corresponding errors, which represent the difference between the actual and desired system outputs at steady state, and examine conditions under which these errors can be reduced or even eliminated.

In Section 6.1 we find analytically the response of a second-order system due to a unit step input. The obtained result is used in Section 6.2 to define important parameters that characterize the system transient response. For higher-order systems, only approximations for the transient response parameters can be obtained using a computer. The steady state errors of linear control systems are defined in Section 6.4, and the feedback elements which help to reduce the steady state errors to zero are identified.

6.1 Response of Second-Order Systems

Consider the second-order feedback system represented, in general, by the block diagram given in Figure 6.1, where K represents the system static gain and T is the system time constant. It is quite easy to find the closed-loop transfer function of this system, that is

$$M(s) = \frac{Y(s)}{U(s)} = \frac{\frac{K}{T}}{s^2 + \frac{1}{T}s + \frac{K}{T}}$$

The closed-loop transfer function can be written in the following form

$$\frac{Y(s)}{U(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad \zeta = \frac{1}{2\omega_n T}, \qquad \omega_n^2 = \frac{K}{T}$$

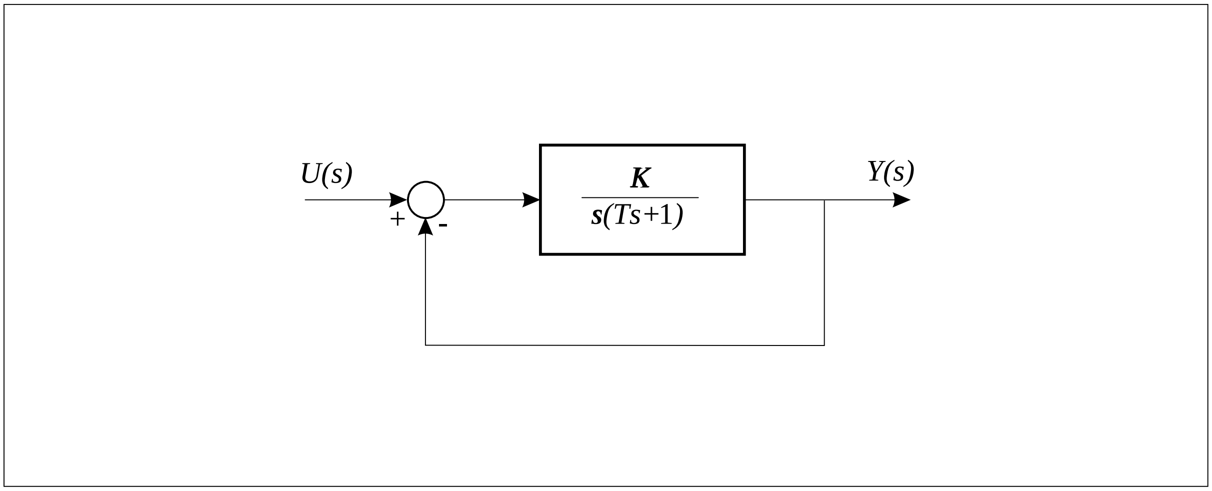

Figure 6.1: Block diagram of a general second-order system

Quantities ζ and ω_n are called, respectively, the *system damping ratio* and the *system natural frequency*. The system eigenvalues are given by

$$\lambda_{1,2} = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2} = -\zeta\omega_n \pm j\omega_d$$

where ω_d is the *system damped frequency*.

The location of the system poles and the relation between damping ratio, natural and damped frequencies are given in Figure 6.2.

