

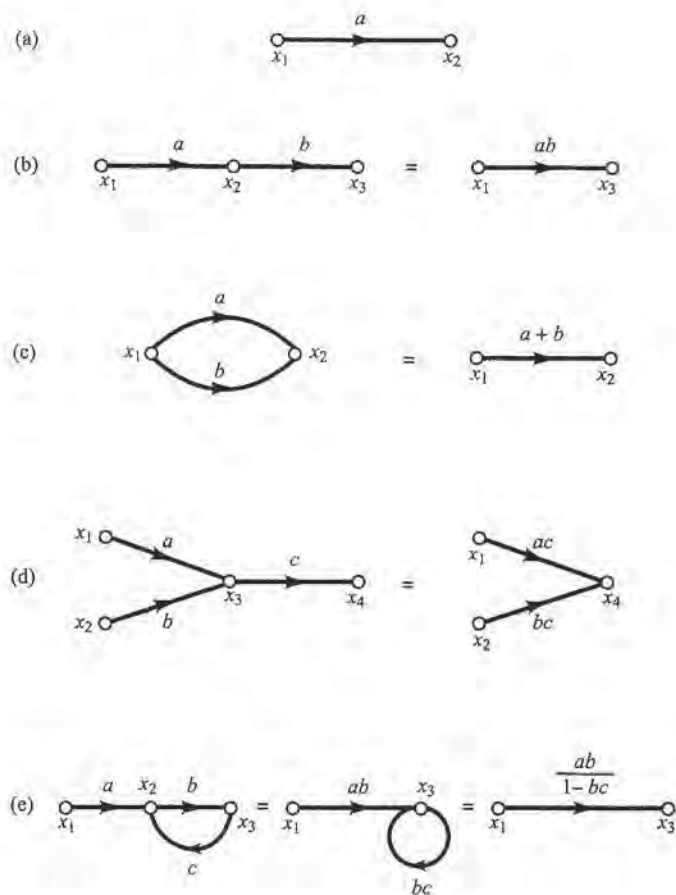


Figure 3-36
Signal flow graphs
and simplifications.

Hence

$$x_3 = abx_1 + bcx_3 \quad (3-77)$$

or

$$x_3 = \frac{ab}{1-bc} x_1 \quad (3-78)$$

Equation (3-77) corresponds to a diagram having a self-loop of transmittance bc . Elimination of the self-loop yields Equation (3-78), which clearly shows that the overall transmittance is $ab/(1-bc)$.

Signal Flow Graph Representation of Linear Systems. Signal flow graphs are widely applied to linear-system analysis. Here the graph can be drawn from the system equations or, with practice, can be drawn by inspection of the physical system. Routine reduction by use of the foregoing rules gives the relation between an input and output variable.

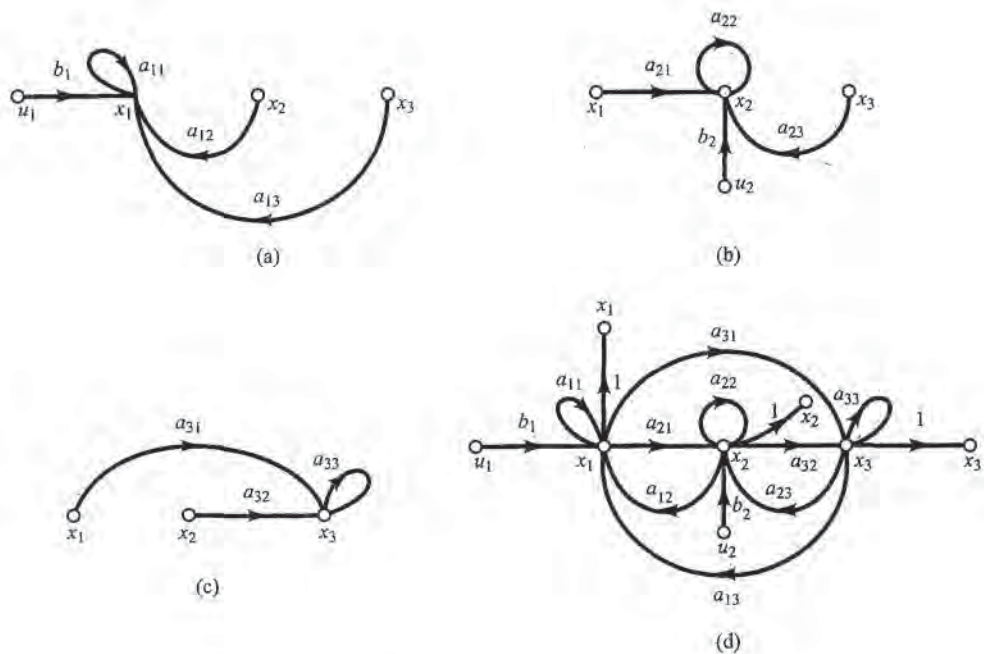


Figure 3-37
Signal flow graphs
representing
(a) Equation (3-79),
(b) Equation (3-80),
and
(c) Equation (3-81);
(d) complete signal
flow graph for the
system described by
Equations
(3-79)-(3-81).

Consider a system defined by the following set of equations:

$$x_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1u_1 \tag{3-79}$$

$$x_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2u_2 \tag{3-80}$$

$$x_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \tag{3-81}$$

where u_1 and u_2 are input variables and x_1 , x_2 , and x_3 are output variables. A signal flow graph for this system, a graphical representation of these three simultaneous equations, indicating the interdependence of the variables, can be obtained as follows: First locate the nodes x_1 , x_2 , and x_3 as shown in Figure 3-37(a). Note that a_{ij} is the transmittance between x_j and x_i . Equation (3-79) states that x_1 is equal to the sum of the four signals $a_{11}x_1$, $a_{12}x_2$, $a_{13}x_3$, and b_1u_1 . The signal flow graph representing Equation (3-79) is shown in Figure 3-37(a). Equation (3-80) states that x_2 is equal to the sum of $a_{21}x_1$, $a_{22}x_2$, $a_{23}x_3$, and b_2u_2 . The corresponding signal flow graph is shown in Figure 3-37(b). The signal flow graph representing Equation (3-81) is shown in Figure 3-37(c).

The signal flow graph representing Equations (3-79), (3-80), and (3-81) is then obtained by combining Figures 3-37(a), (b), and (c). Finally, the complete signal flow graph for the given simultaneous equations is shown in Figure 3-37(d).

In dealing with a signal flow graph, the input nodes (sources) may be considered one at a time. The output signal is then equal to the sum of the individual contributions of each input.

The overall gain from an input to an output may be obtained directly from the signal flow graph by inspection, by use of Mason's formula, or by a reduction of the graph to a simpler form.

Signal Flow Graphs of Control Systems. Some signal flow graphs of simple control systems are shown in Figure 3-38. For such simple graphs, the closed-loop transfer function $C(s)/R(s)$ [or $C(s)/N(s)$] can be obtained easily by inspection. For more complicated signal flow graphs, Mason's gain formula is quite useful.

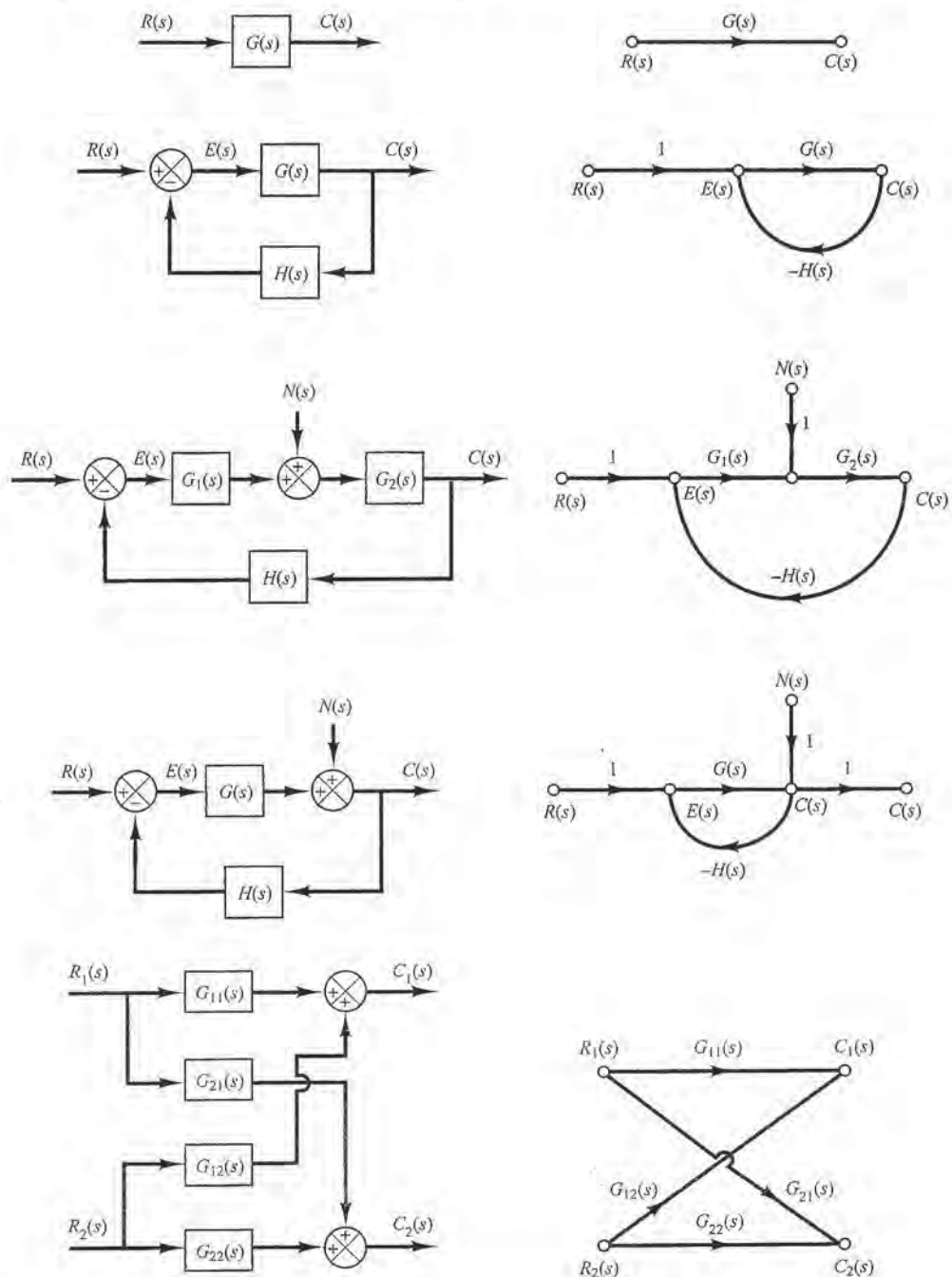


Figure 3-38
Block diagrams and
corresponding signal
flow graphs.

Mason's Gain Formula for Signal Flow Graphs. In many practical cases, we wish to determine the relationship between an input variable and an output variable of the signal flow graph. The transmittance between an input node and an output node is the overall gain, or overall transmittance, between these two nodes.

Mason's gain formula, which is applicable to the overall gain, is given by

$$P = \frac{1}{\Delta} \sum_k P_k \Delta_k$$

where

P_k = path gain or transmittance of k th forward path

Δ = determinant of graph

= $1 - (\text{sum of all individual loop gains}) + (\text{sum of gain products of all possible combinations of two nontouching loops}) - (\text{sum of gain products of all possible combinations of three nontouching loops}) + \dots$

$$= 1 - \sum_a L_a + \sum_{b,c} L_b L_c - \sum_{d,e,f} L_d L_e L_f + \dots$$

$\sum_a L_a$ = sum of all individual loop gains

$\sum_{b,c} L_b L_c$ = sum of gain products of all possible combinations of two nontouching loops

$\sum_{d,e,f} L_d L_e L_f$ = sum of gain products of all possible combinations of three nontouching loops

Δ_k = cofactor of the k th forward path determinant of the graph with the loops touching the k th forward path removed, that is, the cofactor Δ_k is obtained from Δ by removing the loops that touch path P_k

(Note that the summations are taken over all possible paths from input to output.)

In the following, we shall illustrate the use of Mason's gain formula by means of two examples.

EXAMPLE 3-13 Consider the system shown in Figure 3-39. A signal flow graph for this system is shown in Figure 3-40. Let us obtain the closed-loop transfer function $C(s)/R(s)$ by use of Mason's gain formula.

In this system there is only one forward path between the input $R(s)$ and the output $C(s)$. The forward path gain is

$$P_1 = G_1 G_2 G_3$$

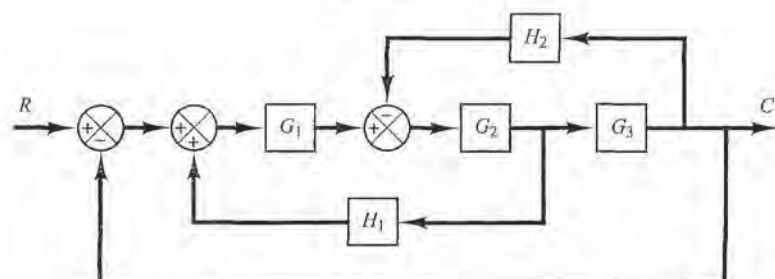
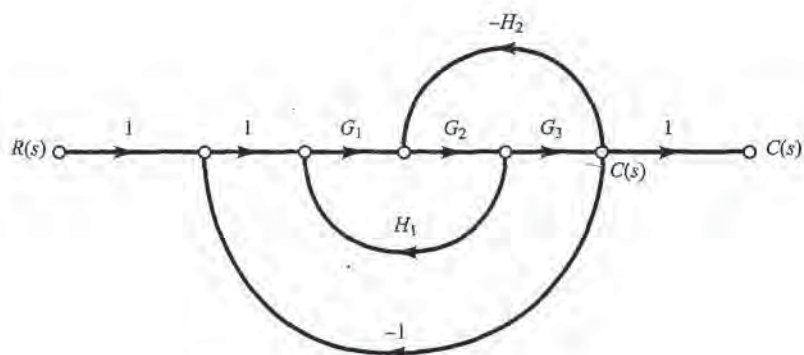


Figure 3-39
Multiple-loop
system.

Figure 3-40
Signal flow graph
for the system in
Figure 3-39.



From Figure 3-40, we see that there are three individual loops. The gains of these loops are

$$\begin{aligned} L_1 &= G_1 G_2 H_1 \\ L_2 &= -G_2 G_3 H_2 \\ L_3 &= -G_1 G_2 G_3 \end{aligned}$$

Note that since all three loops have a common branch, there are no nontouching loops. Hence, the determinant Δ is given by

$$\begin{aligned} \Delta &= 1 - (L_1 + L_2 + L_3) \\ &= 1 - G_1 G_2 H_1 + G_2 G_3 H_2 + G_1 G_2 G_3 \end{aligned}$$

The cofactor Δ_1 of the determinant along the forward path connecting the input node and output node is obtained from Δ by removing the loops that touch this path. Since path P_1 touches all three loops, we obtain

$$\Delta_1 = 1$$

Therefore, the overall gain between the input $R(s)$ and the output $C(s)$, or the closed-loop transfer function, is given by

$$\begin{aligned} \frac{C(s)}{R(s)} &= P = \frac{P_1 \Delta_1}{\Delta} \\ &= \frac{G_1 G_2 G_3}{1 - G_1 G_2 H_1 + G_2 G_3 H_2 + G_1 G_2 G_3} \end{aligned}$$

which is the same as the closed-loop transfer function obtained by block diagram reduction. Mason's gain formula thus gives the overall gain $C(s)/R(s)$ without a reduction of the graph.

EXAMPLE 3-14 Consider the system shown in Figure 3-41. Obtain the closed-loop transfer function $C(s)/R(s)$ by use of Mason's gain formula.

In this system, there are three forward paths between the input $R(s)$ and the output $C(s)$. The forward path gains are

$$\begin{aligned} P_1 &= G_1 G_2 G_3 G_4 G_5 \\ P_2 &= G_1 G_6 G_4 G_5 \\ P_3 &= G_1 G_2 G_7 \end{aligned}$$

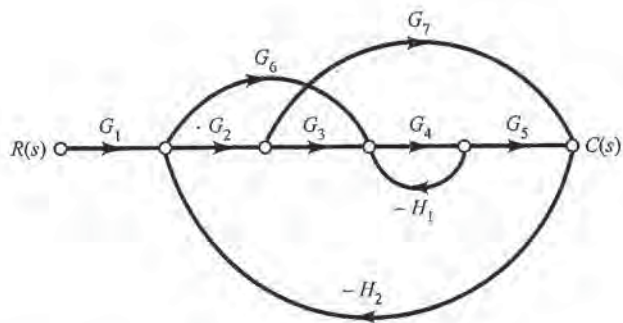


Figure 3-41
Signal flow graph for
a system.

There are four individual loops. The gains of these loops are

$$L_1 = -G_4H_1$$

$$L_2 = -G_2G_7H_2$$

$$L_3 = -G_6G_4G_5H_2$$

$$L_4 = -G_2G_3G_4G_5H_2$$

Loop L_1 does not touch loop L_2 . Hence, the determinant Δ is given by

$$\Delta = 1 - (L_1 + L_2 + L_3 + L_4) + L_1L_2 \quad (3-82)$$

The cofactor Δ_1 , is obtained from Δ by removing the loops that touch path P_1 . Therefore, by removing L_1, L_2, L_3, L_4 , and L_1L_2 from Equation (3-82), we obtain

$$\Delta_1 = 1$$

Similarly, the cofactor Δ_2 is

$$\Delta_2 = 1$$

The cofactor Δ_3 is obtained by removing L_2, L_3, L_4 , and L_1L_2 from Equation (3-82), giving

$$\Delta_3 = 1 - L_1$$

The closed-loop transfer function $C(s)/R(s)$ is then

$$\begin{aligned} \frac{C(s)}{R(s)} &= P = \frac{1}{\Delta} (P_1\Delta_1 + P_2\Delta_2 + P_3\Delta_3) \\ &= \frac{G_1G_2G_3G_4G_5 + G_1G_6G_4G_5 + G_1G_2G_7(1 + G_4H_1)}{1 + G_4H_1 + G_2G_7H_2 + G_6G_4G_5H_2 + G_2G_3G_4G_5H_2 + G_4H_1G_2G_7H_2} \end{aligned}$$

Comments. The usual application of signal flow graphs is in system diagramming. The set of equations describing a linear system is represented by a signal flow graph by establishing nodes that represent the system variables and by interconnecting the nodes with weighted, directed, transmittances, which represent the relationships among the variables. Mason's gain formula may be used to establish the relationship between an input and an output. (Alternatively, the variables in the system may be eliminated one by one with reduction techniques.) Mason's gain formula is especially useful in reducing large and complex system diagrams in one step, without requiring step-by-step reductions.

Finally, it is noted that in applying the Mason's gain formula to a given system, one must be careful not to make mistakes in calculating the cofactors of the forward paths, Δ_k , since any errors, if they exist, may not easily be detected.

3-10 LINEARIZATION OF NONLINEAR MATHEMATICAL MODELS

Nonlinear Systems. A system is nonlinear if the principle of superposition does not apply. Thus, for a nonlinear system the response to two inputs cannot be calculated by treating one input at a time and adding the results.

Although many physical relationships are often represented by linear equations, in most cases actual relationships are not quite linear. In fact, a careful study of physical systems reveals that even so-called "linear systems" are really linear only in limited operating ranges. In practice, many electromechanical systems, hydraulic systems, pneumatic systems, and so on, involve nonlinear relationships among the variables. For example, the output of a component may saturate for large input signals. There may be a dead space that affects small signals. (The dead space of a component is a small range of input variations to which the component is insensitive.) Square-law nonlinearity may occur in some components. For instance, dampers used in physical systems may be linear for low-velocity operations but may become nonlinear at high velocities, and the damping force may become proportional to the square of the operating velocity.

Linearization of Nonlinear Systems. In control engineering a normal operation of the system may be around an equilibrium point, and the signals may be considered small signals around the equilibrium. (It should be pointed out that there are many exceptions to such a case.) However, if the system operates around an equilibrium point and if the signals involved are small signals, then it is possible to approximate the nonlinear system by a linear system. Such a linear system is equivalent to the nonlinear system considered within a limited operating range. Such a linearized model (linear, time-invariant model) is very important in control engineering.

The linearization procedure to be presented in the following is based on the expansion of nonlinear function into a Taylor series about the operating point and the retention of only the linear term. Because we neglect higher-order terms of Taylor series expansion, these neglected terms must be small enough; that is, the variables deviate only slightly from the operating condition.

Linear Approximation of Nonlinear Mathematical Models. To obtain a linear mathematical model for a nonlinear system, we assume that the variables deviate only slightly from some operating condition. Consider a system whose input is $x(t)$ and output is $y(t)$. The relationship between $y(t)$ and $x(t)$ is given by

$$y = f(x) \quad (3-83)$$

If the normal operating condition corresponds to $\bar{x}$, $\bar{y}$, then Equation (3-83) may be expanded into a Taylor series about this point as follows:

$$\begin{aligned}
y &= f(x) \\
&= f(\bar{x}) + \frac{df}{dx}(x - \bar{x}) + \frac{1}{2!} \frac{d^2f}{dx^2}(x - \bar{x})^2 + \dots
\end{aligned} \tag{3-84}$$

where the derivatives $df/dx, d^2f/dx^2, \dots$ are evaluated at $x = \bar{x}$. If the variation $x - \bar{x}$ is small, we may neglect the higher-order terms in $x - \bar{x}$. Then Equation (3-84) may be written as

$$y = \bar{y} + K(x - \bar{x}) \tag{3-85}$$

where

$$\begin{aligned}
\bar{y} &= f(\bar{x}) \\
K &= \left. \frac{df}{dx} \right|_{x=\bar{x}}
\end{aligned}$$

Equation (3-85) may be rewritten as

$$y - \bar{y} = K(x - \bar{x}) \tag{3-86}$$

which indicates that $y - \bar{y}$ is proportional to $x - \bar{x}$. Equation (3-86) gives a linear mathematical model for the nonlinear system given by Equation (3-83) near the operating point $x = \bar{x}, y = \bar{y}$.

Next, consider a nonlinear system whose output y is a function of two inputs x_1 and x_2 , so that

$$y = f(x_1, x_2) \tag{3-87}$$

To obtain a linear approximation to this nonlinear system, we may expand Equation (3-87) into a Taylor series about the normal operating point $\bar{x}_1, \bar{x}_2$. Then Equation (3-87) becomes

$$\begin{aligned}
y &= f(\bar{x}_1, \bar{x}_2) + \left[\frac{\partial f}{\partial x_1}(x_1 - \bar{x}_1) + \frac{\partial f}{\partial x_2}(x_2 - \bar{x}_2) \right] \\
&\quad + \frac{1}{2!} \left[\frac{\partial^2 f}{\partial x_1^2}(x_1 - \bar{x}_1)^2 + 2 \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) \right. \\
&\quad \left. + \frac{\partial^2 f}{\partial x_2^2}(x_2 - \bar{x}_2)^2 \right] + \dots
\end{aligned}$$

where the partial derivatives are evaluated at $x_1 = \bar{x}_1, x_2 = \bar{x}_2$. Near the normal operating point, the higher-order terms may be neglected. The linear mathematical model of this nonlinear system in the neighborhood of the normal operating condition is then given by

$$y - \bar{y} = K_1(x_1 - \bar{x}_1) + K_2(x_2 - \bar{x}_2)$$

where

$$\bar{y} = f(\bar{x}_1, \bar{x}_2)$$

$$K_1 = \left. \frac{\partial f}{\partial x_1} \right|_{x_1=\bar{x}_1, x_2=\bar{x}_2}$$

$$K_2 = \left. \frac{\partial f}{\partial x_2} \right|_{x_1=\bar{x}_1, x_2=\bar{x}_2}$$

The linearization technique presented here is valid in the vicinity of the operating condition. If the operating conditions vary widely, however, such linearized equations are not adequate, and nonlinear equations must be dealt with. It is important to remember that a particular mathematical model used in analysis and design may accurately represent the dynamics of an actual system for certain operating conditions, but may not be accurate for other operating conditions.

EXAMPLE 3-15 Linearize the nonlinear equation

$$z = xy$$

in the region $5 \leq x \leq 7$, $10 \leq y \leq 12$. Find the error if the linearized equation is used to calculate the value of z when $x = 5$, $y = 10$.

Since the region considered is given by $5 \leq x \leq 7$, $10 \leq y \leq 12$, choose $\bar{x} = 6$, $\bar{y} = 11$. Then $\bar{z} = \bar{x}\bar{y} = 66$. Let us obtain a linearized equation for the nonlinear equation near a point $\bar{x} = 6$, $\bar{y} = 11$.

Expanding the nonlinear equation into a Taylor series about point $x = \bar{x}$, $y = \bar{y}$ and neglecting the higher-order terms, we have

$$z - \bar{z} = a(x - \bar{x}) + b(y - \bar{y})$$

where

$$a = \left. \frac{\partial(xy)}{\partial x} \right|_{x=\bar{x}, y=\bar{y}} = \bar{y} = 11$$

$$b = \left. \frac{\partial(xy)}{\partial y} \right|_{x=\bar{x}, y=\bar{y}} = \bar{x} = 6$$

Hence the linearized equation is

$$z - 66 = 11(x - 6) + 6(y - 11)$$

or

$$z = 11x + 6y - 66$$

When $x = 5$, $y = 10$, the value of z given by the linearized equation is

$$z = 11x + 6y - 66 = 55 + 60 - 66 = 49$$

The exact value of z is $z = xy = 50$. The error is thus $50 - 49 = 1$. In terms of percentage, the error is 2%.

EXAMPLE PROBLEMS AND SOLUTIONS

A-3-1. Simplify the block diagram shown in Figure 3-42.

Solution. First, move the branch point of the path involving H_1 outside the loop involving H_2 , as shown in Figure 3-43(a). Then eliminating two loops results in Figure 3-43(b). Combining two blocks into one gives Figure 3-43(c).

A-3-2. Simplify the block diagram shown in Figure 3-44. Obtain the transfer function relating $C(s)$ and $R(s)$.

Figure 3-42
Block diagram of a system.

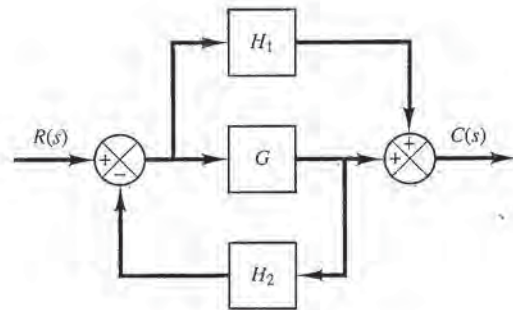


Figure 3-43
Simplified block diagrams for the system shown in Figure 3-42.

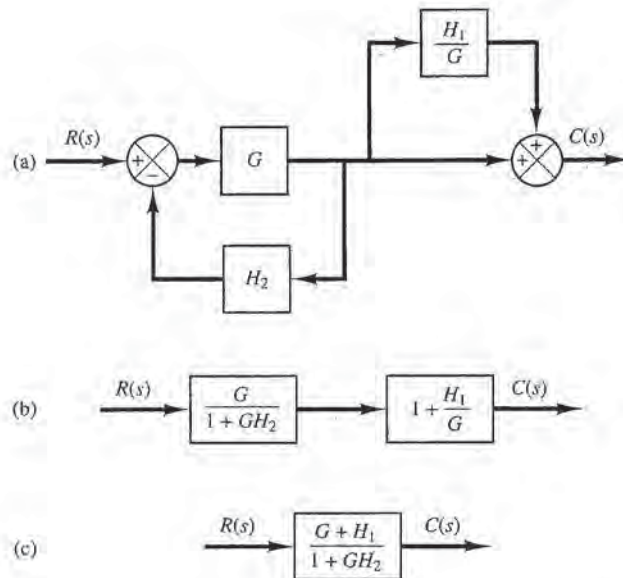
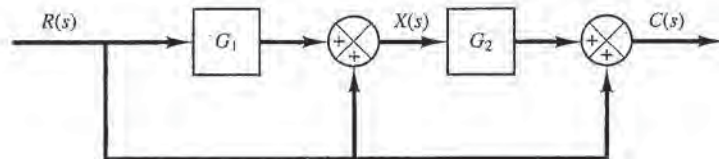


Figure 3-44
Block diagram of a system.



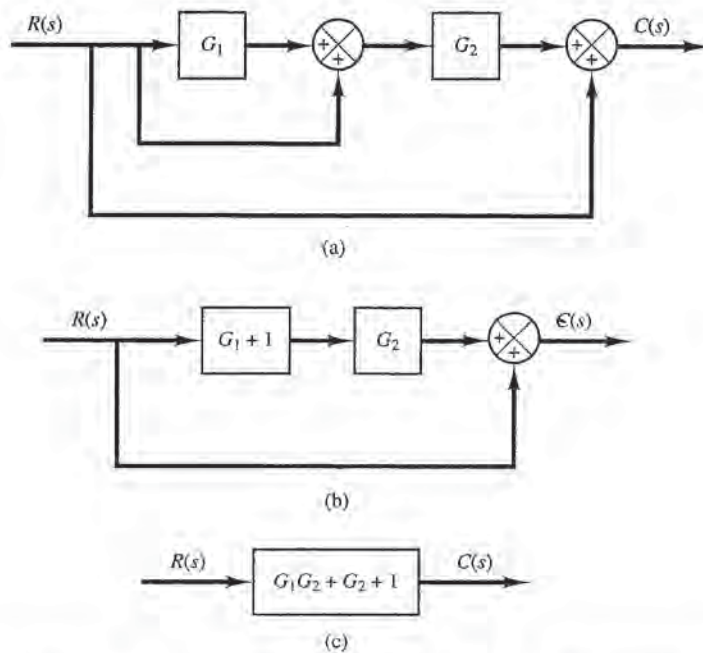


Figure 3-45
Reduction of the
block diagram shown
in Figure 3-44.

Solution. The block diagram of Figure 3-44 can be modified to that shown in Figure 3-45(a). Eliminating the minor feedforward path, we obtain Figure 3-45(b), which can be simplified to that shown in Figure 3-45(c). The transfer function $C(s)/R(s)$ is thus given by

$$\frac{C(s)}{R(s)} = G_1G_2 + G_2 + 1$$

The same result can also be obtained by proceeding as follows: Since signal $X(s)$ is the sum of two signals $G_1R(s)$ and $R(s)$, we have

$$X(s) = G_1R(s) + R(s)$$

The output signal $C(s)$ is the sum of $G_2X(s)$ and $R(s)$. Hence

$$C(s) = G_2X(s) + R(s) = G_2[G_1R(s) + R(s)] + R(s)$$

And so we have the same result as before:

$$\frac{C(s)}{R(s)} = G_1G_2 + G_2 + 1$$

- A-3-3.** Simplify the block diagram shown in Figure 3-46. Then, obtain the closed-loop transfer function $C(s)/R(s)$.

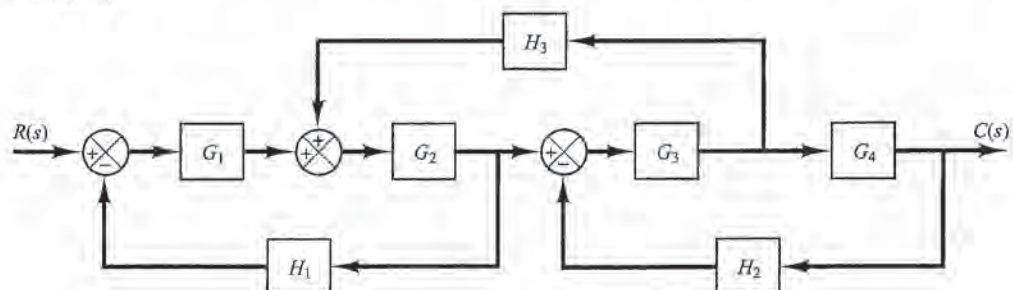


Figure 3-46
Block diagram of a
system.

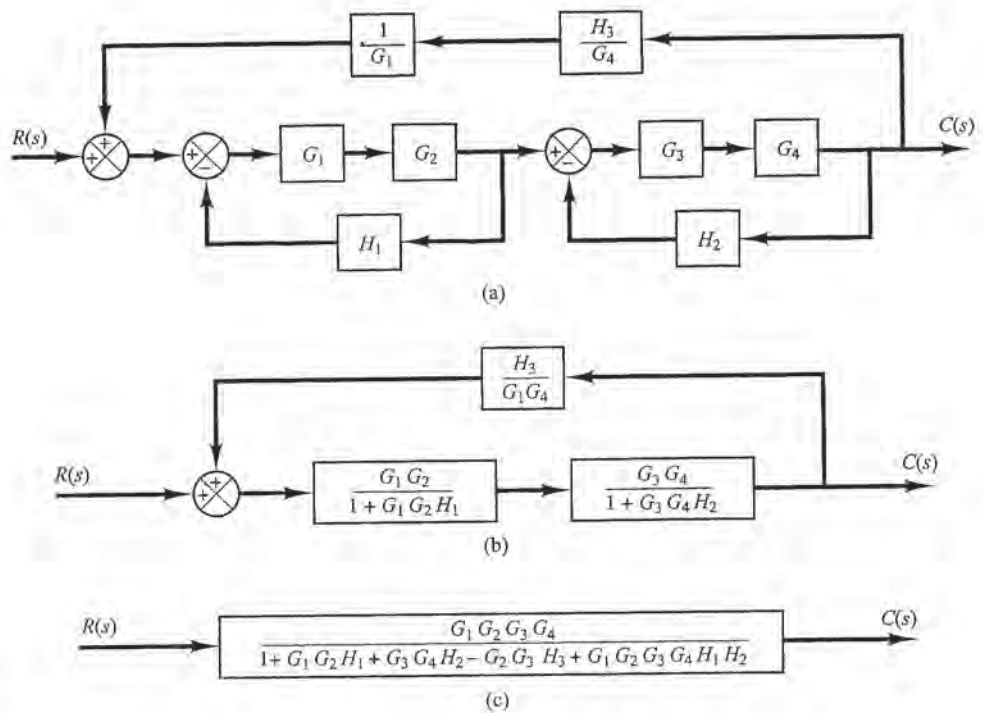


Figure 3-47
Successive
reductions of the
block diagram shown
in Figure 3-46.

Solution. First move the branch point between G_3 and G_4 to the right-hand side of the loop containing G_3 , G_4 , and H_2 . Then move the summing point between G_1 and G_2 to the left-hand side of the first summing point. See Figure 3-47(a). By simplifying each loop, the block diagram can be modified as shown in Figure 3-47(b). Further simplification results in Figure 3-47(c), from which the closed-loop transfer function $C(s)/R(s)$ is obtained as

$$\frac{C(s)}{R(s)} = \frac{G_1 G_2 G_3 G_4}{1 + G_1 G_2 H_1 + G_3 G_4 H_2 - G_2 G_3 H_3 + G_1 G_2 G_3 G_4 H_1 H_2}$$

A-3-4. Obtain transfer functions $C(s)/R(s)$ and $C(s)/D(s)$ of the system shown in Figure 3-48.

Solution. From Figure 3-48 we have

$$U(s) = G_f R(s) + G_c E(s) \quad (3-88)$$

$$C(s) = G_p [D(s) + G_1 U(s)] \quad (3-89)$$

$$E(s) = R(s) - H C(s) \quad (3-90)$$

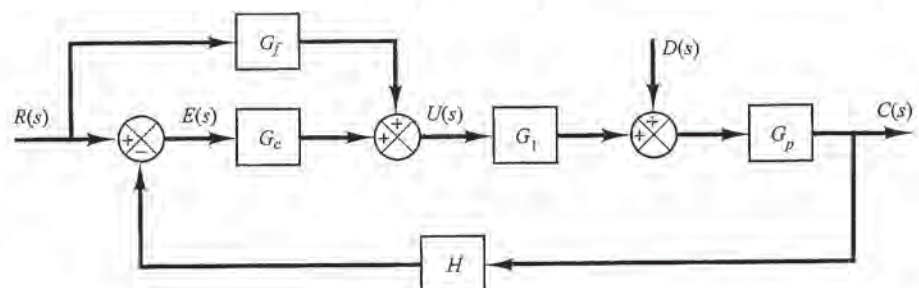


Figure 3-48
Control system with
reference input and
disturbance input.

By substituting Equation (3-88) into Equation (3-89), we get

$$C(s) = G_p D(s) + G_1 G_p [G_f R(s) + G_c E(s)] \quad (3-91)$$

By substituting Equation (3-90) into Equation (3-91), we obtain

$$C(s) = G_p D(s) + G_1 G_p \{G_f R(s) + G_c [R(s) - HC(s)]\}$$

Solving this last equation for $C(s)$, we get

$$C(s) + G_1 G_p G_c H C(s) = G_p D(s) + G_1 G_p (G_f + G_c) R(s)$$

Hence

$$C(s) = \frac{G_p D(s) + G_1 G_p (G_f + G_c) R(s)}{1 + G_1 G_p G_c H} \quad (3-92)$$

Note that Equation (3-92) gives the response $C(s)$ when both reference input $R(s)$ and disturbance input $D(s)$ are present.

To find transfer function $C(s)/R(s)$, we let $D(s) = 0$ in Equation (3-92). Then we obtain

$$\frac{C(s)}{R(s)} = \frac{G_1 G_p (G_f + G_c)}{1 + G_1 G_p G_c H}$$

Similarly, to obtain transfer function $C(s)/D(s)$, we let $R(s) = 0$ in Equation (3-92). Then $C(s)/D(s)$ can be given by

$$\frac{C(s)}{D(s)} = \frac{G_p}{1 + G_1 G_p G_c H}$$

- A-3-5.** Figure 3-49 shows a system with two inputs and two outputs. Derive $C_1(s)/R_1(s)$, $C_1(s)/R_2(s)$, $C_2(s)/R_1(s)$, and $C_2(s)/R_2(s)$. (In deriving outputs for $R_1(s)$, assume that $R_2(s)$ is zero, and vice versa.)

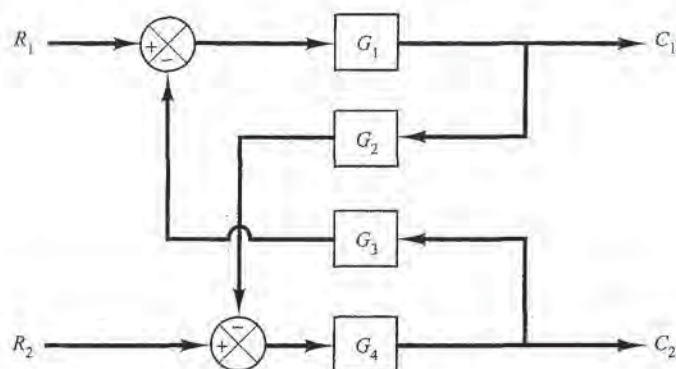


Figure 3-49
System with two
inputs and two
outputs.

Solution. From the figure, we obtain

$$C_1 = G_1(R_1 - G_3C_2) \quad (3-93)$$

$$C_2 = G_4(R_2 - G_2C_1) \quad (3-94)$$

By substituting Equation (3-94) into Equation (3-93), we obtain

$$C_1 = G_1[R_1 - G_3G_4(R_2 - G_2C_1)] \quad (3-95)$$

By substituting Equation (3-93) into Equation (3-94), we get

$$C_2 = G_4[R_2 - G_2G_1(R_1 - G_3C_2)] \quad (3-96)$$

Solving Equation (3-95) for C_1 , we obtain

$$C_1 = \frac{G_1R_1 - G_1G_3G_4R_2}{1 - G_1G_2G_3G_4} \quad (3-97)$$

Solving Equation (3-96) for C_2 gives

$$C_2 = \frac{-G_1G_2G_4R_1 + G_4R_2}{1 - G_1G_2G_3G_4} \quad (3-98)$$

Equations (3-97) and (3-98) can be combined in the form of the transfer matrix as follows:

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} \frac{G_1}{1 - G_1G_2G_3G_4} & -\frac{G_1G_3G_4}{1 - G_1G_2G_3G_4} \\ -\frac{G_1G_2G_4}{1 - G_1G_2G_3G_4} & \frac{G_4}{1 - G_1G_2G_3G_4} \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

Then the transfer functions $C_1(s)/R_1(s)$, $C_1(s)/R_2(s)$, $C_2(s)/R_1(s)$ and $C_2(s)/R_2(s)$ can be obtained as follows:

$$\begin{aligned} \frac{C_1(s)}{R_1(s)} &= \frac{G_1}{1 - G_1G_2G_3G_4}, & \frac{C_1(s)}{R_2(s)} &= -\frac{G_1G_3G_4}{1 - G_1G_2G_3G_4} \\ \frac{C_2(s)}{R_1(s)} &= -\frac{G_1G_2G_4}{1 - G_1G_2G_3G_4}, & \frac{C_2(s)}{R_2(s)} &= \frac{G_4}{1 - G_1G_2G_3G_4} \end{aligned}$$

Note that Equations (3-97) and (3-98) give responses C_1 and C_2 , respectively, when both inputs R_1 and R_2 are present.

Notice that when $R_2(s) = 0$, the original block diagram can be simplified to those shown in Figures 3-50(a) and (b). Similarly, when $R_1(s) = 0$, the original block diagram can be simplified to those shown in Figures 3-50(c) and (d). From these simplified block diagrams we can also obtain $C_1(s)/R_1(s)$, $C_2(s)/R_1(s)$, $C_1(s)/R_2(s)$, and $C_2(s)/R_2(s)$, as shown to the right of each corresponding block diagram.

Example Problems and Solutions

119

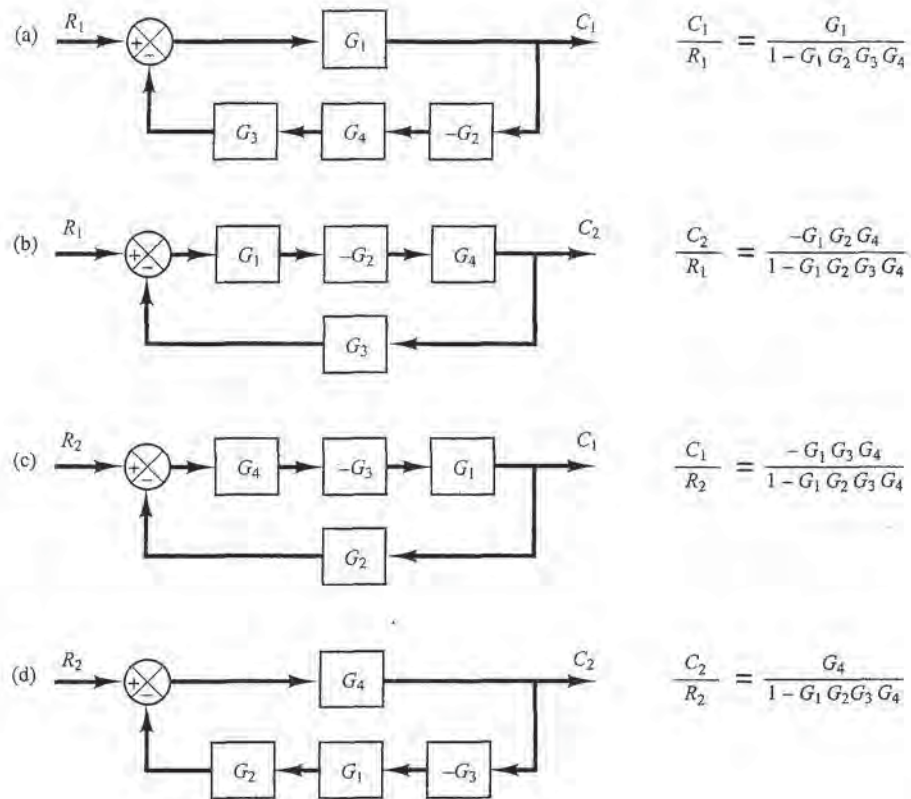


Figure 3-50
Simplified block
diagrams and
corresponding
closed-loop transfer
functions.

A-3-6. Show that for the differential equation system

$$\ddot{y} + a_1 \dot{y} + a_2 y = b_0 \ddot{u} + b_1 \dot{u} + b_2 \dot{u} + b_3 u \quad (3-99)$$

state and output equations can be given, respectively, by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} u \quad (3-100)$$

and

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \beta_0 u \quad (3-101)$$

where state variables are defined by

$$\begin{aligned} x_1 &= y - \beta_0 u \\ x_2 &= \dot{y} - \beta_0 \dot{u} - \beta_1 u = \dot{x}_1 - \beta_1 u \\ x_3 &= \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u = \dot{x}_2 - \beta_2 u \end{aligned}$$

and

$$\begin{aligned}\beta_0 &= b_0 \\ \beta_1 &= b_1 - a_1\beta_0 \\ \beta_2 &= b_2 - a_1\beta_1 - a_2\beta_0 \\ \beta_3 &= b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0\end{aligned}$$

Solution. From the definition of state variables x_2 and x_3 , we have

$$\dot{x}_1 = x_2 + \beta_1 u \quad (3-102)$$

$$\dot{x}_2 = x_3 + \beta_2 u \quad (3-103)$$

To derive the equation for $\dot{x}_3$, we first note from Equation (3-99) that

$$\ddot{y} = -a_1\dot{y} - a_2\dot{y} - a_3y + b_0\ddot{u} + b_1\dot{u} + b_2\dot{u} + b_3u$$

Since

$$x_3 = \ddot{y} - \beta_0\ddot{u} - \beta_1\dot{u} - \beta_2u$$

we have

$$\begin{aligned}\dot{x}_3 &= \ddot{y} - \beta_0\ddot{u} - \beta_1\dot{u} - \beta_2\dot{u} \\ &= (-a_1\dot{y} - a_2\dot{y} - a_3y) + b_0\ddot{u} + b_1\dot{u} + b_2\dot{u} + b_3u - \beta_0\ddot{u} - \beta_1\dot{u} - \beta_2\dot{u} \\ &= -a_1(\dot{y} - \beta_0\ddot{u} - \beta_1\dot{u} - \beta_2u) - a_1\beta_0\ddot{u} - a_1\beta_1\dot{u} - a_1\beta_2u \\ &\quad - a_2(\dot{y} - \beta_0\ddot{u} - \beta_1\dot{u} - \beta_2u) - a_2\beta_0\ddot{u} - a_2\beta_1\dot{u} - a_2\beta_2u \\ &\quad + b_0\ddot{u} + b_1\dot{u} + b_2\dot{u} + b_3u - \beta_0\ddot{u} - \beta_1\dot{u} - \beta_2\dot{u} \\ &= -a_1x_3 - a_2x_2 - a_3x_1 + (b_0 - \beta_0)\ddot{u} + (b_1 - \beta_1 - a_1\beta_0)\dot{u} \\ &\quad + (b_2 - \beta_2 - a_1\beta_1 - a_2\beta_0)\dot{u} + (b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0)u \\ &= -a_1x_3 - a_2x_2 - a_3x_1 + (b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0)u \\ &= -a_1x_3 - a_2x_2 - a_3x_1 + \beta_3u\end{aligned}$$

Hence, we get

$$\dot{x}_3 = -a_3x_1 - a_2x_2 - a_1x_3 + \beta_3u \quad (3-104)$$

Combining Equations (3-102), (3-103), and (3-104) into a vector-matrix equation, we obtain Equation (3-100). Also, from the definition of state variable x_1 , we get the output equation given by Equation (3-101).

A-3-7. Obtain a state-space equation and output equation for the system defined by

$$\frac{Y(s)}{U(s)} = \frac{2s^3 + s^2 + s + 2}{s^3 + 4s^2 + 5s + 2}$$

Solution. From the given transfer function, the differential equation for the system is

$$\ddot{y} + 4\dot{y} + 5\dot{y} + 2y = 2\ddot{u} + \dot{u} + \dot{u} + 2u$$

Comparing this equation with the standard equation given by Equation (3-33), rewritten

$$\ddot{y} + a_1\dot{y} + a_2\dot{y} + a_3y = b_0\ddot{u} + b_1\dot{u} + b_2\dot{u} + b_3u$$

Example Problems and Solutions

121

we find

$$\begin{aligned} a_1 &= 4, & a_2 &= 5, & a_3 &= 2 \\ b_0 &= 2, & b_1 &= 1, & b_2 &= 1, & b_3 &= 2 \end{aligned}$$

Referring to Equation (3-35), we have

$$\begin{aligned} \beta_0 &= b_0 = 2 \\ \beta_1 &= b_1 - a_1\beta_0 = 1 - 4 \times 2 = -7 \\ \beta_2 &= b_2 - a_1\beta_1 - a_2\beta_0 = 1 - 4 \times (-7) - 5 \times 2 = 19 \\ \beta_3 &= b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0 \\ &= 2 - 4 \times 19 - 5 \times (-7) - 2 \times 2 = -43 \end{aligned}$$

Referring to Equation (3-34), we define

$$\begin{aligned} x_1 &= y - \beta_0 u = y - 2u \\ x_2 &= \dot{x}_1 - \beta_1 u = \dot{x}_1 + 7u \\ x_3 &= \dot{x}_2 - \beta_2 u = \dot{x}_2 - 19u \end{aligned}$$

Then referring to Equation (3-36),

$$\begin{aligned} \dot{x}_1 &= x_2 - 7u \\ \dot{x}_2 &= x_3 + 19u \\ \dot{x}_3 &= -a_3 x_1 - a_2 x_2 - a_1 x_3 + \beta_3 u \\ &= -2x_1 - 5x_2 - 4x_3 - 43u \end{aligned}$$

Hence, the state-space representation of the system is

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -7 \\ 19 \\ -43 \end{bmatrix} u \\ y &= [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 2u \end{aligned}$$

This is one possible state-space representation of the system. There are many (infinitely many) others. If we use MATLAB, it produces the following state-space representation:

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} &= \begin{bmatrix} -4 & -5 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u \\ y &= [-7 \quad -9 \quad -2] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 2u \end{aligned}$$

See MATLAB Program 3-4. (Note that all state-space representations for the same system are equivalent.)

MATLAB Program 3-4

```
num = [2 1 1 2];
den = [1 4 5 2];
[A,B,C,D] = tf2ss(num, den)
```

A =

```
-4 -5 -2
 1  0  0
 0  1  0
```

B =

```
1
0
0
```

C =

```
-7 -9 -2
```

D =

```
2
```

A-3-8. Obtain a state-space model of the system shown in Figure 3-51.

Solution. The system involves one integrator and two delayed integrators. The output of each integrator or delayed integrator can be a state variable. Let us define the output of the plant as x_1 , the output of the controller as x_2 , and the output of the sensor as x_3 . Then we obtain

$$\begin{aligned}\frac{X_1(s)}{X_2(s)} &= \frac{10}{s+5} \\ \frac{X_2(s)}{U(s) - X_3(s)} &= \frac{1}{s} \\ \frac{X_3(s)}{X_1(s)} &= \frac{1}{s+1} \\ Y(s) &= X_1(s)\end{aligned}$$

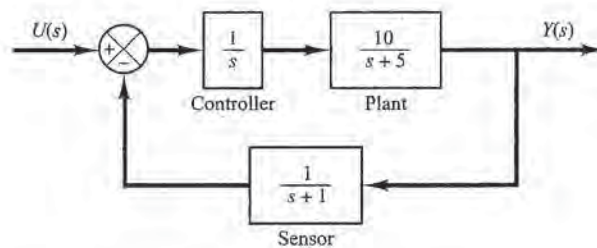


Figure 3-51
Control system.

which can be rewritten as

$$sX_1(s) = -5X_1(s) + 10X_2(s)$$

$$sX_2(s) = -X_3(s) + U(s)$$

$$sX_3(s) = X_1(s) - X_3(s)$$

$$Y(s) = X_1(s)$$

By taking the inverse Laplace transforms of the preceding four equations, we obtain

$$\dot{x}_1 = -5x_1 + 10x_2$$

$$\dot{x}_2 = -x_3 + u$$

$$\dot{x}_3 = x_1 - x_3$$

$$y = x_1$$

Thus, a state-space model of the system in the standard form is given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -5 & 10 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u$$

$$y = [1 \ 0 \ 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

It is important to note that this is not the only state-space representation of the system. Many other state-space representations are possible. However, the number of state variables is the same in any state-space representation of the same system. In the present system, the number of state variables is three, regardless of what variables are chosen as state variables.

A-3-9. Obtain a state-space model for the system shown in Figure 3-52(a).

Solution. First, notice that $(as + b)/s^2$ involves a derivative term. Such a derivative term may be avoided if we modify $(as + b)/s^2$ as

$$\frac{as + b}{s^2} = \left(a + \frac{b}{s}\right) \frac{1}{s}$$

Using this modification, the block diagram of Figure 3-52(a) can be modified to that shown in Figure 3-52(b).

Define the outputs of the integrators as state variables, as shown in Figure 3-52(b). Then from Figure 3-52(b) we obtain

$$\frac{X_1(s)}{X_2(s) + a[U(s) - X_1(s)]} = \frac{1}{s}$$

$$\frac{X_2(s)}{U(s) - X_1(s)} = \frac{b}{s}$$

$$Y(s) = X_1(s)$$

which may be modified to

$$sX_1(s) = X_2(s) + a[U(s) - X_1(s)]$$

$$sX_2(s) = -bX_1(s) + bU(s)$$

$$Y(s) = X_1(s)$$

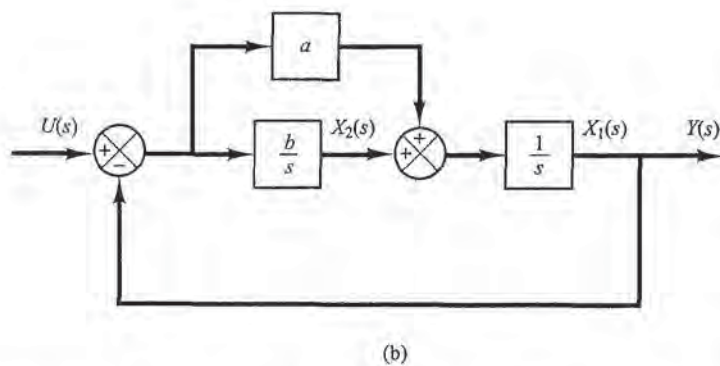
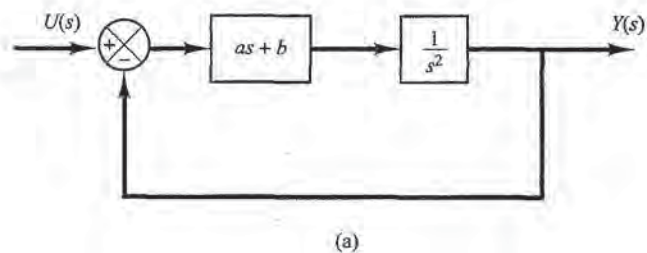


Figure 3-52
(a) Control system;
(b) modified block
diagram.

Taking the inverse Laplace transforms of the preceding three equations, we obtain

$$\dot{x}_1 = -ax_1 + x_2 + au$$

$$\dot{x}_2 = -bx_1 + bu$$

$$y = x_1$$

Rewriting the state and output equations in the standard vector-matrix form, we obtain

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -a & 1 \\ -b & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} u$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

A-3-10. Obtain a state-space representation of the system shown in Figure 3-53(a).

Solution. In this problem, first expand $(s + z)/(s + p)$ into partial fractions.

$$\frac{s + z}{s + p} = 1 + \frac{z - p}{s + p}$$

Next, convert $K/[s(s + a)]$ into the product of K/s and $1/(s + a)$. Then redraw the block diagram, as shown in Figure 3-53(b). Defining a set of state variables, as shown in Figure 3-53(b), we obtain the following equations:

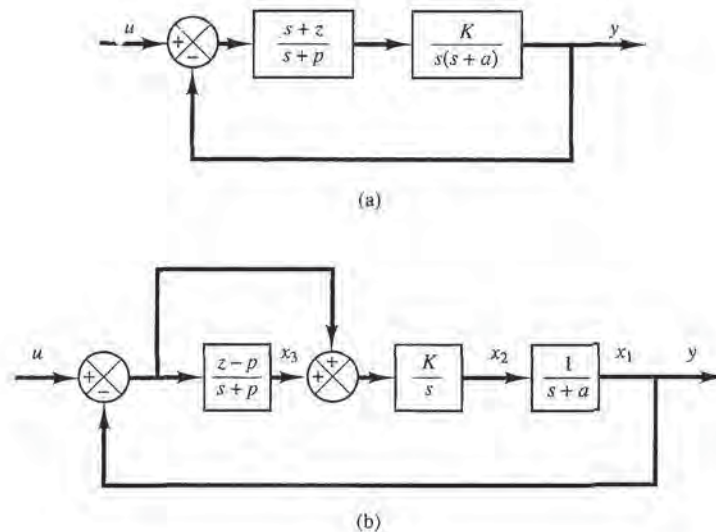
$$\dot{x}_1 = -ax_1 + x_2$$

$$\dot{x}_2 = -Kx_1 + Kx_3 + Ku$$

$$\dot{x}_3 = -(z - p)x_1 - px_3 + (z - p)u$$

$$y = x_1$$

Figure 3-53
a) Control system;
b) block diagram
defining state
variables for the
system.



Rewriting gives

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -a & 1 & 0 \\ -K & 0 & K \\ -(z-p) & 0 & -p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ K \\ z-p \end{bmatrix} u$$

$$y = [1 \ 0 \ 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Notice that the output of the integrator and the outputs of the first-order delayed integrators $[1/(s+a)]$ and $[(z-p)/(s+p)]$ are chosen as state variables. It is important to remember that the output of the block $(s+z)/(s+p)$ in Figure 3-53(a) cannot be a state variable, because this block involves a derivative term, $s+z$.

A-3-11. Obtain the transfer function of the system defined by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \ 0 \ 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Solution. Referring to Equation (3-29), the transfer function $G(s)$ is given by

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + D$$

In this problem, matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and D are

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [1 \ 0 \ 0], \quad D = 0$$

Hence

$$\begin{aligned}
 G(s) &= [1 \ 0 \ 0] \begin{bmatrix} s+1 & -1 & 0 \\ 0 & s+1 & -1 \\ 0 & 0 & s+2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
 &= [1 \ 0 \ 0] \begin{bmatrix} \frac{1}{s+1} & \frac{1}{(s+1)^2} & \frac{1}{(s+1)^2(s+2)} \\ 0 & \frac{1}{s+1} & \frac{1}{(s+1)(s+2)} \\ 0 & 0 & \frac{1}{s+2} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
 &= \frac{1}{(s+1)^2(s+2)} = \frac{1}{s^3 + 4s^2 + 5s + 2}
 \end{aligned}$$

A-3-12. Obtain a state-space representation of the system shown in Figure 3-54.

Solution. The system equations are

$$m_1 \ddot{y}_1 + b \dot{y}_1 + k(y_1 - y_2) = 0$$

$$m_2 \ddot{y}_2 + k(y_2 - y_1) = u$$

The output variables for this system are y_1 and y_2 . Define state variables as

$$x_1 = y_1$$

$$x_2 = \dot{y}_1$$

$$x_3 = y_2$$

$$x_4 = \dot{y}_2$$

Then we obtain the following equations:

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = \frac{1}{m_1} [-b \dot{y}_1 - k(y_1 - y_2)] = -\frac{k}{m_1} x_1 - \frac{b}{m_1} x_2 + \frac{k}{m_1} x_3$$

$$\dot{x}_3 = x_4$$

$$\dot{x}_4 = \frac{1}{m_2} [-k(y_2 - y_1) + u] = \frac{k}{m_2} x_1 - \frac{k}{m_2} x_3 + \frac{1}{m_2} u$$

Hence, the state equation is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k}{m_1} & -\frac{b}{m_1} & \frac{k}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k}{m_2} & 0 & -\frac{k}{m_2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{m_2} \end{bmatrix} u$$

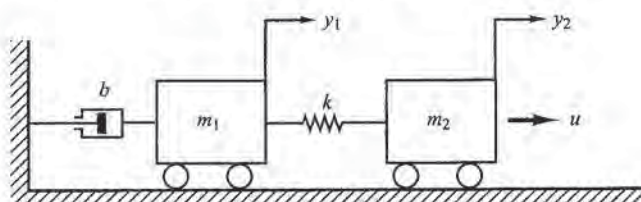


Figure 3-54
Mechanical system.

and the output equation is

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

A-3-13. Consider a system with multiple inputs and multiple outputs. When the system has more than one output, the command

$$[\text{NUM}, \text{den}] = \text{ss2tf}(\text{A}, \text{B}, \text{C}, \text{D}, iu)$$

produces transfer functions for all outputs to each input. (The numerator coefficients are returned to matrix NUM with as many rows as there are outputs.)

Consider the system defined by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -25 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

This system involves two inputs and two outputs. Four transfer functions are involved: $Y_1(s)/U_1(s)$, $Y_2(s)/U_1(s)$, $Y_1(s)/U_2(s)$, and $Y_2(s)/U_2(s)$. (When considering input u_1 , we assume that input u_2 is zero and vice versa.)

Solution. MATLAB Program 3-5 produces four transfer functions.

MATLAB Program 3-5			
A = [0 1;-25 -4];			
B = [1 1;0 1];			
C = [1 0;0 1];			
D = [0 0;0 0];			
[NUM,den] = ss2tf(A,B,C,D,1)			
NUM =			
0 1 4			
0 0 -25			
den =			
1 4 25			
[NUM,den] = ss2tf(A,B,C,D,2)			
NUM =			
0 1.0000 5.0000			
0 1.0000 -25.0000			
den =			
1 4 25			

This is the MATLAB representation of the following four transfer functions:

$$\begin{aligned}\frac{Y_1(s)}{U_1(s)} &= \frac{s+4}{s^2+4s+25}, & \frac{Y_2(s)}{U_1(s)} &= \frac{-25}{s^2+4s+25} \\ \frac{Y_1(s)}{U_2(s)} &= \frac{s+5}{s^2+4s+25}, & \frac{Y_2(s)}{U_2(s)} &= \frac{s-25}{s^2+4s+25}\end{aligned}$$

- A-3-14.** Obtain the equivalent spring constants for the systems shown in Figures 3-55(a) and (b), respectively.

Solution. For the springs in parallel [Figure 3-55(a)] the equivalent spring constant k_{eq} is obtained from

$$k_1x + k_2x = F = k_{eq}x$$

or

$$k_{eq} = k_1 + k_2$$

For the springs in series [Figure-55(b)], the force in each spring is the same. Thus

$$k_1y = F, \quad k_2(x-y) = F$$

Elimination of y from these two equations results in

$$k_2\left(x - \frac{F}{k_1}\right) = F$$

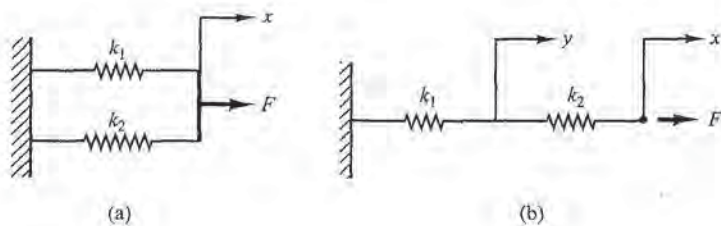
or

$$k_2x = F + \frac{k_2}{k_1}F = \frac{k_1 + k_2}{k_1}F$$

The equivalent spring constant k_{eq} for this case is then found as

$$k_{eq} = \frac{F}{x} = \frac{k_1k_2}{k_1 + k_2} = \frac{1}{\frac{1}{k_1} + \frac{1}{k_2}}$$

Figure 3-55
(a) System consisting of two springs in parallel;
(b) system consisting of two springs in series.



A-3-15. Obtain the equivalent viscous-friction coefficient b_{eq} for each of the systems shown in Figure 3-56(a) and (b).

Solution.

(a) The force f due to the dampers is

$$f = b_1(\dot{y} - \dot{x}) + b_2(\dot{y} - \dot{x}) = (b_1 + b_2)(\dot{y} - \dot{x})$$

In terms of the equivalent viscous friction coefficient b_{eq} , force f is given by

$$f = b_{eq}(\dot{y} - \dot{x})$$

Hence

$$b_{eq} = b_1 + b_2$$

(b) The force f due to the dampers is

$$f = b_1(\dot{z} - \dot{x}) = b_2(\dot{y} - \dot{z}) \quad (3-105)$$

where z is the displacement of a point between damper b_1 and damper b_2 . (Note that the same force is transmitted through the shaft.) From Equation (3-105), we have

$$(b_1 + b_2)\dot{z} = b_2\dot{y} + b_1\dot{x}$$

or

$$\dot{z} = \frac{1}{b_1 + b_2}(b_2\dot{y} + b_1\dot{x}) \quad (3-106)$$

In terms of the equivalent viscous friction coefficient b_{eq} , force f is given by

$$f = b_{eq}(\dot{y} - \dot{x})$$

By substituting Equation (3-106) into Equation (3-105), we have

$$\begin{aligned} f &= b_2(\dot{y} - \dot{z}) = b_2\left[\dot{y} - \frac{1}{b_1 + b_2}(b_2\dot{y} + b_1\dot{x})\right] \\ &= \frac{b_1b_2}{b_1 + b_2}(\dot{y} - \dot{x}) \end{aligned}$$

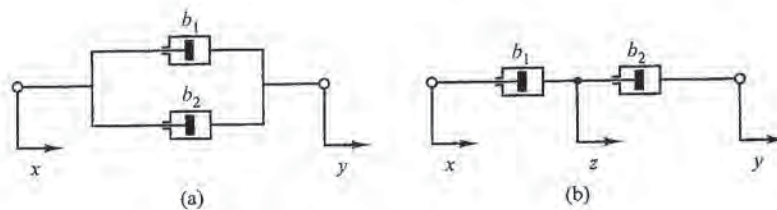
Thus,

$$f = b_{eq}(\dot{y} - \dot{x}) = \frac{b_1b_2}{b_1 + b_2}(\dot{y} - \dot{x})$$

Hence,

$$b_{eq} = \frac{b_1b_2}{b_1 + b_2} = \frac{1}{\frac{1}{b_1} + \frac{1}{b_2}}$$

Figure 3-56
(a) Two dampers connected in parallel;
(b) two dampers connected in series.



A-3-16. Figure 3-57(a) shows a schematic diagram of an automobile suspension system. As the car moves along the road, the vertical displacements at the tires act as the motion excitation to the automobile suspension system. The motion of this system consists of a translational motion of the center of mass and a rotational motion about the center of mass. Mathematical modeling of the complete system is quite complicated.

A very simplified version of the suspension system is shown in Figure 3-57(b). Assuming that the motion x_i at point P is the input to the system and the vertical motion x_o of the body is the output, obtain the transfer function $X_o(s)/X_i(s)$. (Consider the motion of the body only in the vertical direction.) Displacement x_o is measured from the equilibrium position in the absence of input x_i .

Solution. The equation of motion for the system shown in Figure 3-57(b) is

$$m\ddot{x}_o + b(\dot{x}_o - \dot{x}_i) + k(x_o - x_i) = 0$$

or

$$m\ddot{x}_o + b\dot{x}_o + kx_o = b\dot{x}_i + kx_i$$

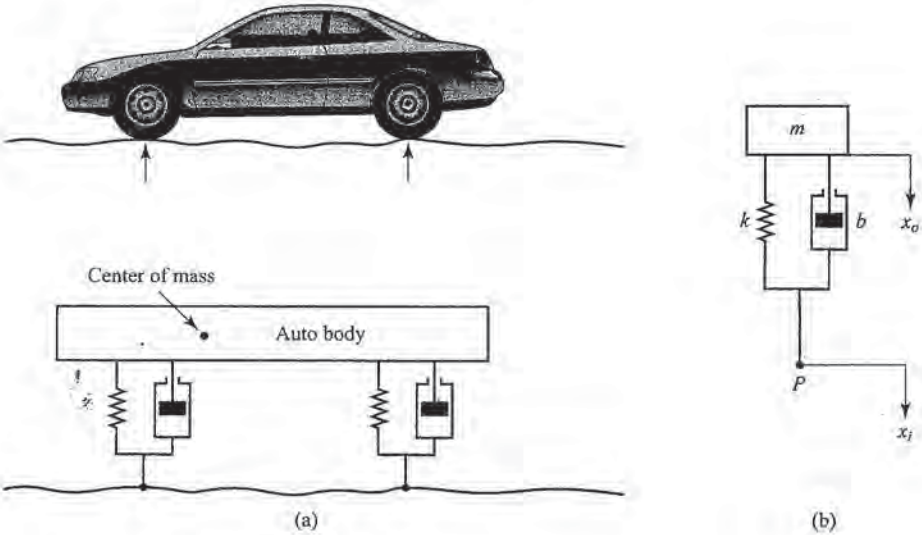
Taking the Laplace transform of this last equation, assuming zero initial conditions, we obtain

$$(ms^2 + bs + k)X_o(s) = (bs + k)X_i(s)$$

Hence the transfer function $X_o(s)/X_i(s)$ is given by

$$\frac{X_o(s)}{X_i(s)} = \frac{bs + k}{ms^2 + bs + k}$$

Figure 3-57
(a) Automobile suspension system;
(b) simplified suspension system.



- A-3-17.** Obtain the transfer function $Y(s)/U(s)$ of the system shown in Figure 3-58. The input u is a displacement input. (Like the system of Problem A-3-16, this is also a simplified version of an automobile or motorcycle suspension system.)

Solution. Assume that displacements x and y are measured from respective steady-state positions in the absence of the input u . Applying the Newton's second law to this system, we obtain

$$m_1 \ddot{x} = k_2(y - x) + b(\dot{y} - \dot{x}) + k_1(u - x)$$

$$m_2 \ddot{y} = -k_2(y - x) - b(\dot{y} - \dot{x})$$

Hence, we have

$$m_1 \ddot{x} + b\dot{x} + (k_1 + k_2)x = b\dot{y} + k_2y + k_1u$$

$$m_2 \ddot{y} + b\dot{y} + k_2y = b\dot{x} + k_2x$$

Taking Laplace transforms of these two equations, assuming zero initial conditions, we obtain

$$[m_1 s^2 + bs + (k_1 + k_2)]X(s) = (bs + k_2)Y(s) + k_1U(s)$$

$$[m_2 s^2 + bs + k_2]Y(s) = (bs + k_2)X(s)$$

Eliminating $X(s)$ from the last two equations, we have

$$(m_1 s^2 + bs + k_1 + k_2) \frac{m_2 s^2 + bs + k_2}{bs + k_2} Y(s) = (bs + k_2)Y(s) + k_1U(s)$$

which yields

$$\frac{Y(s)}{U(s)} = \frac{k_1(bs + k_2)}{m_1 m_2 s^4 + (m_1 + m_2)bs^3 + [k_1 m_2 + (m_1 + m_2)k_2]s^2 + k_1 bs + k_1 k_2}$$

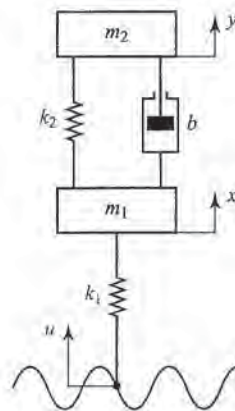


Figure 3-58
Suspension system.

- A-3-18.** Obtain the transfer function of the mechanical system shown in Figure 3-59(a). Also obtain the transfer function of the electrical system shown in Figure 3-59(b). Show that the transfer functions of the two systems are of identical form and thus they are analogous systems.

Solution. In Figure 3-59(a) we assume that displacements x_i , x_o , and y are measured from their respective steady-state positions. Then the equations of motion for the mechanical system shown in Figure 3-59(a) are

$$b_1(\dot{x}_i - \dot{x}_o) + k_1(x_i - x_o) = b_2(\dot{x}_o - \dot{y})$$

$$b_2(\dot{x}_o - \dot{y}) = k_2 y$$

By taking the Laplace transforms of these two equations, assuming zero initial conditions, we have

$$b_1[sX_i(s) - sX_o(s)] + k_1[X_i(s) - X_o(s)] = b_2[sX_o(s) - sY(s)]$$

$$b_2[sX_o(s) - sY(s)] = k_2 Y(s)$$

If we eliminate $Y(s)$ from the last two equations, then we obtain

$$b_1[sX_i(s) - sX_o(s)] + k_1[X_i(s) - X_o(s)] = b_2 s X_o(s) - b_2 s \frac{b_2 s X_o(s)}{b_2 s + k_2}$$

or

$$(b_1 s + k_1)X_i(s) = \left(b_1 s + k_1 + b_2 s - b_2 s \frac{b_2 s}{b_2 s + k_2}\right)X_o(s)$$

Hence the transfer function $X_o(s)/X_i(s)$ can be obtained as

$$\frac{X_o(s)}{X_i(s)} = \frac{\left(\frac{b_1}{k_1} s + 1\right)\left(\frac{b_2}{k_2} s + 1\right)}{\left(\frac{b_1}{k_1} s + 1\right)\left(\frac{b_2}{k_2} s + 1\right) + \frac{b_2}{k_1} s}$$

For the electrical system shown in Figure 3-59(b), the transfer function $E_o(s)/E_i(s)$ is found to be

$$\frac{E_o(s)}{E_i(s)} = \frac{R_1 + \frac{1}{C_1 s}}{\frac{1}{(1/R_2) + C_2 s} + R_1 + \frac{1}{C_1 s}}$$

$$= \frac{(R_1 C_1 s + 1)(R_2 C_2 s + 1)}{(R_1 C_1 s + 1)(R_2 C_2 s + 1) + R_2 C_1 s}$$

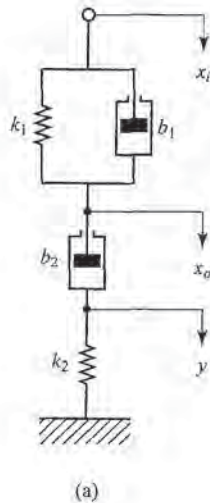
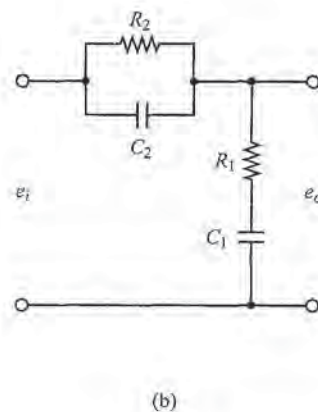


Figure 3-59
(a) Mechanical system;
(b) analogous electrical system.



A comparison of the transfer functions shows that the systems shown in Figures 3-59(a) and (b) are analogous.

A-3-19. Obtain the transfer functions $E_o(s)/E_i(s)$ of the bridged T networks shown in Figures 3-60(a) and (b).

Solution. The bridged T networks shown can both be represented by the network of Figure 3-61(a), where we used complex impedances. This network may be modified to that, shown in Figure 3-61(b).

In Figure 3-61(b), note that

$$I_1 = I_2 + I_3, \quad I_2 Z_1 = (Z_3 + Z_4) I_3$$

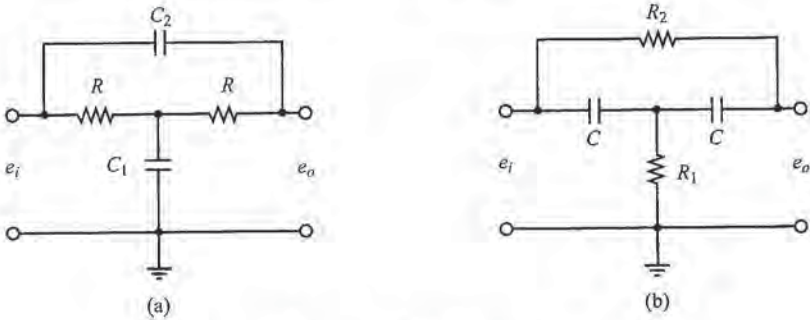


Figure 3-60
Bridged T networks.

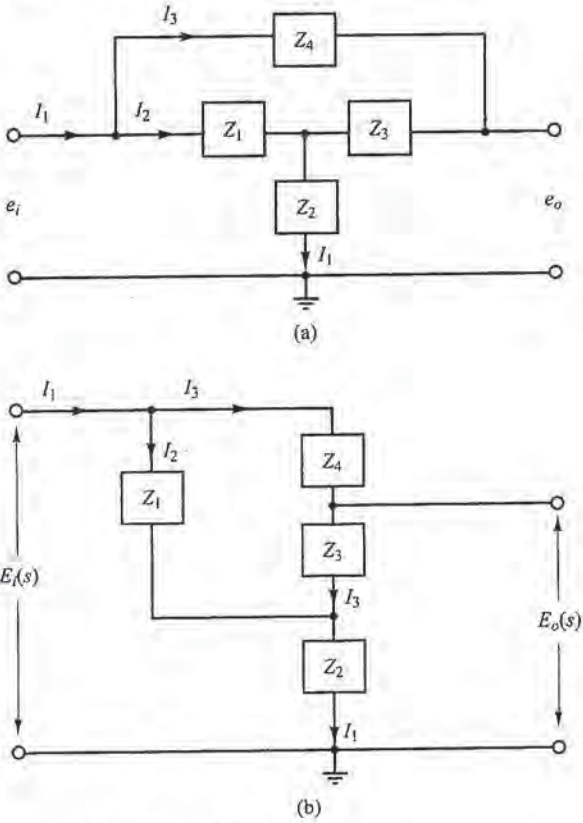


Figure 3-61
(a) Bridged T network in terms of complex impedances; (b) equivalent network.

Hence

$$I_2 = \frac{Z_3 + Z_4}{Z_1 + Z_3 + Z_4} I_1, \quad I_3 = \frac{Z_1}{Z_1 + Z_3 + Z_4} I_1$$

Then the voltages $E_i(s)$ and $E_o(s)$ can be obtained as

$$\begin{aligned} E_i(s) &= Z_1 I_2 + Z_2 I_1 \\ &= \left[Z_2 + \frac{Z_1(Z_3 + Z_4)}{Z_1 + Z_3 + Z_4} \right] I_1 \\ &= \frac{Z_2(Z_1 + Z_3 + Z_4) + Z_1(Z_3 + Z_4)}{Z_1 + Z_3 + Z_4} I_1 \\ E_o(s) &= Z_3 I_3 + Z_2 I_1 \\ &= \frac{Z_3 Z_1}{Z_1 + Z_3 + Z_4} I_1 + Z_2 I_1 \\ &= \frac{Z_3 Z_1 + Z_2(Z_1 + Z_3 + Z_4)}{Z_1 + Z_3 + Z_4} I_1 \end{aligned}$$

Hence, the transfer function $E_o(s)/E_i(s)$ of the network shown in Figure 3-61(a) is obtained as

$$\frac{E_o(s)}{E_i(s)} = \frac{Z_3 Z_1 + Z_2(Z_1 + Z_3 + Z_4)}{Z_2(Z_1 + Z_3 + Z_4) + Z_1 Z_3 + Z_1 Z_4} \quad (3-107)$$

For the bridged T network shown in Figure 3-60(a), substitute

$$Z_1 = R, \quad Z_2 = \frac{1}{C_1 s}, \quad Z_3 = R, \quad Z_4 = \frac{1}{C_2 s}$$

into Equation (3-107). Then, we obtain the transfer function $E_o(s)/E_i(s)$ to be

$$\begin{aligned} \frac{E_o(s)}{E_i(s)} &= \frac{R^2 + \frac{1}{C_1 s} \left(R + R + \frac{1}{C_2 s} \right)}{\frac{1}{C_1 s} \left(R + R + \frac{1}{C_2 s} \right) + R^2 + R \frac{1}{C_2 s}} \\ &= \frac{RC_1 RC_2 s^2 + 2RC_2 s + 1}{RC_1 RC_2 s^2 + (2RC_2 + RC_1)s + 1} \end{aligned}$$

Similarly, for the bridged T network shown in Figure 3-60(b), we substitute

$$Z_1 = \frac{1}{Cs}, \quad Z_2 = R_1, \quad Z_3 = \frac{1}{Cs}, \quad Z_4 = R_2$$

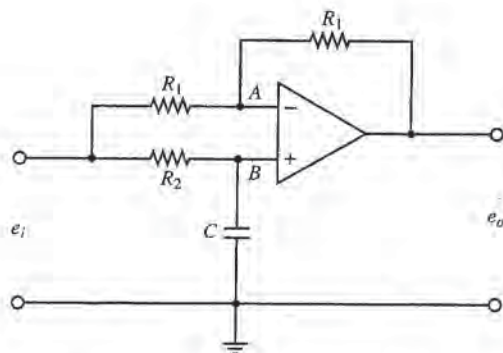
into Equation (3-107). Then the transfer function $E_o(s)/E_i(s)$ can be obtained as follows:

$$\begin{aligned} \frac{E_o(s)}{E_i(s)} &= \frac{\frac{1}{Cs} \frac{1}{Cs} + R_1 \left(\frac{1}{Cs} + \frac{1}{Cs} + R_2 \right)}{R_1 \left(\frac{1}{Cs} + \frac{1}{Cs} + R_2 \right) + \frac{1}{Cs} \frac{1}{Cs} + R_2 \frac{1}{Cs}} \\ &= \frac{R_1 CR_2 Cs^2 + 2R_1 Cs + 1}{R_1 CR_2 Cs^2 + (2R_1 C + R_2 C)s + 1} \end{aligned}$$

Example Problems and Solutions

135

Figure 3-62
Operational-
amplifier circuit.



A-3-20. Obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp circuit shown in Figure 3-62.

Solution. The voltage at point A is

$$e_A = \frac{1}{2}(e_i - e_o) + e_o$$

The Laplace-transformed version of this last equation is

$$E_A(s) = \frac{1}{2}[E_i(s) + E_o(s)]$$

The voltage at point B is

$$E_B(s) = \frac{\frac{1}{Cs}}{R_2 + \frac{1}{Cs}} E_i(s) = \frac{1}{R_2Cs + 1} E_i(s)$$

Since $[E_B(s) - E_A(s)]K = E_o(s)$ and $K \gg 1$, we must have $E_A(s) = E_B(s)$. Thus

$$\frac{1}{2}[E_i(s) + E_o(s)] = \frac{1}{R_2Cs + 1} E_i(s)$$

Hence

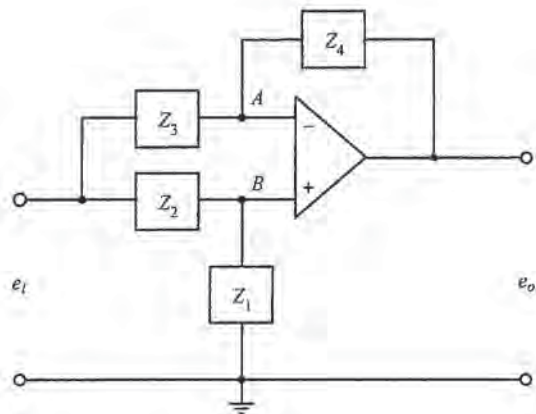
$$\frac{E_o(s)}{E_i(s)} = -\frac{R_2Cs - 1}{R_2Cs + 1} = -\frac{s - \frac{1}{R_2C}}{s + \frac{1}{R_2C}}$$

A-3-21. Obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp system shown in Figure 3-63 in terms of complex impedances Z_1 , Z_2 , Z_3 , and Z_4 . Using the equation derived, obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp system shown in Figure 3-62.

Solution. From Figure 3-63, we find

$$\frac{E_i(s) - E_A(s)}{Z_3} = \frac{E_A(s) - E_o(s)}{Z_4}$$

Figure 3-63
Operational-
amplifier circuit.



or

$$E_i(s) - \left(1 + \frac{Z_3}{Z_4}\right)E_A(s) = -\frac{Z_3}{Z_4}E_o(s) \tag{3-108}$$

Since

$$E_A(s) = E_B(s) = \frac{Z_1}{Z_1 + Z_2}E_i(s) \tag{3-109}$$

by substituting Equation (3-109) into Equation (3-108), we obtain

$$\left[\frac{Z_4Z_1 + Z_4Z_2 - Z_4Z_1 - Z_3Z_1}{Z_4(Z_1 + Z_2)}\right]E_i(s) = -\frac{Z_3}{Z_4}E_o(s)$$

from which we get the transfer function $E_o(s)/E_i(s)$ to be

$$\frac{E_o(s)}{E_i(s)} = -\frac{Z_4Z_2 - Z_3Z_1}{Z_3(Z_1 + Z_2)} \tag{3-110}$$

To find the transfer function $E_o(s)/E_i(s)$ of the circuit shown in Figure 3-62, we substitute

$$Z_1 = \frac{1}{Cs}, \quad Z_2 = R_2, \quad Z_3 = R_1, \quad Z_4 = R_1$$

into Equation (3-110). The result is

$$\frac{E_o(s)}{E_i(s)} = -\frac{R_1R_2 - R_1\frac{1}{Cs}}{R_1\left(\frac{1}{Cs} + R_2\right)} = -\frac{R_2Cs - 1}{R_2Cs + 1}$$

which is, as a matter of course, the same as that obtained in Problem A-3-20.

A-3-22. Obtain the transfer function $E_o(s)/E_i(s)$ of the operational-amplifier circuit shown in Figure 3-64.

Solution. We will first obtain currents i_1, i_2, i_3, i_4 , and i_5 . Then we will use node equations at nodes A and B .

$$i_1 = \frac{e_i - e_A}{R_1}; \quad i_2 = \frac{e_A - e_o}{R_3}; \quad i_3 = C_1 \frac{de_A}{dt}$$

$$i_4 = \frac{e_A}{R_2}; \quad i_5 = C_2 \frac{-de_o}{dt}$$

At node A , we have $i_1 = i_2 + i_3 + i_4$, or

$$\frac{e_i - e_A}{R_1} = \frac{e_A - e_o}{R_3} + C_1 \frac{de_A}{dt} + \frac{e_A}{R_2} \quad (3-111)$$

At node B , we get $i_4 = i_5$, or

$$\frac{e_A}{R_2} = C_2 \frac{-de_o}{dt} \quad (3-112)$$

By rewriting Equation (3-111), we have

$$C_1 \frac{de_A}{dt} + \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \right) e_A = \frac{e_i}{R_1} + \frac{e_o}{R_3} \quad (3-113)$$

From Equation (3-112), we get

$$e_A = -R_2 C_2 \frac{de_o}{dt} \quad (3-114)$$

By substituting Equation (3-114) into Equation (3-113), we obtain

$$C_1 \left(-R_2 C_2 \frac{d^2 e_o}{dt^2} \right) + \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \right) (-R_2 C_2) \frac{de_o}{dt} = \frac{e_i}{R_1} + \frac{e_o}{R_3}$$

Taking the Laplace transform of this last equation, assuming zero initial conditions, we obtain

$$-C_1 C_2 R_2 s^2 E_o(s) + \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \right) (-R_2 C_2) s E_o(s) - \frac{1}{R_3} E_o(s) = \frac{E_i(s)}{R_1}$$

from which we get the transfer function $E_o(s)/E_i(s)$ as follows:

$$\frac{E_o(s)}{E_i(s)} = - \frac{1}{R_1 C_1 R_2 C_2 s^2 + [R_2 C_2 + R_1 C_2 + (R_1/R_3) R_2 C_2] s + (R_1/R_3)}$$

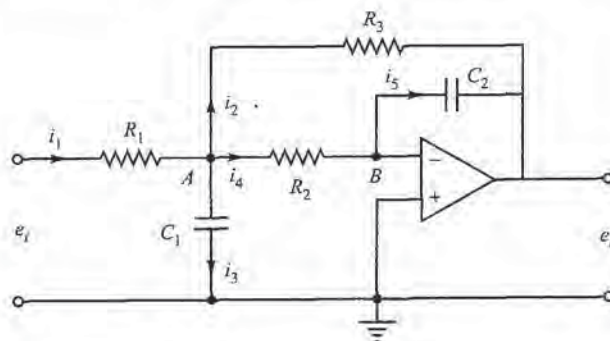


Figure 3-64
Operational-
amplifier circuit.

A-3-23. Consider the servo system shown in Figure 3-65(a). The motor shown is a servomotor, a dc motor designed specifically to be used in a control system. The operation of this system is as follows: A pair of potentiometers acts as an error-measuring device. They convert the input and output positions into proportional electric signals. The command input signal determines the angular position r of the wiper arm of the input potentiometer. The angular position r is the reference input to the system, and the electric potential of the arm is proportional to the angular position of the arm. The output shaft position determines the angular position c of the wiper arm of the output potentiometer. The difference between the input angular position r and the output angular position c is the error signal e , or

$$e = r - c$$

The potential difference $e_r - e_c = e_v$ is the error voltage, where e_r is proportional to r and e_c is proportional to c ; that is, $e_r = K_0 r$ and $e_c = K_0 c$, where K_0 is a proportionality constant. The error voltage that appears at the potentiometer terminals is amplified by the amplifier whose gain constant is K_1 . The output voltage of this amplifier is applied to the armature circuit of the dc motor. A fixed voltage is applied to the field winding. If an error exists, the motor develops a torque to rotate the output load in such a way as to reduce the error to zero. For constant field current, the torque developed by the motor is

$$T = K_2 i_a$$

where K_2 is the motor torque constant and i_a is the armature current.

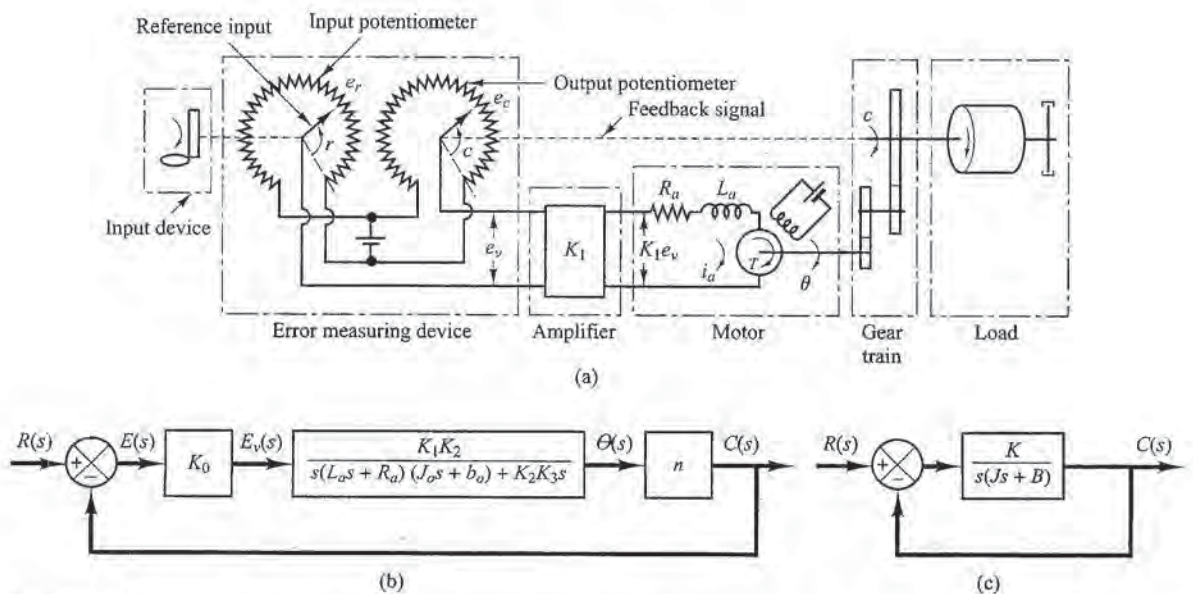


Figure 3-65
(a) Schematic diagram of servo system; (b) block diagram for the system; (c) simplified block diagram.

