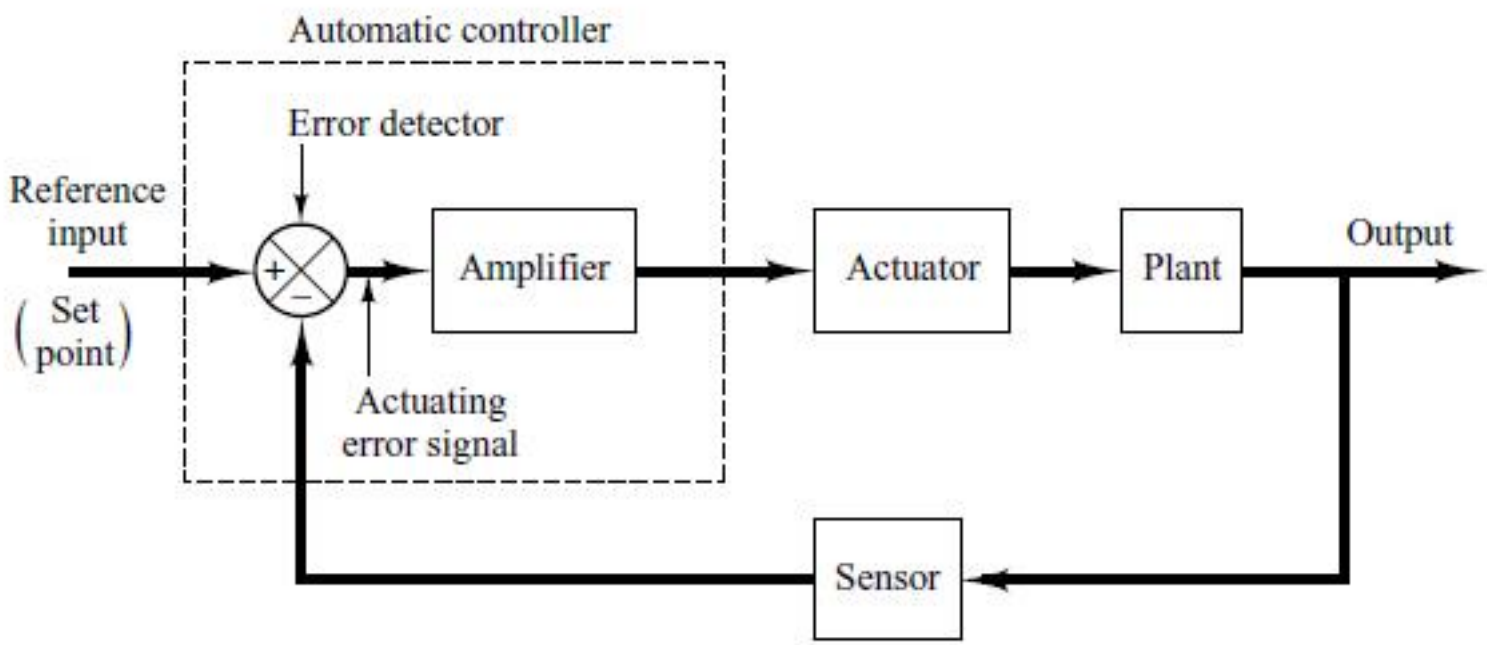
Mathematical Modeling of Control Systems



Mathematical Models of a System

In studying control systems the reader must be able to model dynamic systems in mathematical terms and analyze their dynamic characteristics. A mathematical model of a dynamic system is defined as a set of equations that represents the dynamics of the system accurately, or at least fairly well.

A mathematical model is not unique to a given system. A system may be represented in many different ways and, therefore, may have many mathematical models, depending on one's perspective.

The dynamics of many systems, whether they are mechanical, electrical, thermal, economic, biological, and so on, may be described in terms of differential equations.

Such differential equations may be obtained by using physical laws governing a particular system—for example, Newton's laws for mechanical systems and Kirchhoff's laws for electrical systems. We must always keep in mind that deriving reasonable mathematical models is the most important part of the entire analysis of control systems.

Simplicity Versus Accuracy (1/2)

In obtaining a mathematical model, we must make a compromise between the simplicity of the model and the accuracy of the results of the analysis. In deriving a reasonably simplified mathematical model, we frequently find it necessary to ignore certain inherent physical properties of the system. In particular, if a linear lumped-parameter mathematical model (that is, one employing ordinary differential equations) is desired, it is always necessary to ignore certain nonlinearities and distributed parameters that may be present in the physical system. If the effects that these ignored properties have on the response are small, good agreement will be obtained between the results of the analysis of a mathematical model and the results of the experimental study of the physical system.