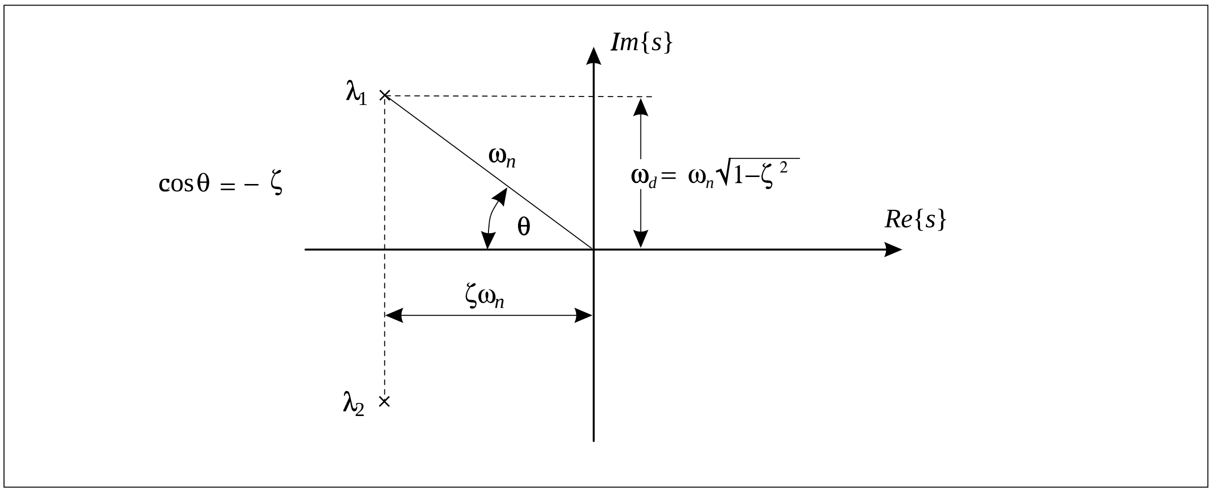

Figure 6.2: Second-order system eigenvalues in terms of parameters $\zeta, \omega_n, \omega_d$

In the following we find the closed-loop response of this second-order system due to a unit step input.

Since the Laplace transform of a unit step is $1/s$ we have

$$Y(s) = \frac{\omega_n^2}{s\left(s^2 + 2\zeta\omega_n s + \omega_n^2\right)}$$

Depending on the value of the damping ratio ζ three interesting cases appear: (a) the critically damped case, $\zeta = 1$; (b) the over-damped case, $\zeta > 1$; and (c) the under-damped case, $\zeta < 1$.

These cases are distinguished by the nature of the system eigenvalues. In case (a) the eigenvalues are multiple and real, in (b) they are real and distinct, and in case (c) the eigenvalues are complex conjugate.

(a) Critically Damped Case

For $\zeta = 1$, we have a double pole at $-\omega_n$. The corresponding output is

$$Y(s) = \frac{\omega_n^2}{s(s+\omega_n)^2} = \frac{1}{s} - \frac{1}{s+\omega_n} - \frac{\omega_n}{(s+\omega_n)^2}$$

which after taking the Laplace inverse produces

$$y(t) = 1 - e^{-\omega_n t} - \omega_n t e^{-\omega_n t}$$

The shape of this response is given in Figure 6.3a, where the location of the system poles $(\lambda_1 = p_1, \lambda_2 = p_2)$ is also presented.

(b) Over-Damped Case

For the over-damped case, we have two real and asymptotically stable poles at $-\zeta\omega_n \pm \omega_d$. The corresponding closed-loop response is easily obtained from

$$Y(s) = \frac{1}{s} + \frac{k_1}{s + \zeta\omega_n + \omega_d} + \frac{k_2}{s + \zeta\omega_n - \omega_d}$$

as

$$y(t) = 1 + k_1 e^{-(\zeta\omega_n + \omega_d)t} + k_2 e^{-(\zeta\omega_n - \omega_d)t}$$

It is represented in Figure 6.3b.