When the armature is rotating, a voltage proportional to the product of the flux and angular velocity is induced in the armature. For a constant flux, the induced voltage e_b is directly proportional to the angular velocity $d\theta/dt$, or

$$e_b = K_3 \frac{d\theta}{dt}$$

where e_b is the back emf, K_3 is the back emf constant of the motor, and θ is the angular displacement of the motor shaft.

Obtain the transfer function between the motor shaft angular displacement θ and the error voltage e_v . Obtain also a block diagram for this system and a simplified block diagram when L_a is negligible.

Solution. The speed of an armature-controlled dc servomotor is controlled by the armature voltage e_a . (The armature voltage $e_a = K_1 e_v$ is the output of the amplifier.) The differential equation for the armature circuit is

$$L_a \frac{di_a}{dt} + R_a i_a + e_b = e_a$$

or

$$L_a \frac{di_a}{dt} + R_a i_a + K_3 \frac{d\theta}{dt} = K_1 e_v \quad (3-115)$$

The equation for torque equilibrium is

$$J_0 \frac{d^2\theta}{dt^2} + b_0 \frac{d\theta}{dt} = T = K_2 i_a \quad (3-116)$$

where J_0 is the inertia of the combination of the motor, load, and gear train referred to the motor shaft and b_0 is the viscous-friction coefficient of the combination of the motor, load, and gear train referred to the motor shaft.

By eliminating i_a from Equations (3-115) and (3-116), we obtain

$$\frac{\Theta(s)}{E_v(s)} = \frac{K_1 K_2}{s(L_a s + R_a)(J_0 s + b_0) + K_2 K_3 s} \quad (3-117)$$

We assume that the gear ratio of the gear train is such that the output shaft rotates n times for each revolution of the motor shaft. Thus,

$$C(s) = n\Theta(s) \quad (3-118)$$

The relationship among $E_v(s)$, $R(s)$, and $C(s)$ is

$$E_v(s) = K_0[R(s) - C(s)] = K_0 E(s) \quad (3-119)$$

The block diagram of this system can be constructed from Equations (3-117), (3-118), and (3-119), as shown in Figure 3-65(b). The transfer function in the feedforward path of this system is

$$G(s) = \frac{C(s)}{\Theta(s)} \frac{\Theta(s)}{E_v(s)} \frac{E_v(s)}{E(s)} = \frac{K_0 K_1 K_2 n}{s[(L_a s + R_a)(J_0 s + b_0) + K_2 K_3]}$$

When L_a is small, it can be neglected, and the transfer function $G(s)$ in the feedforward path becomes

$$\begin{aligned} G(s) &= \frac{K_0 K_1 K_2 n}{s[R_a(J_0 s + b_0) + K_2 K_3]} \\ &= \frac{K_0 K_1 K_2 n / R_a}{J_0 s^2 + \left(b_0 + \frac{K_2 K_3}{R_a}\right)s} \end{aligned} \quad (3-120)$$

The term $[b_0 + (K_2 K_3 / R_a)]s$ indicates that the back emf of the motor effectively increases the viscous friction of the system. The inertia J_0 and viscous friction coefficient $b_0 + (K_2 K_3 / R_a)$ are referred to the motor shaft. When J_0 and $b_0 + (K_2 K_3 / R_a)$ are multiplied by $1/n^2$, the inertia and viscous-friction coefficient are expressed in terms of the output shaft. Introducing new parameters defined by

$$J = J_0 / n^2 = \text{moment of inertia referred to the output shaft}$$

$$B = [b_0 + (K_2 K_3 / R_a)] / n^2 = \text{viscous-friction coefficient referred to the output shaft}$$

$$K = K_0 K_1 K_2 / n R_a$$

the transfer function $G(s)$ given by Equation (3-120) can be simplified, yielding

$$G(s) = \frac{K}{Js^2 + Bs}$$

or

$$G(s) = \frac{K_m}{s(T_m s + 1)}$$

where

$$K_m = \frac{K}{B}, \quad T_m = \frac{J}{B} = \frac{R_a J_0}{R_a b_0 + K_2 K_3}$$

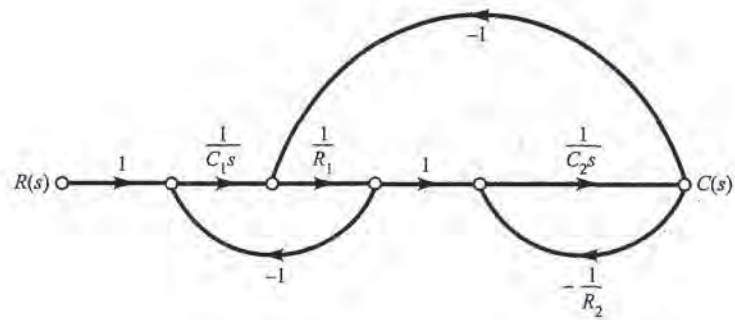
The block diagram of the system shown in Figure 3-65(b) can thus be simplified as shown in Figure 3-65(c).

A-3-24. Consider the system shown in Figure 3-66. Obtain the closed-loop transfer function $C(s)/R(s)$.

Solution. In this system there is only one forward path that connects the input $R(s)$ and the Output $C(s)$. Thus,

$$P_1 = \frac{1}{C_1 s} \frac{1}{R_1} \frac{1}{C_2 s}$$

Figure 3-66
Signal flow graph of
a control system.



There are three individual loops. Thus,

$$L_1 = -\frac{1}{C_1 s} \frac{1}{R_1}$$

$$L_2 = -\frac{1}{C_2 s} \frac{1}{R_2}$$

$$L_3 = -\frac{1}{R_1} \frac{1}{C_2 s}$$

Loop L_1 does not touch loop L_2 . (Loop L_1 touches loop L_3 , and loop L_2 touches loop L_3 .) Hence the determinant Δ is given by

$$\begin{aligned} \Delta &= 1 - (L_1 + L_2 + L_3) + (L_1 L_2) \\ &= 1 + \frac{1}{R_1 C_1 s} + \frac{1}{R_2 C_2 s} + \frac{1}{R_1 C_2 s} + \frac{1}{R_1 C_1 R_2 C_2 s^2} \end{aligned}$$

Since all three loops touch the forward path P_1 , we remove L_1 , L_2 , L_3 , and $L_1 L_2$ from Δ and evaluate the cofactor Δ_1 as follows:

$$\Delta_1 = 1$$

Thus we obtain the closed-loop transfer function to be

$$\begin{aligned} \frac{C(s)}{R(s)} &= \frac{P_1 \Delta_1}{\Delta} \\ &= \frac{\frac{1}{R_1 C_1 C_2 s^2}}{1 + \frac{1}{R_1 C_1 s} + \frac{1}{R_2 C_2 s} + \frac{1}{R_1 C_2 s} + \frac{1}{R_1 C_1 R_2 C_2 s^2}} \\ &= \frac{R_2}{R_1 C_1 R_2 C_2 s^2 + (R_1 C_1 + R_2 C_2 + R_2 C_1)s + 1} \end{aligned}$$

A-3-25. Obtain the transfer function $Y(s)/X(s)$ of the system shown in Figure 3-67.

Solution. The signal flow graph shown in Figure 3-67 can be successively simplified as shown in Figures 3-68 (a), (b), and (c). From Figure 3-68(c), X_3 can be written as

$$X_3 = \frac{1}{s^2} X + \frac{a_1 s + a_2}{s^2} X_3$$

This last equation can be simplified as

$$(s^2 + a_1 s + a_2) X_3 = X$$

from which we obtain

$$\frac{Y(s)}{X(s)} = \frac{b X_3}{X} = \frac{b}{s^2 + a_1 s + a_2}$$

Figure 3-67
Signal flow graph of
a system.

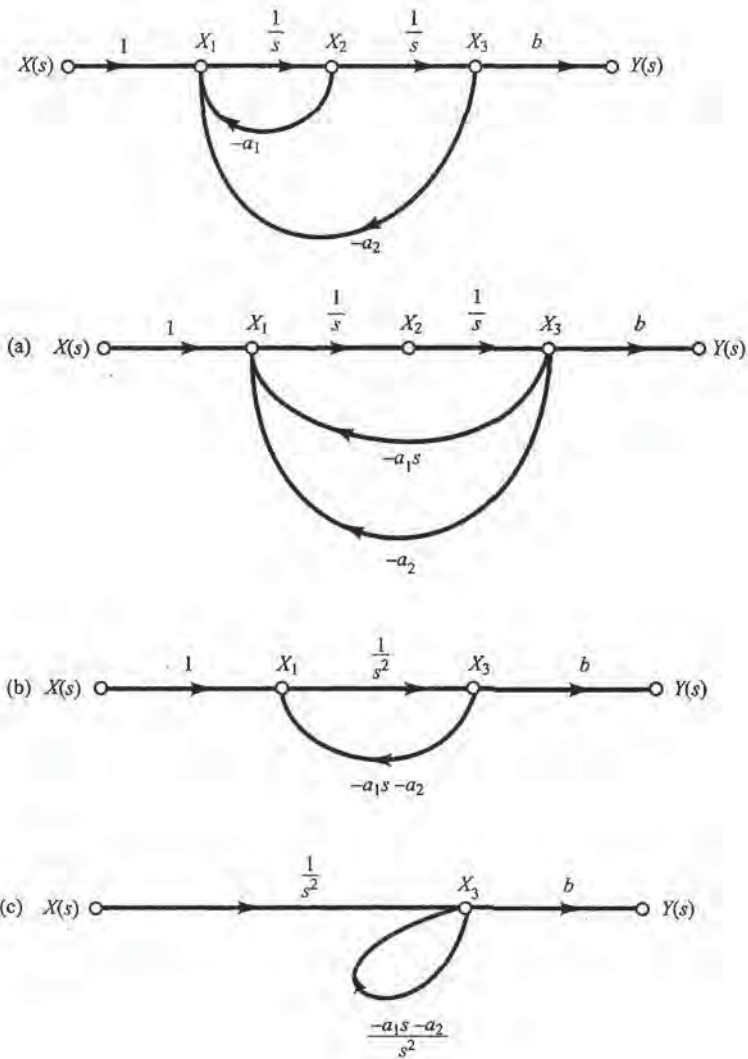


Figure 3-68
Successive
simplifications of the
signal flow graph of
Figure 3-67.

A-3-26. Figure 3-69 is the block diagram of an engine-speed control system. The speed is measured by a set of flyweights. Draw a signal flow graph for this system.

Solution. Referring to Figure 3-36(e), a signal flow graph for

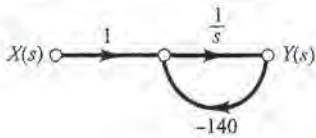
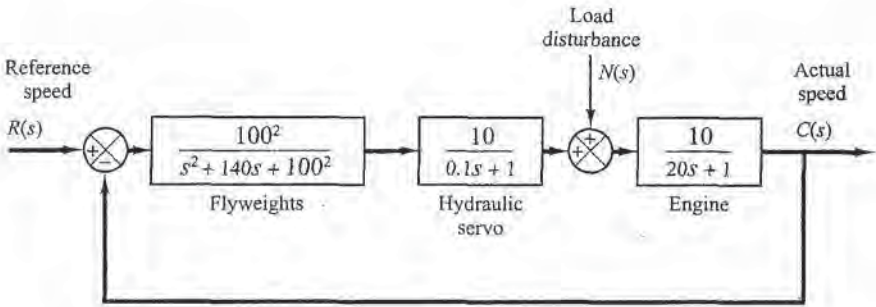
$$\frac{Y(s)}{X(s)} = \frac{1}{s + 140}$$

may be drawn as shown in Figure 3-70(a). Similarly, a signal flow graph for

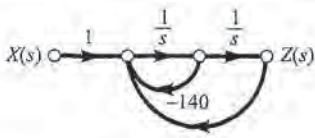
$$\frac{Z(s)}{X(s)} = \frac{1}{s^2 + 140s + 100^2} = \frac{\frac{1}{s + 140} \frac{1}{s}}{1 + \frac{100^2}{s + 140} \frac{1}{s}}$$

may be drawn as shown in Figure 3-70(b). Drawing a signal flow graph for each of the system components and combining them together, a signal flow graph for the complete system may be obtained as shown in Figure 3-70(c).

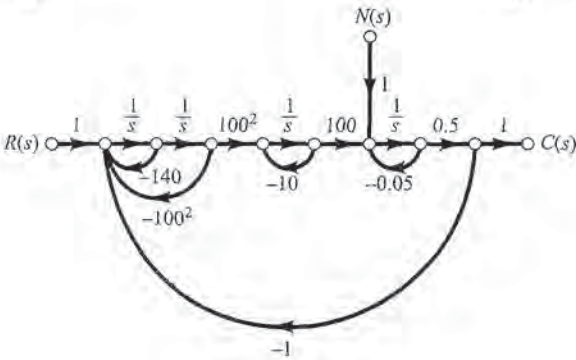
Figure 3-69
Block diagram of an engine-speed control system.



(a)



(b)



(c)

Figure 3-70
(a) Signal flow graph for $Y(s)/X(s) = 1/(s + 140)$;
(b) signal flow graph for $Z(s)/X(s) = 1/(s^2 + 140s + 100^2)$;
(c) signal flow graph for the system shown in Fig. 3-69.

A-3-27. Linearize the nonlinear equation

$$z = x^2 + 4xy + 6y^2$$

in the region defined by $8 \leq x \leq 10, 2 \leq y \leq 4$.

Solution. Define

$$f(x, y) = z = x^2 + 4xy + 6y^2$$

Then

$$z = f(x, y) = f(\bar{x}, \bar{y}) + \left[\frac{\partial f}{\partial x} (x - \bar{x}) + \frac{\partial f}{\partial y} (y - \bar{y}) \right]_{x=\bar{x}, y=\bar{y}} + \dots$$

where $\bar{x} = 9, \bar{y} = 3$.

Since the higher-order terms in the expanded equation are small, neglecting these higher-order terms, we obtain

$$z - \bar{z} = K_1(x - \bar{x}) + K_2(y - \bar{y})$$

where

$$K_1 = \left. \frac{\partial f}{\partial x} \right|_{x=\bar{x}, y=\bar{y}} = 2\bar{x} + 4\bar{y} = 2 \times 9 + 4 \times 3 = 30$$

$$K_2 = \left. \frac{\partial f}{\partial y} \right|_{x=\bar{x}, y=\bar{y}} = 4\bar{x} + 12\bar{y} = 4 \times 9 + 12 \times 3 = 72$$

$$\bar{z} = \bar{x}^2 + 4\bar{x}\bar{y} + 6\bar{y}^2 = 9^2 + 4 \times 9 \times 3 + 6 \times 9 = 243$$

Thus

$$z - 243 = 30(x - 9) + 72(y - 3)$$

Hence a linear approximation of the given nonlinear equation near the operating point is

$$z - 30x - 72y + 243 = 0$$

PROBLEMS

B-3-1. Simplify the block diagram shown in Figure 3-71 and obtain the closed-loop transfer function $C(s)/R(s)$.

B-3-2. Simplify the block diagram shown in Figure 3-72 and obtain the transfer function $C(s)/R(s)$.

B-3-3. Simplify the block diagram shown in Figure 3-73 and obtain the closed-loop transfer function $C(s)/R(s)$.

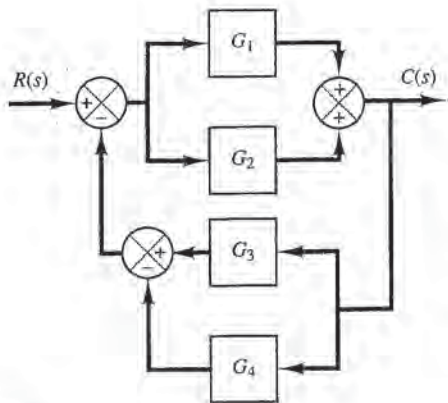


Figure 3-71
Block diagram of a system.

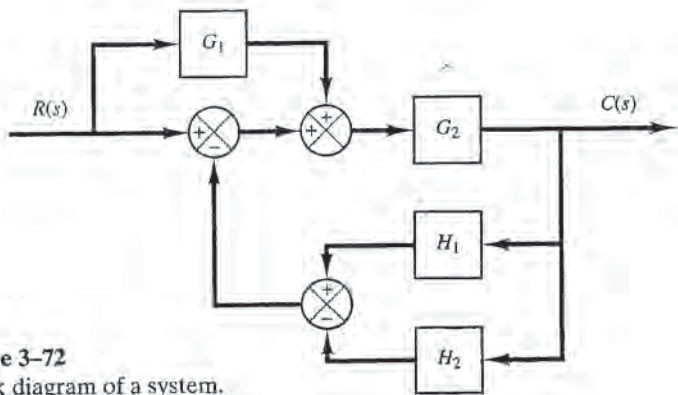


Figure 3-72
Block diagram of a system.

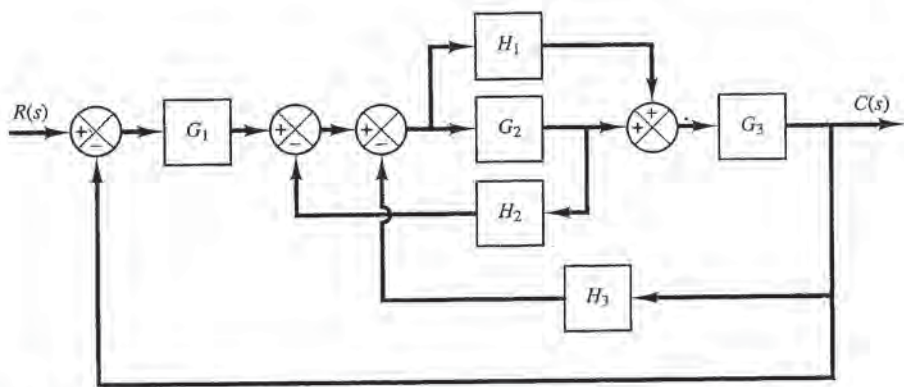


Figure 3-73
Block diagram of a system.

B-3-4. Consider industrial automatic controllers whose control actions are proportional, integral, proportional-plus-integral, proportional-plus-derivative, and proportional-plus-integral-plus-derivative. The transfer functions of these controllers can be given, respectively, by

$$\begin{aligned}\frac{U(s)}{E(s)} &= K_p \\ \frac{U(s)}{E(s)} &= \frac{K_i}{s} \\ \frac{U(s)}{E(s)} &= K_p \left(1 + \frac{1}{T_i s} \right) \\ \frac{U(s)}{E(s)} &= K_p (1 + T_d s) \\ \frac{U(s)}{E(s)} &= K_p \left(1 + \frac{1}{T_i s} + T_d s \right)\end{aligned}$$

where $U(s)$ is the Laplace transform of $u(t)$, the controller output, and $E(s)$ the Laplace transform of $e(t)$, the actuating error signal.

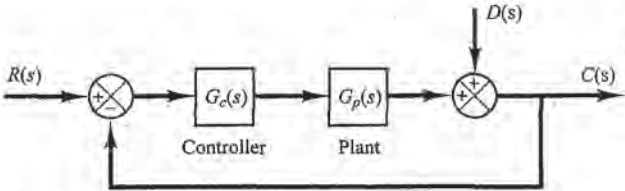


Figure 3-74
Closed-loop system.

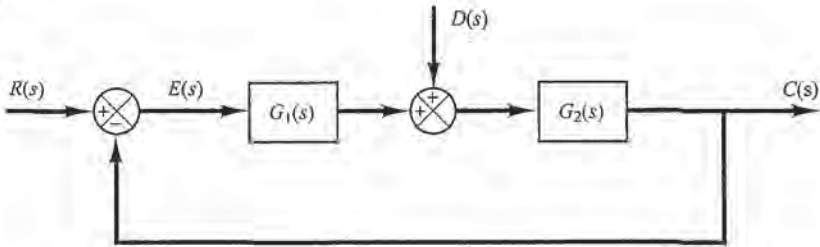


Figure 3-75
Control system.

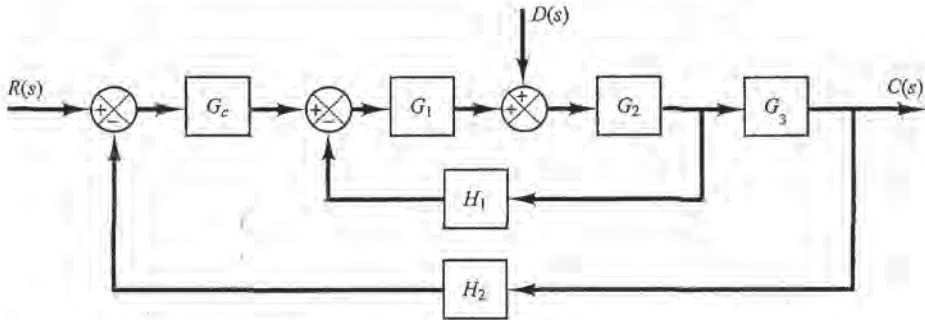


Figure 3-76
Control system.

ing error signal. Sketch $u(t)$ versus t curves for each of the five types of controllers when the actuating error signal is

- (a) $e(t)$ = unit-step function
- (b) $e(t)$ = unit-ramp function

In sketching curves, assume that the numerical values of K_p , K_i , T_i , and T_d are given as

$$\begin{aligned}K_p &= \text{proportional gain} = 4 \\ K_i &= \text{integral gain} = 2 \\ T_i &= \text{integral time} = 2 \text{ sec} \\ T_d &= \text{derivative time} = 0.8 \text{ sec}\end{aligned}$$

B-3-5. Figure 3-74 shows a closed-loop system with a reference input and disturbance input. Obtain the expression for the output $C(s)$ when both the reference input and disturbance input are present.

B-3-6. Consider the system shown in Figure 3-75. Derive the expression for the steady-state error when both the reference input $R(s)$ and disturbance input $D(s)$ are present.

B-3-7. Obtain the transfer functions $C(s)/R(s)$ and $C(s)/D(s)$ of the system shown in Figure 3-76.

B-3-8. Obtain a state-space representation of the system shown in Figure 3-77.

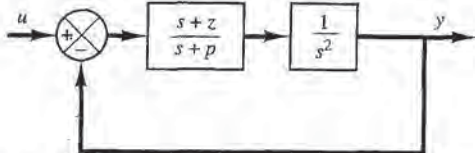


Figure 3-77
Control system.

B-3-9. Consider the system described by

$$\ddot{y} + 3\dot{y} + 2y = u$$

Derive a state-space representation of the system.

B-3-10. Consider the system described by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -4 & -1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Obtain the transfer function of the system.

B-3-11. Consider a system defined by the following state-space equations:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Obtain the transfer function $G(s)$ of the system.

B-3-12. Obtain the transfer matrix of the system defined by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -4 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

B-3-13. Obtain the equivalent viscous-friction coefficient b_{eq} of the system shown in Figure 3-78.

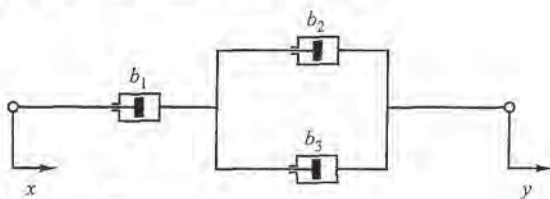


Figure 3-78
Damper system.

B-3-14. Obtain mathematical models of the mechanical systems shown in Figure 3-79(a) and (b).

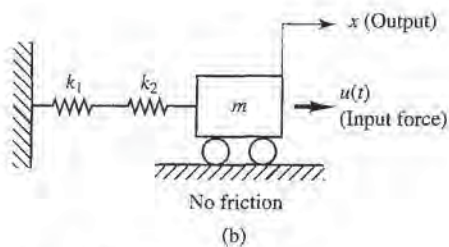
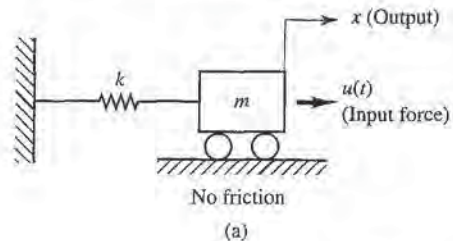


Figure 3-79
Mechanical systems.

B-3-15. Obtain a state-space representation of the mechanical system shown in Figure 3-80, where u_1 and u_2 are the inputs and y_1 and y_2 are the outputs.

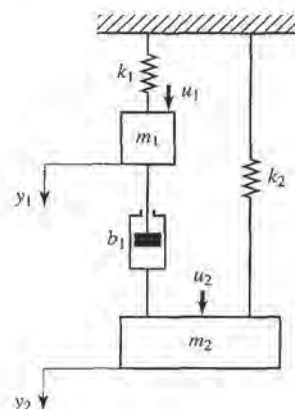


Figure 3-80
Mechanical system.

B-3-16. Consider the spring-loaded pendulum system shown in Figure 3-81. Assume that the spring force acting on the pendulum is zero when the pendulum is vertical, or $\theta = 0$. Assume also that the friction involved is negligible and the angle of oscillation θ is small. Obtain a mathematical model of the system.

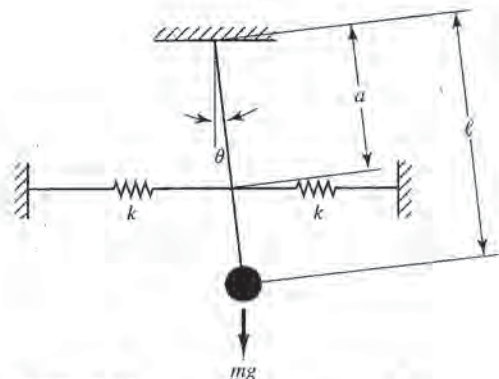


Figure 3-81
Spring-loaded pendulum system.

B-3-17. Referring to Examples 3-8 and 3-9, consider the inverted pendulum system shown in Figure 3-82. Assume that the mass of the inverted pendulum is m and is evenly distributed along the length of the rod. (The center of gravity of the pendulum is located at the center of the rod.) Assuming that θ is small, derive mathematical models for the system in the forms of differential equations, transfer functions, and state-space equations.

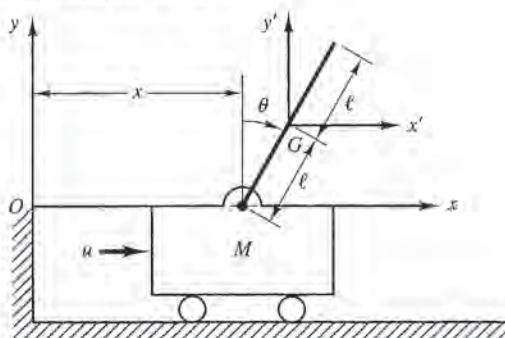


Figure 3-82
Inverted pendulum system.

B-3-18. Obtain the transfer functions $X_1(s)/U(s)$ and $X_2(s)/U(s)$ of the mechanical system shown in Figure 3-83.

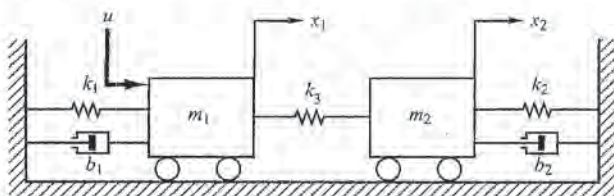


Figure 3-83
Mechanical system.

B-3-19. Obtain the transfer function $E_o(s)/E_i(s)$ of the electrical circuit shown in Figure 3-84.

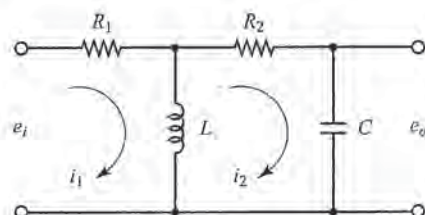


Figure 3-84
Electrical circuit.

B-3-20. Consider the electrical circuit shown in Figure 3-85. Obtain the transfer function $E_o(s)/E_i(s)$ by use of the block diagram approach.

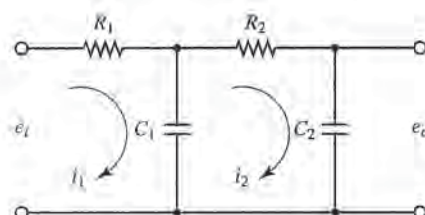


Figure 3-85
Electrical circuit.

B-3-21. Derive the transfer function of the electrical circuit shown in Figure 3-86. Draw a schematic diagram of an analogous mechanical system.

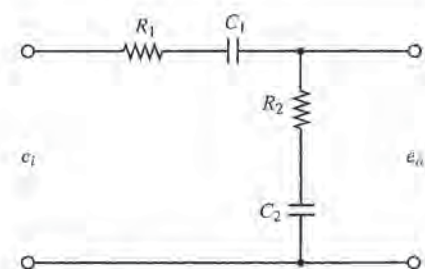


Figure 3-86
Electrical circuit.

B-3-22. Obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp circuit shown in Figure 3-87.

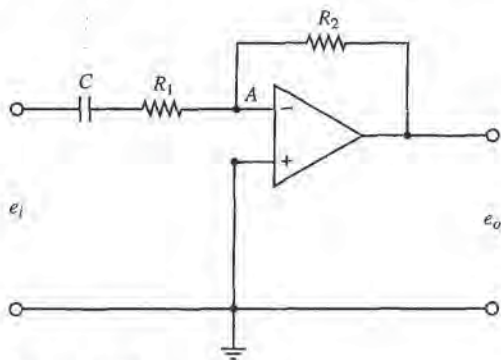


Figure 3-87
Operational-amplifier circuit.

B-3-23. Obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp circuit shown in Figure 3-88.

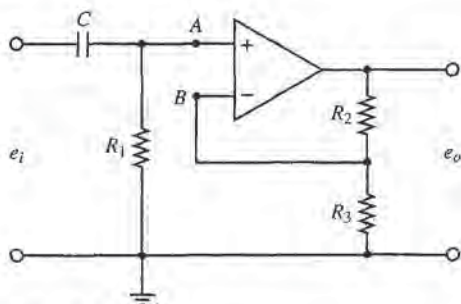


Figure 3-88
Operational-amplifier circuit.

B-3-24. Using the impedance approach, obtain the transfer function $E_o(s)/E_i(s)$ of the op-amp circuit shown in Figure 3-89.

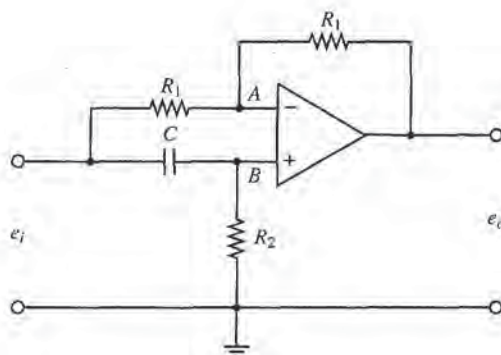


Figure 3-89
Operational-amplifier circuit.

B-3-25. Consider the system shown in Figure 3-90. An armature-controlled dc servomotor drives a load consisting of the moment of inertia J_L . The torque developed by the motor is T . The moment of inertia of the motor rotor is J_m . The angular displacements of the motor rotor and the load element are θ_m and θ , respectively. The gear ratio is $n = \theta/\theta_m$. Obtain the transfer function $\Theta(s)/E_i(s)$.

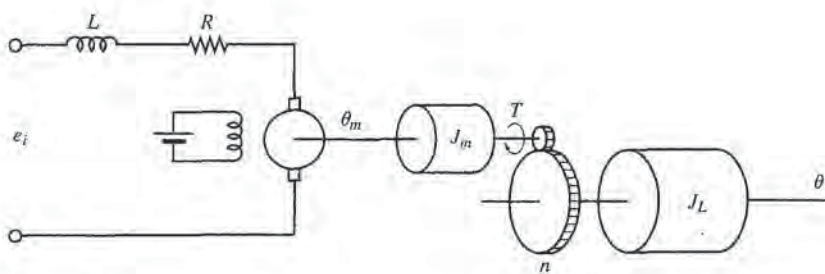


Figure 3-90
Armature-controlled dc servomotor system.

B-3-26. Obtain the transfer function $Y(s)/X(s)$ of the system shown, in Figure 3-91.

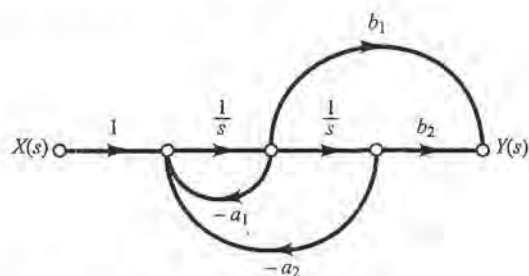


Figure 3-91
Signal flow graph of a system.

B-3-27. Obtain the transfer function $Y(s)/X(s)$ of the system shown in Figure 3-92.

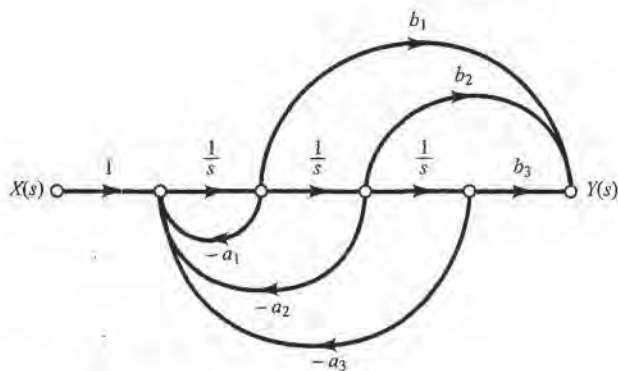


Figure 3-92
Signal flow graph of a system.

B-3-28. Linearize the nonlinear equation

$$z = x^2 + 8xy + 3y^2$$

in the region defined by $2 \leq x \leq 4, 10 \leq y \leq 12$.

B-3-29. Find a linearized equation for

$$y = 0.2x^3$$

about a point $x = 2$.



Mathematical Modeling of Fluid Systems and Thermal Systems*

4-1 INTRODUCTION

This chapter treats mathematical modeling of fluid systems and thermal systems. As the most versatile medium for transmitting signals and power, fluids—liquids and gases—have wide usage in industry. Liquids and gases can be distinguished basically by their relative incompressibilities and the fact that a liquid may have a free surface, whereas a gas expands to fill its vessel. In the engineering field the term *pneumatic* describes fluid systems that use air or gases and *hydraulic* applies to those using oil.

We first discuss liquid-level systems that are frequently used in process control. Here we introduce the concepts of resistance and capacitance to describe the dynamics of such systems. Then we treat pneumatic systems. Such systems are extensively used in the automation of production machinery and in the field of automatic controllers. For instance, pneumatic circuits that convert the energy of compressed air into mechanical energy enjoy wide usage. Also, various types of pneumatic controllers are widely used in industry. Next, we present hydraulic servo systems. These are widely used in machine tool systems, aircraft control systems, etc. We discuss basic aspects of hydraulic servo systems and hydraulic controllers. Both pneumatic systems and hydraulic systems can be modeled easily by using the concepts of resistance and capacitance. Finally, we treat simple thermal systems. Such systems involve heat transfer from one substance to another. Mathematical models of such systems can be obtained by using thermal resistance and thermal capacitance.

*This chapter assumes an introductory background of fluid dynamics, thermodynamics, and heat transfer as normally required for mechanical engineering curriculum. If the students using this text do not have such background, this chapter may be skipped without losing continuity of the analysis and design of control systems presented in this book.

Outline of the Chapter. Section 4-1 has just presented introductory material for the chapter. Section 4-2 discusses liquid-level systems. Section 4-3 treats pneumatic systems—in particular, the basic principles of pneumatic controllers. Section 4-4 first discusses hydraulic servo systems and then presents hydraulic controllers. Finally, Section 4-5 analyzes thermal systems and obtains mathematical models of such systems.

4-2 LIQUID-LEVEL SYSTEMS

In analyzing systems involving fluid flow, we find it necessary to divide flow regimes into laminar flow and turbulent flow, according to the magnitude of the Reynolds number. If the Reynolds number is greater than about 3000 to 4000, then the flow is turbulent. The flow is laminar if the Reynolds number is less than about 2000. In the laminar case, fluid flow occurs in streamlines with no turbulence. Systems involving laminar flow may be represented by linear differential equations.

Industrial processes often involve flow of liquids through connecting pipes and tanks. The flow in such processes is often turbulent and not laminar. Systems involving turbulent flow often have to be represented by nonlinear differential equations. If the region of operation is limited, however, such nonlinear differential equations can be linearized. We shall discuss such linearized mathematical models of liquid-level systems in this section. Note that the introduction of concepts of resistance and capacitance for such liquid-level systems enables us to describe their dynamic characteristics in simple forms.

Resistance and Capacitance of Liquid-Level Systems. Consider the flow through a short pipe connecting two tanks. The resistance R for liquid flow in such a pipe or restriction is defined as the change in the level difference (the difference of the liquid levels of the two tanks) necessary to cause a unit change in flow rate; that is,

$$R = \frac{\text{change in level difference, m}}{\text{change in flow rate, m}^3/\text{sec}}$$

Since the relationship between the flow rate and level difference differs for the laminar flow and turbulent flow, we shall consider both cases in the following.

Consider the liquid-level system shown in Figure 4-1(a). In this system the liquid spouts through the load valve in the side of the tank. If the flow through this restriction is laminar, the relationship between the steady-state flow rate and steady-state head at the level of the restriction is given by

$$Q = KH$$

where Q = steady-state liquid flow rate, m^3/sec

K = coefficient, m^2/sec

H = steady-state head, m

For laminar flow, the resistance R_l is obtained as

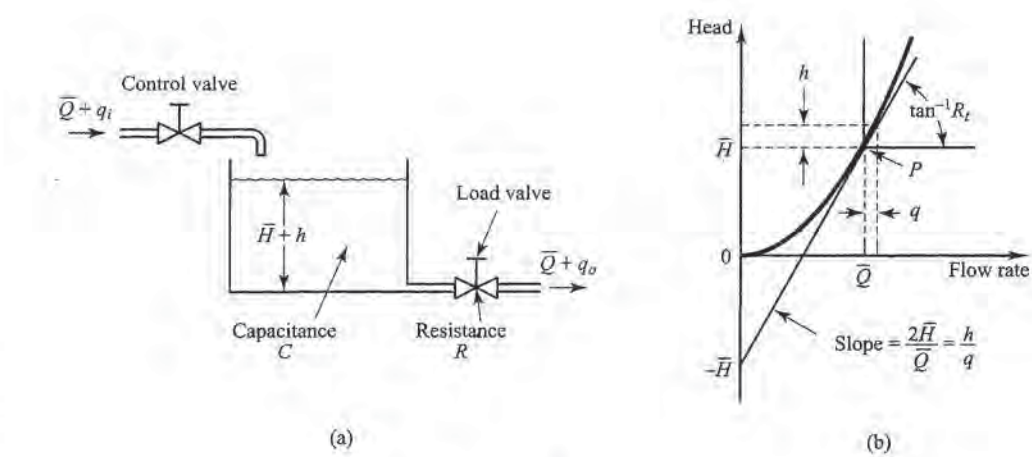
$$R_l = \frac{dH}{dQ} = \frac{H}{Q}$$

The laminar-flow resistance is constant and is analogous to the electrical resistance.

If the flow through the restriction is turbulent, the steady-state flow rate is given by

$$Q = K\sqrt{H} \quad (4-1)$$

Figure 4-1
(a) Liquid-level system; (b) head versus flow rate curve.



where Q = steady-state liquid flow rate, m^3/sec
 K = coefficient, $\text{m}^{2.5}/\text{sec}$
 H = steady-state head, m

The resistance R_t for turbulent flow is obtained from

$$R_t = \frac{dH}{dQ}$$

Since from Equation (4-1) we obtain

$$dQ = \frac{K}{2\sqrt{H}} dH$$

we have

$$\frac{dH}{dQ} = \frac{2\sqrt{H}}{K} = \frac{2\sqrt{H}\sqrt{H}}{Q} = \frac{2H}{Q}$$

Thus,

$$R_t = \frac{2H}{Q}$$

The value of the turbulent-flow resistance R_t depends on the flow rate and the head. The value of R_t , however, may be considered constant if the changes in head and flow rate are small.

By use of the turbulent-flow resistance, the relationship between Q and H can be given by

$$Q = \frac{2H}{R_t}$$

Such linearization is valid, provided that changes in the head and flow rate from their respective steady-state values are small.

In many practical cases, the value of the coefficient K in Equation (4-1), which depends on the flow coefficient and the area of restriction, is not known. Then the resistance may be determined by plotting the head versus flow rate curve based on experimental data and measuring the slope of the curve at the operating condition. An example of such a plot

is shown in Figure 4-1(b). In the figure, point P is the steady-state operating point. The tangent line to the curve at point P intersects the ordinate at point $(0, -\bar{H})$. Thus, the slope of this tangent line is $2\bar{H}/\bar{Q}$. Since the resistance R_i at the operating point P is given by $2\bar{H}/\bar{Q}$, the resistance R_i is the slope of the curve at the operating point.

Consider the operating condition in the neighborhood of point P . Define a small deviation of the head from the steady-state value as h and the corresponding small change of the flow rate as q . Then the slope of the curve at point P can be given by

$$\text{Slope of curve at point } P = \frac{h}{q} = \frac{2\bar{H}}{\bar{Q}} = R_i$$

The linear approximation is based on the fact that the actual curve does not differ much from its tangent line if the operating condition does not vary too much.

The capacitance C of a tank is defined to be the change in quantity of stored liquid necessary to cause a unit change in the potential (head). (The potential is the quantity that indicates the energy level of the system.)

$$C = \frac{\text{change in liquid stored, m}^3}{\text{change in head, m}}$$

It should be noted that the capacity (m^3) and the capacitance (m^2) are different. The capacitance of the tank is equal to its cross-sectional area. If this is constant, the capacitance is constant for any head.

Liquid-Level Systems. Consider the system shown in Figure 4-1(a). The variables are defined as follows:

- $\bar{Q}$ = steady-state flow rate (before any change has occurred), m^3/sec
- q_i = small deviation of inflow rate from its steady-state value, m^3/sec
- q_o = small deviation of outflow rate from its steady-state value, m^3/sec
- $\bar{H}$ = steady-state head (before any change has occurred), m
- h = small deviation of head from its steady-state value, m

As stated previously, a system can be considered linear if the flow is laminar. Even if the flow is turbulent, the system can be linearized if changes in the variables are kept small. Based on the assumption that the system is either linear or linearized, the differential equation of this system can be obtained as follows: Since the inflow minus outflow during the small time interval dt is equal to the additional amount stored in the tank, we see that

$$C dh = (q_i - q_o) dt$$

From the definition of resistance, the relationship between q_o and h is given by

$$q_o = \frac{h}{R}$$

The differential equation for this system for a constant value of R becomes

$$RC \frac{dh}{dt} + h = Rq_i \quad (4-2)$$

Note that RC is the time constant of the system. Taking the Laplace transforms of both sides of Equation (4-2), assuming the zero initial condition, we obtain

$$(RCs + 1)H(s) = RQ_i(s)$$

where

$$H(s) = \mathcal{L}[h] \quad \text{and} \quad Q_i(s) = \mathcal{L}[q_i]$$

If q_i is considered the input and h the output, the transfer function of the system is

$$\frac{H(s)}{Q_i(s)} = \frac{R}{RCs + 1}$$

If, however, q_o is taken as the output, the input being the same, then the transfer function is

$$\frac{Q_o(s)}{Q_i(s)} = \frac{1}{RCs + 1}$$

where we have used the relationship

$$Q_o(s) = \frac{1}{R} H(s)$$

Liquid-Level Systems with Interaction. Consider the system shown in Figure 4-2. In this system, the two tanks interact. Thus the transfer function of the system is not the product of two first-order transfer functions.

In the following, we shall assume only small variations of the variables from the steady-state values. Using the symbols as defined in Figure 4-2, we can obtain the following equations for this system:

$$\frac{h_1 - h_2}{R_1} = q_1 \quad (4-3)$$

$$C_1 \frac{dh_1}{dt} = q - q_1 \quad (4-4)$$

$$\frac{h_2}{R_2} = q_2 \quad (4-5)$$

$$C_2 \frac{dh_2}{dt} = q_1 - q_2 \quad (4-6)$$

If q is considered the input and q_2 the output, the transfer function of the system is

$$\frac{Q_2(s)}{Q(s)} = \frac{1}{R_1 C_1 R_2 C_2 s^2 + (R_1 C_1 + R_2 C_2 + R_2 C_1)s + 1} \quad (4-7)$$

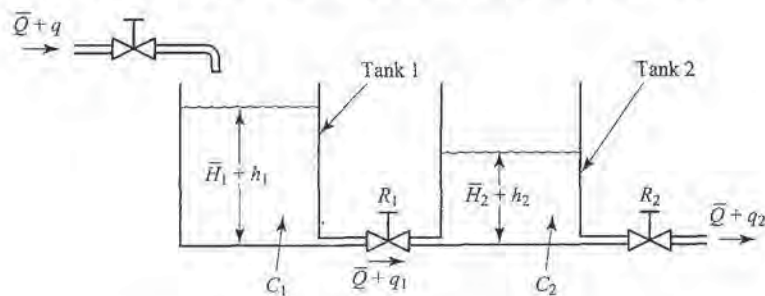


Figure 4-2
Liquid-level system
with interaction.

$\bar{Q}$: Steady-state flow rate
 $\bar{H}_1$: Steady-state liquid level of tank 1
 $\bar{H}_2$: Steady-state liquid level of tank 2

It is instructive to obtain Equation (4-7), the transfer function of the interacted system, by block diagram reduction. From Equations (4-3) through (4-6), we obtain the elements of the block diagram, as shown in Figure 4-3(a). By connecting signals properly, we can construct a block diagram, as shown in Figure 4-3(b). This block diagram can be simplified, as shown in Figure 4-3(c). Further simplifications result in Figures 4-3(d) and (e). Figure 4-3(e) is equivalent to Equation (4-7).

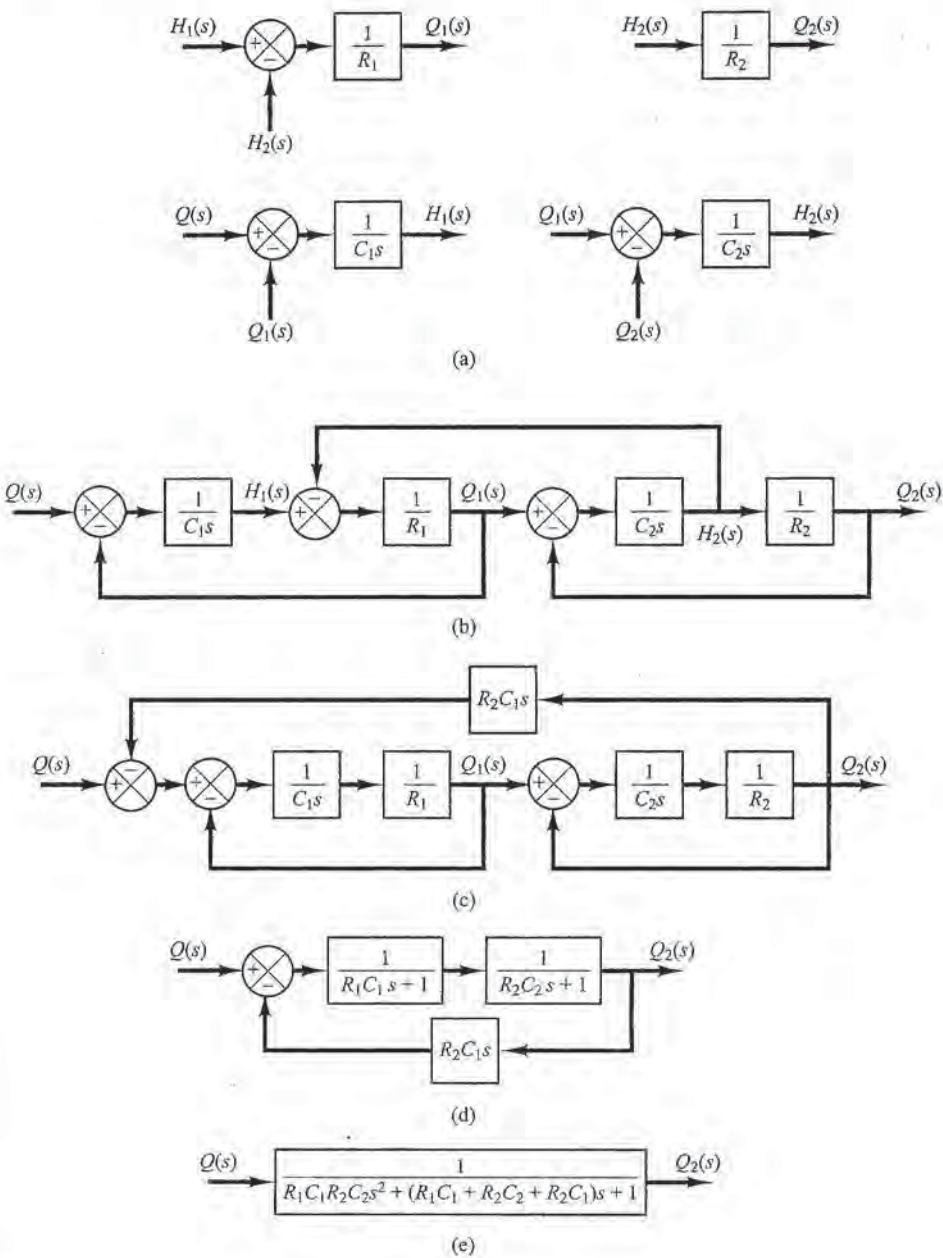


Figure 4-3
(a) Elements of the block diagram of the system shown in Figure 4-2; (b) block diagram of the system; (c)-(e) successive reductions of the block diagram.

Notice the similarity and difference between the transfer function given by Equation (4-7) and that given by Equation (3-72). The term $R_2 C_1 s$ that appears in the denominator of Equation (4-7) exemplifies the interaction between the two tanks. Similarly, the term $R_1 C_2 s$ in the denominator of Equation (3-72) represents the interaction between the two RC circuits shown in Figure 3-23.

4-3 PNEUMATIC SYSTEMS

In industrial applications pneumatic systems and hydraulic systems are frequently compared. Therefore, before we discuss pneumatic systems in detail, we shall give a brief comparison of these two kinds of systems.

Comparison Between Pneumatic Systems and Hydraulic Systems. The fluid generally found in pneumatic systems is air; in hydraulic systems it is oil. And it is primarily the different properties of the fluids involved that characterize the differences between the two systems. These differences can be listed as follows:

1. Air and gases are compressible, whereas oil is incompressible, (except at high pressure).
2. Air lacks lubricating property and always contains water vapor. Oil functions as a hydraulic fluid as well as a lubricator.
3. The normal operating pressure of pneumatic systems is very much lower than that of hydraulic systems.
4. Output powers of pneumatic systems are considerably less than those of hydraulic systems.
5. Accuracy of pneumatic actuators is poor at low velocities, whereas accuracy of hydraulic actuators may be made satisfactory at all velocities.
6. In pneumatic systems, external leakage is permissible to a certain extent, but internal leakage must be avoided because the effective pressure difference is rather small. In hydraulic systems internal leakage is permissible to a certain extent, but external leakage must be avoided.
7. No return pipes are required in pneumatic systems when air is used, whereas they are always needed in hydraulic systems.
8. Normal operating temperature for pneumatic systems is 5° to 60°C (41° to 140°F). The pneumatic system, however, can be operated in the 0° to 200°C (32° to 392°F) range. Pneumatic systems are insensitive to temperature changes, in contrast to hydraulic systems, in which fluid friction due to viscosity depends greatly on temperature. Normal operating temperature for hydraulic systems is 20° to 70°C (68° to 158°F).
9. Pneumatic systems are fire- and explosion-proof, whereas hydraulic systems are not, unless nonflammable liquid is used.

In what follows we begin with a mathematical modeling of pneumatic systems. Then we shall present pneumatic proportional controllers.

We shall first give detailed discussions of the principle by which proportional controllers operate. Then we shall treat methods for obtaining derivative and integral control actions. Throughout the discussions, we shall place emphasis on the

fundamental principles, rather than on the details of the operation of the actual mechanisms.

Pneumatic Systems. The past decades have seen a great development in low-pressure pneumatic controllers for industrial control systems, and today they are used extensively in industrial processes. Reasons for their broad appeal include an explosion-proof character, simplicity, and ease of maintenance.

Resistance and Capacitance of Pressure Systems. Many industrial processes and pneumatic controllers involve the flow of a gas or air through connected pipelines and pressure vessels.

Consider the pressure system shown in Figure 4-4(a). The gas flow through the restriction is a function of the gas pressure difference $p_i - p_o$. Such a pressure system may be characterized in terms of a resistance and a capacitance.

The gas flow resistance R may be defined as follows:

$$R = \frac{\text{change in gas pressure difference, lb}_f/\text{ft}^2}{\text{change in gas flow rate, lb/sec}}$$

or

$$R = \frac{d(\Delta P)}{dq} \tag{4-8}$$

where $d(\Delta P)$ is a small change in the gas pressure difference and dq is a small change in the gas flow rate. Computation of the value of the gas flow resistance R may be quite time consuming. Experimentally, however, it can be easily determined from a plot of the pressure difference versus flow rate by calculating the slope of the curve at a given operating condition, as shown in Figure 4-4(b).

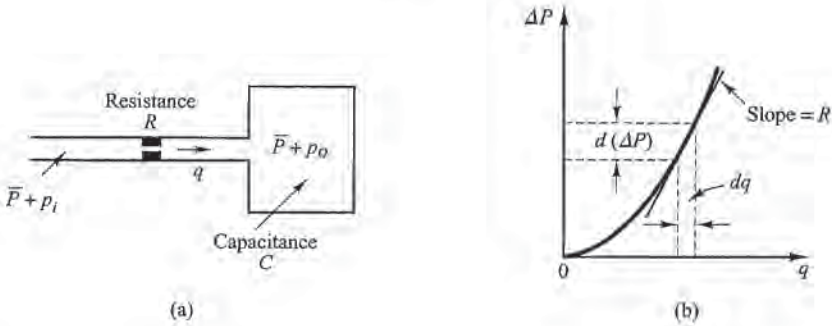
The capacitance of the pressure vessel may be defined by

$$C = \frac{\text{change in gas stored, lb}}{\text{change in gas pressure, lb}_f/\text{ft}^2}$$

or

$$C = \frac{dm}{dp} = V \frac{dp}{dp} \tag{4-9}$$

Figure 4-4
(a) Schematic diagram of a pressure system;
(b) pressure difference versus flow rate curve.



where C = capacitance, lb-ft²/lb_f
 m = mass of gas in vessel, lb
 p = gas pressure, lb_f/ft²
 V = volume of vessel, ft³
 ρ = density, lb/ft³

The capacitance of the pressure system depends on the type of expansion process involved. The capacitance can be calculated by use of the ideal gas law. If the gas expansion process is polytropic and the change of state of the gas is between isothermal and adiabatic, then

$$p\left(\frac{V}{m}\right)^n = \frac{p}{\rho^n} = \text{constant} = K \quad (4-10)$$

where n = polytropic exponent.
For ideal gases,

$$p\bar{v} = \bar{R}T \quad \text{or} \quad pv = \frac{\bar{R}}{M}T$$

where p = absolute pressure, lb_f/ft²
 $\bar{v}$ = volume occupied by 1 mole of a gas, ft³/lb-mole
 $\bar{R}$ = universal gas constant, ft-lb_f/lb-mole °R
 T = absolute temperature, °R
 v = specific volume of gas, ft³/lb
 M = molecular weight of gas per mole, lb/lb-mole

Thus

$$pv = \frac{p}{\rho} = \frac{\bar{R}}{M}T = R_{\text{gas}}T \quad (4-11)$$

where R_{gas} = gas constant, ft-lb_f/lb °R.

The polytropic exponent n is unity for isothermal expansion. For adiabatic expansion, n is equal to the ratio of specific heats c_p/c_v , where c_p is the specific heat at constant pressure and c_v is the specific heat at constant volume. In many practical cases, the value of n is approximately constant, and thus the capacitance may be considered constant.

The value of $d\rho/dp$ is obtained from Equations (4-10) and (4-11). From Equation (4-10) we have

$$dp = Kn\rho^{n-1} d\rho$$

or

$$\frac{d\rho}{dp} = \frac{1}{Kn\rho^{n-1}} = \frac{\rho^n}{pn\rho^{n-1}} = \frac{\rho}{pn}$$

Substituting Equation (4-11) into this last equation, we get

$$\frac{d\rho}{dp} = \frac{1}{nR_{\text{gas}}T}$$

The capacitance C is then obtained as

$$C = \frac{V}{nR_{\text{gas}}T} \quad (4-12)$$

The capacitance of a given vessel is constant if the temperature stays constant. (In many practical cases, the polytropic exponent n is approximately 1.0 ~ 1.2 for gases in uninsulated metal vessels.)

Pressure Systems. Consider the system shown in Figure 4-4(a). If we assume only small deviations in the variables from their respective steady-state values, then this system may be considered linear.

Let us define

- $\bar{P}$ = gas pressure in the vessel at steady state (before changes in pressure have occurred), lb_f/ft²
- p_i = small change in inflow gas pressure, lb_f/ft²
- p_o = small change in gas pressure in the vessel, lb_f/ft²
- V = volume of the vessel, ft³
- m = mass of gas in vessel, lb
- q = gas flow rate, lb/sec
- ρ = density of gas, lb/ft³

For small values of p_i and p_o , the resistance R given by Equation (4-8) becomes constant and may be written as

$$R = \frac{p_i - p_o}{q}$$

The capacitance C is given by Equation (4-9), or

$$C = \frac{dm}{dp}$$

Since the pressure change dp_o times the capacitance C is equal to the gas added to the vessel during dt seconds, we obtain

$$C dp_o = q dt$$

or

$$C \frac{dp_o}{dt} = \frac{p_i - p_o}{R}$$

which can be written as

$$RC \frac{dp_o}{dt} + p_o = p_i$$

If p_i and p_o are considered the input and output, respectively, then the transfer function of the system is

$$\frac{P_o(s)}{P_i(s)} = \frac{1}{RCs + 1}$$

where RC has the dimension of time and is the time constant of the system.

Pneumatic Nozzle-Flapper Amplifiers. A schematic diagram of a pneumatic nozzle-flapper amplifier is shown in Figure 4-5(a). The power source for this amplifier is a supply of air at constant pressure. The nozzle-flapper amplifier converts small changes in the position of the flapper into large changes in the back pressure in the nozzle. Thus a large power output can be controlled by the very little power that is needed to position the flapper.

In Figure 4-5(a), pressurized air is fed through the orifice, and the air is ejected from the nozzle toward the flapper. Generally, the supply pressure P_s for such a controller is 20 psig (1.4 kgf/cm² gage). The diameter of the orifice is on the order of 0.01 in. (0.25 mm) and that of the nozzle is on the order of 0.016 in. (0.4 mm). To ensure proper functioning of the amplifier, the nozzle diameter must be larger than the orifice diameter.

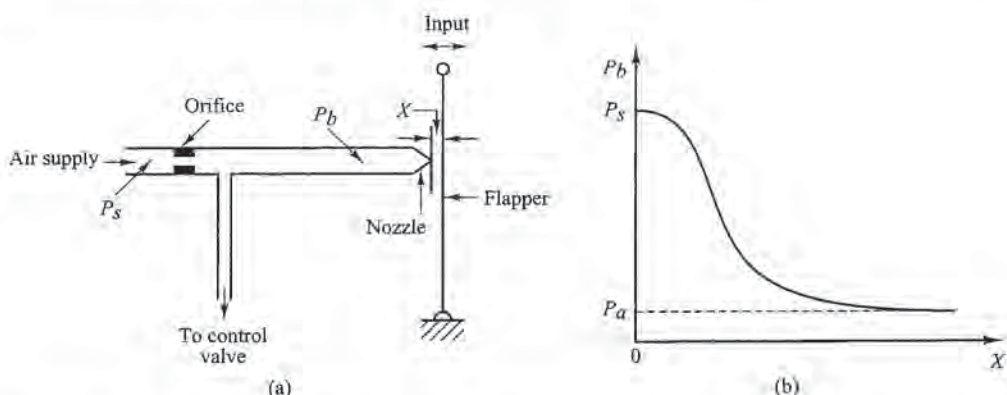
In operating this system, the flapper is positioned against the nozzle opening. The nozzle back pressure P_b is controlled by the nozzle-flapper distance X . As the flapper approaches the nozzle, the opposition to the flow of air through the nozzle increases, with the result that the nozzle back pressure P_b increases. If the nozzle is completely closed by the flapper, the nozzle back pressure P_b becomes equal to the supply pressure P_s . If the flapper is moved away from the nozzle, so that the nozzle-flapper distance is wide (on the order of 0.01 in.), then there is practically no restriction to flow, and the nozzle back pressure P_b takes on a minimum value that depends on the nozzle-flapper device. (The lowest possible pressure will be the ambient pressure P_a .)

Note that, because the air jet puts a force against the flapper, it is necessary to make the nozzle diameter as small as possible.

A typical curve relating the nozzle back pressure P_b to the nozzle-flapper distance X is shown in Figure 4-5(b). The steep and almost linear part of the curve is utilized in the actual operation of the nozzle-flapper amplifier. Because the range of flapper displacements is restricted to a small value, the change in output pressure is also small, unless the curve is very steep.

The nozzle-flapper amplifier converts displacement into a pressure signal. Since industrial process control systems require large output power to operate large pneumatic actuating valves, the power amplification of the nozzle-flapper amplifier is usually insufficient. Consequently, a pneumatic relay is often needed as a power amplifier in connection with the nozzle-flapper amplifier.

Figure 4-5
(a) Schematic diagram of a pneumatic nozzle-flapper amplifier;
(b) characteristic curve relating nozzle back pressure and nozzle-flapper distance.



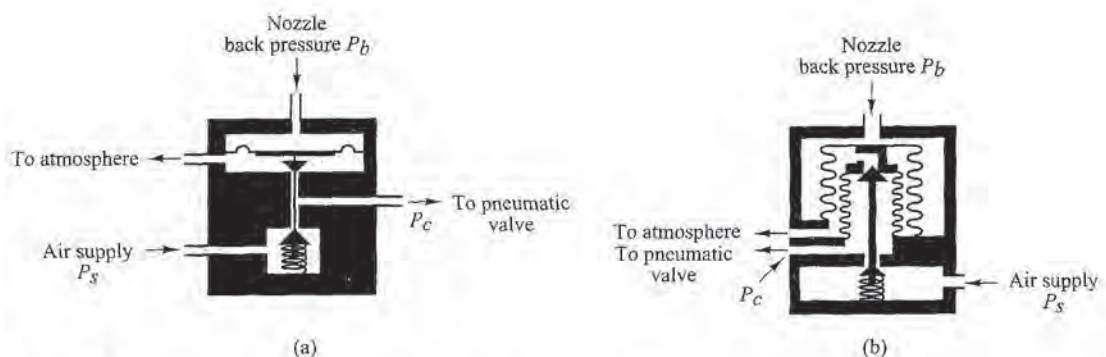


Figure 4-6
(a) Schematic diagram of a bleed-type relay; (b) schematic diagram of a nonbleed-type relay.

Pneumatic Relays. In practice, in a pneumatic controller, a nozzle-flapper amplifier acts as the first-stage amplifier and a pneumatic relay as the second-stage amplifier. The pneumatic relay is capable of handling a large quantity of airflow.

A schematic diagram of a pneumatic relay is shown in Figure 4-6(a). As the nozzle back pressure P_b increases, the diaphragm valve moves downward. The opening to the atmosphere decreases and the opening to the pneumatic valve increases, thereby increasing the control pressure P_c . When the diaphragm valve closes the opening to the atmosphere, the control pressure P_c becomes equal to the supply pressure P_s . When the nozzle back pressure P_b decreases and the diaphragm valve moves upward and shuts off the air supply, the control pressure P_c drops to the ambient pressure P_a . The control pressure P_c can thus be made to vary from 0 psig to full supply pressure, usually 20 psig.

The total movement of the diaphragm valve is very small. In all positions of the valve, except at the position to shut off the air supply, air continues to bleed into the atmosphere, even after the equilibrium condition is attained between the nozzle back pressure and the control pressure. Thus the relay shown in Figure 4-6(a) is called a bleed-type relay.

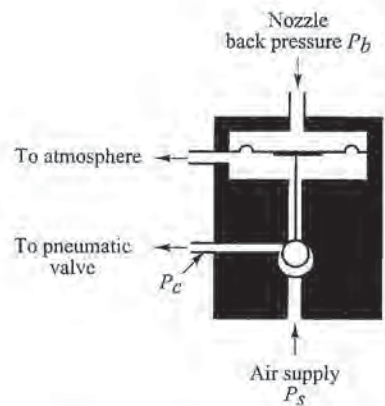
There is another type of relay, the nonbleed type. In this one the air bleed stops when the equilibrium condition is obtained and, therefore, there is no loss of pressurized air at steady-state operation. Note, however, that the nonbleed-type relay must have an atmospheric relief to release the control pressure P_c from the pneumatic actuating valve. A schematic diagram of a nonbleed-type relay is shown in Figure 4-6(b).

In either type of relay, the air supply is controlled by a valve, which is in turn controlled by the nozzle back pressure. Thus, the nozzle back pressure is converted into the control pressure with power amplification.

Since the control pressure P_c changes almost instantaneously with changes in the nozzle back pressure P_b , the time constant of the pneumatic relay is negligible compared with the other larger time constants of the pneumatic controller and the plant.

It is noted that some pneumatic relays are reverse acting. For example, the relay shown in Figure 4-7 is a reverse-acting relay. Here, as the nozzle back pressure P_b increases, the ball valve is forced toward the lower seat, thereby decreasing the control pressure P_c . Thus, this relay is a reverse-acting relay.

Figure 4-7
Reverse-acting relay.



Pneumatic Proportional Controllers (Force-Distance Type). Two types of pneumatic controllers, one called the force-distance type and the other the force-balance type, are used extensively in industry. Regardless of how differently industrial pneumatic controllers may appear, careful study will show the close similarity in the functions of the pneumatic circuit. Here we shall consider the force-distance type of pneumatic controllers.

Figure 4-8(a) shows a schematic diagram of such a proportional controller. The nozzle-flapper amplifier constitutes the first-stage amplifier, and the nozzle back pressure is controlled by the nozzle-flapper distance. The relay-type amplifier constitutes the second-stage amplifier. The nozzle back pressure determines the position of the diaphragm valve for the second-stage amplifier, which is capable of handling a large quantity of airflow.

In most pneumatic controllers, some type of pneumatic feedback is employed. Feedback of the pneumatic output reduces the amount of actual movement of the flapper. Instead of mounting the flapper on a fixed point, as shown in Figure 4-8(b), it is often pivoted on the feedback bellows, as shown in Figure 4-8(c). The amount of feedback can be regulated by introducing a variable linkage between the feedback bellows and the flapper connecting point. The flapper then becomes a floating link. It can be moved by both the error signal and the feedback signal.

The operation of the controller shown in Figure 4-8(a) is as follows. The input signal to the two-stage pneumatic amplifier is the actuating error signal. Increasing the actuating error signal moves the flapper to the left. This move will, in turn, increase the nozzle back pressure, and the diaphragm valve moves downward. This results in an increase of the control pressure. This increase will cause bellows *F* to expand and move the flapper to the right, thus opening the nozzle. Because of this feedback, the nozzle-flapper displacement is very small, but the change in the control pressure can be large.

It should be noted that proper operation of the controller requires that the feedback bellows move the flapper less than that movement caused by the error signal alone. (If these two movements were equal, no control action would result.)

Equations for this controller can be derived as follows. When the actuating error is zero, or $e = 0$, an equilibrium state exists with the nozzle-flapper distance equal to $\bar{X}$, the displacement of bellows equal to $\bar{Y}$, the displacement of the diaphragm equal to $\bar{Z}$, the nozzle back pressure equal to $\bar{P}_b$, and the control pressure equal to $\bar{P}_c$. When an

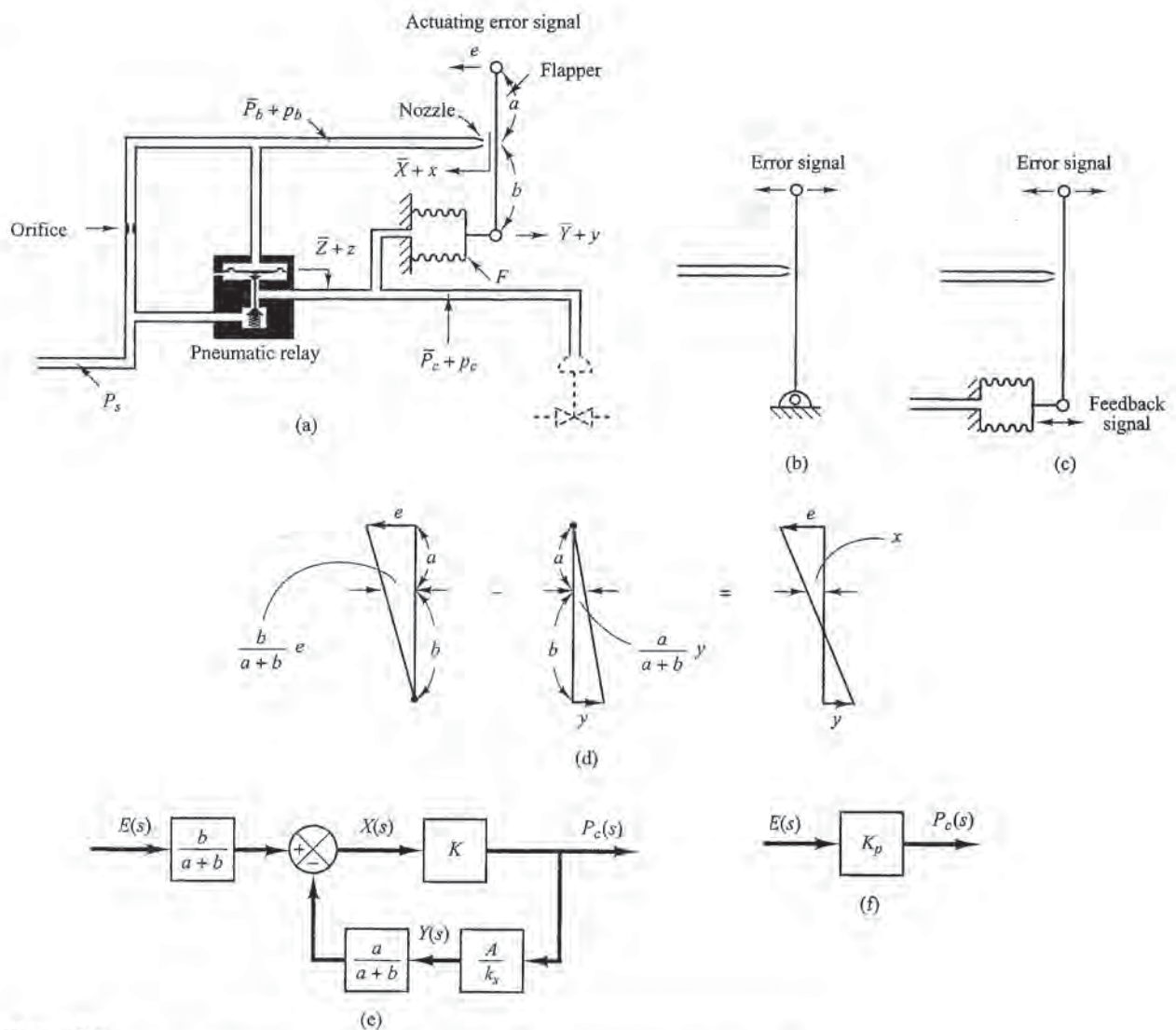


Figure 4-8
 (a) Schematic diagram of a force-distance type of pneumatic proportional controller;
 (b) flapper mounted on a fixed point; (c) flapper mounted on a feedback bellows;
 (d) displacement x as a result of addition of two small displacements;
 (e) block diagram for the controller; (f) simplified block diagram for the controller.

actuating error exists, the nozzle-flapper distance, the displacement of the bellows, the displacement of the diaphragm, the nozzle back pressure, and the control pressure deviate from their respective equilibrium values. Let these deviations be x , y , z , p_b , and p_c , respectively. (The positive direction for each displacement variable is indicated by an arrowhead in the diagram.)

Assuming that the relationship between the variation in the nozzle back pressure and the variation in the nozzle-flapper distance is linear, we have

$$p_b = K_1 x \quad (4-13)$$

where K_1 is a positive constant. For the diaphragm valve,

$$p_b = K_2 z \quad (4-14)$$

where K_2 is a positive constant. The position of the diaphragm valve determines the control pressure. If the diaphragm valve is such that the relationship between p_c and z is linear, then

$$p_c = K_3 z \quad (4-15)$$

where K_3 is a positive constant. From Equations (4-13), (4-14), and (4-15), we obtain

$$p_c = \frac{K_3}{K_2} p_b = \frac{K_1 K_3}{K_2} x = K x \quad (4-16)$$

where $K = K_1 K_3 / K_2$ is a positive constant. For the flapper, since there are two small movements (e and y) in opposite directions, we can consider such movements separately and add up the results of two movements into one displacement x . See Figure 4-8(d). Thus, for the flapper movement, we have

$$x = \frac{b}{a+b} e - \frac{a}{a+b} y \quad (4-17)$$

The bellows acts like a spring, and the following equation holds true:

$$A p_c = k_s y \quad (4-18)$$

where A is the effective area of the bellows and k_s is the equivalent spring constant, that is, the stiffness due to the action of the corrugated side of the bellows.

Assuming that all variations in the variables are within a linear range, we can obtain a block diagram for this system from Equations (4-16), (4-17), and (4-18) as shown in Figure 4-8(e). From Figure 4-8(e), it can be clearly seen that the pneumatic controller shown in Figure 4-8(a) itself is a feedback system. The transfer function between p_c and e is given by

$$\frac{P_c(s)}{E(s)} = \frac{\frac{b}{a+b} K}{1 + K \frac{a}{a+b} \frac{A}{k_s}} = K_p \quad (4-19)$$

A simplified block diagram is shown in Figure 4-8(f). Since p_c and e are proportional, the pneumatic controller shown in Figure 4-8(a) is a *pneumatic proportional controller*. As seen from Equation (4-19), the gain of the pneumatic proportional controller can be widely varied by adjusting the flapper connecting linkage. [The flapper connecting linkage is not shown in Figure 4-8(a).] In most commercial proportional controllers an adjusting knob or other mechanism is provided for varying the gain by adjusting this linkage.

As noted earlier, the actuating error signal moved the flapper in one direction, and the feedback bellows moved the flapper in the opposite direction, but to a smaller degree.

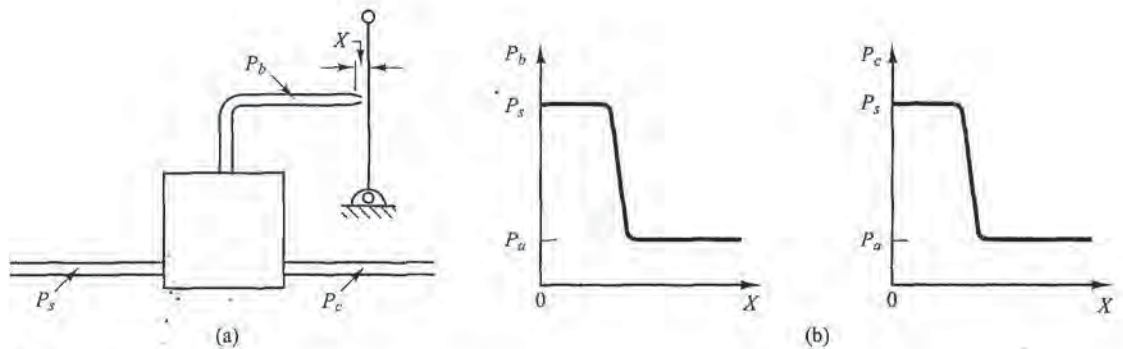


Figure 4-9
(a) Pneumatic controller without a feedback mechanism; (b) curves P_b versus X and P_c versus X .

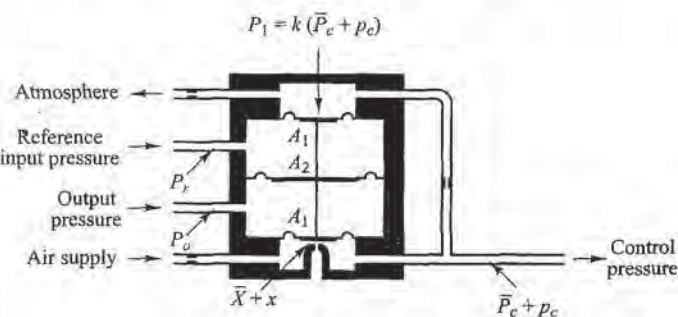
The effect of the feedback bellows is thus to reduce the sensitivity of the controller. The principle of feedback is commonly used to obtain wide proportional-band controllers.

Pneumatic controllers that do not have feedback mechanisms [which means that one end of the flapper is fixed, as shown in Figure 4-9(a)] have high sensitivity and are called *pneumatic two-position controllers* or *pneumatic on-off controllers*. In such a controller, only a small motion between the nozzle and the flapper is required to give a complete change from the maximum to the minimum control pressure. The curves relating P_b to X and P_c to X are shown in Figure 4-9(b). Notice that a small change in X can cause a large change in P_b , which causes the diaphragm valve to be completely open or completely closed.

Pneumatic Proportional Controllers (Force-Balance Type). Figure 4-10 shows a schematic diagram of a force-balance pneumatic proportional controller. Force-balance controllers are in extensive use in industry. Such controllers are called *stack controllers*. The basic principle of operation does not differ from that of the force-distance controller. The main advantage of the force-balance controller is that it eliminates many mechanical linkages and pivot joints, thereby reducing the effects of friction.

In what follows, we shall consider the principle of the force-balance controller. In the controller shown in Figure 4-10, the reference input pressure P_r and the output pressure P_o are fed to large diaphragm chambers. Note that a force-balance pneumatic controller operates only on pressure signals. Therefore, it is necessary to convert the reference input and system output to corresponding pressure signals.

Figure 4-10
Schematic diagram of a force-balance type of pneumatic proportional controller.



As in the case of the force-distance controller, this controller employs a flapper, nozzle, and orifices. In Figure 4-10, the drilled opening in the bottom chamber is the nozzle. The diaphragm just above the nozzle acts as a flapper.

The operation of the force-balance controller shown in Figure 4-10 may be summarized as follows: 20-psig air from an air supply flows through an orifice, causing a reduced pressure in the bottom chamber. Air in this chamber escapes to the atmosphere through the nozzle. The flow through the nozzle depends on the gap and the pressure drop across it. An increase in the reference input pressure P_r , while the output pressure P_o remains the same, causes the valve stem to move down, decreasing the gap between the nozzle and the flapper diaphragm. This causes the control pressure P_c to increase. Let

$$p_e = P_r - P_o \quad (4-20)$$

If $p_e = 0$, there is an equilibrium state with the nozzle-flapper distance equal to $\bar{X}$ and the control pressure equal to $\bar{P}_c$. At this equilibrium state, $P_1 = \bar{P}_c k$ (where $k < 1$) and

$$\bar{X} = \alpha(\bar{P}_c A_1 - \bar{P}_c k A_1) \quad (4-21)$$

where α is a constant.

Let us assume that $p_e \neq 0$ and define small variations in the nozzle-flapper distance and control pressure as x and p_c , respectively. Then we obtain the following equation:

$$\bar{X} + x = \alpha[(\bar{P}_c + p_c)A_1 - (\bar{P}_c + p_c)kA_1 - p_e(A_2 - A_1)] \quad (4-22)$$

From Equations (4-21) and (4-22), we obtain

$$x = \alpha[p_c(1 - k)A_1 - p_e(A_2 - A_1)] \quad (4-23)$$

At this point, we must examine the quantity x . In the design of pneumatic controllers, the nozzle-flapper distance is made quite small. In view of the fact that x/α is very much smaller than $p_c(1 - k)A_1$ or $p_e(A_2 - A_1)$, that is, for $p_e \neq 0$

$$\frac{x}{\alpha} \ll p_c(1 - k)A_1$$

$$\frac{x}{\alpha} \ll p_e(A_2 - A_1)$$

we may neglect the term x in our analysis. Equation (4-23) can then be rewritten to reflect this assumption as follows:

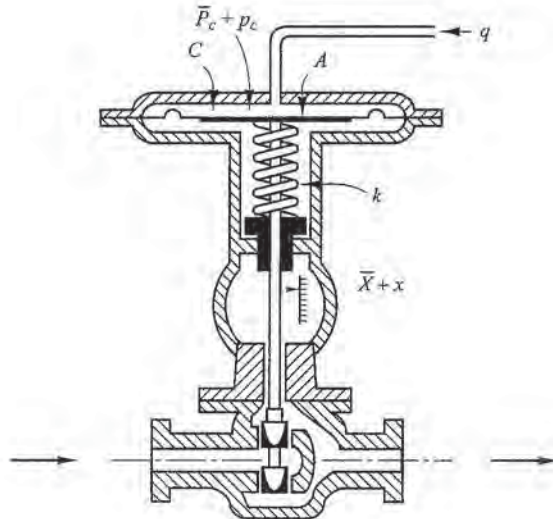
$$p_c(1 - k)A_1 = p_e(A_2 - A_1)$$

and the transfer function between p_c and p_e becomes

$$\frac{P_c(s)}{P_e(s)} = \frac{A_2 - A_1}{A_1} \frac{1}{1 - k} = K_p$$

where p_e is defined by Equation (4-20). The controller shown in Figure 4-10 is a proportional controller. The value of gain K_p increases as k approaches unity. Note that the value of k depends on the diameters of the orifices in the inlet and outlet pipes of the feedback chamber. (The value of k approaches unity as the resistance to flow in the orifice of the inlet pipe is made smaller.)

Figure 4-11
Schematic diagram
of a pneumatic
actuating valve.



Pneumatic Actuating Valves. One characteristic of pneumatic controls is that they almost exclusively employ pneumatic actuating valves. A pneumatic actuating valve can provide a large power output. (Since a pneumatic actuator requires a large power input to produce a large power output, it is necessary that a sufficient quantity of pressurized air be available.) In practical pneumatic actuating valves, the valve characteristics may not be linear; that is, the flow may not be directly proportional to the valve stem position, and also there may be other nonlinear effects, such as hysteresis.

Consider the schematic diagram of a pneumatic actuating valve shown in Figure 4-11. Assume that the area of the diaphragm is A . Assume also that when the actuating error is zero the control pressure is equal to $\bar{P}_c$ and the valve displacement is equal to $\bar{X}$.

In the following analysis, we shall consider small variations in the variables and linearize the pneumatic actuating valve. Let us define the small variation in the control pressure and the corresponding valve displacement to be p_c and x , respectively. Since a small change in the pneumatic pressure force applied to the diaphragm repositions the load, consisting of the spring, viscous friction, and mass, the force balance equation becomes

$$Ap_c = m\ddot{x} + b\dot{x} + kx$$

where m = mass of the valve and valve stem

b = viscous-friction coefficient

k = spring constant

If the force due to the mass and viscous friction are negligibly small, then this last equation can be simplified to

$$Ap_c = kx$$

The transfer function between x and p_c thus becomes

$$\frac{X(s)}{P_c(s)} = \frac{A}{k} = K_c$$

where $X(s) = \mathcal{L}[x]$ and $P_c(s) = \mathcal{L}[p_c]$. If q_i , the change in flow through the pneumatic actuating valve, is proportional to x , the change in the valve-stem displacement, then

$$\frac{Q_i(s)}{X(s)} = K_q$$

where $Q_i(s) = \mathcal{L}[q_i]$ and K_q is a constant. The transfer function between q_i and p_c becomes

$$\frac{Q_i(s)}{P_c(s)} = K_c K_q = K_v$$

where K_v is a constant.

The standard control pressure for this kind of a pneumatic actuating valve is between 3 and 15 psig. The valve-stem displacement is limited by the allowable stroke of the diaphragm and is only a few inches. If a longer stroke is needed, a piston-spring combination may be employed.

In pneumatic actuating valves, the static-friction force must be limited to a low value so that excessive hysteresis does not result. Because of the compressibility of air, the control action may not be positive; that is, an error may exist in the valve-stem position. The use of a valve positioner results in improvements in the performance of a pneumatic actuating valve.

Basic Principle for Obtaining Derivative Control Action. We shall now present methods for obtaining derivative control action. We shall again place the emphasis on the principle and not on the details of the actual mechanisms.

The basic principle for generating a desired control action is to insert the inverse of the desired transfer function in the feedback path. For the system shown in Figure 4-12, the closed-loop transfer function is

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

If $|G(s)H(s)| \gg 1$, then $C(s)/R(s)$ can be modified to

$$\frac{C(s)}{R(s)} = \frac{1}{H(s)}$$

Thus, if proportional-plus-derivative control action is desired, we insert an element having the transfer function $1/(Ts + 1)$ in the feedback path.

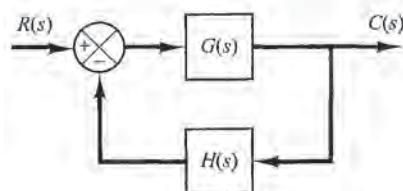


Figure 4-12
Control system.

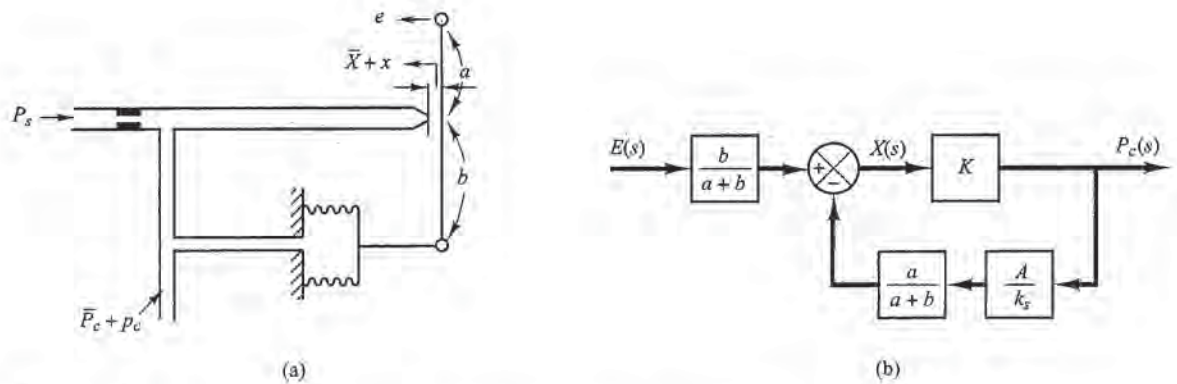


Figure 4-13
(a) Pneumatic proportional controller; (b) block diagram of the controller.

Consider the pneumatic controller shown in Figure 4-13(a). Considering small changes in the variables, we can draw a block diagram of this controller as shown in Figure 4-13(b). From the block diagram we see that the controller is of proportional type.

We shall now show that the addition of a restriction in the negative feedback path will modify the proportional controller to a proportional-plus-derivative controller, or a PD controller.

Consider the pneumatic controller shown in Figure 4-14(a). Assuming again small changes in the actuating error, nozzle-flapper distance, and control pressure, we can summarize the operation of this controller as follows: Let us first assume a small step change in e . Then the change in the control pressure p_c will be instantaneous. The restriction R

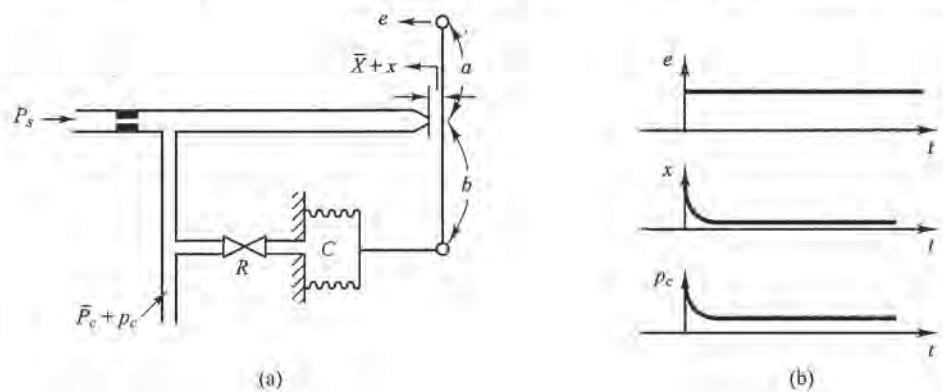
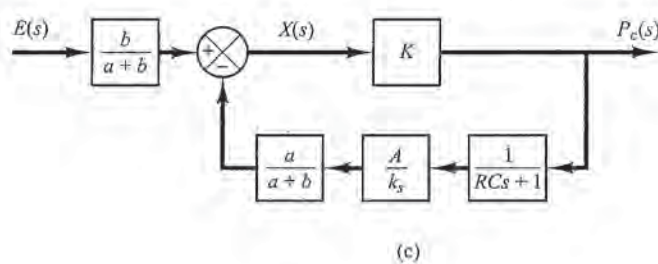


Figure 4-14
(a) Pneumatic proportional-plus-derivative controller; (b) step change in e and the corresponding changes in x and p_c plotted versus t ; (c) block diagram of the controller.



will momentarily prevent the feedback bellows from sensing the pressure change p_c . Thus the feedback bellows will not respond momentarily, and the pneumatic actuating valve will feel the full effect of the movement of the flapper. As time goes on, the feedback bellows will expand. The change in the nozzle-flapper distance x and the change in the control pressure p_c can be plotted against time t , as shown in Figure 4-14(b). At steady state, the feedback bellows acts like an ordinary feedback mechanism. The curve p_c versus t clearly shows that this controller is of the proportional plus-derivative type.

A block diagram corresponding to this pneumatic controller is shown in Figure 4-14(c). In the block diagram, K is a constant, A is the area of the bellows, and k_s is the equivalent spring constant of the bellows. The transfer function between p_c and e can be obtained from the block diagram as follows:

$$\frac{P_c(s)}{E(s)} = \frac{\frac{b}{a+b} K}{1 + \frac{Ka}{a+b} \frac{A}{k_s} \frac{1}{RCs + 1}}$$

In such a controller the loop gain $|KaA/[(a+b)k_s(RCs + 1)]|$ is made much greater than unity. Thus the transfer function $P_c(s)/E(s)$ can be simplified to give

$$\frac{P_c(s)}{E(s)} = K_p(1 + T_d s)$$

where

$$K_p = \frac{bk_s}{aA}, \quad T_d = RC$$

Thus, delayed negative feedback, or the transfer function $1/(RCs + 1)$ in the feedback path, modifies the proportional controller to a proportional-plus-derivative controller.

Note that if the feedback valve is fully opened the control action becomes proportional. If the feedback valve is fully closed, the control action becomes narrow-band proportional (on-off).

Obtaining Pneumatic Proportional-Plus-Integral Control Action. Consider the proportional controller shown in Figure 4-13(a). Considering small changes in the variables, we can show that the addition of delayed positive feedback will modify this proportional controller to a proportional-plus-integral controller, or a PI controller.

Consider the pneumatic controller shown in Figure 4-15(a). The operation of this controller is as follows: The bellows denoted by I is connected to the control pressure source without any restriction. The bellows denoted by II is connected to the control pressure source through a restriction. Let us assume a small step change in the actuating error. This will cause the back pressure in the nozzle to change instantaneously. Thus a change in the control pressure p_c also occurs instantaneously. Due to the restriction of the valve in the path to bellows II, there will be a pressure drop across the valve. As time goes on, air will flow across the valve in such a way that the change in pressure in bellows II attains the value p_c . Thus bellows II will expand or contract as time elapses in such a way as to move the flapper an additional amount in the direction of the original displacement e . This will cause the back pressure p_c in the nozzle to change continuously, as shown in Figure 4-15(b).