Simplicity Versus Accuracy (2/2)

In general, in solving a new problem, it is desirable to build a simplified model so that we can get a general feeling for the solution. A more complete mathematical model may then be built and used for a more accurate analysis.

We must be well aware that a linear lumped-parameter model, which may be valid in low-frequency operations, may not be valid at sufficiently high frequencies, since the neglected property of distributed parameters may become an important factor in the dynamic behavior of the system.

For example, the mass of a spring may be neglected in low frequency operations, but it becomes an important property of the system at high frequencies.

For the case where a mathematical model involves considerable errors, robust control theory may be applied.

Linear Systems

A system is called linear if the principle of superposition applies.

The principle of superposition states that the response produced by the simultaneous application of two different forcing functions is the sum of the two individual responses.

Hence, for the linear system, the response to several inputs can be calculated by treating one input at a time and adding the results. It is this principle that allows one to build up complicated solutions to the linear differential equation from simple solutions.

In an experimental investigation of a dynamic system, if cause and effect are proportional, thus implying that the principle of superposition holds, then the system can be considered linear.

Linear Time-Invariant Systems and Linear Time-Varying Systems

A differential equation is linear if the coefficients are constants or functions only of the independent variable.

Dynamic systems that are composed of linear time-invariant lumped-parameter components may be described by linear time-invariant differential equations—that is, constant-coefficient differential equations.

Such systems are called *linear time-invariant* (or *linear constant-coefficient*) systems. Systems that are represented by differential equations whose coefficients are functions of time are called *linear time-varying* systems.

An example of a time-varying control system is a spacecraft control system. (The mass of a spacecraft changes due to fuel consumption.)

TRANSFER FUNCTION AND IMPULSE RESPONSE FUNCTION

In control theory, functions called transfer functions are commonly used to characterize the input-output relationships of components or systems that can be described by linear, time-invariant, differential equations.

We begin by defining the transfer function and follow with a derivation of the transfer function of a differential equation system.

Then we discuss the impulse-response function.

Transfer Function

The transfer function of a linear, time-invariant, differential equation system is defined as the ratio of the Laplace transform of the output (response function) to that of the input (driving function) assumed that all initial conditions are zero.

Consider the linear time-invariant system defined by the following differential equation:

$$a_0 \overset{(n)}{y} + a_1 \overset{(n-1)}{y} + \cdots + a_{n-1}\dot{y} + a_n y$$

$$= b_0 \overset{(m)}{x} + b_1 \overset{(m-1)}{x} + \cdots + b_{m-1}\dot{x} + b_m x \qquad (n \geq m)$$

y the output of the system and x the input. The transfer function:

$$\text{Transfer function} = G(s) = \left.\frac{\mathscr{L}[\text{output}]}{\mathscr{L}[\text{input}]}\right|_{\text{zero initial conditions}}$$

$$= \frac{Y(s)}{X(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots + b_{m-1}s + b_m}{a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}$$

Comments on Transfer Function

By using the concept of transfer function, it is possible to represent system dynamics by algebraic equations in s.

If the highest power of s in the denominator of the transfer function is equal to n, the system is called an nth-order system.

applicability of the concept of the transfer function is limited to linear, time-invariant, differential equation systems. The transfer function approach, however, is extensively used in the analysis and design of such systems. In what follows, we shall list important comments concerning the transfer function.

(Note that a system referred to in the list is one described by a linear, time-invariant, differential equation.)

Comments on Transfer Function

1. The transfer function of a system is a mathematical model in that it is an operational method of expressing the differential equation that relates the output variable to the input variable.
2. The transfer function is a property of a system itself, independent of the magnitude and nature of the input or driving function.
3. The transfer function includes the units necessary to relate the input to the output; however, it does not provide any information concerning the physical structure of the system. (The transfer functions of many physically different systems can be identical.)
4. If the transfer function of a system is known, the output or response can be studied for various forms of inputs with a view toward understanding the nature of the system.
5. If the transfer function of a system is unknown, it may be established experimentally by introducing known inputs and studying the output of the system. Once established, a transfer function gives a full description of the dynamic characteristics of the system, as distinct from its physical description.