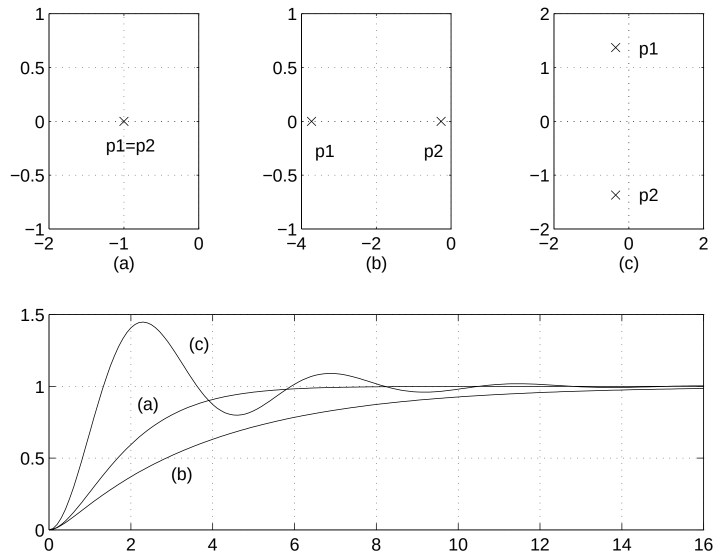

Figure 6.3: Responses of second-order systems and locations of system poles

(c) Under-Damped Case

This case is the most interesting and important one. The system has a pair of complex conjugate poles so that in the s-domain we have

$$Y(s) = \frac{k_1}{s} + \frac{k_2}{s + \zeta\omega_n + j\omega_d} + \frac{k_2^*}{s + \zeta\omega_n - j\omega_d}$$

Applying the Laplace transform it is easy to show (see Problem 6.1) that the system output in the time domain is given by

$$y(t) = 1 + \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin\left[\left(\omega_n\sqrt{1-\zeta^2}\right)t - \theta\right]$$

where from Figure 6.2 we have

$$\cos\theta = -\zeta, \qquad \sin\theta = \sqrt{1-\zeta^2}, \qquad \tan\theta = \frac{\sqrt{1-\zeta^2}}{-\zeta}$$

The response of this system is presented in Figure 6.3c.

The under-damped case is the most common in control system applications. A magnified figure of the system step response for the under-damped case is presented in Figure 6.4.

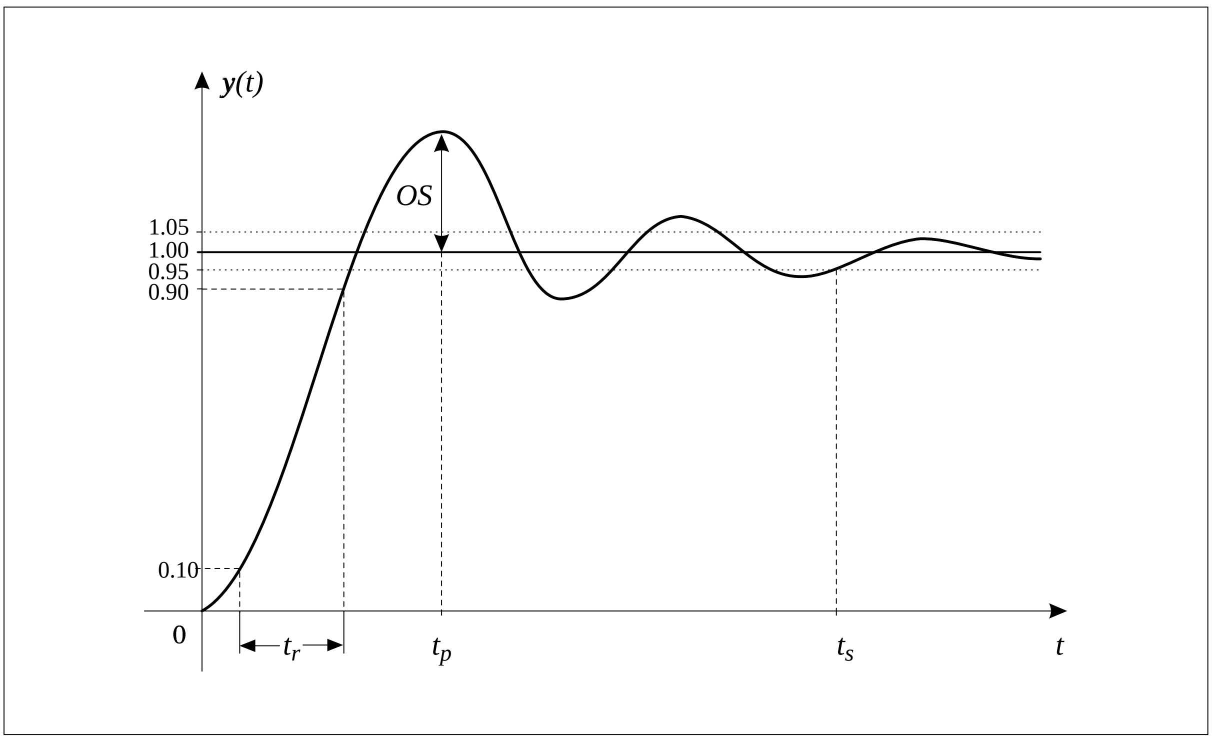

Figure 6.4: Response of an under-damped second-order system

6.2 Transient Response Parameters

The most important transient response parameters are denoted in Figure 6.4 as response overshoot, settling time, peak time, and rise time.