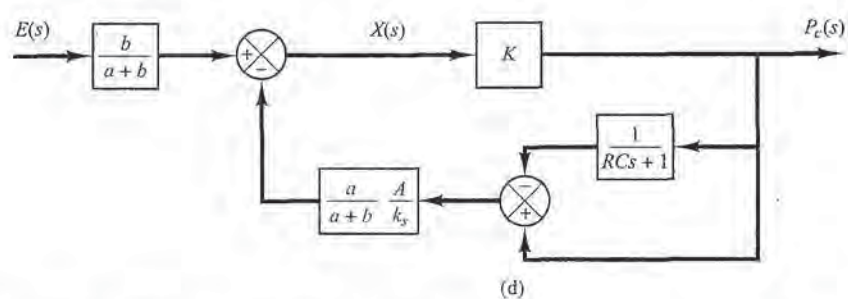
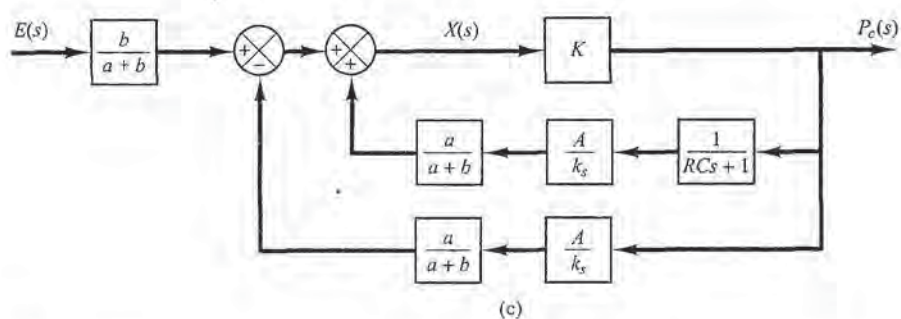
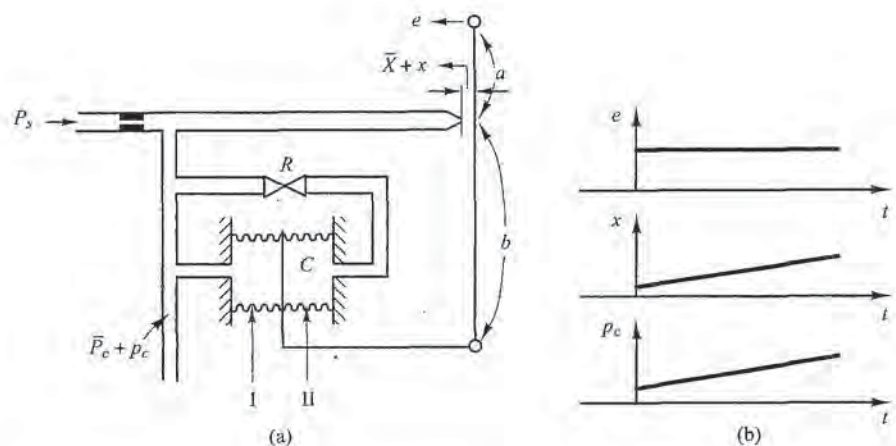


Figure 4-15
(a) Pneumatic proportional-plus-integral controller; (b) step change in e and the corresponding changes in x and p_c plotted versus t ; (c) block diagram of the controller; (d) simplified block diagram.

Note that the integral control action in the controller takes the form of slowly canceling the feedback that the proportional control originally provided.

A block diagram of this controller under the assumption of small variations in the variables is shown in Figure 4-15(c). A simplification of this block diagram yields Figure 4-15(d). The transfer function of this controller is

$$\frac{P_c(s)}{E(s)} = \frac{\frac{b}{a+b} K}{1 + \frac{Ka}{a+b} \frac{A}{k_x} \left(1 - \frac{1}{RCs+1}\right)}$$

where K is a constant, A is the area of the bellows, and k_s is the equivalent spring constant of the combined bellows. If $|KaARCs/[(a + b)k_s(RCs + 1)]| \gg 1$, which is usually the case, the transfer function can be simplified to

$$\frac{P_c(s)}{E(s)} = K_p \left(1 + \frac{1}{T_i s} \right)$$

where

$$K_p = \frac{bk_s}{aA}, \quad T_i = RC$$

Obtaining Pneumatic Proportional-Plus-Integral-Plus-Derivative Control Action. A combination of the pneumatic controllers shown in Figures 4-14(a) and 4-15(a) yields a proportional-plus-integral-plus-derivative controller, or a PID controller. Figure 4-16(a) shows a schematic diagram of such a controller. Figure 4-16(b) shows a block diagram of this controller under the assumption of small variations in the variables.

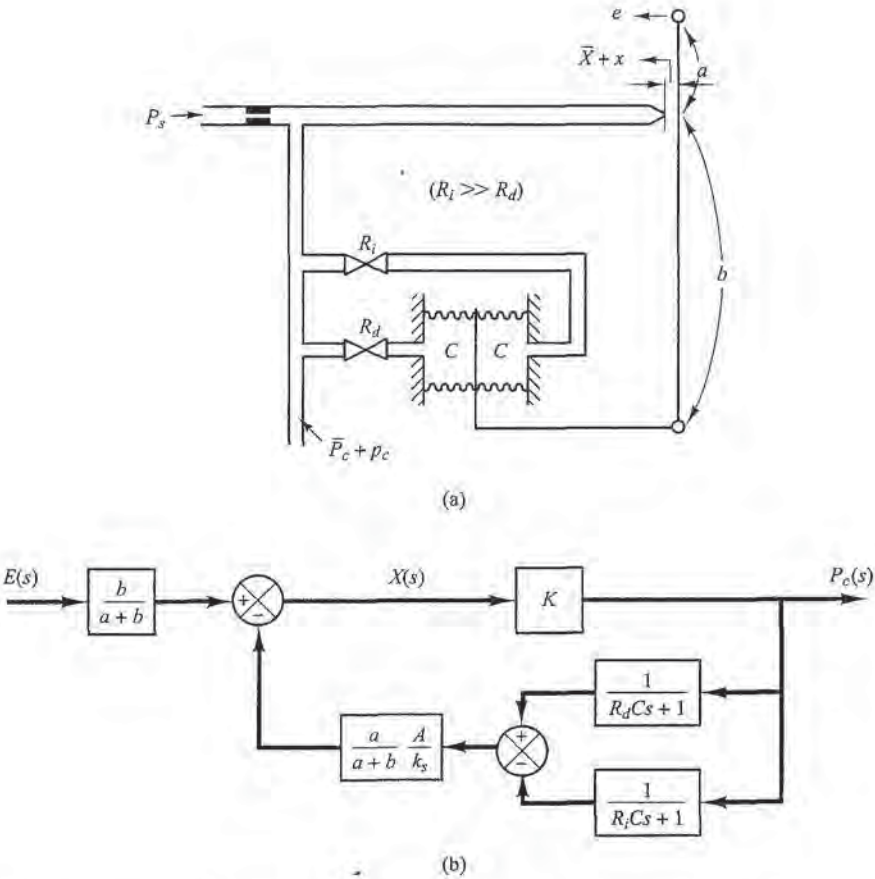


Figure 4-16
(a) Pneumatic proportional-plus-integral-plus-derivative controller; (b) block diagram of the controller.

The transfer function of this controller is

$$\frac{P_c(s)}{E(s)} = \frac{\frac{bK}{a+b}}{1 + \frac{Ka}{a+b} \frac{A}{k_s} \frac{(R_i C - R_d C)s}{(R_d C s + 1)(R_i C s + 1)}}$$

By defining

$$T_i = R_i C, \quad T_d = R_d C$$

and noting that under normal operation $|K a A (T_i - T_d) s / [(a + b) k_s (T_d s + 1)(T_i s + 1)]| \gg 1$ and $T_i \gg T_d$, we obtain

$$\begin{aligned} \frac{P_c(s)}{E(s)} &\doteq \frac{b k_s (T_d s + 1)(T_i s + 1)}{a A (T_i - T_d) s} \\ &\doteq \frac{b k_s T_d T_i s^2 + T_i s + 1}{a A T_i s} \\ &= K_p \left(1 + \frac{1}{T_i s} + T_d s \right) \end{aligned} \quad (4-24)$$

where

$$K_p = \frac{b k_s}{a A}$$

Equation (4-24) indicates that the controller shown in Figure 4-16(a) is a proportional-plus-integral-plus-derivative controller or a PID controller.

4-4 HYDRAULIC SYSTEMS

Except for low-pressure pneumatic controllers, compressed air has seldom been used for the continuous control of the motion of devices having significant mass under external load forces. For such a case, hydraulic controllers are generally preferred.

Hydraulic Systems. The widespread use of hydraulic circuitry in machine tool applications, aircraft control systems, and similar operations occurs because of such factors as positiveness, accuracy, flexibility, high horsepower-to-weight ratio, fast starting, stopping, and reversal with smoothness and precision, and simplicity of operations.

The operating pressure in hydraulic systems is somewhere between 145 and 5000 lb_f/in.² (between 1 and 35 MPa). In some special applications, the operating pressure may go up to 10,000 lb_f/in.² (70 MPa). For the same power requirement, the weight and size of the hydraulic unit can be made smaller by increasing the supply pressure. With high-pressure hydraulic systems, very large force can be obtained. Rapid-acting, accurate positioning of heavy loads is possible with hydraulic systems. A combination of electronic and hydraulic systems is widely used because it combines the advantages of both electronic control and hydraulic power.

Advantages and Disadvantages of Hydraulic Systems. There are certain advantages and disadvantages in using hydraulic systems rather than other systems. Some of the advantages are the following:

1. Hydraulic fluid acts as a lubricant, in addition to carrying away heat generated in the system to a convenient heat exchanger.
2. Comparatively small sized hydraulic actuators can develop large forces or torques.
3. Hydraulic actuators have a higher speed of response with fast starts, stops, and speed reversals.
4. Hydraulic actuators can be operated under continuous, intermittent, reversing, and stalled conditions without damage.
5. Availability of both linear and rotary actuators gives flexibility in design.
6. Because of low leakages in hydraulic actuators, speed drop when loads are applied is small.

On the other hand, several disadvantages tend to limit their use.

1. Hydraulic power is not readily available compared to electric power.
2. Cost of a hydraulic system may be higher than a comparable electrical system performing a similar function.
3. Fire and explosion hazards exist unless fire-resistant fluids are used.
4. Because it is difficult to maintain a hydraulic system that is free from leaks, the system tends to be messy.
5. Contaminated oil may cause failure in the proper functioning of a hydraulic system.
6. As a result of the nonlinear and other complex characteristics involved, the design of sophisticated hydraulic systems is quite involved.
7. Hydraulic circuits have generally poor damping characteristics. If a hydraulic circuit is not designed properly, some unstable phenomena may occur or disappear, depending on the operating condition.

Comments. Particular attention is necessary to ensure that the hydraulic system is stable and satisfactory under all operating conditions. Since the viscosity of hydraulic fluid can greatly affect damping and friction effects of the hydraulic circuits, stability tests must be carried out at the highest possible operating temperature.

Note that most hydraulic systems are nonlinear. Sometimes, however, it is possible to linearize nonlinear systems so as to reduce their complexity and permit solutions that are sufficiently accurate for most purposes. A useful linearization technique for dealing with nonlinear systems was presented in Section 3–10.

Hydraulic Servo System. Figure 4–17(a) shows a hydraulic servomotor. It is essentially a pilot-valve-controlled hydraulic power amplifier and actuator. The pilot valve is a balanced valve, in the sense that the pressure forces acting on it are all balanced. A very large power output can be controlled by a pilot valve, which can be positioned with very little power.

In practice, the ports shown in Figure 4–17(a) are often made wider than the corresponding valves. In such a case, there is always leakage through the valves. Such leak-

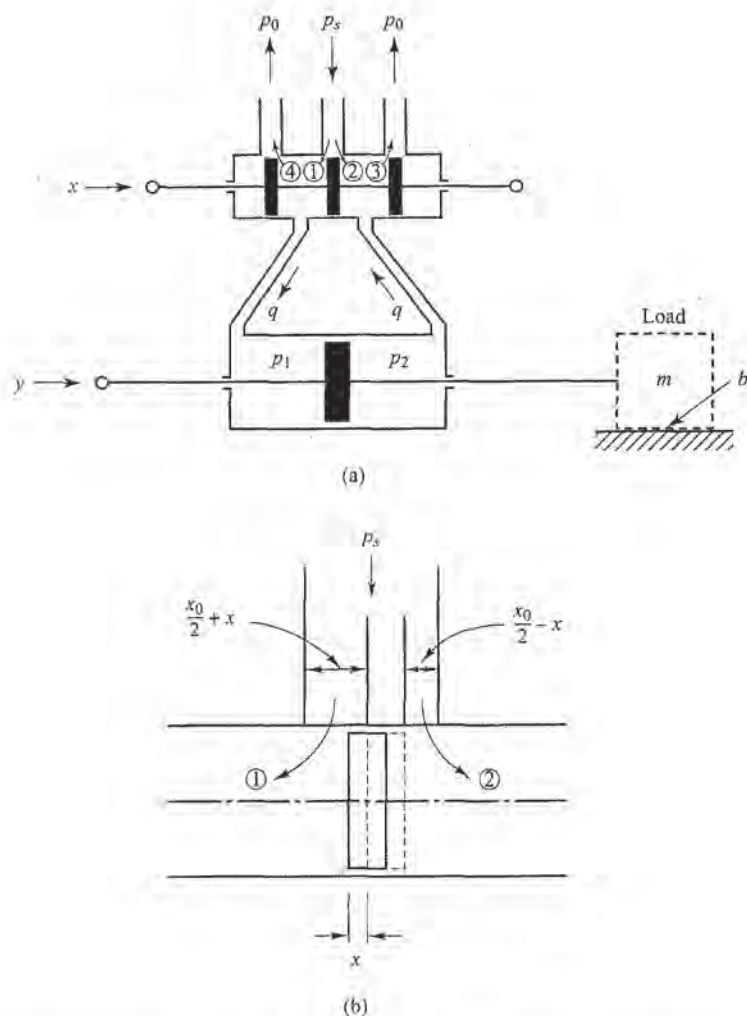


Figure 4-17
(a) Hydraulic servo system; (b) enlarged diagram of the valve orifice area.

age improves both the sensitivity and the linearity of the hydraulic servomotor. In the following analysis we shall make the assumption that the ports are made wider than the valves, that is, the valves are underlapped. [Note that sometimes a dither signal, a high-frequency signal of very small amplitude (with respect to the maximum displacement of the valve), is superimposed on the motion of the pilot valve. This also improves the sensitivity and linearity. In this case also there is leakage through the valve.]

We shall apply the linearization technique presented in Section 3-10 to obtain a linearized mathematical model of the hydraulic servomotor. We assume that the valve is underlapped and symmetrical and admits hydraulic fluid under high pressure into a power cylinder that contains a large piston, so that a large hydraulic force is established to move a load.

In Figure 4-17(b) we have an enlarged diagram of the valve orifice area. Let us define the valve orifice areas of ports 1, 2, 3, 4 as A_1, A_2, A_3, A_4 , respectively. Also, define the flow rates through ports 1, 2, 3, 4 as q_1, q_2, q_3, q_4 , respectively. Note that, since the

valve is symmetrical, $A_1 = A_3$ and $A_2 = A_4$. Assuming the displacement x to be small, we obtain

$$A_1 = A_3 = k \left(\frac{x_0}{2} + x \right)$$

$$A_2 = A_4 = k \left(\frac{x_0}{2} - x \right)$$

where k is a constant.

Furthermore, we shall assume that the return pressure p_o in the return line is small and thus can be neglected. Then, referring to Figure 4-17(a), flow rates through valve orifices are

$$q_1 = c_1 A_1 \sqrt{\frac{2g}{\gamma} (p_s - p_1)} = C_1 \sqrt{p_s - p_1} \left(\frac{x_0}{2} + x \right)$$

$$q_2 = c_2 A_2 \sqrt{\frac{2g}{\gamma} (p_s - p_2)} = C_2 \sqrt{p_s - p_2} \left(\frac{x_0}{2} - x \right)$$

$$q_3 = c_1 A_3 \sqrt{\frac{2g}{\gamma} (p_2 - p_o)} = C_1 \sqrt{p_2 - p_o} \left(\frac{x_0}{2} + x \right) = C_1 \sqrt{p_2} \left(\frac{x_0}{2} + x \right)$$

$$q_4 = c_2 A_4 \sqrt{\frac{2g}{\gamma} (p_1 - p_o)} = C_2 \sqrt{p_1 - p_o} \left(\frac{x_0}{2} - x \right) = C_2 \sqrt{p_1} \left(\frac{x_0}{2} - x \right)$$

where $C_1 = c_1 k \sqrt{2g/\gamma}$ and $C_2 = c_2 k \sqrt{2g/\gamma}$, and γ is the specific weight and is given by $\gamma = \rho g$, where ρ is mass density and g is the acceleration of gravity. The flow rate q to the left-hand side of the power piston is

$$q = q_1 - q_4 = C_1 \sqrt{p_s - p_1} \left(\frac{x_0}{2} + x \right) - C_2 \sqrt{p_1} \left(\frac{x_0}{2} - x \right) \quad (4-25)$$

The flow rate from the right-hand side of the power piston to the drain is the same as this q and is given by

$$q = q_3 - q_2 = C_1 \sqrt{p_2} \left(\frac{x_0}{2} + x \right) - C_2 \sqrt{p_s - p_2} \left(\frac{x_0}{2} - x \right)$$

In the present analysis we assume that the fluid is incompressible. Since the valve is symmetrical we have $q_1 = q_3$ and $q_2 = q_4$. By equating q_1 and q_3 , we obtain

$$p_s - p_1 = p_2$$

or

$$p_s = p_1 + p_2$$

If we define the pressure difference across the power piston as Δp or

$$\Delta p = p_1 - p_2$$

then

$$p_1 = \frac{p_s + \Delta p}{2}, \quad p_2 = \frac{p_s - \Delta p}{2}$$

For the symmetrical valve shown in Figure 4-17(a), the pressure in each side of the power piston is $(1/2)p_s$, when no load is applied, or $\Delta p = 0$. As the spool valve is displaced, the pressure in one line increases as the pressure in the other line decreases by the same amount.

In terms of p_s and Δp , we can rewrite the flow rate q given by Equation (4-25) as

$$q = q_1 - q_4 = C_1 \sqrt{\frac{p_s - \Delta p}{2}} \left(\frac{x_0}{2} + x \right) - C_2 \sqrt{\frac{p_s + \Delta p}{2}} \left(\frac{x_0}{2} - x \right)$$

Noting that the supply pressure p_s is constant, the flow rate q can be written as a function of the valve displacement x and pressure difference Δp , or

$$q = C_1 \sqrt{\frac{p_s - \Delta p}{2}} \left(\frac{x_0}{2} + x \right) - C_2 \sqrt{\frac{p_s + \Delta p}{2}} \left(\frac{x_0}{2} - x \right) = f(x, \Delta p)$$

By applying the linearization technique presented in Section 3-10 to this case, the linearized equation about point $x = \bar{x}$, $\Delta p = \Delta \bar{p}$, $q = \bar{q}$ is

$$q - \bar{q} = a(x - \bar{x}) + b(\Delta p - \Delta \bar{p}) \quad (4-26)$$

where

$$\bar{q} = f(\bar{x}, \Delta \bar{p})$$

$$a = \left. \frac{\partial f}{\partial x} \right|_{x=\bar{x}, \Delta p=\Delta \bar{p}} = C_1 \sqrt{\frac{p_s - \Delta \bar{p}}{2}} + C_2 \sqrt{\frac{p_s + \Delta \bar{p}}{2}}$$

$$b = \left. \frac{\partial f}{\partial \Delta p} \right|_{x=\bar{x}, \Delta p=\Delta \bar{p}} = - \left[\frac{C_1}{2\sqrt{2}\sqrt{p_s - \Delta \bar{p}}} \left(\frac{x_0}{2} + \bar{x} \right) + \frac{C_2}{2\sqrt{2}\sqrt{p_s + \Delta \bar{p}}} \left(\frac{x_0}{2} - \bar{x} \right) \right] < 0$$

Coefficients a and b here are called *valve coefficients*. Equation (4-26) is a linearized mathematical model of the spool valve near an operating point $x = \bar{x}$, $\Delta p = \Delta \bar{p}$, $q = \bar{q}$. The values of valve coefficients a and b vary with the operating point. Note that $\partial f / \partial \Delta p$ is negative and so b is negative.

Since the normal operating point is the point where $\bar{x} = 0$, $\Delta \bar{p} = 0$, $\bar{q} = 0$, near the normal operating point, Equation (4-26) becomes

$$q = K_1 x - K_2 \Delta p \quad (4-27)$$

where

$$K_1 = (C_1 + C_2) \sqrt{\frac{p_s}{2}} > 0$$

$$K_2 = (C_1 + C_2) \frac{x_0}{4\sqrt{2}\sqrt{p_s}} > 0$$

Equation (4-27) is a linearized mathematical model of the spool valve near the origin ($\bar{x} = 0$, $\Delta \bar{p} = 0$, $\bar{q} = 0$.) Note that the region near the origin is most important in this kind of system, because the system operation usually occurs near this point.

Figure 4-18
Characteristic curves
of the linearized
hydraulic
servomotor.

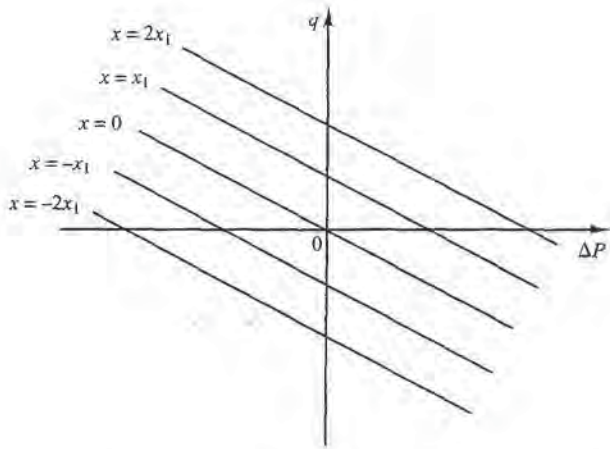


Figure 4-18 shows this linearized relationship among q , x , and ΔP . The straight lines shown are the characteristic curves of the linearized hydraulic servomotor. This family of curves consists of equidistant parallel straight lines, parametrized by x .
In the present analysis, we assume that the load reactive forces are small so that the leakage flow rate and oil compressibility can be ignored. (For the case where the load reactive forces are large, see Problem A-4-13.)

Referring to Figure 4-17(a), we see that the rate of flow of oil q times dt is equal to the power piston displacement dy times the piston area A times the density of oil ρ . Thus, we obtain

$$A\rho\,dy = q\,dt$$

Notice that for a given flow rate q the larger the piston area A is, the lower will be the velocity dy/dt . Hence, if the piston area A is made smaller, the other variables remaining constant, the velocity dy/dt will become higher. Also, an increased flow rate q will cause an increased velocity of the power piston and will make the response time shorter.

Equation (4-27) can now be written as

$$\Delta P = \frac{1}{K_2}\left(K_1x - A\rho\frac{dy}{dt}\right)$$

The force developed by the power piston is equal to the pressure difference ΔP times the piston area A or

$$\begin{aligned}\text{Force developed by the power piston} &= A\,\Delta P \\ &= \frac{A}{K_2}\left(K_1x - A\rho\frac{dy}{dt}\right)\end{aligned}$$

For a given maximum force, if the pressure difference is sufficiently high, the piston area, or the volume of oil in the cylinder, can be made small. Consequently, to minimize the weight of the controller, we must make the supply pressure sufficiently high.

Assume that the power piston moves a load consisting of a mass and viscous friction. Then the force developed by the power piston is applied to the load mass and friction, and we obtain

$$m\ddot{y} + b\dot{y} = \frac{A}{K_2} (K_1 x - A\rho\dot{y})$$

or

$$m\ddot{y} + \left(b + \frac{A^2\rho}{K_2}\right)\dot{y} = \frac{AK_1}{K_2} x \quad (4-28)$$

where m is the mass of the load and b is the viscous-friction coefficient.

Assuming that the pilot valve displacement x is the input and the power piston displacement y is the output, we find that the transfer function for the hydraulic servomotor is, from Equation (4-28),

$$\begin{aligned} \frac{Y(s)}{X(s)} &= \frac{1}{s \left[\left(\frac{mK_2}{AK_1} \right) s + \frac{bK_2}{AK_1} + \frac{A\rho}{K_1} \right]} \\ &= \frac{K}{s(Ts + 1)} \end{aligned} \quad (4-29)$$

where

$$K = \frac{1}{\frac{bK_2}{AK_1} + \frac{A\rho}{K_1}} \quad \text{and} \quad T = \frac{mK_2}{bK_2 + A^2\rho}$$

From Equation (4-29) we see that this transfer function is of the second order. If the ratio $mK_2/(bK_2 + A^2\rho)$ is negligibly small or the time constant T is negligible, the transfer function $Y(s)/X(s)$ can be simplified to give

$$\frac{Y(s)}{X(s)} = \frac{K}{s}$$

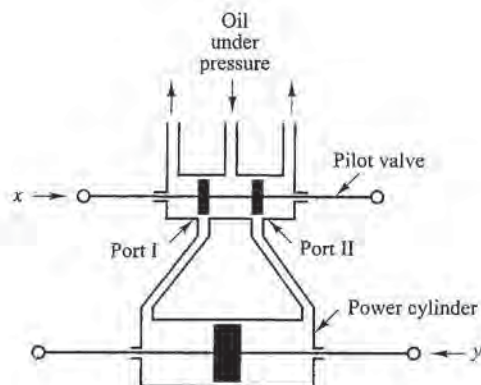
It is noted that a more detailed analysis shows that if oil leakage, compressibility (including the effects of dissolved air), expansion of pipelines, and the likes are taken into consideration, the transfer function becomes

$$\frac{Y(s)}{X(s)} = \frac{K}{s(T_1s + 1)(T_2s + 1)}$$

where T_1 and T_2 are time constants. As a matter of fact, these time constants depend on the volume of oil in the operating circuit. The smaller the volume, the smaller the time constants.

Hydraulic Integral Controller. The hydraulic servomotor shown in Figure 4-19 is a pilot-valve-controlled hydraulic power amplifier and actuator. Similar to the hydraulic servo system shown in Figure 4-17, for negligibly small load mass the servomotor shown in Figure 4-19 acts as an integrator or an integral controller. Such a servomotor constitutes the basis of the hydraulic control circuit.

Figure 4-19
Hydraulic
servomotor.



In the hydraulic servomotor shown in Figure 4-19, the pilot valve (a four-way valve) has two lands on the spool. If the width of the land is smaller than the port in the valve sleeve, the valve is said to be *underlapped*. *Overlapped* valves have a land width greater than the port width. A *zero-lapped* valve has a land width that is identical to the port width. (If the pilot valve is a zero-lapped valve, analyses of hydraulic servomotors become simpler.)

In the present analysis, we assume that hydraulic fluid is incompressible and that the inertia force of the power piston and load is negligible compared to the hydraulic force at the power piston. We also assume that the pilot valve is a zero-lapped valve, and the oil flow rate is proportional to the pilot valve displacement.

Operation of this hydraulic servomotor is as follows. If input x moves the pilot valve to the right, port II is uncovered, and so high-pressure oil enters the right-hand side of the power piston. Since port I is connected to the drain port, the oil in the left-hand side of the power piston is returned to the drain. The oil flowing into the power cylinder is at high pressure; the oil flowing out from the power cylinder into the drain is at low pressure. The resulting difference in pressure on both sides of the power piston will cause it to move to the left.

Note that the rate of flow of oil q (kg/sec) times dt (sec) is equal to the power piston displacement dy (m) times the piston area A (m²) times the density of oil ρ (kg/m³). Therefore,

$$A\rho\,dy = q\,dt \quad (4-30)$$

Because of the assumption that the oil flow rate q is proportional to the pilot valve displacement x , we have

$$q = K_1x \quad (4-31)$$

where K_1 is a positive constant. From Equations (4-30) and (4-31) we obtain

$$A\rho\,\frac{dy}{dt} = K_1x$$

The Laplace transform of this last equation, assuming a zero initial condition, gives

$$A\rho sY(s) = K_1X(s)$$

or

$$\frac{Y(s)}{X(s)} = \frac{K_1}{Aps} = \frac{K}{s}$$

where $K = K_1/(A\rho)$. Thus the hydraulic servomotor shown in Figure 4-19 acts as an integral controller.

Hydraulic Proportional Controller. It has been shown that the servomotor in Figure 4-19 acts as an integral controller. This servomotor can be modified to a proportional controller by means of a feedback link. Consider the hydraulic controller shown in Figure 4-20(a). The left-hand side of the pilot valve is joined to the left-hand side of the power piston by a link ABC . This link is a floating link rather than one moving about a fixed pivot.

The controller here operates in the following way. If input e moves the pilot valve to the right, port II will be uncovered and high-pressure oil will flow through port II into the right-hand side of the power piston and force this piston to the left. The power piston, in moving to the left, will carry the feedback link ABC with it, thereby moving the pilot valve to the left. This action continues until the pilot piston again covers ports I and II. A block diagram of the system can be drawn as in Figure 4-20(b). The transfer function between $Y(s)$ and $E(s)$ is given by

$$\frac{Y(s)}{E(s)} = \frac{\frac{b}{a+b} \frac{K}{s}}{1 + \frac{K}{s} \frac{a}{a+b}}$$

Noting that under the normal operating conditions we have $|Ka/[s(a+b)]| \gg 1$, this last equation can be simplified to

$$\frac{Y(s)}{E(s)} = \frac{b}{a} = K_p$$

The transfer function between y and e becomes a constant. Thus, the hydraulic controller shown in Figure 4-20(a) acts as a proportional controller, the gain of which is K_p . This gain

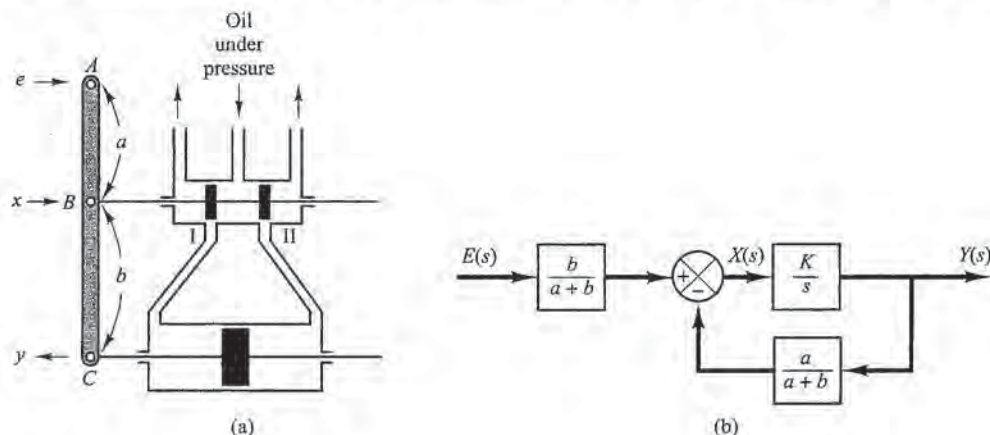


Figure 4-20
(a) Servomotor that acts as a proportional controller; (b) block diagram of the servomotor.

can be adjusted by effectively changing the lever ratio b/a . (The adjusting mechanism is not shown in the diagram.)

We have thus seen that the addition of a feedback link will cause the hydraulic servomotor to act as a proportional controller.

Dashpots. The dashpot (also called a damper) shown in Figure 4-21(a) acts as a differentiating element. Suppose that we introduce a step displacement to the piston position y . Then the displacement z becomes equal to y momentarily. Because of the spring force, however, the oil will flow through the resistance R and the cylinder will come back to the original position. The curves y versus t and z versus t are shown in Figure 4-21(b).

Let us derive the transfer function between the displacement z and displacement y . Define the pressures existing on the right and left sides of the piston as P_1 (lb_f/in.²) and P_2 (lb_f/in.²), respectively. Suppose that the inertia force involved is negligible. Then the force acting on the piston must balance the spring force. Thus

$$A(P_1 - P_2) = kz$$

where A = piston area, in.²

k = spring constant, lb_f/in.

The flow rate q is given by

$$q = \frac{P_1 - P_2}{R}$$

where q = flow rate through the restriction, lb/sec

R = resistance to flow at the restriction, lb_f-sec/in.²-lb

Since the flow through the restriction during dt seconds must equal the change in the mass of oil to the left of the piston during the same dt seconds, we obtain

$$q dt = A\rho(dy - dz)$$

where ρ = density, lb/in.³. (We assume that the fluid is incompressible or ρ = constant.)

This last equation can be rewritten as

$$\frac{dy}{dt} - \frac{dz}{dt} = \frac{q}{A\rho} = \frac{P_1 - P_2}{RA\rho} = \frac{kz}{RA^2\rho}$$

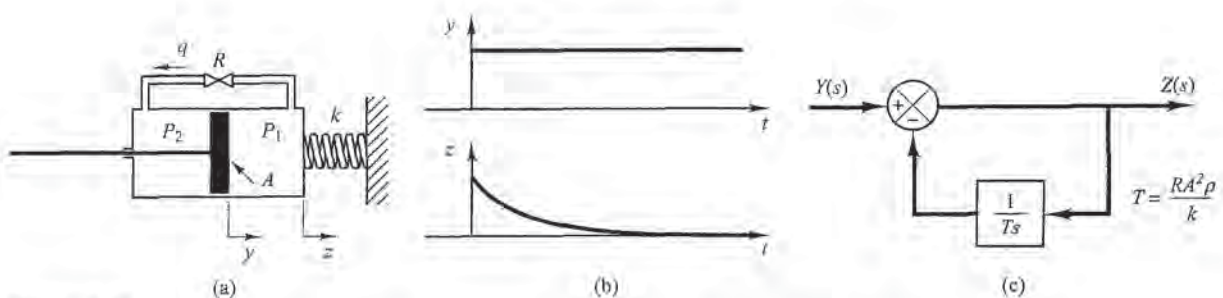


Figure 4-21

(a) Dashpot; (b) step change in y and the corresponding change in z plotted versus t ; (c) block diagram of the dashpot.

or

$$\frac{dy}{dt} = \frac{dz}{dt} + \frac{kz}{RA^2\rho}$$

Taking the Laplace transforms of both sides of this last equation, assuming zero initial conditions, we obtain

$$sY(s) = sZ(s) + \frac{k}{RA^2\rho} Z(s)$$

The transfer function of this system thus becomes

$$\frac{Z(s)}{Y(s)} = \frac{s}{s + \frac{k}{RA^2\rho}}$$

Let us define $RA^2\rho/k = T$. (Note that $RA^2\rho/k$ has the dimension of time.) Then

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

Clearly, the dashpot is a differentiating element. Figure 4-21(c) shows a block diagram representation for this system.

Obtaining Hydraulic Proportional-Plus-Integral Control Action. Figure 4-22(a) shows a schematic diagram of a hydraulic proportional-plus-integral controller. A block diagram of this controller is shown in Figure 4-22(b). The transfer function $Y(s)/E(s)$ is given by

$$\frac{Y(s)}{E(s)} = \frac{\frac{b}{a+b} \frac{K}{s}}{1 + \frac{Ka}{a+b} \frac{T}{Ts + 1}}$$

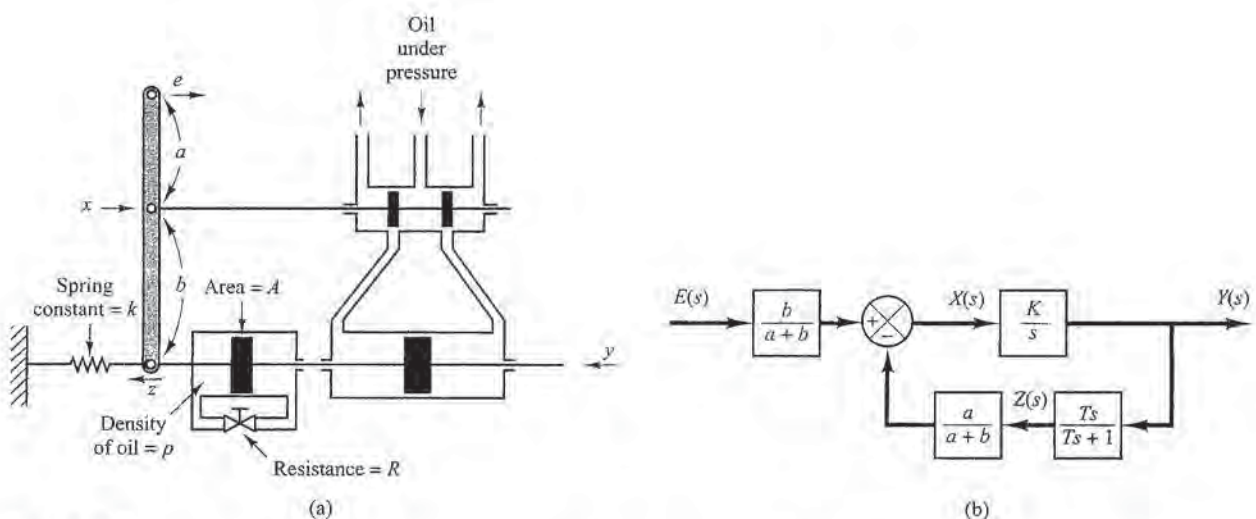


Figure 4-22

(a) Schematic diagram of a hydraulic proportional-plus-integral controller; (b) block diagram of the controller.

In such a controller, under normal operation $|KaT/[(a + b)(Ts + 1)]| \gg 1$, with the result that

$$\frac{Y(s)}{E(s)} = K_p \left(1 + \frac{1}{T_i s} \right)$$

where

$$K_p = \frac{b}{a}, \quad T_i = T = \frac{RA^2\rho}{k}$$

Thus the controller shown in Figure 4-23(a) is a proportional-plus-integral controller (PI controller).

Obtaining Hydraulic Proportional-Plus-Derivative Control Action. Figure 4-23(a) shows a schematic diagram of a hydraulic proportional-plus-derivative controller. The cylinders are fixed in space and the pistons can move. For this system, notice that

$$\begin{aligned} k(y - z) &= A(P_2 - P_1) \\ q &= \frac{P_2 - P_1}{R} \\ q \, dt &= \rho A \, dz \end{aligned}$$

Hence

$$y = z + \frac{A}{k} qR = z + \frac{RA^2\rho}{k} \frac{dz}{dt}$$

or

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

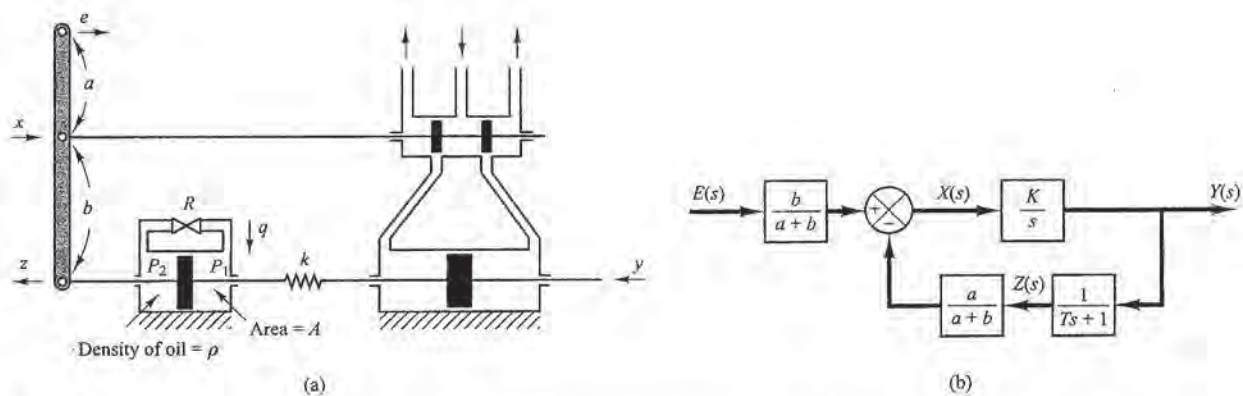


Figure 4-23
(a) Schematic diagram of a hydraulic proportional-plus-derivative controller; (b) block diagram of the controller.

where

$$T = \frac{RA^2\rho}{k}$$

A block diagram for this system is shown in Figure 4-23(b). From the block diagram the transfer function $Y(s)/E(s)$ can be obtained as

$$\frac{Y(s)}{E(s)} = \frac{\frac{b}{a+b} \frac{K}{s}}{1 + \frac{a}{a+b} \frac{K}{s} \frac{1}{Ts+1}}$$

Under normal operation we have $|aK/[(a+b)s(Ts+1)]| \gg 1$. Hence

$$\frac{Y(s)}{E(s)} = K_p(1 + Ts)$$

where

$$K_p = \frac{b}{a}, \quad T = \frac{RA^2\rho}{k}$$

Thus the controller shown in Figure 4-23(a) is a proportional-plus-derivative controller (PD controller).

Obtaining Hydraulic Proportional-Plus-Integral-Plus-Derivative Control Action. Figure 4-24 shows a schematic diagram of a hydraulic proportional-plus-integral-plus-derivative controller. It is a combination of the proportional-plus-integral controller and proportional-plus derivative controller.

If the two dashpots are identical, the transfer function $Z(s)/Y(s)$ can be obtained as follows:

$$\frac{Z(s)}{Y(s)} = \frac{T_1 s}{T_1 T_2 s^2 + (T_1 + 2T_2)s + 1}$$

(For the derivation of this transfer function, refer to Problem A-4-12.)

Figure 4-24
Schematic diagram
of a hydraulic
proportional-plus-
integral-plus-
derivative controller.

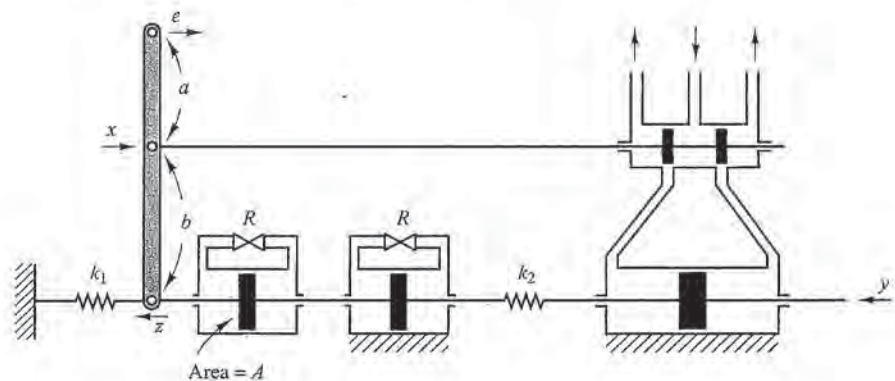
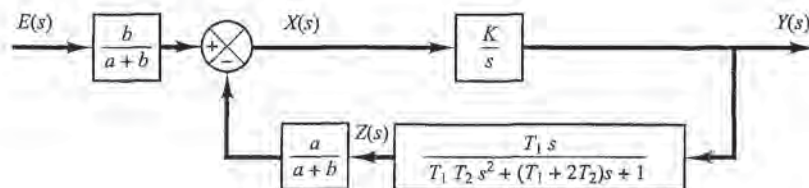


Figure 4-25
Block diagram for
the system shown in
Figure 4-24.



A block diagram for this system is shown in Figure 4-25. The transfer function $Y(s)/E(s)$ can be obtained as

$$\frac{Y(s)}{E(s)} = \frac{b}{a+b} \frac{\frac{K}{s}}{1 + \frac{a}{a+b} \frac{K}{s} \frac{T_1 s}{T_1 T_2 s^2 + (T_1 + 2T_2)s + 1}}$$

Under normal operation of the system we have

$$\left| \frac{a}{a+b} \frac{K}{s} \frac{T_1 s}{T_1 T_2 s^2 + (T_1 + 2T_2)s + 1} \right| \gg 1$$

Hence

$$\begin{aligned} \frac{Y(s)}{E(s)} &= \frac{b}{a} \frac{T_1 T_2 s^2 + (T_1 + 2T_2)s + 1}{T_1 s} \\ &= K_p + \frac{K_i}{s} + K_d s \end{aligned}$$

where

$$K_p = \frac{b}{a} \frac{T_1 + 2T_2}{T_1}, \quad K_i = \frac{b}{a} \frac{1}{T_1}, \quad K_d = \frac{b}{a} T_2$$

Thus, the controller shown in Figure 4-24 is a proportional-plus-integral-plus-derivative controller (PID controller).

4-5 THERMAL SYSTEMS

Thermal systems are those that involve the transfer of heat from one substance to another. Thermal systems may be analyzed in terms of resistance and capacitance, although the thermal capacitance and thermal resistance may not be represented accurately as lumped parameters since they are usually distributed throughout the substance. For precise analysis, distributed-parameter models must be used. Here, however, to simplify the analysis we shall assume that a thermal system can be represented by a lumped-parameter model, that substances that are characterized by resistance to heat flow have negligible heat capacitance, and that substances that are characterized by heat capacitance have negligible resistance to heat flow.

There are three different ways heat can flow from one substance to another: conduction, convection, and radiation. Here we consider only conduction and convection. (Radiation heat transfer is appreciable only if the temperature of the emitter is very

high compared to that of the receiver. Most thermal processes in process control systems do not involve radiation heat transfer.)

For conduction or convection heat transfer,

$$q = K \Delta\theta$$

where q = heat flow rate, kcal/sec

$\Delta\theta$ = temperature difference, °C

K = coefficient, kcal/sec °C

The coefficient K is given by

$$K = \frac{kA}{\Delta X}, \quad \text{for conduction}$$

$$= HA, \quad \text{for convection}$$

where k = thermal conductivity, kcal/m sec °C

A = area normal to heat flow, m²

ΔX = thickness of conductor, m

H = convection coefficient, kcal/m² sec °C

Thermal Resistance and Thermal Capacitance. The thermal resistance R for heat transfer between two substances may be defined as follows:

$$R = \frac{\text{change in temperature difference, } ^\circ\text{C}}{\text{change in heat flow rate, kcal/sec}}$$

The thermal resistance for conduction or convection heat transfer is given by

$$R = \frac{d(\Delta\theta)}{dq} = \frac{1}{K}$$

Since the thermal conductivity and convection coefficients are almost constant, the thermal resistance for either conduction or convection is constant.

The thermal capacitance C is defined by

$$C = \frac{\text{change in heat stored, kcal}}{\text{change in temperature, } ^\circ\text{C}}$$

or

$$C = mc$$

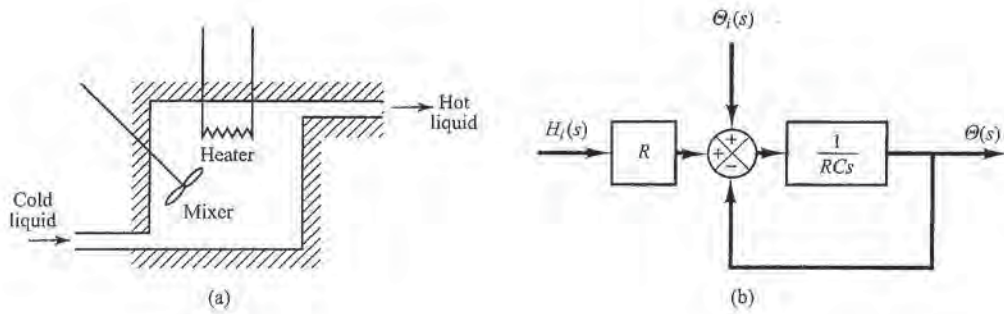
where m = mass of substance considered, kg

c = specific heat of substance, kcal/kg °C

Thermal System. Consider the system shown in Figure 4-26(a). It is assumed that the tank is insulated to eliminate heat loss to the surrounding air. It is also assumed that there is no heat storage in the insulation and that the liquid in the tank is perfectly mixed so that it is at a uniform temperature. Thus, a single temperature is used to describe the temperature of the liquid in the tank and of the outflowing liquid.

Figure 4-26

(a) Thermal system:
(b) block diagram of
the system.



Let us define

$\bar{\theta}_i$ = steady-state temperature of inflowing liquid, °C

$\bar{\theta}_o$ = steady-state temperature of outflowing liquid, °C

G = steady-state liquid flow rate, kg/sec

M = mass of liquid in tank, kg

c = specific heat of liquid, kcal/kg °C

R = thermal resistance, °C sec/kcal

C = thermal capacitance, kcal/°C

$\bar{H}$ = steady-state heat input rate, kcal/sec

Assume that the temperature of the inflowing liquid is kept constant and that the heat input rate to the system (heat supplied by the heater) is suddenly changed from $\bar{H}$ to $\bar{H} + h_i$, where h_i represents a small change in the heat input rate. The heat outflow rate will then change gradually from $\bar{H}$ to $\bar{H} + h_o$. The temperature of the outflowing liquid will also be changed from $\bar{\theta}_o$ to $\bar{\theta}_o + \theta$. For this case, h_o , C , and R are obtained, respectively, as

$$h_o = Gc\theta$$

$$C = Mc$$

$$R = \frac{\theta}{h_o} = \frac{1}{Gc}$$

The heat balance equation for this system is

$$Cd\theta = (h_i - h_o)dt$$

or

$$C \frac{d\theta}{dt} = h_i - h_o$$

which may be rewritten as

$$RC \frac{d\theta}{dt} + \theta = Rh_i$$

Note that the time constant of the system is equal to RC or M/G seconds. The transfer function relating θ and h_i is given by

$$\frac{\Theta(s)}{H_i(s)} = \frac{R}{RCs + 1}$$

where $\Theta(s) = \mathcal{L}[\theta(t)]$ and $H_i(s) = \mathcal{L}[h_i(t)]$.

In practice, the temperature of the inflowing liquid may fluctuate and may act as a load disturbance. (If a constant outflow temperature is desired, an automatic controller may be installed to adjust the heat inflow rate to compensate for the fluctuations in the temperature of the inflowing liquid.) If the temperature of the inflowing liquid is suddenly changed from $\bar{\theta}_i$ to $\bar{\theta}_i + \theta_i$, while the heat input rate H and the liquid flow rate G are kept constant, then the heat outflow rate will be changed from $\bar{H}$ to $\bar{H} + h_o$, and the temperature of the outflowing liquid will be changed from $\bar{\theta}_o$ to $\bar{\theta}_o + \theta$. The heat balance equation for this case is

$$C d\theta = (Gc\theta_i - h_o) dt$$

or

$$C \frac{d\theta}{dt} = Gc\theta_i - h_o$$

which may be rewritten

$$RC \frac{d\theta}{dt} + \theta = \theta_i$$

The transfer function relating θ and θ_i is given by

$$\frac{\Theta(s)}{\Theta_i(s)} = \frac{1}{RCs + 1}$$

where $\Theta(s) = \mathcal{L}[\theta(t)]$ and $\Theta_i(s) = \mathcal{L}[\theta_i(t)]$.

If the present thermal system is subjected to changes in both the temperature of the inflowing liquid and the heat input rate, while the liquid flow rate is kept constant, the change θ in the temperature of the outflowing liquid can be given by the following equation:

$$RC \frac{d\theta}{dt} + \theta = \theta_i + Rh_i$$

A block diagram corresponding to this case is shown in Figure 4-26(b). Notice that the system involves two inputs.

EXAMPLE PROBLEMS AND SOLUTIONS

- A-4-1.** In the liquid-level system of Figure 4-27 assume that the outflow rate Q m³/sec through the outflow valve is related to the head H m by

$$Q = K\sqrt{H} = 0.01\sqrt{H}$$

Assume also that when the inflow rate Q_i is 0.015 m³/sec the head stays constant. For $t < 0$ the system is at steady state ($Q_i = 0.015$ m³/sec). At $t = 0$ the inflow valve is closed and so there is no inflow for $t \geq 0$. Find the time necessary to empty the tank to half the original head. The capacitance C of the tank is 2 m².

Solution. When the head is stationary, the inflow rate equals the outflow rate. Thus head H_o at $t = 0$ is obtained from

$$0.015 = 0.01\sqrt{H_o}$$

or

$$H_o = 2.25 \text{ m}$$

The equation for the system for $t > 0$ is

$$-C \, dH = Q \, dt$$

or

$$\frac{dH}{dt} = -\frac{Q}{C} = \frac{-0.01\sqrt{H}}{2}$$

Hence

$$\frac{dH}{\sqrt{H}} = -0.005 \, dt$$

Assume that, at $t = t_1$, $H = 1.125$ m. Integrating both sides of this last equation, we obtain

$$\int_{2.25}^{1.125} \frac{dH}{\sqrt{H}} = \int_0^{t_1} (-0.005) \, dt = -0.005t_1$$

It follows that

$$2\sqrt{H} \Big|_{2.25}^{1.125} = 2\sqrt{1.125} - 2\sqrt{2.25} = -0.005t_1$$

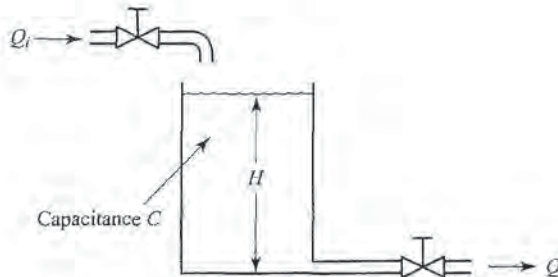


Figure 4-27
Liquid-level system.

or

$$t_1 = 175.7$$

Thus, the head becomes half the original value (2.25 m) in 175.7 sec.

- A-4-2.** Consider the liquid-level system shown in Figure 4-28. At steady state, the inflow rate and outflow rate are both $\bar{Q}$ and the flow rate between the tanks is zero. The heads of tanks 1 and 2 are both $\bar{H}$. At $t = 0$, the inflow rate is changed from $\bar{Q}$ to $\bar{Q} + q$, where q is a small change in the inflow rate. The resulting changes in the heads (h_1 and h_2) and flow rates (q_1 and q_2) are assumed to be small. The capacitances of tanks 1 and 2 are C_1 and C_2 , respectively. The resistance of the valve between the tanks is R_1 and that of the outflow valve is R_2 .

Derive mathematical models for the system when (a) q is the input and h_2 the output, (b) q is the input and q_2 the output, and (c) q is the input and h_1 the output.

Solution. (a) For tank 1, we have

$$C_1 dh_1 = q_1 dt$$

where

$$q_1 = \frac{h_2 - h_1}{R_1}$$

Consequently,

$$R_1 C_1 \frac{dh_1}{dt} + h_1 = h_2 \quad (4-32)$$

For tank 2, we get

$$C_2 dh_2 = (q - q_1 - q_2) dt$$

where

$$q_1 = \frac{h_2 - h_1}{R_1}, \quad q_2 = \frac{h_2}{R_2}$$

It follows that

$$R_2 C_2 \frac{dh_2}{dt} + \frac{R_2}{R_1} h_2 + h_2 = R_2 q + \frac{R_2}{R_1} h_1 \quad (4-33)$$

By eliminating h_1 from Equations (4-32) and (4-33), we have

$$R_1 C_1 R_2 C_2 \frac{d^2 h_2}{dt^2} + (R_1 C_1 + R_2 C_2 + R_2 C_1) \frac{dh_2}{dt} + h_2 = R_1 R_2 C_1 \frac{dq}{dt} + R_2 q \quad (4-34)$$

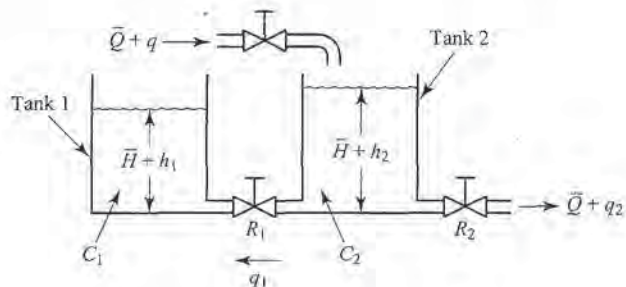


Figure 4-28
Liquid-level system.

In terms of the transfer function, we have

$$\frac{H_2(s)}{Q(s)} = \frac{R_2(R_1C_1s + 1)}{R_1C_1R_2C_2s^2 + (R_1C_1 + R_2C_2 + R_2C_1)s + 1}$$

This is the desired mathematical model in which q is considered the input and h_2 is the output.

(b) Substitution of $h_2 = R_2q_2$ into Equation (4-34) gives

$$R_1C_1R_2C_2 \frac{d^2q_2}{dt^2} + (R_1C_1 + R_2C_2 + R_2C_1) \frac{dq_2}{dt} + q_2 = R_1C_1 \frac{dq}{dt} + q$$

This equation is a mathematical model of the system when q is considered the input and q_2 is the output. In terms of the transfer function, we obtain

$$\frac{Q_2(s)}{Q(s)} = \frac{R_1C_1s + 1}{R_1C_1R_2C_2s^2 + (R_1C_1 + R_2C_2 + R_2C_1)s + 1}$$

(c) Elimination of h_2 from Equations (4-32) and (4-33), yields

$$R_1C_1R_2C_2 \frac{d^2h_1}{dt^2} + (R_1C_1 + R_2C_2 + R_2C_1) \frac{dh_1}{dt} + h_1 = R_2q$$

which is a mathematical model of the system in which q is considered the input and h_1 is the output. In terms of the transfer function, we get

$$\frac{H_1(s)}{Q(s)} = \frac{R_2}{R_1C_1R_2C_2s^2 + (R_1C_1 + R_2C_2 + R_2C_1)s + 1}$$

A-4-3. Consider the liquid-level system shown in Figure 4-29. In the system, $\bar{Q}_1$ and $\bar{Q}_2$ are steady-state inflow rates and $\bar{H}_1$ and $\bar{H}_2$ are steady-state heads. The quantities q_{i1} , q_{i2} , h_1 , h_2 , q_1 , and q_o are considered small. Obtain a state-space representation for the system when h_1 and h_2 are the outputs and q_{i1} and q_{i2} are the inputs.

Solution. The equations for the system are

$$C_1 dh_1 = (q_{i1} - q_1) dt \quad (4-35)$$

$$\frac{h_1 - h_2}{R_1} = q_1 \quad (4-36)$$

$$C_2 dh_2 = (q_1 + q_{i2} - q_o) dt \quad (4-37)$$

$$\frac{h_2}{R_2} = q_o \quad (4-38)$$

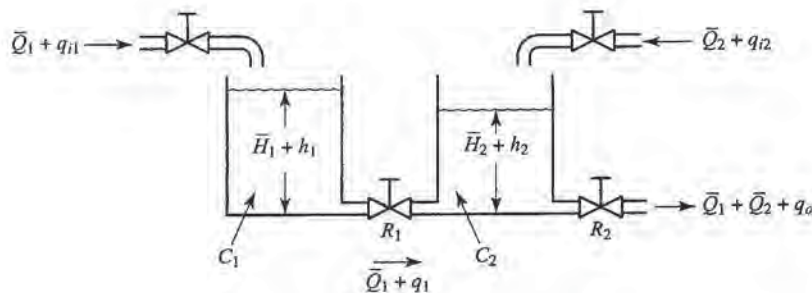


Figure 4-29
Liquid-level system.

Elimination of q_1 from Equation (4-35) using Equation (4-36) results in

$$\frac{dh_1}{dt} = \frac{1}{C_1} \left(q_{i1} - \frac{h_1 - h_2}{R_1} \right) \quad (4-39)$$

Eliminating q_1 and q_o from Equation (4-37) by using Equations (4-36) and (4-38) gives

$$\frac{dh_2}{dt} = \frac{1}{C_2} \left(\frac{h_1 - h_2}{R_1} + q_{i2} - \frac{h_2}{R_2} \right) \quad (4-40)$$

Define state variables x_1 and x_2 by

$$x_1 = h_1$$

$$x_2 = h_2$$

the input variables u_1 and u_2 by

$$u_1 = q_{i1}$$

$$u_2 = q_{i2}$$

and the output variables y_1 and y_2 by

$$y_1 = h_1 = x_1$$

$$y_2 = h_2 = x_2$$

Then Equations (4-39) and (4-40) can be written as

$$\begin{aligned} \dot{x}_1 &= -\frac{1}{R_1 C_1} x_1 + \frac{1}{R_1 C_1} x_2 + \frac{1}{C_1} u_1 \\ \dot{x}_2 &= \frac{1}{R_1 C_2} x_1 - \left(\frac{1}{R_1 C_2} + \frac{1}{R_2 C_2} \right) x_2 + \frac{1}{C_2} u_2 \end{aligned}$$

In the form of the standard vector-matrix representation, we have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{R_1 C_1} & \frac{1}{R_1 C_1} \\ \frac{1}{R_1 C_2} & -\left(\frac{1}{R_1 C_2} + \frac{1}{R_2 C_2} \right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \frac{1}{C_1} & 0 \\ 0 & \frac{1}{C_2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

which is the state equation, and

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

which is the output equation.

Example Problems and Solutions

195

- A-4-4.** Consider the liquid-level system shown in Figure 4-30. At steady state the inflow rate is $Q_i = \bar{Q}$, the outflow rate is $Q_o = \bar{Q}$, and head is $H = \bar{H}$. If the flow is turbulent, then we have

$$\bar{Q} = K\sqrt{\bar{H}}$$

Assume that at $t = 0$ the inflow rate is changed from $Q_i = \bar{Q}$ to $Q_i = \bar{Q} + q_i$. This change causes the head to change from $H = \bar{H}$ to $H = \bar{H} + h$, which, in turn, causes the outflow rate to change from $Q_o = \bar{Q}$ to $Q_o = \bar{Q} + q_o$. For this system we have

$$C \frac{dH}{dt} = Q_i - Q_o = Q_i - K\sqrt{H}$$

where C is the capacitance of the tank. Let us define

$$\frac{dH}{dt} = f(H, Q_i) = \frac{1}{C} Q_i - \frac{K\sqrt{H}}{C} \quad (4-41)$$

Note that the steady-state operating condition is $(\bar{H}, \bar{Q})$ and $H = \bar{H} + h$, $Q_i = \bar{Q} + q_i$. Since at steady-state operation $dH/dt = 0$, we have $f(\bar{H}, \bar{Q}) = 0$.

Linearize Equation (4-41) near the operating point $(\bar{H}, \bar{Q})$.

Solution. Using the linearization technique presented in Section 3-10, a linearized equation for Equation (4-41) can be obtained as follows:

$$\frac{dH}{dt} - f(\bar{H}, \bar{Q}) = \frac{\partial f}{\partial H} (H - \bar{H}) + \frac{\partial f}{\partial Q_i} (Q_i - \bar{Q}) \quad (4-42)$$

where

$$\begin{aligned} f(\bar{H}, \bar{Q}) &= 0 \\ \left. \frac{\partial f}{\partial H} \right|_{H=\bar{H}, Q_i=\bar{Q}} &= -\frac{K}{2C\sqrt{\bar{H}}} = -\frac{\bar{Q}}{\sqrt{\bar{H}}} \frac{1}{2C\sqrt{\bar{H}}} = -\frac{\bar{Q}}{2C\bar{H}} = -\frac{1}{RC} \end{aligned}$$

where we used the resistance R defined by

$$R = \frac{2\bar{H}}{\bar{Q}}$$

Also,

$$\left. \frac{\partial f}{\partial Q_i} \right|_{H=\bar{H}, Q_i=\bar{Q}} = \frac{1}{C}$$

Then Equation (4-42) can be written as

$$\frac{dH}{dt} = -\frac{1}{RC} (H - \bar{H}) + \frac{1}{C} (Q_i - \bar{Q}) \quad (4-43)$$

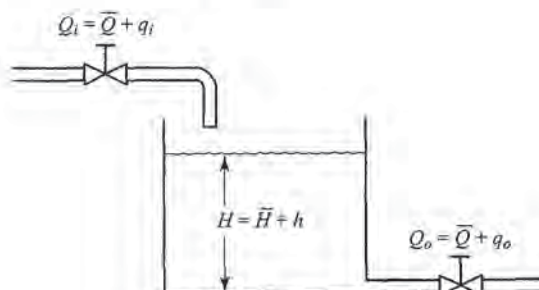


Figure 4-30
Liquid-level system.

Since $H - \bar{H} = h$ and $Q_i - \bar{Q} = q_i$, Equation (4-43) can be written as

$$\frac{dh}{dt} = -\frac{1}{RC}h + \frac{1}{C}q_i$$

or

$$RC \frac{dh}{dt} + h = Rq_i$$

which is the linearized equation for the liquid-level system and is the same as Equation (4-2) that we obtained in Section 4-2.

A-4-5. The value of the gas constant for any gas may be determined from accurate experimental observations of simultaneous values of p , v , and T .

Obtain the gas constant R_{air} for air. Note that at 32°F and 14.7 psia the specific volume of air is 12.39 ft³/lb. Then obtain the capacitance of a 20-ft³ pressure vessel that contains air at 160°F. Assume that the expansion process is isothermal.

Solution.

$$R_{\text{air}} = \frac{pv}{T} = \frac{14.7 \times 144 \times 12.39}{460 + 32} = 53.3 \text{ ft-lb}_f/\text{lb}^\circ\text{R}$$

Referring to Equation (4-12), the capacitance of a 20-ft³ pressure vessel is

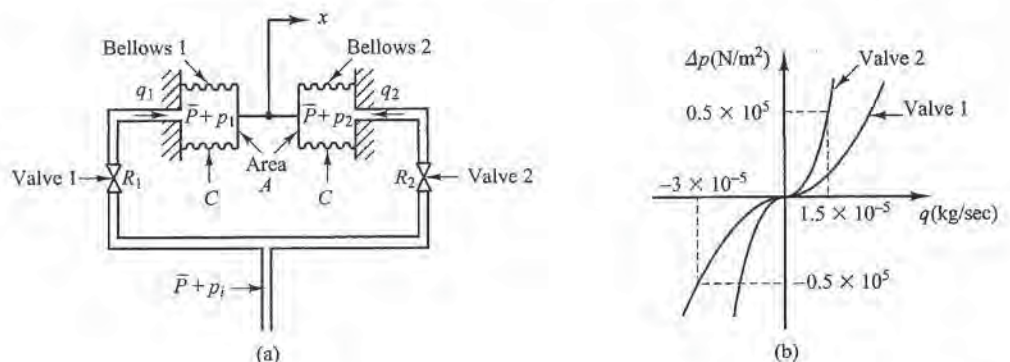
$$C = \frac{V}{nR_{\text{air}}T} = \frac{20}{1 \times 53.3 \times 620} = 6.05 \times 10^{-4} \frac{\text{lb}}{\text{lb}_f/\text{ft}^2}$$

Note that in terms of SI units, R_{air} is given by

$$R_{\text{air}} = 287 \text{ N-m/kg K}$$

A-4-6. In the pneumatic pressure system of Figure 4-31(a), assume that, for $t < 0$, the system is at steady state and that the pressure of the entire system is $\bar{P}$. Also, assume that the two bellows are identical. At $t = 0$, the input pressure is changed from $\bar{P}$ to $\bar{P} + p_i$. Then the pressures in bellows 1 and 2 will change from $\bar{P}$ to $\bar{P} + p_1$ and from $\bar{P}$ to $\bar{P} + p_2$, respectively. The capacity (volume) of each bellows is $5 \times 10^{-4} \text{ m}^3$, and the operating pressure difference Δp (difference between p_i and p_1 or difference between p_1 and p_2) is between $-0.5 \times 10^5 \text{ N/m}^2$ and $0.5 \times 10^5 \text{ N/m}^2$. The

Figure 4-31
(a) Pneumatic pressure system;
(b) pressure difference versus mass flow rate curves.



Example Problems and Solutions

197

corresponding mass flow rates (kg/sec) through the valves are shown in Figure 4-31(b). Assume that the bellows expand or contract linearly with the air pressures applied to them, that the equivalent spring constant of the bellows system is $k = 1 \times 10^5$ N/m, and that each bellows has area $A = 15 \times 10^{-4}$ m².

Defining the displacement of the midpoint of the rod that connects two bellows as x , find the transfer function $X(s)/P_i(s)$. Assume that the expansion process is isothermal and that the temperature of the entire system stays at 30°C.

Solution. Referring to Section 4-3, transfer function $P_1(s)/P_i(s)$ can be obtained as

$$\frac{P_1(s)}{P_i(s)} = \frac{1}{R_1Cs + 1} \quad (4-44)$$

Similarly, transfer function $P_2(s)/P_i(s)$ is

$$\frac{P_2(s)}{P_i(s)} = \frac{1}{R_2Cs + 1} \quad (4-45)$$

The force acting on bellows 1 in the x direction is $A(\bar{P} + p_1)$, and the force acting on bellows 2 in the negative x direction is $A(\bar{P} + p_2)$. The resultant force balances with kx , the equivalent spring force of the corrugated sides of the bellows.

$$A(p_1 - p_2) = kx$$

or

$$A[P_1(s) - P_2(s)] = kX(s) \quad (4-46)$$

Referring to Equations (4-44) and (4-45), we see that

$$\begin{aligned} P_1(s) - P_2(s) &= \left(\frac{1}{R_1Cs + 1} - \frac{1}{R_2Cs + 1} \right) P_i(s) \\ &= \frac{R_2Cs - R_1Cs}{(R_1Cs + 1)(R_2Cs + 1)} P_i(s) \end{aligned}$$

By substituting this last equation into Equation (4-46) and rewriting, the transfer function $X(s)/P_i(s)$ is obtained as

$$\frac{X(s)}{P_i(s)} = \frac{A}{k} \frac{(R_2C - R_1C)s}{(R_1Cs + 1)(R_2Cs + 1)} \quad (4-47)$$

The numerical values of average resistances R_1 and R_2 are

$$\begin{aligned} R_1 &= \frac{d \Delta p}{dq_1} = \frac{0.5 \times 10^5}{3 \times 10^{-5}} = 0.167 \times 10^{10} \frac{\text{N/m}^2}{\text{kg/sec}} \\ R_2 &= \frac{d \Delta p}{dq_2} = \frac{0.5 \times 10^5}{1.5 \times 10^{-5}} = 0.333 \times 10^{10} \frac{\text{N/m}^2}{\text{kg/sec}} \end{aligned}$$

The numerical value of capacitance C of each bellows is

$$C = \frac{V}{nR_{\text{air}}T} = \frac{5 \times 10^{-4}}{1 \times 287 \times (273 + 30)} = 5.75 \times 10^{-9} \frac{\text{kg}}{\text{N/m}^2}$$

where $R_{\text{air}} = 287 \text{ N-m/kg K}$. (See Problem A-4-5.) Consequently,

$$R_1 C = 0.167 \times 10^{10} \times 5.75 \times 10^{-9} = 9.60 \text{ sec}$$

$$R_2 C = 0.333 \times 10^{10} \times 5.75 \times 10^{-9} = 19.2 \text{ sec}$$

By substituting the numerical values for A , k , $R_1 C$, and $R_2 C$ into Equation (4-47), we obtain

$$\frac{X(s)}{P_i(s)} = \frac{1.44 \times 10^{-7}s}{(9.6s + 1)(19.2s + 1)}$$

A-4-7. Draw a block diagram of the pneumatic controller shown in Figure 4-32. Then derive the transfer function of this controller. Assume that $R_d \ll R_i$.

If the resistance R_d is removed (replaced by the line-sized tubing), what control action do we get? If the resistance R_i is removed (replaced by the line-sized tubing), what control action do we get?

Solution. Let us assume that when $e = 0$ the nozzle-flapper distance is equal to $\bar{X}$ and the control pressure is equal to $\bar{P}_c$. In the present analysis, we shall assume small deviations from the respective reference values as follows:

e = small error signal

x = small change in the nozzle-flapper distance

p_c = small change in the control pressure

p_I = small pressure change in bellows I due to small change in the control pressure

p_{II} = small pressure change in bellows II due to small change in the control pressure

y = small displacement at the lower end of the flapper

In this controller, p_c is transmitted to bellows I through the resistance R_d . Similarly, p_c is transmitted to bellows II through the series of resistances R_d and R_i . The relationship between p_I and p_c is

$$\frac{P_I(s)}{P_c(s)} = \frac{1}{R_d C s + 1} = \frac{1}{T_d s + 1}$$

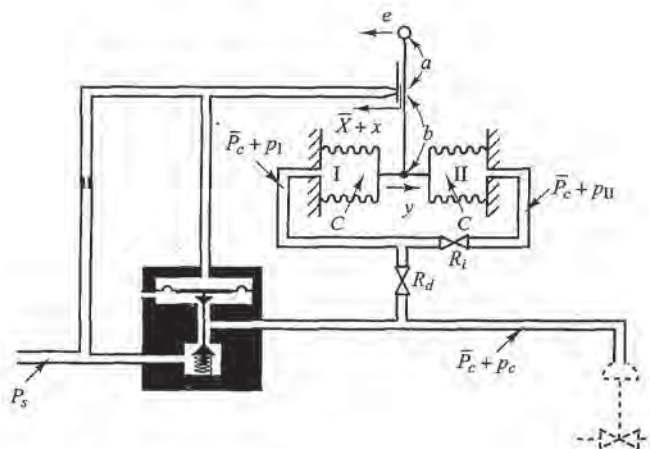


Figure 4-32
Schematic diagram
of a pneumatic
controller.

where $T_d = R_d C =$ derivative time. Similarly, p_{II} and p_I are related by the transfer function

$$\frac{P_{II}(s)}{P_I(s)} = \frac{1}{R_i C s + 1} = \frac{1}{T_i s + 1}$$

where $T_i = R_i C =$ integral time. The force-balance equation for the two bellows is

$$(p_I - p_{II})A = k_s y$$

where k_s is the stiffness of the two connected bellows and A is the cross-sectional area of the bellows. The relationship among the variables e , x , and y is

$$x = \frac{b}{a + b} e - \frac{a}{a + b} y$$

The relationship between p_c and x is

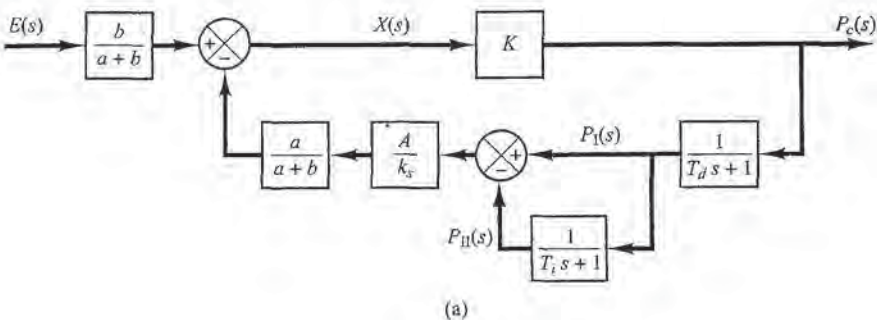
$$p_c = K x \quad (K > 0)$$

From the equations just derived, a block diagram of the controller can be drawn, as shown in Figure 4-33(a). Simplification of this block diagram results in Figure 4-33(b).

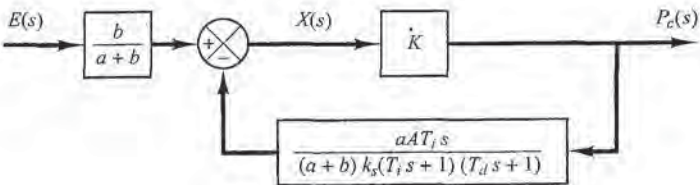
The transfer function between $P_c(s)$ and $E(s)$ is

$$\frac{P_c(s)}{E(s)} = \frac{\frac{b}{a + b} K}{1 + K \frac{a}{a + b} \frac{A}{k_s} \left(\frac{T_i s}{T_i s + 1} \right) \left(\frac{1}{T_d s + 1} \right)}$$

For a practical controller, under normal operation $|K a A T_i s / [(a + b) k_s (T_i s + 1) (T_d s + 1)]|$ is very much greater than unity and $T_i \gg T_d$. Therefore, the transfer function can be simplified as follows:



(a)



(b)

Figure 4-33
(a) Block diagram of the pneumatic controller shown in Figure 4-32;
(b) simplified block diagram.

$$\begin{aligned}\frac{P_c(s)}{E(s)} &\doteq \frac{bk_s(T_i s + 1)(T_d s + 1)}{aAT_i s} \\ &= \frac{bk_s}{aA} \left(\frac{T_i + T_d}{T_i} + \frac{1}{T_i s} + T_d s \right) \\ &\doteq K_p \left(1 + \frac{1}{T_i s} + T_d s \right)\end{aligned}$$

where

$$K_p = \frac{bk_s}{aA}$$

Thus the controller shown in Figure 4-32 is a proportional-plus-integral-plus-derivative one.

If the resistance R_d is removed, or $R_d = 0$, the action becomes that of a proportional-plus-integral controller.

If the resistance R_i is removed, or $R_i = 0$, the action becomes that of a narrow-band proportional, or two-position, controller. (Note that the actions of two feedback bellows cancel each other, and there is no feedback.)

A-4-8. Actual spool valves are either overlapped or underlapped because of manufacturing tolerances. Consider the overlapped and underlapped spool valves shown in Figures 4-34(a) and (b). Sketch curves relating the uncovered port area A versus displacement x .

Solution. For the overlapped valve, a dead zone exists between $-\frac{1}{2}x_0$ and $\frac{1}{2}x_0$, or $-\frac{1}{2}x_0 < x < \frac{1}{2}x_0$. The uncovered port area A versus displacement x curve is shown in Figure 4-35(a). Such an overlapped valve is unfit as a control valve.

For the underlapped valve, the port area A versus displacement x curve is shown in Figure 4-35(b). The effective curve for the underlapped region has a higher slope, meaning a higher sensitivity. Valves used for controls are usually underlapped.

Figure 4-34
(a) Overlapped spool valve;
(b) underlapped spool valve.

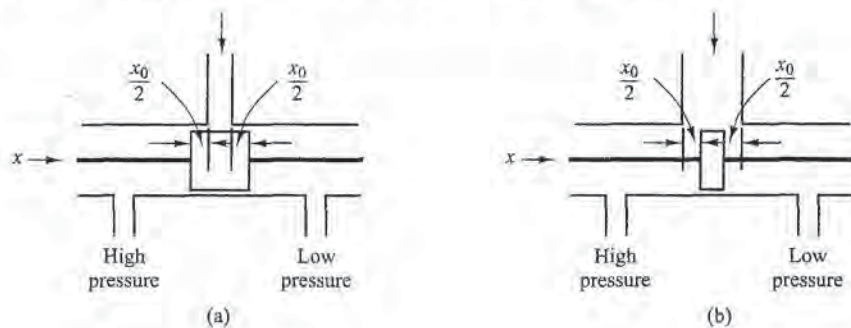
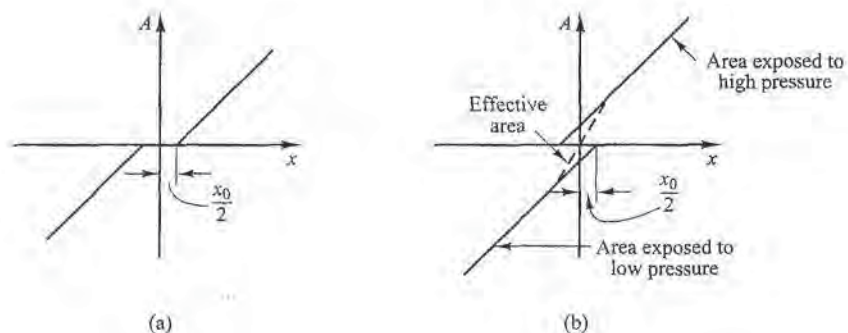


Figure 4-35
(a) Uncovered port area A versus displacement x curve for the overlapped valve; (b) uncovered port area A versus displacement x curve for the underlapped valve.



- A-4-9.** Figure 4-36 shows a hydraulic jet-pipe controller. Hydraulic fluid is ejected from the jet pipe. If the jet pipe is shifted to the right from the neutral position, the power piston moves to the left, and vice versa. The jet pipe valve is not used as much as the flapper valve because of large null flow, slower response, and rather unpredictable characteristics. Its main advantage lies in its insensitivity to dirty fluids.

Suppose that the power piston is connected to a light load so that the inertia force of the load element is negligible compared to the hydraulic force developed by the power piston. What type of control action does this controller produce?

Solution. Define the displacement of the jet nozzle from the neutral position as x and the displacement of the power piston as y . If the jet nozzle is moved to the right by a small displacement x , the oil flows to the right side of the power piston, and the oil in the left side of the power piston is returned to the drain. The oil flowing into the power cylinder is at high pressure; the oil flowing out from the power cylinder into the drain is at low pressure. The resulting pressure difference causes the power piston to move to the left.

For a small jet nozzle displacement x , the flow rate q to the power cylinder is proportional to x ; that is,

$$q = K_1 x$$

For the power cylinder,

$$A\rho dy = q dt$$

where A is the power piston area and ρ is the density of oil. Hence

$$\frac{dy}{dt} = \frac{q}{A\rho} = \frac{K_1}{A\rho} x = Kx$$

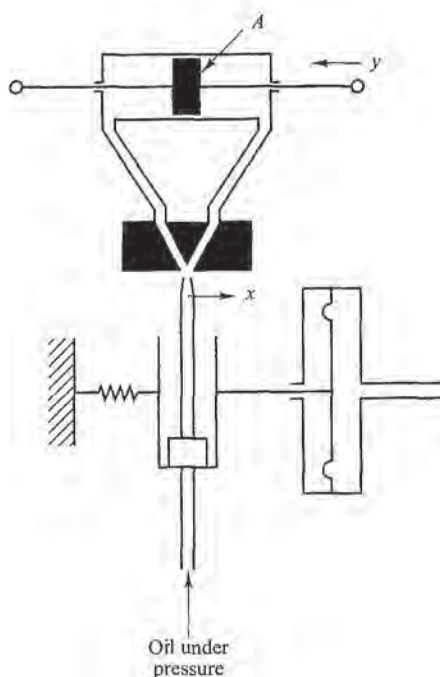


Figure 4-36
Hydraulic jet-pipe
controller.

where $K = K_1/(A\rho) = \text{constant}$. The transfer function $Y(s)/X(s)$ is thus

$$\frac{Y(s)}{X(s)} = \frac{K}{s}$$

The controller produces the integral control action.

- A-4-10.** Figure 4-37 shows a hydraulic jet-pipe applied to a flow control system. The jet-pipe controller governs the position of the butterfly valve. Discuss the operation of this system. Plot a possible curve relating the displacement x of the nozzle to the total force F acting on the power piston.

Solution. The operation of this system is as follows: The flow rate is measured by the orifice, and the pressure difference produced by this orifice is transmitted to the diaphragm of the pressure measuring device. The diaphragm is connected to the free swinging nozzle, or jet pipe, through a linkage. High-pressure oil ejects from the nozzle all the time. When the nozzle is at a neutral position, no oil flows through either of the pipes to move the power piston. If the nozzle is displaced by the motion of the balance arm to one side, the high-pressure oil flows through the corresponding pipe, and the oil in the power cylinder flows back to the sump through the other pipe.

Assume that the system is initially at rest. If the reference input is changed suddenly to a higher flow rate, then the nozzle is moved in such a direction as to move the power piston and open the butterfly valve. Then the flow rate will increase, the pressure difference across the orifice becomes larger, and the nozzle will move back to the neutral position. The movement of the power piston stops when x , the displacement of the nozzle, comes back to and stays at the neutral position. (The jet pipe controller thus possesses an integrating property.)

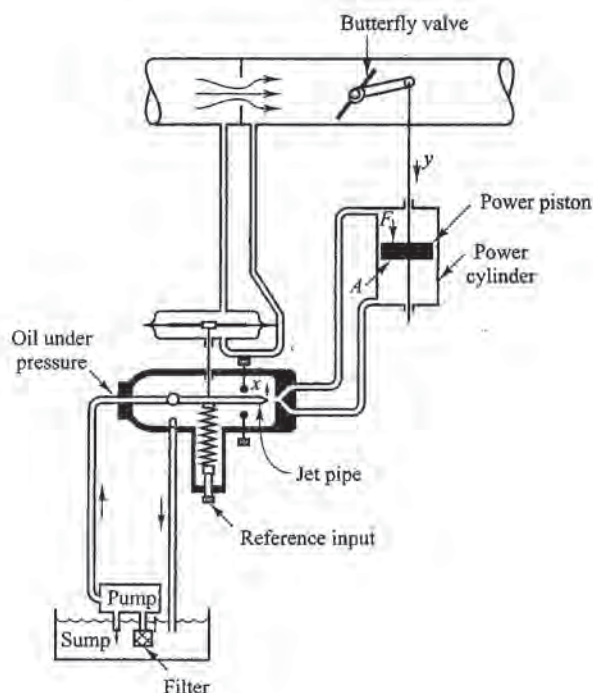
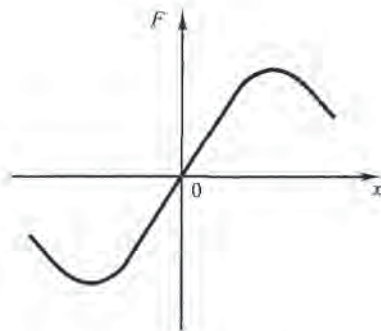


Figure 4-37
Schematic diagram
of a flow control
system using a
hydraulic jet-pipe
controller.

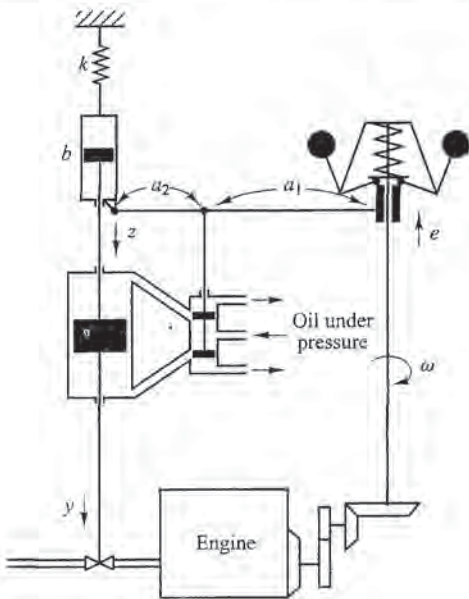
Figure 4-38
Force versus
displacement curve.



The relationship between the total force F acting on the power piston and the displacement x of the nozzle is shown in Figure 4-38. The total force is equal to the pressure difference ΔP across the piston times the area A of the power piston. For a small displacement x of the nozzle, the total force F and displacement x may be considered proportional.

A-4-11. Explain the operation of the speed control system shown in Figure 4-39.

Figure 4-39
Speed control
system.

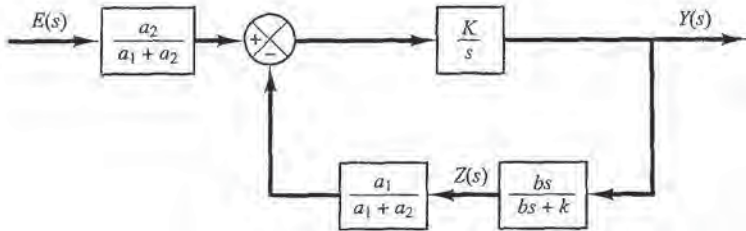


Solution. If the engine speed increases, the sleeve of the fly-ball governor moves upward. This movement acts as the input to the hydraulic controller. A positive error signal (upward motion of the sleeve) causes the power piston to move downward, reduces the fuel-valve opening, and decreases the engine speed. A block diagram for the system is shown in Figure 4-40.

From the block diagram the transfer function $Y(s)/E(s)$ can be obtained as

$$\frac{Y(s)}{E(s)} = \frac{a_2}{a_1 + a_2} \frac{\frac{K}{s}}{1 + \frac{a_1}{a_1 + a_2} \frac{bs}{bs + k} \frac{K}{s}}$$

Figure 4-40
Block diagram for
the speed control
system shown in
Figure 4-39.



If the following condition applies,

$$\left| \frac{a_1}{a_1 + a_2} \frac{bs}{bs + k} \frac{K}{s} \right| \gg 1$$

the transfer function $Y(s)/E(s)$ becomes

$$\frac{Y(s)}{E(s)} \doteq \frac{a_2}{a_1 + a_2} \frac{a_1 + a_2}{a_1} \frac{bs + k}{bs} = \frac{a_2}{a_1} \left(1 + \frac{k}{bs} \right)$$

The speed controller has proportional-plus-integral control action.

A-4-12. Derive the transfer function $Z(s)/Y(s)$ of the hydraulic system shown in Figure 4-41. Assume that the two dashpots in the system are identical ones.

Solution. In deriving the equations for the system, we assume that force F is applied at the right end of the shaft causing displacement y . (All displacements y , w , and z are measured from respective equilibrium positions when no force is applied at the right end of the shaft.) When force F is applied, pressure P_1 becomes higher than pressure P'_1 , or $P_1 > P'_1$. Similarly, $P_2 > P'_2$.

For the force balance, we have the following equation:

$$k_2(y - w) = A(P_1 - P'_1) + A(P_2 - P'_2) \tag{4-48}$$

Since

$$k_1 z = A(P_1 - P'_1) \tag{4-49}$$

and

$$q_1 = \frac{P_1 - P'_1}{R}$$

we have

$$k_1 z = ARq_1$$

Also, since

$$q_1 \, dt = A(dw - dz)\rho$$

we have

$$q_1 = A(\dot{w} - \dot{z})\rho$$

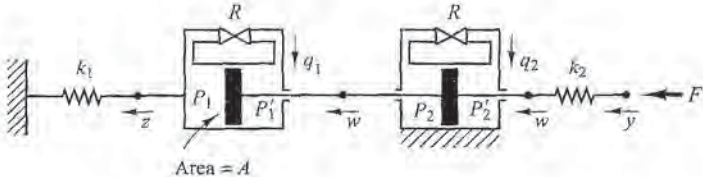


Figure 4-41
Hydraulic system.

or

$$\dot{w} - \dot{z} = \frac{k_1 z}{A^2 R \rho}$$

Define $A^2 R \rho = B$. (B is the viscous friction coefficient.) Then

$$\dot{w} - \dot{z} = \frac{k_1}{B} z \quad (4-50)$$

Also, for the right-hand side dashpot we have

$$q_2 dt = A dw \rho$$

Hence

$$\dot{w} = \frac{q_2}{A \rho} = \frac{A(P_2 - P'_2)}{A^2 R \rho}$$

or

$$A(P_2 - P'_2) = B \dot{w} \quad (4-51)$$

Substituting Equations (4-49) and (4-51) into Equation (4-48) we have

$$k_2 y - k_2 w = k_1 z + B \dot{w}$$

Taking the Laplace transform of this last equation, assuming zero initial condition, we obtain

$$k_2 Y(s) = (k_2 + Bs)W(s) + k_1 Z(s) \quad (4-52)$$

Taking the Laplace transform of Equation (4-50), assuming zero initial condition, we obtain

$$W(s) = \frac{k_1 + Bs}{Bs} Z(s) \quad (4-53)$$

By using Equation (4-53) to eliminate $W(s)$ from Equation (4-52), we obtain

$$k_2 Y(s) = (k_2 + Bs) \frac{k_1 + Bs}{Bs} Z(s) + k_1 Z(s)$$

from which we obtain the transfer function $Z(s)/Y(s)$ to be

$$\frac{Z(s)}{Y(s)} = \frac{k_2 s}{Bs^2 + (2k_1 + k_2)s + \frac{k_1 k_2}{B}}$$

Multiplying $B/(k_1 k_2)$ to both the numerator and denominator of this last equation, we get

$$\frac{Z(s)}{Y(s)} = \frac{\frac{B}{k_1} s}{\frac{B^2}{k_1 k_2} s^2 + \left(\frac{2B}{k_2} + \frac{B}{k_1} \right) s + 1}$$

Define $B/k_1 = T_1$, $B/k_2 = T_2$. Then the transfer function $Z(s)/Y(s)$ becomes as follows:

$$\frac{Z(s)}{Y(s)} = \frac{T_1 s}{T_1 T_2 s^2 + (T_1 + 2T_2)s + 1}$$

A-4-13. Consider the hydraulic servo system shown in Figure 4-42. Assuming that the load reaction forces are not negligible, derive a mathematical model of the system. Assume also that the mass of the power piston is included in the load mass m .