Convolution Integral

For a linear, time-invariant system the transfer function is $G(s) = \frac{Y(s)}{X(s)}$

where X(s) is the Laplace transform of the input to the system and Y(s) is the Laplace transform of the output of the system, where we assume that all initial conditions involved are zero. It follows that the output Y(s) can be written as the product of G(s) and X(s), or (2-1): $Y(s) = G(s)X(s)$

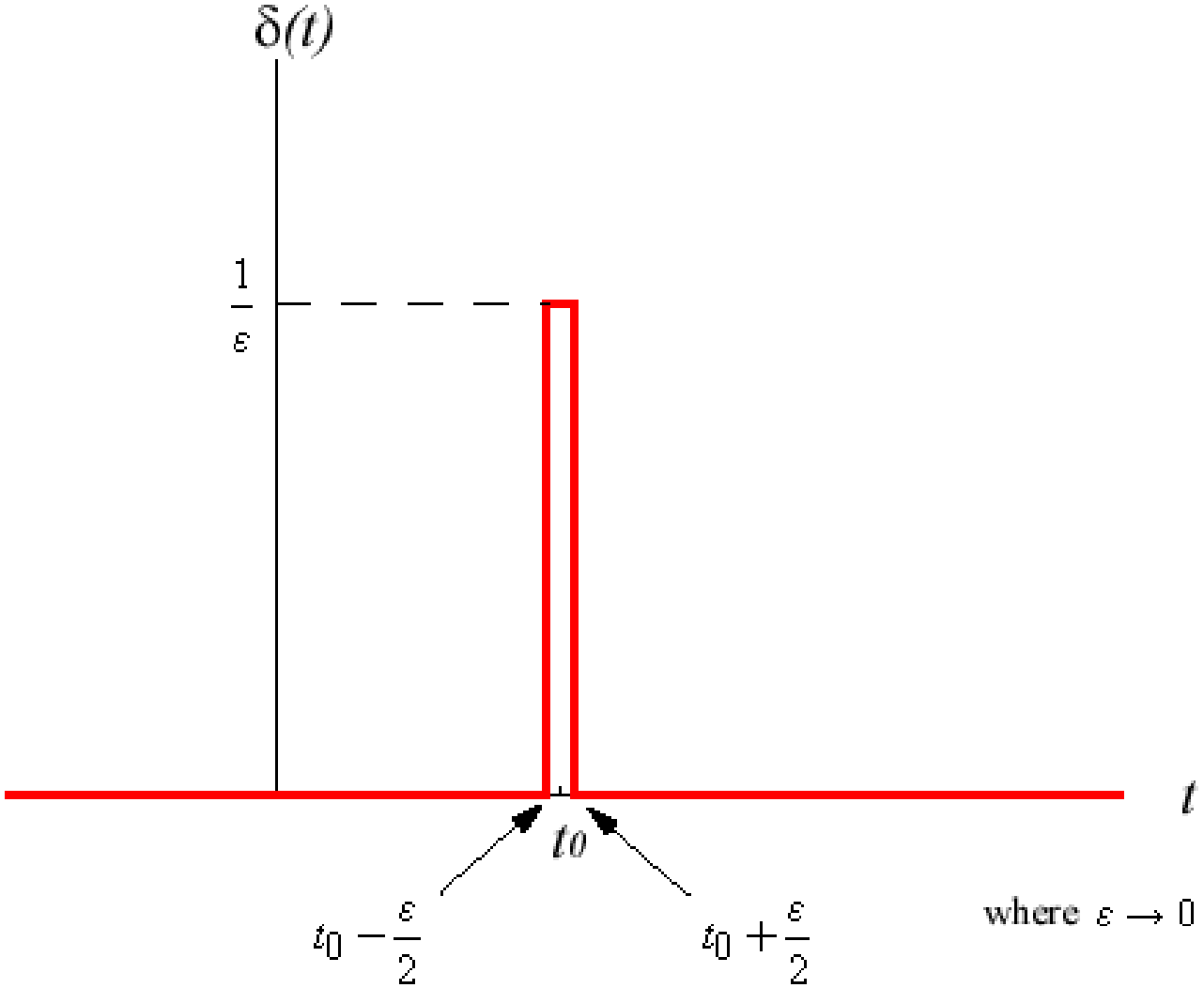
Note that multiplication in the complex domain is equivalent to convolution in the time domain (see Appendix A), so the inverse Laplace transform of Equation (2–1) is given by the following convolution integral:

$$y(t) = \int_0^t x(\tau)g(t-\tau)\,d\tau$$

$$= \int_0^t g(\tau)x(t-\tau)\,d\tau$$

where both g(t) and x(t) are 0 for t<0.