The response overshoot can be obtained by finding the maximum of the function $y(t)$ with respect to time. This leads to

$$\frac{dy(t)}{dt} = 0$$

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

or

$$\zeta\omega_n \sin(\omega_d t - \theta) - \omega_d \cos(\omega_d t - \theta) = 0$$

which implies

$$\sin \omega_d t = 0$$

From this equation we have

$$\omega_d t = i\pi, \quad i = 0, 1, 2, \ldots$$

The *peak time* is obtained for $i = 1$, i.e. as

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

and times for other minima and maxima are given by

$$t_{ip} = \frac{i\pi}{\omega_d} = \frac{i\pi}{\omega_n \sqrt{1 - \zeta^2}}, \quad i = 2, 3, 4, \ldots$$

Since the steady state value of $y(t)$ is $y_{ss}(t) = 1$, it follows that the *response overshoot* is given by

$$OS = y(t_p) - y_{ss}(t) = 1 + e^{-\zeta\omega_n t_p} - 1 = e^{-\zeta\omega_n t_p} = e^{-\frac{\zeta\pi}{\sqrt{1-\zeta^2}}}$$

Overshoot is very often expressed in percent, so that we can define the *maximum percent overshoot* as

$$MPOS = OS(\%) = e^{-\frac{\zeta\pi}{\sqrt{1-\zeta^2}}\times} 100(\%)$$

From Figure 6.4, the expression for the response 5 percent *settling time* can be obtained as

$$y(t_s) = 1 + \frac{e^{-\zeta\omega_n t_s}}{\sqrt{1-\zeta^2}} = 1.05$$

which for the standard values of ζ leads to

$$t_s = -\frac{1}{\zeta\omega_n}\ln\left(0.05\sqrt{1-\zeta^2}\right) \approx \frac{3}{\zeta\omega_n}$$

Note that in practice $0.5 < \zeta < 0.8$.

The response *rise time* is defined as the time required for the unit step response to change from 0.1 to 0.9 of its steady state value.

The rise time is inversely proportional to the system bandwidth, i.e. the wider bandwidth, the smaller the rise time. However, designing systems with wide bandwidth is costly, which indicates that systems with very fast response are expensive to design.

Example 6.1: Consider the following second-order system

$$\frac{Y(s)}{U(s)} = \frac{4}{s^2 + 2s + 4}$$

We have

$$\omega_n^2 = 4 \Rightarrow \omega_n = 2\,\mathrm{rad/s}, \quad 2\zeta\omega_n = 2 \Rightarrow \zeta = 0.5$$

$$\omega_d = \omega_n\sqrt{1-\zeta^2} = \sqrt{3}\ \mathrm{rad/s}$$

The peak time is obtained as

$$t_p = \frac{\pi}{\omega_d} = \frac{\pi}{\sqrt{3}} = 1.82\,\mathrm{s}$$

and the settling time is found to be

$$t_s \approx \frac{3}{\zeta\omega_n} = 3\,\mathrm{s}$$

The maximum percent overshoot is equal to

$$MPOS = e^{-\frac{\zeta\pi}{\sqrt{1-\zeta^2}}} 100(\%) = 16.3\%$$

The step response obtained using the MATLAB functions `[y,x]=step(num,den,t); t=0:0.1:5` is given in Figure 6.5.