Solution. In deriving a mathematical model of the system when the load reactive forces are not negligible, such effects as the pressure drop across the orifice, the leakage of oil around the valve and around the piston, and the compressibility of the oil must be considered.

The pressure drop across the orifice is a function of the supply pressure p_s and the pressure difference $\Delta p = p_1 - p_2$. Thus the flow rate q is a nonlinear function of valve displacement x and pressure difference Δp or

$$q = f(x, \Delta p)$$

Linearizing this nonlinear equation about the origin ($x = 0, \Delta p = 0, q = 0$), we obtain, referring to Equation (4-27),

$$q = K_1 x - K_2 \Delta p \quad (4-54)$$

The flow rate q can be considered as consisting of three parts

$$q = q_0 + q_L + q_C \quad (4-55)$$

where q_0 = useful flow rate to the power cylinder causing power piston to move, kg/sec

q_L = leakage flow rate, kg/sec

q_C = equivalent compressibility flow rate, kg/sec

Let us obtain specific expressions for q_0, q_L , and q_C . The flow $q_0 dt$ to the left-hand side of the power piston causes the piston to move to the right by dy . So we have

$$A\rho dy = q_0 dt$$

where A (m^2) is the power piston area, ρ (kg/m^3) the density of oil, and dy (m) the displacement of the power piston. Then

$$q_0 = A\rho \frac{dy}{dt} \quad (4-56)$$

The leakage component q_L can be written

$$q_L = L \Delta p \quad (4-57)$$

where L is the leakage coefficient of the system.

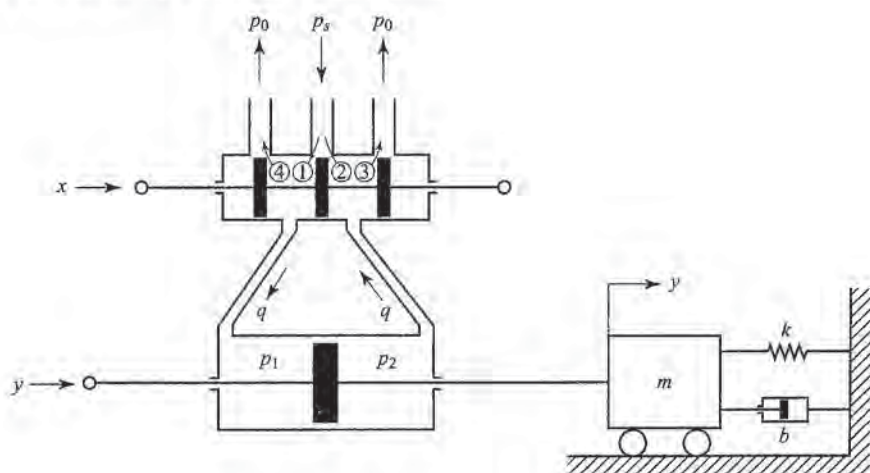


Figure 4-42
Hydraulic servo
system.

The equivalent compressibility flow rate q_c can be expressed in terms of the effective bulk modulus K of oil (including the effects of entrapped air, expansion of pipes, etc.), where

$$K = \frac{d \Delta p}{-dV/V}$$

(Here dV is negative and so $-dV$ is positive.) Rewriting this last equation gives

$$-dV = \frac{V}{K} d \Delta p$$

or

$$\rho \frac{-dV}{dt} = \frac{\rho V}{K} \frac{d \Delta p}{dt}$$

Noting that $q_c = \rho(-dV)/dt$, we find

$$q_c = \frac{\rho V}{K} \frac{d \Delta p}{dt} \quad (4-58)$$

where V is the effective volume of oil under compression (that is, approximately half the total power cylinder volume).

Using Equations (4-54) through (4-58),

$$\begin{aligned} q &= K_1 x - K_2 \Delta p \\ &= q_0 + q_L + q_c \\ &= A \rho \frac{dy}{dt} + L \Delta p + \frac{\rho V}{K} \frac{d \Delta p}{dt} \end{aligned}$$

or

$$A \rho \frac{dy}{dt} + \frac{\rho V}{K} \frac{d \Delta p}{dt} + (L + K_2) \Delta p = K_1 x \quad (4-59)$$

The force developed by the power piston is $A \Delta p$, and this force is applied to the load elements. Thus

$$m \frac{d^2 y}{dt^2} + b \frac{dy}{dt} + k y = A \Delta p \quad (4-60)$$

Eliminating Δp from Equations (4-59) and (4-60) results in

$$\begin{aligned} \frac{\rho V m}{K A} \frac{d^3 y}{dt^3} + \left[\frac{\rho V b}{K A} + \frac{(L + K_2) m}{A} \right] \frac{d^2 y}{dt^2} \\ + \left[A \rho + \frac{\rho V k}{K A} + \frac{(L + K_2) b}{A} \right] \frac{dy}{dt} + \frac{(L + K_2) k}{A} y = K_1 x \end{aligned}$$

This is a mathematical model of the system relating the valve spool displacement x and the power piston displacement y when the load reactive forces are not negligible.

- A-4-14.** Considering small deviations from steady-state operation, draw a block diagram of the air heating system shown in Figure 4-43. Assume that the heat loss to the surroundings and the heat capacitance of the metal parts of the heater are negligible.

Solution. Let us define

$\bar{\theta}_i$ = steady-state temperature of inlet air, °C

$\bar{\theta}_o$ = steady-state temperature of outlet air, °C

G = mass flow rate of air through the heating chamber, kg/sec

M = mass of air contained in the heating chamber, kg

c = specific heat of air, kcal/kg °C

R = thermal resistance, °C sec/kcal

C = thermal capacitance of air contained in the heating chamber = Mc , kcal/°C

$\bar{H}$ = steady-state heat input, kcal/sec

Let us assume that the heat input is suddenly changed from $\bar{H}$ to $\bar{H} + h$ and the inlet air temperature is suddenly changed from $\bar{\theta}_i$ to $\bar{\theta}_i + \theta_i$. Then the outlet air temperature will be changed from $\bar{\theta}_o$ to $\bar{\theta}_o + \theta_o$.

The equation describing the system behavior is

$$C d\theta_o = [h + Gc(\theta_i - \theta_o)] dt$$

or

$$C \frac{d\theta_o}{dt} = h + Gc(\theta_i - \theta_o)$$

Noting that

$$Gc = \frac{1}{R}$$

we obtain

$$C \frac{d\theta_o}{dt} = h + \frac{1}{R}(\theta_i - \theta_o)$$

or

$$RC \frac{d\theta_o}{dt} + \theta_o = Rh + \theta_i$$

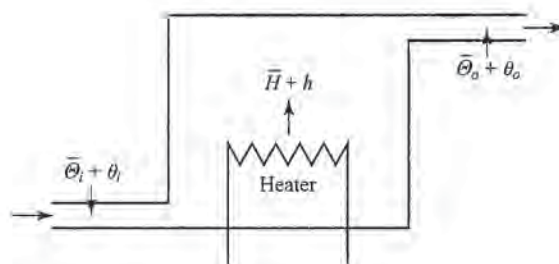
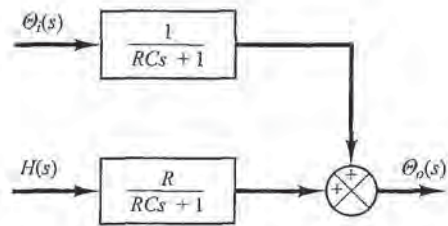


Figure 4-43
Air heating system.

Figure 4-44
Block diagram of the
air heating system
shown in
Figure 4-43.



Taking the Laplace transforms of both sides of this last equation and substituting the initial condition that $\theta_0(0) = 0$, we obtain

$$\Theta_o(s) = \frac{R}{RCs + 1} H(s) + \frac{1}{RCs + 1} \Theta_i(s)$$

The block diagram of the system corresponding to this equation is shown in Figure 4-44.

A-4-15. Consider the thin, glass-wall, mercury thermometer system shown in Figure 4-45. Assume that the thermometer is at a uniform temperature $\bar{\theta}$ (ambient temperature) and that at $t = 0$ it is immersed in a bath of temperature $\bar{\theta} + \theta_b$, where θ_b is the bath temperature (which may be constant or changing) measured from the ambient temperature $\bar{\theta}$. Define the instantaneous thermometer temperature by $\bar{\theta} + \theta$ so that θ is the change in the thermometer temperature satisfying the condition that $\theta(0) = 0$. Obtain a mathematical model for the system. Also obtain an electrical analog of the thermometer system.

Solution. A mathematical model for the system can be derived by considering heat balance as follows: The heat entering the thermometer during dt sec is $q dt$, where q is the heat flow rate to the thermometer. This heat is stored in the thermal capacitance C of the thermometer, thereby raising its temperature by $d\theta$. Thus the heat-balance equation is

$$C d\theta = q dt \tag{4-61}$$

Since thermal resistance R may be written as

$$R = \frac{d(\Delta\theta)}{dq} = \frac{\Delta\theta}{q}$$

heat flow rate q may be given, in terms of thermal resistance R , as

$$q = \frac{(\bar{\theta} + \theta_b) - (\bar{\theta} + \theta)}{R} = \frac{\theta_b - \theta}{R}$$

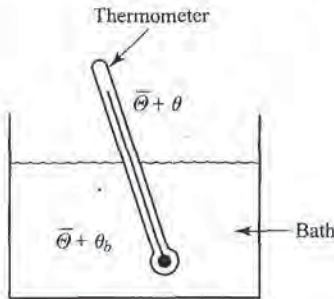


Figure 4-45
Thin, glass-wall,
mercury thermo-
meter system.

where $\bar{\theta} + \theta_b$ is the bath temperature and $\bar{\theta} + \theta$ is the thermometer temperature. Hence, we can rewrite Equation (4-61) as

$$C \frac{d\theta}{dt} = \frac{\theta_b - \theta}{R}$$

or

$$RC \frac{d\theta}{dt} + \theta = \theta_b \quad (4-62)$$

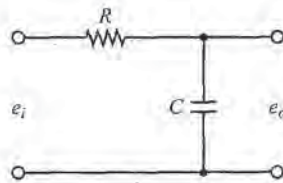
Equation (4-62) is a mathematical model of the thermometer system.

Referring to Equation (4-62), an electrical analog for the thermometer system can be written as

$$RC \frac{de_o}{dt} + e_o = e_i$$

An electrical circuit represented by this last equation is shown in Figure 4-46.

Figure 4-46
Electrical analog of the thermometer system shown in Figure 4-45.



PROBLEMS

B-4-1. Consider the liquid-level system shown in Figure 4-47. Assuming that $\bar{H} = 3$ m, $\bar{Q} = 0.02$ m³/sec, and the cross-sectional area of the tank is equal to 5 m², obtain the time constant of the system at the operating point $(\bar{H}, \bar{Q})$. Assume that the flow through the valve is turbulent.

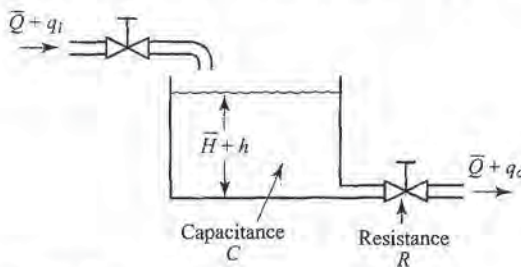


Figure 4-47
Liquid-level system.

B-4-2. Consider the conical water tank system shown in Figure 4-48; The flow through the valve is turbulent and is related to the head H by

$$Q = 0.005\sqrt{H}$$

where Q is the flow rate measured in m³/sec and H is in meters.

Suppose that the head is 2 m at $t = 0$. What will be the head at $t = 60$ sec?

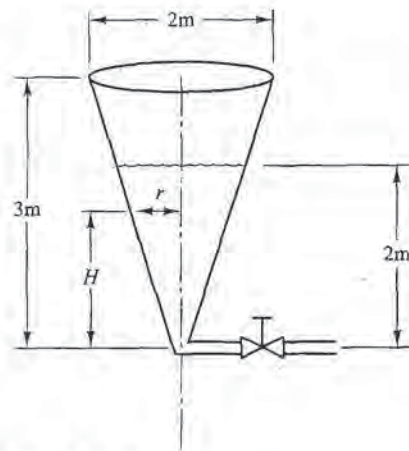


Figure 4-48
Conical water tank system.

B-4-3. Consider the liquid-level system shown in Figure 4-49. At steady state the inflow rate is $\bar{Q}$ and the outflow rate is also $\bar{Q}$. Assume that at $t = 0$ the inflow rate is changed from $\bar{Q}$ to $\bar{Q} + q_i$, where q_i is a small quantity. The disturbance input is q_d , which is also a small quantity. Draw a block diagram of the system and simplify it to obtain $H_2(s)$ as a function of $Q_i(s)$ and $Q_d(s)$, where $H_2(s) = \mathcal{L}[h_2(t)]$, $Q_i(s) = \mathcal{L}[q_i(t)]$, and $Q_d(s) = \mathcal{L}[q_d(t)]$. The capacitances of tanks 1 and 2 are C_1 and C_2 , respectively.

B-4-4. Consider the liquid-level control system shown in Figure 4-50. The controller is of the proportional type. The set point of the controller is fixed. Draw a block diagram of the system, assuming that changes in the variables are small. Obtain the transfer function between the level of the second tank and the disturbance input q_d . Obtain the steady-state error when the disturbance q_d is a unit-step function.

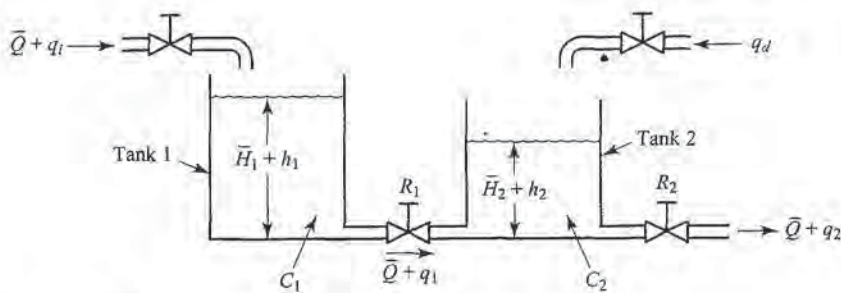


Figure 4-49
Liquid-level system.

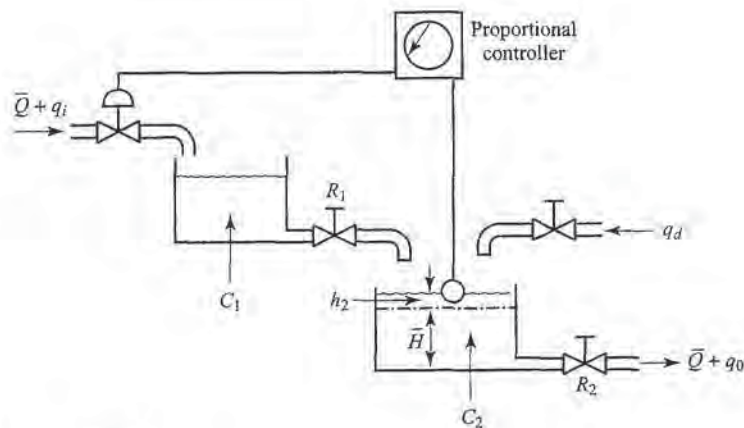


Figure 4-50
Liquid-level control system.

B-4-5. For the pneumatic system shown in Figure 4-51, assume that steady-state values of the air pressure and the displacement of the bellows are $\bar{P}$ and $\bar{X}$, respectively. Assume also that the input pressure is changed from $\bar{P}$ to $\bar{P} + p_i$, where p_i is a small change in the input pressure. This change will cause the displacement of the bellows to change a small amount x . Assuming that the capacitance of the bellows is C and the resistance of the valve is R , obtain the transfer function relating x and p_i .

B-4-6. Figure 4-52 shows a pneumatic controller. The pneumatic relay has the characteristic that $p_c = K p_b$, where $K > 0$. What kind of control action does this controller produce? Derive the transfer function $P_c(s)/E(s)$.

B-4-7. Consider the pneumatic controller shown in Figure 4-53. Assuming that the pneumatic relay has the characteristics that $p_c = K p_b$ (where $K > 0$), determine the control action of this controller. The input to the controller is e and the output is p_c .

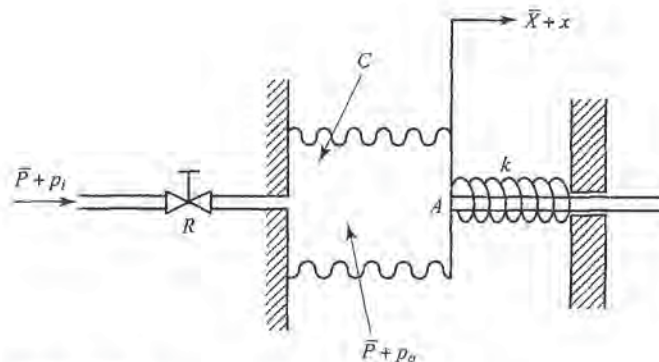


Figure 4-51
Pneumatic system.

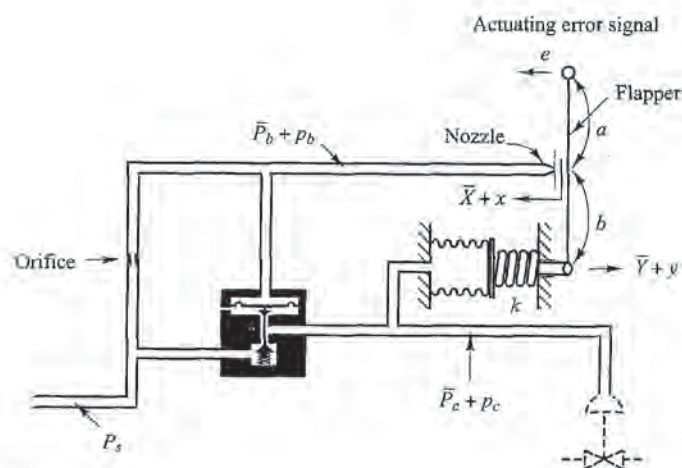


Figure 4-52
Pneumatic controller.

B-4-8. Figure 4-54 shows a pneumatic controller. The signal e is the input and the change in the control pressure p_c is the output. Obtain the transfer function $P_c(s)/E(s)$.

Assume that the pneumatic relay has the characteristics that $p_c = K p_b$, where $K > 0$.

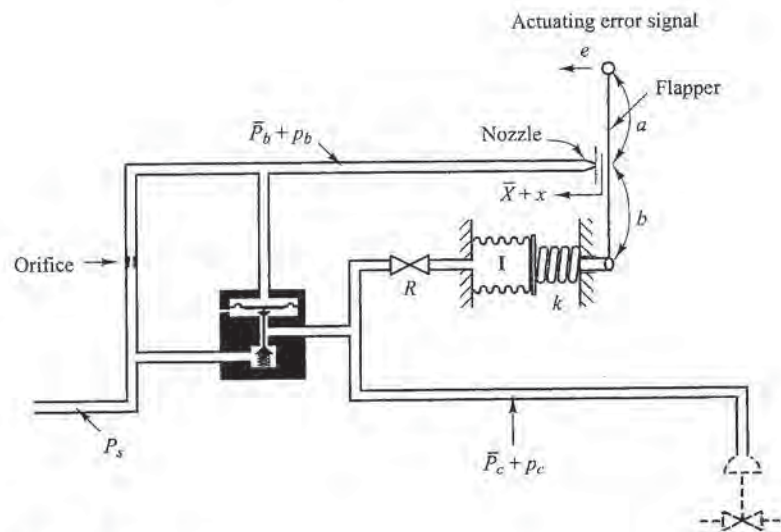


Figure 4-53
Pneumatic controller.

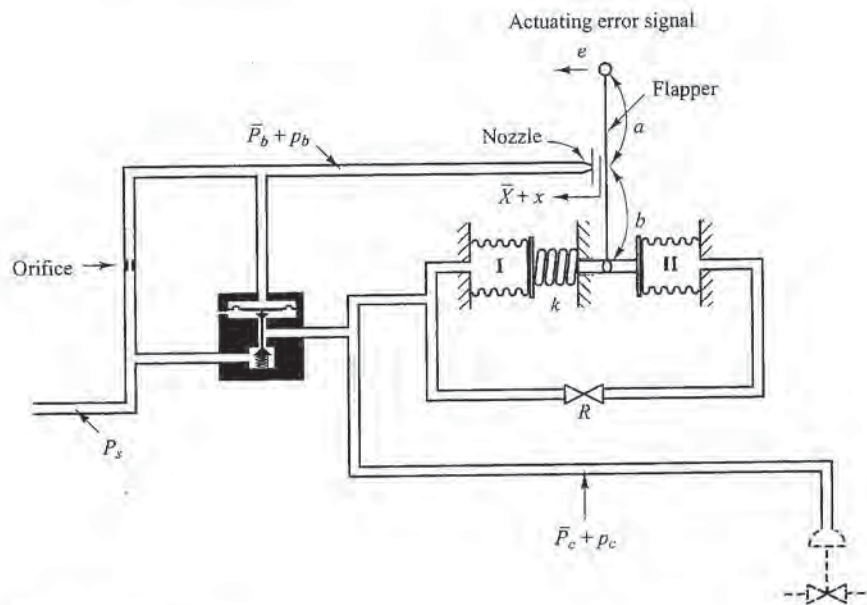


Figure 4-54
Pneumatic controller.

B-4-9. Consider the pneumatic controller shown in Figure 4-55. What control action does this controller produce? Assume that the pneumatic relay has the characteristics that $p_c = K p_b$, where $K > 0$.

B-4-10. Figure 4-56 shows a flapper valve. It is placed between two opposing nozzles. If the flapper is moved slightly to the right, the pressure unbalance occurs in the nozzles and the power piston moves to the left, and vice versa. Such

a device is frequently used in hydraulic servos as the first-stage valve in two-stage servovalves. This usage occurs because considerable force may be needed to stroke larger spool valves that result from the steady-state flow force. To reduce or compensate this force, two-stage valve configuration is often employed; a flapper valve or jet pipe is used as the first-stage valve to provide a necessary force to stroke the second-stage spool valve.

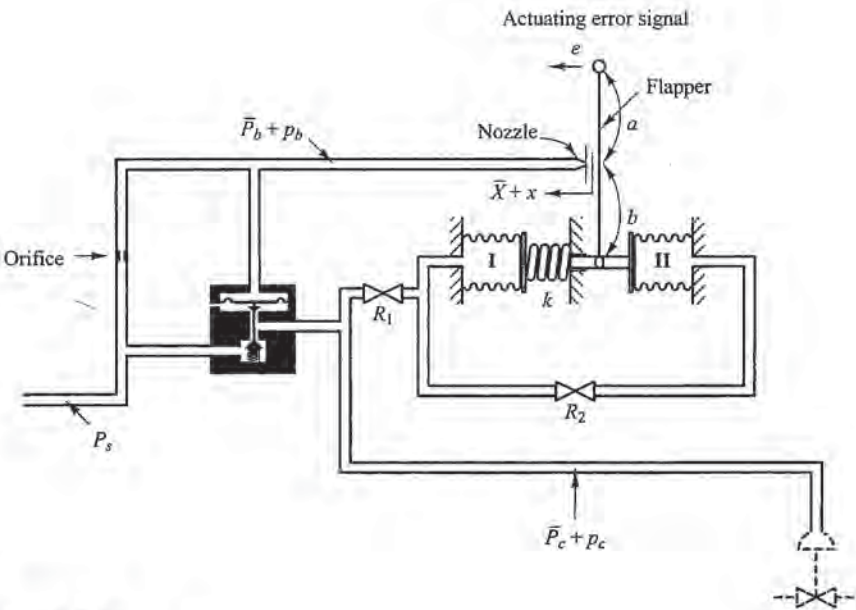


Figure 4-55
Pneumatic controller.

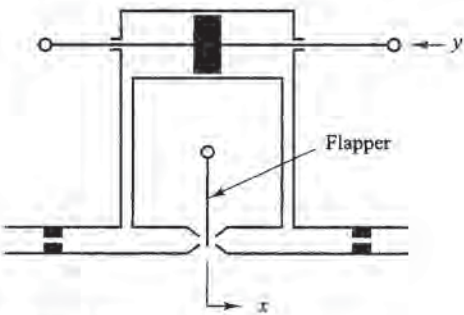


Figure 4-56
Flapper valve.

Figure 4-57 shows a schematic diagram of a hydraulic servomotor in which the error signal is amplified in two stages using a jet pipe and a pilot valve. Draw a block diagram of the system of Figure 4-57 and then find the transfer function between y and x , where x is the air pressure and y is the displacement of the power piston.

B-4-11. Figure 4-58 is a schematic diagram of an aircraft elevator control system. The input to the system is the deflection angle θ of the control lever, and the output is the elevator angle ϕ . Assume that angles θ and ϕ are relatively small. Show that for each angle θ of the control lever there is a corresponding (steady-state) elevator angle ϕ .

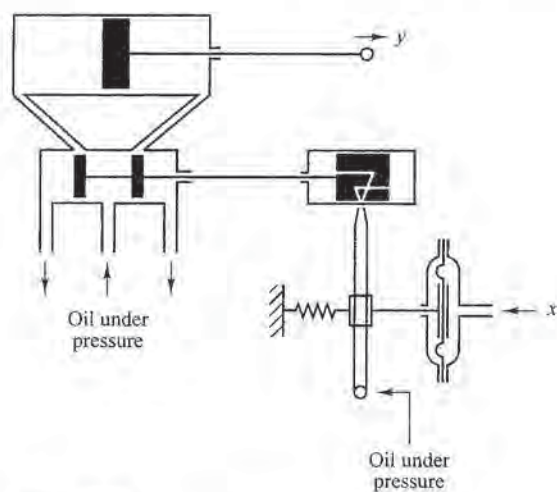


Figure 4-57
Schematic diagram of a hydraulic servomotor.

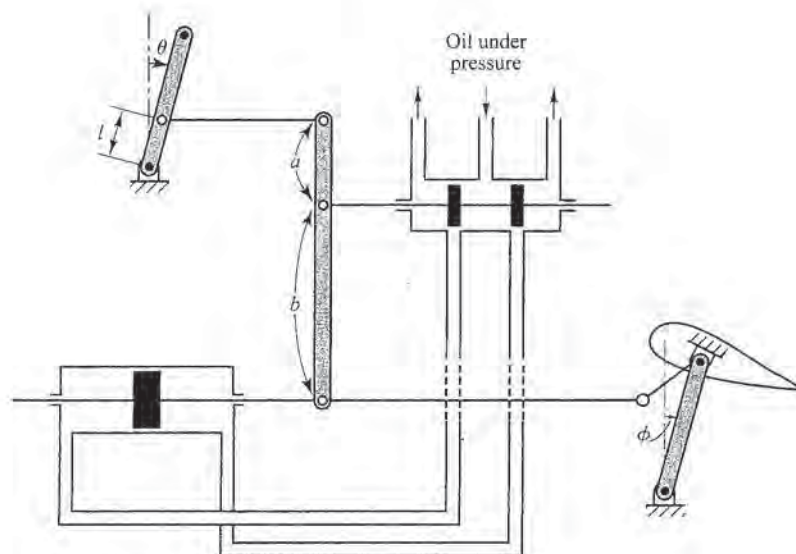


Figure 4-58
Aircraft elevator control system.

B-4-12. Consider the liquid-level control system shown in Figure 4-59. The inlet valve is controlled by a hydraulic integral controller. Assume that the steady-state inflow rate is $\bar{Q}$ and steady-state outflow rate is also $\bar{Q}$, the steady-state head is $\bar{H}$, steady-state pilot valve displacement is $\bar{X} = 0$, and steady-state valve position is $\bar{Y}$. We assume that the set point $\bar{R}$ corresponds to the steady-state head $\bar{H}$. The set point is fixed. Assume also that the disturbance inflow rate q_d , which is a small quantity, is applied to the water tank at $t = 0$. This disturbance causes the head to change from $\bar{H}$ to $\bar{H} + h$. This change results in a change in the outflow rate by q_o . Through the hydraulic controller, the change in head causes a change in the inflow rate from $\bar{Q}$ to $\bar{Q} + q_i$. (The integral controller tends to keep the head constant as much as possible in the presence of disturbances.) We assume that all changes are of small quantities.

We assume that the velocity of the power piston (valve) is proportional to pilot valve displacement x , or

$$\frac{dy}{dt} = K_1 x$$

where K_1 is a positive constant. We also assume that the change in the inflow rate q_i is negatively proportional to the change in the valve opening y , or

$$q_i = -K_v y$$

where K_v is a positive constant.

Assuming the following numerical values for the system,

$$\begin{aligned} C &= 2 \text{ m}^2, & R &= 0.5 \text{ sec/m}^2, & K_v &= 1 \text{ m}^2/\text{sec} \\ a &= 0.25 \text{ m}, & b &= 0.75 \text{ m}, & K_1 &= 4 \text{ sec}^{-1} \end{aligned}$$

obtain the transfer function $H(s)/Q_d(s)$.

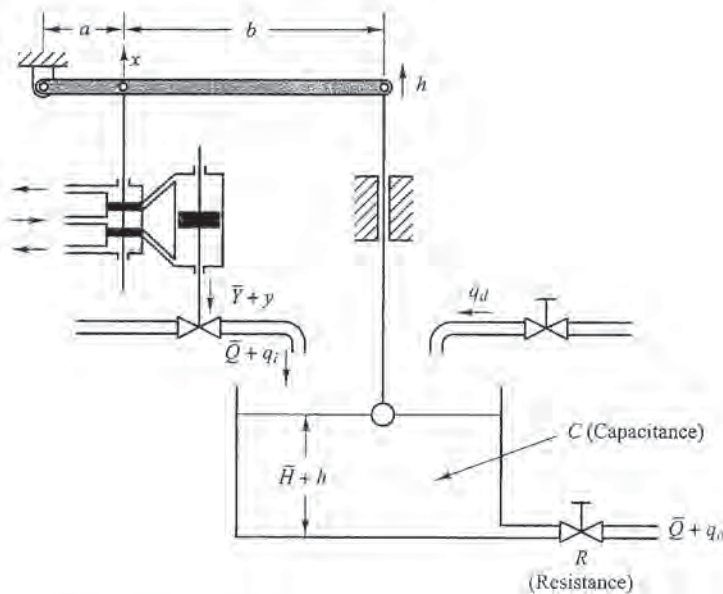


Figure 4-59
Liquid-level control system.

B-4-13. Consider the controller shown in Figure 4-60. The input is the air pressure p_i measured from some steady-state reference pressure $\bar{P}$ and the output is the displacement y of the power piston. Obtain the transfer function $Y(s)/P_i(s)$.

B-4-14. A thermocouple has a time constant of 2 sec. A thermal well has a time constant of 30 sec. When the thermocouple is inserted into the well, this temperature-

measuring device can be considered a two-capacitance system.

Determine the time constants of the combined thermocouple-thermal well system. Assume that the weight of the thermocouple is 8 g and the weight of the thermal well is 40 g. Assume also that the specific heats of the thermocouple and thermal well are the same.

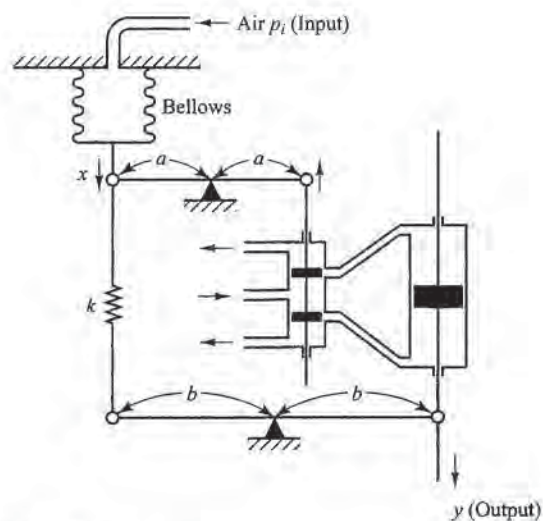


Figure 4-60
Controller.



Transient and Steady-State Response Analyses

5-1 INTRODUCTION

It was stated in Chapter 3 that the first step in analyzing a control system was to derive a mathematical model of the system. Once such a model is obtained, various methods are available for the analysis of system performance.

In practice, the input signal to a control system is not known ahead of time but is random in nature, and the instantaneous input cannot be expressed analytically. Only in some special cases is the input signal known in advance and expressible analytically or by curves, such as in the case of the automatic control of cutting tools.

In analyzing and designing control systems, we must have a basis of comparison of performance of various control systems. This basis may be set up by specifying particular test input signals and by comparing the responses of various systems to these input signals.

Many design criteria are based on the response to such signals or on the response of systems to changes in initial conditions (without any test signals). The use of test signals can be justified because of a correlation existing between the response characteristics of a system to a typical test input signal and the capability of the system to cope with actual input signals.

Typical Test Signals. The commonly used test input signals are those of step functions, ramp functions, acceleration functions, impulse functions, sinusoidal functions, and the like. With these test signals, mathematical and experimental analyses of control systems can be carried out easily since the signals are very simple functions of time.

Which of these typical input signals to use for analyzing system characteristics may be determined by the form of the input that the system will be subjected to most frequently under normal operation. If the inputs to a control system are gradually changing functions of time, then a ramp function of time may be a good test signal. Similarly, if a system is subjected to sudden disturbances, a step function of time may be a good test signal; and for a system subjected to shock inputs, an impulse function may be best. Once a control system is designed on the basis of test signals, the performance of the system in response to actual inputs is generally satisfactory. The use of such test signals enables one to compare the performance of all systems on the same basis.

Transient Response and Steady-State Response. The time response of a control system consists of two parts: the transient response and the steady-state response. By transient response, we mean that which goes from the initial state to the final state. By steady-state response, we mean the manner in which the system output behaves as t approaches infinity. Thus the system response $c(t)$ may be written as

$$c(t) = c_{tr}(t) + c_{ss}(t)$$

where the first term on the right-hand side of the equation is the transient response and the second term is the steady-state response.

Absolute Stability, Relative Stability, and Steady-State Error. In designing a control system, we must be able to predict the dynamic behavior of the system from a knowledge of the components. The most important characteristic of the dynamic behavior of a control system is absolute stability, that is, whether the system is stable or unstable. A control system is in equilibrium if, in the absence of any disturbance or input, the output stays in the same state. A linear time-invariant control system is stable if the output eventually comes back to its equilibrium state when the system is subjected to an initial condition. A linear time-invariant control system is critically stable if oscillations of the output continue forever. It is unstable if the output diverges without bound from its equilibrium state when the system is subjected to an initial condition. Actually, the output of a physical system may increase to a certain extent but may be limited by mechanical "stops," or the system may break down or become nonlinear after the output exceeds a certain magnitude so that the linear differential equations no longer apply.

Important system behavior (other than absolute stability) to which we must give careful consideration includes relative stability and steady-state error. Since a physical control system involves energy storage, the output of the system, when subjected to an input, cannot follow the input immediately but exhibits a transient response before a steady state can be reached. The transient response of a practical control system often exhibits damped oscillations before reaching a steady state. If the output of a system at steady state does not exactly agree with the input, the system is said to have steady-state error. This error is indicative of the accuracy of the system. In analyzing a control system, we must examine transient-response behavior and steady-state behavior.

Outline of the Chapter. This chapter is concerned with system responses to aperiodic signals (such as step, ramp, acceleration, and impulse functions of time). The outline of the chapter is as follows: Section 5-1 has presented introductory material for the chapter. Section 5-2 treats the response of first-order systems to aperiodic inputs. Section 5-3 deals with the transient response of the second-order systems. Detailed

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

$$\frac{C(s)}{R(s)} = \frac{1}{Ts + 1} \quad (5-1)$$

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

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

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

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

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.

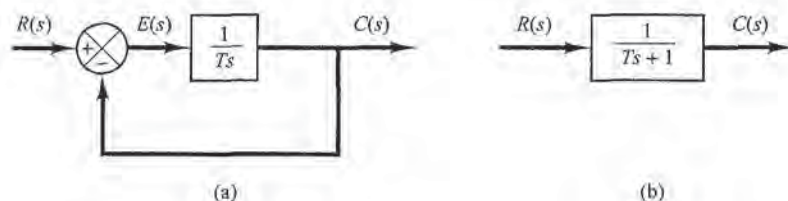
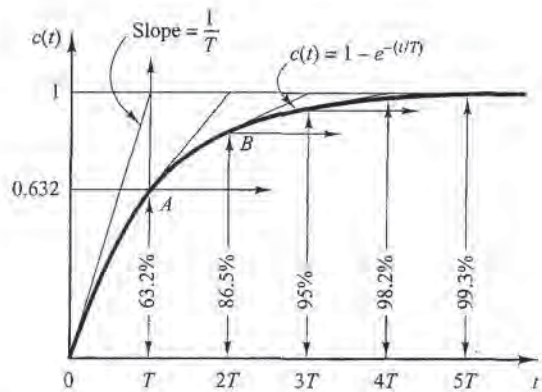


Figure 5-2
Exponential
response curve.



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

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

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

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

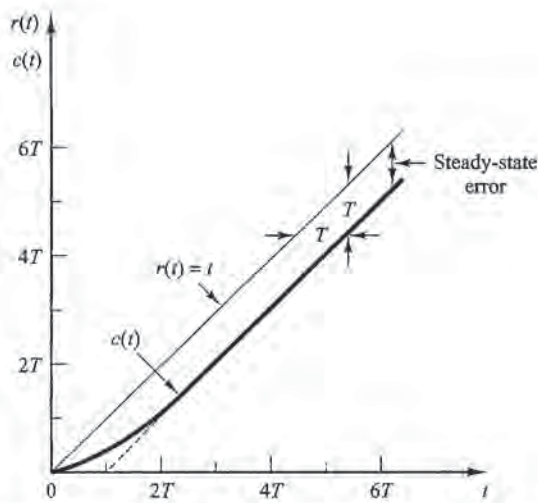
Taking the inverse Laplace transform of Equation (5-5), we obtain

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

The error signal $e(t)$ is then

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

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

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

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

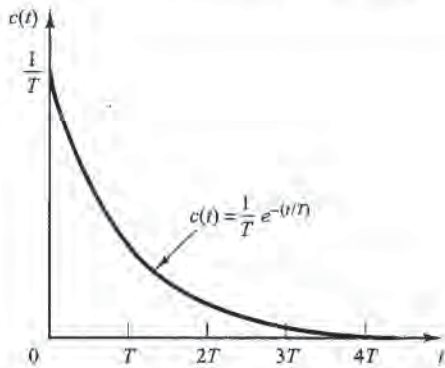


Figure 5-4
Unit-impulse
response of the
system shown in
Figure 5-1(a).

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

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

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

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

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

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

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.)

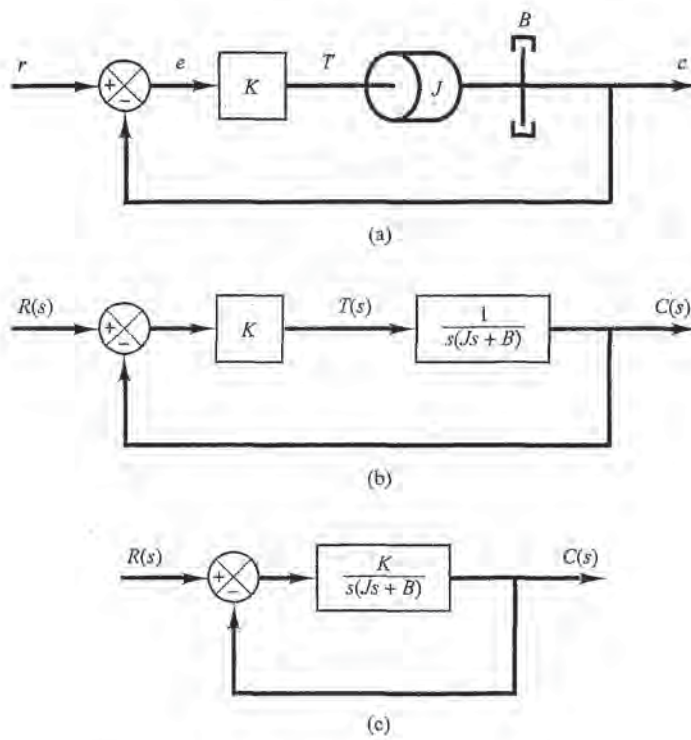


Figure 5-5
 (a) Servo system;
 (b) block diagram;
 (c) simplified block diagram.

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

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

which can be rewritten as

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

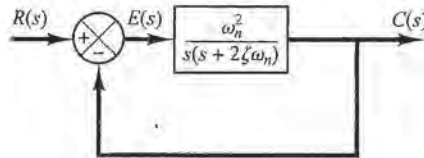
The closed-loop poles are complex conjugates if $B^2 - 4JK < 0$ and they are real if $B^2 - 4JK \geq 0$. In the transient-response analysis, it is convenient to write

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

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

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

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

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

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

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

In Chapter 2 it was shown that

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

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

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

This result can be obtained directly by using a table of Laplace transforms. 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

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

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

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

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

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

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

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

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

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

(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

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

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

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

where $s_1 = (\zeta + \sqrt{\zeta^2 - 1})\omega_n$ and $s_2 = (\zeta - \sqrt{\zeta^2 - 1})\omega_n$. 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

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

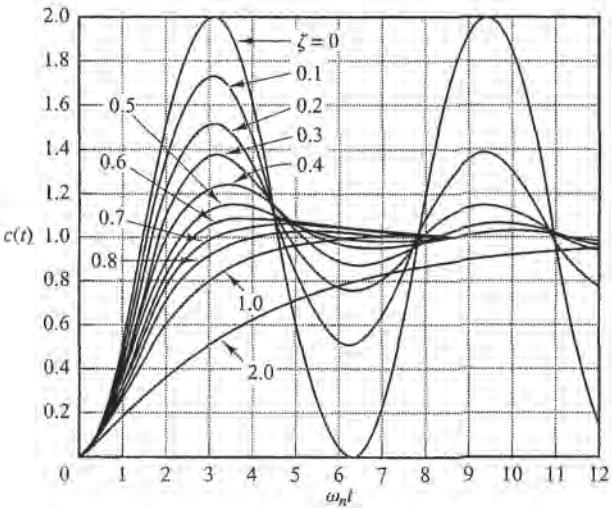
The time response $c(t)$ is then

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

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

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

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



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.

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.

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.

Definitions of Transient-Response Specifications. In many practical cases, the desired performance characteristics of control systems are specified in terms of time-domain quantities. Systems with energy storage cannot respond instantaneously and will exhibit transient responses whenever they are subjected to inputs or disturbances.

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.

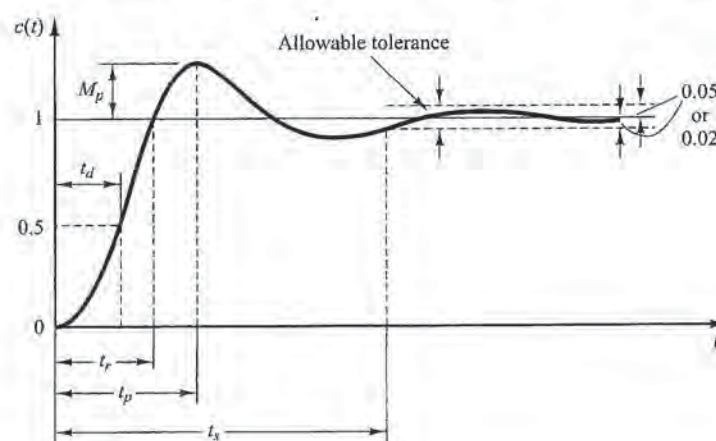


Figure 5-8
Unit-step response
curve showing t_d , t_r ,
 t_p , M_p , and t_s .

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.)

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 5-9.)

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 ($\zeta < 0.4$) yield excessive overshoot in the transient response, and a system with a large value of ζ ($\zeta > 0.8$) 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 $c(t_r) = 1$.

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

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

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

or

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

Thus, the rise time t_r is

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

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

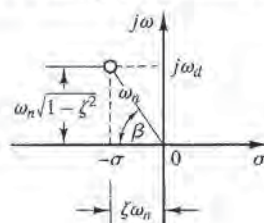


Figure 5-9
Definition of the
angle β .

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

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

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

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

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

$$t_p = \frac{\pi}{\omega_d} \quad (5-20)$$

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

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

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:

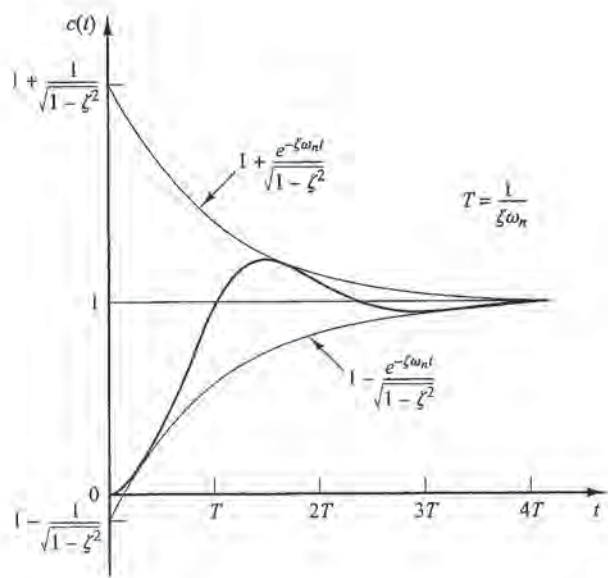
$$M_p = \frac{c(t_p) - c(\infty)}{c(\infty)}$$

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

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

The curves $1 \pm (e^{-\zeta \omega_n t}/\sqrt{1-\zeta^2})$ 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 $1/\zeta \omega_n$.

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



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

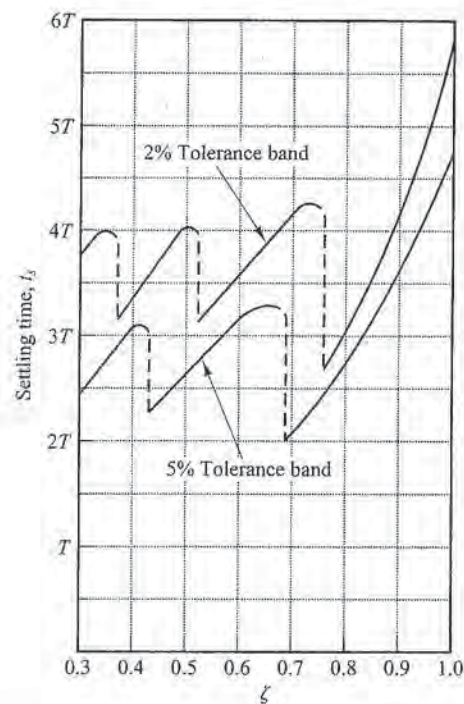
$$t_s = 4T = \frac{4}{\sigma} = \frac{4}{\zeta\omega_n} \quad (2\% \text{ criterion}) \quad (5-22)$$

or

$$t_s = 3T = \frac{3}{\sigma} = \frac{3}{\zeta\omega_n} \quad (5\% \text{ criterion}) \quad (5-23)$$

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 is determined primarily by the undamped natural frequency ω_n . This means that the

Figure 5-11
Settling time t_s
versus ζ curves.



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%.

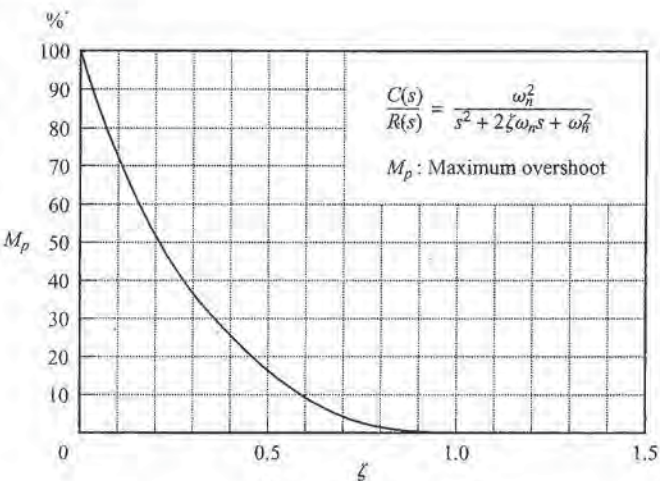


Figure 5-12
 M_p versus ζ curve.

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

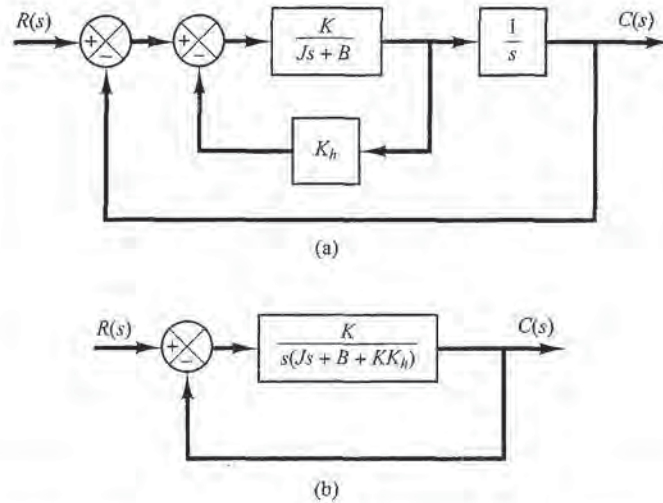
$$t_s = \frac{4}{\sigma} = \frac{4}{3} = 1.33 \text{ sec}$$

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

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

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

$$\zeta = \frac{B + KK_h}{2\sqrt{KJ}} \quad (5-25)$$

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.

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 $J = 1 \text{ kg-m}^2$ and $B = 1 \text{ N-m/rad/sec}$.

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 $0 \leq \zeta < 1$,

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

For $\zeta = 1$,

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

For $\zeta > 1$,

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

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, $c(t) \geq 0$. 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.

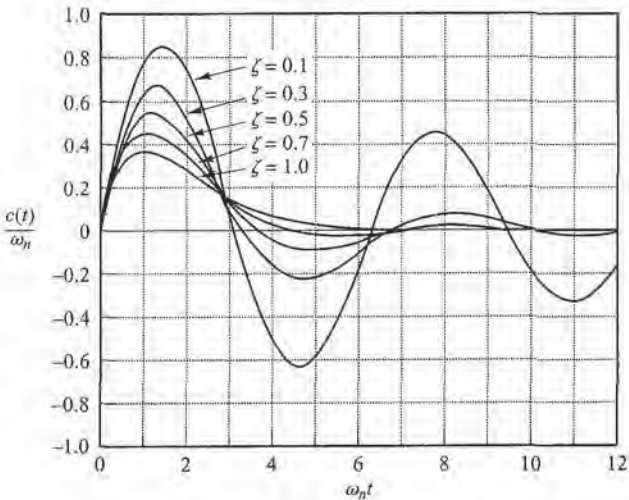
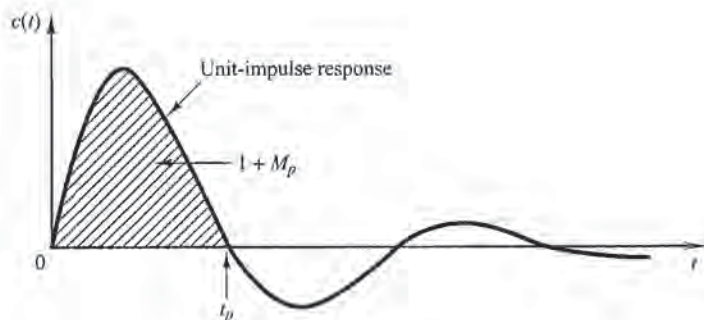


Figure 5-14
Unit-impulse
response curves of
the system shown in
Figure 5-6.

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

$$t = \frac{\tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}}{\omega_n \sqrt{1-\zeta^2}}, \quad \text{where } 0 < \zeta < 1 \quad (5-29)$$

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

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

[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 higher-order systems is the sum of the responses of first-order and second-order systems.

Transient Response of Higher-Order Systems. Consider the system shown in Figure 5-16. The closed-loop transfer function is

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \quad (5-31)$$

In general, $G(s)$ and $H(s)$ are given as ratios of polynomials in s , or

$$G(s) = \frac{p(s)}{q(s)} \quad \text{and} \quad H(s) = \frac{n(s)}{d(s)}$$

where $p(s)$, $q(s)$, $n(s)$, and $d(s)$ are polynomials in s . The closed-loop transfer function given by Equation (5-31) may then be written

$$\begin{aligned} \frac{C(s)}{R(s)} &= \frac{p(s)d(s)}{q(s)d(s) + p(s)n(s)} \\ &= \frac{b_0s^m + b_1s^{m-1} + \cdots + b_{m-1}s + b_m}{a_0s^n + a_1s^{n-1} + \cdots + a_{n-1}s + a_n} \quad (m \leq n) \end{aligned}$$

The transient response of this system to any given input can be obtained by a computer simulation. (See Section 5-5.) If an analytical expression for the transient response is desired, then it is necessary to factor the denominator polynomial. [MATLAB may be used for finding the roots of the denominator polynomial. Use the command `roots(den)`.] Once the numerator and the denominator have been factored, $C(s)/R(s)$ can be written in the form

$$\frac{C(s)}{R(s)} = \frac{K(s + z_1)(s + z_2) \cdots (s + z_m)}{(s + p_1)(s + p_2) \cdots (s + p_n)} \quad (5-32)$$

Let us examine the response behavior of this system to a unit-step input. Consider first the case where the closed-loop poles are all real and distinct. For a unit-step input, Equation (5-32) can be written

$$C(s) = \frac{a}{s} + \sum_{i=1}^n \frac{a_i}{s + p_i} \quad (5-33)$$

where a_i is the residue of the pole at $s = -p_i$. (If the system involves multiple poles, then $C(s)$ will have multiple-pole terms.) [The partial-fraction expansion of $C(s)$, as given by Equation (5-33), can be obtained easily with MATLAB. Use the `residue` command.]

If all closed-loop poles lie in the left-half s plane, the relative magnitudes of the residues determine the relative importance of the components in the expanded form of

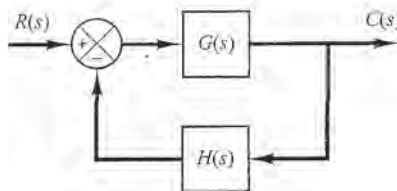


Figure 5-16
Control system.

$C(s)$. If there is a closed-loop zero close to a closed-loop pole, then the residue at this pole is small and the coefficient of the transient-response term corresponding to this pole becomes small. A pair of closely located poles and zeros will effectively cancel each other. If a pole is located very far from the origin, the residue at this pole may be small. The transients corresponding to such a remote pole are small and last a short time. Terms in the expanded form of $C(s)$ having very small residues contribute little to the transient response, and these terms may be neglected. If this is done, the higher-order system may be approximated by a lower-order one. (Such an approximation often enables us to estimate the response characteristics of a higher-order system from those of a simplified one.)

Next, consider the case where the poles of $C(s)$ consist of real poles and pairs of complex-conjugate poles. A pair of complex-conjugate poles yields a second-order term in s . Since the factored form of the higher-order characteristic equation consists of first- and second-order terms, Equation (5-33) can be rewritten

$$C(s) = \frac{a}{s} + \sum_{j=1}^q \frac{a_j}{s + p_j} + \sum_{k=1}^r \frac{b_k(s + \zeta_k \omega_k) + c_k \omega_k \sqrt{1 - \zeta_k^2}}{s^2 + 2\zeta_k \omega_k s + \omega_k^2} \quad (q + 2r = n)$$

where we assumed all closed-loop poles are distinct. [If the closed-loop poles involve multiple poles, $C(s)$ must have multiple-pole terms.] From this last equation, we see that the response of a higher-order system is composed of a number of terms involving the simple functions found in the responses of first- and second-order systems. The unit-step response $c(t)$, the inverse Laplace transform of $C(s)$, is then

$$c(t) = a + \sum_{j=1}^q a_j e^{-p_j t} + \sum_{k=1}^r b_k e^{-\zeta_k \omega_k t} \cos \omega_k \sqrt{1 - \zeta_k^2} t + \sum_{k=1}^r c_k e^{-\zeta_k \omega_k t} \sin \omega_k \sqrt{1 - \zeta_k^2} t, \quad \text{for } t \geq 0 \quad (5-34)$$

Thus the response curve of a stable higher-order system is the sum of a number of exponential curves and damped sinusoidal curves.

If all closed-loop poles lie in the left-half s plane, then the exponential terms and the damped exponential terms in Equation (5-34) will approach zero as time t increases. The steady-state output is then $c(\infty) = a$.

Let us assume that the system considered is a stable one. Then the closed-loop poles that are located far from the $j\omega$ axis have large negative real parts. The exponential terms that correspond to these poles decay very rapidly to zero. (Note that the horizontal distance from a closed-loop pole to the $j\omega$ axis determines the settling time of transients due to that pole. The smaller the distance is, the longer the settling time.)

Remember that the type of transient response is determined by the closed-loop poles, while the shape of the transient response is primarily determined by the closed-loop zeros. As we have seen earlier, the poles of the input $R(s)$ yield the steady-state response terms in the solution, while the poles of $C(s)/R(s)$ enter into the exponential transient-response terms and/or damped sinusoidal transient-response terms. The zeros of $C(s)/R(s)$ do not affect the exponents in the exponential terms, but they do affect the magnitudes and signs of the residues.

Dominant Closed-Loop Poles. The relative dominance of closed-loop poles is determined by the ratio of the real parts of the closed-loop poles, as well as by the relative magnitudes of the residues evaluated at the closed-loop poles. The magnitudes of the residues depend on both the closed-loop poles and zeros.

If the ratios of the real parts exceed 5 and there are no zeros nearby, then the closed-loop poles nearest the $j\omega$ axis will dominate in the transient-response behavior because these poles correspond to transient-response terms that decay slowly. Those closed-loop poles that have dominant effects on the transient-response behavior are called *dominant closed-loop poles*. Quite often the dominant closed-loop poles occur in the form of a complex-conjugate pair. The dominant closed-loop poles are most important among all closed-loop poles.

Note that the gain of a higher-order system is often adjusted so that there will exist a pair of dominant complex-conjugate closed-loop poles. The presence of such poles in a stable system reduces the effects of such nonlinearities as dead zone, backlash, and coulomb-friction.

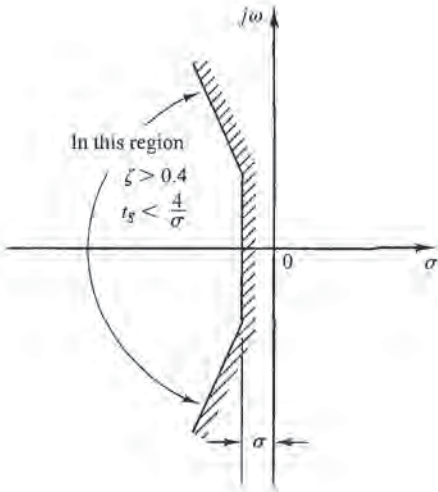
Stability Analysis in the Complex Plane. The stability of a linear closed-loop system can be determined from the location of the closed-loop poles in the s plane. If any of these poles lie in the right-half s plane, then with increasing time they give rise to the dominant mode, and the transient response increases monotonically or oscillates with increasing amplitude. This represents an unstable system. For such a system, as soon as the power is turned on, the output may increase with time. If no saturation takes place in the system and no mechanical stop is provided, then the system may eventually be subjected to damage and fail since the response of a real physical system cannot increase indefinitely. Therefore, closed-loop poles in the right-half s plane are not permissible in the usual linear control system. If all closed-loop poles lie to the left of the $j\omega$ axis, any transient response eventually reaches equilibrium. This represents a stable system.

Whether a linear system is stable or unstable is a property of the system itself and does not depend on the input or driving function of the system. The poles of the input, or driving function, do not affect the property of stability of the system, but they contribute only to steady-state response terms in the solution. Thus, the problem of absolute stability can be solved readily by choosing no closed-loop poles in the right-half s plane, including the $j\omega$ axis. (Mathematically, closed-loop poles on the $j\omega$ axis will yield oscillations, the amplitude of which is neither decaying nor growing with time. In practical cases, where noise is present, however, the amplitude of oscillations may increase at a rate determined by the noise power level. Therefore, a control system should not have closed-loop poles on the $j\omega$ axis.)

Note that the mere fact that all closed-loop poles lie in the left-half s plane does not guarantee satisfactory transient-response characteristics. If dominant complex-conjugate closed-loop poles lie close to the $j\omega$ axis, the transient response may exhibit excessive oscillations or may be very slow. Therefore, to guarantee fast, yet well-damped, transient-response characteristics, it is necessary that the closed-loop poles of the system lie in a particular region in the complex plane, such as the region bounded by the shaded area in Figure 5-17.

Since the relative stability and transient-response performance of a closed-loop control system are directly related to the closed-loop pole-zero configuration in the s plane,

Figure 5-17
Region in the
complex plane
satisfying the
conditions $\zeta > 0.4$
and $t_s < 4/\sigma$.



it is frequently necessary to adjust one or more system parameters in order to obtain suitable configurations. The effects of varying system parameters on the closed-loop poles will be discussed in detail in Chapter 6.

5-5 TRANSIENT-RESPONSE ANALYSIS WITH MATLAB

Introduction. The practical procedure for plotting time response curves of systems higher than second-order is through computer simulation. In this section we present the computational approach to the transient-response analysis with MATLAB. In particular, we discuss step response, impulse response, ramp response, and responses to other simple inputs.

MATLAB Representation of Linear Systems. The transfer function of a system is represented by two arrays of numbers. Consider the system

$$\frac{C(s)}{R(s)} = \frac{2s + 25}{s^2 + 4s + 25} \tag{5-35}$$

This system can be represented as two arrays, each containing the coefficients of the polynomials in decreasing powers of s as follows:

$$\begin{aligned} \text{num} &= [2 \ 25] \\ \text{den} &= [1 \ 4 \ 25] \end{aligned}$$

An alternative representation is

$$\begin{aligned} \text{num} &= [0 \ 2 \ 25] \\ \text{den} &= [1 \ 4 \ 25] \end{aligned}$$

In this expression a zero is padded. Note that if zeros are padded, the dimensions of “num” vector and “den” vector become the same. An advantage of padding zeros is that the “num” vector and “den” vector can be directly added. For example,

$$\begin{aligned}\text{num} + \text{den} &= [0 \ 2 \ 25] + [1 \ 4 \ 25] \\ &= [1 \ 6 \ 50]\end{aligned}$$

If num and den (the numerator and denominator of the closed-loop transfer function) are known, commands such as

$$\text{step}(\text{num},\text{den}), \quad \text{step}(\text{num},\text{den},t)$$

will generate plots of unit-step responses (t in the step command is the user-specified time.)

For a control system defined in a state-space form, where state matrix **A**, control matrix **B**, output matrix **C**, and direct transmission matrix **D** of state-space equations are known, the command

$$\text{step}(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D})$$

will generate plots of unit-step responses. The time vector is automatically determined when t is not explicitly included in the step commands.

Note that the command $\text{step}(\text{sys})$ may be used to obtain the unit-step response of a system. First, define the system by

$$\text{sys} = \text{tf}(\text{num},\text{den})$$

or

$$\text{sys} = \text{ss}(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D})$$

Then, to obtain, for example, the unit-step response, enter

$$\text{step}(\text{sys})$$

into the computer.

When step commands have left-hand arguments such as

$$\begin{aligned}[y,x,t] &= \text{step}(\text{num},\text{den},t) \\ [y,x,t] &= \text{step}(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\text{iu}) \\ [y,x,t] &= \text{step}(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\text{iu},t)\end{aligned}\tag{5-36}$$

no plot is shown on the screen. Hence it is necessary to use a plot command to see the response curves. The matrices y and x contain the output and state response of the system, respectively, evaluated at the computation time points t. (y has as many columns as outputs and one row for each element in t. x has as many columns as states and one row for each element in t.)

Note in Equation (5-36) that the scalar iu is an index into the inputs of the system and specifies which input is to be used for the response, and t is the user-specified time. If the system involves multiple inputs and multiple outputs, the step command, such as given by Equation (5-36), produces a series of step response plots, one for each input and output combination of

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} &= \mathbf{Cx} + \mathbf{Du}\end{aligned}$$

(For details, see Example 5-3.)

EXAMPLE 5-3 Consider the following system:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 6.5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Obtain the unit-step response curves.

Although it is not necessary to obtain the transfer matrix expression for the system to obtain the unit-step response curves with MATLAB, we shall derive such an expression for reference. For the system defined by

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}$$

$$\mathbf{y} = \mathbf{Cx} + \mathbf{Du}$$

the transfer matrix $\mathbf{G}(s)$ is a matrix that relates $\mathbf{Y}(s)$ and $\mathbf{U}(s)$ as follows:

$$\mathbf{Y}(s) = \mathbf{G}(s)\mathbf{U}(s)$$

Taking Laplace transforms of the state-space equations, we obtain

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{AX}(s) + \mathbf{BU}(s) \quad (5-37)$$

$$\mathbf{Y}(s) = \mathbf{CX}(s) + \mathbf{DU}(s) \quad (5-38)$$

In deriving the transfer matrix, we assume that $\mathbf{x}(0) = \mathbf{0}$. Then, from Equation (5-37), we get

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{BU}(s) \quad (5-39)$$

Substituting Equation (5-39) into Equation (5-38), we obtain

$$\mathbf{Y}(s) = [\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\mathbf{U}(s)$$

Thus the transfer matrix $\mathbf{G}(s)$ is given by

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$$

The transfer matrix $\mathbf{G}(s)$ for the given system becomes

$$\begin{aligned} \mathbf{G}(s) &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s+1 & 1 \\ -6.5 & s \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \frac{1}{s^2 + s + 6.5} \begin{bmatrix} s & -1 \\ 6.5 & s+1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \frac{1}{s^2 + s + 6.5} \begin{bmatrix} s-1 & s \\ s+7.5 & 6.5 \end{bmatrix} \end{aligned}$$

Hence

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \begin{bmatrix} \frac{s-1}{s^2 + s + 6.5} & \frac{s}{s^2 + s + 6.5} \\ \frac{s+7.5}{s^2 + s + 6.5} & \frac{6.5}{s^2 + s + 6.5} \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix}$$

Since this system involves two inputs and two outputs, four transfer functions may be defined depending on which signals are considered as input and output. Note that, when considering the

signal u_1 as the input, we assume that signal u_2 is zero, and vice versa. The four transfer functions are

$$\begin{aligned} \frac{Y_1(s)}{U_1(s)} &= \frac{s-1}{s^2+s+6.5}, & \frac{Y_1(s)}{U_2(s)} &= \frac{s}{s^2+s+6.5} \\ \frac{Y_2(s)}{U_1(s)} &= \frac{s+7.5}{s^2+s+6.5}, & \frac{Y_2(s)}{U_2(s)} &= \frac{6.5}{s^2+s+6.5} \end{aligned}$$

The four individual step-response curves can be plotted by use of the command

```
step(A,B,C,D)
```

MATLAB Program 5-1 produces four such step-response curves. The curves are shown in Figure 5-18.

MATLAB Program 5-1

A = [-1 -1;6.5 0];
B = [1 1;1 0];
C = [1 0;0 1];
D = [0 0;0 0];
step(A,B,C,D)

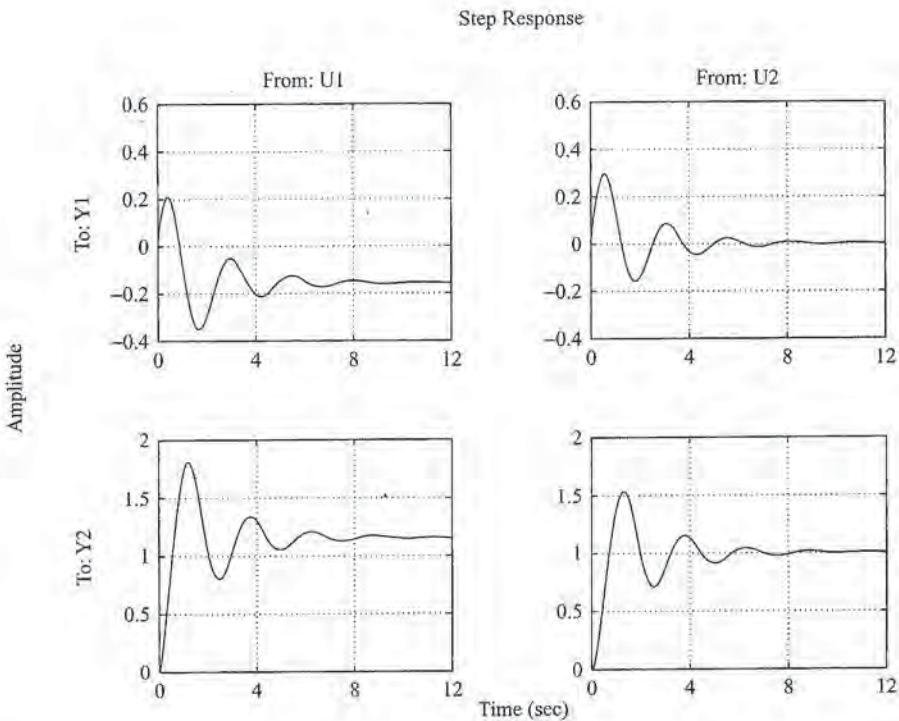


Figure 5-18
Unit-step response
curves.

To plot two step-response curves for the input u_1 in one diagram and two step-response curves for the input u_2 in another diagram, we may use the commands

`step(A,B,C,D,1)`

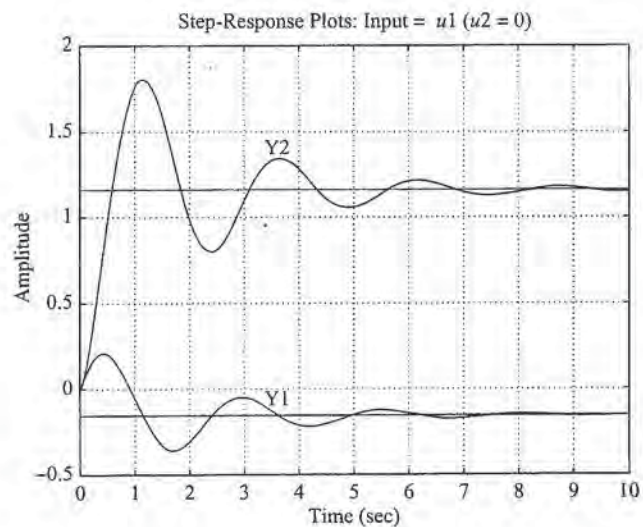
and

`step(A,B,C,D,2)`

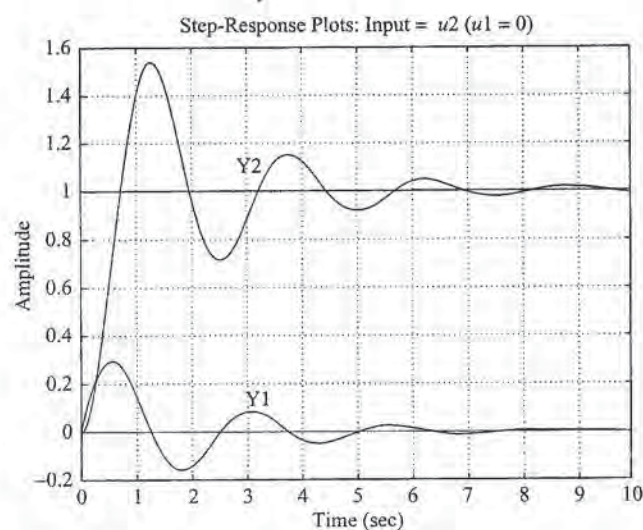
respectively. MATLAB Program 5-2 is a program to plot two step-response curves for the input u_1 in one diagram and two step-response curves for the input u_2 in another diagram. Figure 5-19 shows the two diagrams, each consisting of two step-response curves.

MATLAB Program 5-2

```
% ----- Step-response curves for system defined in state
% space -----
% ***** In this program we plot step-response curves of a system
% having two inputs (u1 and u2) and two outputs (y1 and y2) *****
% ***** We shall first plot step-response curves when the input is
% u1. Then we shall plot step-response curves when the input is
% u2 *****
% ***** Enter matrices A, B, C, and D *****
A = [-1 -1;6.5 0];
B = [1 1;1 0];
C = [1 0;0 1];
D = [0 0;0 0];
% ***** To plot step-response curves when the input is u1, enter
% the command 'step(A,B,C,D,1)' *****
step(A,B,C,D,1)
grid
title ('Step-Response Plots: Input = u1 (u2 = 0)')
text(3.4, -0.06, 'Y1')
text(3.4, 1.4, 'Y2')
% ***** Next, we shall plot step-response curves when the input
% is u2. Enter the command 'step(A,B,C,D,2)' *****
step(A,B,C,D,2)
grid
title ('Step-Response Plots: Input = u2 (u1 = 0)')
text(3,0.14, 'Y1')
text(2.8,1.1, 'Y2')
```



(a)



(b)

Figure 5-19
Unit-step response
curves. (a) u_1 is the
input ($u_2 = 0$); (b) u_2
is the input ($u_1 = 0$).

Writing Text on the Graphics Screen. To write text on the graphics screen, enter, for example, the following statements:

```
text(3.4, -0.06, 'Y1')
```

and

```
text(3.4, 1.4, 'Y2')
```

The first statement tells the computer to write 'Y1' beginning at the coordinates $x = 3.4$, $y = -0.06$. Similarly, the second statement tells the computer to write 'Y2' beginning at the coordinates $x = 3.4$, $y = 1.4$. [See MATLAB Program 5-2 and Figure 5-19(a).]

Another way to write a text or texts in the plot is to use the gtext command. The syntax is

gtext('text')

When gtext is executed, the computer waits until the cursor is positioned (using a mouse) at the desired position in the screen. When the left mouse button is pressed, the text enclosed in simple quotes is written on the plot at the cursor's position. Any number of gtext commands can be used in a plot. (See, for example, MATLAB Program 5-15.)

MATLAB Description of Standard Second-Order System. As noted earlier, the second-order system

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (5-40)$$

is called the standard second-order system. Given ω_n and ζ , the command

printsys(num,den) or printsys(num,den,s)

prints num/den as a ratio of polynomials in s .

Consider, for example, the case where $\omega_n = 5$ rad/sec and $\zeta = 0.4$. MATLAB Program 5-3 generates the standard second-order system where $\omega_n = 5$ rad/sec and $\zeta = 0.4$.

MATLAB Program 5-3
<pre> wn = 5; damping_ratio = 0.4; [num0,den] = ord2(wn,damping_ratio); num = 5^2*num0; printsys(num,den,'s') num/den = </pre>
$\frac{25}{s^2 + 4s + 25}$

Obtaining the Unit-Step Response of the Transfer-Function System. Let us consider the unit-step response of the system given by

$$G(s) = \frac{25}{s^2 + 4s + 25}$$

MATLAB Program 5-4 will yield a plot of the unit-step response of this system. A plot of the unit-step response curve is shown in Figure 5-20.

Notice in Figure 5-20 (and many others) that the x axis and y axis labels are automatically determined. If it is desired to label the x axis and y axis differently, we need to modify the step command. For example, if it is desired to label the x axis as 't Sec' and the y axis as 'Input and Output,' then use step-response commands with left-hand arguments, such as

c = step(num,den,t)

or, more generally,

$$[y,x,t] = \text{step}(\text{num},\text{den},t)$$

See, for example, MATLAB Program 5-6.

MATLAB Program 5-4

```
% ----- Unit-step response -----  
% ***** Enter the numerator and denominator of the transfer  
% function *****  
num = [0 0 25];  
den = [1 4 25];  
  
% ***** Enter the following step-response command *****  
step(num,den)  
  
% ***** Enter grid and title of the plot *****  
grid  
title (' Unit-Step Response of G(s) = 25/(s^2+4s+25)')
```

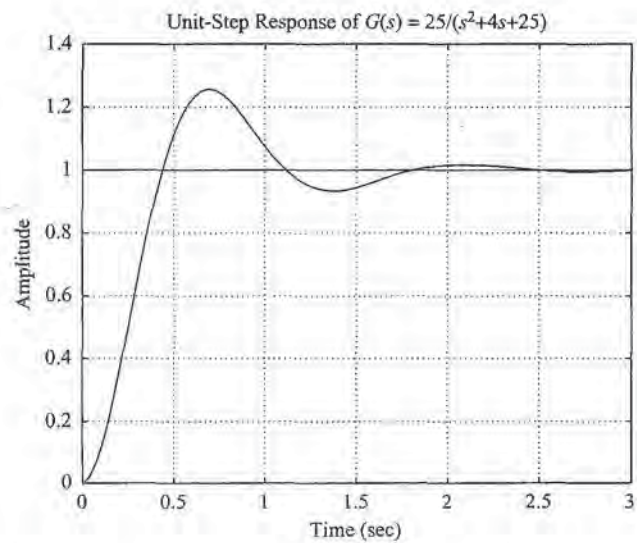


Figure 5-20
Unit-step response
curve.

Obtaining Three-Dimensional Plot of Unit-Step Response Curves with MATLAB. MATLAB enables us to plot three-dimensional plots easily. The command to obtain three-dimensional plots is “mesh.”

EXAMPLE 5-4 Consider the closed-loop system defined by

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

(The undamped natural frequency ω_n is normalized to 1.) Plot unit-step response curves $c(t)$ when ζ assumes the following values:

$$\zeta = 0, 0.2, 0.4, 0.6, 0.8, 1.0$$

Also plot a three-dimensional plot.

An illustrative MATLAB program for plotting a three-dimensional diagram of unit-step response curves of this second-order system is given in MATLAB 5-5. This program includes a program to obtain a two-dimensional plot (ordinary unit-step response curves) and a program to obtain a three-dimensional plot. The resulting plots are shown in Figures 5-21(a) and (b). [Note that mesh(y) will produce a three-dimensional plot the same as Figure 5-21(b) except that x axis and y axis are interchanged. See Problem A-5-17.]

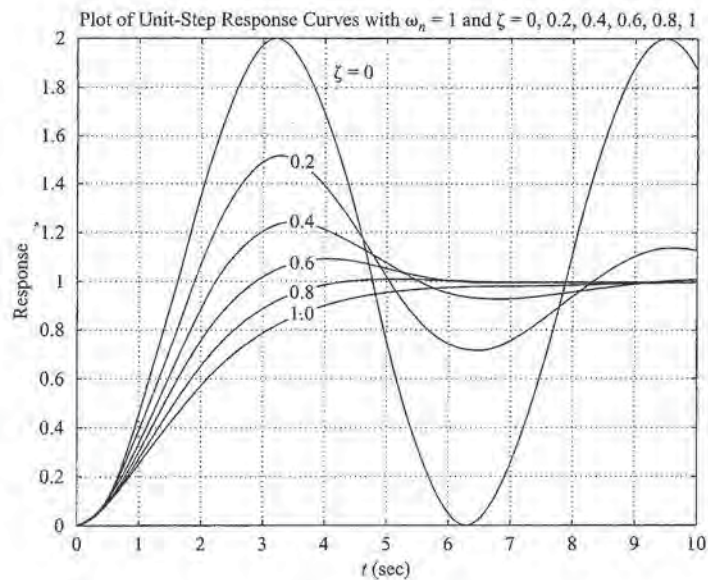
When a MATLAB program involves repetitive computations, many different MATLAB programs can be written. In this book many different MATLAB programs using loops such as "while loop" and "for loop" are presented for illustration purposes. The reader is advised to study all those programs and try to improve them if possible.

MATLAB Program 5-5

```
% ----- Two-dimensional plot and three-dimensional plot of unit-step
% response curves for the standard second-order system with  $\omega_n = 1$ 
% and  $\zeta = 0, 0.2, 0.4, 0.6, 0.8$ , and  $1$ . -----
t = 0:0.2:10;
zeta = [0 0.2 0.4 0.6 0.8 1];
for n = 1:6;
    num = [0 0 1];
    den = [1 2*zeta(n) 1];
    [y(1:51,n),x,t] = step(num,den,t);
end

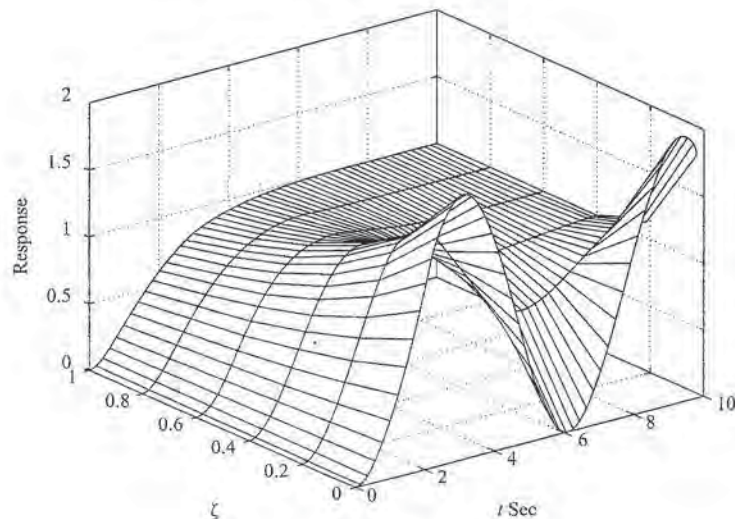
% To plot a two-dimensional diagram, enter the command plot(t,y).
plot(t,y)
grid
title('Plot of Unit-Step Response Curves with  $\omega_n = 1$  and  $\zeta = 0, 0.2, 0.4, 0.6, 0.8, 1$ ')
xlabel('t (sec)')
ylabel('Response')
text(4.1,1.86,'\zeta = 0')
text(3.5,1.5,'0.2')
text(3.5,1.24,'0.4')
text(3.5,1.08,'0.6')
text(3.5,0.95,'0.8')
text(3.5,0.86,'1.0')

% To plot a three-dimensional diagram, enter the command mesh(t,zeta,y').
mesh(t,zeta,y')
title('Three-Dimensional Plot of Unit-Step Response Curves')
xlabel('t Sec')
ylabel('\zeta')
zlabel('Response')
```



(a)

Three-Dimensional Plot of Unit-Step Response Curves



(b)

Figure 5-21
(a) Two-dimensional plots of unit-step response curves for $\zeta = 0, 0.2, 0.4, 0.6, 0.8$, and 1.0 ; (b) three-dimensional plots of unit-step response curves.

Obtaining Rise Time, Peak Time, Maximum Overshoot, and Settling Time with MATLAB. MATLAB can conveniently be used to obtain the rise time, peak time, maximum overshoot, and settling time. Consider the system defined by

$$\frac{C(s)}{R(s)} = \frac{25}{s^2 + 6s + 25}$$

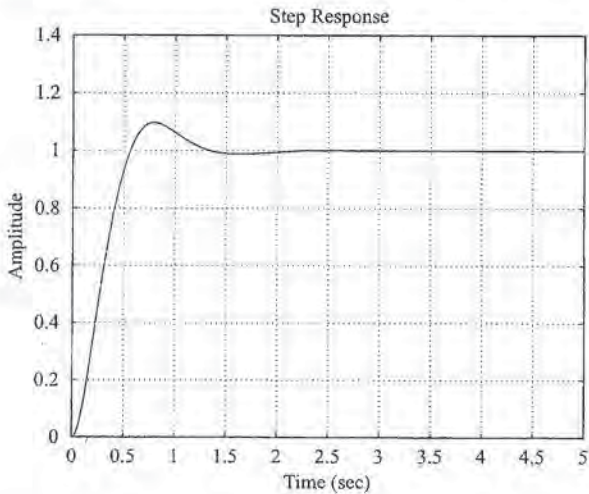
MATLAB Program 5-6 yields the rise time, peak time, maximum overshoot, and settling time. A unit-step response curve for this system is given in Figure 5-22 to verify the

results obtained with MATLAB Program 5-6. (Note that this program can also be applied to higher-order systems. See Problem A-5-12.)

MATLAB Program 5-6

```
% ----- This is a MATLAB program to find the rise time, peak time,
% maximum overshoot, and settling time of the second-order system
% and higher-order system -----
% ----- In this example, we assume zeta = 0.6 and wn = 5 -----
num = [0 0 25];
den = [1 6 25];
t = 0:0.005:5;
[y,x,t] = step(num,den,t);
r = 1; while y(r) < 1.0001; r = r + 1; end;
rise_time = (r - 1)*0.005
rise_time =
    0.5550
[ymax,tp] = max(y);
peak_time = (tp - 1)*0.005
peak_time =
    0.7850
max_overshoot = ymax-1
max_overshoot =
    0.0948
s = 1001; while y(s) > 0.98 & y(s) < 1.02; s = s - 1; end;
settling_time = (s - 1)*0.005
settling_time =
    1.1850
```

Figure 5-22
Unit-step response
curve.



Impulse Response. The unit-impulse response of a control system may be obtained by using one of the following MATLAB commands:

$$\begin{aligned} &\text{impulse}(\text{num},\text{den}) \\ &\text{impulse}(\text{A},\text{B},\text{C},\text{D}) \\ &[y,x,t] = \text{impulse}(\text{num},\text{den}) \\ &[y,x,t] = \text{impulse}(\text{num},\text{den},t) \end{aligned} \tag{5-41}$$

$$\begin{aligned} &[y,x,t] = \text{impulse}(\text{A},\text{B},\text{C},\text{D}) \\ &[y,x,t] = \text{impulse}(\text{A},\text{B},\text{C},\text{D},\text{iu}) \end{aligned} \tag{5-42}$$

$$[y,x,t] = \text{impulse}(\text{A},\text{B},\text{C},\text{D},\text{iu},t) \tag{5-43}$$

The command `impulse(num,den)` plots the unit-impulse response on the screen. The command `impulse(A,B,C,D)` produces a series of unit-impulse-response plots, one for each input and output combination of the system

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} &= \mathbf{Cx} + \mathbf{Du} \end{aligned}$$

with the time vector automatically determined. Note that in Equations (5-42) and (5-43) the scalar `iu` is an index into the inputs of the system and specifies which input to be used for the impulse response.

Note also that in Equations (5-41) and (5-43) `t` is the user-supplied time vector. The vector `t` specifies the times at which the impulse response is to be computed.

If MATLAB is invoked with the left-hand argument `[y,x,t]`, such as in the case of `[y,x,t] = impulse(A,B,C,D)`, the command returns the output and state responses of the system and the time vector `t`. No plot is drawn on the screen. The matrices `y` and `x` contain the output and state responses of the system evaluated at the time points `t`. (`y` has as many columns as outputs and one row for each element in `t`. `x` has as many columns as state variables and one row for each element in `t`.)

EXAMPLE 5-5 Obtain the unit-impulse response of the following system:

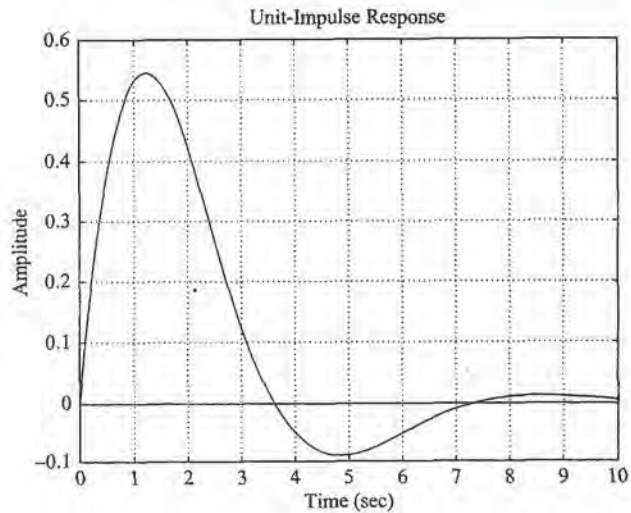
$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y &= [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [0]u \end{aligned}$$

A possible MATLAB program is shown in MATLAB Program 5-7. The resulting response curve is shown in Figure 5-23.

MATLAB Program 5-7

A = [0 1;-1 -1];
B = [0;1];
C = [1 0];
D = [0];
impulse(A,B,C,D);
grid;
title('Unit-Impulse Response')

Figure 5-23
Unit-impulse
response curve.



EXAMPLE 5-6 Obtain the unit-impulse response of the following system:

$$\frac{C(s)}{R(s)} = G(s) = \frac{1}{s^2 + 0.2s + 1}$$

MATLAB Program 5-8 will produce the unit-impulse response. The resulting plot is shown in Figure 5-24.

MATLAB Program 5-8

```
num = [0 0 1];  
den = [1 0.2 1];  
impulse(num,den);  
grid  
title('Unit-Impulse Response of G(s) = 1/(s^2 + 0.2s + 1)')
```

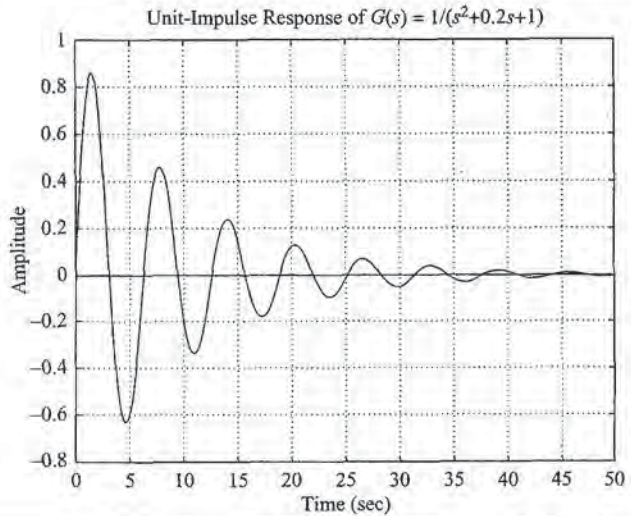


Figure 5-24
Unit-impulse
response curve.

Alternative Approach to Obtain Impulse Response. Note that when the initial conditions are zero the unit-impulse response of $G(s)$ is the same as the unit-step response of $sG(s)$.

Consider the unit-impulse response of the system considered in Example 5–6. Since $R(s) = 1$ for the unit-impulse input, we have

$$\begin{aligned}\frac{C(s)}{R(s)} &= C(s) = G(s) = \frac{1}{s^2 + 0.2s + 1} \\ &= \frac{s}{s^2 + 0.2s + 1} \frac{1}{s}\end{aligned}$$

We can thus convert the unit-impulse response of $G(s)$ to the unit-step response of $sG(s)$.