Dirac δ – Function (unit impulse)



Impulse-Response Function

Consider the output (response) of a linear time invariant system to a unit-impulse input when the initial conditions are zero. Since the Laplace transform of the unit-impulse function is unity, the Laplace transform of the output of the system is

$$Y(s) = G(s)$$

The inverse Laplace transform of the output given by this Eqn gives the impulse response of the system. The inverse Laplace transform of G(s), or

$$\mathscr{L}^{-1}[G(s)] = g(t)$$

is called the impulse-response function. This function g(t) is also called the weighting function of the system. The impulse-response function g(t) is thus the response of a linear time-invariant system to a unit-impulse input when the initial conditions are zero. The Laplace transform of this function gives the transfer function.

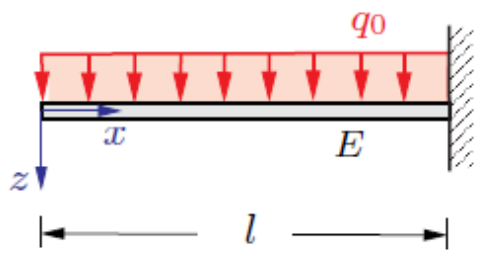
Impulse-Response Function

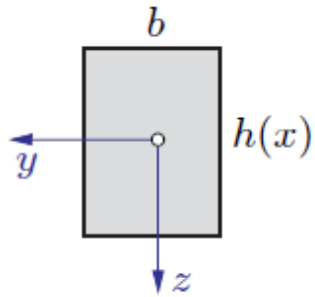
Therefore, the transfer function and impulse-response function of a linear, time-invariant system contain the same information about the system dynamics.

It is hence possible to obtain complete information about the dynamic characteristics of the system by exciting it with an impulse input and measuring the response.

(In practice, a pulse input with a very short duration compared with the significant time constants of the system can be considered an impulse.)

π.χ. Διέγερση δοκού σε ταλάντωση





$$E \cdot I \cdot w_{max} = \frac{q_0 \cdot l^4}{8}$$

Η εξίσωση της ελαστικής γραμμής:
Η διαφορική εξίσωση που περιγράφει την ταλάντωση της δοκού (χωρίς απόσβεση):

$$EI\frac{\partial^4 W}{\partial x^4}(x,t) + \rho A\frac{\partial^2 w}{\partial t^2}(x,t) = f(x,t)$$

Για ελεύθερη ταλάντωση:

$$c^2\frac{\partial^4 w}{\partial x^4}(x,t) + \frac{\partial^2 w}{\partial t^2}(x,t) = 0 \qquad c = \sqrt{\frac{EI}{\rho A}}$$

Η λύση βρίσκεται με τη μέθοδο χωρισμού μεταβλητών: $w(x,t) = W(x)T(t)$

Οπότε καταλήγουμε στη σχέση:

$$\frac{c^2}{W(x)}\frac{d^4W(x)}{dx^4} = -\frac{1}{T(t)}\frac{dy}{dx} = \alpha = \omega^2$$

Οι ιδιοσυχνότητες της δοκού:

$$\omega = \beta^2\sqrt{\frac{EI}{\rho A}} = (\beta l)^2\sqrt{\frac{EI}{\rho A l^4}}$$

$\beta_1 l = 3.926602$
$\beta_2 l = 7.068583$
$\beta_3 l = 10.210176$
$\beta_4 l = 13.351768$