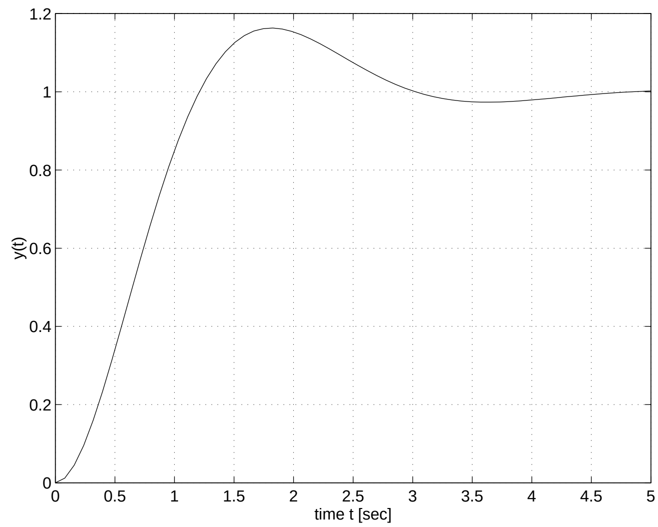

Figure 6.5: System step response for Example 6.1

It can be seen that the analytically obtained results agree with the results presented in Figure 6.5. From Figure 6.5 we are able to estimate the rise time, which in this case is approximately equal to $\boldsymbol{t_r \approx 0.8}\,\mathbf{s}$.

6.3 Transient Response of High-Order Systems

In the previous section we have been able to precisely define and determine parameters that characterize the system transient response. This has been possible due to the fact that the system under consideration has been of order two only. For higher-order systems, analytical expressions for the system response are not generally available. However, in some cases of high-order systems one is able to determine approximately the transient response parameters.

A particularly important is the case in which an asymptotically stable system has a pair of complex conjugate poles (eigenvalues) much closer to the imaginary axis than the remaining poles. This situation is represented in Figure 6.6.

The system poles far to the left of the imaginary axis have large negative real parts so that they decay very quickly to zero (as a matter of fact, they decay exponentially with $e^{\sigma_i t}$, where σ_i are negative real parts of the corresponding poles). Thus, the system response is dominated by the pair of complex conjugate poles closest to the imaginary axis. These poles are called the *dominant system poles*.

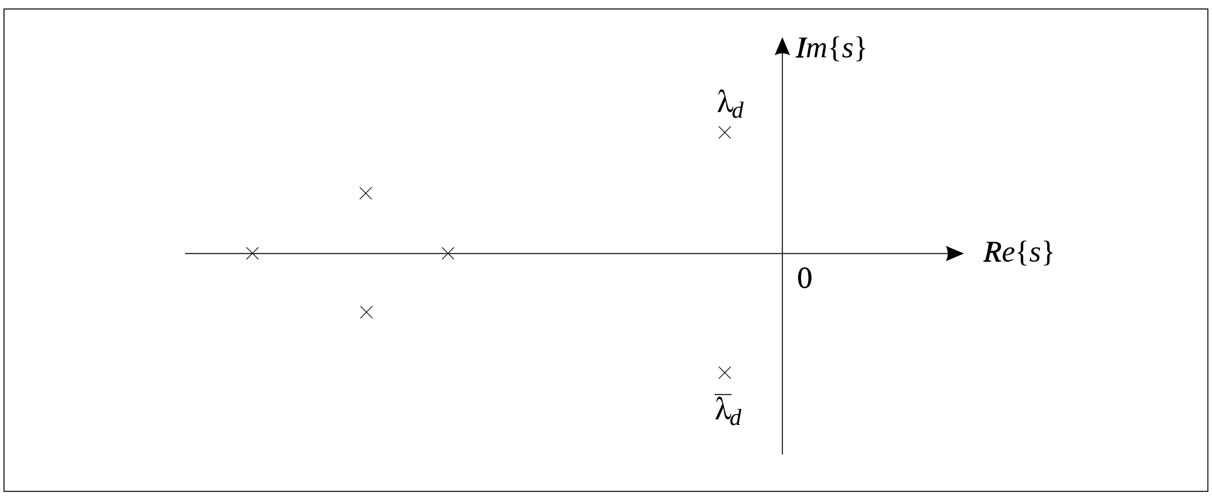

Figure 6.6: Complex conjugate dominant system poles

This analysis can be also justified by using the closed-loop system transfer function. Consider, for example, a system described by its transfer function as

$$M(s) = \frac{Y(s)}{U(s)} = \frac{12600(s+1)}{(s+3)(s+10)(s+60)(s+70)}$$

Since the poles at –60 and –70 are far to the left, their contribution to the system response is negligible (they decay very quickly to zero as e^{-60t} and e^{-70t}). The transfer function can be formally simplified as follows

$$M(s) = \frac{12600(s+1)}{(s+3)(s+10)60\left(\frac{s}{60}+1\right)70\left(\frac{s}{70}+1\right)}$$
$$\approx \frac{3(s+1)}{(s+3)(s+10)} = M_r(s)$$