If we enter the following num and den into MATLAB,

```
num = [0 1 0]
den = [1 0.2 1]
```

and use the step-response command; as given in MATLAB Program 5–9, we obtain a plot of the unit-impulse response of the system as shown in Figure 5–25.

```
MATLAB Program 5–9
num = [0 1 0];
den = [1 0.2 1];
step(num,den);
grid
title('Unit-Step Response of sG(s) = s/(s^2 + 0.2s + 1)')
```

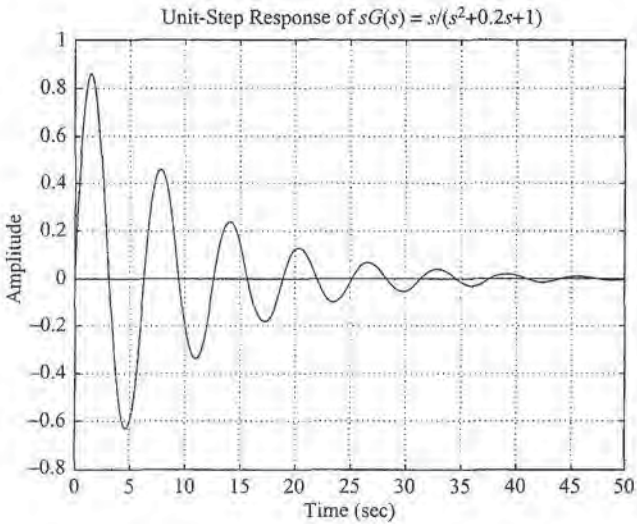


Figure 5–25
Unit-impulse response curve obtained as the unit-step response of $sG(s) = s/(s^2 + 0.2s + 1)$.

Ramp Response. There is no ramp command in MATLAB. Therefore, we need to use the step command or the lsim command (presented later) to obtain the ramp response. Specifically, to obtain the ramp response of the transfer-function system $G(s)$, divide $G(s)$ by s and use the step-response command. For example, consider the closed-loop system

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

For a unit-ramp input, $R(s) = 1/s^2$. Hence

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

To obtain the unit-ramp response of this system, enter the following numerator and denominator into the MATLAB program,

```
num = [0 0 0 1];
den = [1 1 1 0];
```

and use the step-response command. See MATLAB Program 5-10. The plot obtained by using this program is shown in Figure 5-26.

MATLAB Program 5-10

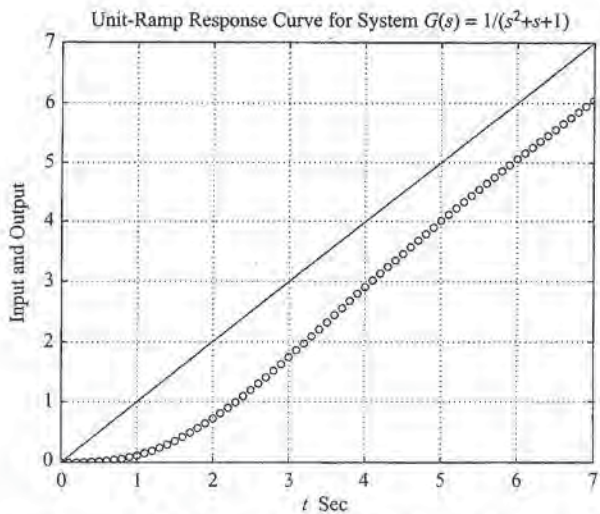
```
% ----- Unit-ramp response -----
% ***** The unit-ramp response is obtained as the unit-step
% response of G(s)/s *****
% ***** Enter the numerator and denominator of G(s)/s *****
num = [0 0 0 1];
den = [1 1 1 0];

% ***** Specify the computing time points (such as t = 0:0.1:7)
% and then enter step-response command: c = step(num,den,t) *****
t = 0:0.1:7;
c = step(num,den,t);

% ***** In plotting the ramp-response curve, add the reference
% input to the plot. The reference input is t. Add to the
% argument of the plot command with the following: t,t,'-'. Thus
% the plot command becomes as follows: plot(t,c,'o',t,t,'-') *****
plot(t,c,'o',t,t,'-')

% ***** Add grid, title, xlabel, and ylabel *****
grid
title('Unit-Ramp Response Curve for System G(s) = 1/(s^2 + s + 1)')
xlabel('t Sec')
ylabel('Input and Output')
```

Figure 5-26
Unit-ramp response curve.



Unit-Ramp Response of a System Defined in State Space. Next, we shall treat the unit-ramp response of the system in state-space form. Consider the system described by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$$

$$y = \mathbf{C}\mathbf{x} + Du$$

where u is the unit-ramp function. In what follows, we shall consider a simple example to explain the method. Consider the case where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}(0) = \mathbf{0}$$

$$\mathbf{C} = [1 \ 0], \quad D = [0]$$

When the initial conditions are zeros, the unit-ramp response is the integral of the unit-step response. Hence the unit-ramp response can be given by

$$z = \int_0^t y \, dt \tag{5-44}$$

From Equation (5-44), we obtain

$$\dot{z} = y = x_1 \tag{5-45}$$

Let us define

$$z = x_3$$

Then Equation (5-45) becomes

$$\dot{x}_3 = x_1 \tag{5-46}$$

Combining Equation (5-46) with the original state-space equation, we obtain

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u \tag{5-47}$$

$$z = [0 \ 0 \ 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{5-48}$$

where u appearing in Equation (5-47) is the unit-step function. These equations can be written as

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{AAx} + \mathbf{BB}u \\ z &= \mathbf{CCx} + \mathbf{DD}u\end{aligned}$$

where

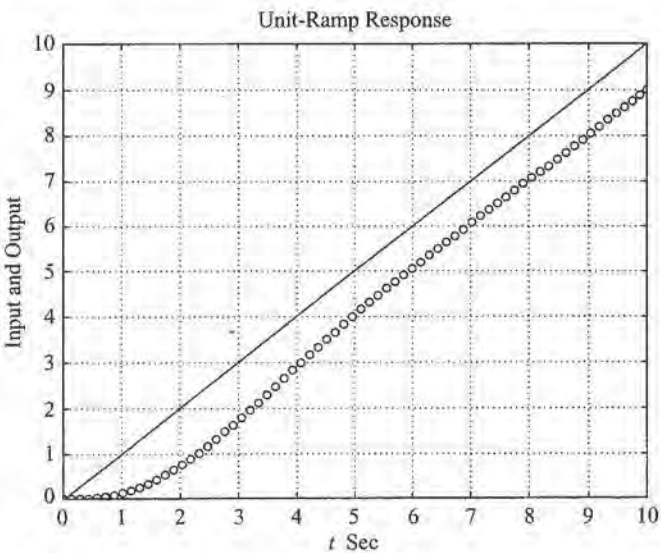
$$\begin{aligned}\mathbf{AA} &= \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \left[\begin{array}{c|c} \mathbf{A} & 0 \\ \hline \mathbf{C} & 0 \end{array} \right] \\ \mathbf{BB} &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{B} \\ 0 \end{bmatrix}, \quad \mathbf{CC} = [0 \ 0 \ 1], \quad \mathbf{DD} = [0]\end{aligned}$$

Note that x_3 is the third element of $\mathbf{x}$. A plot of the unit-ramp response curve $z(t)$ can be obtained by entering MATLAB Program 5-11 into the computer. A plot of the unit-ramp response curve obtained from this MATLAB program is shown in Figure 5-27.

MATLAB Program 5-11

```
% ----- Unit-ramp response -----
% ***** The unit-ramp response is obtained by adding a new
% state variable x3. The dimension of the state equation
% is enlarged by one *****
% ***** Enter matrices A, B, C, and D of the original state
% equation and output equation *****
A = [0 1;-1 -1];
B = [0; 1];
C = [1 0];
D = [0];
% ***** Enter matrices AA, BB, CC, and DD of the new,
% enlarged state equation and output equation *****
AA = [A zeros(2,1);C 0];
BB = [B;0];
CC = [0 0 1];
DD = [0];
% ***** Enter step-response command: [z,x,t] = step(AA,BB,CC,DD) *****
[z,x,t] = step(AA,BB,CC,DD);
% ***** In plotting x3 add the unit-ramp input t in the plot
% by entering the following command: plot(t,x3,'o',t,t,'-') *****
x3 = [0 0 1]*x'; plot(t,x3,'o',t,t,'-')
grid
title('Unit-Ramp Response')
xlabel('t Sec')
ylabel('Input and Output')
```

Figure 5-27
Unit-ramp response
curve.



Obtaining Response to Arbitrary Input. To obtain the response to an arbitrary input, the command `lsim` may be used. The commands like

```
lsim(num,den,r,t)
lsim(A,B,C,D,u,t)
y = lsim(num,den,r,t)
y = lsim(A,B,C,D,u,t)
```

will generate the response to input time function `r` or `u`. See the following two examples. (Also, see Problems **A-5-14** through **A-5-16**.)

EXAMPLE 5-7 Using the `lsim` command, obtain the unit-ramp response of the following system:

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

We may enter MATLAB Program 5-12 into the computer to obtain the unit-ramp response. The resulting plot is shown in Figure 5-28.

MATLAB Program 5-12

```
% ----- Ramp Response -----
num = [0 0 1];
den = [1 1 1];
t = 0:0.1:8;
r = t;
y = lsim(num,den,r,t);
plot(t,r,'-',t,y,'o')
grid
title('Unit-Ramp Response Obtained by Use of Command "lsim"')
xlabel('t Sec')
ylabel('Unit-Ramp Input and System Output')
text(2.1,4.65,'Unit-Ramp Input')
text(4.5,2.0,'Output')
```

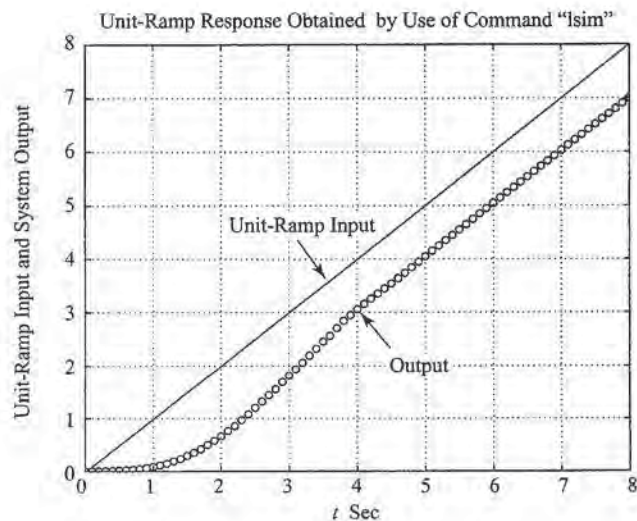


Figure 5-28
Unit-ramp response.

EXAMPLE 5-8 Consider the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0.5 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Using MATLAB, obtain the response curves $y(t)$ when the input u is given by

1. $u = \text{unit-step input}$
2. $u = e^{-t}$

Assume that the initial state is $\mathbf{x}(0) = \mathbf{0}$.

A possible MATLAB program to produce the responses of this system to the unit-step input $[u = 1(t)]$ and the exponential input $[u = e^{-t}]$ is shown in MATLAB Program 5-13. The resulting response curves are shown in Figures 5-29(a) and (b), respectively.

MATLAB Program 5-13

```
t = 0:0.1:12;
A = [-1 0.5;-1 0];
B = [0;1];
C = [1 0];
D = [0];

% For the unit-step input u = 1(t), use the command "y = step(A,B,C,D,1,t)".
y = step(A,B,C,D,1,t);
plot(t,y)
grid
title('Unit-Step Response')
xlabel('t Sec')
ylabel('Output')

% For the response to exponential input u = exp(-t), use the command
% z = lsim(A,B,C,D,u,t).
u = exp(-t);
z = lsim(A,B,C,D,u,t);
plot(t,u,'-o',t,z,'o')
grid
title('Response to Exponential Input u = exp(-t)')
xlabel('t Sec')
ylabel('Exponential Input and System Output')
text(2.3,0.49,'Exponential input')
text(6.4,0.28,'Output')
```

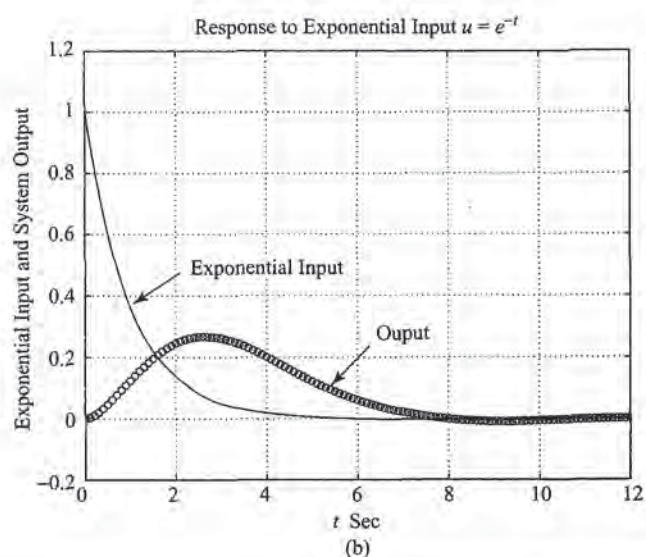
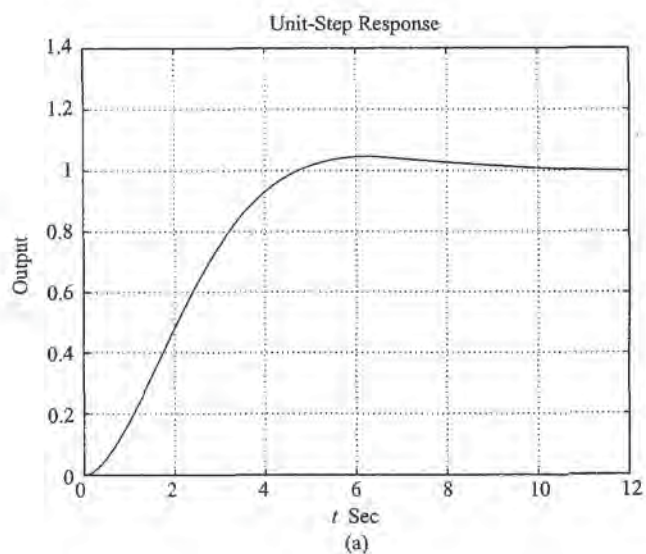


Figure 5-29
 (a) Unit-step
 response;
 (b) response to input
 $u = e^{-t}$.

Response to Initial Condition. In what follows we shall present a few methods for obtaining the response to an initial condition. Commands that we may use are “step” or “initial”. We shall first present a method to obtain the response to initial condition using a simple example. Then we shall discuss the response to initial condition when the system is given in state-space form. Finally, we shall present a command initial to obtain the response of a system given in a state-space form.

EXAMPLE 5-9

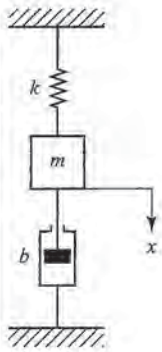


Figure 5-30
Mechanical system.

Consider the mechanical system shown in Figure 5-30, where $m = 1$ kg, $b = 3$ N-sec/m, and $k = 2$ N/m. Assume that at $t = 0$ the mass m is pulled downward such that $x(0) = 0.1$ m and $\dot{x}(0) = 0.05$ m/sec. The displacement $x(t)$ is measured from the equilibrium position before the mass is pulled down. Obtain the motion of the mass subjected to the initial condition. (Assume no external forcing function.)

The system equation is

$$m\ddot{x} + b\dot{x} + kx = 0$$

with the initial conditions $x(0) = 0.1$ m and $\dot{x}(0) = 0.05$ m/sec. (x is measured from the equilibrium position.) The Laplace transform of the system equation gives

$$m[s^2X(s) - sx(0) - \dot{x}(0)] + b[sX(s) - x(0)] + kX(s) = 0$$

or

$$(ms^2 + bs + k)X(s) = mx(0)s + m\dot{x}(0) + bx(0)$$

Solving this last equation for $X(s)$ and substituting the given numerical values, we obtain

$$\begin{aligned} X(s) &= \frac{mx(0)s + m\dot{x}(0) + bx(0)}{ms^2 + bs + k} \\ &= \frac{0.1s + 0.35}{s^2 + 3s + 2} \end{aligned}$$

This equation can be written as

$$X(s) = \frac{0.1s^2 + 0.35s}{s^2 + 3s + 2} \frac{1}{s}$$

Hence the motion of the mass m may be obtained as the unit-step response of the following system:

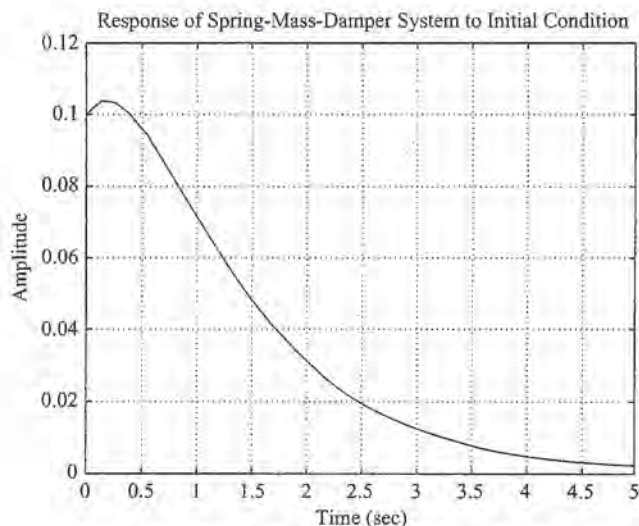
$$G(s) = \frac{0.1s^2 + 0.35s}{s^2 + 3s + 2}$$

MATLAB Program 5-14 will give a plot of the motion of the mass. The plot is shown in Figure 5-31.

MATLAB Program 5-14

```
% ----- Response to initial condition -----
% ***** System response to initial condition is converted to
% a unit-step response by modifying the numerator polynomial *****
% ***** Enter the numerator and denominator of the transfer
% function G(s) *****
num = [0.1 0.35 0];
den = [1 3 2];
% ***** Enter the following step-response command *****
step(num,den)
% ***** Enter grid and title of the plot *****
grid
title('Response of Spring-Mass-Damper System to Initial Condition')
```

Figure 5-31
Response of the
mechanical system
considered in
Example 5-9.



Response to Initial Condition (State-Space Approach, Case 1). Consider the system defined by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (5-49)$$

Let us obtain the response $\mathbf{x}(t)$ when the initial condition $\mathbf{x}(0)$ is specified. (There is no external input function acting on this system.) Assume that $\mathbf{x}$ is an n -vector.

First, take Laplace transforms of both sides of Equation (5-49).

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s)$$

This equation can be rewritten as

$$s\mathbf{X}(s) = \mathbf{A}\mathbf{X}(s) + \mathbf{x}(0) \quad (5-50)$$

Taking the inverse Laplace transform of Equation (5-50), we obtain

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{x}(0)\delta(t) \quad (5-51)$$

(Notice that by taking the Laplace transform of a differential equation and then by taking the inverse Laplace transform of the Laplace-transformed equation we generate a differential equation that involves the initial condition.)

Now define

$$\dot{\mathbf{z}} = \mathbf{x} \quad (5-52)$$

Then Equation (5-51) can be written as

$$\dot{\mathbf{z}} = \mathbf{A}\dot{\mathbf{z}} + \mathbf{x}(0)\delta(t) \quad (5-53)$$

By integrating Equation (5-53) with respect to t , we obtain

$$\dot{\mathbf{z}} = \mathbf{A}\dot{\mathbf{z}} + \mathbf{x}(0)1(t) = \mathbf{A}\dot{\mathbf{z}} + \mathbf{B}u \quad (5-54)$$

where

$$\mathbf{B} = \mathbf{x}(0), \quad u = 1(t)$$

Referring to Equation (5-52), the state $\mathbf{x}(t)$ is given by $\dot{\mathbf{z}}(t)$. Thus,

$$\dot{\mathbf{x}} = \dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}u \quad (5-55)$$

The solution of Equations (5-54) and (5-55) gives the response to the initial condition.

Summarizing, the response of Equation (5-49) to the initial condition $\mathbf{x}(0)$ is obtained by solving the following state-space equations:

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}u$$

$$\mathbf{x} = \mathbf{A}\mathbf{z} + \mathbf{B}u$$

where

$$\mathbf{B} = \mathbf{x}(0), \quad u = 1(t)$$

MATLAB commands to obtain the response curves in one diagram are given next.

$$[x,z,t] = \text{step}(A,B,A,B);$$

$$x1 = [1 \ 0 \ 0 \ \dots \ 0] * x';$$

$$x2 = [0 \ 1 \ 0 \ \dots \ 0] * x';$$

.

.

.

$$xn = [0 \ 0 \ 0 \ \dots \ 1] * x';$$

$$\text{plot}(t,x1,t,x2, \dots ,t,xn)$$

Response to Initial Condition (State-Space Approach, Case 2). Consider the system defined by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (5-56)$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} \quad (5-57)$$

(Assume that $\mathbf{x}$ is an n -vector and $\mathbf{y}$ is an m -vector.)

Similar to case 1, by defining

$$\dot{\mathbf{z}} = \mathbf{x}$$

we can obtain the following equation:

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{x}(0)1(t) = \mathbf{A}\mathbf{z} + \mathbf{B}u \quad (5-58)$$

where

$$\mathbf{B} = \mathbf{x}(0), \quad u = 1(t)$$

Noting that $\mathbf{x} = \dot{\mathbf{z}}$, Equation (5-57) can be written as

$$\mathbf{y} = \mathbf{C}\dot{\mathbf{z}} \quad (5-59)$$

By substituting Equation (5-58) into Equation (5-59), we obtain

$$\mathbf{y} = \mathbf{C}(\mathbf{A}\mathbf{z} + \mathbf{B}u) = \mathbf{C}\mathbf{A}\mathbf{z} + \mathbf{C}\mathbf{B}u \quad (5-60)$$

The solution of Equations (5-58) and (5-60) gives the response of the system to a given initial condition. MATLAB commands to obtain the response curves (output curves y1 versus t, y2 versus t, ... , ym versus t) are shown next.

```
[y,z,t] = step(A,B,C*A,C*B);
y1 = [1 0 0 ... 0]*y';
y2 = [0 1 0 ... 0]*y';
.
.
.
ym = [0 0 0 ... 1]*y';
plot(t,y1,t,y2, ... ,t,ym)
```

EXAMPLE 5-10 Obtain the response of the system subjected to the given initial condition.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -10 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

or

$$\dot{\mathbf{x}} = \mathbf{Ax}, \quad \mathbf{x}(0) = \mathbf{x}_0$$

Obtaining the response of the system to the given initial condition becomes that of solving the unit-step response of the following system:

$$\begin{aligned} \dot{\mathbf{z}} &= \mathbf{Az} + \mathbf{Bu} \\ \mathbf{x} &= \mathbf{Az} + \mathbf{Bu} \end{aligned}$$

where

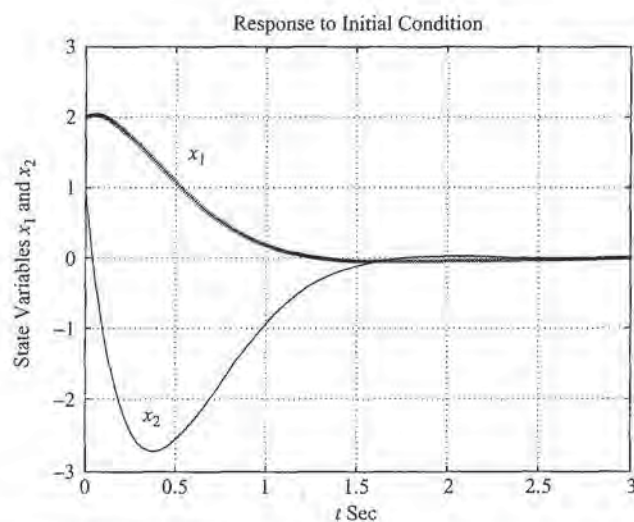
$$\mathbf{B} = \mathbf{x}(0), \quad u = 1(t)$$

Hence a possible MATLAB program for obtaining the response may be given as shown in MATLAB Program 5-15. The resulting response curves are shown in Figure 5-32.

MATLAB Program 5-15

```
t = 0:0.01:3;
A = [0 1;-10 -5];
B = [2;1];
[x,z,t] = step(A,B,A,B,1,t);
x1 = [1 0]*x';
x2 = [0 1]*x';
plot(t,x1,'x',t,x2,'-')
grid
title('Response to Initial Condition')
xlabel('t Sec')
ylabel('State Variables x1 and x2')
gtext('x1')
gtext('x2')
```

Figure 5-32
Response of system
in Example 5-10 to
initial condition.



Obtaining Response to Initial Condition by Use of Command `initial`. If the system is given in the state-space form, then the following command

`initial(A,B,C,D,[initial condition],t)`

will produce the response to the initial condition.

Suppose that we have the system defined by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \quad \mathbf{x}(0) = \mathbf{x}_0$$

$$y = \mathbf{C}\mathbf{x} + Du$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -10 & -5 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{C} = [0 \ 0], \quad D = 0$$

$$\mathbf{x}_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Then the command “initial” can be used as shown in MATLAB Program 5-16 to obtain the response to the initial condition. The response curves $x_1(t)$ and $x_2(t)$ are shown in Figure 5-33. They are the same as those shown in Figure 5-32.

MATLAB Program 5-16

```
t = 0:0.05:3;
A = [0 1;-10 -5];
B = [0;0];
C = [0 0];
D = [0];
[y,x] = initial(A,B,C,D,[2;1],t);
x1 = [1 0]*x';
x2 = [0 1]*x';
plot(t,x1,'o',t,x1,t,x2,'x',t,x2)
grid
title('Response to Initial Condition')
xlabel('t Sec')
ylabel('State Variables x1 and x2')
gtext('x1')
gtext('x2')
```

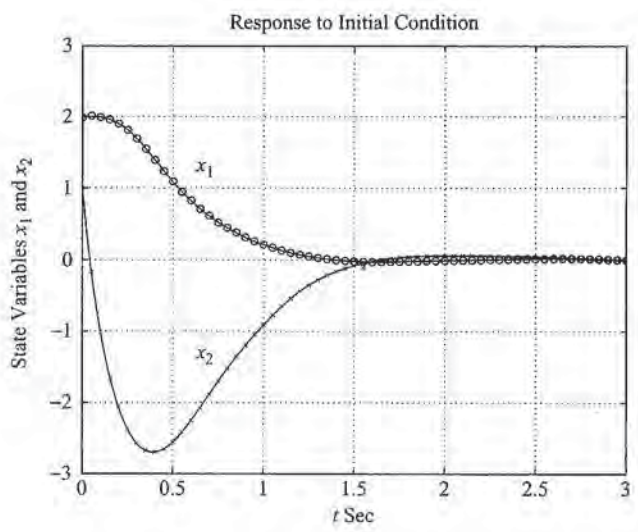


Figure 5-33
Response curves to
initial condition.

EXAMPLE 5-11 Consider the following system that is subjected to the initial condition. (No external forcing function is present.)

$$\ddot{y} + 8\dot{y} + 17y = 0$$
$$y(0) = 2, \quad \dot{y}(0) = 1, \quad \ddot{y}(0) = 0.5$$

Obtain the response $y(t)$ to the given initial condition.

By defining the state variables as

$$\begin{aligned}x_1 &= y \\x_2 &= \dot{y} \\x_3 &= \ddot{y}\end{aligned}$$

we obtain the following state-space representation for the system:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -10 & -17 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \begin{bmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0.5 \end{bmatrix}$$
$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

A possible MATLAB program to obtain the response $y(t)$ is given in MATLAB Program 5-17. The resulting response curve is shown in Figure 5-34.

MATLAB Program 5-17

```
t = 0:0.05:10;
A = [0 1 0; 0 0 1; -10 -17 -8];
B = [0; 0; 0];
C = [1 0 0];
D = [0];
y = initial(A,B,C,D,[2;1;0.5],t);
plot(t,y)
grid
title('Response to Initial Condition')
xlabel('t (sec)')
ylabel('Output y')
```

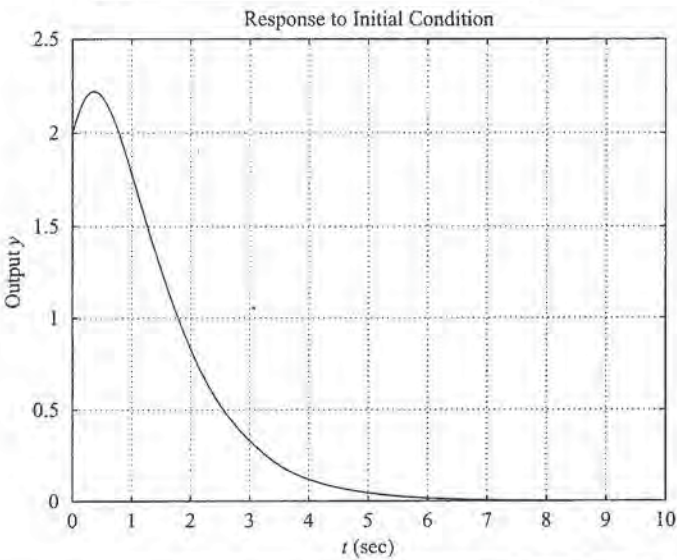


Figure 5-34
Response $y(t)$ to
initial condition.

5-6 AN EXAMPLE PROBLEM SOLVED WITH MATLAB

The purpose of this section is to present a MATLAB solution to the response of a mechanical vibratory system.

Mechanical Vibratory System. Consider the mechanical vibratory system shown in Figure 5-35(a). A wheel has a spring-mass-damper system hanging from it. The wheel is in a track that contains a flat (horizontal) portion, a slanted (downward at 45°) portion, and another flat (horizontal) portion. We start the motion of the system by nudging the wheel over the edge of the ramp. As the wheel drops down the ramp for a total of 0.707 m (vertically measured), the mass m hanging from the spring and damper drops with it, and the mass gains momentum that dissipates gradually. In this problem the wheel is assumed to slide on the slanted portion of the track without friction. On the second flat portion of the track, the wheel slides and rolls. The wheel continues to move on the flat portion of the track until it is stopped by an external means.

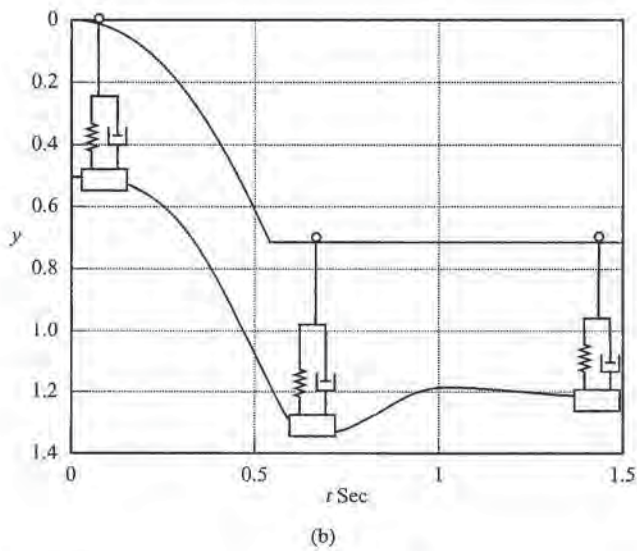
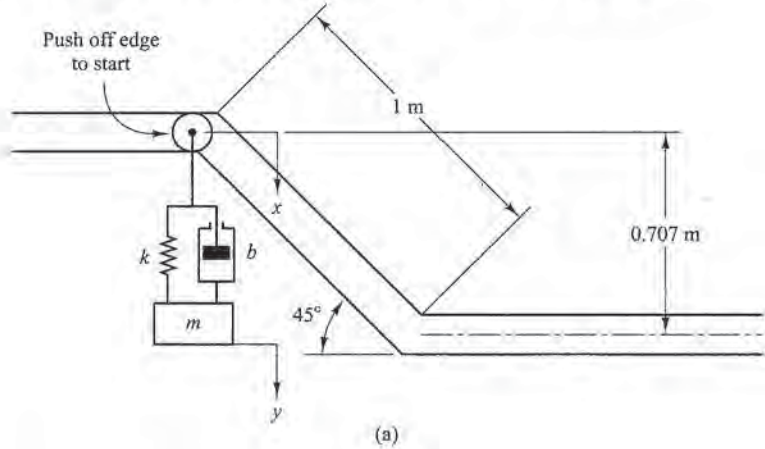


Figure 5-35
(a) Wheel with hanging mass-damper system;
(b) dynamic response of the system.

Assume the following numerical values for m , b , and k :

$$m = 4 \text{ kg}, \quad b = 40 \text{ N-sec/m}, \quad k = 400 \text{ N/m}$$

Assume also that the mass m_p of the wheel is negligible compared with the mass m . Obtain $x(t)$, the vertical motion of the wheel. Then obtain $Y(s)$, the Laplace transform of $y(t)$, which represents the up and down motion of mass m . The coordinate y is attached to the spring-mass-damper system as shown in Figure 5-35(a) and is measured from the equilibrium position of the system. The initial conditions are that $y(0) = 0$ and $\dot{y}(0) = 0$. Note that in this problem we are interested only in the vertical motions of the spring-mass-damper system. Note also that the system is frictionless with the exception of the damper, which relies on viscosity for its operation.

As the spring-mass-damper component travels down the ramp, it will undergo an acceleration produced as a result of the gravity force. When the spring-mass-damper reaches the level region at the bottom of the ramp, a shock will immediately be imposed on the spring-mass-damper component. It will, however, eventually come to a state of equilibrium following the impact due to the settling effects of the damper and spring. The dynamic response of this system is shown in Figure 5-35(b).

Determination of $x(t)$. The system starts with zero initial velocity and follows the track. The input to the system is the vertical position x along the track, and the output is the vertical position y of the mass. Since we assume no sliding friction, referring to Figure 5-36(a) we have in the z direction the following equation:

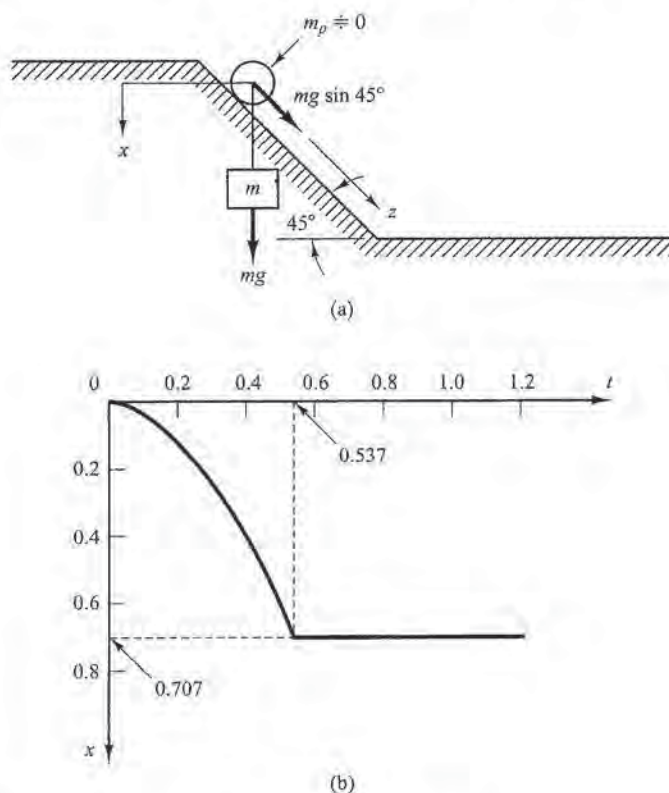


Figure 5-36
(a) Wheel with mass m slides on inclined plane; (b) curve $x(t)$ versus t .

$$m\ddot{z} = mg \sin 45^\circ$$

or

$$\ddot{z} = 9.81 \times 0.707 = 6.9357$$

Let us define the time it takes for the wheel to move from $z = 0$ to $z = 1$ m as t_1 . Then

$$z = 6.9357 \frac{t_1^2}{2} = 1$$

which yields

$$t_1 = 0.537 \text{ sec}$$

Thus $x(t)$ can be given as follows:

$$\begin{aligned} x(t) &= 0.707z = 0.707 \times 3.4678 t^2 = 2.452 t^2, & \text{for } 0 \leq t \leq 0.537 \\ &= 0.707, & \text{for } 0.537 < t \end{aligned}$$

It follows that from $t = 0.537$ sec to $t = \infty$ we have an input defined by a constant of 0.707. The position x at the end of the ramp is 0.707 and it takes 0.537 sec to get there. A curve $x(t)$ versus t is shown in Figure 5-36(b). Note that the positive direction of $x(t)$ is vertically downward.

To get a better picture of the events taking place in the system, we need to look at the input, shown in Figure 5-36(b). The effects of gravity do not allow us to model the behavior of the system with an ordinary ramp, but rather a parabolic function, which is followed by the input.

Determination of Transfer Function $Y(s)/X(s)$. Next, we shall obtain the equation of motion for the system and then the transfer function $Y(s)/X(s)$. Since y is measured from the equilibrium position, the system equation becomes

$$m\ddot{y} + b(\dot{y} - \dot{x}) + k(y - x) = 0$$

or

$$m\ddot{y} + b\dot{y} + ky = b\dot{x} + kx$$

where x is the input to the system and y is the output. By substituting the given numerical values for m , b , and k , we obtain

$$4\ddot{y} + 40\dot{y} + 400y = 40\dot{x} + 400x$$

or

$$\ddot{y} + 10\dot{y} + 100y = 10\dot{x} + 100x$$

The transfer function for the system can now be given by

$$\frac{Y(s)}{X(s)} = \frac{10s + 100}{s^2 + 10s + 100}$$

where the input $x(t)$ is given by $2.452t^2$ (for $0 \leq t \leq 0.537$) and 0.707 (for $0.537 < t$). Our problem here is to plot the response curve with MATLAB.

MATLAB Solution. Since the input consists of two different time functions (one is an acceleration input and the other is a step input,) we shall use “lsim” command in this problem. Since

$$\begin{aligned} x(t) &= 2.452t^2, & \text{for } 0 \leq t \leq 0.537 \\ x(t) &= 0.707, & \text{for } 0.537 < t \end{aligned}$$

we choose the computing time interval to be 0.001 sec. We define x as an array of points in MATLAB. This array initially follows $x(t) = 2.452t^2$ and, after $t = 0.537$ sec, follows $x(t) = 0.707$. We assume that the time region is $0 \leq t \leq 1.5$.

The acceleration input in the first part can be written as

$$\begin{aligned} t1 &= 0:0.001:0.537 \\ x1 &= 2.452*(t1.^2) \end{aligned}$$

where $t1$ represents a time count and $x1$ is the first part of the complete input function. (There are 538 calculation points from the initial position until the input reaches 0.707 m.) For the second part of the input, we need a step function with magnitude 0.707. After time 0.537 sec,

$$\begin{aligned} t2 &= 0.538:0.001:1.5 \\ x2 &= 0.707*\text{ones}(\text{size}(t2)) \end{aligned}$$

(There are 963 points from 0.538 sec through 1.5 sec, inclusive.) The next step is to transform both inputs to one complete input:

$$x = [x1 \ x2]$$

(The two input equations are transformed into a single vector, in order to appear as a single entry in the lsim command argument.)

Now we can use the lsim command

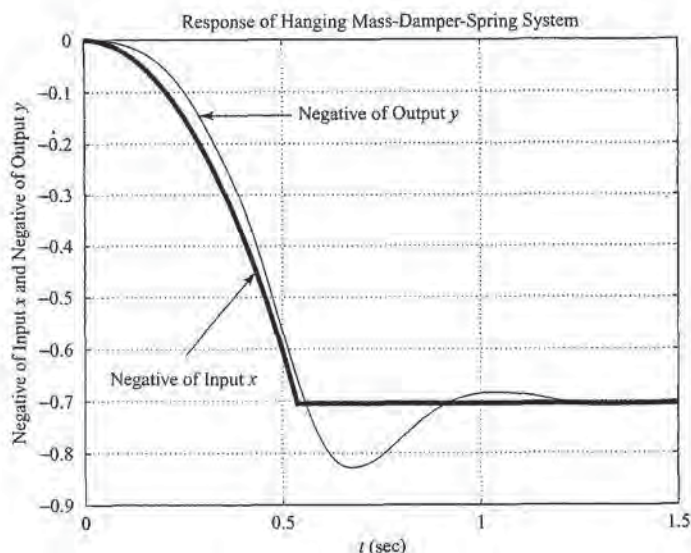
$$y = \text{lsim}(\text{num}, \text{den}, x, t)$$

and plot the response $y(t)$ as well as the input $x(t)$. A possible MATLAB program is shown in MATLAB Program 5–18. Note that the plotted input and output functions are negated. The resulting curves $x(t)$ and $y(t)$ are shown in Figure 5–37.

MATLAB Program 5–18

```
num = [0 10 100];
den = [1 10 100];
t1 = 0:0.001:0.537;
x1 = 2.452*(t1.^2);
t2 = 0.538:0.001:1.5;
x2 = 0.707*ones(size(t2));
t = [t1 t2];
x = [x1 x2];
y = lsim(num,den,x,t);
plot(t,-x,'x',t,-y)
grid
title('Response of Hanging Mass-Damper-Spring System')
xlabel('t (sec)')
ylabel('Negative of Input x and Negative of Output y')
text(0.07,-0.66,'Negative of Input x')
text(0.55,-0.14,'Negative of Output y')
```

Figure 5-37
Curves $-x(t)$ versus t
and $-y(t)$ versus t .



5-7 ROUTH'S STABILITY CRITERION

The most important problem in linear control systems concerns stability. That is, under what conditions will a system become unstable? If it is unstable, how should we stabilize the system? In Section 5-4 it was stated that a control system is stable if and only if all closed-loop poles lie in the left-half s plane. Most linear closed-loop systems have closed-loop transfer functions of the form

$$\frac{C(s)}{R(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1} s + b_m}{a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n} = \frac{B(s)}{A(s)}$$

where the a 's and b 's are constants and $m \leq n$. A simple criterion, known as Routh's stability criterion, enables us to determine the number of closed-loop poles that lie in the right-half s plane without having to factor the denominator polynomial. (The polynomial may include parameters that MATLAB cannot handle.)

Routh's Stability Criterion. Routh's stability criterion tells us whether or not there are unstable roots in a polynomial equation without actually solving for them. This stability criterion applies to polynomials with only a finite number of terms. When the criterion is applied to a control system, information about absolute stability can be obtained directly from the coefficients of the characteristic equation.

The procedure in Routh's stability criterion is as follows:

1. Write the polynomial in s in the following form:

$$a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n = 0 \quad (5-61)$$

where the coefficients are real quantities. We assume that $a_n \neq 0$; that is, any zero root has been removed.

2. If any of the coefficients are zero or negative in the presence of at least one positive coefficient, there is a root or roots that are imaginary or that have positive

real parts. Therefore, in such a case, the system is not stable. If we are interested in only the absolute stability, there is no need to follow the procedure further. Note that all the coefficients must be positive. This is a necessary condition, as may be seen from the following argument: A polynomial in s having real coefficients can always be factored into linear and quadratic factors, such as $(s + a)$ and $(s^2 + bs + c)$, where a , b , and c are real. The linear factors yield real roots and the quadratic factors yield complex-conjugate roots of the polynomial. The factor $(s^2 + bs + c)$ yields roots having negative real parts only if b and c are both positive. For all roots to have negative real parts, the constants a , b , c , and so on, in all factors must be positive. The product of any number of linear and quadratic factors containing only positive coefficients always yields a polynomial with positive coefficients. It is important to note that the condition that all the coefficients be positive is not sufficient to assure stability. The necessary but not sufficient condition for stability is that the coefficients of Equation (5-61) all be present and all have a positive sign. (If all a 's are negative, they can be made positive by multiplying both sides of the equation by -1 .)

3. If all coefficients are positive, arrange the coefficients of the polynomial in rows and columns according to the following pattern:

s^n	a_0	a_2	a_4	a_6	...
s^{n-1}	a_1	a_3	a_5	a_7	...
s^{n-2}	b_1	b_2	b_3	b_4	...
s^{n-3}	c_1	c_2	c_3	c_4	...
s^{n-4}	d_1	d_2	d_3	d_4	...
.	.	.			
.	.	.			
.	.	.			
s^2	e_1	e_2			
s^1	f_1				
s^0	g_1				

The process of forming rows continues until we run out of elements. (The total number of rows is $n + 1$.) The coefficients b_1 , b_2 , b_3 , and so on, are evaluated as follows:

$$\begin{aligned}
 b_1 &= \frac{a_1 a_2 - a_0 a_3}{a_1} \\
 b_2 &= \frac{a_1 a_4 - a_0 a_5}{a_1} \\
 b_3 &= \frac{a_1 a_6 - a_0 a_7}{a_1} \\
 &\vdots \\
 &\vdots \\
 &\vdots
 \end{aligned}$$

The evaluation of the b 's is continued until the remaining ones are all zero. The same pattern of cross-multiplying the coefficients of the two previous rows is followed in evaluating the c 's, d 's, e 's, and so on. That is,

$$\begin{aligned} c_1 &= \frac{b_1 a_3 - a_1 b_2}{b_1} \\ c_2 &= \frac{b_1 a_5 - a_1 b_3}{b_1} \\ c_3 &= \frac{b_1 a_7 - a_1 b_4}{b_1} \\ &\vdots \end{aligned}$$

and

$$\begin{aligned} d_1 &= \frac{c_1 b_2 - b_1 c_2}{c_1} \\ d_2 &= \frac{c_1 b_3 - b_1 c_3}{c_1} \\ &\vdots \end{aligned}$$

This process is continued until the n th row has been completed. The complete array of coefficients is triangular. Note that in developing the array an entire row may be divided or multiplied by a positive number in order to simplify the subsequent numerical calculation without altering the stability conclusion.

Routh's stability criterion states that the number of roots of Equation (5-61) with positive real parts is equal to the number of changes in sign of the coefficients of the first column of the array. It should be noted that the exact values of the terms in the first column need not be known; instead, only the signs are needed. The necessary and sufficient condition that all roots of Equation (5-61) lie in the left-half s plane is that all the coefficients of Equation (5-61) be positive and all terms in the first column of the array have positive signs.

EXAMPLE 5-12 Let us apply Routh's stability criterion to the following third-order polynomial:

$$a_0 s^3 + a_1 s^2 + a_2 s + a_3 = 0$$

where all the coefficients are positive numbers. The array of coefficients becomes

$$\begin{array}{ccc} s^3 & a_0 & a_2 \\ s^2 & a_1 & a_3 \\ s^1 & \frac{a_1 a_2 - a_0 a_3}{a_1} & \\ s^0 & a_3 & \end{array}$$

The condition that all roots have negative real parts is given by

$$a_1 a_2 > a_0 a_3$$

EXAMPLE 5-13 Consider the following polynomial:

$$s^4 + 2s^3 + 3s^2 + 4s + 5 = 0$$

Let us follow the procedure just presented and construct the array of coefficients. (The first two rows can be obtained directly from the given polynomial. The remaining terms are obtained from these. If any coefficients are missing, they may be replaced by zeros in the array.)

$$\begin{array}{r|rrrr} s^4 & 1 & 3 & 5 & \\ s^3 & 2 & 4 & 0 & \\ s^2 & 1 & 5 & & \\ s^1 & -6 & & & \\ s^0 & 5 & & & \end{array} \quad \begin{array}{r|rrrr} s^4 & 1 & 3 & 5 & \\ s^3 & 2 & 4 & 0 & \\ s^2 & 1 & 5 & & \\ s^1 & -6 & & & \\ s^0 & 5 & & & \end{array} \quad \begin{array}{l} \text{The second row is divided} \\ \text{by 2.} \end{array}$$

In this example, the number of changes in sign of the coefficients in the first column is 2. This means that there are two roots with positive real parts. Note that the result is unchanged when the coefficients of any row are multiplied or divided by a positive number in order to simplify the computation.

Special Cases. If a first-column term in any row is zero, but the remaining terms are not zero or there is no remaining term, then the zero term is replaced by a very small positive number ϵ and the rest of the array is evaluated. For example, consider the following equation:

$$s^3 + 2s^2 + s + 2 = 0 \quad (5-62)$$

The array of coefficients is

$$\begin{array}{r|rr} s^3 & 1 & 1 \\ s^2 & 2 & 2 \\ s^1 & 0 \approx \epsilon & \\ s^0 & 2 & \end{array}$$

If the sign of the coefficient above the zero (ϵ) is the same as that below it, it indicates that there are a pair of imaginary roots. Actually, Equation (5-62) has two roots at $s = \pm j$.

If, however, the sign of the coefficient above the zero (ϵ) is opposite that below it, it indicates that there is one sign change. For example, for the equation

$$s^3 - 3s + 2 = (s - 1)^2(s + 2) = 0$$

the array of coefficients is

$$\begin{array}{r|rr} s^3 & 1 & -3 \\ s^2 & 0 \approx \epsilon & 2 \\ s^1 & -3 & -\frac{2}{\epsilon} \\ s^0 & 2 & \end{array}$$

One sign change: $\rightarrow$ (from s^3 to s^2)

One sign change: $\rightarrow$ (from s^2 to s^1)

There are two sign changes of the coefficients in the first column. This agrees with the correct result indicated by the factored form of the polynomial equation.

If all the coefficients in any derived row are zero, it indicates that there are roots of equal magnitude lying radially opposite in the s plane, that is, two real roots with equal magnitudes and opposite signs and/or two conjugate imaginary roots. In such a case, the evaluation of the rest of the array can be continued by forming an auxiliary polynomial with the coefficients of the last row and by using the coefficients of the derivative of this polynomial in the next row. Such roots with equal magnitudes and lying radially opposite in the s plane can be found by solving the auxiliary polynomial, which is always even. For a $2n$ -degree auxiliary polynomial, there are n pairs of equal and opposite roots. For example, consider the following equation:

$$s^5 + 2s^4 + 24s^3 + 48s^2 - 25s - 50 = 0$$

The array of coefficients is

$$\begin{array}{rcll} s^5 & 1 & 24 & -25 \\ s^4 & 2 & 48 & -50 \\ s^3 & 0 & 0 & \end{array} \quad \leftarrow \text{Auxiliary polynomial } P(s)$$

The terms in the s^3 row are all zero. (Note that such a case occurs only in an odd-numbered row.) The auxiliary polynomial is then formed from the coefficients of the s^4 row. The auxiliary polynomial $P(s)$ is

$$P(s) = 2s^4 + 48s^2 - 50$$

which indicates that there are two pairs of roots of equal magnitude and opposite sign (that is, two real roots with the same magnitude but opposite signs or two complex-conjugate roots on the imaginary axis). These pairs are obtained by solving the auxiliary polynomial equation $P(s) = 0$. The derivative of $P(s)$ with respect to s is

$$\frac{dP(s)}{ds} = 8s^3 + 96s$$

The terms in the s^3 row are replaced by the coefficients of the last equation, that is, 8 and 96. The array of coefficients then becomes

$$\begin{array}{rcll} s^5 & 1 & 24 & -25 \\ s^4 & 2 & 48 & -50 \\ s^3 & 8 & 96 & \\ s^2 & 24 & -50 & \\ s^1 & 112.7 & 0 & \\ s^0 & -50 & & \end{array} \quad \leftarrow \text{Coefficients of } dP(s)/ds$$

We see that there is one change in sign in the first column of the new array. Thus, the original equation has one root with a positive real part. By solving for roots of the auxiliary polynomial equation,

$$2s^4 + 48s^2 - 50 = 0$$

we obtain

$$s^2 = 1, \quad s^2 = -25$$

or

$$s = \pm 1, \quad s = \pm j5$$

These two pairs of roots of $P(s)$ are a part of the roots of the original equation. As a matter of fact, the original equation can be written in factored form as follows:

$$(s + 1)(s - 1)(s + j5)(s - j5)(s + 2) = 0$$

Clearly, the original equation has one root with a positive real part.

Relative Stability Analysis. Routh's stability criterion provides the answer to the question of absolute stability. This, in many practical cases, is not sufficient. We usually require information about the relative stability of the system. A useful approach for examining relative stability is to shift the s -plane axis and apply Routh's stability criterion. That is, we substitute

$$s = \hat{s} - \sigma \quad (\sigma = \text{constant})$$

into the characteristic equation of the system, write the polynomial in terms of $\hat{s}$; and apply Routh's stability criterion to the new polynomial in $\hat{s}$. The number of changes of sign in the first column of the array developed for the polynomial in $\hat{s}$ is equal to the number of roots that are located to the right of the vertical line $s = -\sigma$. Thus, this test reveals the number of roots that lie to the right of the vertical line $s = -\sigma$.

Application of Routh's Stability Criterion to Control System Analysis. Routh's stability criterion is of limited usefulness in linear control system analysis mainly because it does not suggest how to improve relative stability or how to stabilize an unstable system. It is possible, however, to determine the effects of changing one or two parameters of a system by examining the values that cause instability. In the following, we shall consider the problem of determining the stability range of a parameter value.

Consider the system shown in Figure 5-38. Let us determine the range of K for stability. The closed-loop transfer function is

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

The characteristic equation is

$$s^4 + 3s^3 + 3s^2 + 2s + K = 0$$

The array of coefficients becomes

s^4	1	3	K
s^3	3	2	0
s^2	$\frac{7}{3}$	K	
s^1	$2 - \frac{9}{7}K$		
s^0	K		

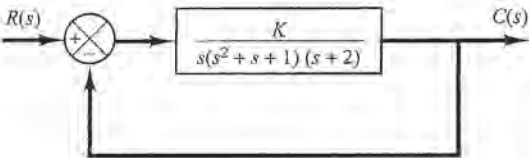


Figure 5-38
Control system.

For stability, K must be positive, and all coefficients in the first column must be positive. Therefore,

$$\frac{14}{9} > K > 0$$

When $K = \frac{14}{9}$, the system becomes oscillatory and, mathematically, the oscillation is sustained at constant amplitude.

Note that the ranges of design parameters that lead to stability may be determined by use of Routh's stability criterion.

5-8 EFFECTS OF INTEGRAL AND DERIVATIVE CONTROL ACTIONS ON SYSTEM PERFORMANCE

In this section, we shall investigate the effects of integral and derivative control actions on the system performance. Here we shall consider only simple systems so that the effects of integral and derivative control actions on system performance can be clearly seen.

Integral Control Action. In the proportional control of a plant whose transfer function does not possess an integrator $1/s$, there is a steady-state error, or offset, in the response to a step input. Such an offset can be eliminated if the integral control action is included in the controller.

In the integral control of a plant, the control signal, the output signal from the controller, at any instant is the area under the actuating error signal curve up to that instant. The control signal $u(t)$ can have a nonzero value when the actuating error signal $e(t)$ is zero, as shown in Figure 5-39(a). This is impossible in the case of the proportional controller since a nonzero control signal requires a nonzero actuating error signal. (A nonzero actuating error signal at steady state means that there is an offset.) Figure 5-39(b) shows the curve $e(t)$ versus t and the corresponding curve $u(t)$ versus t when the controller is of the proportional type.

Note that integral control action, while removing offset or steady-state error, may lead to oscillatory response of slowly decreasing amplitude or even increasing amplitude, both of which are usually undesirable.

Figure 5-39
(a) Plots of $e(t)$ and $u(t)$ curves showing nonzero control signal when the actuating error signal is zero (integral control); (b) plots of $e(t)$ and $u(t)$ curves showing zero control signal when the actuating error signal is zero (proportional control).

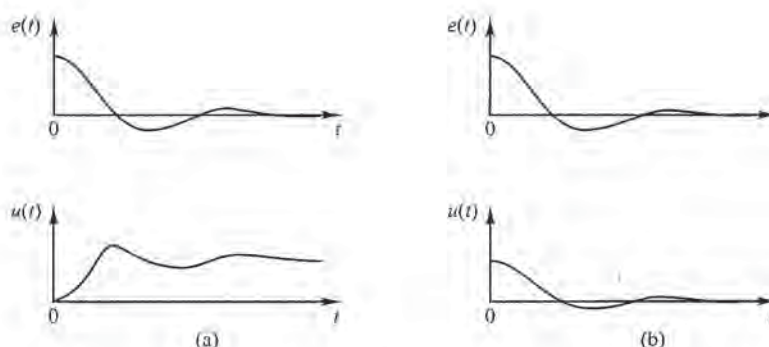
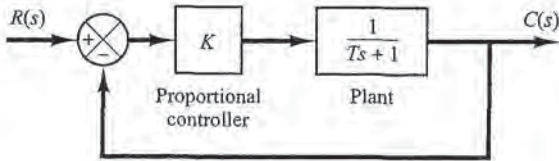


Figure 5-40
Proportional control system.



Proportional Control of Systems. We shall show that the proportional control of a system without an integrator will result in a steady-state error with a step input. We shall then show that such an error can be eliminated if integral control action is included in the controller.

Consider the system shown in Figure 5-40. Let us obtain the steady-state error in the unit-step response of the system. Define

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

Since

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

the error $E(s)$ is given by

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

For the unit-step input $R(s) = 1/s$, we have

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

The steady-state error is

$$e_{ss} = \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} sE(s) = \lim_{s \rightarrow 0} \frac{Ts + 1}{Ts + 1 + K} = \frac{1}{K + 1}$$

Such a system without an integrator in the feedforward path always has a steady-state error in the step response. Such a steady-state error is called an offset. Figure 5-41 shows the unit-step response and the offset.

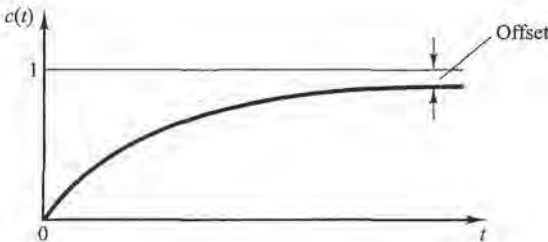
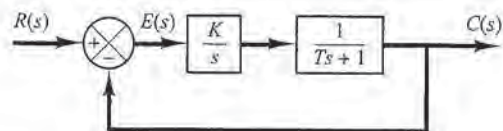


Figure 5-41
Unit-step response
and offset.

Figure 5-42
Integral control
system.



Integral Control of Systems. Consider the system shown in Figure 5-42. The controller is an integral controller. The closed-loop transfer function of the system is

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

Hence

$$\frac{E(s)}{R(s)} = \frac{R(s) - C(s)}{R(s)} = \frac{s(Ts + 1)}{s(Ts + 1) + K}$$

Since the system is stable, the steady-state error for the unit-step response can be obtained by applying the final-value theorem, as follows:

$$\begin{aligned} e_{ss} &= \lim_{s \rightarrow 0} sE(s) \\ &= \lim_{s \rightarrow 0} \frac{s^2(Ts + 1)}{Ts^2 + s + K} \cdot \frac{1}{s} \\ &= 0 \end{aligned}$$

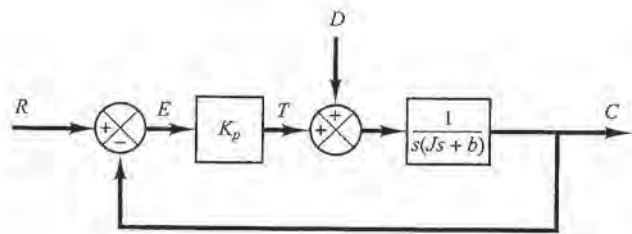
Integral control of the system thus eliminates the steady-state error in the response to the step input. This is an important improvement over the proportional control alone, which gives offset.

Response to Torque Disturbances (Proportional Control). Let us investigate the effect of a torque disturbance occurring at the load element. Consider the system shown in Figure 5-43. The proportional controller delivers torque T to position the load element, which consists of moment of inertia and viscous friction. Torque disturbance is denoted by D .

Assuming that the reference input is zero or $R(s) = 0$, the transfer function between $C(s)$ and $D(s)$ is given by

$$\frac{C(s)}{D(s)} = \frac{1}{Js^2 + bs + K_p}$$

Figure 5-43
Control system with
a torque disturbance.



Hence

$$\frac{E(s)}{D(s)} = -\frac{C(s)}{D(s)} = -\frac{1}{Js^2 + bs + K_p}$$

The steady-state error due to a step disturbance torque of magnitude T_d is given by

$$\begin{aligned} e_{ss} &= \lim_{s \rightarrow 0} sE(s) \\ &= \lim_{s \rightarrow 0} \frac{-s}{Js^2 + bs + K_p} \frac{T_d}{s} \\ &= -\frac{T_d}{K_p} \end{aligned}$$

At steady state, the proportional controller provides the torque $-T_d$, which is equal in magnitude but opposite in sign to the disturbance torque T_d . The steady-state output due to the step disturbance torque is

$$c_{ss} = -e_{ss} = \frac{T_d}{K_p}$$

The steady-state error can be reduced by increasing the value of the gain K_p . Increasing this value, however, will cause the system response to be more oscillatory.

Response to Torque Disturbances (Proportional-Plus-Integral Control). To eliminate offset due to torque disturbance, the proportional controller may be replaced by a proportional-plus-integral controller.

If integral control action is added to the controller, then, as long as there is an error signal, a torque is developed by the controller to reduce this error, provided the control system is a stable one.

Figure 5-44 shows the proportional-plus-integral control of the load element, consisting of moment of inertia and viscous friction.

The closed-loop transfer function between $C(s)$ and $D(s)$ is

$$\frac{C(s)}{D(s)} = \frac{s}{Js^3 + bs^2 + K_p s + \frac{K_p}{T_i}}$$

In the absence of the reference input, or $r(t) = 0$, the error signal is obtained from

$$E(s) = -\frac{s}{Js^3 + bs^2 + K_p s + \frac{K_p}{T_i}} D(s)$$

Figure 5-44
Proportional-plus-integral control of a load element consisting of moment of inertia and viscous friction.

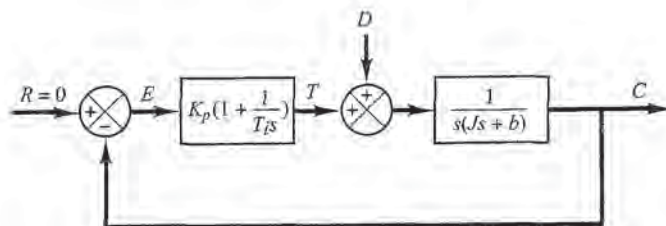
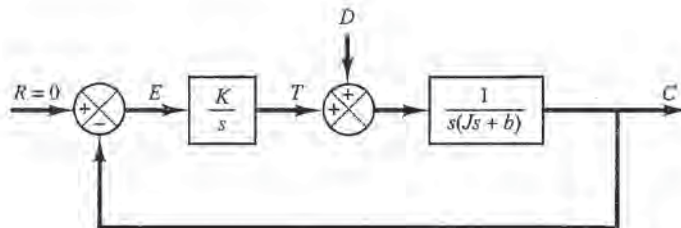


Figure 5-45
Integral control of a
load element
consisting of moment
of inertia and viscous
friction.



If this control system is stable, that is, if the roots of the characteristic equation

$$Js^3 + bs^2 + K_p s + \frac{K_p}{T_i} = 0$$

have negative real parts, then the steady-state error in the response to a unit-step disturbance torque can be obtained by applying the final-value theorem as follows:

$$\begin{aligned} e_{ss} &= \lim_{s \rightarrow 0} sE(s) \\ &= \lim_{s \rightarrow 0} \frac{-s^2}{Js^3 + bs^2 + K_p s + \frac{K_p}{T_i}} \frac{1}{s} \\ &= 0 \end{aligned}$$

Thus steady-state error to the step disturbance torque can be eliminated if the controller is of the proportional-plus-integral type.

Note that the integral control action added to the proportional controller has converted the originally second-order system to a third-order one. Hence the control system may become unstable for a large value of K_p since the roots of the characteristic equation may have positive real parts. (The second-order system is always stable if the coefficients in the system differential equation are all positive.)

It is important to point out that if the controller were an integral controller, as in Figure 5-45, then the system always becomes unstable because the characteristic equation

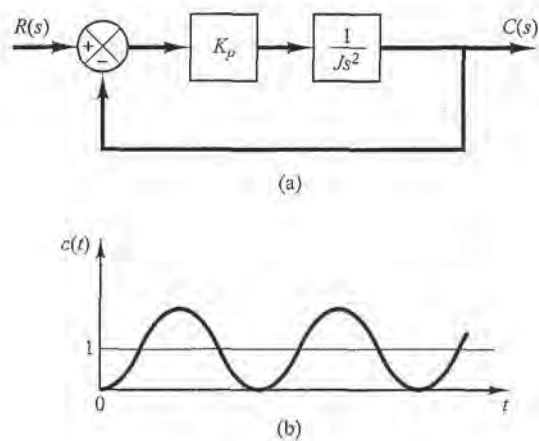
$$Js^3 + bs^2 + K = 0$$

will have roots with positive real parts. Such an unstable system cannot be used in practice.

Note that in the system of Figure 5-44 the proportional control action tends to stabilize the system, while the integral control action tends to eliminate or reduce steady-state error in response to various inputs.

Derivative Control Action. Derivative control action, when added to a proportional controller, provides a means of obtaining a controller with high sensitivity. An advantage of using derivative control action is that it responds to the rate of change of the actuating error and can produce a significant correction before the magnitude of the actuating error becomes too large. Derivative control thus anticipates the actuating error, initiates an early corrective action, and tends to increase the stability of the system.

Figure 5-46
 (a) Proportional control of a system with inertia load;
 (b) response to a unit-step input.



Although derivative control does not affect the steady-state error directly, it adds damping to the system and thus permits the use of a larger value of the gain K , which will result in an improvement in the steady-state accuracy.

Because derivative control operates on the rate of change of the actuating error and not the actuating error itself, this mode is never used alone. It is always used in combination with proportional or proportional-plus-integral control action.

Proportional Control of Systems with Inertia Load. Before we discuss the effect of derivative control action on system performance, we shall consider the proportional control of an inertia load.

Consider the system shown in Figure 5-46(a). The closed-loop transfer function is obtained as

$$\frac{C(s)}{R(s)} = \frac{K_p}{Js^2 + K_p}$$

Since the roots of the characteristic equation

$$Js^2 + K_p = 0$$

are imaginary, the response to a unit-step input continues to oscillate indefinitely, as shown in Figure 5-46(b).

Control systems exhibiting such response characteristics are not desirable. We shall see that the addition of derivative control will stabilize the system.

Proportional-Plus-Derivative Control of a System with Inertia Load. Let us modify the proportional controller to a proportional-plus-derivative controller whose transfer function is $K_p(1 + T_d s)$. The torque developed by the controller is proportional to $K_p(e + T_d \dot{e})$. Derivative control is essentially anticipatory, measures the instantaneous error velocity, and predicts the large overshoot ahead of time and produces an appropriate counteraction before too large an overshoot occurs.

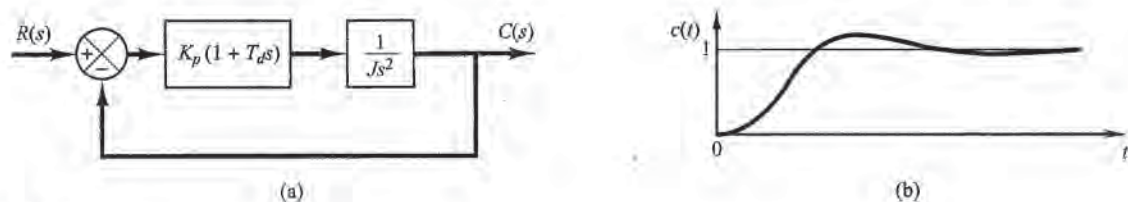


Figure 5-47
(a) Proportional-plus-derivative control of a system with inertia load; (b) response to a unit-step input.

Consider the system shown in Figure 5-47(a). The closed-loop transfer function is given by

$$\frac{C(s)}{R(s)} = \frac{K_p(1 + T_d s)}{J s^2 + K_p T_d s + K_p}$$

The characteristic equation

$$J s^2 + K_p T_d s + K_p = 0$$

now has two roots with negative real parts for positive values of J , K_p , and T_d . Thus derivative control introduces a damping effect. A typical response curve $c(t)$ to a unit-step input is shown in Figure 5-47(b). Clearly, the response curve shows a marked improvement over the original response curve shown in Figure 5-46(b).

Proportional-Plus-Derivative Control of Second-Order Systems. A compromise between acceptable transient-response behavior and acceptable steady-state behavior may be achieved by use of proportional-plus-derivative control action.

Consider the system shown in Figure 5-48. The closed-loop transfer function is

$$\frac{C(s)}{R(s)} = \frac{K_p + K_d s}{J s^2 + (B + K_d) s + K_p}$$

The steady-state error for a unit-ramp input is

$$e_{ss} = \frac{B}{K_p}$$

The characteristic equation is

$$J s^2 + (B + K_d) s + K_p = 0$$

The effective damping coefficient of this system is thus $B + K_d$ rather than B . Since the damping ratio ζ of this system is

$$\zeta = \frac{B + K_d}{2 \sqrt{K_p J}}$$

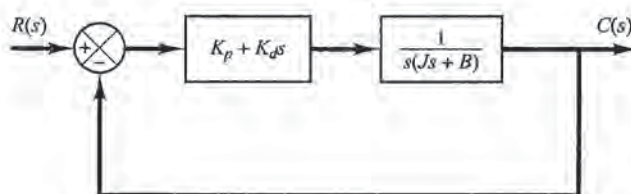


Figure 5-48
Control system.

it is possible to make both the steady-state error e_{ss} for a ramp input and the maximum overshoot for a step input small by making B small, K_p large, and K_d large enough so that ζ is between 0.4 and 0.7.

5-9 STEADY-STATE ERRORS IN UNITY-FEEDBACK CONTROL SYSTEMS

Errors in a control system can be attributed to many factors. Changes in the reference input will cause unavoidable errors during transient periods and may also cause steady-state errors. Imperfections in the system components, such as static friction, backlash, and amplifier drift, as well as aging or deterioration, will cause errors at steady state. In this section, however, we shall not discuss errors due to imperfections in the system components. Rather, we shall investigate a type of steady-state error that is caused by the incapability of a system to follow particular types of inputs.

Any physical control system inherently suffers steady-state error in response to certain types of inputs. A system may have no steady-state error to a step input, but the same system may exhibit nonzero steady-state error to a ramp input. (The only way we may be able to eliminate this error is to modify the system structure.) Whether a given system will exhibit steady-state error for a given type of input depends on the type of open-loop transfer function of the system, to be discussed in what follows.

Classification of Control Systems. Control systems may be classified according to their ability to follow step inputs, ramp inputs, parabolic inputs, and so on. This is a reasonable classification scheme because actual inputs may frequently be considered combinations of such inputs. The magnitudes of the steady-state errors due to these individual inputs are indicative of the goodness of the system.

Consider the unity-feedback control system with the following open-loop transfer function $G(s)$:

$$G(s) = \frac{K(T_a s + 1)(T_b s + 1) \cdots (T_m s + 1)}{s^N (T_1 s + 1)(T_2 s + 1) \cdots (T_p s + 1)}$$

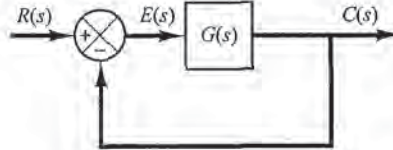
It involves the term s^N in the denominator, representing a pole of multiplicity N at the origin. The present classification scheme is based on the number of integrations indicated by the open-loop transfer function. A system is called type 0, type 1, type 2, ..., if $N = 0$, $N = 1$, $N = 2$, ..., respectively. Note that this classification is different from that of the order of a system. As the type number is increased, accuracy is improved; however, increasing the type number aggravates the stability problem. A compromise between steady-state accuracy and relative stability is always necessary.

We shall see later that, if $G(s)$ is written so that each term in the numerator and denominator, except the term s^N , approaches unity as s approaches zero, then the open-loop gain K is directly related to the steady-state error.

Steady-State Errors. Consider the system shown in Figure 5-49. The closed-loop transfer function is

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

Figure 5-49
Control system.



The transfer function between the error signal $e(t)$ and the input signal $r(t)$ is

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

where the error $e(t)$ is the difference between the input signal and the output signal.

The final-value theorem provides a convenient way to find the steady-state performance of a stable system. Since $E(s)$ is

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

the steady-state error is

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

The static error constants defined in the following are figures of merit of control systems. The higher the constants, the smaller the steady-state error. In a given system, the output may be the position, velocity, pressure, temperature, or the like. The physical form of the output, however, is immaterial to the present analysis. Therefore, in what follows, we shall call the output “position,” the rate of change of the output “velocity,” and so on. This means that in a temperature control system “position” represents the output temperature, “velocity” represents the rate of change of the output temperature, and so on.

Static Position Error Constant K_p . The steady-state error of the system for a unit-step input is

$$\begin{aligned} e_{ss} &= \lim_{s \rightarrow 0} \frac{s}{1 + G(s)} \frac{1}{s} \\ &= \frac{1}{1 + G(0)} \end{aligned}$$

The static position error constant K_p is defined by

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

Thus, the steady-state error in terms of the static position error constant K_p is given by

$$e_{ss} = \frac{1}{1 + K_p}$$

For a type 0 system,

$$K_p = \lim_{s \rightarrow 0} \frac{K(T_a s + 1)(T_b s + 1) \cdots}{(T_1 s + 1)(T_2 s + 1) \cdots} = K$$

For a type 1 or higher system,

$$K_p = \lim_{s \rightarrow 0} \frac{K(T_a s + 1)(T_b s + 1) \cdots}{s^N (T_1 s + 1)(T_2 s + 1) \cdots} = \infty, \quad \text{for } N \geq 1$$

Hence, for a type 0 system, the static position error constant K_p is finite, while for a type 1 or higher system, K_p is infinite.

For a unit-step input, the steady-state error e_{ss} may be summarized as follows:

$$\begin{aligned} e_{ss} &= \frac{1}{1 + K}, & \text{for type 0 systems} \\ e_{ss} &= 0, & \text{for type 1 or higher systems} \end{aligned}$$

From the foregoing analysis, it is seen that the response of a feedback control system to a step input involves a steady-state error if there is no integration in the feedforward path. (If small errors for step inputs can be tolerated, then a type 0 system may be permissible, provided that the gain K is sufficiently large. If the gain K is too large, however, it is difficult to obtain reasonable relative stability.) If zero steady-state error for a step input is desired, the type of the system must be one or higher.

Static Velocity Error Constant K_v . The steady-state error of the system with a unit-ramp input is given by

$$\begin{aligned} e_{ss} &= \lim_{s \rightarrow 0} \frac{s}{1 + G(s)} \frac{1}{s^2} \\ &= \lim_{s \rightarrow 0} \frac{1}{sG(s)} \end{aligned}$$

The static velocity error constant K_v is defined by

$$K_v = \lim_{s \rightarrow 0} sG(s)$$

Thus, the steady-state error in terms of the static velocity error constant K_v is given by

$$e_{ss} = \frac{1}{K_v}$$

The term *velocity error* is used here to express the steady-state error for a ramp input. The dimension of the velocity error is the same as the system error. That is, velocity error is not an error in velocity, but it is an error in position due to a ramp input.

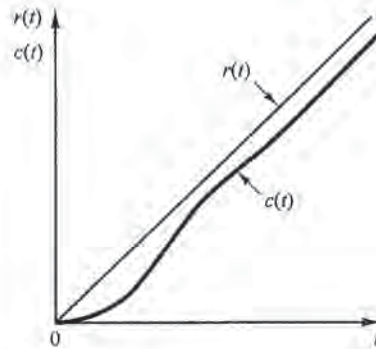
For a type 0 system,

$$K_v = \lim_{s \rightarrow 0} \frac{sK(T_a s + 1)(T_b s + 1) \cdots}{(T_1 s + 1)(T_2 s + 1) \cdots} = 0$$

For a type 1 system,

$$K_v = \lim_{s \rightarrow 0} \frac{sK(T_a s + 1)(T_b s + 1) \cdots}{s(T_1 s + 1)(T_2 s + 1) \cdots} = K$$

Figure 5-50
Response of a type 1
unity-feedback
system to a ramp
input.



For a type 2 or higher system,

$$K_v = \lim_{s \rightarrow 0} \frac{sK(T_a s + 1)(T_b s + 1) \cdots}{s^N(T_1 s + 1)(T_2 s + 1) \cdots} = \infty, \quad \text{for } N \geq 2$$

The steady-state error e_{ss} for the unit-ramp input can be summarized as follows:

$$e_{ss} = \frac{1}{K_v} = \infty, \quad \text{for type 0 systems}$$

$$e_{ss} = \frac{1}{K_v} = \frac{1}{K}, \quad \text{for type 1 systems}$$

$$e_{ss} = \frac{1}{K_v} = 0, \quad \text{for type 2 or higher systems}$$

The foregoing analysis indicates that a type 0 system is incapable of following a ramp input in the steady state. The type 1 system with unity feedback can follow the ramp input with a finite error. In steady-state operation, the output velocity is exactly the same as the input velocity, but there is a positional error. This error is proportional to the velocity of the input and is inversely proportional to the gain K . Figure 5-50 shows an example of the response of a type 1 system with unity feedback to a ramp input. The type 2 or higher system can follow a ramp input with zero error at steady state.

Static Acceleration Error Constant K_a . The steady-state error of the system with a unit-parabolic input (acceleration input), which is defined by

$$r(t) = \frac{t^2}{2}, \quad \text{for } t \geq 0$$

$$= 0, \quad \text{for } t < 0$$

is given by

$$e_{ss} = \lim_{s \rightarrow 0} \frac{s}{1 + G(s)} \frac{1}{s^3}$$

$$= \frac{1}{\lim_{s \rightarrow 0} s^2 G(s)}$$

The static acceleration error constant K_a is defined by the equation

$$K_a = \lim_{s \rightarrow 0} s^2 G(s)$$

The steady-state error is then

$$e_{ss} = \frac{1}{K_a}$$

Note that the acceleration error, the steady-state error due to a parabolic input, is an error in position.

The values of K_a are obtained as follows:

For a type 0 system,

$$K_a = \lim_{s \rightarrow 0} \frac{s^2 K (T_a s + 1)(T_b s + 1) \cdots}{(T_1 s + 1)(T_2 s + 1) \cdots} = 0$$

For a type 1 system,

$$K_a = \lim_{s \rightarrow 0} \frac{s^2 K (T_a s + 1)(T_b s + 1) \cdots}{s (T_1 s + 1)(T_2 s + 1) \cdots} = 0$$

For a type 2 system,

$$K_a = \lim_{s \rightarrow 0} \frac{s^2 K (T_a s + 1)(T_b s + 1) \cdots}{s^2 (T_1 s + 1)(T_2 s + 1) \cdots} = K$$

For a type 3 or higher system,

$$K_a = \lim_{s \rightarrow 0} \frac{s^2 K (T_a s + 1)(T_b s + 1) \cdots}{s^N (T_1 s + 1)(T_2 s + 1) \cdots} = \infty, \quad \text{for } N \geq 3$$

Thus, the steady-state error for the unit parabolic input is

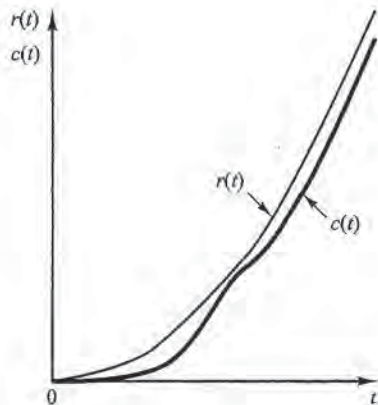
$$e_{ss} = \infty, \quad \text{for type 0 and type 1 systems}$$

$$e_{ss} = \frac{1}{K}, \quad \text{for type 2 systems}$$

$$e_{ss} = 0, \quad \text{for type 3 or higher systems}$$

Note that both type 0 and type 1 systems are incapable of following a parabolic input in the steady state. The type 2 system with unity feedback can follow a parabolic input with a finite error signal. Figure 5-51 shows an example of the response of a type 2 system with unity feedback to a parabolic input. The type 3 or higher system with unity feedback follows a parabolic input with zero error at steady state.

Figure 5-51
Response of a type 2
unity-feedback
system to a parabolic
input.



Summary. Table 5-1 summarizes the steady-state errors for type 0, type 1, and type 2 systems when they are subjected to various inputs. The finite values for steady-state errors appear on the diagonal line. Above the diagonal, the steady-state errors are infinity; below the diagonal, they are zero.

Table 5-1 Steady-State Error in Terms of Gain K

	Step Input $r(t) = 1$	Ramp Input $r(t) = t$	Acceleration Input $r(t) = \frac{1}{2}t^2$
Type 0 system	$\frac{1}{1 + K}$	∞	∞
Type 1 system	0	$\frac{1}{K}$	∞
Type 2 system	0	0	$\frac{1}{K}$

Remember that the terms *position error*, *velocity error*, and *acceleration error* mean steady-state deviations in the output position. A finite velocity error implies that after transients have died out the input and output move at the same velocity but have a finite position difference.

The error constants K_p , K_v , and K_a describe the ability of a unity-feedback system to reduce or eliminate steady-state error. Therefore, they are indicative of the steady-state performance. It is generally desirable to increase the error constants, while maintaining the transient response within an acceptable range. It is noted that to improve the steady-state performance we can increase the type of the system by adding an integrator or integrators to the feedforward path. This, however, introduces an additional stability problem. The design of a satisfactory system with more than two integrators in series in the feedforward path is generally not easy.

EXAMPLE PROBLEMS AND SOLUTIONS

A-5-1. In the system of Figure 5-52, $x(t)$ is the input displacement and $\theta(t)$ is the output angular displacement. Assume that the masses involved are negligibly small and that all motions are restricted to be small; therefore, the system can be considered linear. The initial conditions for x and θ are zeros, or $x(0-) = 0$ and $\theta(0-) = 0$. Show that this system is a differentiating element. Then obtain the response $\theta(t)$ when $x(t)$ is a unit-step input.

Solution. The equation for the system is

$$b(\dot{x} - L\dot{\theta}) = kL\theta$$

or

$$L\dot{\theta} + \frac{k}{b}L\theta = \dot{x}$$

The Laplace transform of this last equation, using zero initial conditions, gives

$$\left(Ls + \frac{k}{b}L\right)\Theta(s) = sX(s)$$

And so

$$\frac{\Theta(s)}{X(s)} = \frac{1}{L} \frac{s}{s + (k/b)}$$

Thus the system is a differentiating system.

For the unit-step input $X(s) = 1/s$, the output $\Theta(s)$ becomes

$$\Theta(s) = \frac{1}{L} \frac{1}{s + (k/b)}$$

The inverse Laplace transform of $\Theta(s)$ gives

$$\theta(t) = \frac{1}{L} e^{-(k/b)t}$$

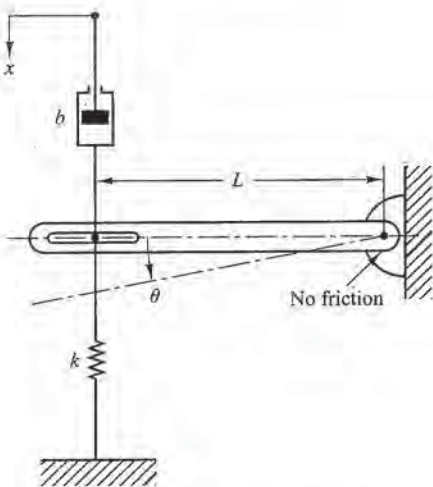
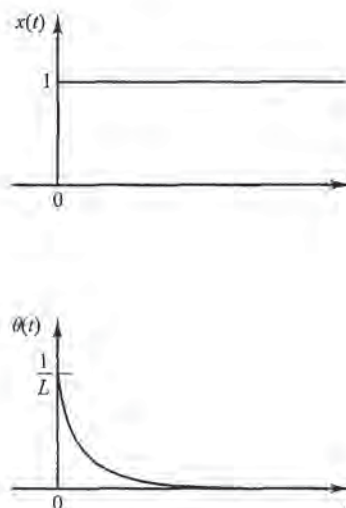


Figure 5-52
Mechanical system.

Figure 5-53
Unit-step input and
the response of the
mechanical system
shown in Figure
5-52.



Note that if the value of k/b is large the response $\theta(t)$ approaches a pulse signal as shown in Figure 5-53.

- A-5-2.** Consider the mechanical system shown in Figure 5-54. Suppose that the system is at rest initially [$x(0) = 0, \dot{x}(0) = 0$], and at $t = 0$ it is set into motion by a unit-impulse force. Obtain a mathematical model for the system. Then find the motion of the system.

Solution. The system is excited by a unit-impulse input. Hence

$$m\ddot{x} + kx = \delta(t)$$

This is a mathematical model for the system.

Taking the Laplace transform of both sides of this last equation gives

$$m[s^2X(s) - sx(0) - \dot{x}(0)] + kX(s) = 1$$

By substituting the initial conditions $x(0) = 0$ and $\dot{x}(0) = 0$ into this last equation and solving for $X(s)$, we obtain

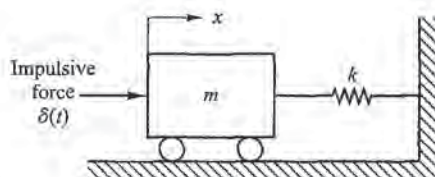
$$X(s) = \frac{1}{ms^2 + k}$$

The inverse Laplace transform of $X(s)$ becomes

$$x(t) = \frac{1}{\sqrt{mk}} \sin \sqrt{\frac{k}{m}} t$$

The oscillation is a simple harmonic motion. The amplitude of the oscillation is $1/\sqrt{mk}$.

Figure 5-54
Mechanical system.



- A-5-3.** Gear trains are often used in servo systems to reduce speed, to magnify torque, or to obtain the most efficient power transfer by matching the driving member to the given load.

Consider the gear train system shown in Figure 5-55. In this system, a load is driven by a motor through the gear train. Assuming that the stiffness of the shafts of the gear train is infinite (there is neither backlash nor elastic deformation) and that the number of teeth on each gear is proportional to the radius of the gear, obtain the equivalent moment of inertia and equivalent viscous-friction coefficient referred to the motor shaft and referred to the load shaft.

In Figure 5-55 the numbers of teeth on gears 1, 2, 3, and 4 are N_1, N_2, N_3 , and N_4 , respectively. The angular displacements of shafts 1, 2, and 3 are θ_1, θ_2 , and θ_3 , respectively. Thus, $\theta_2/\theta_1 = N_1/N_2$ and $\theta_3/\theta_2 = N_3/N_4$. The moment of inertia and viscous-friction coefficient of each gear train component are denoted by $J_1, b_1; J_2, b_2$; and J_3, b_3 , respectively. (J_3 and b_3 include the moment of inertia and friction of the load.)

Solution. For this gear train system, we can obtain the following equations: For shaft 1,

$$J_1 \ddot{\theta}_1 + b_1 \dot{\theta}_1 + T_1 = T_m \quad (5-63)$$

where T_m is the torque developed by the motor and T_1 is the load torque on gear 1 due to the rest of the gear train. For shaft 2,

$$J_2 \ddot{\theta}_2 + b_2 \dot{\theta}_2 + T_3 = T_2 \quad (5-64)$$

where T_2 is the torque transmitted to gear 2 and T_3 is the load torque on gear 3 due to the rest of the gear train. Since the work done by gear 1 is equal to that of gear 2,

$$T_1 \theta_1 = T_2 \theta_2 \quad \text{or} \quad T_2 = T_1 \frac{N_2}{N_1}$$

If $N_1/N_2 < 1$, the gear ratio reduces the speed as well as magnifies the torque. For shaft 3,

$$J_3 \ddot{\theta}_3 + b_3 \dot{\theta}_3 + T_L = T_4 \quad (5-65)$$

where T_L is the load torque and T_4 is the torque transmitted to gear 4. T_3 and T_4 are related by

$$T_4 = T_3 \frac{N_4}{N_3}$$

and θ_3 and θ_1 are related by

$$\theta_3 = \theta_2 \frac{N_3}{N_4} = \theta_1 \frac{N_1}{N_2} \frac{N_3}{N_4}$$

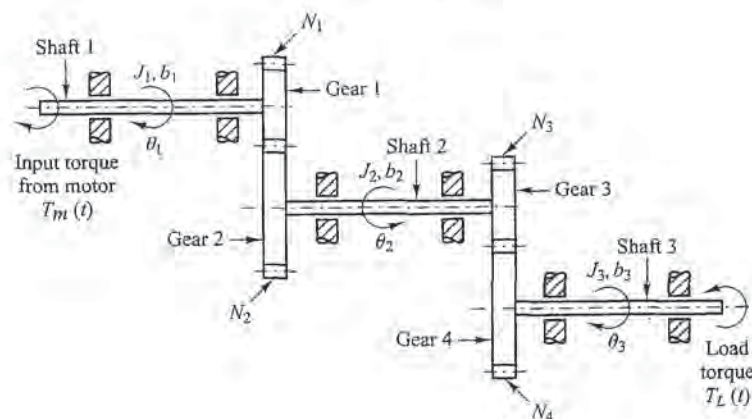


Figure 5-55
Gear train system.

Eliminating T_1 , T_2 , T_3 , and T_4 from Equations (5-63), (5-64), and (5-65) yields

$$J_1 \ddot{\theta}_1 + b_1 \dot{\theta}_1 + \frac{N_1}{N_2} (J_2 \ddot{\theta}_2 + b_2 \dot{\theta}_2) + \frac{N_1 N_3}{N_2 N_4} (J_3 \ddot{\theta}_3 + b_3 \dot{\theta}_3 + T_L) = T_m$$

Eliminating θ_2 and θ_3 from this last equation and writing the resulting equation in terms of θ_1 and its time derivatives, we obtain

$$\begin{aligned} & \left[J_1 + \left(\frac{N_1}{N_2} \right)^2 J_2 + \left(\frac{N_1}{N_2} \right)^2 \left(\frac{N_3}{N_4} \right)^2 J_3 \right] \ddot{\theta}_1 \\ & + \left[b_1 + \left(\frac{N_1}{N_2} \right)^2 b_2 + \left(\frac{N_1}{N_2} \right)^2 \left(\frac{N_3}{N_4} \right)^2 b_3 \right] \dot{\theta}_1 + \left(\frac{N_1}{N_2} \right) \left(\frac{N_3}{N_4} \right) T_L = T_m \end{aligned} \quad (5-66)$$

Thus, the equivalent moment of inertia and viscous-friction coefficient of the gear train referred to shaft 1 are given, respectively, by

$$\begin{aligned} J_{1eq} &= J_1 + \left(\frac{N_1}{N_2} \right)^2 J_2 + \left(\frac{N_1}{N_2} \right)^2 \left(\frac{N_3}{N_4} \right)^2 J_3 \\ b_{1eq} &= b_1 + \left(\frac{N_1}{N_2} \right)^2 b_2 + \left(\frac{N_1}{N_2} \right)^2 \left(\frac{N_3}{N_4} \right)^2 b_3 \end{aligned}$$

Similarly, the equivalent moment of inertia and viscous-friction coefficient of the gear train referred to the load shaft (shaft 3) are given, respectively, by

$$\begin{aligned} J_{3eq} &= J_3 + \left(\frac{N_4}{N_3} \right)^2 J_2 + \left(\frac{N_2}{N_1} \right)^2 \left(\frac{N_4}{N_3} \right)^2 J_1 \\ b_{3eq} &= b_3 + \left(\frac{N_4}{N_3} \right)^2 b_2 + \left(\frac{N_2}{N_1} \right)^2 \left(\frac{N_4}{N_3} \right)^2 b_1 \end{aligned}$$

The relationship between J_{1eq} and J_{3eq} is thus

$$J_{1eq} = \left(\frac{N_1}{N_2} \right)^2 \left(\frac{N_3}{N_4} \right)^2 J_{3eq}$$

and that between b_{1eq} and b_{3eq} is

$$b_{1eq} = \left(\frac{N_1}{N_2} \right)^2 \left(\frac{N_3}{N_4} \right)^2 b_{3eq}$$

The effect of J_2 and J_3 on an equivalent moment of inertia is determined by the gear ratios N_1/N_2 and N_3/N_4 . For speed-reducing gear trains, the ratios, N_1/N_2 and N_3/N_4 are usually less than unity. If $N_1/N_2 \ll 1$ and $N_3/N_4 \ll 1$, then the effect of J_2 and J_3 on the equivalent moment of inertia J_{1eq} is negligible. Similar comments apply to the equivalent viscous-friction coefficient b_{1eq} of the gear train. In terms of the equivalent moment of inertia J_{1eq} and equivalent viscous-friction coefficient b_{1eq} , Equation (5-66) can be simplified to give

$$J_{1eq} \ddot{\theta}_1 + b_{1eq} \dot{\theta}_1 + n T_L = T_m$$

where

$$n = \frac{N_1 N_3}{N_2 N_4}$$

A-5-4. When the system shown in Figure 5-56(a) is subjected to a unit-step input, the system output responds as shown in Figure 5-56(b). Determine the values of K and T from the response curve.