Example 6.2: In this example we use MATLAB to compare the step responses of the original and reduced-order systems. The results obtained for $y(t)$ and $y_r(t)$ are given in Figure 6.7. It can be seen from this figure that step responses for the original and reduced-order (approximate) systems almost overlap.

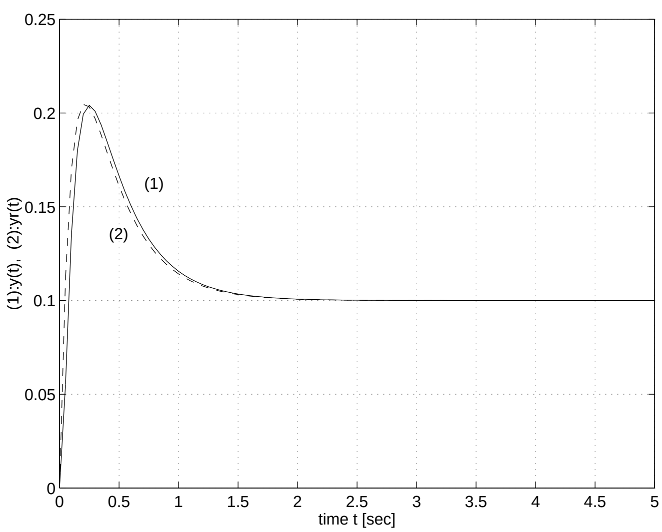

Figure 6.7: System step responses for the original (1) and reduced-order approximate (2) systems

The corresponding MATLAB program is:

```
z=-1; p=[-3 –10 –60 –70]; k=12600;
[num,den]=zp2tf(z,p,k);
t=0:0.05:5; [y,x]=step(num,den,t);
zr=-1; pr=[-3 –10]; kr=3;
[numr,denr]=zp2tf(zr,pr,kr);
[yr,xr]=step(numr,denr,t);
plot(t,y,t,yr,’- –’);
xlabel(’time t [sec]’);
ylabel(’(1):y(t), (2):yr(t)’);
grid; text(0.71,0.16,’(1)’);
text(0.41,0.13,’(2)’);
```

6.4 Steady State Errors

The response of an asymptotically stable linear system is in the long run determined by its steady state component. During the initial time interval the transient response decays to zero so that in the remaining part of the time interval the system response is represented by its steady state component only. Control engineers are interested in having steady state responses as close as possible to the desired ones so that we define the so-called steady state errors, which represent the differences at steady state of the actual and desired system responses (outputs).

Before we proceed to steady state error analysis, we introduce a simplified version of the basic linear control system problem defined in Section 1.1.

Simplified Basic Linear Control Problem

As defined in Section 1.1 the basic linear control problem is still very difficult to solve. A simplified version of this problem can be formulated as follows. Apply to the system input a time function equal to the desired system output. This time function is known as the system's *reference input* and is denoted by $r(t)$. Note that $r(t) = u(t)$. Compare the actual and desired outputs by feeding back the actual output variable. The difference $y(t) - r(t) = e(t)$ represents the error signal. Use the error signal together with simple controllers (if necessary) to drive the system under consideration such that $e(t)$ is reduced as much as possible, at least at steady state. If a simple controller is used in the feedback loop (Figure 6.11) the error signal has to be slightly redefined.

In the following we use this simplified basic linear control problem in order to identify the structure of controllers (feedback elements) that for certain types of reference inputs (desired outputs) produce zero steady state errors.

Consider the simplest feedback configuration of a single-input single-output system given in Figure 6.11.