Solution. The maximum overshoot of 25.4% corresponds to $\zeta = 0.4$. From the response curve we have

$$t_p = 3$$

Consequently,

$$t_p = \frac{\pi}{\omega_d} = \frac{\pi}{\omega_n \sqrt{1 - \zeta^2}} = \frac{\pi}{\omega_n \sqrt{1 - 0.4^2}} = 3$$

It follows that

$$\omega_n = 1.14$$

From the block diagram we have

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

from which

$$\omega_n = \sqrt{\frac{K}{T}}, \quad 2\zeta\omega_n = \frac{1}{T}$$

Therefore, the values of T and K are determined as

$$T = \frac{1}{2\zeta\omega_n} = \frac{1}{2 \times 0.4 \times 1.14} = 1.09$$

$$K = \omega_n^2 T = 1.14^2 \times 1.09 = 1.42$$

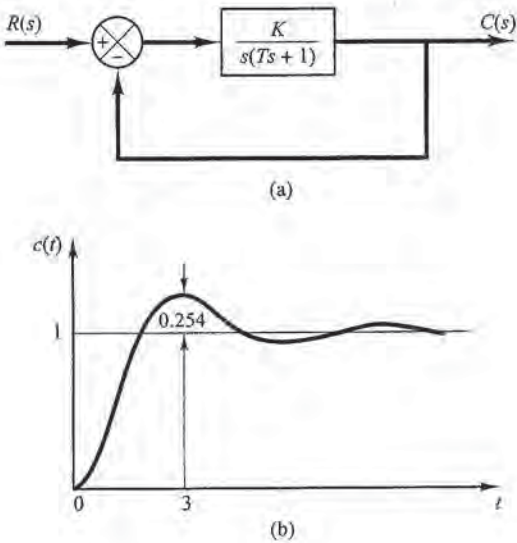


Figure 5-56
(a) Closed-loop system; (b) unit-step response curve.

- A-5-5.** Determine the values, of K and k of the closed-loop system shown in Figure 5-57 so that the maximum overshoot in unit-step response is 25% and the peak time is 2 sec. Assume that $J = 1 \text{ kg-m}^2$.

Solution. The closed-loop transfer function is

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + Kks + K}$$

By substituting $J = 1 \text{ kg-m}^2$ into this last equation, we have

$$\frac{C(s)}{R(s)} = \frac{K}{s^2 + Kks + K}$$

Note that in this problem

$$\omega_n = \sqrt{K}, \quad 2\zeta\omega_n = Kk$$

The maximum overshoot M_p is

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

which is specified as 25%. Hence

$$e^{-\zeta\pi/\sqrt{1-\zeta^2}} = 0.25$$

from which

$$\frac{\zeta\pi}{\sqrt{1-\zeta^2}} = 1.386$$

or

$$\zeta = 0.404$$

The peak time t_p is specified as 2 sec. And so

$$t_p = \frac{\pi}{\omega_d} = 2$$

or

$$\omega_d = 1.57$$

Then the undamped natural frequency ω_n is

$$\omega_n = \frac{\omega_d}{\sqrt{1-\zeta^2}} = \frac{1.57}{\sqrt{1-0.404^2}} = 1.72$$

Therefore, we obtain

$$K = \omega_n^2 = 1.72^2 = 2.95 \text{ N-m}$$

$$k = \frac{2\zeta\omega_n}{K} = \frac{2 \times 0.404 \times 1.72}{2.95} = 0.471 \text{ sec}$$

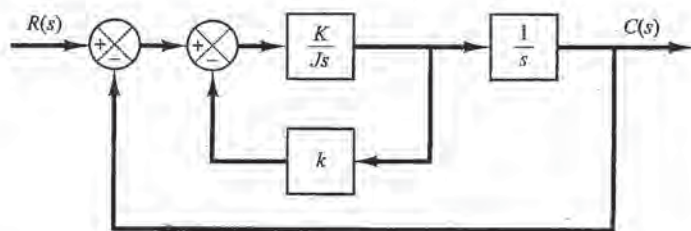


Figure 5-57
Closed-loop system.

A-5-6. Figures 5-58(a) shows a mechanical vibratory system. When 2 lb of force (step input) is applied to the system, the mass oscillates, as shown in Figure 5-58(b). Determine m , b , and k of the system from this response curve. The displacement x is measured from the equilibrium position.

Solution. The transfer function of this system is

$$\frac{X(s)}{P(s)} = \frac{1}{ms^2 + bs + k}$$

Since

$$P(s) = \frac{2}{s}$$

we obtain

$$X(s) = \frac{2}{s(ms^2 + bs + k)}$$

It follows that the steady-state value of x is

$$x(\infty) = \lim_{s \rightarrow 0} sX(s) = \frac{2}{k} = 0.1 \text{ ft}$$

Hence

$$k = 20 \text{ lb}_f/\text{ft}$$

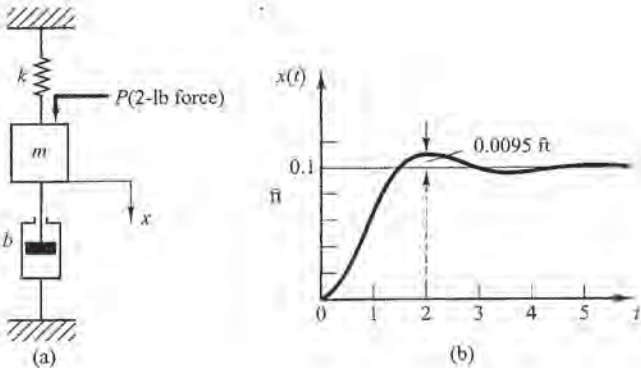
Note that $M_p = 9.5\%$ corresponds to $\zeta = 0.6$. The peak time t_p is given by

$$t_p = \frac{\pi}{\omega_d} = \frac{\pi}{\omega_n \sqrt{1 - \zeta^2}} = \frac{\pi}{0.8\omega_n}$$

The experimental curve shows that $t_p = 2$ sec. Therefore,

$$\omega_n = \frac{3.14}{2 \times 0.8} = 1.96 \text{ rad/sec}$$

Figure 5-58
(a) Mechanical vibratory system;
(b) step-response curve.



Since $\omega_n^2 = k/m = 20/m$, we obtain

$$m = \frac{20}{\omega_n^2} = \frac{20}{1.96^2} = 5.2 \text{ slugs} = 166 \text{ lb}$$

(Note that 1 slug = 1 lb_f-sec²/ft.) Then b is determined from

$$2\zeta\omega_n = \frac{b}{m}$$

or

$$b = 2\zeta\omega_n m = 2 \times 0.6 \times 1.96 \times 5.2 = 12.2 \text{ lb}_f/\text{ft}/\text{sec}$$

A-5-7. Consider the unit-step response of the second-order system

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

The amplitude of the exponentially-damped sinusoid changes as a geometric series. At time $t = t_p = \pi/\omega_d$, the amplitude is equal to $e^{-(\sigma/\omega_d)\pi}$. After one oscillation, or at $t = t_p + 2\pi/\omega_d = 3\pi/\omega_d$, the amplitude is equal to $e^{-(\sigma/\omega_d)3\pi}$; after another cycle of oscillation, the amplitude is $e^{-(\sigma/\omega_d)5\pi}$. The logarithm of the ratio of successive amplitudes is called the *logarithmic decrement*. Determine the logarithmic decrement for this second-order system. Describe a method for experimental determination of the damping ratio from the rate of decay of the oscillation.

Solution. Let us define the amplitude of the output oscillation at $t = t_i$ to be x_i , where $t_i = t_p + (i - 1)T$ (T = period of oscillation). The amplitude ratio per one period of damped oscillation is

$$\frac{x_1}{x_2} = \frac{e^{-(\sigma/\omega_d)\pi}}{e^{-(\sigma/\omega_d)3\pi}} = e^{2(\sigma/\omega_d)\pi} = e^{2\zeta\pi/\sqrt{1-\zeta^2}}$$

Thus, the logarithmic decrement δ is

$$\delta = \ln \frac{x_1}{x_2} = \frac{2\zeta\pi}{\sqrt{1-\zeta^2}}$$

It is a function only of the damping ratio ζ . Thus, the damping ratio ζ can be determined by use of the logarithmic decrement.

In the experimental determination of the damping ratio ζ from the rate of decay of the oscillation, we measure the amplitude x_1 at $t = t_p$ and amplitude x_n at $t = t_p + (n - 1)T$. Note that it is necessary to choose n large enough so that the ratio of x_1/x_n is not near unity. Then

$$\frac{x_1}{x_n} = e^{(n-1)2\zeta\pi/\sqrt{1-\zeta^2}}$$

or

$$\ln \frac{x_1}{x_n} = (n - 1) \frac{2\zeta\pi}{\sqrt{1-\zeta^2}}$$

Hence

$$\zeta = \frac{\frac{1}{n-1} \left(\ln \frac{x_1}{x_n} \right)}{\sqrt{4\pi^2 + \left[\frac{1}{n-1} \left(\ln \frac{x_1}{x_n} \right) \right]^2}}$$

Example Problems and Solutions

301

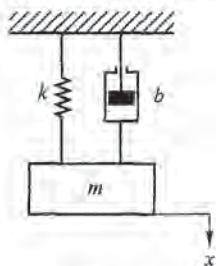


Figure 5-59
Spring-mass-damper system.

- A-5-8.** In the system shown in Figure 5-59, the numerical values of m , b , and k are given as $m = 1$ kg, $b = 2$ N-sec/m, and $k = 100$ N/m. The mass is displaced 0.05 m and released without initial velocity. Find the frequency observed in the vibration. In addition, find the amplitude four cycles later. The displacement x is measured from the equilibrium position.

Solution. The equation of motion for the system is

$$m\ddot{x} + b\dot{x} + kx = 0$$

Substituting the numerical values for m , b , and k into this equation gives

$$\ddot{x} + 2\dot{x} + 100x = 0$$

where the initial conditions are $x(0) = 0.05$ and $\dot{x}(0) = 0$. From this last equation the undamped natural frequency ω_n and the damping ratio ζ are found to be

$$\omega_n = 10, \quad \zeta = 0.1$$

The frequency actually observed in the vibration is the damped natural frequency ω_d .

$$\omega_d = \omega_n \sqrt{1 - \zeta^2} = 10 \sqrt{1 - 0.01} = 9.95 \text{ rad/sec}$$

In the present analysis, $\dot{x}(0)$ is given as zero. Thus, solution $x(t)$ can be written as

$$x(t) = x(0)e^{-\zeta\omega_n t} \left(\cos \omega_d t + \frac{\zeta}{\sqrt{1 - \zeta^2}} \sin \omega_d t \right)$$

It follows that at $t = nT$, where $T = 2\pi/\omega_d$,

$$x(nT) = x(0)e^{-\zeta\omega_n nT}$$

Consequently, the amplitude four cycles later becomes

$$\begin{aligned} x(4T) &= x(0)e^{-\zeta\omega_n 4T} = x(0)e^{-(0.1)(10)(4)(0.6315)} \\ &= 0.05e^{-2.526} = 0.05 \times 0.07998 = 0.004 \text{ m} \end{aligned}$$

- A-5-9.** Obtain both analytically and computationally the unit-step response of the following higher-order system:

$$\frac{C(s)}{R(s)} = \frac{3s^3 + 25s^2 + 72s + 80}{s^4 + 8s^3 + 40s^2 + 96s + 80}$$

[Obtain the partial-fraction expansion of $C(s)$ with MATLAB when $R(s)$ is a unit-step function.]

Solution. MATLAB Program 5-19 yields the unit-step response curve shown in Figure 5-60. It also yields the partial-fraction expansion of $C(s)$ as follows:

$$\begin{aligned} C(s) &= \frac{3s^3 + 25s^2 + 72s + 80}{s^4 + 8s^3 + 40s^2 + 96s + 80} \cdot \frac{1}{s} \\ &= \frac{-0.2813 - j0.1719}{s + 2 - j4} + \frac{-0.2813 + j0.1719}{s + 2 + j4} \\ &\quad + \frac{-0.4375}{s + 2} + \frac{-0.375}{(s + 2)^2} + \frac{1}{s} \\ &= \frac{-0.5626(s + 2)}{(s + 2)^2 + 4^2} + \frac{(0.3438) \times 4}{(s + 2)^2 + 4^2} \\ &\quad - \frac{0.4375}{s + 2} - \frac{0.375}{(s + 2)^2} + \frac{1}{s} \end{aligned}$$

MATLAB Program 5-19

% ----- Unit-Step Response of C(s)/R(s) and Partial-Fraction Expansion of C(s) -----

```
num = [0 3 25 72 80];  
den = [1 8 40 96 80];  
step(num,den);  
v = [0 3 0 1.2]; axis(v), grid
```

% To obtain the partial-fraction expansion of C(s), enter commands

```
% num1 = [0 0 3 25 72 80];  
% den1 = [1 8 40 96 80 0];  
% [r,p,k] = residue(num1,den1)
```

```
num1 = [0 0 3 25 72 80];  
den1 = [1 8 40 96 80 0];  
[r,p,k] = residue(num1,den1)
```

r =

```
-0.2813- 0.1719i  
-0.2813+ 0.1719i  
-0.4375  
-0.3750  
1.0000
```

p =

```
-2.0000+ 4.0000i  
-2.0000- 4.0000i  
-2.0000  
-2.0000  
0
```

k =

```
[]
```

Hence, the time response $c(t)$ can be given by

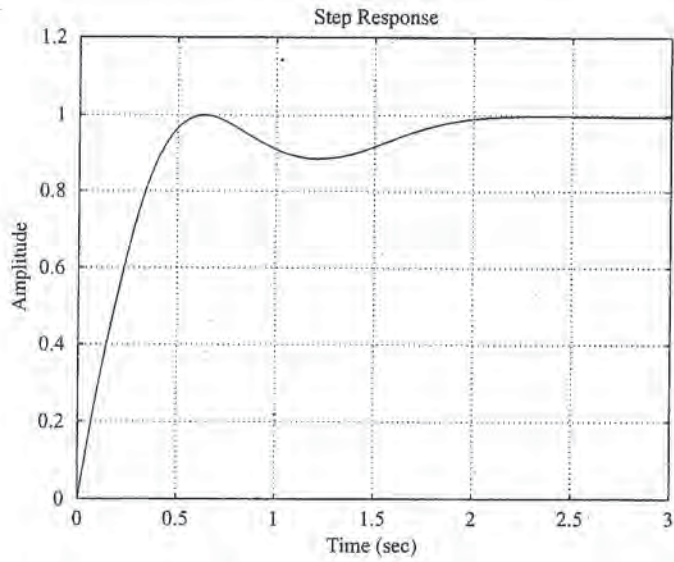
$$c(t) = -0.5626e^{-2t} \cos 4t + 0.3438e^{-2t} \sin 4t \\ - 0.4375e^{-2t} - 0.375te^{-2t} + 1$$

The fact that the response curve is an exponential curve superimposed by damped sinusoidal curves can be seen from Figure 5-60.

Example Problems and Solutions

303

Figure 5-60
Unit-step response
curve.



- A-5-10.** Obtain both analytical and computational solutions of the unit-step response of a unity-feedback system whose open-loop transfer function is

$$G(s) = \frac{5(s + 20)}{s(s + 4.59)(s^2 + 3.41s + 16.35)}$$

Solution. The closed-loop transfer function is

$$\begin{aligned} \frac{C(s)}{R(s)} &= \frac{5(s + 20)}{s(s + 4.59)(s^2 + 3.41s + 16.35) + 5(s + 20)} \\ &= \frac{5s + 100}{s^4 + 8s^3 + 32s^2 + 80s + 100} \\ &= \frac{5(s + 20)}{(s^2 + 2s + 10)(s^2 + 6s + 10)} \end{aligned}$$

The unit-step response of this system is then

$$\begin{aligned} C(s) &= \frac{5(s + 20)}{s(s^2 + 2s + 10)(s^2 + 6s + 10)} \\ &= \frac{1}{s} + \frac{\frac{3}{8}(s + 1) - \frac{17}{8}}{(s + 1)^2 + 3^2} + \frac{-\frac{11}{8}(s + 3) - \frac{13}{8}}{(s + 3)^2 + 1^2} \end{aligned}$$

The time response $c(t)$ can be found by taking the inverse Laplace transform of $C(s)$ as follows:

$$c(t) = 1 + \frac{3}{8}e^{-t}\cos 3t - \frac{17}{24}e^{-t}\sin 3t - \frac{11}{8}e^{-3t}\cos t - \frac{13}{8}e^{-3t}\sin t, \quad \text{for } t \geq 0$$

A MATLAB program to obtain the unit-step response of this system is shown in MATLAB Program 5-20. The resulting unit-step response curve is shown in Figure 5-61.

MATLAB Program 5-20

```
% ----- Unit-step-response -----
num = [0 0 0 5 100];
den = [1 8 32 80 100];
step(num,den)
grid
title('Unit-Step Response of C(s)/R(s) = (5s + 100)/(s^4 + 8s^3 + 32s^2 + 80s + 100)')
```

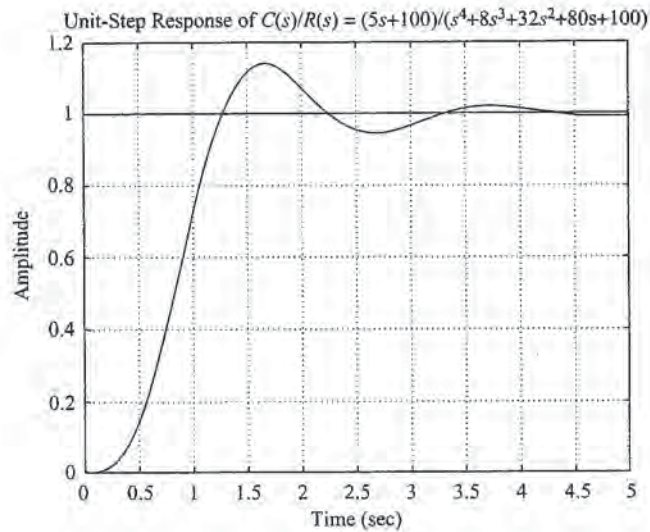


Figure 5-61
Unit-step response curve.

- A-5-11.** When the closed-loop system involves a numerator dynamics, the unit-step response curve may exhibit a large overshoot. Obtain the unit-step response of the following system with MATLAB:

$$\frac{C(s)}{R(s)} = \frac{10s + 4}{s^2 + 4s + 4}$$

Obtain also the unit-ramp response with MATLAB.

Solution. MATLAB Program 5-21 produces the unit-step response as well as the unit-ramp response of the system. The unit-step response curve and unit-ramp response curve, together with the unit-ramp input, are shown in Figures 5-62(a) and (b), respectively.

Notice that the unit-step response curve exhibits over 215% of overshoot. The unit-ramp response curve leads the input curve. These phenomena occurred because of the presence of a large derivative term in the numerator.

MATLAB Program 5-21

```
num = [0 10 4];
den = [1 4 4];
t = 0:0.02:10;
y = step(num,den,t);
plot(t,y)
grid
title('Unit-Step Response')
xlabel('t (sec)')
ylabel('Output')

num1 = [0 0 10 4];
den1 = [1 4 4 0];
y1 = step(num1,den1,t);
plot(t,t,'--',t,y1)
v = [0 10 0 10]; axis(v);
grid
title('Unit-Ramp Response')
xlabel('t (sec)')
ylabel('Unit-Ramp Input and Output')
text(6.1,5.0,'Unit-Ramp Input')
text(3.5,7.1,'Output')
```

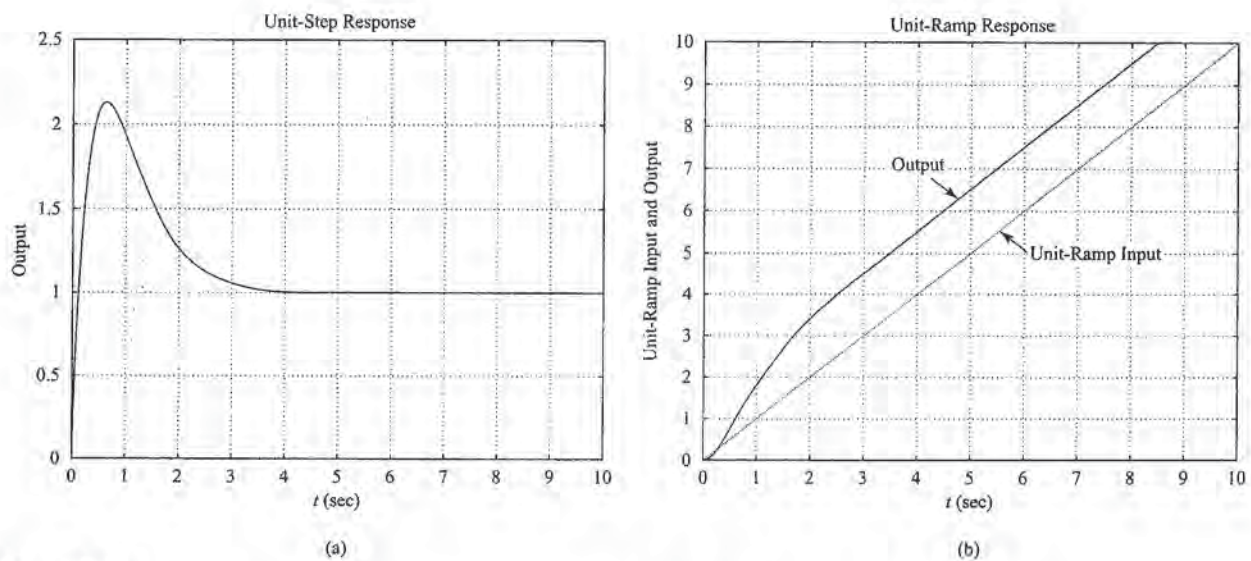


Figure 5-62
(a) Unit-step response curve; (b) unit-ramp response curve plotted with unit-ramp input.

A-5-12. Consider a higher-order system defined by

$$\frac{C(s)}{R(s)} = \frac{6.3223s^2 + 18s + 12.811}{s^4 + 6s^3 + 11.3223s^2 + 18s + 12.811}$$

Using MATLAB, plot the unit-step response curve of this system. Using MATLAB, obtain the rise time, peak time, maximum overshoot, and settling time.

Solution. MATLAB Program 5-22 plots the unit-step response curve as well as to give the rise time, peak time, maximum overshoot, and settling time. The unit-step response curve is shown in Figure 5-63.

MATLAB Program 5-22

```
% ----- This program is to plot the unit-step response curve, as well as to
% find the rise time, peak time, maximum overshoot, and settling time.
% In this program the rise time is calculated as the time required for the
% response to rise from 10% to 90% of its final value. -----

num = [0 0 6.3223 18 12.811];
den = [1 6 11.3223 18 12.811];
t = 0:0.02:20;
[y,x,t] = step(num,den,t);
plot(t,y)
grid
title('Unit-Step Response')
xlabel('t (sec)')
ylabel('Output y(t)')

r1 = 1; while y(r1) < 0.1, r1 = r1+1; end;
r2 = 1; while y(r2) < 0.9, r2 = r2+1; end;
rise_time = (r2-r1)*0.02

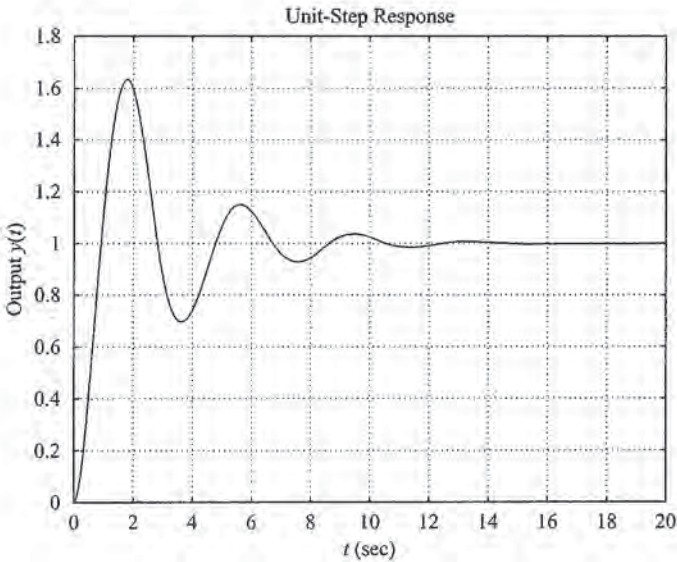
rise_time =
    0.5800

[ymax,tp] = max(y);
peak_time = (tp-1)*0.02
peak_time =
    1.6600

max_overshoot = ymax-1
max_overshoot =
    0.6182

s = 1001; while y(s) > 0.98 & y(s) < 1.02; s = s-1; end;
settling_time = (s-1)*0.02
settling_time =
    10.0200
```

Figure 5-63
Unit-step response
curve.



A-5-13. Consider the closed-loop system defined by

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

Using a “for loop,” write a MATLAB program to obtain unit-step response of this system for the following four cases:

- Case 1: $\zeta = 0.3, \quad \omega_n = 1$
- Case 2: $\zeta = 0.5, \quad \omega_n = 2$
- Case 3: $\zeta = 0.7, \quad \omega_n = 4$
- Case 4: $\zeta = 0.8, \quad \omega_n = 6$

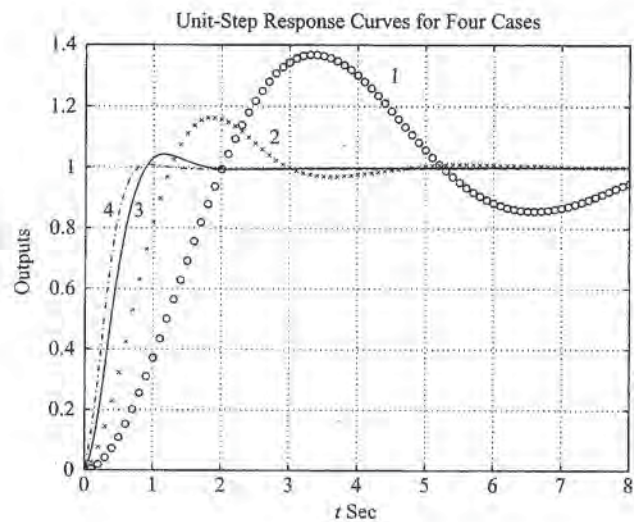
Solution. Define $\omega_n^2 = a$ and $2\zeta\omega_n = b$. Then, a and b each have four elements as follows:

$$\begin{aligned} a &= [1 \quad 4 \quad 16 \quad 36] \\ b &= [0.6 \quad 2 \quad 5.6 \quad 9.6] \end{aligned}$$

Using vectors a and b , MATLAB Program 5-23 will produce the unit-step response curves as shown in Figure 5-64.

MATLAB Program 5-23
<pre> a = [1 4 16 36]; b = [0.6 2 5.6 9.6]; t = 0:0.1:8; y = zeros(81,4); for i = 1:4; num = [0 0 a(i)]; den = [1 b(i) a(i)]; y(:,i) = step(num,den,t); end plot(t,y(:,1),'o',t,y(:,2),'x',t,y(:,3),'-',t,y(:,4),'-.-') grid title('Unit-Step Response Curves for Four Cases') xlabel('t Sec') ylabel('Outputs') gtext('1') gtext('2') gtext('3') gtext('4') </pre>

Figure 5-64
Unit-step response
curves for four cases.



- A-5-14.** Using MATLAB, obtain the unit-ramp response of the closed-loop control system whose closed-loop transfer function is

$$\frac{C(s)}{R(s)} = \frac{s + 10}{s^3 + 6s^2 + 9s + 10}$$

Also, obtain the response of this system when the input is given by

$$r = e^{-0.5t}$$

Solution. MATLAB Program 5-24 produces the unit-ramp response and the response to the exponential input $r = e^{-0.5t}$. The resulting response curves are shown in Figures 5-65(a) and (b), respectively.

MATLAB Program 5-24

```
% ----- Unit-Ramp Response -----

num = [0 0 1 10];
den = [1 6 9 10];
t = 0:0.1:10;
r = t;
y = lsim(num,den,r,t);
plot(t,r,'-t,y,'o')
grid
title('Unit-Ramp Response by Use of Command "lsim"')
xlabel('t Sec')
ylabel('Output')
text(3.2,6.5,'Unit-Ramp Input')
text(6.0,3.1,'Output')

% ----- Response to Input r1 = exp(-0.5t). -----

num = [0 0 1 10];
den = [1 6 9 10];
t = 0:0.1:12;
r1 = exp(-0.5*t);
y1 = lsim(num,den,r1,t);
plot(t,r1,'-t,y1,'o')
grid
title('Response to Input r1 = exp(-0.5t)')
xlabel('t Sec')
ylabel('Input and Output')
text(1.4,0.75,'Input r1 = exp(-0.5t)')
text(6.2,0.34,'Output')
```

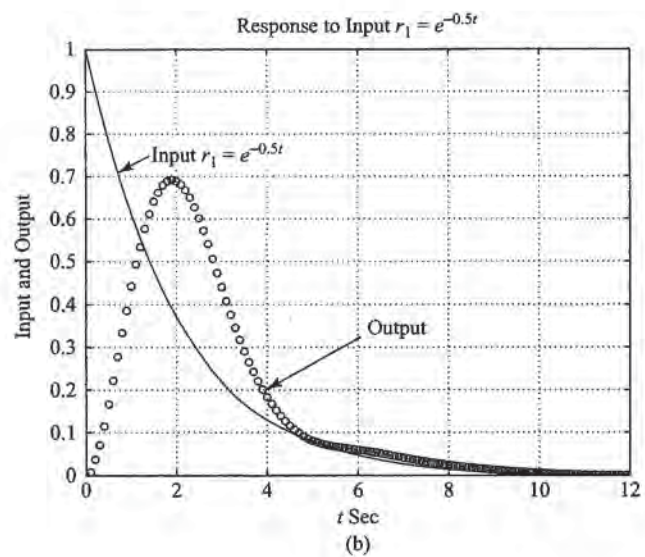
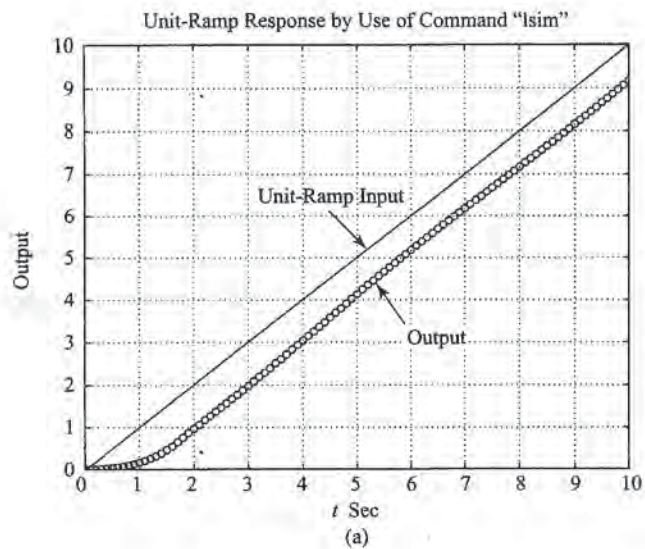


Figure 5-65
(a) Unit-ramp response curve;
(b) response to exponential input $r = e^{-0.5t}$.

A-5-15. Obtain the response of the closed-loop system defined by

$$\frac{C(s)}{R(s)} = \frac{5}{s^2 + s + 5}$$

when the input $r(t)$ is given by

$$r(t) = 2 + t$$

[The input $r(t)$ is a step input of magnitude 2 plus unit-ramp input.]

Example Problems and Solutions

311

Solution. A possible MATLAB program is shown in MATLAB Program 5-25. The resulting response curve, together with a plot of the input function, is shown in Figure 5-66.

MATLAB Program 5-25

```
num = [0 0 5];
den = [1 1 5];
t = 0:0.05:10;
r = 2+t;
c = lsim(num,den,r,t);
plot(t,r,'-',t,c,'o')
grid
title('Response to Input r(t) = 2 + t')
xlabel('t Sec')
ylabel('Output c(t) and Input r(t) = 2 + t')
```

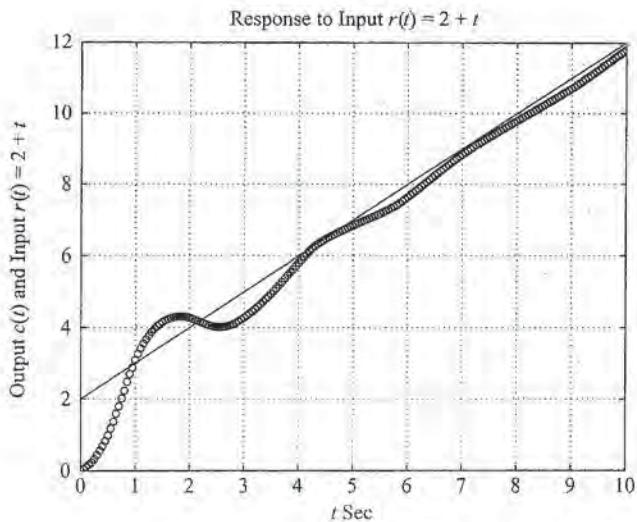


Figure 5-66
Response to input $r(t) = 2 + t$.

A-5-16. Obtain the response of the system shown in Figure 5-67 when the input $r(t)$ is given by

$$r(t) = \frac{1}{2}t^2$$

[The input $r(t)$ is the unit-acceleration input.]

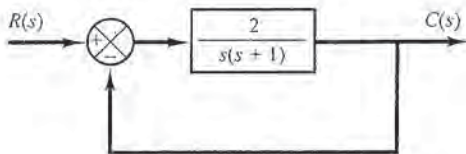


Figure 5-67
Control system.

Solution. The closed-loop transfer function is

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

MATLAB Program 5-26 produces the unit-accerelation response. The resulting response, together with the unit-acceleration input, is shown in Figure 5-68.

MATLAB Program 5-26

```
num = [0 0 2];
den = [1 1 2];
t = 0:0.2:10;
r = 0.5*t.^2;
y = lsim(num,den,r,t);
plot(t,r,'-',t,y,'o','t,y','-')
grid
title('Unit-Acceleration Response')
xlabel('t Sec')
ylabel('Input and Output')
text(2.1,27.5,'Unit-Acceleration Input')
text(7.2,7.5,'Output')
```

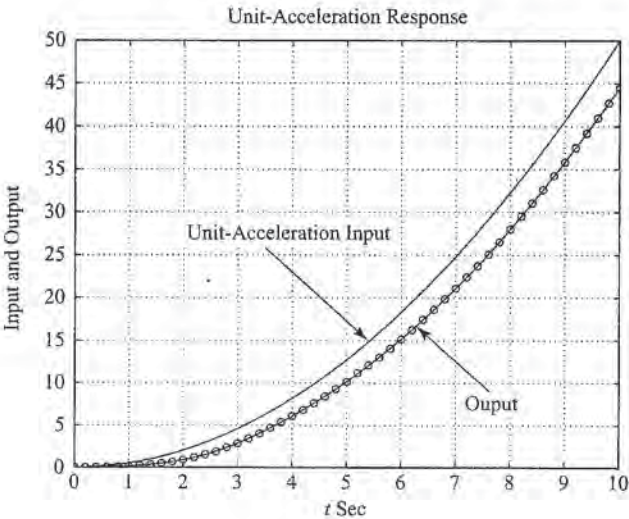


Figure 5-68
Response to unit-acceleration input.

A-5-17. Consider the system defined by

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

where $\zeta = 0, 0.2, 0.4, 0.6, 0.8$, and 1.0 . Write a MATLAB program using a “for loop” to obtain the two-dimensional and three-dimensional plots of the system output. The input is the unit-step function.

Solution. MATLAB Program 5–27 is a possible program to obtain two-dimensional and three-dimensional plots. Figure 5–69(a) is the two-dimensional plot of the unit-step response curves for various values of ζ . Figure 5–69(b) is the three-dimensional plot obtained by use of the command “mesh(y)” and Figure 5–69(c) is obtained by use of the command “mesh(y’).” (These two three-dimensional plots are basically the same. The only difference is that x axis and y axis are interchanged.)

MATLAB Program 5–27

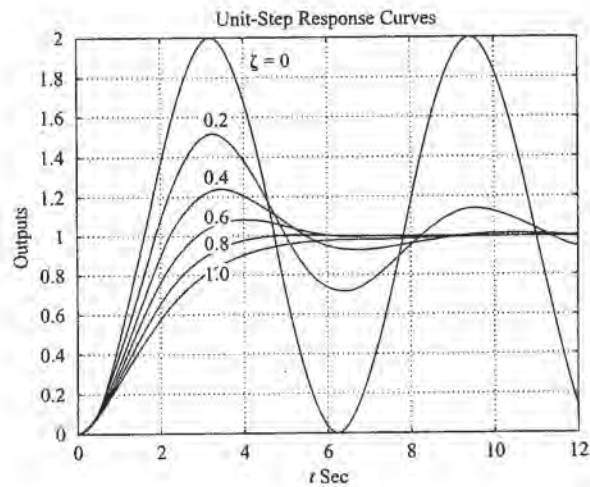
```
t = 0:0.2:12;
for n = 1:6;
    num = [0 0 1];
    den = [1 2*(n-1)*0.2 1];
    [y(1:61,n),x,t] = step(num,den,t);
end
plot(t,y)
grid
title('Unit-Step Response Curves')
xlabel('t Sec')
ylabel('Outputs')
gtext('\zeta = 0'),
gtext('0.2')
gtext('0.4')
gtext('0.6')
gtext('0.8')
gtext('1.0')

% To draw a three-dimensional plot, enter the following command: mesh(y) or mesh(y').
% We shall show two three-dimensional plots, one using "mesh(y)" and the other using
% "mesh(y')". These two plots are the same, except that the x axis and y axis are
% interchanged.

mesh(y)
title('Three-Dimensional Plot of Unit-Step Response Curves using Command "mesh(y)"')
xlabel('n, where n = 1,2,3,4,5,6')
ylabel('Computation Time Points')
zlabel('Outputs')

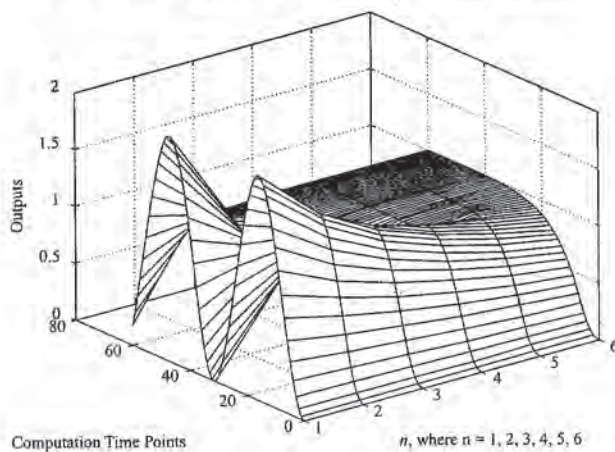
mesh(y')
title('Three-Dimensional Plot of Unit-Step Response Curves using Command "mesh(y transpose)"')
xlabel('Computation Time Points')
ylabel('n, where n = 1,2,3,4,5,6')
zlabel('Outputs')
```

Figure 5-69
 (a) Two-dimensional plot of unit-step response curves;
 (b) three-dimensional plot of unit-step response curves using command “mesh(y)”;
 (c) three-dimensional plot of unit-step response curves using command “mesh(y’)”.



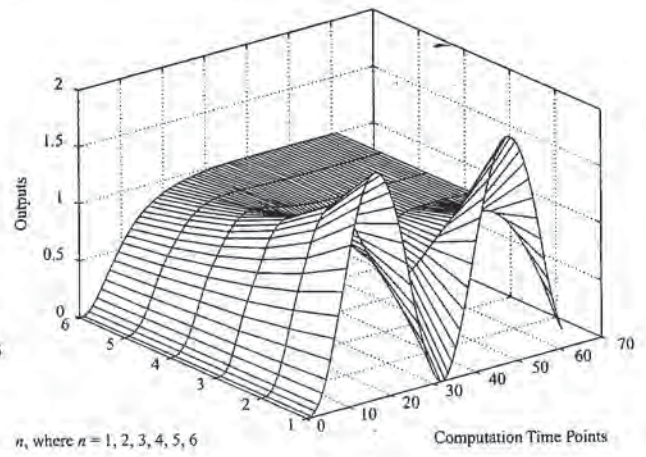
(a)

Three-Dimensional Plot of Unit-Step Response Curves using Command “mesh(y)”



(b)

Three-Dimensional Plot of Unit-Step Response Curves using Command “mesh(y transpose)”



(c)

A-5-18. Consider the following characteristic equation:

$$s^4 + Ks^3 + s^2 + s + 1 = 0$$

Determine the range of K for stability.

Solution. The Routh array of coefficients is

s^4	1	1	1
s^3	K	1	0
s^2	$\frac{K-1}{K}$	1	0
s^1	$1 - \frac{K^2}{K-1}$	0	
s^0	1		

For stability, we require that

$$\begin{aligned} K &> 0 \\ \frac{K-1}{K} &> 0 \\ 1 - \frac{K^2}{K-1} &> 0 \end{aligned}$$

From the first and second conditions, K must be greater than 1. For $K > 1$, notice that the term $1 - [K^2/(K-1)]$ is always negative, since

$$\frac{K-1-K^2}{K-1} = \frac{-1+K(1-K)}{K-1} < 0$$

Thus, the three conditions cannot be fulfilled simultaneously. Therefore, there is no value of K that allows stability of the system.

A-5-19. Consider the characteristic equation given by

$$a_0 s^n + a_1 s^{n-1} + a_2 s^{n-2} + \cdots + a_{n-1} s + a_n = 0 \quad (5-67)$$

The Hurwitz stability criterion, given next, gives conditions for all the roots to have negative real parts in terms of the coefficients of the polynomial. As stated in the discussions of Routh's stability criterion in Section 5-7, for all the roots to have negative real parts, all the coefficients a 's must be positive. This is a necessary condition but not a sufficient condition. If this condition is not satisfied, it indicates that some of the roots have positive real parts or are imaginary or zero. A sufficient condition for all the roots to have negative real parts is given in the following Hurwitz stability criterion: If all the coefficients of the polynomial are positive, arrange these coefficients in the following determinant:

$$\Delta_n = \begin{vmatrix} a_1 & a_3 & a_5 & \cdots & 0 & 0 & 0 \\ a_0 & a_2 & a_4 & \cdots & 0 & 0 & 0 \\ 0 & a_1 & a_3 & \cdots & a_n & 0 & 0 \\ 0 & a_0 & a_2 & \cdots & a_{n-1} & 0 & 0 \\ \cdot & \cdot & \cdot & & a_{n-2} & a_n & 0 \\ \cdot & \cdot & \cdot & & a_{n-3} & a_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & a_{n-4} & a_{n-2} & a_n \end{vmatrix}$$

where we substituted zero for a_s if $s > n$. For all the roots to have negative real parts, it is necessary and sufficient that successive principal minors of Δ_n be positive. The successive principal minors are the following determinants:

$$\Delta_i = \begin{vmatrix} a_1 & a_3 & \cdots & a_{2i-1} \\ a_0 & a_2 & \cdots & a_{2i-2} \\ 0 & a_1 & \cdots & a_{2i-3} \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & a_i \end{vmatrix} \quad (i = 1, 2, \dots, n-1)$$

where $a_s = 0$ if $s > n$. (It is noted that some of the conditions for the lower-order determinants are included in the conditions for the higher-order determinants.) If all these determinants are positive, and $a_0 > 0$ as already assumed, the equilibrium state of the system whose characteristic

equation is given by Equation (5-67) is asymptotically stable. Note that exact values of determinants are not needed; instead, only signs of these determinants are needed for the stability criterion. Now consider the following characteristic equation:

$$a_0 s^4 + a_1 s^3 + a_2 s^2 + a_3 s + a_4 = 0$$

Obtain the conditions for stability using the Hurwitz stability criterion.

Solution. The conditions for stability are that all the a 's be positive and that

$$\begin{aligned}\Delta_2 &= \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} = a_1 a_2 - a_0 a_3 > 0 \\ \Delta_3 &= \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} \\ &= a_1(a_2 a_3 - a_1 a_4) - a_0 a_3^2 \\ &= a_3(a_1 a_2 - a_0 a_3) - a_1^2 a_4 > 0\end{aligned}$$

It is clear that, if all the a 's are positive and if the condition $\Delta_3 > 0$ is satisfied, the condition $\Delta_2 > 0$ is also satisfied. Therefore, for all the roots of the given characteristic equation to have negative real parts, it is necessary and sufficient that all the coefficients a 's are positive and $\Delta_3 > 0$.

A-5-20. Show that the first column of the Routh array of

$$s^n + a_1 s^{n-1} + a_2 s^{n-2} + \cdots + a_{n-1} s + a_n = 0$$

is given by

$$1, \quad \Delta_1, \quad \frac{\Delta_2}{\Delta_1}, \quad \frac{\Delta_3}{\Delta_2}, \dots, \quad \frac{\Delta_n}{\Delta_{n-1}}$$

where

$$\Delta_r = \begin{vmatrix} a_1 & 1 & 0 & 0 & \cdots & 0 \\ a_3 & a_2 & a_1 & 1 & \cdots & 0 \\ a_5 & a_4 & a_3 & a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{2r-1} & \vdots & \vdots & \vdots & \vdots & a_r \end{vmatrix}, \quad (n \geq r \geq 1)$$

$$a_k = 0 \quad \text{if } k > n$$

Solution. The Routh array of coefficients has the form

$$\begin{array}{ccccccc} 1 & a_2 & a_4 & a_6 & \cdots & a_n \\ a_1 & a_3 & a_5 & \cdots & & \\ b_1 & b_2 & b_3 & \cdots & & \\ c_1 & c_2 & \cdots & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

The first term in the first column of the Routh array is 1. The next term in the first column is a_1 , which is equal to Δ_1 . The next term is b_1 , which is equal to

$$\frac{a_1a_2-a_3}{a_1}=\frac{\Delta_2}{\Delta_1}$$

The next term in the first column is c_1 , which is equal to

$$\begin{aligned}\frac{b_1a_3-a_1b_2}{b_1}&=\frac{\left[\frac{a_1a_2-a_3}{a_1}\right]a_3-a_1\left[\frac{a_1a_4-a_5}{a_1}\right]}{\left[\frac{a_1a_2-a_3}{a_1}\right]}\\&=\frac{a_1a_2a_3-a_3^2-a_1^2a_4+a_1a_5}{a_1a_2-a_3}\\&=\frac{\Delta_3}{\Delta_2}.\end{aligned}$$

In a similar manner the remaining terms in the first column of the Routh array can be found. The Routh array has the property that the last nonzero terms of any columns are the same; that is, if the array is given by

$$\begin{array}{cccc}a_0 & a_2 & a_4 & a_6 \\a_1 & a_3 & a_5 & a_7 \\b_1 & b_2 & b_3 & \\c_1 & c_2 & c_3 & \\d_1 & d_2 & & \\e_1 & e_2 & & \\f_1 & & & \\g_1 & & & \end{array}$$

then

$$a_7=c_3=e_2=g_1$$

and if the array is given by

$$\begin{array}{cccc}a_0 & a_2 & a_4 & a_6 \\a_1 & a_3 & a_5 & 0 \\b_1 & b_2 & b_3 & \\c_1 & c_2 & 0 & \\d_1 & d_2 & & \\e_1 & 0 & & \\f_1 & & & \end{array}$$

then

$$a_6=b_3=d_2=f_1$$

In any case, the last term of the first column is equal to a_n , or

$$a_n=\frac{\Delta_{n-1}a_n}{\Delta_{n-1}}=\frac{\Delta_n}{\Delta_{n-1}}$$

For example, if $n = 4$, then

$$\Delta_4 = \begin{vmatrix} a_1 & 1 & 0 & 0 \\ a_3 & a_2 & a_1 & 1 \\ a_5 & a_4 & a_3 & a_2 \\ a_7 & a_6 & a_5 & a_4 \end{vmatrix} = \begin{vmatrix} a_1 & 1 & 0 & 0 \\ a_3 & a_2 & a_1 & 1 \\ 0 & a_4 & a_3 & a_2 \\ 0 & 0 & 0 & a_4 \end{vmatrix} = \Delta_3 a_4$$

Thus it has been shown that the first column of the Routh array is given by

$$1, \quad \Delta_1, \quad \frac{\Delta_2}{\Delta_1}, \quad \frac{\Delta_3}{\Delta_2}, \quad \dots, \quad \frac{\Delta_n}{\Delta_{n-1}}$$

A-5-21. Show that the Routh's stability criterion and Hurwitz stability criterion are equivalent.

Solution. If we write Hurwitz determinants in the triangular form

$$\Delta_i = \begin{vmatrix} a_{11} & & & * \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{ii} \end{vmatrix}, \quad (i = 1, 2, \dots, n)$$

where the elements below the diagonal line are all zeros and the elements above the diagonal line any numbers, then the Hurwitz conditions for asymptotic stability become

$$\Delta_i = a_{11}a_{22}\cdots a_{ii} > 0, \quad (i = 1, 2, \dots, n)$$

which are equivalent to the conditions

$$a_{11} > 0, \quad a_{22} > 0, \quad \dots, \quad a_{nn} > 0$$

We shall show that these conditions are equivalent to

$$a_1 > 0, \quad b_1 > 0, \quad c_1 > 0, \quad \dots$$

where $a_1, b_1, c_1, \dots$, are the elements of the first column in the Routh array.

Consider, for example, the following Hurwitz determinant, which corresponds to $i = 4$:

$$\Delta_4 = \begin{vmatrix} a_1 & a_3 & a_5 & a_7 \\ a_0 & a_2 & a_4 & a_6 \\ 0 & a_1 & a_3 & a_5 \\ 0 & a_0 & a_2 & a_4 \end{vmatrix}$$

The determinant is unchanged if we subtract from the i th row k times the j th row. By subtracting from the second row a_0/a_1 times the first row, we obtain

$$\Delta_4 = \begin{vmatrix} a_{11} & a_3 & a_5 & a_7 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & a_1 & a_3 & a_5 \\ 0 & a_0 & a_2 & a_4 \end{vmatrix}$$

Example Problems and Solutions

319

where

$$\begin{aligned} a_{11} &= a_1 \\ a_{22} &= a_2 - \frac{a_0}{a_1} a_3 \\ a_{23} &= a_4 - \frac{a_0}{a_1} a_5 \\ a_{24} &= a_6 - \frac{a_0}{a_1} a_7 \end{aligned}$$

Similarly, subtracting from the fourth row a_0/a_1 times the third row yields

$$\Delta_4 = \begin{vmatrix} a_{11} & a_3 & a_5 & a_7 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & a_1 & a_3 & a_5 \\ 0 & 0 & \hat{a}_{43} & \hat{a}_{44} \end{vmatrix}$$

where

$$\begin{aligned} \hat{a}_{43} &= a_2 - \frac{a_0}{a_1} a_3 \\ \hat{a}_{44} &= a_4 - \frac{a_0}{a_1} a_5 \end{aligned}$$

Next, subtracting from the third row a_1/a_{22} times the second row yields

$$\Delta_4 = \begin{vmatrix} a_{11} & a_3 & a_5 & a_7 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & \hat{a}_{43} & \hat{a}_{44} \end{vmatrix}$$

where

$$\begin{aligned} a_{33} &= a_3 - \frac{a_1}{a_{22}} a_{23} \\ a_{34} &= a_5 - \frac{a_1}{a_{22}} a_{24} \end{aligned}$$

Finally, subtracting from the last row $\hat{a}_{43}/a_{33}$ times the third row yields

$$\Delta_4 = \begin{vmatrix} a_{11} & a_3 & a_5 & a_7 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{vmatrix}$$

where

$$a_{44} = \hat{a}_{44} - \frac{\hat{a}_{43}}{a_{33}} a_{34}$$

From this analysis, we see that

$$\Delta_4 = a_{11}a_{22}a_{33}a_{44}$$

$$\Delta_3 = a_{11}a_{22}a_{33}$$

$$\Delta_2 = a_{11}a_{22}$$

$$\Delta_1 = a_{11}$$

The Hurwitz conditions for asymptotic stability

$$\Delta_1 > 0, \quad \Delta_2 > 0, \quad \Delta_3 > 0, \quad \Delta_4 > 0, \quad \dots$$

reduce to the conditions

$$a_{11} > 0, \quad a_{22} > 0, \quad a_{33} > 0, \quad a_{44} > 0, \quad \dots$$

The Routh array for the polynomial

$$a_0s^4 + a_1s^3 + a_2s^2 + a_3s + a_4 = 0$$

where $a_0 > 0$ and $n = 4$, is given by

$$\begin{array}{ccc} a_0 & a_2 & a_4 \\ a_1 & a_3 & \\ b_1 & b_2 & \\ c_1 & & \\ d_1 & & \end{array}$$

From this Routh array, we see that

$$a_{11} = a_1$$

$$a_{22} = a_2 - \frac{a_0}{a_1}a_3 = b_1$$

$$a_{33} = a_3 - \frac{a_1}{a_{22}}a_{23} = \frac{a_3b_1 - a_1b_2}{b_1} = c_1$$

$$a_{44} = \hat{a}_{44} - \frac{\hat{a}_{43}}{a_{33}}a_{34} = a_4 = d_1$$

(The last equation is obtained using the fact that $a_{34} = 0$, $\hat{a}_{44} = a_4$, and $a_4 = b_2 = d_1$.) Hence the Hurwitz conditions for asymptotic stability become

$$a_1 > 0, \quad b_1 > 0, \quad c_1 > 0, \quad d_1 > 0$$

Thus we have demonstrated that Hurwitz conditions for asymptotic stability can be reduced to Routh's conditions for asymptotic stability. The same argument can be extended to Hurwitz determinants of any order, and the equivalence of Routh's stability criterion and Hurwitz stability criterion can be established.

A-5-22. Consider the characteristic equation

$$s^4 + 2s^3 + (4 + K)s^2 + 9s + 25 = 0$$

Using the Hurwitz stability criterion, determine the range of K for stability.

Solution. Comparing the given characteristic equation

$$s^4 + 2s^3 + (4 + K)s^2 + 9s + 25 = 0$$

Example Problems and Solutions

321

with the following standard fourth-order characteristic equation:

$$a_0 s^4 + a_1 s^3 + a_2 s^2 + a_3 s + a_4 = 0$$

we find

$$a_0 = 1, \quad a_1 = 2, \quad a_2 = 4 + K, \quad a_3 = 9, \quad a_4 = 25$$

The Hurwitz stability criterion states that Δ_4 is given by

$$\Delta_4 = \begin{vmatrix} a_1 & a_3 & 0 & 0 \\ a_0 & a_2 & a_4 & 0 \\ 0 & a_1 & a_3 & 0 \\ 0 & a_0 & a_2 & a_4 \end{vmatrix}$$

For all the roots to have negative real parts, it is necessary and sufficient that successive principal minors of Δ_4 be positive. The successive principal minors are

$$\Delta_1 = |a_1| = 2$$

$$\Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} = \begin{vmatrix} 2 & 9 \\ 1 & 4 + K \end{vmatrix} = 2K - 1$$

$$\Delta_3 = \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} = \begin{vmatrix} 2 & 9 & 0 \\ 1 & 4 + K & 25 \\ 0 & 2 & 9 \end{vmatrix} = 18K - 109$$

For all principal minors to be positive, we require that $\Delta_i (i = 1, 2, 3)$ be positive. Thus, we require

$$2K - 1 > 0$$

$$18K - 109 > 0$$

from which we obtain the region of K for stability to be

$$K > \frac{109}{18}$$

- A-5-23.** Explain why the proportional control of a plant that does not possess an integrating property (which means that the plant transfer function does not include the factor $1/s$) suffers offset in response to step inputs.

Solution. Consider, for example, the system shown in Figure 5-70. At steady state, if c were equal to a nonzero constant r , then $e = 0$ and $u = Ke = 0$, resulting in $c = 0$, which contradicts the assumption that $c = r = \text{nonzero constant}$.

A nonzero offset must exist for proper operation of such a control system. In other words, at steady state, if e were equal to $r/(1 + K)$, then $u = Kr/(1 + K)$ and $c = Kr/(1 + K)$, which results in the assumed error signal $e = r/(1 + K)$. Thus the offset of $r/(1 + K)$ must exist in such a system.

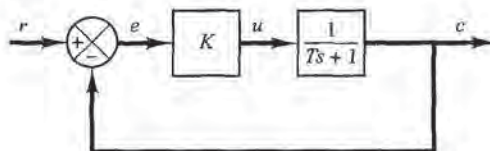


Figure 5-70
Control system.

- A-5-24.** Consider the system shown in Figure 5-71. Show that the steady-state error in following the unit-ramp input is B/K . This error can be made smaller by choosing B small and/or K large. However, making B small and/or K large would have the effect of making the damping ratio small, which is normally not desirable. Describe a method or methods to make B/K small and yet make the damping ratio have reasonable value ($0.5 < \zeta < 0.7$).

Solution. From Figure 5-71 we obtain

$$E(s) = R(s) - C(s) = \frac{Js^2 + Bs}{Js^2 + Bs + K} R(s)$$

The steady-state error for the unit-ramp response can be obtained as follows: For the unit-ramp input, the steady-state error e_{ss} is

$$\begin{aligned} e_{ss} &= \lim_{s \rightarrow 0} sE(s) \\ &= \lim_{s \rightarrow 0} s \frac{Js^2 + Bs}{Js^2 + Bs + K} \frac{1}{s^2} \\ &= \frac{B}{K} \\ &= \frac{2\zeta}{\omega_n} \end{aligned}$$

where

$$\zeta = \frac{B}{2\sqrt{KJ}}, \quad \omega_n = \sqrt{\frac{K}{J}}$$

To assure acceptable transient response and acceptable steady-state error in following a ramp input, ζ must not be too small and ω_n must be sufficiently large. It is possible to make the steady-state error e_{ss} small by making the value of the gain K large. (A large value of K has an additional advantage of suppressing undesirable effects caused by dead zone, backlash, coulomb friction, and the like.) A large value of K would, however, make the value of ζ small and increase the maximum overshoot, which is undesirable.

It is therefore necessary to compromise between the magnitude of the steady-state error to a ramp input and the maximum overshoot to a unit-step input. In the system shown in Figure 5-71, a reasonable compromise may not be reached easily. It is then desirable to consider other types of control action that may improve both the transient-response and steady-state behavior. Two schemes to improve both the transient-response and steady-state behavior are available. One scheme is to use a proportional-plus-derivative controller and the other is to use tachometer feedback.

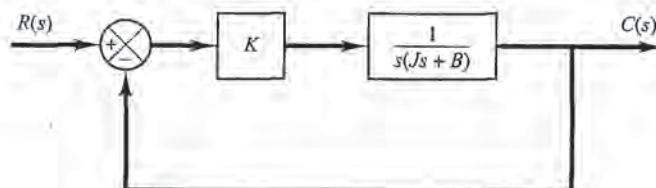
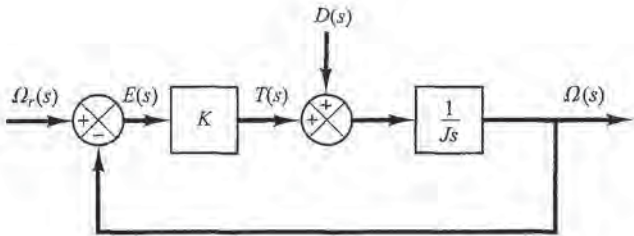


Figure 5-71
Control system.

- A-5-25.** The block diagram of Figure 5-72 shows a speed control system in which the output member of the system is subject to a torque disturbance. In the diagram, $\Omega_r(s)$, $\Omega(s)$, $T(s)$, and $D(s)$ are the Laplace transforms of the reference speed, output speed, driving torque, and disturbance torque, respectively. In the absence of a disturbance torque, the output speed is equal to the reference speed.

Figure 5-72
Block diagram of a speed control system.



Investigate the response of this system to a unit-step disturbance torque. Assume that the reference input is zero, or $\Omega_r(s) = 0$.

Solution. Figure 5-73 is a modified block diagram convenient for the present analysis. The closed-loop transfer function is

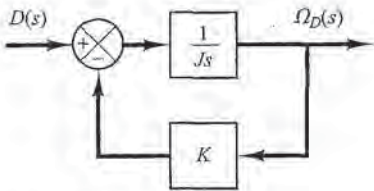
$$\frac{\Omega_D(s)}{D(s)} = \frac{1}{Js + K}$$

where $\Omega_D(s)$ is the Laplace transform of the output speed due to the disturbance torque. For a unit-step disturbance torque, the steady-state output velocity is

$$\begin{aligned} \omega_D(\infty) &= \lim_{s \rightarrow 0} s \Omega_D(s) \\ &= \lim_{s \rightarrow 0} \frac{s}{Js + K} \frac{1}{s} \\ &= \frac{1}{K} \end{aligned}$$

From this analysis, we conclude that, if a step disturbance torque is applied to the output member of the system, an error speed will result so that the ensuing motor torque will exactly cancel the disturbance torque. To develop this motor torque, it is necessary that there be an error in speed so that nonzero torque will result.

Figure 5-73
Block diagram of the speed control system of Figure 5-72 when $\Omega_r(s) = 0$.



A-5-26. In the system considered in Problem A-5-25, it is desired to eliminate as much as possible the speed errors due to torque disturbances.

Is it possible to cancel the effect of a disturbance torque at steady state so that a constant disturbance torque applied to the output member will cause no speed change at steady state?

Solution. Suppose that we choose a suitable controller whose transfer function is $G_c(s)$, as shown in Figure 5-74. Then in the absence of the reference input the closed-loop transfer function between the output velocity $\Omega_D(s)$ and the disturbance torque $D(s)$ is

$$\begin{aligned}\frac{\Omega_D(s)}{D(s)} &= \frac{\frac{1}{Js}}{1 + \frac{1}{Js} G_c(s)} \\ &= \frac{1}{Js + G_c(s)}\end{aligned}$$

The steady-state output speed due to a unit-step disturbance torque is

$$\begin{aligned}\omega_D(\infty) &= \lim_{s \rightarrow 0} s \Omega_D(s) \\ &= \lim_{s \rightarrow 0} \frac{s}{Js + G_c(s)} \frac{1}{s} \\ &= \frac{1}{G_c(0)}\end{aligned}$$

To satisfy the requirement that

$$\omega_D(\infty) = 0$$

we must choose $G_c(0) = \infty$. This can be realized if we choose

$$G_c(s) = \frac{K}{s}$$

Integral control action will continue to correct until the error is zero. This controller, however, presents a stability problem because the characteristic equation will have two imaginary roots.

One method of stabilizing such a system is to add a proportional mode to the controller or choose

$$G_c(s) = K_p + \frac{K}{s}$$

Figure 5-74
Block diagram of a speed control system.

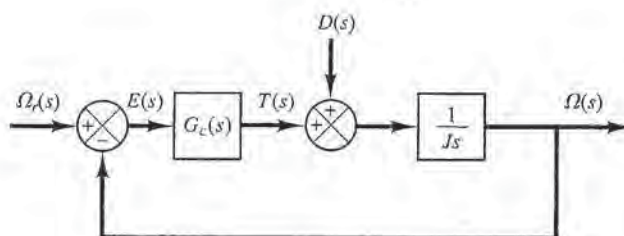
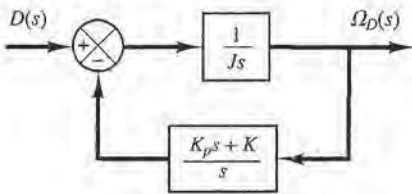


Figure 5-75
Block diagram of the speed control system of Figure 5-74 when $G_c(s) = K_p + (K/s)$ and $\Omega_r(s) = 0$.



With this controller, the block diagram of Figure 5-74 in the absence of the reference input can be modified to that of Figure 5-75. The closed-loop transfer function $\Omega_D(s)/D(s)$ becomes

$$\frac{\Omega_D(s)}{D(s)} = \frac{s}{Js^2 + K_p s + K}$$

For a unit-step disturbance torque, the steady-state output speed is

$$\omega_D(\infty) = \lim_{s \rightarrow 0} s \Omega_D(s) = \lim_{s \rightarrow 0} \frac{s^2}{Js^2 + K_p s + K} \frac{1}{s} = 0$$

Thus, we see that the proportional-plus-integral controller eliminates speed error at steady state. The use of integral control action has increased the order of the system by 1. (This tends to produce an oscillatory response.)

In the present system, a step disturbance torque will cause a transient error in the output speed, but the error will become zero at steady state. The integrator provides a nonzero output with zero error. (The nonzero output of the integrator produces a motor torque that exactly cancels the disturbance torque.)

Note that the integrator in the transfer function of the plant does not eliminate the steady-state error due to a step disturbance torque. To eliminate this, we must have an integrator before the point where the disturbance torque enters.

- A-5-27.** Consider the system shown in Figure 5-76(a). The steady-state error to a unit-ramp input is $e_{ss} = 2\zeta/\omega_n$. Show that the steady-state error for following a ramp input may be eliminated if the input is introduced to the system through a proportional-plus-derivative filter, as shown in Figure 5-76(b), and the value of k is properly set. Note that the error $e(t)$ is given by $r(t) - c(t)$.

Solution. The closed-loop transfer function of the system shown in Figure 5-76(b) is

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

Then

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

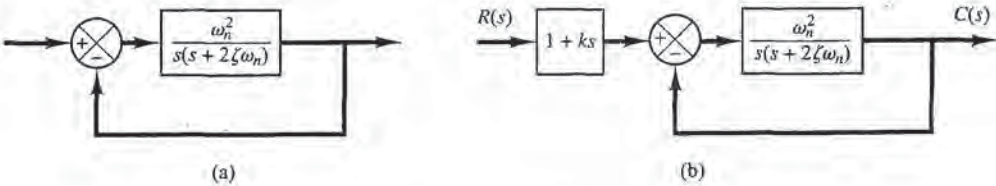


Figure 5-76
(a) Control system;
(b) control system with input filter.

If the input is a unit ramp, then the steady-state error is

$$\begin{aligned} e(\infty) &= r(\infty) - c(\infty) \\ &= \lim_{s \rightarrow 0} s \left(\frac{s^2 + 2\zeta\omega_n s - \omega_n^2 k s}{s^2 + 2\zeta\omega_n s + \omega_n^2} \right) \frac{1}{s^2} \\ &= \frac{2\zeta\omega_n - \omega_n^2 k}{\omega_n^2} \end{aligned}$$

Therefore, if k is chosen as

$$k = \frac{2\zeta}{\omega_n}$$

then the steady-state error for following a ramp input can be made equal to zero. Note that, if there are any variations in the values of ζ and/or ω_n due to environmental changes or aging, then a nonzero steady-state error for a ramp response may result.

- A-5-28.** Consider the servo system with tachometer feedback shown in Figure 5-77. Obtain the error signal $E(s)$ when both the reference input $R(s)$ and disturbance input $D(s)$ are present. Obtain also the steady-state error when the system is subjected to a reference input (unit-ramp input) and disturbance input (step input of magnitude d).

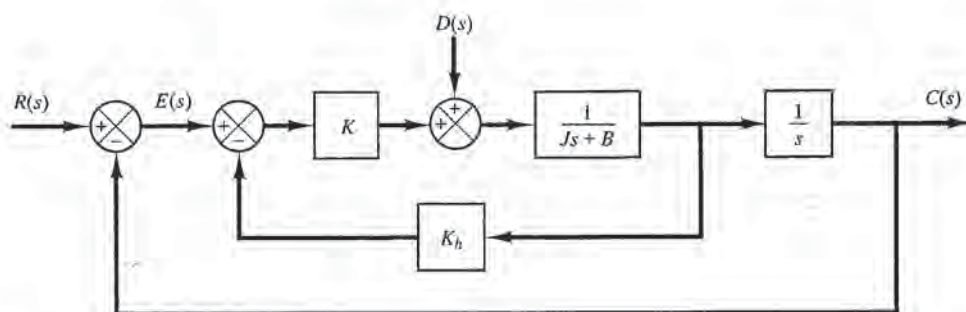


Figure 5-77
Servo system with
tachometer
feedback.

Solution. When we consider the reference input $R(s)$ we can assume that the disturbance input $D(s)$ is zero, and vice versa. Then, a block diagram that relates the reference input $R(s)$ and the output $C(s)$ may be drawn as shown in Figure 5-78(a). Similarly, Figure 5-78(b) relates the disturbance input $D(s)$ and the output $C(s)$.

The closed-loop transfer function $C(s)/R(s)$ can be obtained from Figure 5-78(a) as follows:

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

Similarly, the closed-loop transfer function $C(s)/D(s)$ can be obtained from Figure 5-78(b) as

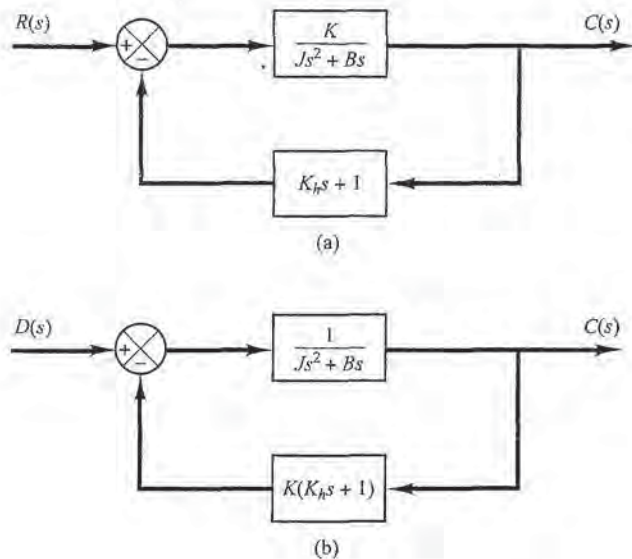
$$\frac{C(s)}{D(s)} = \frac{1}{Js^2 + (B + KK_h)s + K}$$

If both $R(s)$ and $D(s)$ are present, then

$$C(s) = \frac{1}{Js^2 + (B + KK_h)s + K} [KR(s) + D(s)]$$

Figure 5-78

(a) Block diagram that relates reference input $R(s)$ and output $C(s)$;
(b) block diagram that relates disturbance input $D(s)$ and output $C(s)$.



Since

$$\begin{aligned} E(s) &= R(s) - C(s) \\ &= R(s) - \frac{KR(s) + D(s)}{Js^2 + (B + KK_h)s + K} \end{aligned}$$

we obtain

$$E(s) = \frac{1}{Js^2 + (B + KK_h)s + K} \{ [Js^2 + (B + KK_h)s]R(s) - D(s) \}$$

Hence the steady-state error can be obtained as follows:

$$\begin{aligned} \lim_{t \rightarrow \infty} e(t) &= \lim_{s \rightarrow 0} sE(s) \\ &= \lim_{s \rightarrow 0} \frac{s}{Js^2 + (B + KK_h)s + K} [s(Js + B + KK_h)R(s) - D(s)] \\ &= \lim_{s \rightarrow 0} \left[\frac{B + KK_h}{K} s^2 R(s) - \frac{1}{K} sD(s) \right] \end{aligned}$$

Since $R(s) = 1/s^2$ (unit-ramp input) and $D(s) = d/s$ (step input of magnitude d) the steady-state error is

$$\begin{aligned} \lim_{t \rightarrow \infty} e(t) &= \lim_{s \rightarrow 0} \left[\frac{B + KK_h}{K} \frac{s^2}{s^2} - \frac{1}{K} \frac{sd}{s} \right] \\ &= \frac{B + KK_h}{K} - \frac{d}{K} \end{aligned}$$

- A-5-29.** Consider the stable unity-feedback control system with feedforward transfer function $G(s)$. Suppose that the closed-loop transfer function can be written

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)} = \frac{(T_a s + 1)(T_b s + 1) \cdots (T_m s + 1)}{(T_1 s + 1)(T_2 s + 1) \cdots (T_n s + 1)} \quad (m \leq n)$$

Show that

$$\int_0^\infty e(t) dt = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

where $e(t)$ is the error in the unit-step response. Show also that

$$\frac{1}{K_v} = \frac{1}{\lim_{s \rightarrow 0} sG(s)} = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

Solution. Let us define

$$(T_a s + 1)(T_b s + 1) \cdots (T_m s + 1) = P(s)$$

and

$$(T_1 s + 1)(T_2 s + 1) \cdots (T_n s + 1) = Q(s)$$

Then

$$\frac{C(s)}{R(s)} = \frac{P(s)}{Q(s)}$$

and

$$E(s) = \frac{Q(s) - P(s)}{Q(s)} R(s)$$

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

$$E(s) = \frac{Q(s) - P(s)}{sQ(s)}$$

Since the system is stable, $\int_0^\infty e(t) dt$ converges to a constant value. Referring to Table 2-2 (item 9), we have

$$\int_0^\infty e(t) dt = \lim_{s \rightarrow 0} s \frac{E(s)}{s} = \lim_{s \rightarrow 0} E(s)$$

Hence

$$\begin{aligned} \int_0^\infty e(t) dt &= \lim_{s \rightarrow 0} \frac{Q(s) - P(s)}{sQ(s)} \\ &= \lim_{s \rightarrow 0} \frac{Q'(s) - P'(s)}{Q(s) + sQ'(s)} \\ &= \lim_{s \rightarrow 0} [Q'(s) - P'(s)] \end{aligned}$$

Since

$$\lim_{s \rightarrow 0} P'(s) = T_a + T_b + \cdots + T_m$$

$$\lim_{s \rightarrow 0} Q'(s) = T_1 + T_2 + \cdots + T_n$$

we have

$$\int_0^\infty e(t) dt = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

For a unit-step input $r(t)$, since

$$\int_0^\infty e(t) dt = \lim_{s \rightarrow 0} E(s) = \lim_{s \rightarrow 0} \frac{1}{1 + G(s)} R(s) = \lim_{s \rightarrow 0} \frac{1}{1 + G(s)} \frac{1}{s} = \frac{1}{\lim_{s \rightarrow 0} sG(s)} = \frac{1}{K_v}$$

we have

$$\frac{1}{K_v} = \frac{1}{\lim_{s \rightarrow 0} sG(s)} = (T_1 + T_2 + \cdots + T_n) - (T_a + T_b + \cdots + T_m)$$

Note that zeros in the left half-plane (that is, positive $T_a, T_b, \dots, T_m$) will improve K_v . Poles close to the origin cause low velocity-error constants unless there are zeros nearby.

PROBLEMS

B-5-1. A thermometer requires 1 min to indicate 98% of the response to a step input. Assuming the thermometer to be a first-order system, find the time constant.

If the thermometer is placed in a bath, the temperature of which is changing linearly at a rate of $10^\circ/\text{min}$, how much error does the thermometer show?

B-5-2. Consider the unit-step response of a unity-feedback control system whose open-loop transfer function is

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

Obtain the rise time, peak time, maximum overshoot, and settling time.

B-5-3. Consider the closed-loop system given by

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

Determine the values of ζ and ω_n so that the system responds to a step input with approximately 5% overshoot and with a settling time of 2 sec. (Use the 2% criterion.)

B-5-4. Figure 5-79 is a block diagram of a space-vehicle attitude-control system. Assuming the time constant T of the controller to be 3 sec and the ratio K/J to be $\frac{2}{3} \text{ rad}^2/\text{sec}^2$, find the damping ratio of the system.

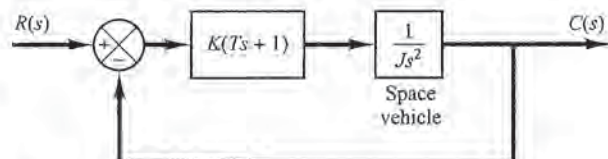


Figure 5-79
Space-vehicle attitude-control system.

B-5-5. Consider the system shown in Figure 5-80. The system is initially at rest. Suppose that the cart is set into motion by an impulsive force whose strength is unity. Can it be stopped by another such impulsive force?

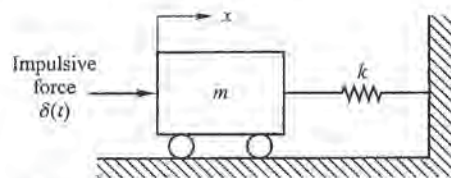


Figure 5-80
Mechanical system.

B-5-6. Obtain the unit-impulse response and the unit-step response of a unity-feedback system whose open-loop transfer function is

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

B-5-7. Consider the system shown in Figure 5-81. Show that the transfer function $Y(s)/X(s)$ has a zero in the right-half s plane. Then obtain $y(t)$ when $x(t)$ is a unit step. Plot $y(t)$ versus t .

B-5-8. An oscillatory system is known to have a transfer function of the following form:

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

Assume that a record of a damped oscillation is available as shown in Figure 5-82. Determine the damping ratio ζ of the system from the graph.

B-5-9. Consider the system shown in Figure 5-83(a). The damping ratio of this system is 0.158 and the undamped natural frequency is 3.16 rad/sec. To improve the relative stability, we employ tachometer feedback. Figure 5-83(b) shows such a tachometer-feedback system.

Determine the value of K_h so that the damping ratio of the system is 0.5. Draw unit-step response curves of both the original and tachometer-feedback systems. Also draw the error-versus-time curves for the unit-ramp response of both systems.

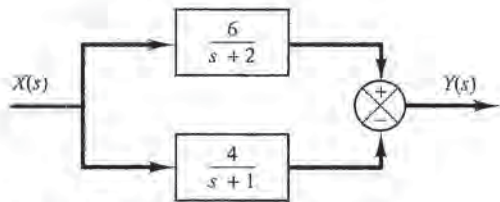


Figure 5-81
System with zero in the right-half s plane.

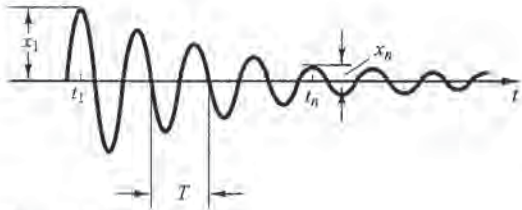


Figure 5-82
Decaying oscillation.

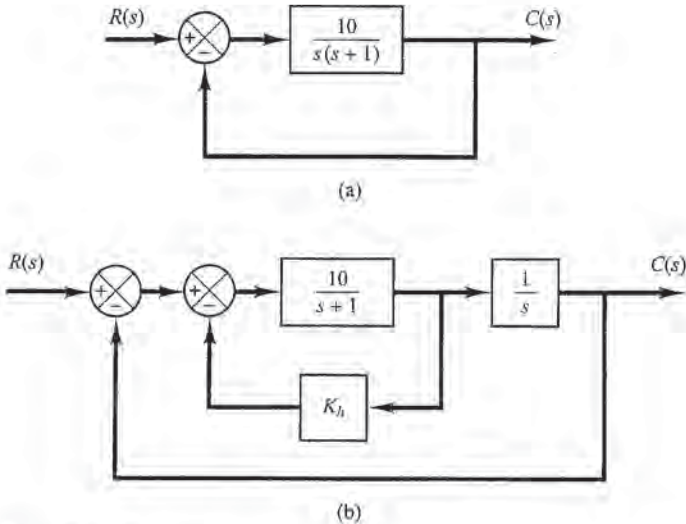


Figure 5-83
(a) Control system; (b) control system with tachometer feedback.

B-5-10. Referring to the system shown in Figure 5-84, determine the values of K and k such that the system has a damping ratio ζ of 0.7 and an undamped natural frequency ω_n of 4 rad/sec.

B-5-11. Consider the system shown in Figure 5-85. Determine the value of k such that the damping ratio ζ is 0.5. Then obtain the rise time t_r , peak time t_p , maximum overshoot M_p , and settling time t_s in the unit-step response.

B-5-12. Using MATLAB, obtain the unit-step response, unit-ramp response, and unit-impulse response of the following system:

$$\frac{C(s)}{R(s)} = \frac{10}{s^2 + 2s + 10}$$

where $R(s)$ and $C(s)$ are Laplace transforms of the input $r(t)$ and output $c(t)$, respectively.

B-5-13. Using MATLAB, obtain the unit-step response, unit-ramp response, and unit-impulse response of the following system:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & -0.5 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} u$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

where u is the input and y is the output.

B-5-14. Obtain both analytically and computationally the rise time, peak time, maximum overshoot, and settling time in the unit-step response of a closed-loop system given by

$$\frac{C(s)}{R(s)} = \frac{36}{s^2 + 2s + 36}$$

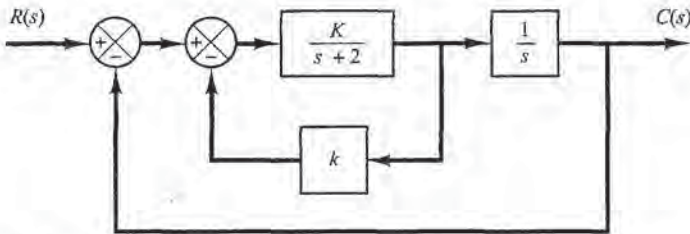


Figure 5-84
Closed-loop system.

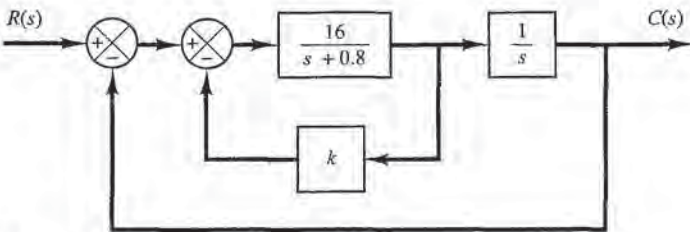


Figure 5-85
Block diagram of a system.

B-5-15. Figure 5-86 shows three systems. System I is a positional servo system. System II is a positional servo system with PD control action. System III is a positional servo system with velocity feedback. Compare the unit-step, unit-impulse, and unit-ramp responses of the three systems. Which system is best with respect to the speed of response and maximum overshoot in the step response?

B-5-16. Consider the position control system shown in Figure 5-87. Write a MATLAB program to obtain a unit-step response and a unit-ramp response of the system. Plot curves $x_1(t)$ versus t , $x_2(t)$ versus t , $x_3(t)$ versus t , and $e(t)$ versus t [where $e(t) = r(t) - x_1(t)$] for both the unit-step response and the unit-ramp response.

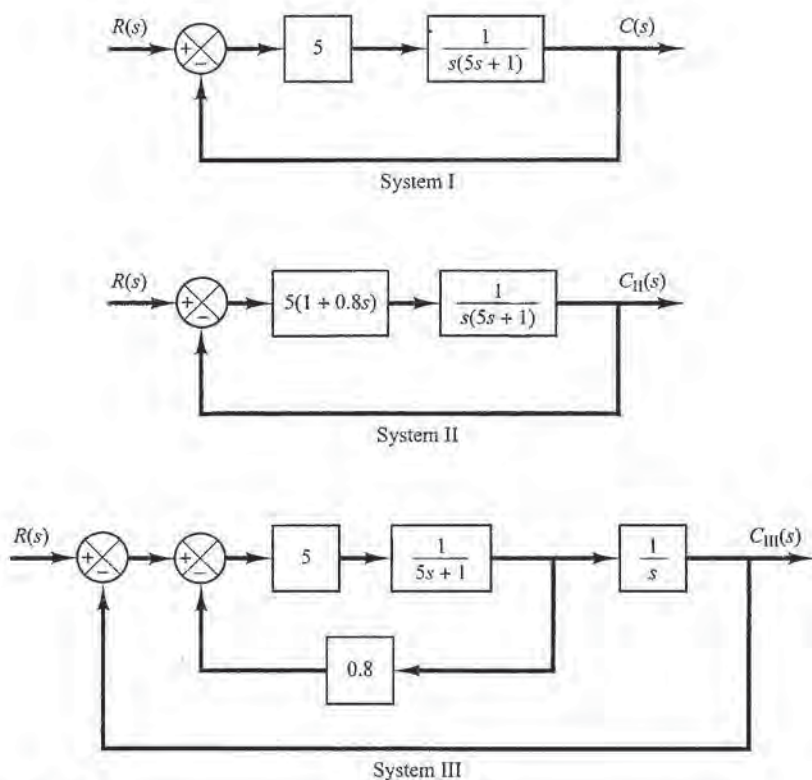


Figure 5-86
Positional servo system (System I), positional servo system with PD control action (System II), and positional servo system with velocity feedback (System III).

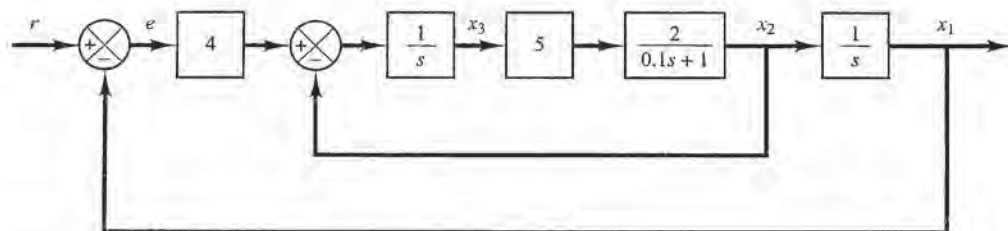


Figure 5-87
Position control system.

B-5-17. Using MATLAB, obtain the unit-step response curve for the unity-feedback control system whose open-loop transfer function is

$$G(s) = \frac{10}{s(s+2)(s+4)}$$

Using MATLAB, obtain also the rise time, peak time, maximum overshoot, and settling time in the unit-step response curve.

B-5-18. Consider the closed-loop system defined by

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

where $\zeta = 0.2, 0.4, 0.6, 0.8$, and 1.0 . Using MATLAB, plot a two-dimensional diagram of unit-impulse response curves. Also plot a three-dimensional plot of the response curves.

B-5-19. Consider the second-order system defined by

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

where $\zeta = 0.2, 0.4, 0.6, 0.8, 1.0$. Plot a three-dimensional diagram of the unit-step response curves.

B-5-20. Obtain the unit-ramp response of the system defined by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

where u is the unit-ramp input. Use `lsim` command to obtain the response.

B-5-21. Using MATLAB obtain the unit acceleration response curve of the unity-feedback control system whose open-loop transfer function is

$$G(s) = \frac{10(s+1)}{s^2(s+4)}$$

The unit acceleration input is defined by

$$r(t) = \frac{1}{2}t^2 \quad (t \geq 0)$$

B-5-22. Consider the differential equation system given by

$$\ddot{y} + 3\dot{y} + 2y = 0, \quad y(0) = 0.1, \quad \dot{y}(0) = 0.05$$

Obtain the response $y(t)$, subject to the given initial condition.

B-5-23. Determine the range of K for stability of a unity-feedback control system whose open-loop transfer function is

$$G(s) = \frac{K}{s(s+1)(s+2)}$$

B-5-24. Consider the unity-feedback control system with the following open-loop transfer function:

$$G(s) = \frac{10}{s(s-1)(2s+3)}$$

Is this system stable?

B-5-25. Consider the following characteristic equation:

$$s^4 + 2s^3 + (4+K)s^2 + 9s + 25 = 0$$

Using Routh stability criterion, determine the range of K for stability.

B-5-26. Consider the closed-loop system shown in Figure 5-88. Determine the range of K for stability. Assume that $K > 0$.

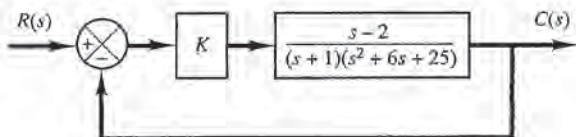


Figure 5-88
Closed-loop system.

B-5-27. Consider the satellite attitude control system shown in Figure 5-89(a). The output of this system exhibits continued oscillations and is not desirable. This system can be stabilized by use of tachometer feedback, as shown in Figure 5-89(b). If $K/J = 4$, what value of K_h will yield the damping ratio to be 0.6?

B-5-28. Consider the servo system with tachometer feedback shown in Figure 5-90. Determine the ranges of stability for K and K_h . (Note that K_h must be positive.)

B-5-29. Consider the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ where matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ -b_3 & 0 & 1 \\ 0 & -b_2 & -b_1 \end{bmatrix}$$

($\mathbf{A}$ is called Schwarz matrix.) Show that the first column of the Routh's array of the characteristic equation $|s\mathbf{I} - \mathbf{A}| = 0$ consists of 1, b_1 , b_2 , and b_1b_3 .

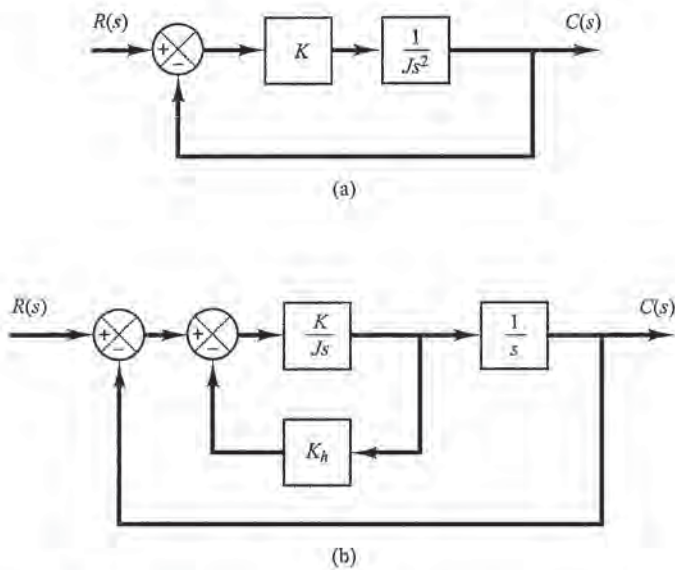


Figure 5-89
(a) Unstable satellite attitude control system; (b) stabilized system.

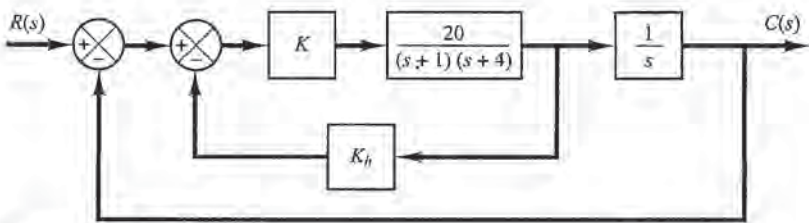


Figure 5-90
Servo system with tachometer feedback.

B-5-30. Consider a unity-feedback control system with the closed-loop transfer function

$$\frac{C(s)}{R(s)} = \frac{Ks + b}{s^2 + as + b}$$

Determine the open-loop transfer function $G(s)$.

Show that the steady-state error in the unit-ramp response is given by

$$e_{ss} = \frac{1}{K_v} = \frac{a - K}{b}$$

B-5-31. Consider a unity-feedback control system whose open-loop transfer function is

$$G(s) = \frac{K}{s(Js + B)}$$

Discuss the effects that varying the values of K and B has on the steady-state error in unit-ramp response. Sketch typical unit-ramp response curves for a small value, medium value, and large value of K , assuming that B is constant.

B-5-32. If the feedforward path of a control system contains at least one integrating element, then the output continues to change as long as an error is present. The output stops when the error is precisely zero. If an external disturbance enters the system, it is desirable to have an integrating element between the error-measuring element and the point where the disturbance enters so that the effect of the external disturbance may be made zero at steady state.

Show that, if the disturbance is a ramp function, then the steady-state error due to this ramp disturbance may be eliminated only if two integrators precede the point where the disturbance enters.



Root-Locus Analysis

6-1 INTRODUCTION

The basic characteristic of the transient response of a closed-loop system is closely related to the location of the closed-loop poles. If the system has a variable loop gain, then the location of the closed-loop poles depends on the value of the loop gain chosen. It is important, therefore, that the designer know how the closed-loop poles move in the s plane as the loop gain is varied.

From the design viewpoint, in some systems simple gain adjustment may move the closed-loop poles to desired locations. Then the design problem may become the selection of an appropriate gain value. If the gain adjustment alone does not yield a desired result, addition of a compensator to the system will become necessary. (This subject is discussed in detail in Chapter 7.)

The closed-loop poles are the roots of the characteristic equation. Finding the roots of the characteristic equation of degree higher than 3 is laborious and will need computer solution. (MATLAB provides a simple solution to this problem.) However, just finding the roots of the characteristic equation may be of limited value, because as the gain of the open-loop transfer function varies the characteristic equation changes and the computations must be repeated.

A simple method for finding the roots of the characteristic equation has been developed by W. R. Evans and used extensively in control engineering. This method, called the *root-locus method*, is one in which the roots of the characteristic equation are

plotted for all values of a system parameter. The roots corresponding to a particular value of this parameter can then be located on the resulting graph. Note that the parameter is usually the gain, but any other variable of the open-loop transfer function may be used. (See Chapter 7.) Unless otherwise stated, we shall assume that the gain of the open-loop transfer function is the parameter to be varied through all values, from zero to infinity.

By using the root-locus method the designer can predict the effects on the location of the closed-loop poles of varying the gain value or adding open-loop poles and/or open-loop zeros. Therefore, it is desired that the designer have a good understanding of the method for generating the root loci of the closed-loop system, both by hand and by use of a computer software like MATLAB.

Root-Locus Method. The basic idea behind the root-locus method is that the values of s that make the transfer function around the loop equal -1 must satisfy the characteristic equation of the system.

The root locus is the locus of roots of the characteristic equation of the closed-loop system as a specific parameter (usually, gain K) is varied from zero to infinity, giving the method its name. Such a plot clearly shows the contributions of each open-loop pole or zero to the locations of the closed-loop poles.

In designing a linear control system, we find that the root-locus method proves quite useful since it indicates the manner in which the open-loop poles and zeros should be modified so that the response meets system performance specifications. This method is particularly suited to obtaining approximate results very quickly.

Because generating the root loci by use of MATLAB is very simple, one may think sketching the root loci by hand is a waste of time and effort. However, experience in sketching the root loci by hand is invaluable for interpreting computer-generated root loci, as well as for getting a rough idea of the root loci very quickly.

By using the root-locus method, it is possible to determine the value of the loop gain K that will make the damping ratio of the dominant closed-loop poles as prescribed. If the location of an open-loop pole or zero is a system variable, then the root-locus method suggests the way to choose the location of an open-loop pole or zero.

Outline of the Chapter. This chapter introduces the basic concept of the root-locus method and presents useful rules for graphically constructing the root loci, as well as the generation of root loci with MATLAB.

The outline of the chapter is as follows: Section 6-1 has presented an introduction to the root-locus method. Section 6-2 details the concepts underlying the root-locus method and presents the general procedure for sketching root loci using illustrative examples. Section 6-3 summarizes general rules for constructing root loci. (If the designer follows the general rules for constructing the root loci, sketching the root loci for a given system will become a simple matter.) Section 6-4 discusses generating root-locus plots with MATLAB. Section 6-5 treats a special case when the closed-loop system has positive feedback. Section 6-6 treats conditionally stable systems. Finally, Section 6-7 extends the root-locus method to treat closed-loop systems with transport lag.

6-2 ROOT-LOCUS PLOTS

Angle and Magnitude Conditions. Consider the system shown in Figure 6-1. The closed-loop transfer function is

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \quad (6-1)$$

The characteristic equation for this closed-loop system is obtained by setting the denominator of the right-hand side of Equation (6-1) equal to zero. That is,

$$1 + G(s)H(s) = 0$$

or

$$G(s)H(s) = -1 \quad (6-2)$$

Here we assume that $G(s)H(s)$ is a ratio of polynomials in s . [Later, in Section 6-7, we extend the analysis to the case when $G(s)H(s)$ involves the transport lag e^{-Ts} .] Since $G(s)H(s)$ is a complex quantity, Equation (6-2) can be split into two equations by equating the angles and magnitudes of both sides, respectively, to obtain the following:

Angle condition:

$$\angle G(s)H(s) = \pm 180^\circ(2k + 1) \quad (k = 0, 1, 2, \dots) \quad (6-3)$$

Magnitude condition:

$$|G(s)H(s)| = 1 \quad (6-4)$$

The values of s that fulfill both the angle and magnitude conditions are the roots of the characteristic equation, or the closed-loop poles. A locus of the points in the complex plane satisfying the angle condition alone is the root locus. The roots of the characteristic equation (the closed-loop poles) corresponding to a given value of the gain can be determined from the magnitude condition. The details of applying the angle and magnitude conditions to obtain the closed-loop poles are presented later in this section.

In many cases, $G(s)H(s)$ involves a gain parameter K , and the characteristic equation may be written as

$$1 + \frac{K(s + z_1)(s + z_2) \cdots (s + z_m)}{(s + p_1)(s + p_2) \cdots (s + p_n)} = 0$$

Then the root loci for the system are the loci of the closed-loop poles as the gain K is varied from zero to infinity.

Note that to begin sketching the root loci of a system by the root-locus method we must know the location of the poles and zeros of $G(s)H(s)$. Remember that the angles

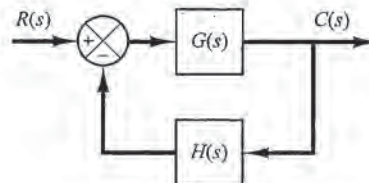
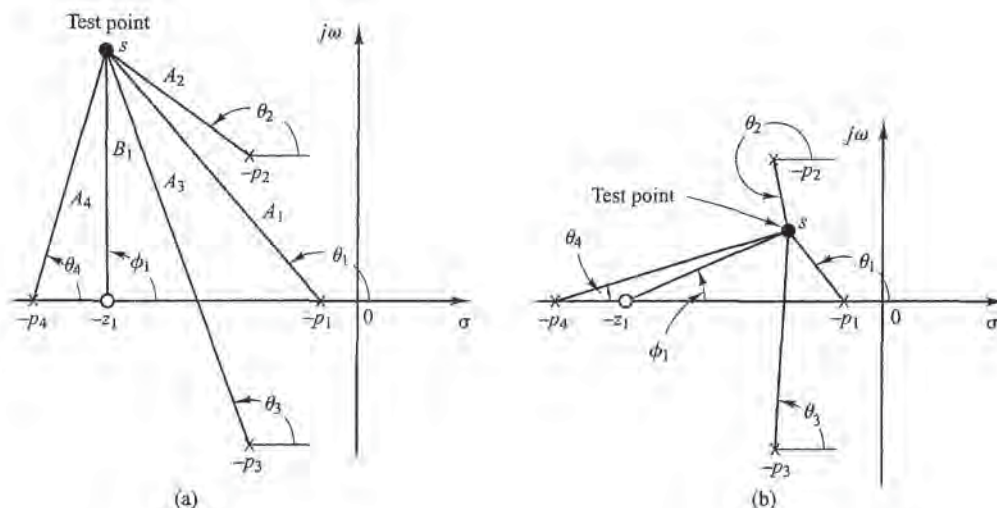


Figure 6-1
Control system.

Figure 6-2
(a) and (b) Diagrams showing angle measurements from open-loop poles and open-loop zero to test point s .



of the complex quantities originating from the open-loop poles and open-loop zeros to the test point s are measured in the counterclockwise direction. For example, if $G(s)H(s)$ is given by

$$G(s)H(s) = \frac{K(s + z_1)}{(s + p_1)(s + p_2)(s + p_3)(s + p_4)}$$

where $-p_2$ and $-p_3$ are complex-conjugate poles, then the angle of $G(s)H(s)$ is

$$\angle G(s)H(s) = \phi_1 - \theta_1 - \theta_2 - \theta_3 - \theta_4$$

where ϕ_1 , θ_1 , θ_2 , θ_3 , and θ_4 are measured counterclockwise as shown in Figures 6-2(a) and (b). The magnitude of $G(s)H(s)$ for this system is

$$|G(s)H(s)| = \frac{KB_1}{A_1A_2A_3A_4}$$

where A_1 , A_2 , A_3 , A_4 , and B_1 are the magnitudes of the complex quantities $s + p_1$, $s + p_2$, $s + p_3$, $s + p_4$, and $s + z_1$, respectively, as shown in Figure 6-2(a).

Note that, because the open-loop complex-conjugate poles and complex-conjugate zeros, if any, are always located symmetrically about the real axis, the root loci are always symmetrical with respect to this axis. Therefore, we only need to construct the upper half of the root loci and draw the mirror image of the upper half in the lower-half s plane.

Illustrative Examples. In what follows, two illustrative examples for constructing root-locus plots will be presented. Although computer approaches to the construction of the root loci are easily available, here we shall use graphical computation, combined with inspection, to determine the root loci upon which the roots of the characteristic equation of the closed-loop system must lie. Such a graphical approach will enhance understanding of how the closed-loop poles move in the complex plane as the open-loop poles and zeros are moved. Although we employ only simple systems for illustrative purposes, the procedure for finding the root loci is no more complicated for higher-order systems.

The first step in the procedure for constructing a root-locus plot is to seek out the loci of possible roots using the angle condition. Then, if necessary, the loci can be scaled, or graduated, in gain using the magnitude condition.

Because graphical measurements of angles and magnitudes are involved in the analysis, we find it necessary to use the same divisions on the abscissa as on the ordinate axis when sketching the root locus on graph paper.

EXAMPLE 6-1 Consider the system shown in Figure 6-3. (We assume that the value of gain K is nonnegative.) For this system,

$$G(s) = \frac{K}{s(s+1)(s+2)}, \quad H(s) = 1$$

Let us sketch the root-locus plot and then determine the value of K such that the damping ratio ζ of a pair of dominant complex-conjugate closed-loop poles is 0.5.

For the given system, the angle condition becomes

$$\begin{aligned} \angle G(s) &= \angle \frac{K}{s(s+1)(s+2)} \\ &= -\angle s - \angle s+1 - \angle s+2 \\ &= \pm 180^\circ(2k+1) \quad (k = 0, 1, 2, \dots) \end{aligned}$$

The magnitude condition is

$$|G(s)| = \left| \frac{K}{s(s+1)(s+2)} \right| = 1$$

A typical procedure for sketching the root-locus plot is as follows:

1. Determine the root loci on the real axis. The first step in constructing a root-locus plot is to locate the open-loop poles, $s = 0$, $s = -1$, and $s = -2$, in the complex plane. (There are no open-loop zeros in this system.) The locations of the open-loop poles are indicated by crosses. (The locations of the open-loop zeros in this book will be indicated by small circles.) Note that the starting points of the root loci (the points corresponding to $K = 0$) are open-loop poles. The number of individual root loci for this system is three, which is the same as the number of open-loop poles.

To determine the root loci on the real axis, we select a test point, s . If the test point is on the positive real axis, then

$$\angle s = \angle s+1 = \angle s+2 = 0^\circ$$

This shows that the angle condition cannot be satisfied. Hence, there is no root locus on the positive real axis. Next, select a test point on the negative real axis between 0 and -1 . Then

$$\angle s = 180^\circ, \quad \angle s+1 = \angle s+2 = 0^\circ$$

Thus

$$-\angle s - \angle s+1 - \angle s+2 = -180^\circ$$

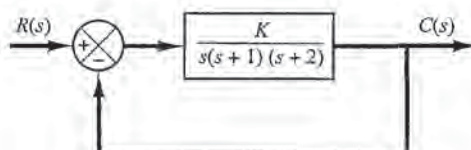


Figure 6-3
Control system.

and the angle condition is satisfied. Therefore, the portion of the negative real axis between 0 and -1 forms a portion of the root locus. If a test point is selected between -1 and -2, then

$$\angle s = \angle s + 1 = 180^\circ, \quad \angle s + 2 = 0^\circ$$

and

$$-\angle s - \angle s + 1 - \angle s + 2 = -360^\circ$$

It can be seen that the angle condition is not satisfied. Therefore, the negative real axis from -1 to -2 is not a part of the root locus. Similarly, if a test point is located on the negative real axis from -2 to $-\infty$, the angle condition is satisfied. Thus, root loci exist on the negative real axis between 0 and -1 and between -2 and $-\infty$.

2. *Determine the asymptotes of the root loci.* The asymptotes of the root loci as s approaches infinity can be determined as follows: If a test point s is selected very far from the origin, then

$$\lim_{s \rightarrow \infty} G(s) = \lim_{s \rightarrow \infty} \frac{K}{s(s+1)(s+2)} = \lim_{s \rightarrow \infty} \frac{K}{s^3}$$

and the angle condition becomes

$$-3\angle s = \pm 180^\circ(2k+1) \quad (k = 0, 1, 2, \dots)$$

or

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k+1)}{3} \quad (k = 0, 1, 2, \dots)$$

Since the angle repeats itself as k is varied, the distinct angles for the asymptotes are determined as 60° , -60° , and 180° . Thus, there are three asymptotes. The one having the angle of 180° is the negative real axis.

Before we can draw these asymptotes in the complex plane, we must find the point where they intersect the real axis. Since

$$G(s) = \frac{K}{s(s+1)(s+2)}$$

if a test point is located very far from the origin, then $G(s)$ may be written as

$$G(s) = \frac{K}{s^3 + 3s^2 + \dots}$$

For large values of s , this last equation may be approximated by

$$G(s) \doteq \frac{K}{(s+1)^3} \quad (6-5)$$

A root-locus diagram of $G(s)$ given by Equation (6-5) consists of three straight lines. This can be seen as follows: The equation of the root locus is

$$\angle \frac{K}{(s+1)^3} = \pm 180^\circ(2k+1)$$

or

$$-3\angle s + 1 = \pm 180^\circ(2k+1)$$

which can be written as

$$\angle s + 1 = \pm 60^\circ(2k+1)$$

By substituting $s = \sigma + j\omega$ into this last equation, we obtain

$$\angle \sigma + j\omega + 1 = \pm 60^\circ (2k + 1)$$

or

$$\tan^{-1} \frac{\omega}{\sigma + 1} = 60^\circ, \quad -60^\circ, \quad 0^\circ$$

Taking the tangent of both sides of this last equation,

$$\frac{\omega}{\sigma + 1} = \sqrt{3}, \quad -\sqrt{3}, \quad 0$$

which can be written as

$$\sigma + 1 - \frac{\omega}{\sqrt{3}} = 0, \quad \sigma + 1 + \frac{\omega}{\sqrt{3}} = 0, \quad \omega = 0$$

These three equations represent three straight lines, as shown in Figure 6-4. The three straight lines shown are the asymptotes. They meet at point $s = -1$. Thus, the abscissa of the intersection of the asymptotes and the real axis is obtained by setting the denominator of the right-hand side of Equation (6-5) equal to zero and solving for s . The asymptotes are almost parts of the root loci in regions very far from the origin.

3. *Determine the breakaway point.* To plot root loci accurately, we must find the breakaway point, where the root-locus branches originating from the poles at 0 and -1 break away (as K is increased) from the real axis and move into the complex plane. The breakaway point corresponds to a point in the s plane where multiple roots of the characteristic equation occur.

A simple method for finding the breakaway point is available. We shall present this method in the following: Let us write the characteristic equation as

$$f(s) = B(s) + KA(s) = 0 \quad (6-6)$$

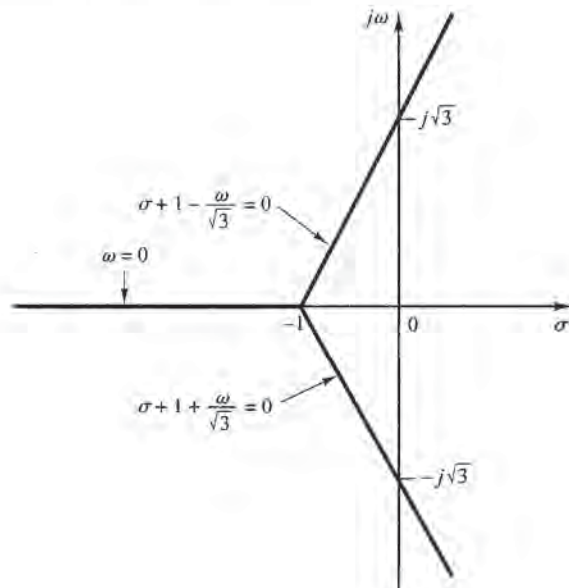


Figure 6-4
Three asymptotes.

where $A(s)$ and $B(s)$ do not contain K . Note that $f(s) = 0$ has multiple roots at points where

$$\frac{df(s)}{ds} = 0$$

This can be seen as follows: Suppose that $f(s)$ has multiple roots of order r . Then $f(s)$ may be written as

$$f(s) = (s - s_1)^r (s - s_2) \cdots (s - s_n)$$

If we differentiate this equation with respect to s and set $s = s_1$, then we get

$$\left. \frac{df(s)}{ds} \right|_{s=s_1} = 0 \quad (6-7)$$

This means that multiple roots of $f(s)$ will satisfy Equation (6-7). From Equation (6-6), we obtain

$$\frac{df(s)}{ds} = B'(s) + KA'(s) = 0 \quad (6-8)$$

where

$$A'(s) = \frac{dA(s)}{ds}, \quad B'(s) = \frac{dB(s)}{ds}$$

The particular value of K that will yield multiple roots of the characteristic equation is obtained from Equation (6-8) as

$$K = -\frac{B'(s)}{A'(s)}$$

If we substitute this value of K into Equation (6-6), we get

$$f(s) = B(s) - \frac{B'(s)}{A'(s)} A(s) = 0$$

or

$$B(s)A'(s) - B'(s)A(s) = 0 \quad (6-9)$$

If Equation (6-9) is solved for s , the points where multiple roots occur can be obtained. On the other hand, from Equation (6-6) we obtain

$$K = -\frac{B(s)}{A(s)}$$

and

$$\frac{dK}{ds} = -\frac{B'(s)A(s) - B(s)A'(s)}{A^2(s)}$$

If dK/ds is set equal to zero, we get the same equation as Equation (6-9). Therefore, the breakaway points can be simply determined from the roots of

$$\frac{dK}{ds} = 0$$

It should be noted that not all the solutions of Equation (6-9) or of $dK/ds = 0$ correspond to actual breakaway points. If a point at which $dK/ds = 0$ is on a root locus, it is an actual breakaway or break-in point. Stated differently, if at a point at which $dK/ds = 0$ the value of K takes a real positive value then that point is an actual breakaway or break-in point.

For the present example, the characteristic equation $G(s) + 1 = 0$ is given by

$$\frac{K}{s(s+1)(s+2)} + 1 = 0$$

or

$$K = -(s^3 + 3s^2 + 2s)$$

By setting $dK/ds = 0$, we obtain

$$\frac{dK}{ds} = -(3s^2 + 6s + 2) = 0$$

or

$$s = -0.4226, \quad s = -1.5774$$

Since the breakaway point must lie on a root locus between 0 and -1 , it is clear that $s = -0.4226$ corresponds to the actual breakaway point. Point $s = -1.5774$ is not on the root locus. Hence, this point is not an actual breakaway or break-in point. In fact, evaluation of the values of K corresponding to $s = -0.4226$ and $s = -1.5774$ yields

$$K = 0.3849, \quad \text{for } s = -0.4226$$

$$K = -0.3849, \quad \text{for } s = -1.5774$$

4. *Determine the points where the root loci cross the imaginary axis.* These points can be found by use of Routh's stability criterion as follows: Since the characteristic equation for the present system is

$$s^3 + 3s^2 + 2s + K = 0$$

the Routh array becomes

$$\begin{array}{ccc} s^3 & 1 & 2 \\ s^2 & 3 & K \\ s^1 & \frac{6-K}{3} & \\ s^0 & K & \end{array}$$

The value of K that makes the s^1 term in the first column equal zero is $K = 6$. The crossing points on the imaginary axis can then be found by solving the auxiliary equation obtained from the s^2 row; that is,

$$3s^2 + K = 3s^2 + 6 = 0$$

which yields

$$s = \pm j\sqrt{2}$$

The frequencies at the crossing points on the imaginary axis are thus $\omega = \pm\sqrt{2}$. The gain value corresponding to the crossing points is $K = 6$.

An alternative approach is to let $s = j\omega$ in the characteristic equation, equate both the real part and the imaginary part to zero, and then solve for ω and K . For the present system, the characteristic equation, with $s = j\omega$, is

$$(j\omega)^3 + 3(j\omega)^2 + 2(j\omega) + K = 0$$

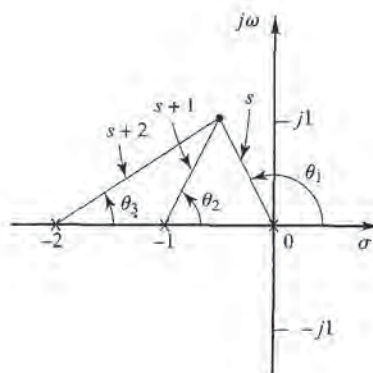
or

$$(K - 3\omega^2) + j(2\omega - \omega^3) = 0$$

Equating both the real and imaginary parts of this last equation to zero, we obtain

$$K - 3\omega^2 = 0, \quad 2\omega - \omega^3 = 0$$

Figure 6-5
Construction of root locus.



from which

$$\omega = \pm\sqrt{2}, \quad K = 6 \quad \text{or} \quad \omega = 0, \quad K = 0$$

Thus, root loci cross the imaginary axis at $\omega = \pm\sqrt{2}$, and the value of K at the crossing points is 6. Also, a root-locus branch on the real axis touches the imaginary axis at $\omega = 0$.

5. Choose a test point in the broad neighborhood of the $j\omega$ axis and the origin, as shown in Figure 6-5, and apply the angle condition. If a test point is on the root loci, then the sum of the three angles, $\theta_1 + \theta_2 + \theta_3$, must be 180° . If the test point does not satisfy the angle condition, select another test point until it satisfies the condition. (The sum of the angles at the test point will indicate which direction the test point should be moved.) Continue this process and locate a sufficient number of points satisfying the angle condition.

6. Draw the root loci, based on the information obtained in the foregoing steps, as shown in Figure 6-6.

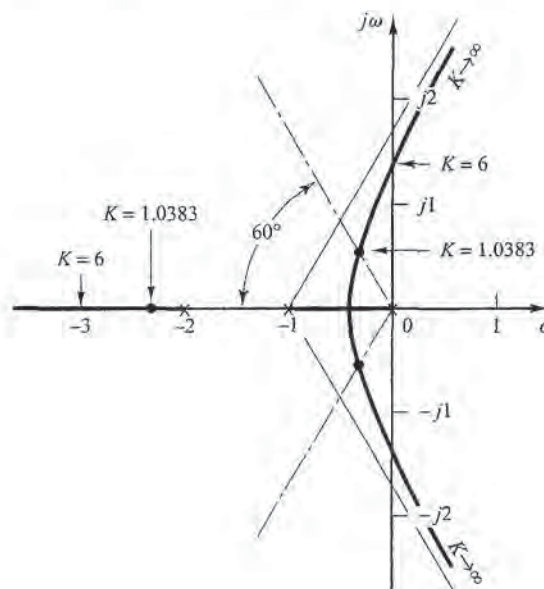


Figure 6-6
Root-locus plot.

7. Determine a pair of dominant complex-conjugate closed-loop poles such that the damping ratio ζ is 0.5. Closed-loop poles with $\zeta = 0.5$ lie on lines passing through the origin and making the angles $\pm \cos^{-1} \zeta = \pm \cos^{-1} 0.5 = \pm 60^\circ$ with the negative real axis. From Figure 6-6, such closed-loop poles having $\zeta = 0.5$ are obtained as follows:

$$s_1 = -0.3337 + j0.5780, \quad s_2 = -0.3337 - j0.5780$$

The value of K that yields such poles is found from the magnitude condition as follows:

$$\begin{aligned} K &= |s(s+1)(s+2)|_{s=-0.3337+j0.5780} \\ &= 1.0383 \end{aligned}$$

Using this value of K , the third pole is found at $s = -2.3326$.

Note that, from step 4, it can be seen that for $K = 6$ the dominant closed-loop poles lie on the imaginary axis at $s = \pm j\sqrt{2}$. With this value of K , the system will exhibit sustained oscillations. For $K > 6$, the dominant closed-loop poles lie in the right-half s plane, resulting in an unstable system.

Finally, note that, if necessary, the root loci can be easily graduated in terms of K by use of the magnitude condition. We simply pick out a point on a root locus, measure the magnitudes of the three complex quantities s , $s+1$, and $s+2$, and multiply these magnitudes; the product is equal to the gain value K at that point, or

$$|s| \cdot |s+1| \cdot |s+2| = K$$

Graduation of the root loci can be done easily by use of MATLAB. (See Section 6-4.)

EXAMPLE 6-2 In this example, we shall sketch the root-locus plot of a system with complex-conjugate open-loop poles. Consider the system shown in Figure 6-7. For this system,

$$G(s) = \frac{K(s+2)}{s^2 + 2s + 3}, \quad H(s) = 1$$

where $K \geq 0$. It is seen that $G(s)$ has a pair of complex conjugate poles at

$$s = -1 + j\sqrt{2}, \quad s = -1 - j\sqrt{2}$$

A typical procedure for sketching the root-locus plot is as follows:

1. *Determine the root loci on the real axis.* For any test point s on the real axis, the sum of the angular contributions of the complex-conjugate poles is 360° , as shown in Figure 6-8. Thus the net effect of the complex-conjugate poles is zero on the real axis. The location of the root locus on the real axis is determined from the open-loop zero on the negative real axis. A simple test reveals that a section of the negative real axis, that between -2 and $-\infty$, is a part of the root locus. It is noted that, since this locus lies between two zeros (at $s = -2$ and $s = -\infty$), it is actually a part of two root loci, each of which starts from one of the two complex-conjugate poles. In other words, two root loci break in the part of the negative real axis between -2 and $-\infty$.

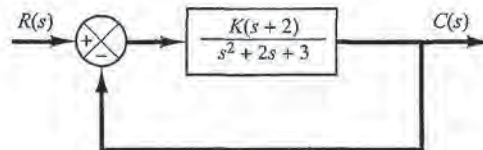
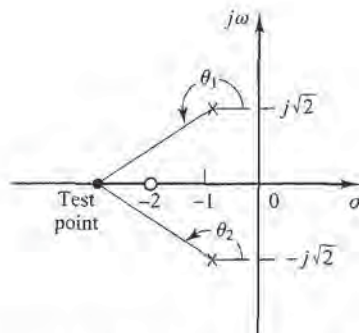


Figure 6-7
Control system.

Figure 6-8
Determination of the
root locus on the real
axis.



Since there are two open-loop poles and one zero, there is one asymptote, which coincides with the negative real axis.

2. *Determine the angle of departure from the complex-conjugate open-loop poles.* The presence of a pair of complex-conjugate open-loop poles requires the determination of the angle of departure from these poles. Knowledge of this angle is important since the root locus near a complex pole yields information as to whether the locus originating from the complex pole migrates toward the real axis or extends toward the asymptote.

Referring to Figure 6-9, if we choose a test point and move it in the very vicinity of the complex open-loop pole at $s = -p_1$, we find that the sum of the angular contributions from the pole at $s = p_2$ and zero at $s = -z_1$ to the test point can be considered remaining the same. If the test point is to be on the root locus, then the sum of ϕ'_1 , $-\theta_1$, and $-\theta'_2$ must be $\pm 180^\circ(2k + 1)$, where $k = 0, 1, 2, \dots$. Thus, in the example,

$$\phi'_1 - (\theta_1 + \theta'_2) = \pm 180^\circ(2k + 1)$$

or

$$\theta_1 = 180^\circ - \theta'_2 + \phi'_1 = 180^\circ - \theta_2 + \phi_1$$

The angle of departure is then

$$\theta_1 = 180^\circ - \theta_2 + \phi_1 = 180^\circ - 90^\circ + 55^\circ = 145^\circ$$

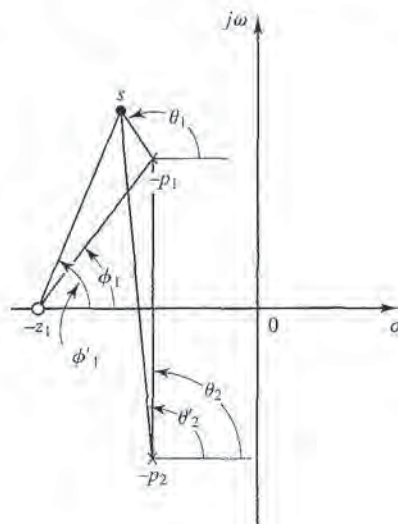


Figure 6-9
Determination of the
angle of departure.

Since the root locus is symmetric about the real axis, the angle of departure from the pole at $s = -p_2$ is -145° .

3. *Determine the break-in point.* A break-in point exists where a pair of root-locus branches coalesces as K is increased. For this problem, the break-in point can be found as follows: Since

$$K = -\frac{s^2 + 2s + 3}{s + 2}$$

we have

$$\frac{dK}{ds} = -\frac{(2s + 2)(s + 2) - (s^2 + 2s + 3)}{(s + 2)^2} = 0$$

which gives

$$s^2 + 4s + 1 = 0$$

or

$$s = -3.7320 \quad \text{or} \quad s = -0.2680$$

Notice that point $s = -3.7320$ is on the root locus. Hence this point is an actual break-in point. (Note that at point $s = -3.7320$ the corresponding gain value is $K = 5.4641$.) Since point $s = -0.2680$ is not on the root locus, it cannot be a break-in point. (For point $s = -0.2680$, the corresponding gain value is $K = -1.4641$.)

4. *Sketch a root-locus plot, based on the information obtained in the foregoing steps.* To determine accurate root loci, several points must be found by trial and error between the break-in point and the complex open-loop poles. (To facilitate sketching the root-locus plot, we should find the direction in which the test point should be moved by mentally summing up the changes on the angles of the poles and zeros.) Figure 6-10 shows a complete root-locus plot for the system considered.

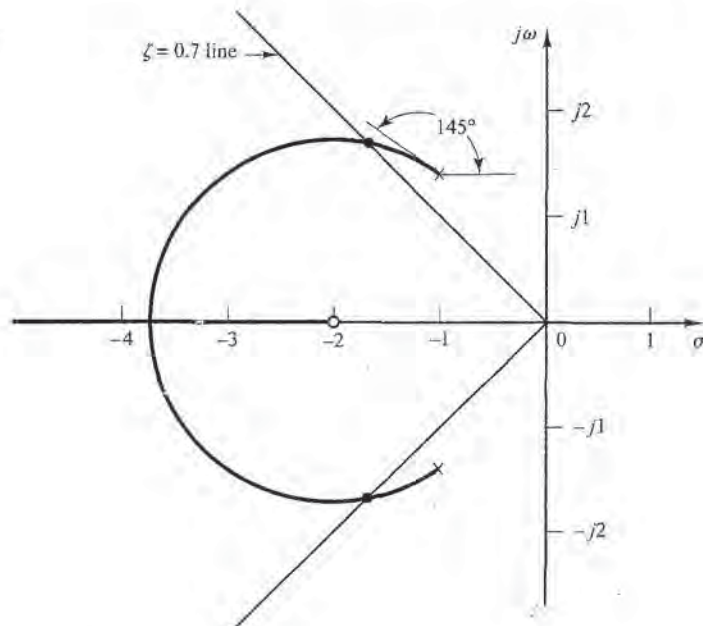


Figure 6-10
Root-locus plot.

The value of the gain K at any point on root locus can be found by applying the magnitude condition or by use of MATLAB (see Section 6-4). For example, the value of K at which the complex-conjugate closed-loop poles have the damping ratio $\zeta = 0.7$ can be found by locating the roots, as shown in Figure 6-10, and computing the value of K as follows:

$$K = \left| \frac{(s+1-j\sqrt{2})(s+1+j\sqrt{2})}{s+2} \right|_{s=-1.67+j1.70} = 1.34$$

Or use MATLAB to find the value of K . (See Section 6-4.)

It is noted that in this system the root locus in the complex plane is a part of a circle. Such a circular root locus will not occur in most systems. Circular root loci may occur in systems that involve two poles and one zero, two poles and two zeros, or one pole and two zeros. Even in such systems, whether circular root loci occur depends on the locations of poles and zeros involved.

To show the occurrence of a circular root locus in the present system, we need to derive the equation for the root locus. For the present system, the angle condition is

$$\angle s+2 - \angle s+1-j\sqrt{2} - \angle s+1+j\sqrt{2} = \pm 180^\circ(2k+1)$$

If $s = \sigma + j\omega$ is substituted into this last equation, we obtain

$$\angle \sigma+2+j\omega - \angle \sigma+1+j\omega-j\sqrt{2} - \angle \sigma+1+j\omega+j\sqrt{2} = \pm 180^\circ(2k+1)$$

which can be written as

$$\tan^{-1}\left(\frac{\omega}{\sigma+2}\right) - \tan^{-1}\left(\frac{\omega-j\sqrt{2}}{\sigma+1}\right) - \tan^{-1}\left(\frac{\omega+j\sqrt{2}}{\sigma+1}\right) = \pm 180^\circ(2k+1)$$

or

$$\tan^{-1}\left(\frac{\omega-j\sqrt{2}}{\sigma+1}\right) + \tan^{-1}\left(\frac{\omega+j\sqrt{2}}{\sigma+1}\right) = \tan^{-1}\left(\frac{\omega}{\sigma+2}\right) \pm 180^\circ(2k+1)$$

Taking tangents of both sides of this last equation using the relationship

$$\tan(x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \tan y} \quad (6-10)$$

we obtain

$$\tan\left[\tan^{-1}\left(\frac{\omega-j\sqrt{2}}{\sigma+1}\right) + \tan^{-1}\left(\frac{\omega+j\sqrt{2}}{\sigma+1}\right)\right] = \tan\left[\tan^{-1}\left(\frac{\omega}{\sigma+2}\right) \pm 180^\circ(2k+1)\right]$$

or

$$\frac{\frac{\omega-j\sqrt{2}}{\sigma+1} + \frac{\omega+j\sqrt{2}}{\sigma+1}}{1 - \left(\frac{\omega-j\sqrt{2}}{\sigma+1}\right)\left(\frac{\omega+j\sqrt{2}}{\sigma+1}\right)} = \frac{\frac{\omega}{\sigma+2} \pm 0}{1 \mp \frac{\omega}{\sigma+2} \times 0}$$

which can be simplified to

$$\frac{2\omega(\sigma+1)}{(\sigma+1)^2 - (\omega^2 - 2)} = \frac{\omega}{\sigma+2}$$

or

$$\omega[(\sigma+2)^2 + \omega^2 - 3] = 0$$

This last equation is equivalent to

$$\omega = 0 \quad \text{or} \quad (\sigma+2)^2 + \omega^2 = (\sqrt{3})^2$$

These two equations are the equations for the root loci for the present system. Notice that the first equation, $\omega = 0$, is the equation for the real axis. The real axis from $s = -2$ to $s = -\infty$ corresponds to a root locus for $K \geq 0$. The remaining part of the real axis corresponds to a root locus when K is negative. (In the present system, K is nonnegative.) The second equation for the root locus is an equation of a circle with center at $\sigma = -2$, $\omega = 0$ and the radius equal to $\sqrt{3}$. That part of the circle to the left of the complex-conjugate poles corresponds to a root locus for $K \geq 0$. The remaining part of the circle corresponds to a root locus when K is negative.

It is important to note that easily interpretable equations for the root locus can be derived for simple systems only. For complicated systems having many poles and zeros, any attempt to derive equations for the root loci is discouraged. Such derived equations are very complicated and their configuration in the complex plane is difficult to visualize.

6-3 SUMMARY OF GENERAL RULES FOR CONSTRUCTING ROOT LOCI

For a complicated system with many open-loop poles and zeros, constructing a root-locus plot may seem complicated, but actually it is not difficult if the rules for constructing the root loci are applied. By locating particular points and asymptotes and by computing angles of departure from complex poles and angles of arrival at complex zeros, we can construct the general form of the root loci without difficulty.

Some of the rules for constructing root loci were given in Section 6-2. The purpose of this section is to summarize the general rules for constructing root loci of the system shown in Figure 6-11. While the root-locus method is essentially based on a trial-and-error technique, the number of trials required can be greatly reduced if we use these rules.

General Rules for Constructing Root Loci. We shall now summarize the general rules and procedure for constructing the root loci of the system shown in Figure 6-11. First, obtain the characteristic equation

$$1 + G(s)H(s) = 0$$

Then rearrange this equation so that the parameter of interest appears as the multiplying factor in the form

$$1 + \frac{K(s + z_1)(s + z_2) \cdots (s + z_m)}{(s + p_1)(s + p_2) \cdots (s + p_n)} = 0 \tag{6-11}$$

In the present discussions, we assume that the parameter of interest is the gain K , where $K > 0$. (If $K < 0$, which corresponds to the positive-feedback case, the angle condition must be modified. See Section 6-5.) Note, however, that the method is still applicable to systems with parameters of interest other than gain. (See Section 7-6.)

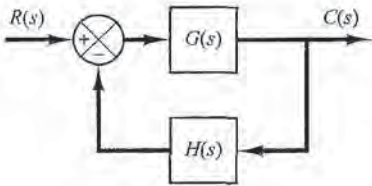


Figure 6-11
Control system.

1. Locate the poles and zeros of $G(s)H(s)$ on the s plane. The root-locus branches start from open-loop poles and terminate at zeros (finite zeros or zeros at infinity). From the factored form of the open-loop transfer function, locate the open-loop poles and zeros in the s plane. [Note that the open-loop zeros are the zeros of $G(s)H(s)$, while the closed-loop zeros consist of the zeros of $G(s)$ and the poles of $H(s)$.]

Note that the root loci are symmetrical about the real axis of the s plane, because the complex poles and complex zeros occur only in conjugate pairs.

A root-locus plot will have just as many branches as there are roots of the characteristic equation. Since the number of open-loop poles generally exceeds that of zeros, the number of branches equals that of poles. If the number of closed-loop poles is the same as the number of open-loop poles, then the number of individual root-locus branches terminating at finite open-loop zeros is equal to the number m of the open-loop zeros. The remaining $n - m$ branches terminate at infinity ($n - m$ implicit zeros at infinity) along asymptotes.

If we include poles and zeros at infinity, the number of open-loop poles is equal to that of open-loop zeros. Hence we can always state that the root loci start at the poles of $G(s)H(s)$ and end at the zeros of $G(s)H(s)$, as K increases from zero to infinity, where the poles and zeros include both those in the finite s plane and those at infinity.

2. Determine the root loci on the real axis. Root loci on the real axis are determined by open-loop poles and zeros lying on it. The complex-conjugate poles and zeros of the open-loop transfer function have no effect on the location of the root loci on the real axis because the angle contribution of a pair of complex-conjugate poles or zeros is 360° on the real axis. Each portion of the root locus on the real axis extends over a range from a pole or zero to another pole or zero. In constructing the root loci on the real axis, choose a test point on it. If the total number of real poles and real zeros to the right of this test point is odd, then this point lies on a root locus. If the open-loop poles and open-loop zeros are simple poles and simple zeros, then the root locus and its complement form alternate segments along the real axis.

3. Determine the asymptotes of root loci. If the test point s is located far from the origin, then the angle of each complex quantity may be considered the same. One open-loop zero and one open-loop pole then cancel the effects of the other. Therefore, the root loci for very large values of s must be asymptotic to straight lines whose angles (slopes) are given by

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k + 1)}{n - m} \quad (k = 0, 1, 2, \dots)$$

where n = number of finite poles of $G(s)H(s)$

m = number of finite zeros of $G(s)H(s)$

Here, $k = 0$ corresponds to the asymptotes with the smallest angle with the real axis. Although k assumes an infinite number of values, as k is increased the angle repeats itself, and the number of distinct asymptotes is $n - m$.

All the asymptotes intersect on the real axis. The point at which they do so is obtained as follows: If both the numerator and denominator of the open-loop transfer function are expanded, the result is

$$G(s)H(s) = \frac{K[s^m + (z_1 + z_2 + \dots + z_m)s^{m-1} + \dots + z_1 z_2 \dots z_m]}{s^n + (p_1 + p_2 + \dots + p_n)s^{n-1} + \dots + p_1 p_2 \dots p_n}$$

If a test point is located very far from the origin, then by dividing the denominator by the numerator, it is possible to write $G(s)H(s)$ as

$$G(s)H(s) = \frac{K}{s^{n-m} + [(p_1 + p_2 + \cdots + p_n) - (z_1 + z_2 + \cdots + z_m)]s^{n-m-1} + \cdots}$$

or

$$G(s)H(s) = \frac{K}{\left[s + \frac{(p_1 + p_2 + \cdots + p_n) - (z_1 + z_2 + \cdots + z_m)}{n - m} \right]^{n-m}} \quad (6-12)$$

The abscissa of the intersection of the asymptotes and the real axis is then obtained by setting the denominator of the right-hand side of Equation (6-12) equal to zero and solving for s , or

$$s = - \frac{(p_1 + p_2 + \cdots + p_n) - (z_1 + z_2 + \cdots + z_m)}{n - m} \quad (6-13)$$

[Example 6-1 shows why Equation (6-13) gives the intersection.] Once this intersection is determined, the asymptotes can be readily drawn in the complex plane.

It is important to note that the asymptotes show the behavior of the root loci for $|s| \gg 1$. A root locus branch may lie on one side of the corresponding asymptote or may cross the corresponding asymptote from one side to the other side.

4. Find the breakaway and break-in points. Because of the conjugate symmetry of the root loci, the breakaway points and break-in points either lie on the real axis or occur in complex-conjugate pairs.

If a root locus lies between two adjacent open-loop poles on the real axis, then there exists at least one breakaway point between the two poles. Similarly, if the root locus lies between two adjacent zeros (one zero may be located at $-\infty$) on the real axis, then there always exists at least one break-in point between the two zeros. If the root locus lies between an open-loop pole and a zero (finite or infinite) on the real axis, then there may exist no breakaway or break-in points or there may exist both breakaway and break-in points.

Suppose that the characteristic equation is given by

$$B(s) + KA(s) = 0$$

The breakaway points and break-in points correspond to multiple roots of the characteristic equation. Hence, as discussed in Example 6-1, the breakaway and break-in points can be determined from the roots of

$$\frac{dK}{ds} = - \frac{B'(s)A(s) - B(s)A'(s)}{A^2(s)} = 0 \quad (6-14)$$

where the prime indicates differentiation with respect to s . It is important to note that the breakaway points and break-in points must be the roots of Equation (6-14), but not all roots of Equation (6-14) are breakaway or break-in points. If a real root of Equation (6-14) lies on the root-locus portion of the real axis, then it is an actual breakaway or break-in point. If a real root of Equation (6-14) is not on the root-locus portion of the real axis, then this root corresponds to neither a breakaway point nor a break-in point.

If two roots $s = s_1$ and $s = -s_1$ of Equation (6-14) are a complex-conjugate pair and if it is not certain whether they are on root loci, then it is necessary to check the corresponding K value. If the value of K corresponding to a root $s = s_1$ of $dK/ds = 0$ is positive, point $s = s_1$ is an actual breakaway or break-in point. (Since K is assumed to be nonnegative, if the value of K thus obtained is negative, or a complex quantity, then point $s = s_1$ is neither a breakaway nor break-in point.)

5. Determine the angle of departure (angle of arrival) of the root locus from a complex pole (at a complex zero). To sketch the root loci with reasonable accuracy, we must find the directions of the root loci near the complex poles and zeros. If a test point is chosen and moved in the very vicinity of a complex pole (or complex zero), the sum of the angular contributions from all other poles and zeros can be considered to remain the same. Therefore, the angle of departure (or angle of arrival) of the root locus from a complex pole (or at a complex zero) can be found by subtracting from 180° the sum of all the angles of vectors from all other poles and zeros to the complex pole (or complex zero) in question, with appropriate signs included.

Angle of departure from a complex pole = 180°

– (sum of the angles of vectors to a complex pole in question from other poles)
+ (sum of the angles of vectors to a complex pole in question from zeros)

Angle of arrival at a complex zero = 180°

– (sum of the angles of vectors to a complex zero in question from other zeros)
+ (sum of the angles of vectors to a complex zero in question from poles)

The angle of departure is shown in Figure 6-12.

6. Find the points where the root loci may cross the imaginary axis. The points where the root loci intersect the $j\omega$ axis can be found easily by (a) use of Routh's stability criterion or (b) letting $s = j\omega$ in the characteristic equation, equating both the real part and the imaginary part to zero, and solving for ω and K . The values of ω thus found give the frequencies at which root loci cross the imaginary axis. The K value corresponding to each crossing frequency gives the gain at the crossing point.

7. Taking a series of test points in the broad neighborhood of the origin of the s plane, sketch the root loci. Determine the root loci in the broad neighborhood of the $j\omega$ axis and the origin. The most important part of the root loci is on neither the real axis nor the asymptotes, but the part in the broad neighborhood of the $j\omega$ axis and the origin. The

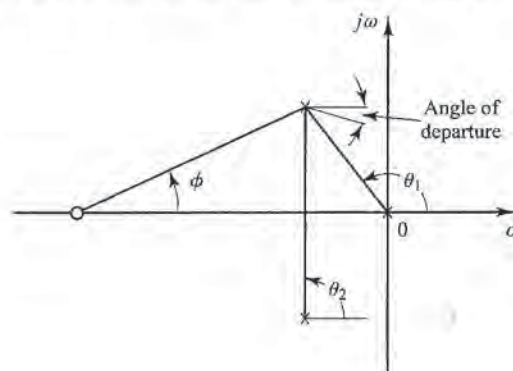


Figure 6-12
Construction of the
root locus. [Angle of
departure
= $180^\circ -$
 $[\theta_1 + \theta_2] + \phi$.]

shape of the root loci in this important region in the s plane must be obtained with reasonable accuracy. (If accurate shape of the root loci is needed, MATLAB may be used rather than hand calculations of the exact shape of the root loci.)

8. Determine closed-loop poles. A particular point on each root-locus branch will be a closed-loop pole if the value of K at that point satisfies the magnitude condition. Conversely, the magnitude condition enables us to determine the value of the gain K at any specific root location on the locus. (If necessary, the root loci may be graduated in terms of K . The root loci are continuous with K .)

The value of K corresponding to any point s on a root locus can be obtained using the magnitude condition, or

$$K = \frac{\text{product of lengths between point } s \text{ to poles}}{\text{product of lengths between point } s \text{ to zeros}}$$

This value can be evaluated either graphically or analytically. (MATLAB can be used for graduating the root loci with K . See Section 6-4.)

If the gain K of the open-loop transfer function is given in the problem, then by applying the magnitude condition, we can find the correct locations of the closed-loop poles for a given K on each branch of the root loci by a trial-and-error approach or by use of MATLAB, which will be presented in Section 6-4.

Comments on the Root-Locus Plots. It is noted that the characteristic equation of the system whose open-loop transfer function is

$$G(s)H(s) = \frac{K(s^m + b_1s^{m-1} + \dots + b_m)}{s^n + a_1s^{n-1} + \dots + a_n} \quad (n \geq m)$$

is an n th-degree algebraic equation in s . If the order of the numerator of $G(s)H(s)$ is lower than that of the denominator by two or more (which means that there are two or more zeros at infinity), then the coefficient a_1 is the negative sum of the roots of the equation and is independent of K . In such a case, if some of the roots move on the locus toward the left as K is increased, then the other roots must move toward the right as K is increased. This information is helpful in finding the general shape of the root loci.

It is also noted that a slight change in the pole-zero configuration may cause significant changes in the root-locus configurations. Figure 6-13 demonstrates the fact that a slight change in the location of a zero or pole will make the root-locus configuration look quite different.

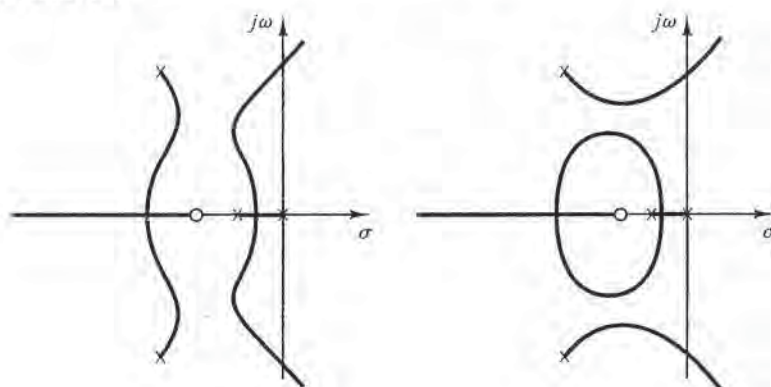


Figure 6-13
Root-locus plots.

Cancellation of Poles of $G(s)$ with Zeros of $H(s)$. It is important to note that if the denominator of $G(s)$ and the numerator of $H(s)$ involve common factors then the corresponding open-loop poles and zeros will cancel each other, reducing the degree of the characteristic equation by one or more. For example, consider the system shown in Figure 6-14(a). (This system has velocity feedback.) By modifying the block diagram of Figure 6-14(a) to that shown in Figure 6-14(b), it is clearly seen that $G(s)$ and $H(s)$ have a common factor $s + 1$. The closed-loop transfer function $C(s)/R(s)$ is

$$\frac{C(s)}{R(s)} = \frac{K}{s(s + 1)(s + 2) + K(s + 1)}$$

The characteristic equation is

$$[s(s + 2) + K](s + 1) = 0$$

Because of the cancellation of the terms $(s + 1)$ appearing in $G(s)$ and $H(s)$, however, we have

$$\begin{aligned} 1 + G(s)H(s) &= 1 + \frac{K(s + 1)}{s(s + 1)(s + 2)} \\ &= \frac{s(s + 2) + K}{s(s + 2)} \end{aligned}$$

The reduced characteristic equation is

$$s(s + 2) + K = 0$$

The root-locus plot of $G(s)H(s)$ does not show all the roots of the characteristic equation, only the roots of the reduced equation.

To obtain the complete set of closed-loop poles, we must add the canceled pole of $G(s)H(s)$ to those closed-loop poles obtained from the root-locus plot of $G(s)H(s)$. The important thing to remember is that the canceled pole of $G(s)H(s)$ is a closed-loop pole of the system, as seen from Figure 6-14(c).

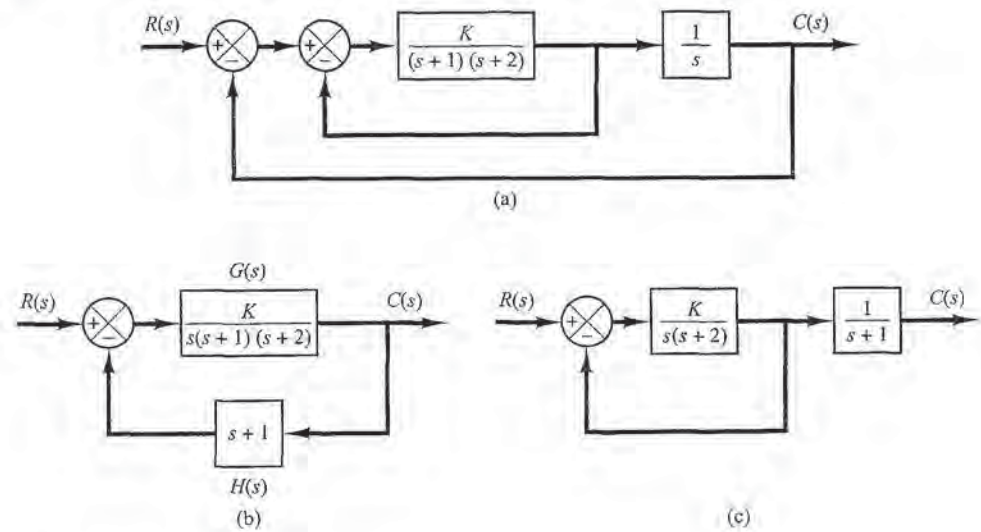
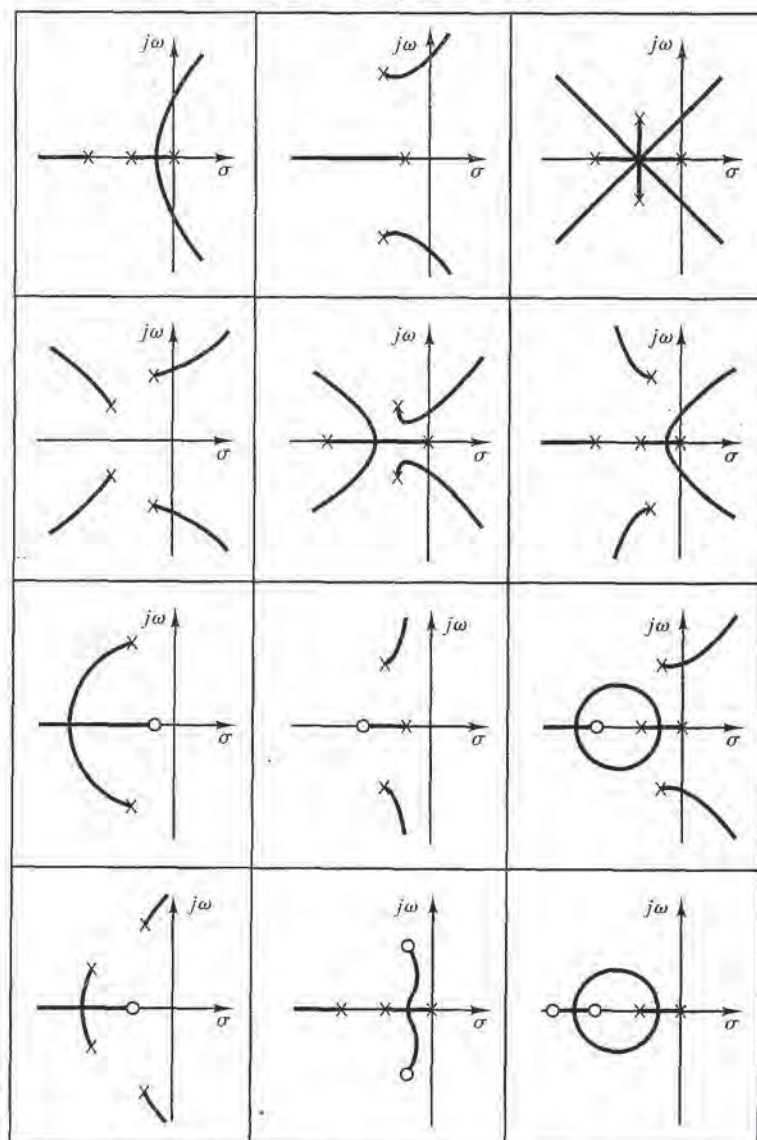


Figure 6-14
 (a) Control system with velocity feedback; (b) and (c) modified block diagrams.

Typical Pole-Zero Configurations and Corresponding Root Loci. In summarizing, we show several open-loop pole-zero configurations and their corresponding root loci in Table 6-1. The pattern of the root loci depends only on the relative separation of the open-loop poles and zeros. If the number of open-loop poles exceeds the number of finite zeros by three or more, there is a value of the gain K beyond which root loci enter the right-half s plane, and thus the system can become unstable. A stable system must have all its closed-loop poles in the left-half s plane.

Table 6-1 Open-Loop Pole-Zero Configurations and the Corresponding Root Loci



Note that once we have some experience with the method, we can easily evaluate the changes in the root loci due to the changes in the number and location of the open-loop poles and zeros by visualizing the root-locus plots resulting from various pole-zero configurations.

Summary. From the preceding discussions, it should be clear that it is possible to sketch a reasonably accurate root-locus diagram for a given system by following simple rules. (The reader should study the various root-locus diagrams shown in the solved problems at the end of the chapter.) At preliminary design stages, we may not need the precise locations of the closed-loop poles. Often their approximate locations are all that is needed to make an estimate of system performance. Thus, it is important that the designer have the capability of quickly sketching the root loci for a given system.

6-4 ROOT-LOCUS PLOTS WITH MATLAB

In this section we present the MATLAB approach to the generation of root-locus plots and finding relevant information from the root-locus plots.

Plotting Root Loci with MATLAB. In plotting root loci with MATLAB we deal with the system equation given in the form of Equation (6-11), which may be written as

$$1 + K \frac{\text{num}}{\text{den}} = 0$$

where num is the numerator polynomial and den is the denominator polynomial. That is,

$$\begin{aligned} \text{num} &= (s + z_1)(s + z_2) \cdots (s + z_m) \\ &= s^m + (z_1 + z_2 + \cdots + z_m)s^{m-1} + \cdots + z_1 z_2 \cdots z_m \\ \text{den} &= (s + p_1)(s + p_2) \cdots (s + p_n) \\ &= s^n + (p_1 + p_2 + \cdots + p_n)s^{n-1} + \cdots + p_1 p_2 \cdots p_n \end{aligned}$$

Note that both vectors num and den must be written in descending powers of s .

A MATLAB command commonly used for plotting root loci is

rlocus(num,den)

Using this command, the root-locus plot is drawn on the screen. The gain vector K is automatically determined. (The vector K contains all the gain values for which the closed-loop poles are to be computed.)

For the systems defined in state space, rlocus(A,B,C,D) plots the root locus of the system with the gain vector automatically determined.

Note that commands

rlocus(num,den,K) and rlocus(A,B,C,D,K)

use the user-supplied gain vector K .

If invoked with left-hand arguments

$$\begin{aligned} [r,K] &= \text{rlocus}(\text{num},\text{den}) \\ [r,K] &= \text{rlocus}(\text{num},\text{den},K) \end{aligned}$$

```
[r,K] = rlocus(A,B,C,D)
[r,K] = rlocus(A,B,C,D,K)
[r,K] = rlocus(sys)
```

the screen will show the matrix r and gain vector K . (r has length K rows and length $\text{den} - 1$ columns containing the complex root locations. Each row of the matrix corresponds to a gain from vector K .) The plot command

```
plot(r,'-')
```

plots the root loci.

If it is desired to plot the root loci with marks 'o' or 'x', it is necessary to use the following command:

```
r = rlocus(num,den)
plot(r,'o') or plot(r,'x')
```

Plotting root loci using marks o or x is instructive, since each calculated closed-loop pole is graphically shown; in some portion of the root loci those marks are densely placed and in another portion of the root loci they are sparsely placed. MATLAB supplies its own set of gain values used to calculate a root-locus plot. It does so by an internal adaptive step-size routine. Also, MATLAB uses the automatic axis-scaling feature of the plot command.

Finally, note that, since the gain vector is automatically determined, root-locus plots of

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

$$G(s)H(s) = \frac{10K(s+1)}{s(s+2)(s+3)}$$

$$G(s)H(s) = \frac{200K(s+1)}{s(s+2)(s+3)}$$

are all the same. The num and den set of the system is the same for all three systems. The num and den are

$$\text{num} = [0 \ 0 \ 1 \ 1]$$

$$\text{den} = [1 \ 5 \ 6 \ 0]$$

EXAMPLE 6-3 Consider the system shown in Figure 6-15. Plot root loci with a square aspect ratio so that a line with slope 1 is a true 45° line. Choose the region of root-locus plot to be

$$-6 \leq x \leq 6, \quad -6 \leq y \leq 6$$

where x and y are the real-axis coordinate and imaginary-axis coordinate, respectively.

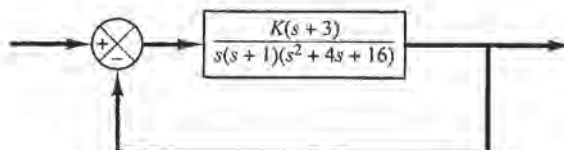


Figure 6-15
Control system.

To set the given plot region on the screen to be square, enter the command

```
v = [-6 6 -6 6]; axis(v); axis('square')
```

With this command, the region of the plot is as specified and a line with slope 1 is at a true 45°, not skewed by the irregular shape of the screen.

For this problem, the denominator is given as a product of first- and second-order terms. So we must multiply these terms to get a polynomial in s . The multiplication of these terms can be done easily by use of the convolution command, as shown next.

Define

```
a = s(s + 1);      a = [1 1 0]
b = s^2 + 4s + 16; b = [1 4 16]
```

Then we use the following command:

```
c = conv(a, b)
```

Note that `conv(a, b)` gives the product of two polynomials a and b . See the following computer output:

```
a = [1 1 0];
b = [1 4 16];
c = conv(a,b)
c =
    1    5   20   16    0
```

The denominator polynomial is thus found to be

```
den = [1 5 20 16 0]
```

To find the complex-conjugate open-loop poles (the roots of $s^2 + 4s + 16 = 0$), we may enter the `roots` command as follows:

```
r = roots(b)
r =
   -2.0000 + 3.4641i
   -2.0000 - 3.4641i
```

Thus, the system has the following open-loop zero and open-loop poles:

Open-loop zero: $s = -3$

Open-loop poles: $s = 0, s = -1, s = -2 \pm j3.4641$

MATLAB Program 6-1 will plot the root-locus diagram for this system. The plot is shown in Figure 6-16.

MATLAB Program 6-1

```
% ----- Root-locus plot -----
num = [0 0 0 1 3];
den = [1 5 20 16 0];
rlocus(num,den)
v = [-6 6 -6 6];
axis(v); axis('square')
grid;
title('Root-Locus Plot of G(s) = K(s + 3)/[s(s + 1)(s^2 + 4s + 16)]')
```

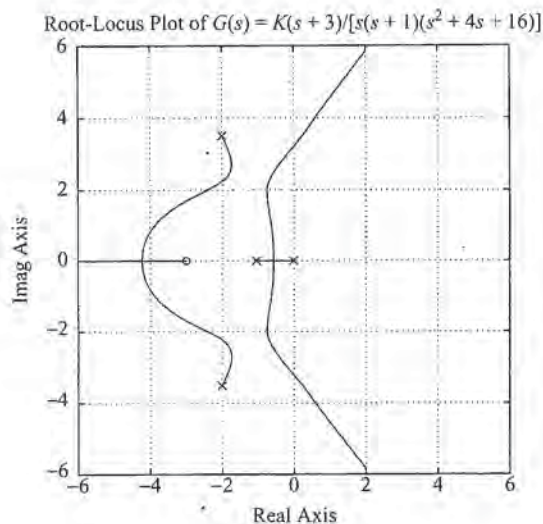


Figure 6-16
Root-locus plot.

Note that in MATLAB Program 6-1, instead of

`den = [1 5 20 16 0]`

we may enter

`den = conv([1 1 0], [1 4 16])`

The results are the same.

EXAMPLE 6-4 Consider the system whose open-loop transfer function $G(s)H(s)$ is

$$G(s)H(s) = \frac{K}{s(s+0.5)(s^2+0.6s+10)}$$

$$= \frac{K}{s^4 + 1.1s^3 + 10.3s^2 + 5s}$$

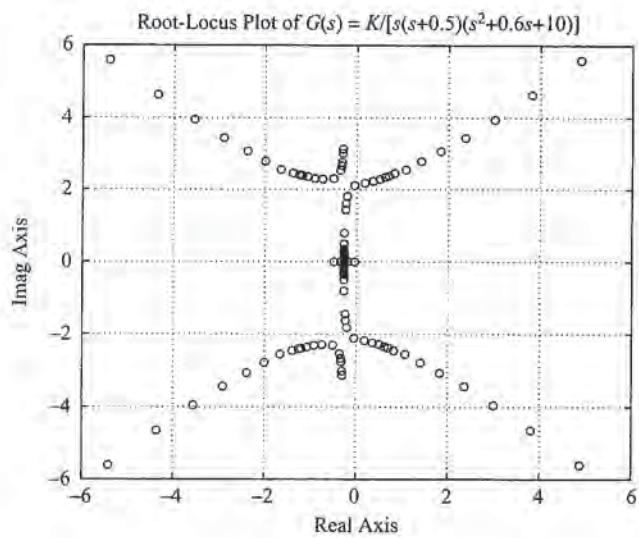
There are no open-loop zeros. Open-loop poles are located at $s = -0.3 + j3.1480$, $s = -0.3 - j3.1480$, $s = -0.5$, and $s = 0$.

Entering MATLAB Program 6-2 into the computer, we obtain the root-locus plot shown in Figure 6-17.

MATLAB Program 6-2

```
% ----- Root-locus plot -----
num = [0 0 0 0 1];
den = [1 1.1 10.3 5 0];
r = rlocus(num,den);
plot(r,'o')
v = [-6 6 -6 6]; axis(v)
grid
title('Root-Locus Plot of G(s) = K/[s(s+0.5)(s^2+0.6s+10)]')
xlabel('Real Axis')
ylabel('Imag Axis')
```

Figure 6-17
Root-locus plot.



Notice that in the regions near $x = -0.3$, $y = 2.3$ and $x = -0.3$, $y = -2.3$ two loci approach each other. We may wonder if these two branches should touch or not. To explore this situation, we may plot the root loci using smaller increments of K in the critical region.

By a conventional trial-and-error approach or using the command `rlocfind` to be presented later in this section, we find the particular region of interest to be $20 \leq K \leq 30$. By entering MATLAB Program 6-3, we obtain the root-locus plot shown in Figure 6-18. From this plot, it is clear that the two branches that approach in the upper half-plane (or in the lower half-plane) do not touch.

MATLAB Program 6-3
<pre>% ----- Root-locus plot ----- num = [0 0 0 0 1]; den = [1 1.1 10.3 5 0]; K1 = 0:0.2:20; K2 = 20:0.1:30; K3 = 30:5:1000; K = [K1 K2 K3]; r = rlocus(num,den,K); plot(r, 'o') v = [-4 4 -4 4]; axis(v) grid title('Root-Locus Plot of $G(s) = K/[s(s + 0.5)(s^2 + 0.6s + 10)]$') xlabel('Real Axis') ylabel('Imag Axis')</pre>

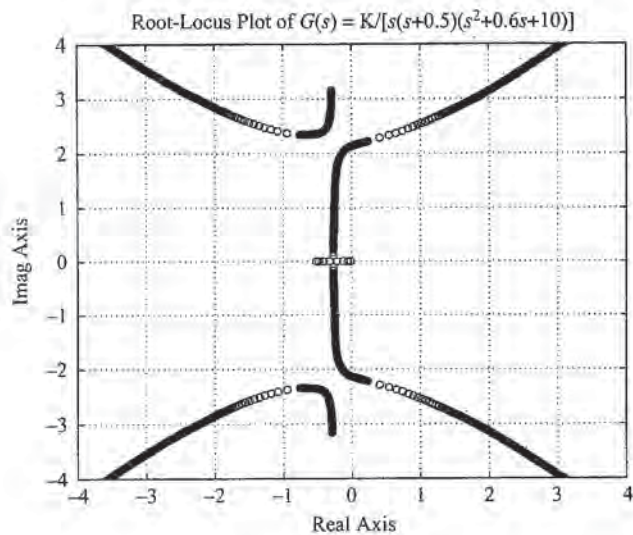


Figure 6-18
Root-locus plot.

EXAMPLE 6-5 Consider the system shown in Figure 6-19. The system equations are

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$$

$$y = \mathbf{C}\mathbf{x} + \mathbf{D}u$$

$$u = r - y$$

In this example problem we shall obtain the root-locus diagram of the system defined in state space. As an example let us consider the case where matrices **A**, **B**, **C**, and **D** are

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -160 & -56 & -14 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \\ -14 \end{bmatrix} \quad (6-15)$$

$$\mathbf{C} = [1 \ 0 \ 0], \quad \mathbf{D} = [0]$$

The root-locus plot for this system can be obtained with MATLAB by use of the following command:

`rlocus(A,B,C,D)`

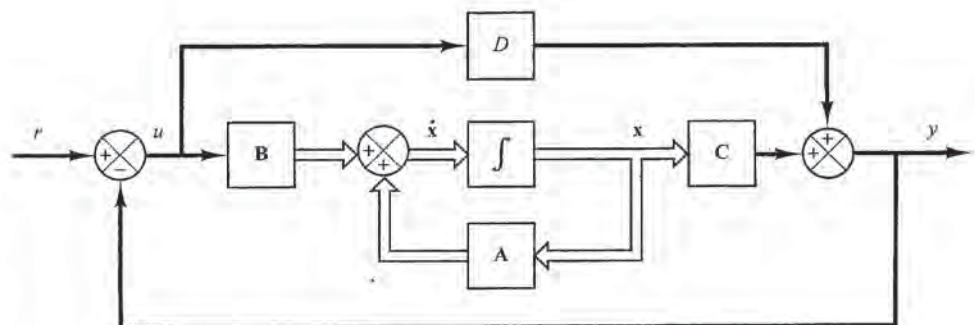


Figure 6-19
Closed-loop control system.

This command will produce the same root-locus plot as can be obtained by use of the rlocus (num,den) command, where num and den are obtained from

$$[\text{num}, \text{den}] = \text{ss2tf}(\text{A}, \text{B}, \text{C}, \text{D})$$

as follows:

$$\begin{aligned} \text{num} &= [0 \ 0 \ 1 \ 0] \\ \text{den} &= [1 \ 14 \ 56 \ 160] \end{aligned}$$

MATLAB Program 6-4 is a program that will generate the root-locus plot as shown in Figure 6-20.

MATLAB Program 6-4
<pre>% ----- Root-locus plot ----- A = [0 1 0; 0 0 1; -160 -56 -14]; B = [0; 1; -14]; C = [1 0 0]; D = [0]; K = 0:0.1:400; rlocus(A,B,C,D,K); v = [-20 20 -20 20]; axis(v) grid title('Root-Locus Plot of System Defined in State Space')</pre>

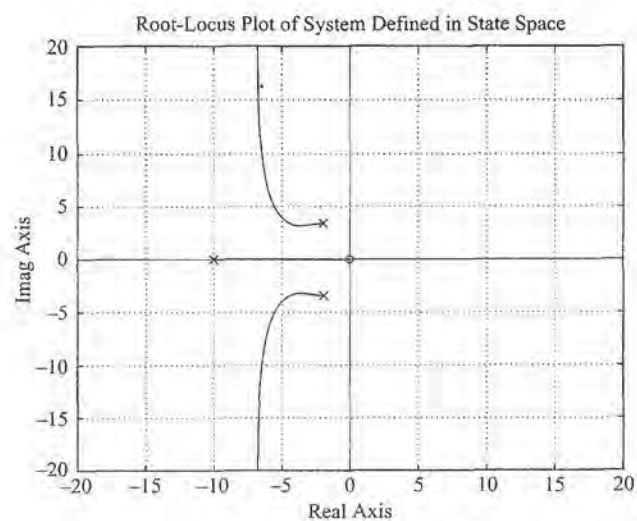
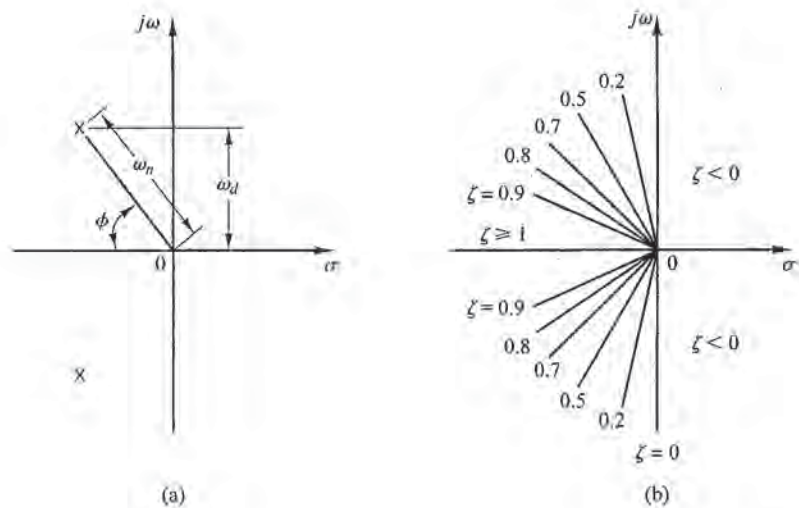


Figure 6-20
Root-locus plot of system defined in state space, where **A**, **B**, **C**, and **D** are as given by Equation (6-15).

Figure 6-21
(a) Complex poles;
(b) lines of constant damping ratio ζ .



Constant ζ Loci and Constant ω_n Loci. Recall that in the complex plane the damping ratio ζ of a pair of complex-conjugate poles can be expressed in terms of the angle ϕ , which is measured from the negative real axis, as shown in Figure 6-21(a) with

$$\zeta = \cos \phi$$

In other words, lines of constant damping ratio ζ are radial lines passing through the origin as shown in Figure 6-21(b). For example, a damping ratio of 0.5 requires that the complex poles lie on the lines drawn through the origin making angles of $\pm 60^\circ$ with the negative real axis. (If the real part of a pair of complex poles is positive, which means that the system is unstable, the corresponding ζ is negative.) the damping ratio determines the angular location of the poles, while the distance of the pole from the origin is determined by the undamped natural frequency ω_n . The constant ω_n loci are circles.

To draw constant ζ lines and constant ω_n circles on the root-locus diagram with MATLAB, use the command `sgrid`.

Plotting Polar Grids in the Root-Locus Diagram. The command

`sgrid`

overlays lines of constant damping ratio ($\zeta = 0 \sim 1$ with 0.1 increment) and circles of constant ω_n on the root-locus plot. See MATLAB Program 6-5 and the resulting diagram shown in Figure 6-22.

If only particular constant ζ lines (such as the $\zeta = 0.5$ line and $\zeta = 0.707$ line) and particular constant ω_n circles (such as the $\omega_n = 0.5$ circle, $\omega_n = 1$ circle, and $\omega_n = 2$ circle) are desired, use the following command:

`sgrid([0.5, 0.707], [0.5, 1, 2])`

MATLAB Program 6-5

```
sgrid
v = [-2 2 -2 2]; axis(v); axis('square')
title('Constant \zeta Lines and Constant \omega_n Circles')
xlabel('Real Axis')
ylabel('Imag Axis')
gtext('\zeta = 0.9')
gtext('0.8')
gtext('0.7')
gtext('0.6')
gtext('0.5')
gtext('0.4')
gtext('0.3')
gtext('0.2')
gtext('0.1')
gtext('\omega_n = 1')
gtext('\omega_n = 2')
```

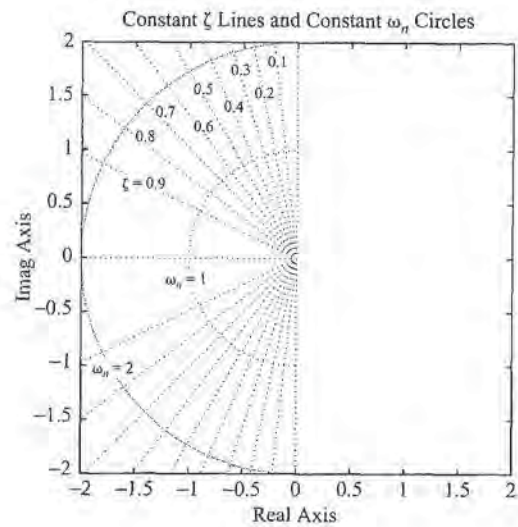


Figure 6-22
Constant ζ lines and
constant ω_n circles.

If we wish to overlay lines of constant ζ and circles of constant ω_n as given above to a root-locus plot of a system with

$$\begin{aligned} \text{num} &= [0 \ 0 \ 0 \ 1] \\ \text{den} &= [1 \ 4 \ 5 \ 0] \end{aligned}$$

then enter MATLAB Program 6-6 into the computer. The resulting root-locus plot is shown in Figure 6-23.

MATLAB Program 6-6

```
num = [0 0 0 1];
den = [1 4 5 0];
rlocus(num, den);
v = [-3 1 -2 2]; axis(v); axis('square')
sgrid([0.5,0.707], [0.5,1,2])
title('Root-Locus Plot with \zeta = 0.5 and 0.707 Lines and \omega_n = 0.5, 1, and 2 Circles')
gtext('\zeta = 0.5')
gtext('\zeta = 0.707')
gtext('\omega_n = 2')
gtext('\omega_n = 1')
gtext('\omega_n = 0.5')
```

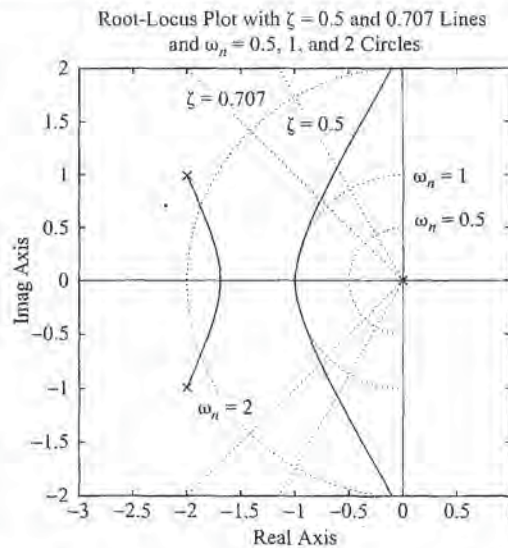


Figure 6-23
Constant ζ lines and constant ω_n circles superimposed on a root-locus plot.

If we want to omit either the entire constant ζ lines or entire constant ω_n circles, we may use empty brackets `[]` in the arguments of the `sgrid` command. For example, if we want to overlay only the constant damping ratio line corresponding to $\zeta = 0.5$ and no constant ω_n circles to the root-locus plot shown in Figure 6-23, then we may use the command

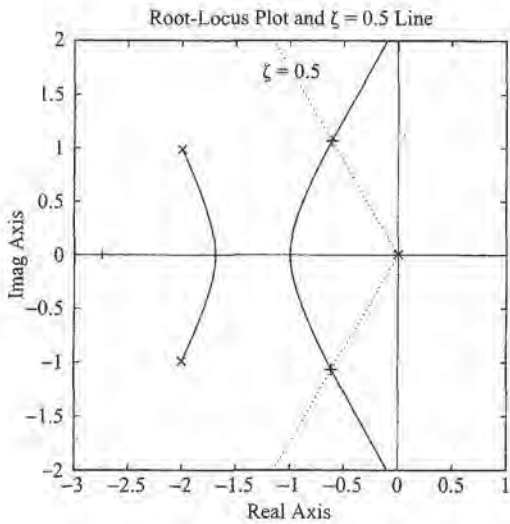
```
sgrid(0.5, [])
```

See MATLAB Program 6-7 and the resulting plot shown in Figure 6-24.

MATLAB Program 6-7

```
num = [0 0 0 1];
den = [1 4 5 0];
rlocus(num, den)
v = [-3 1 -2 2]; axis(v); axis('square')
sgrid(0.5, [])
title('Root-Locus Plot and \zeta = 0.5 Line')
gtext('\zeta = 0.5')
```

Figure 6-24
Root-locus plot with $\zeta = 0.5$ line.



Orthogonality of Root Loci and Constant-Gain Loci. Consider the system whose open-loop transfer function is $G(s)H(s)$. In the $G(s)H(s)$ plane, the loci of $|G(s)H(s)| = \text{constant}$ are circles centered at the origin, and the loci corresponding to $\angle G(s)H(s) = \pm 180^\circ(2k + 1)$ ($k = 0, 1, 2, \dots$) lie on the negative real axis of the $G(s)H(s)$ plane, as shown in Figure 6-25. [Note that the complex plane employed here is not the s plane, but the $G(s)H(s)$ plane.]

The root loci and constant-gain loci in the s plane are conformal mappings of the loci of $\angle G(s)H(s) = \pm 180^\circ(2k + 1)$ and of $|G(s)H(s)| = \text{constant}$ in the $G(s)H(s)$ plane.

Since the constant-phase and constant-gain loci in the $G(s)H(s)$ plane are orthogonal, the root loci and constant-gain loci in the s plane are orthogonal. Figure 6-26(a) shows the root loci and constant-gain loci for the following system:

$$G(s) = \frac{K(s + 2)}{s^2 + 2s + 3}, \quad H(s) = 1$$

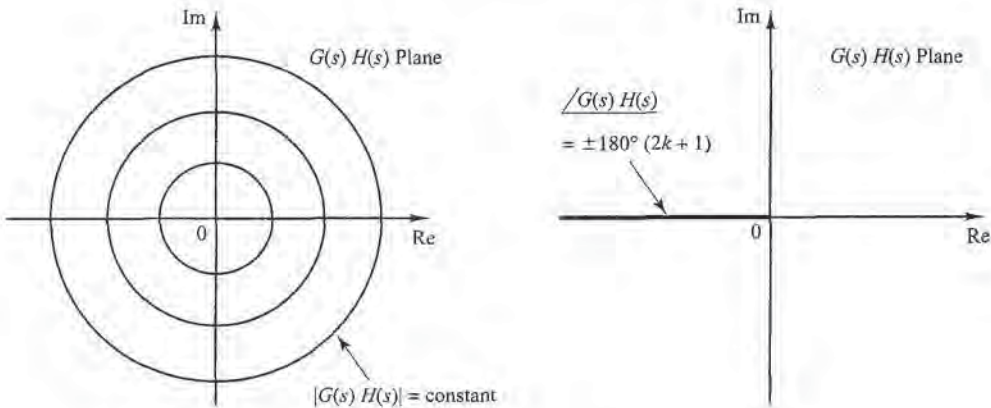


Figure 6-25
Plots of constant-gain and constant-phase loci in the $G(s)H(s)$ plane.

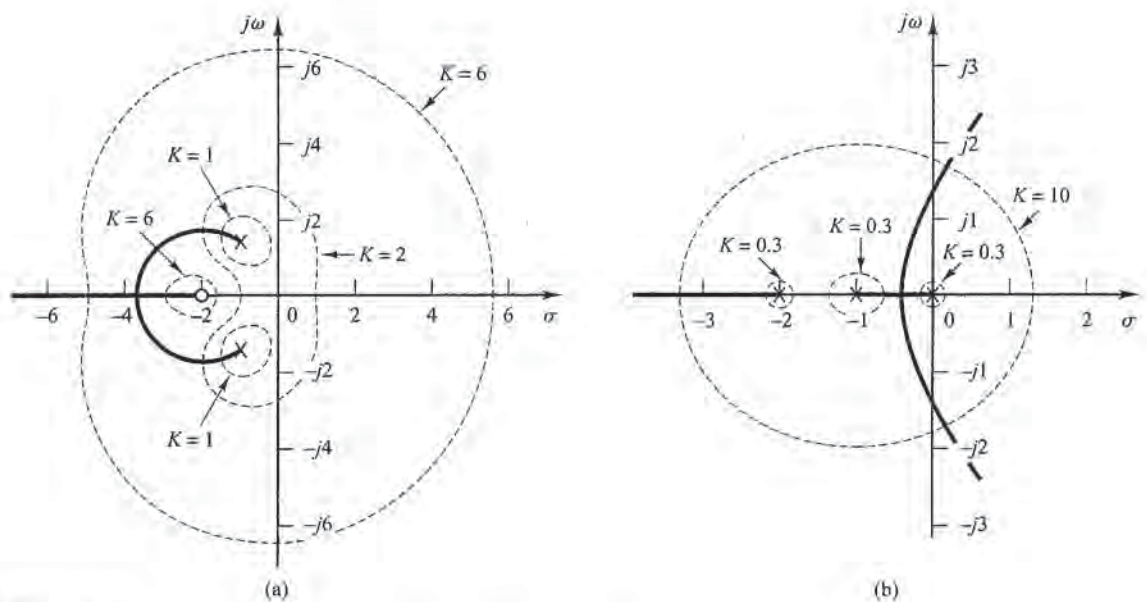


Figure 6-26
Plots of root loci and constant-gain loci. (a) System with $G(s) = K(s + 2)/(s^2 + 2s + 3)$, $H(s) = 1$; (b) system with $G(s) = K/[s(s + 1)(s + 2)]$, $H(s) = 1$.

Notice that since the pole-zero configuration is symmetrical about the real axis the constant-gain loci are also symmetrical about the real axis.

Figure 6-26(b) shows the root loci and constant-gain loci for the system:

$$G(s) = \frac{K}{s(s + 1)(s + 2)}, \quad H(s) = 1$$

Notice that since the configuration of the poles in the s plane is symmetrical about the real axis and the line parallel to the imaginary axis passing through point ($\sigma = -1$, $\omega = 0$), the constant-gain loci are symmetrical about the $\omega = 0$ line (real axis) and the $\sigma = -1$ line.

From Figures 6-26(a) and (b), notice that every point in the s plane has the corresponding K value. If we use a command `rlocfind` (presented next), MATLAB will give the K value of the specified point as well as the nearest closed-loop poles corresponding to this K value.

Finding the Gain Value K at an Arbitrary Point on the Root Loci. In MATLAB analysis of closed-loop systems, it is frequently desired to find the gain value K at an arbitrary point on the root locus. This can be accomplished by using the following `rlocfind` command:

$$[K, r] = \text{rlocfind}(\text{num}, \text{den})$$

The `rlocfind` command, which must follow an `rlocus` command, overlays movable x - y coordinates on the screen. Using the mouse, we position the origin of the x - y coordinates over the desired point on the root locus and press the mouse button. Then MATLAB

displays on the screen the coordinates of that point, the gain value at that point, and the closed-loop poles corresponding to this gain value.

If the selected point is not on the root locus, the rlocfind command gives the coordinates of this selected point, the gain value of this point, and the locations of the closed-loop poles corresponding to this K value. [Note that every point on the s plane has a gain value. See, for example, Figures 6–26 (a) and (b).]

EXAMPLE 6–6 Consider the unity-feedback control system with the following feedforward transfer function:

$$G(s) = \frac{K}{s(s^2 + 4s + 5)}$$

Plot the root loci with MATLAB. Determine closed-loop poles that have the damping ratio of 0.5. Find the gain value K at this point.

We first plot a root-locus diagram as shown in Figure 6–27. Then enter the rlocfind command as shown in MATLAB Program 6–8. Position the origin of the x - y coordinates over the intersection of the upper root-locus branch and the $\zeta = 0.5$ line. Then press the button of the mouse. The screen shows the coordinates of this point, the gain value at this point, and the closed-loop poles corresponding to this gain value.

The plot shows the closed-loop poles by a plus sign (+). The three closed-loop poles obtained are

$$s = -2.7474, \quad s = -0.6263 + j1.0800, \quad s = -0.6263 - j1.0800$$

Note that the three closed-loop poles are slightly off the exact locations obtained by the analytic method. The reason is that we cannot position the origin of the movable x - y coordinates exactly at the intersection of the upper root-locus branch and the $\zeta = 0.5$ line.

MATLAB Program 6–8

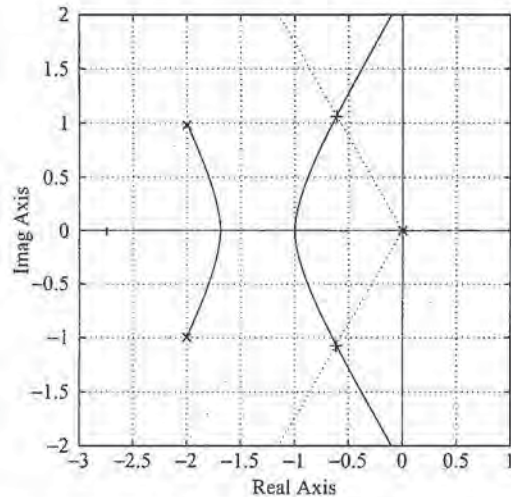
```
num = [0 0 0 1];
den = [1 4 5 0];
rlocus(num, den);
v = [-3 1 -2 2]; axis(v); axis('square')
sgrid(0.5, [])
[K,r] = rlocfind(num, den)
Select a point in the graphics window

selected_point =
-0.6246 + 1.0792i

K =
4.2823

r =
-2.7474
-0.6263 + 1.0800i
-0.6263 - 1.0800i
```

Figure 6-27
Root-locus plot with
 $\zeta = 0.5$ line.



Nonminimum-Phase Systems. If all the poles and zeros of a system lie in the left-half s plane, then the system is called *minimum phase*. If a system has at least one pole or zero in the right-half s plane, then the system is called *nonminimum phase*. The term nonminimum phase comes from the phase-shift characteristics of such a system when subjected to sinusoidal inputs.

Consider the system shown in Figure 6-28(a). For this system

$$G(s) = \frac{K(1 - T_a s)}{s(Ts + 1)} \quad (T_a > 0), \quad H(s) = 1$$

This is a nonminimum-phase system since there is one zero in the right-half s plane. For this system, the angle condition becomes

$$\begin{aligned} \angle G(s) &= \angle \left[-\frac{K(T_a s - 1)}{s(Ts + 1)} \right] \\ &= \angle \frac{K(T_a s - 1)}{s(Ts + 1)} + 180^\circ \\ &= \pm 180^\circ(2k + 1) \quad (k = 0, 1, 2, \dots) \end{aligned}$$

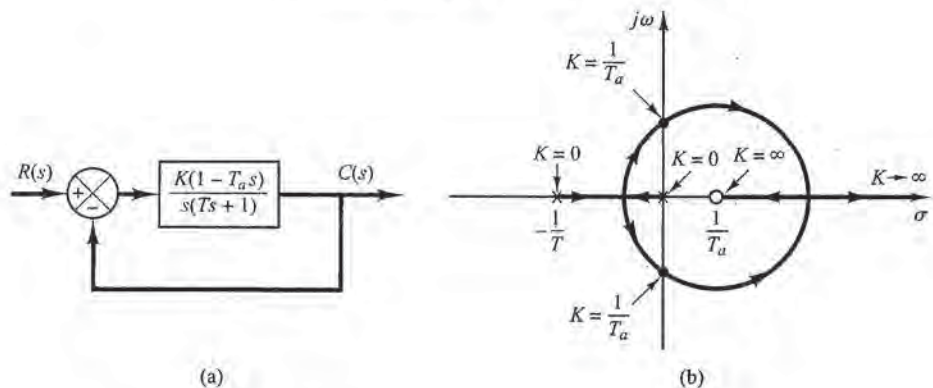


Figure 6-28
(a) Nonminimum-
phase system;
(b) root-locus plot.

or

$$\angle \frac{K(T_a s - 1)}{s(Ts + 1)} = 0^\circ \tag{6-16}$$

The root loci can be obtained from Equation (6-16). Figure 6-28(b) shows a root-locus plot for this system. From the diagram, we see that the system is stable if the gain K is less than $1/T_a$.

To obtain a root-locus plot with MATLAB, enter the numerator and denominator as usual. For example, if $T = 1$ sec and $T_a = 0.5$ sec, enter the following num and den in the program:

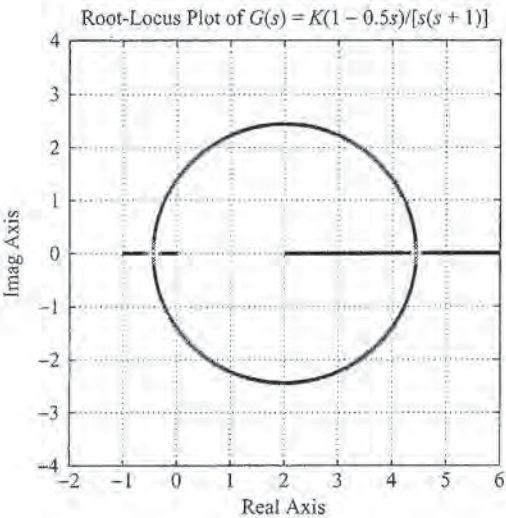
```
num = [0 -0.5 1]
den = [1 1 0]
```

MATLAB Program 6-9 gives the plot of the root loci shown in Figure 6-29.