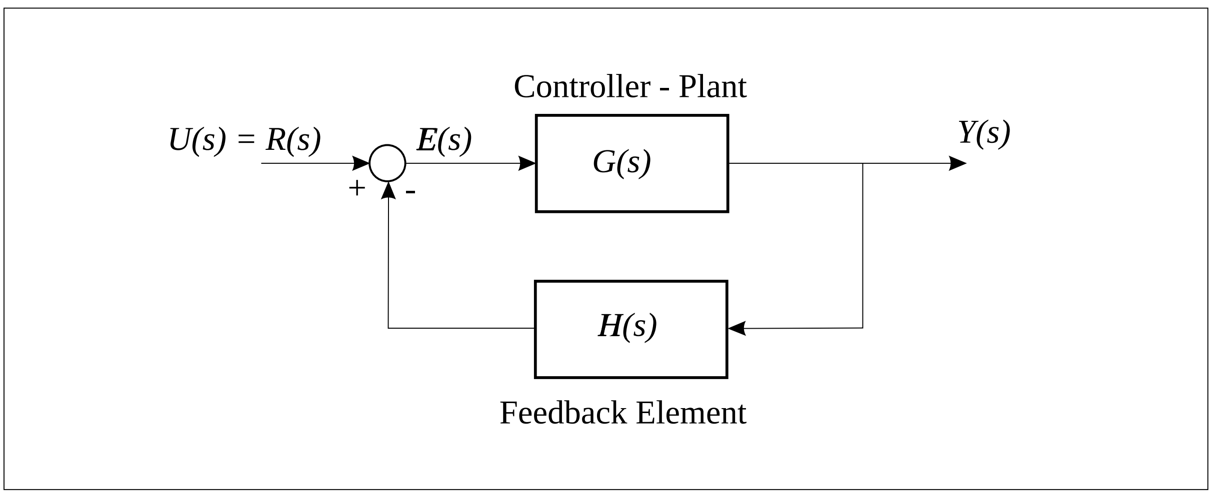

Figure 6.11: Feedback system and steady state errors

Let the input signal $U(s) = R(s)$ represent the Laplace transform of the desired output (in this feedback configuration the desired output signal is used as an input signal); then for $H(s) = 1$, we see that in Figure 6.11 the quantity $E(s)$ represents the difference between the desired output $R(s) = U(s)$ and the actual output $Y(s)$. In order to be able to reduce this error as much as possible, we allow dynamic elements in the feedback loop. Thus, $H(s)$ as a function of s has to be chosen such that for the given type of reference input, the error, now defined by

$$E(s) = R(s) - H(s)Y(s)$$

is eliminated or reduced to its minimal value at steady state.

From the block diagram given in Figure 6.11 we have

$$E(s) = R(s) - H(s)G(s)E(s)$$

so that the expression for the error is given by

$$E(s) = \frac{R(s)}{1 + H(s)G(s)}$$

The steady state error component can be obtained by using the final value theorem of the Laplace transform as

$$e_{ss} = \lim_{t \to \infty} e(t) = \lim_{s \to 0} \{sE(s)\} = \lim_{s \to 0} \left\{ \frac{sR(s)}{1 + H(s)G(s)} \right\}$$

This expression will be used in order to determine the nature of the feedback element $H(s)$ such that the steady state error is reduced to zero for different types of desired outputs. We will particularly consider step, ramp, and parabolic functions as desired system outputs.

Before we proceed to the actual steady state error analysis, we introduce one additional definition.

Definition 6.1 *The type of feedback control system* is determined by the number of poles of the open-loop feedback system transfer function located at the origin, i.e. it is equal to j, where j is obtained from

$$G(s)H(s) = \frac{K(s+z_1)\cdots(s+z_m)}{s^j(s+p_1)(s+p_2)\cdots\left(s+p_{n-j}\right)}$$

Now we consider the steady state errors for different desired outputs, namely unit step, unit ramp, and unit parabolic outputs.

Unit Step Function as Desired Output

Assuming that our goal is that the system output follows as close as possible the unit step function, i.e. $U(s) = R(s) = 1/s$, we have

$$e_{ss} = \lim_{s \to 0} \left\{ \frac{s}{1 + H(s)G(s)} \frac{1}{s} \right\}$$
$$= \frac{1}{1 + \lim_{s \to 0} \{H(s)G(s)\}} = \frac{1}{1 + K_p}$$

where K_p is known as the *position constant*

$$K_p = \lim_{s \to 0} \{H(s)G(s)\}$$

It follows that the steady state error for the unit step reference is reduced to zero for $K_p = \infty$. We see that this condition is satisfied for $j \geq 1$.