MATLAB Program 6-9

```
num = [0 -0.5 1];
den = [1 1 0];
k1 = 0:0.01:30;
k2 = 30:1:100;
K3 = 100:5:500;
K = [k1 k2 k3];
rlocus(num,den,K)
v = [-2 6 -4 4]; axis(v); axis('square')
grid
title('Root-Locus Plot of G(s) = K(1 - 0.5s)/[s(s + 1)]')
```

Figure 6-29
Root-locus plot of
 $G(s) = \frac{K(1 - 0.5s)}{s(s + 1)}$



6-5 POSITIVE-FEEDBACK SYSTEMS

Root Loci for Positive-Feedback Systems.* In a complex control system, there may be a positive-feedback inner loop as shown in Figure 6-30. Such a loop is usually stabilized by the outer loop. In what follows, we shall be concerned only with the positive-feedback inner loop. The closed-loop transfer function of the inner loop is

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 - G(s)H(s)}$$

The characteristic equation is

$$1 - G(s)H(s) = 0 \tag{6-17}$$

This equation can be solved in a manner similar to the development of the root-locus method in Section 6-2. The angle condition, however, must be altered.

Equation (6-17) can be rewritten as

$$G(s)H(s) = 1$$

which is equivalent to the following two equations:

$$\angle G(s)H(s) = 0^\circ \pm k360^\circ \quad (k = 0, 1, 2, \dots)$$

$$|G(s)H(s)| = 1$$

The total sum of all angles from the open-loop poles and zeros must be equal to $0^\circ \pm k360^\circ$. Thus the root locus follows a 0° locus in contrast to the 180° locus considered previously. The magnitude condition remains unaltered.

To illustrate the root-locus plot for the positive-feedback system, we shall use the following transfer functions $G(s)$ and $H(s)$ as an example.

$$G(s) = \frac{K(s + 2)}{(s + 3)(s^2 + 2s + 2)}, \quad H(s) = 1$$

The gain K is assumed to be positive.

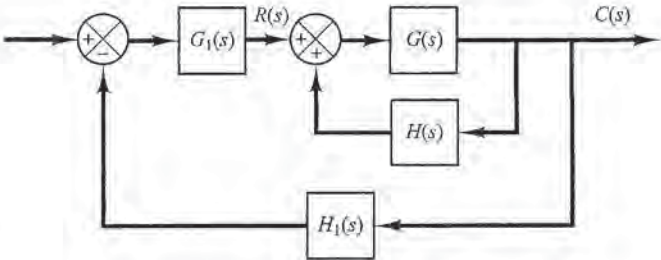


Figure 6-30
Control system.

* Reference W-4

The general rules for constructing root loci given in Section 6-3 must be modified in the following way:

Rule 2 is Modified as Follows: If the total number of real poles and real zeros to the right of a test point on the real axis is even, then this test point lies on the root locus.

Rule 3 is Modified as Follows:

$$\text{Angles of asymptotes} = \frac{\pm k 360^\circ}{n - m} \quad (k = 0, 1, 2, \dots)$$

where n = number of finite poles of $G(s)H(s)$

m = number of finite zeros of $G(s)H(s)$

Rule 5 is Modified as Follows: When calculating the angle of departure (or angle of arrival) from a complex open-loop pole (or at a complex zero), subtract from 0° the sum of all angles of the vectors from all the other poles and zeros to the complex pole (or complex zero) in question, with appropriate signs included.

Other rules for constructing the root-locus plot remain the same. We shall now apply the modified rules to construct the root-locus plot.

1. Plot the open-loop poles ($s = -1 + j$, $s = -1 - j$, $s = -3$) and zero ($s = -2$) in the complex plane. As K is increased from 0 to ∞ , the closed-loop poles start at the open-loop poles and terminate at the open-loop zeros (finite or infinite), just as in the case of negative-feedback systems.
2. Determine the root loci on the real axis. Root loci exist on the real axis between -2 and $+\infty$ and between -3 and $-\infty$.
3. Determine the asymptotes of the root loci. For the present system,

$$\text{Angles of asymptote} = \frac{\pm k 360^\circ}{3 - 1} = \pm 180^\circ$$

This simply means that asymptotes are on the real axis.

4. Determine the breakaway and break-in points. Since the characteristic equation is

$$(s + 3)(s^2 + 2s + 2) - K(s + 2) = 0$$

we obtain

$$K = \frac{(s + 3)(s^2 + 2s + 2)}{s + 2}$$

By differentiating K with respect to s , we obtain

$$\frac{dK}{ds} = \frac{2s^3 + 11s^2 + 20s + 10}{(s + 2)^2}$$

Note that

$$\begin{aligned} 2s^3 + 11s^2 + 20s + 10 &= 2(s + 0.8)(s^2 + 4.7s + 6.24) \\ &= 2(s + 0.8)(s + 2.35 + j0.77)(s + 2.35 - j0.77) \end{aligned}$$

Point $s = -0.8$ is on the root locus. Since this point lies between two zeros (a finite zero and an infinite zero), it is an actual break-in point. Points $s = -2.35 \pm j0.77$ do not satisfy the angle condition and, therefore, they are neither breakaway nor break-in points.

5. Find the angle of departure of the root locus from a complex pole. For the complex pole at $s = -1 + j$, the angle of departure θ is

$$\theta = 0^\circ - 27^\circ - 90^\circ + 45^\circ$$

or

$$\theta = -72^\circ$$

(The angle of departure from the complex pole at $s = -1 - j$ is 72° .)

6. Choose a test point in the broad neighborhood of the $j\omega$ axis and the origin and apply the angle condition. Locate a sufficient number of points that satisfy the angle condition.

Figure 6-31 shows the root loci for the given positive-feedback system. The root loci are shown with dashed lines and a curve.

Note that if

$$K > \frac{(s + 3)(s^2 + 2s + 2)}{s + 2} \bigg|_{s=0} = 3$$

one real root enters the right-half s plane. Hence, for values of K greater than 3, the system becomes unstable. (For $K > 3$, the system must be stabilized with an outer loop.)

Figure 6-31
Root-locus plot for the positive-feedback system with
 $G(s) = K(s + 2)/[(s + 3)(s^2 + 2s + 2)]$,
 $H(s) = 1$.

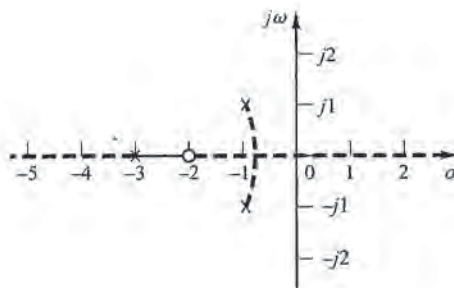
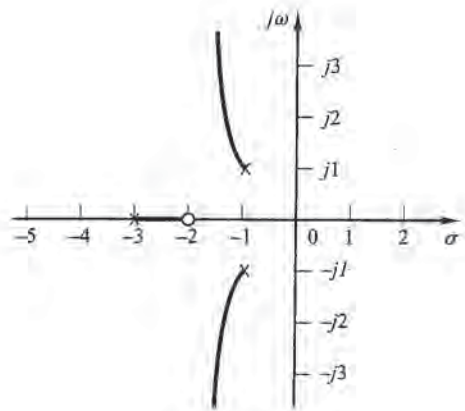


Figure 6-32
 Root-locus plot for the negative-feedback system with
 $G(s) = K(s + 2)/[(s + 3)(s^2 + 2s + 2)]$,
 $H(s) = 1$.



Note that the closed-loop transfer function for the positive-feedback system is given by

$$\begin{aligned} \frac{C(s)}{R(s)} &= \frac{G(s)}{1 - G(s)H(s)} \\ &= \frac{K(s + 2)}{(s + 3)(s^2 + 2s + 2) - K(s + 2)} \end{aligned}$$

To compare this root-locus plot with that of the corresponding negative-feedback system, we show in Figure 6-32 the root loci for the negative-feedback system whose closed-loop transfer function is

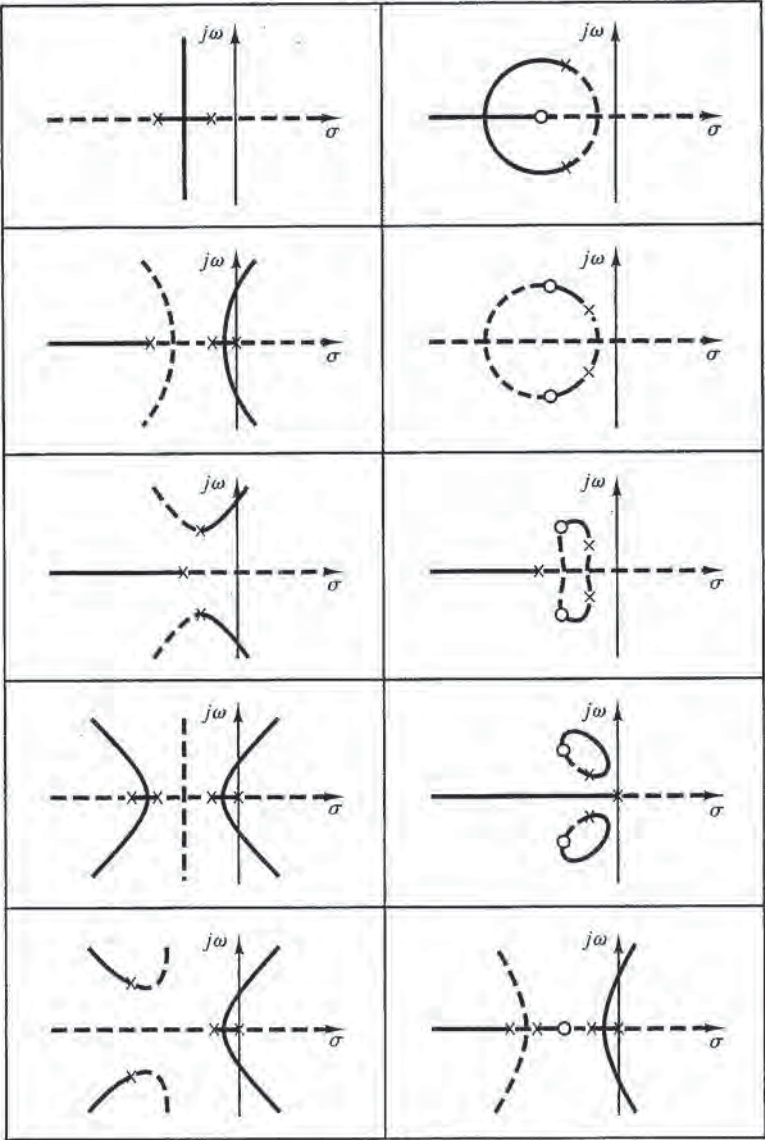
$$\frac{C(s)}{R(s)} = \frac{K(s + 2)}{(s + 3)(s^2 + 2s + 2) + K(s + 2)}$$

Table 6-2 shows various root-locus plots of negative-feedback and positive-feedback systems. The closed-loop transfer functions are given by

$$\begin{aligned} \frac{C}{R} &= \frac{G}{1 + GH}, & \text{for negative-feedback systems} \\ \frac{C}{R} &= \frac{G}{1 - GH}, & \text{for positive-feedback systems} \end{aligned}$$

where GH is the open-loop transfer function. In Table 6-2, the root loci for negative-feedback systems are drawn with heavy lines and curves, and those for positive-feedback systems are drawn with dashed lines and curves.

Table 6-2 Root-Locus Plots of Negative-Feedback and Positive-Feedback Systems



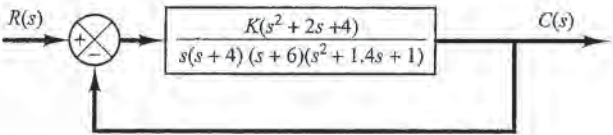
Heavy lines and curves correspond to negative-feedback systems; dashed lines and curves correspond to positive-feedback systems.

6-6 CONDITIONALLY STABLE SYSTEMS

Consider the system shown in Figure 6-33. We can plot the root loci for this system by applying the general rules and procedure for constructing root loci, or use MATLAB to get root-locus plots. MATLAB Program 6-10 will plot the root-locus diagram for the system. The plot is shown in Figure 6-34.

It can be seen from the root-locus plot of Figure 6-34 that this system is stable only for limited ranges of the value of K —that is, $0 < K < 12$ and $73 < K < 154$. The system becomes unstable for $12 < K < 73$ and $154 < K$. (If K assumes a value corre-

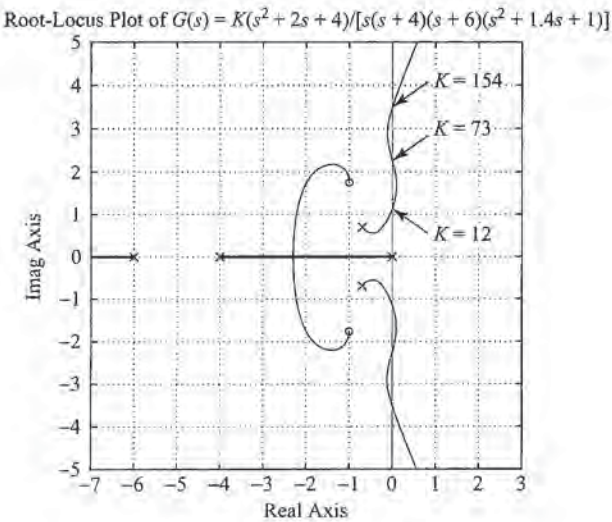
Figure 6-33
Control system.



MATLAB Program 6-10

```
num = [0 0 0 1 2 4];
den = conv(conv([1 4 0],[1 6]), [1 1.4 1]);
rlocus(num, den)
v = [-7 3 -5 5]; axis(v); axis('square')
grid
title('Root-Locus Plot of G(s) = K(s^2 + 2s + 4)/[s(s + 4)(s + 6)(s^2 + 1.4s + 1)]')
text(1.0, 0.55, 'K = 12')
text(1.0, 3.0, 'K = 73')
text(1.0, 4.15, 'K = 154')
```

Figure 6-34
Root-locus plot of
conditionally stable
system.



sponding to unstable operation, the system may break down or may become nonlinear due to a saturation nonlinearity that may exist.) Such a system is called conditionally stable.

In practice, conditionally stable systems are not desirable. Conditional stability is dangerous but does occur in certain systems, in particular, a system that has an unstable feedforward path. Such an unstable feedforward path may occur if the system has a minor loop. It is advisable to avoid such conditional stability since, if the gain drops beyond the critical value for any reason, the system becomes unstable. Note that the addition of a proper compensating network will eliminate conditional stability. [An addition of a zero will cause the root loci to bend to the left. (See Section 7-2.) Hence conditional stability may be eliminated by adding proper compensation.]

6-7 ROOT LOCI FOR SYSTEMS WITH TRANSPORT LAG

Figure 6-35 shows a thermal system in which hot air is circulated to keep the temperature of a chamber constant. In this system, the measuring element is placed downstream a distance L ft from the furnace, the air velocity is v ft/sec, and $T = L/v$ sec would elapse before any change in the furnace temperature is sensed by the thermometer. Such a delay in measuring, delay in controller action, or delay in actuator operation, and the like, is called *transport lag* or *dead time*. Dead time is present in most process control systems.

The input $x(t)$ and the output $y(t)$ of a transport-lag or dead-time element are related by

$$y(t) = x(t - T)$$

where T is dead time. The transfer function of transport lag or dead time is given by

$$\begin{aligned} \text{Transfer function of transport lag or dead time} &= \frac{\mathcal{L}[x(t - T)1(t - T)]}{\mathcal{L}[x(t)1(t)]} \\ &= \frac{X(s)e^{-Ts}}{X(s)} = e^{-Ts} \end{aligned}$$

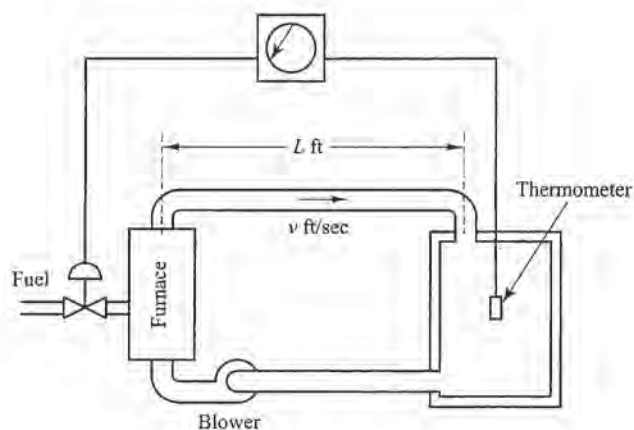
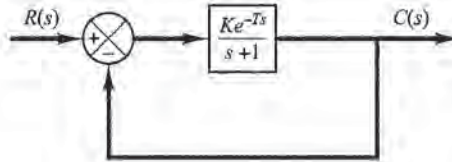


Figure 6-35
Thermal system.

Figure 6-36
Block diagram of the
system shown in
Figure 6-35.



Suppose that the feedforward transfer function of this thermal system can be approximated by

$$G(s) = \frac{Ke^{-Ts}}{s+1}$$

as shown in Figure 6-36. Let us construct a root-locus plot for this system. The characteristic equation for this closed-loop system is

$$1 + \frac{Ke^{-Ts}}{s+1} = 0 \quad (6-18)$$

It is noted that for systems with transport lag the rules of construction presented earlier need to be modified. For example, the number of the root-locus branches is infinite, since the characteristic equation has an infinite number of roots. The number of asymptotes is infinite. They are all parallel to the real axis of the s plane, as will be seen later.

From Equation (6-18), we obtain

$$\frac{Ke^{-Ts}}{s+1} = -1$$

Thus, the angle condition becomes

$$\angle \frac{Ke^{-Ts}}{s+1} = \angle e^{-Ts} - \angle s+1 = \pm 180^\circ(2k+1) \quad (k = 0, 1, 2, \dots) \quad (6-19)$$

To find the angle of e^{-Ts} , substitute $s = \sigma + j\omega$. Then we obtain

$$e^{-Ts} = e^{-T\sigma - j\omega T}$$

Since $e^{-T\sigma}$ is a real quantity, the angle of $e^{-T\sigma}$ is zero. Hence

$$\begin{aligned} \angle e^{-Ts} &= \angle e^{-j\omega T} = \angle \cos \omega T - j \sin \omega T \\ &= -\omega T \quad (\text{radians}) \\ &= -57.3\omega T \quad (\text{degrees}) \end{aligned}$$

Since T is a given constant, the angle of e^{-Ts} is a function of ω only. The angle condition, Equation (6-19), then becomes

$$-57.3\omega T - \angle s+1 = \pm 180^\circ(2k+1)$$

We shall next determine the angle contribution due to e^{-Ts} as given by Equation (6-19). For $k = 0$, the angle condition may be written

$$\angle s+1 = \pm 180^\circ - 57.3\omega T \quad (6-20)$$

Since the angle contribution of e^{-Ts} is zero for $\omega = 0$, the real axis from -1 to $-\infty$ forms a part of the root loci. Now assume a value ω_1 for ω and compute $57.3^\circ\omega_1 T$. At point -1 on the negative real axis, draw a line that makes an angle of $180^\circ - 57.3^\circ\omega_1 T$ with the real axis. Find the intersection of this line and the horizontal line $\omega = \omega_1$. This intersection, point P in Figure 6-37(a), is a point satisfying Equation (6-20) and hence is on a root locus. Continuing the same process, we obtain the root-locus plot as shown in Figure 6-37(b).

Note that as s approaches minus infinity, the open-loop transfer function

$$\frac{Ke^{-Ts}}{s+1}$$

approaches minus infinity since

$$\begin{aligned}\lim_{s \rightarrow -\infty} \frac{Ke^{-Ts}}{s+1} &= \frac{\frac{d}{ds}(Ke^{-Ts})}{\frac{d}{ds}(s+1)} \bigg|_{s=-\infty} \\ &= -KTe^{-Ts} \bigg|_{s=-\infty} \\ &= -\infty\end{aligned}$$

Therefore, $s = -\infty$ is a pole of the open-loop transfer function. Thus, root loci start from $s = -1$ or $s = -\infty$ and terminate at $s = \infty$, as K increases from zero to infinity. Since the right-hand side of the angle condition given by Equation (6-19) has an infinite number of values, there are an infinite number of root loci, as the value of k ($k = 0, 1, 2, \dots$) goes from zero to infinity. For example, if $k = 1$, the angle condition becomes

$$\begin{aligned}\angle s+1 &= \pm 540^\circ - 57.3^\circ\omega T \quad (\text{degrees}) \\ &= \pm 3\pi - \omega T \quad (\text{radians})\end{aligned}$$

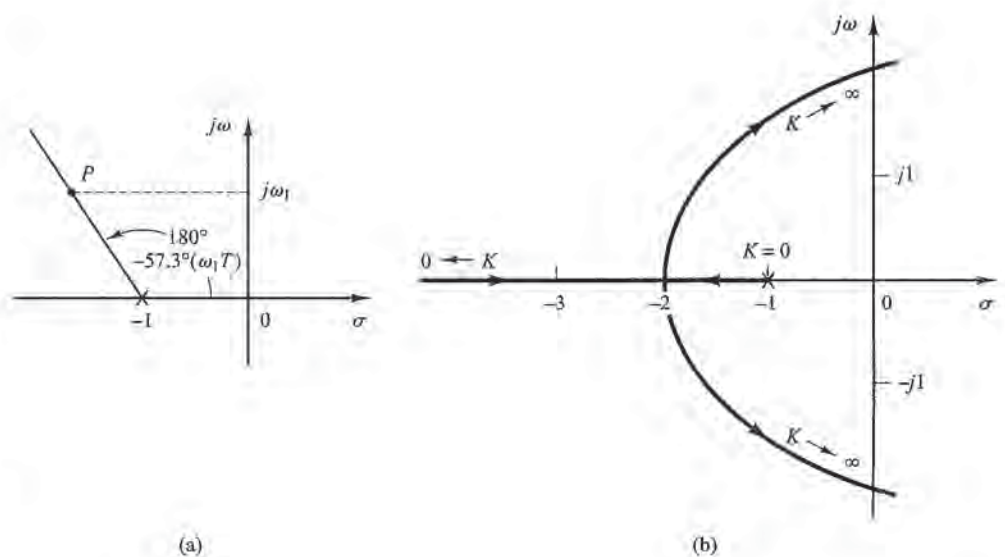


Figure 6-37
(a) Construction of the root locus;
(b) root-locus plot.

The construction of the root loci for $k = 1$ is the same as that for $k = 0$. A plot of root loci for $k = 0, 1$, and 2 when $T = 1$ sec is shown in Figure 6-38.

The magnitude condition states that

$$\left| \frac{Ke^{-Ts}}{s+1} \right| = 1$$

Since the magnitude of e^{-Ts} is equal to that of $e^{-T\sigma}$ or

$$|e^{-Ts}| = |e^{-T\sigma}| \cdot |e^{-j\omega T}| = e^{-T\sigma}$$

the magnitude condition becomes

$$|s+1| = Ke^{-T\sigma}$$

The root loci shown in Figure 6-38 are graduated in terms of K when $T = 1$ sec.

Although there are an infinite number of root-locus branches, the primary branch that lies between $-j\pi$ and $j\pi$ is most important. Referring to Figure 6-38, the critical value of K at the primary branch is equal to 2, while the critical values of K at other branches are much higher (8, 14, ...). Therefore, the critical value $K = 2$ on the primary branch is most significant from the stability viewpoint. The transient response of the system is determined by the roots located closest to the $j\omega$ axis and lie on the primary branch. In summary, the root-locus branch corresponding to $k = 0$ is the dominant one; other branches corresponding to $k = 1, 2, 3, \dots$ are not so important and may be neglected.

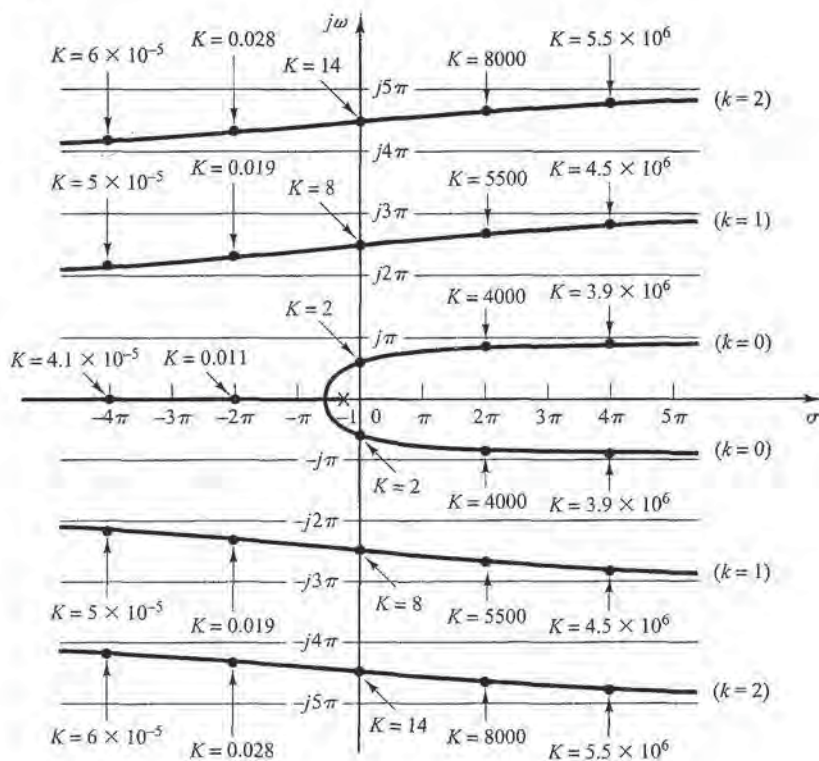


Figure 6-38
Root-locus plot for
the system shown in
Figure 6-36
($T = 1$ sec).

This example illustrates the fact that dead time can cause instability even in the first-order system because the root loci enter the right-half s plane for large values of K . Therefore, although the gain K of the first-order system can be set at a high value in the absence of dead time, it cannot be set too high if dead time is present. (For the system considered here, the value of gain K must be considerably less than 2 for a satisfactory operation.)

Approximation of Transport Lag or Dead Time. If the dead time T is very small, then e^{-Ts} is frequently approximated by

$$e^{-Ts} \doteq 1 - Ts$$

or

$$e^{-Ts} \doteq \frac{1}{Ts + 1}$$

Such approximations are good if the dead time is very small and, in addition, the input time function $f(t)$ to the dead-time element is smooth and continuous. [This means that the second- and higher-order derivatives of $f(t)$ are small.]

A more elaborate expression to approximate e^{-Ts} is available and is

$$e^{-Ts} = \frac{1 - \frac{Ts}{2} + \frac{(Ts)^2}{8} - \frac{(Ts)^3}{48} + \dots}{1 + \frac{Ts}{2} + \frac{(Ts)^2}{8} + \frac{(Ts)^3}{48} + \dots}$$

If only the first two terms in the numerator and denominator are taken, then

$$e^{-Ts} \doteq \frac{1 - \frac{Ts}{2}}{1 + \frac{Ts}{2}} = \frac{2 - Ts}{2 + Ts}$$

This approximation is also used frequently.

MATLAB Approximation of Dead Time. To handle dead time e^{-sT} , MATLAB uses the pate approximation. For example, if $T = 0.1$ sec, then using the third-order transfer function as an approximation to e^{-sT} , enter the following MATLAB program into the computer.

```
[num,den] = pate(0.1, 3);
printsys(num, den, 's')
num/den =
-1s^3 + 120s^2 - 6000s + 120000
s^3 + 120s^2 + 6000s + 120000
```

Similarly, the program for the fourth-order transfer function approximation with $T = 0.1$ sec is

```
[num,den] = pade(0.1, 4);
printsys(num, den, 's')

num/den =
```

$$\frac{s^4 - 200s^3 + 18000s^2 - 840000s + 16800000}{s^4 + 200s^3 + 18000s^2 + 840000s + 16800000}$$

Notice that the pade approximation depends on the dead time T and the desired order for the approximating transfer function.

EXAMPLE PROBLEMS AND SOLUTIONS

A-6-1. Sketch the root loci for the system shown in Figure 6-39(a). (The gain K is assumed to be positive.) Observe that for small or large values of K the system is overdamped and for medium values of K it is underdamped.

Solution. The procedure for plotting the root loci is as follows:

1. Locate the open-loop poles and zeros on the complex plane. Root loci exist on the negative real axis between 0 and -1 and between -2 and -3 .
2. The number of open-loop poles and that of finite zeros are the same. This means that there are no asymptotes in the complex region of the s plane.

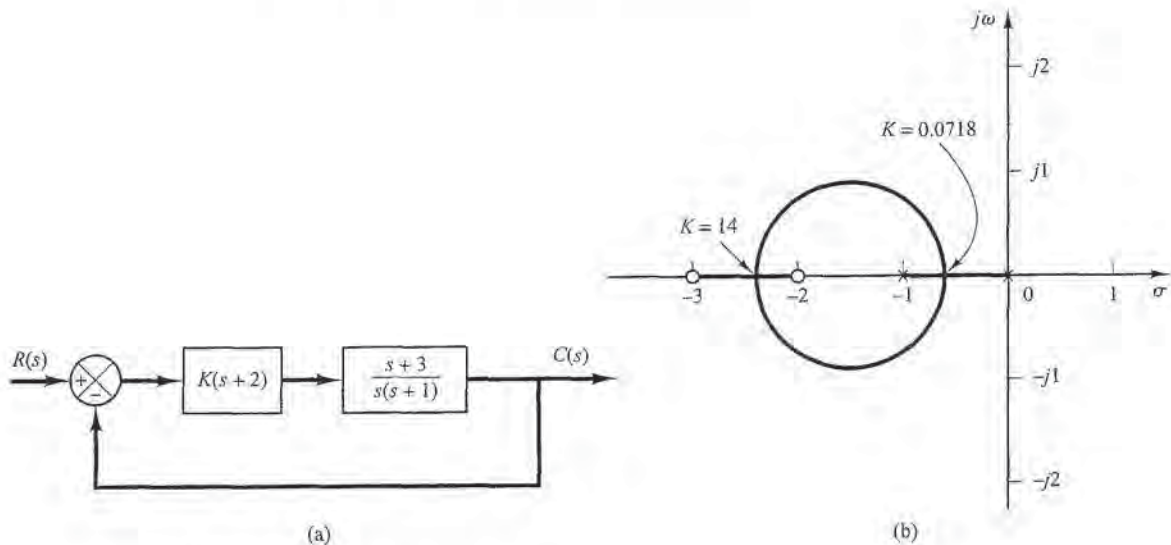


Figure 6-39
(a) Control system; (b) root-locus plot.

3. Determine the breakaway and break-in points. The characteristic equation for the system is

$$1 + \frac{K(s+2)(s+3)}{s(s+1)} = 0$$

or

$$K = -\frac{s(s+1)}{(s+2)(s+3)}$$

The breakaway and break-in points are determined from

$$\begin{aligned} \frac{dK}{ds} &= -\frac{(2s+1)(s+2)(s+3) - s(s+1)(2s+5)}{[(s+2)(s+3)]^2} \\ &= -\frac{4(s+0.634)(s+2.366)}{[(s+2)(s+3)]^2} \\ &= 0 \end{aligned}$$

as follows:

$$s = -0.634, \quad s = -2.366$$

Notice that both points are on root loci. Therefore, they are actual breakaway or break-in points. At point $s = -0.634$, the value of K is

$$K = -\frac{(-0.634)(0.366)}{(1.366)(2.366)} = 0.0718$$

Similarly, at $s = -2.366$,

$$K = -\frac{(-2.366)(-1.366)}{(-0.366)(0.634)} = 14$$

(Because point $s = -0.634$ lies between two poles, it is a breakaway point, and because point $s = -2.366$ lies between two zeros, it is a break-in point.)

4. Determine a sufficient number of points that satisfy the angle condition. (It can be found that the root loci involve a circle with center at -1.5 that passes through the breakaway and break-in points.) The root-locus plot for this system is shown in Figure 6-39(b).

Note that this system is stable for any positive value of K since all the root loci lie in the left-half s plane.

Small values of K ($0 < K < 0.0718$) correspond to an overdamped system. Medium values of K ($0.0718 < K < 14$) correspond to an underdamped system. Finally, large values of K ($14 < K$) correspond to an overdamped system. With a large value of K , the steady state can be reached in much shorter time than with a small value of K .

The value of K should be adjusted so that system performance is optimum according to a given performance index.

A-6-2. Sketch the root loci of the control system shown in Figure 6-40(a).

Solution. The open-loop poles are located at $s = 0$, $s = -3 + j4$, and $s = -3 - j4$. A root locus branch exists on the real axis between the origin and $-\infty$. There are three asymptotes for the root loci. The angles of asymptotes are

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k + 1)}{3} = 60^\circ, -60^\circ, 180^\circ$$

Referring to Equation (6-13), the intersection of the asymptotes and the real axis is obtained as

$$s = -\frac{0 + 3 + 3}{3} = -2$$

Next we check the breakaway and break-in points. For this system we have

$$K = -s(s^2 + 6s + 25)$$

Now we set

$$\frac{dK}{ds} = -(3s^2 + 12s + 25) = 0$$

which yields

$$s = -2 + j2.0817, \quad s = -2 - j2.0817$$

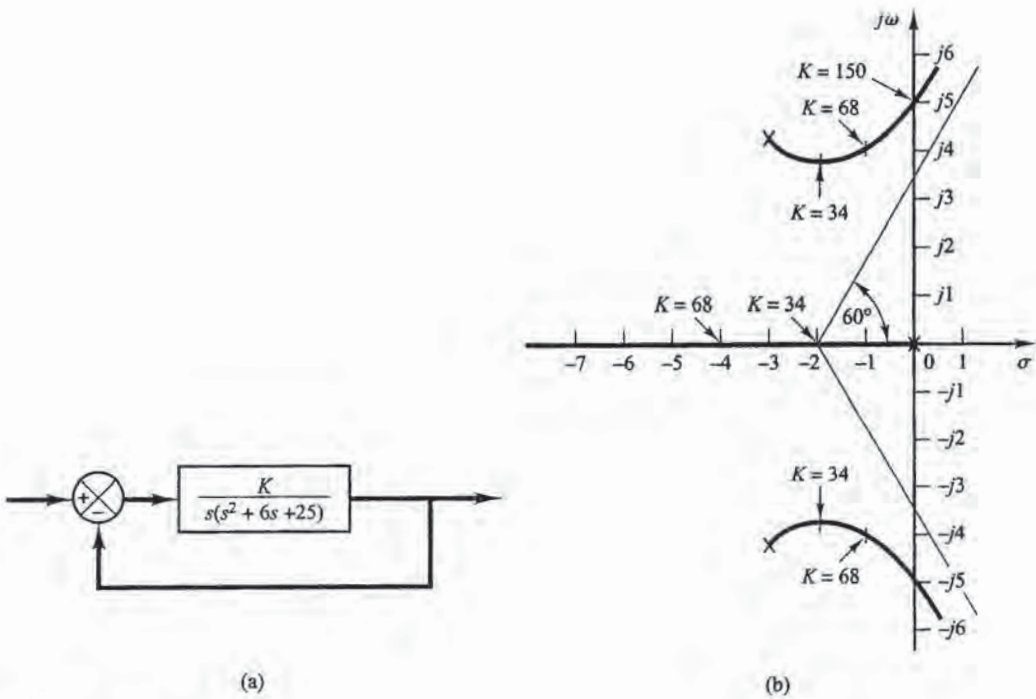


Figure 6-40
(a) Control system; (b) root-locus plot.

Notice that at points $s = -2 \pm j2.0817$ the angle condition is not satisfied. Hence, they are neither breakaway nor break-in points. In fact, if we calculate the value of K , we obtain

$$K = -s(s^2 + 6s + 25) \Big|_{s=-2 \pm j2.0817} = 34 \pm j18.04$$

(To be an actual breakaway or break-in point, the corresponding value of K must be real and positive.)

The angle of departure from the complex pole in the upper half s plane is

$$\theta_p = 180^\circ - 126.87^\circ - 90^\circ$$

or

$$\theta = -36.87^\circ$$

The points where root-locus branches cross the imaginary axis may be found by substituting $s = j\omega$ into the characteristic equation and solving the equation for ω and K as follows: Noting that the characteristic equation is

$$s^3 + 6s^2 + 25s + K = 0$$

we have

$$(j\omega)^3 + 6(j\omega)^2 + 25(j\omega) + K = (-6\omega^2 + K) + j\omega(25 - \omega^2) = 0$$

which yields

$$\omega = \pm 5, \quad K = 150 \quad \text{or} \quad \omega = 0, \quad K = 0$$

Root-locus branches cross the imaginary axis at $\omega = 5$ and $\omega = -5$. The value of gain K at the crossing points is 150. Also, the root-locus branch on the real axis touches the imaginary axis at $\omega = 0$. Figure 6-40(b) shows a root-locus plot for the system.

It is noted that if the order of the numerator of $G(s)H(s)$ is lower than that of the denominator by two or more, and if some of the closed-loop poles move on the root locus toward the right as gain K is increased, then other closed-loop poles must move toward the left as gain K is increased. This fact can be seen clearly in this problem. If the gain K is increased from $K = 34$ to $K = 68$, the complex-conjugate closed-loop poles are moved from $s = -2 + j3.65$ to $s = -1 + j4$; the third pole is moved from $s = -2$ (which corresponds to $K = 34$) to $s = -4$ (which corresponds to $K = 68$). Thus, the movements of two complex-conjugate closed-loop poles to the right by one unit cause the remaining closed-loop pole (real pole in this case) to move to the left by two units.

- A-6-3.** Consider the system shown in Figure 6-41(a). Sketch the root loci for the system. Observe that for small or large values of K the system is underdamped and for medium values of K it is overdamped.

Solution. A root locus exists on the real axis between the origin and $-\infty$. The angles of asymptotes of the root-locus branches are obtained as

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k + 1)}{3} = 60^\circ, -60^\circ, -180^\circ$$

The intersection of the asymptotes and the real axis is located on the real axis at

$$s = -\frac{0 + 2 + 2}{3} = -1.3333$$

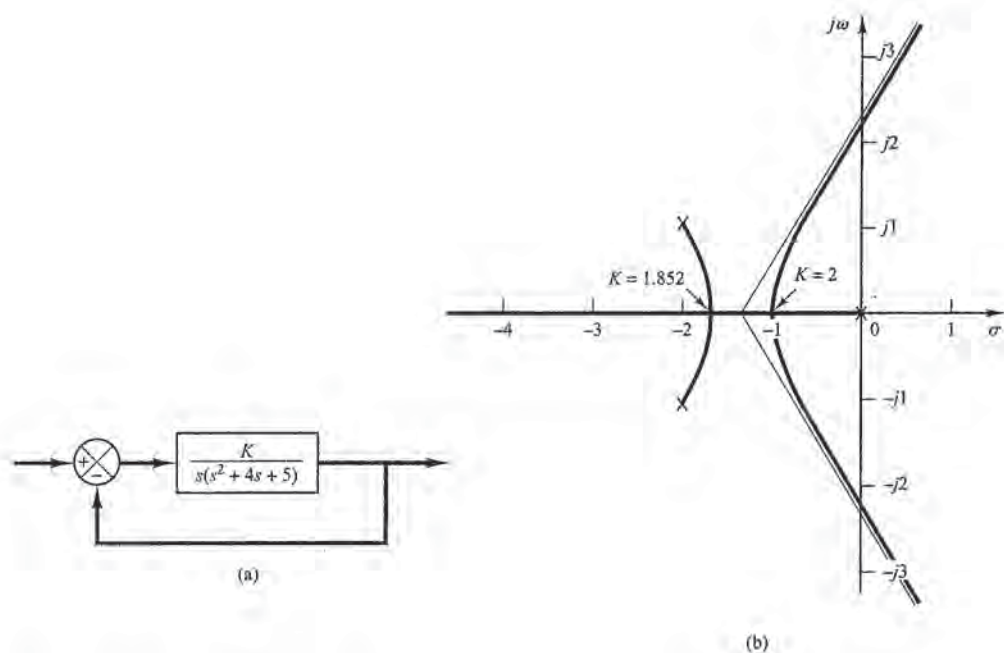


Figure 6-41
(a) Control system;
(b) root-locus plot.

The breakaway and break-in points are found from $dK/ds = 0$. Since the characteristic equation is

$$s^3 + 4s^2 + 5s + K = 0$$

we have

$$K = -(s^3 + 4s^2 + 5s)$$

Now we set

$$\frac{dK}{ds} = -(3s^2 + 8s + 5) = 0$$

which yields

$$s = -1, \quad s = -1.6667$$

Since these points are on root loci, they are actual breakaway or break-in points. (At point $s = -1$, the value of K is 2, and at point $s = -1.6667$, the value of K is 1.852.)

The angle of departure from a complex pole in the upper half s plane is obtained from

$$\theta = 180^\circ - 153.43^\circ = 90^\circ$$

or

$$\theta = -63.43^\circ$$

The root-locus branch from the complex pole in the upper half s plane breaks into the real axis at $s = -1.6667$.

Next we determine the points where root-locus branches cross the imaginary axis. By substituting $s = j\omega$ into the characteristic equation, we have

$$(j\omega)^3 + 4(j\omega)^2 + 5(j\omega) + K = 0$$

or

$$(K - 4\omega^2) + j\omega(5 - \omega^2) = 0$$

from which we obtain

$$\omega = \pm \sqrt{5}, \quad K = 20 \quad \text{or} \quad \omega = 0, \quad K = 0$$

Root-locus branches cross the imaginary axis at $\omega = \sqrt{5}$ and $\omega = -\sqrt{5}$. The root-locus branch on the real axis touches the $j\omega$ axis at $\omega = 0$. A sketch of the root loci for the system is shown in Figure 6-41(b).

Note that since this system is of third order, there are three closed-loop poles. The nature of the system response to a given input depends on the locations of the closed-loop poles.

For $0 < K < 1.852$, there are a set of complex-conjugate closed-loop poles and a real closed-loop pole. For $1.852 \leq K \leq 2$, there are three real closed-loop poles. For example, the closed-loop poles are located at

$$\begin{array}{llll} s = -1.667, & s = -1.667, & s = -0.667, & \text{for } K = 1.852 \\ s = -1, & s = -1, & s = -2, & \text{for } K = 2 \end{array}$$

For $2 < K$, there are a set of complex-conjugate closed-loop poles and a real closed-loop pole. Thus, small values of K ($0 < K < 1.852$) correspond to an underdamped system. (Since the real closed-loop pole dominates, only a small ripple may show up in the transient response.) Medium values of K ($1.852 \leq K \leq 2$) correspond to an overdamped system. Large values of K ($2 < K$) correspond to an underdamped system. With a large value of K , the system responds much faster than with a smaller value of K .

A-6-4. Sketch the root loci for the system shown in Figure 6-42(a).

Solution. The open-loop poles are located at $s = 0$, $s = -1$, $s = -2 + j3$, and $s = -2 - j3$. A root locus exists on the real axis between points $s = 0$ and $s = -1$. The angles of the asymptotes are found as follows:

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k + 1)}{4} = 45^\circ, -45^\circ, 135^\circ, -135^\circ$$

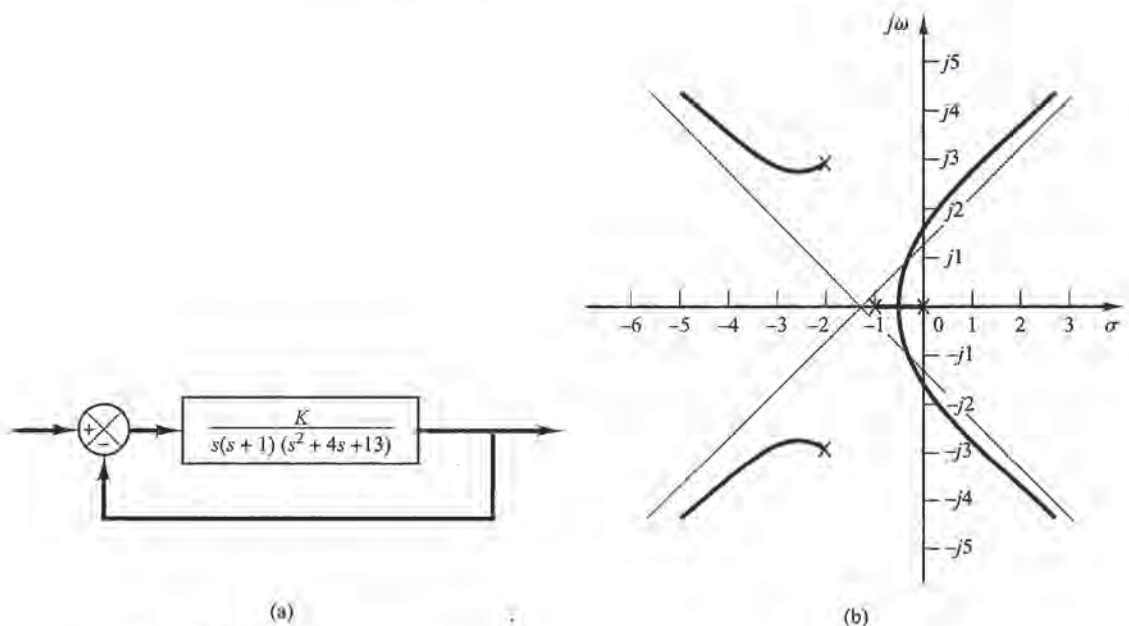


Figure 6-42
(a) Control system; (b) root-locus plot.

The intersection of the asymptotes and the real axis is found from

$$s = -\frac{0 + 1 + 2 + 2}{4} = -1.25$$

The breakaway and break-in points are found from $dK/ds = 0$. Noting that

$$K = -s(s + 1)(s^2 + 4s + 13) = -(s^4 + 5s^3 + 17s^2 + 13s)$$

we have

$$\frac{dK}{ds} = -(4s^3 + 15s^2 + 34s + 13) = 0$$

from which we get

$$s = -0.467, \quad s = -1.642 + j2.067, \quad s = -1.642 - j2.067$$

Point $s = -0.467$ is on a root locus. Therefore, it is an actual breakaway point. The gain values K corresponding to points $s = -1.642 \pm j2.067$ are complex quantities. Since the gain values are not real positive, these points are neither breakaway nor break-in points.

The angle of departure from the complex pole in the upper half s plane is

$$\theta = 180^\circ - 123.69^\circ - 108.44^\circ - 90^\circ$$

or

$$\theta = -142.13^\circ$$

Next we shall find the points where root loci may cross the $j\omega$ axis. Since the characteristic equation is

$$s^4 + 5s^3 + 17s^2 + 13s + K = 0$$

by substituting $s = j\omega$ into it we obtain

$$(j\omega)^4 + 5(j\omega)^3 + 17(j\omega)^2 + 13(j\omega) + K = 0$$

or

$$(K + \omega^4 - 17\omega^2) + j\omega(13 - 5\omega^2) = 0$$

from which we obtain

$$\omega = \pm 1.6125, \quad K = 37.44 \quad \text{or} \quad \omega = 0, \quad K = 0$$

The root-locus branches that extend to the right-half s plane cross the imaginary axis at $\omega = \pm 1.6125$. Also, the root-locus branch on the real axis touches the imaginary axis at $\omega = 0$. Figure 6-42(b) shows a sketch of the root loci for the system. Notice that each root-locus branch that extends to the right half s plane crosses its own asymptote.

A-6-5. Sketch the root loci for the system shown in Figure 6-43(a).

Solution. A root locus exists on the real axis between points $s = -1$ and $s = -3.6$. The asymptotes can be determined as follows:

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k + 1)}{3 - 1} = 90^\circ, -90^\circ$$

The intersection of the asymptotes and the real axis is found from

$$s = -\frac{0 + 0 + 3.6 - 1}{3 - 1} = -1.3$$

Since the characteristic equation is

$$s^3 + 3.6s^2 + K(s + 1) = 0$$

we have

$$K = -\frac{s^3 + 3.6s^2}{s + 1}$$

The breakaway and break-in points are found from

$$\frac{dK}{ds} = -\frac{(3s^2 + 7.2s)(s + 1) - (s^3 + 3.6s^2)}{(s + 1)^2} = 0$$

or

$$s^3 + 3.3s^2 + 3.6s = 0$$

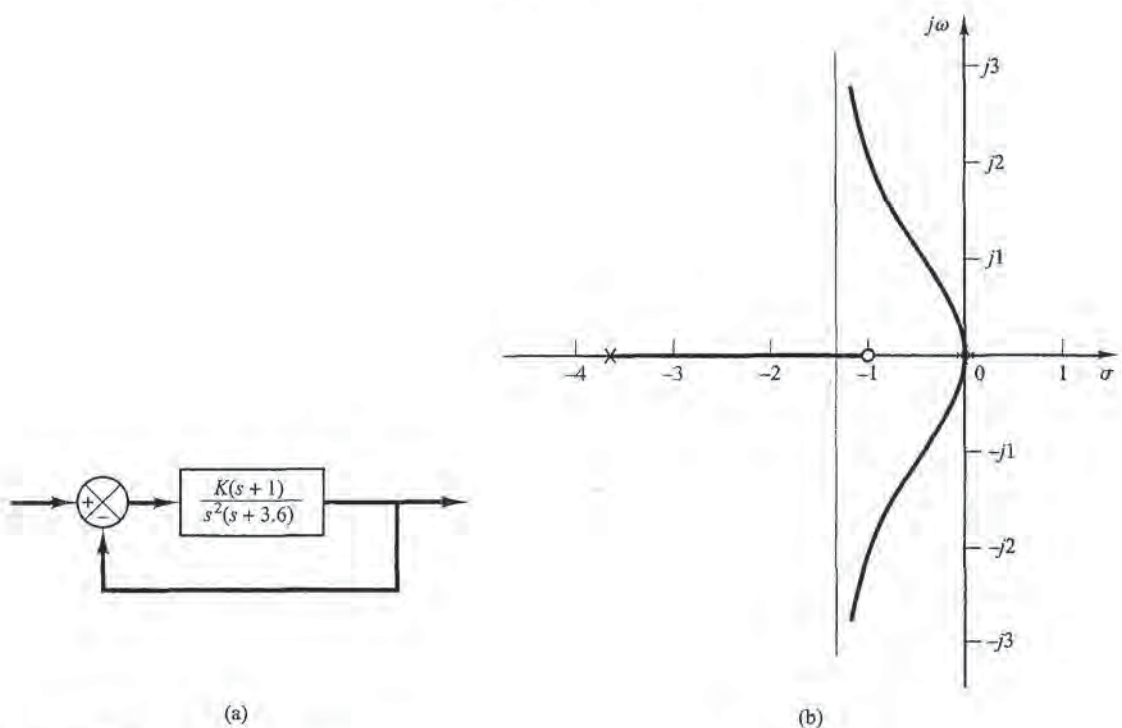


Figure 6-43
(a) Control system; (b) root-locus plot.

from which we get

$$s = 0, \quad s = -1.65 + j0.9367, \quad s = -1.65 - j0.9367$$

Point $s = 0$ corresponds to the actual breakaway point. But points $s = 1.65 \pm j0.9367$ are neither breakaway nor break-in points, because the corresponding gain values K become complex quantities.

To check the points where root-locus branches may cross the imaginary axis, substitute $s = j\omega$ into the characteristic equation, yielding.

$$(j\omega)^3 + 3.6(j\omega)^2 + Kj\omega + K = 0$$

or

$$(K - 3.6\omega^2) + j\omega(K - \omega^2) = 0$$

Notice that this equation can be satisfied only if $\omega = 0$, $K = 0$. Because of the presence of a double pole at the origin, the root locus is tangent to the $j\omega$ axis at $\omega = 0$. The root-locus branches do not cross the $j\omega$ axis. Figure 6-43(b) is a sketch of the root loci for this system.

A-6-6. Sketch the root loci for the system shown in Figure 6-44(a).

Solution. A root locus exists on the real axis between point $s = -0.4$ and $s = -3.6$. The angles of asymptotes can be found as follows:

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k + 1)}{3 - 1} = 90^\circ, -90^\circ$$

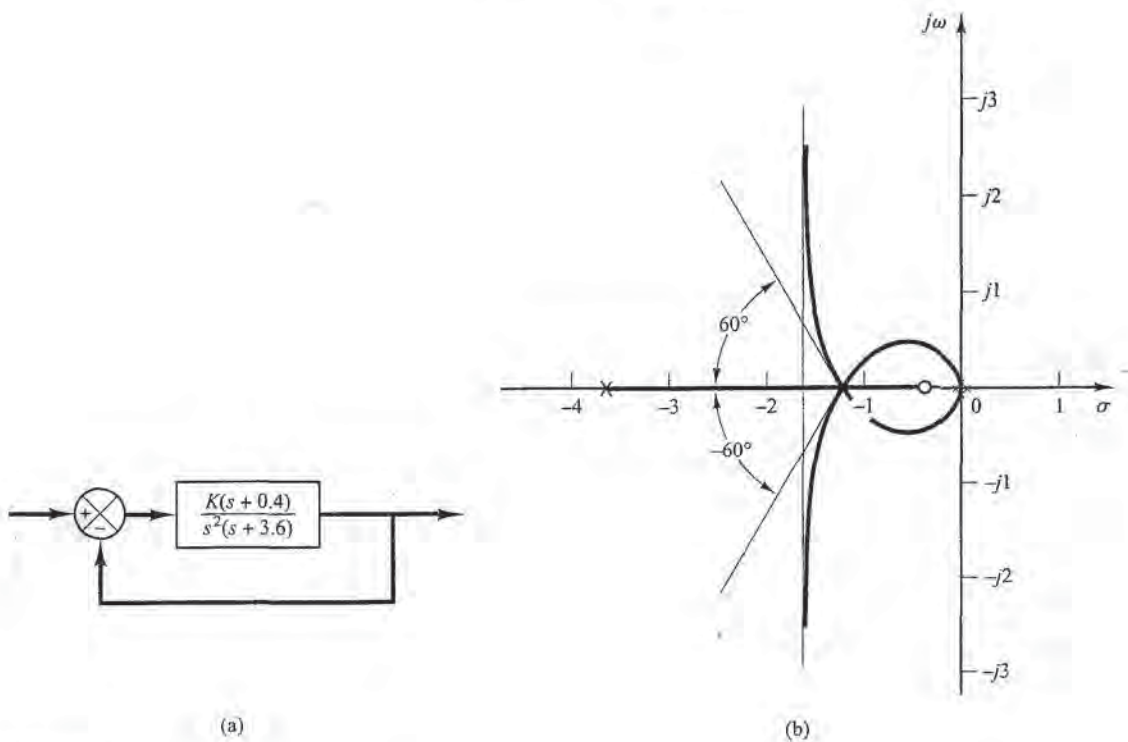


Figure 6-44
(a) Control system; (b) root-locus plot.

The intersection of the asymptotes and the real axis is obtained from

$$s = -\frac{0 + 0 + 3.6 - 0.4}{3 - 1} = -1.6$$

Next we shall find the breakaway points. Since the characteristic equation is

$$s^3 + 3.6s^2 + Ks + 0.4K = 0$$

we have

$$K = -\frac{s^3 + 3.6s^2}{s + 0.4}$$

The breakaway and break-in points are found from

$$\frac{dK}{ds} = -\frac{(3s^2 + 7.2s)(s + 0.4) - (s^3 + 3.6s^2)}{(s + 0.4)^2} = 0$$

from which we get

$$s^3 + 2.4s^2 + 1.44s = 0$$

or

$$s(s + 1.2)^2 = 0$$

Thus, the breakaway or break-in points are at $s = 0$ and $s = -1.2$. Note that $s = -1.2$ is a double root. When a double root occurs in $dK/ds = 0$ at point $s = -1.2$, $d^2K/(ds^2) = 0$ at this point. The value of gain K at point $s = -1.2$ is

$$K = -\frac{s^3 + 3.6s^2}{s + 0.4} \bigg|_{s=-1.2} = 4.32$$

This means that with $K = 4.32$ the characteristic equation has a triple root at point $s = -1.2$. This can be easily verified as follows:

$$s^3 + 3.6s^2 + 4.32s + 1.728 = (s + 1.2)^3 = 0$$

Hence, three root-locus branches meet at point $s = -1.2$. The angles of departures at point $s = -1.2$ of the root locus branches that approach the asymptotes are $\pm 180^\circ/3$, that is, 60° and -60° . (See Problem A-6-7.)

Finally, we shall examine if root-locus branches cross the imaginary axis. By substituting $s = j\omega$ into the characteristic equation, we have

$$(j\omega)^3 + 3.6(j\omega)^2 + K(j\omega) + 0.4K = 0$$

or

$$(0.4K - 3.6\omega^2) + j\omega(K - \omega^2) = 0$$

This equation can be satisfied only if $\omega = 0$, $K = 0$. At point $\omega = 0$, the root locus is tangent to the $j\omega$ axis because of the presence of a double pole at the origin. There are no points that root-locus branches cross the imaginary axis.

A sketch of the root loci for this system is shown in Figure 6-44(b).

- A-6-7.** Referring to Problem A-6-6, obtain the equations for the root-locus branches for the system shown in Figure 6-44(a). Show that the root-locus branches cross the real axis at the breakaway point at angles $\pm 60^\circ$.

Solution. The equations for the root-locus branches can be obtained from the angle condition

$$\angle \frac{K(s + 0.4)}{s^2(s + 3.6)} = \pm 180^\circ(2k + 1)$$

which can be rewritten as

$$\angle s + 0.4 - 2\angle s - \angle s + 3.6 = \pm 180^\circ(2k + 1)$$

By substituting $s = \sigma + j\omega$, we obtain

$$\angle \sigma + j\omega + 0.4 - 2\angle \sigma + j\omega - \angle \sigma + j\omega + 3.6 = \pm 180^\circ(2k + 1)$$

or

$$\tan^{-1}\left(\frac{\omega}{\sigma + 0.4}\right) - 2\tan^{-1}\left(\frac{\omega}{\sigma}\right) - \tan^{-1}\left(\frac{\omega}{\sigma + 3.6}\right) = \pm 180^\circ(2k + 1)$$

By rearranging, we have

$$\tan^{-1}\left(\frac{\omega}{\sigma + 0.4}\right) - \tan^{-1}\left(\frac{\omega}{\sigma}\right) = \tan^{-1}\left(\frac{\omega}{\sigma}\right) + \tan^{-1}\left(\frac{\omega}{\sigma + 3.6}\right) \pm 180^\circ(2k + 1)$$

Taking tangents of both sides of this last equation, and noting that

$$\tan\left[\tan^{-1}\left(\frac{\omega}{\sigma + 3.6}\right) \pm 180^\circ(2k + 1)\right] = \frac{\omega}{\sigma + 3.6}$$

we obtain

$$\frac{\frac{\omega}{\sigma + 0.4} - \frac{\omega}{\sigma}}{1 + \frac{\omega}{\sigma + 0.4} \frac{\omega}{\sigma}} = \frac{\frac{\omega}{\sigma} + \frac{\omega}{\sigma + 3.6}}{1 - \frac{\omega}{\sigma} \frac{\omega}{\sigma + 3.6}}$$

which can be simplified to

$$\frac{\omega\sigma - \omega(\sigma + 0.4)}{(\sigma + 0.4)\sigma + \omega^2} = \frac{\omega(\sigma + 3.6) + \omega\sigma}{\sigma(\sigma + 3.6) - \omega^2}$$

or

$$\omega(\sigma^3 + 2.4\sigma^2 + 1.44\sigma + 1.6\omega^2 + \sigma\omega^2) = 0$$

which can be further simplified to

$$\omega[\sigma(\sigma + 1.2)^2 + (\sigma + 1.6)\omega^2] = 0$$

For $\sigma \neq -1.6$, we may write this last equation as

$$\omega\left[\omega - (\sigma + 1.2)\sqrt{\frac{-\sigma}{\sigma + 1.6}}\right]\left[\omega + (\sigma + 1.2)\sqrt{\frac{-\sigma}{\sigma + 1.6}}\right] = 0$$

which gives the equations for the root-locus as follows:

$$\omega = 0$$

$$\omega = (\sigma + 1.2)\sqrt{\frac{-\sigma}{\sigma + 1.6}}$$

$$\omega = -(\sigma + 1.2)\sqrt{\frac{-\sigma}{\sigma + 1.6}}$$

The equation $\omega = 0$ represents the real axis. The root locus for $0 \leq K \leq \infty$ is between points $s = -0.4$ and $s = -3.6$. (The real axis other than this line segment and the origin $s = 0$ corresponds to the root locus for $-\infty \leq K < 0$.)

The equations

$$\omega = \pm(\sigma + 1.2)\sqrt{\frac{-\sigma}{\sigma + 1.6}} \quad (6-21)$$

represent the complex branches for $0 \leq K \leq \infty$. These two branches lie between $\sigma = -1.6$ and $\sigma = 0$. [See Figure 6-44(b).] The slopes of the complex root-locus branches at the breakaway point ($\sigma = -1.2$) can be found by evaluating $d\omega/d\sigma$ of Equation (6-21) at point $\sigma = -1.2$.

$$\left. \frac{d\omega}{d\sigma} \right|_{\sigma=-1.2} = \pm \sqrt{\frac{-\sigma}{\sigma + 1.6}} \bigg|_{\sigma=-1.2} = \pm \sqrt{\frac{1.2}{0.4}} = \pm \sqrt{3}$$

Since $\tan^{-1} \sqrt{3} = 60^\circ$, the root-locus branches intersect the real axis with angles $\pm 60^\circ$.

A-6-8. Consider the system shown in Figure 6-45(a), which has an unstable feedforward transfer function. Sketch the root-locus plot and locate the closed-loop poles. Show that, although the closed-loop poles lie on the negative real axis and the system is not oscillatory, the unit-step response curve will exhibit overshoot.

Solution. The root-locus plot for this system is shown in Figure 6-45(b). The closed-loop poles are located at $s = -2$ and $s = -5$.

The closed-loop transfer function becomes

$$\frac{C(s)}{R(s)} = \frac{10(s+1)}{s^2 + 7s + 10}$$

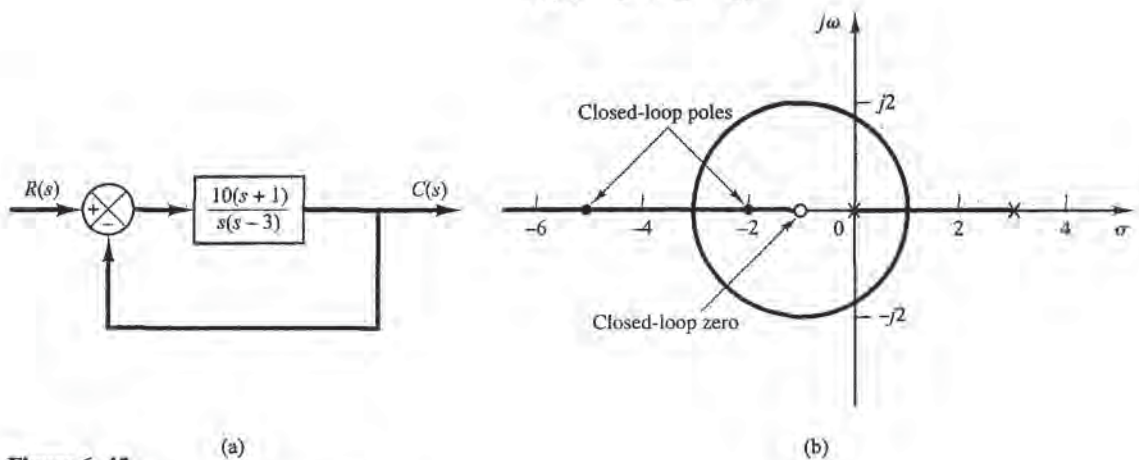
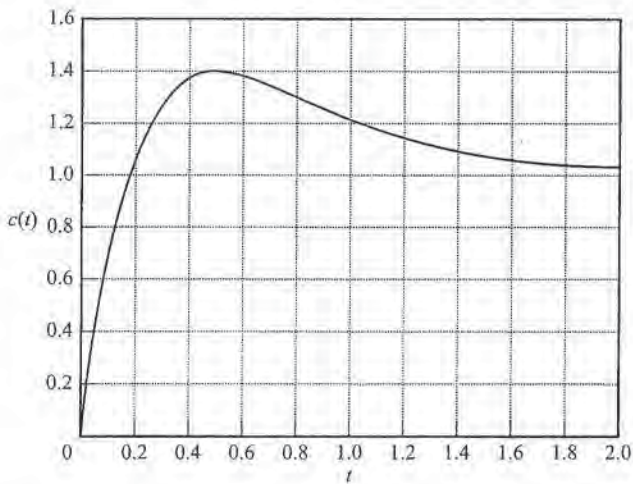


Figure 6-45
(a) Control system; (b) root-locus plot.

Figure 6-46
Unit-step response curve for the system shown in Figure 6-45(a).



The unit-step response of this system is

$$C(s) = \frac{10(s + 1)}{s(s + 2)(s + 5)}$$

The inverse Laplace transform of $C(s)$ gives

$$c(t) = 1 + 1.666e^{-2t} - 2.666e^{-5t}, \quad \text{for } t \geq 0$$

The unit-step response curve is shown in Figure 6-46. Although the system is not oscillatory, the unit-step response curve exhibits overshoot. (This is due to the presence of a zero at $s = -1$.)

- A-6-9.** Sketch the root loci of the control system shown in Figure 6-47(a). Determine the range of gain K for stability.

Solution. Open-loop poles are located at $s = 1$, $s = -2 + j\sqrt{3}$, and $s = -2 - j\sqrt{3}$. A root locus exists on the real axis between points $s = 1$ and $s = -\infty$. The asymptotes of the root-locus branches are found as follows:

$$\text{Angles of asymptotes} = \frac{\pm 180^\circ(2k + 1)}{3} = 60^\circ, -60^\circ, 180^\circ$$

The intersection of the asymptotes and the real axis is obtained as

$$s = -\frac{-1 + 2 + 2}{3} = -1$$

The breakaway and break-in points can be located from $dK/ds = 0$. Since

$$K = -(s - 1)(s^2 + 4s + 7) = -(s^3 + 3s^2 + 3s - 7)$$

we have

$$\frac{dK}{ds} = -(3s^2 + 6s + 3) = 0$$

which yields

$$(s + 1)^2 = 0$$

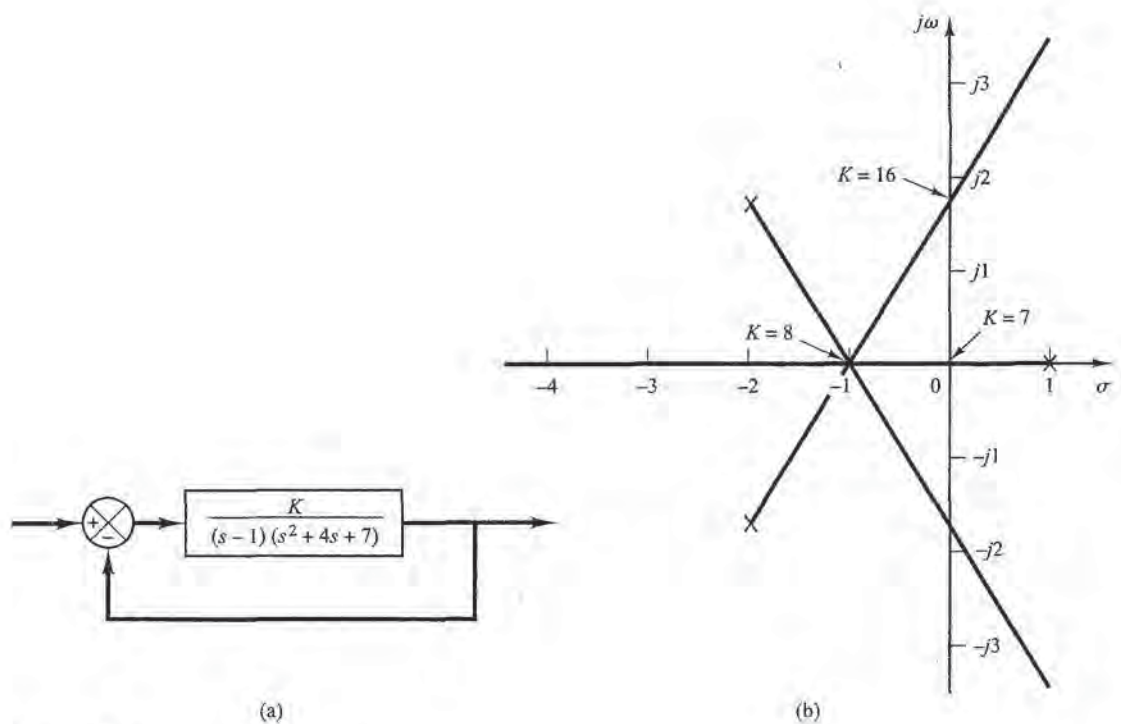


Figure 6-47
(a) Control system; (b) root-locus plot.

Thus the equation $dK/ds = 0$ has a double root at $s = -1$. (This means that the characteristic equation has a triple root at $s = -1$.) The breakaway point is located at $s = -1$. Three root-locus branches meet at this breakaway point. The angles of departure of the branches at the breakaway point are $\pm 180^\circ/3$, that is, 60° and -60° .

We shall next determine the points where root-locus branches may cross the imaginary axis. Noting that the characteristic equation is

$$(s - 1)(s^2 + 4s + 7) + K = 0$$

or

$$s^3 + 3s^2 + 3s - 7 + K = 0$$

we substitute $s = j\omega$ into it and obtain

$$(j\omega)^3 + 3(j\omega)^2 + 3(j\omega) - 7 + K = 0$$

By rewriting this last equation, we have

$$(K - 7 - 3\omega^2) + j\omega(3 - \omega^2) = 0$$

This equation is satisfied when

$$\omega = \pm\sqrt{3}, \quad K = 7 + 3\omega^2 = 16 \quad \text{or} \quad \omega = 0, \quad K = 7$$

Example Problems and Solutions

397

The root-locus branches cross the imaginary axis at $\omega = \pm\sqrt{3}$ (where $K = 16$) and $\omega = 0$ (where $K = 7$). Since the value of gain K at the origin is 7, the range of gain value K for stability is

$$7 < K < 16$$

Figure 6-47(b) shows a sketch of the root loci for the system. Notice that all branches consist of parts of straight lines.

The fact that the root-locus branches consist of straight lines can be verified as follows: Since the angle condition is

$$\angle \frac{K}{(s-1)(s+2+j\sqrt{3})(s+2-j\sqrt{3})} = \pm 180^\circ(2k+1)$$

we have

$$-\angle s-1 - \angle s+2+j\sqrt{3} - \angle s+2-j\sqrt{3} = \pm 180^\circ(2k+1)$$

By substituting $s = \sigma + j\omega$ into this last equation,

$$\angle \sigma-1+j\omega + \angle \sigma+2+j\omega+j\sqrt{3} + \angle \sigma+2+j\omega-j\sqrt{3} = \pm 180^\circ(2k+1)$$

or

$$\angle \sigma+2+j(\omega+\sqrt{3}) + \angle \sigma+2+j(\omega-\sqrt{3}) = -\angle \sigma-1+j\omega \pm 180^\circ(2k+1)$$

which can be rewritten as

$$\tan^{-1}\left(\frac{\omega+\sqrt{3}}{\sigma+2}\right) + \tan^{-1}\left(\frac{\omega-\sqrt{3}}{\sigma+2}\right) = -\tan^{-1}\left(\frac{\omega}{\sigma-1}\right) \pm 180^\circ(2k+1)$$

Taking tangents of both sides of this last equation, we obtain

$$\frac{\frac{\omega+\sqrt{3}}{\sigma+2} + \frac{\omega-\sqrt{3}}{\sigma+2}}{1 - \left(\frac{\omega+\sqrt{3}}{\sigma+2}\right)\left(\frac{\omega-\sqrt{3}}{\sigma+2}\right)} = -\frac{\omega}{\sigma-1}$$

or

$$\frac{2\omega(\sigma+2)}{\sigma^2+4\sigma+4-\omega^2+3} = -\frac{\omega}{\sigma-1}$$

which can be simplified to

$$2\omega(\sigma+2)(\sigma-1) = -\omega(\sigma^2+4\sigma+7-\omega^2)$$

or

$$\omega(3\sigma^2+6\sigma+3-\omega^2) = 0$$

Further simplification of this last equation yields

$$\omega\left(\sigma+1+\frac{1}{\sqrt{3}\omega}\right)\left(\sigma+1-\frac{1}{\sqrt{3}\omega}\right) = 0$$

which defines three lines:

$$\omega = 0, \quad \sigma+1+\frac{1}{\sqrt{3}}\omega = 0, \quad \sigma+1-\frac{1}{\sqrt{3}}\omega = 0$$

Thus the root-locus branches consist of three lines. Note that the root loci for $K > 0$ consist of portions of the straight lines as shown in Figure 6-47(b). (Note that each straight line starts from an open-loop pole and extends to infinity in the direction of 180° , 60° , or -60° measured from the real axis.) The remaining portion of each straight line corresponds to $K < 0$.

A-6-10. Consider the system shown in Figure 6-48(a). Sketch the root loci.

Solution. The open-loop zeros of the system are located at $s = \pm j$. The open-loop poles are located at $s = 0$ and $s = -2$. This system involves two poles and two zeros. Hence, there is a possibility that a circular root-locus branch exists. In fact, such a circular root locus exists in this case, as shown in the following. The angle condition is

$$\angle \frac{K(s+j)(s-j)}{s(s+2)} = \pm 180^\circ(2k+1)$$

or

$$\angle s + j + \angle s - j - \angle s - \angle s + 2 = \pm 180^\circ(2k+1)$$

By substituting $s = \sigma + j\omega$ into this last equation, we obtain

$$\angle \sigma + j\omega + j + \angle \sigma + j\omega - j = \angle \sigma + j\omega + \angle \sigma + 2 + j\omega \pm 180^\circ(2k+1)$$

or

$$\tan^{-1}\left(\frac{\omega+1}{\sigma}\right) + \tan^{-1}\left(\frac{\omega-1}{\sigma}\right) = \tan^{-1}\left(\frac{\omega}{\sigma}\right) + \tan^{-1}\left(\frac{\omega}{\sigma+2}\right) \pm 180^\circ(2k+1)$$

Taking tangents of both sides of this equation and noting that

$$\tan\left[\tan^{-1}\left(\frac{\omega}{\sigma+2}\right) \pm 180^\circ\right] = \frac{\omega}{\sigma+2}$$

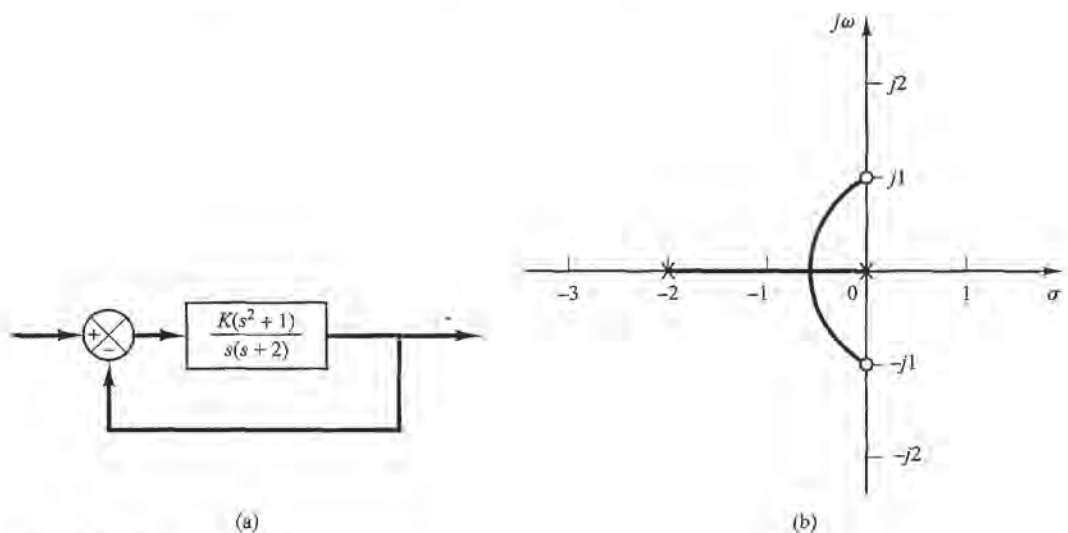


Figure 6-48
(a) Control system; (b) root-locus plot.

we obtain

$$\frac{\frac{\omega + 1}{\sigma} + \frac{\omega - 1}{\sigma}}{1 - \frac{\omega + 1}{\sigma} \frac{\omega - 1}{\sigma}} = \frac{\frac{\omega}{\sigma} + \frac{\omega}{\sigma + 2}}{1 - \frac{\omega}{\sigma} \frac{\omega}{\sigma + 2}}$$

or

$$\omega \left[\left(\sigma - \frac{1}{2} \right)^2 + \omega^2 - \frac{5}{4} \right] = 0$$

which is equivalent to

$$\omega = 0 \quad \text{or} \quad \left(\sigma - \frac{1}{2} \right)^2 + \omega^2 = \frac{5}{4}$$

These two equations are equations for the root loci. The first equation corresponds to the root locus on the real axis. (The segment between $s = 0$ and $s = -2$ corresponds to the root locus for $0 \leq K < \infty$. The remaining parts of the real axis correspond to the root locus for $K < 0$.) The second equation is an equation for a circle. Thus, there exists a circular root locus with center at $\sigma = \frac{1}{2}$, $\omega = 0$ and the radius equal to $\sqrt{5}/2$. The root loci are sketched in Figure 6-48(b). [That part of the circular locus to the left of the imaginary zeros corresponds to $K > 0$. The portion of the circular locus not shown in Figure 6-48(b) corresponds to $K < 0$.]

A-6-11. Consider the control system shown in Figure 6-49. Plot the root loci with MATLAB.

Solution. MATLAB Program 6-11 generates a root-locus plot as shown in Figure 6-50. The root loci must be symmetric about the real axis. However, Figure 6-50 shows otherwise.

MATLAB supplies its own set of gain values that are used to calculate a root-locus plot. It does so by an internal adaptive step-size routine. However, in certain systems, very small changes in the gain cause drastic changes in root locations within a certain range of gains. Thus, MATLAB takes too big a jump in its gain values when calculating the roots, and root locations change by a relatively large amount. When plotting, MATLAB connects these points and causes a strange-looking graph at the location of sensitive gains. Such erroneous root-locus plots typically occur when the loci approach a double pole (or triple or higher pole), since the locus is very sensitive to small gain changes.

MATLAB Program 6-11

```
% ----- Root-locus plot -----
num = [0 0 1 0.4];
den = [1 3.6 0 0];
rlocus(num,den);
v = [-5 1 -3 3]; axis(v)
grid
title('Root-Locus Plot of G(s) = K(s + 0.4)/[s^2(s + 3.6)]')
```

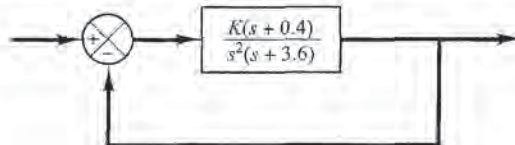


Figure 6-49
Control system.

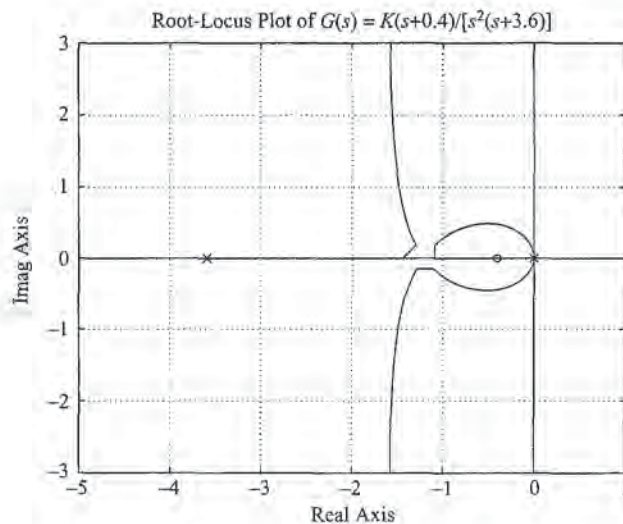


Figure 6-50
Root-locus plot.

In the problem considered here, the critical region of gain K is between 4.2 and 4.4. Thus we need to set the step size small enough in this region. We may divide the region for K as follows:

```
K1 = [0:0.2:4.2];
K2 = [4.2:0.002:4.4];
K3 = [4.4:0.2:10];
K4 = [10:5:200];
K = [K1 K2 K3 K4];
```

Entering MATLAB Program 6-12 into the computer, we obtain the plot as shown in Figure 6-51. If we change the plot command `plot(r,'o')` in MATLAB Program 6-12 to `plot(r,'-')`, we obtain Figure 6-52. Figures 6-51 and 6-52 respectively, show satisfactory root-locus plots.

MATLAB Program 6-12

```
% ----- Root-locus plot -----
num = [0 0 1 0.4];
den = [1 3.6 0 0];
K1 = [0:0.2:4.2];
K2 = [4.2:0.002:4.4];
K3 = [4.4:0.2:10];
K4 = [10:5:200];
K = [K1 K2 K3 K4];
r = rlocus(num,den,K);
plot(r,'o')
v = [-5 1 -5 5]; axis(v)
grid
title('Root-Locus Plot of  $G(s) = K(s + 0.4)/[s^2(s + 3.6)]$ ')
xlabel('Real Axis')
ylabel('Imag Axis')
```

Figure 6-51
Root-locus plot.

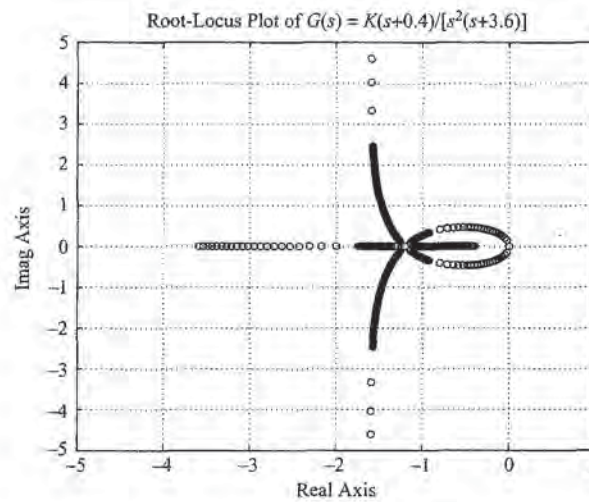
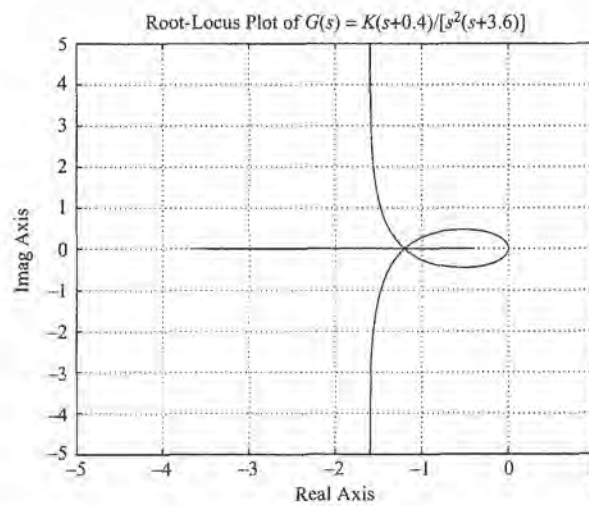


Figure 6-52
Root-locus plot.



A-6-12. Consider the system whose open-loop transfer function $G(s)H(s)$ is given by

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

Using MATLAB, plot root loci and their asymptotes.

Solution. We shall plot the root loci and asymptotes on one diagram. Since the open-loop transfer function is given by

$$\begin{aligned} G(s)H(s) &= \frac{K}{s(s+1)(s+2)} \\ &= \frac{K}{s^3 + 3s^2 + 2s} \end{aligned}$$

the equation for the asymptotes may be obtained as follows: Noting that

$$\lim_{s \rightarrow \infty} \frac{K}{s^3 + 3s^2 + 2s} \neq \lim_{s \rightarrow \infty} \frac{K}{s^3 + 3s^2 + 3s + 1} = \frac{K}{(s+1)^3}$$

the equation for the asymptotes may be given by

$$G_a(s)H_a(s) = \frac{K}{(s + 1)^3}$$

Hence, for the system we have

$$\begin{aligned} \text{num} &= [0 \ 0 \ 0 \ 1] \\ \text{den} &= [1 \ 3 \ 2 \ 0] \end{aligned}$$

and for the asymptotes,

$$\begin{aligned} \text{numa} &= [0 \ 0 \ 0 \ 1] \\ \text{dena} &= [1 \ 3 \ 3 \ 1] \end{aligned}$$

In using the following root-locus and plot commands

$$\begin{aligned} r &= \text{rlocus}(\text{num}, \text{den}) \\ a &= \text{rlocus}(\text{numa}, \text{dena}) \\ \text{plot}([r \ a]) \end{aligned}$$

the number of rows of r and that of a must be the same. To ensure this, we include the gain constant K in the commands. For example,

$$\begin{aligned} K1 &= 0:0.1:0.3; \\ K2 &= 0.3:0.005:0.5; \\ K3 &= 0.5:0.5:10; \\ K4 &= 10:5:100; \\ K &= [K1 \ K2 \ K3 \ K4] \\ r &= \text{rlocus}(\text{num}, \text{den}, K) \\ a &= \text{rlocus}(\text{numa}, \text{dena}, K) \\ y &= [r \ a] \\ \text{plot}(y, '-') \end{aligned}$$

MATLAB Program 6-13

```
% ----- Root-Locus Plots -----
num = [0 0 0 1];
den = [1 3 2 0];
numa = [0 0 0 1];
dena = [1 3 3 1];
K1 = 0:0.1:0.3;
K2 = 0.3:0.005:0.5;
K3 = 0.5:0.5:10;
K4 = 10:5:100;
K = [K1 K2 K3 K4];
r = rlocus(num,den,K);
a = rlocus(numa,dena,K);
y = [r a];
plot(y, '-')
v = [-4 4 -4 4]; axis(v)
grid
title('Root-Locus Plot of G(s) = K/[s(s + 1)(s + 2)] and Asymptotes')
xlabel('Real Axis')
ylabel('Imag Axis')
% ***** Manually draw open-loop poles in the hard copy *****
```

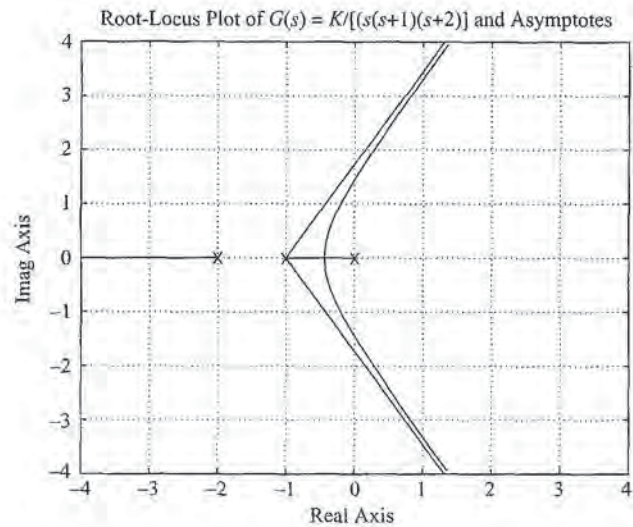


Figure 6-53
Root-locus plot.

Including gain K in rlocus command ensures that the r matrix and a matrix have the same number of rows. MATLAB Program 6-13 will generate a plot of root loci and their asymptotes. See Figure 6-53. Drawing two or more plots in one diagram can also be accomplished by using the hold command. MATLAB Program 6-14 uses the hold command. The resulting root-locus plot is shown in Figure 6-54.

MATLAB Program 6-14
<pre>% ----- Root-Locus Plots ----- num = [0 0 0 1]; den = [1 3 2 0]; numa = [0 0 0 1]; dena = [1 3 3 1]; K1 = 0:0.1:0.3; K2 = 0.3:0.005:0.5; K3 = 0.5:0.5:10; K4 = 10:5:100; K = [K1 K2 K3 K4]; r = rlocus(num,den,K); a = rlocus(numa,dena,K); plot(r,'o') hold Current plot held plot(a,'-') v = [-4 4 -4 4]; axis(v) grid title('Root-Locus Plot of G(s) = K/[s(s+1)(s+2)] and Asymptotes') xlabel('Real Axis') ylabel('Imag Axis')</pre>

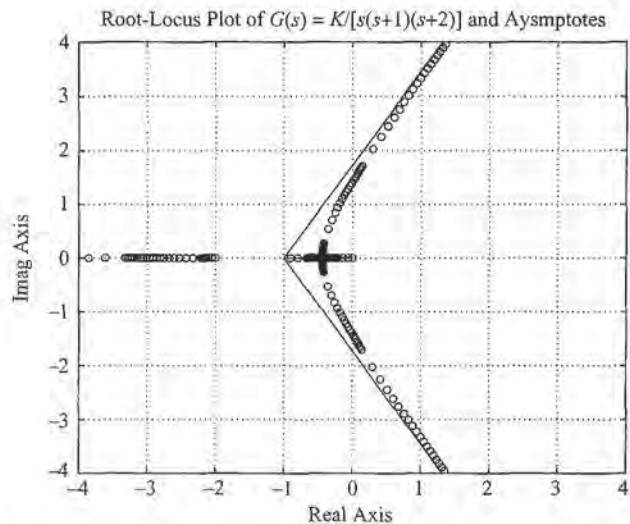


Figure 6-54
Root-locus plot.

A-6-13. Consider a unity-feedback system with the following feedforward transfer function $G(s)$:

$$G(s) = \frac{K(s+2)^2}{(s^2+4)(s+5)^2}$$

Plot root loci for the system with MATLAB.

Solution. A MATLAB program to plot the root loci is given as MATLAB Program 6-15. The resulting root-locus plot is shown in Figure 6-55.

Notice that this is a special case where no root locus exists on the real axis. This means that for any value of $K > 0$ the closed-loop poles of the system are two sets of complex-conjugate poles. (No real closed-loop poles exist.) For example, with $K = 25$, the characteristic equation for the system becomes

$$\begin{aligned} s^4 + 10s^3 + 54s^2 + 140s + 200 \\ &= (s^2 + 4s + 10)(s^2 + 6s + 20) \\ &= (s + 2 + j2.4495)(s + 2 - j2.4495)(s + 3 + j3.3166)(s + 3 - j3.3166) \end{aligned}$$

MATLAB Program 6-15

```
% ----- Root-Locus Plot -----
num = [0 0 1 4 4];
den = [1 10 29 40 100];
r = rlocus(num,den);
plot(r,'o')
hold
current plot held
plot(r,'-')
v = [-8 4 -6 6]; axis(v); axis('square')
grid
title('Root-Locus Plot of G(s) = (s + 2)^2/[(s^2 + 4)(s + 5)^2]')
xlabel('Real Axis')
ylabel('Imag Axis')
```