Thus, we can conclude that the feedback type system of order at least one allows the system output at steady state to track the unit step function perfectly.

Unit Ramp Function as Desired Output

In this case the steady state error is obtained as

$$e_{ss} = \lim_{s \to 0} \{sE(s)\} = \lim_{s \to 0} \left\{ \frac{s}{1 + H(s)G(s)} \frac{1}{s^2} \right\}$$
$$= \frac{1}{\lim\limits_{s \to 0} \{sH(s)G(s)\}} = \frac{1}{K_v}$$

where

$$K_v = \lim_{s \to 0} \{sH(s)G(s)\}$$

is known as the *velocity constant*. It follows that $K_v = \infty$, i.e. $e_{ss} = 0$ for $j \geq 2$. Thus, systems having two and more pure integrators ($1/s$ terms) in the feedback loop will be able to perfectly track the unit ramp function as a desired system output.

Unit Parabolic Function as Desired Output

For a unit parabolic function we have $R(s) = 2/s^3$ so that

$$e_{ss} = \lim_{s \to 0} \left\{ \frac{s}{1 + H(s)G(s)} \frac{2}{s^3} \right\} = \frac{2}{\lim_{s \to 0} \{s^2 H(s)G(s)\}} = \frac{2}{K_a}$$

where the so-called *acceleration constant*, K_a, is defined by

$$K_a = \lim_{s \to 0} \left\{ s^2 H(s)G(s) \right\}$$

We can conclude that $K_a = \infty$ for $j \geq 3$, i.e. the feedback loop must have three pure integrators in order to reduce the corresponding steady state error to zero.

Example 6.5: The steady state errors for a system that has the open-loop transfer function as

$$H(s)G(s) = \frac{20(s+1)}{s(s+2)(s+5)}$$

are

$$K_p = \infty \;\Rightarrow\; e_{ss} = 0 \quad \text{(step)}$$

$$K_v = 2 \;\Rightarrow\; e_{ss} = 0.5 \quad \text{(ramp)}$$

$$K_a = 0 \;\Rightarrow\; e_{ss} = \infty \quad \text{(parabolic)}$$

Since the open-loop transfer function of this system has one integrator the output of the closed-loop system can perfectly track only the unit step.

Example 6.6: Consider the second-order system whose open-loop transfer function is given by

$$\boldsymbol{H(s)G(s) = \frac{(s+3)}{(s+1)(s+2)}}$$

The position constant for this system is $\boldsymbol{K_p = 1.5}$ so that the corresponding steady state error is

$$\boldsymbol{e_{ss} = \frac{1}{1+K_p} = \frac{1}{1+1.5} = 0.4}$$

The unit step response of this system is presented in Figure 6.12, from which it can be clearly seen that the steady state output is equal to 0.6; hence the steady state error is equal to $1 - 0.6 = 0.4$.

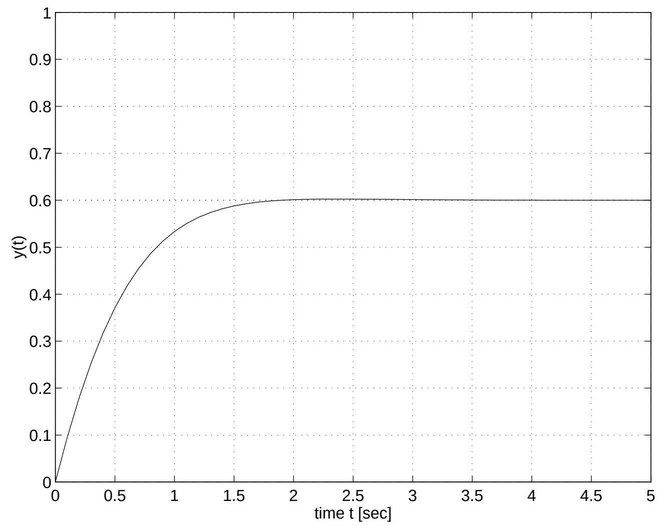

Figure 6.12: System step response for Example 6.6