First order system response

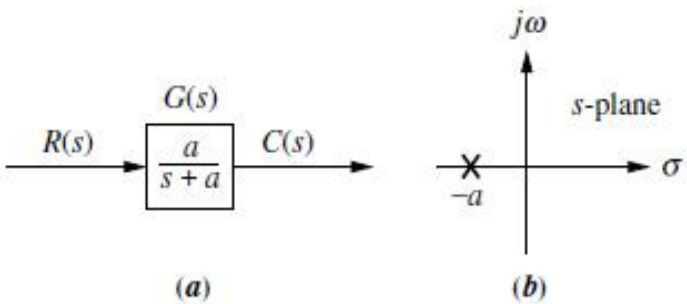


FIGURE 4.4 **a.** First-order system; **b.** pole plot

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

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

Taking the inverse transform, the step response is given by

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

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

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

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

or

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

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

Time constant

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

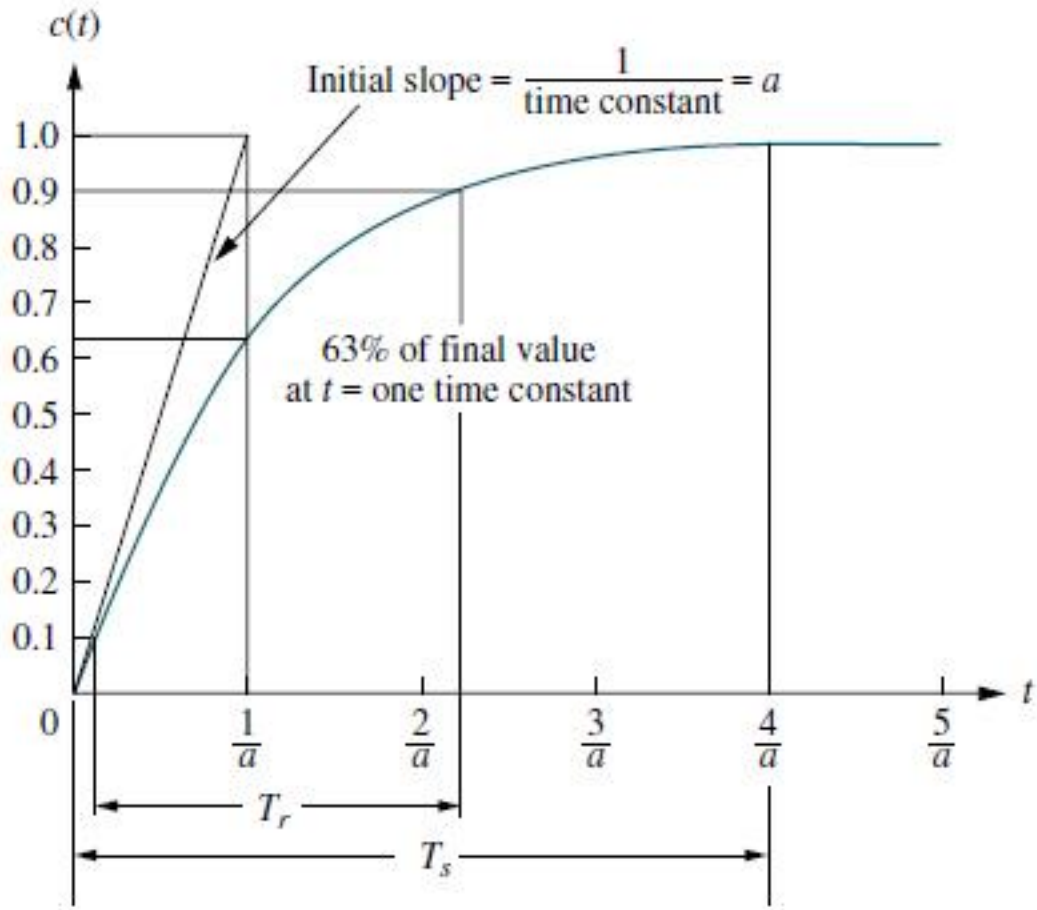


FIGURE 4.5 First-order system response to a unit step

Rise Time, Settling Time

Rise Time, T_r

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

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

Settling Time, T_s

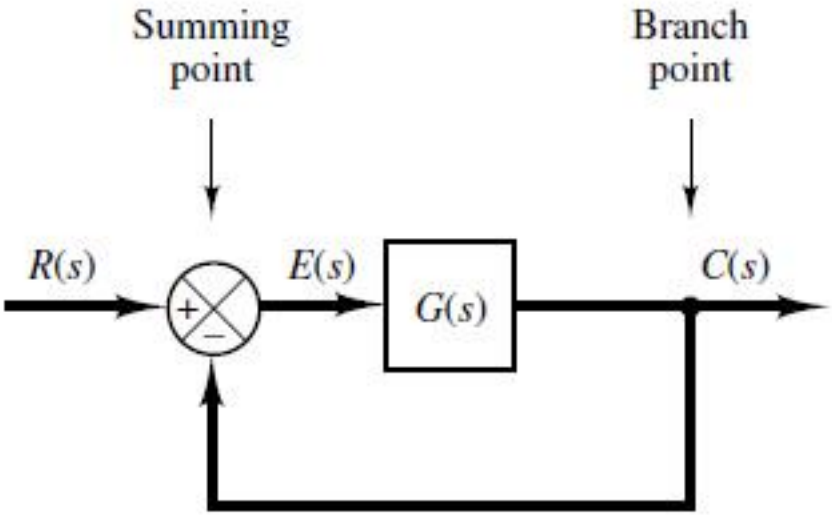
Settling time is defined as the time for the response to reach, and stay within, 2% of its final value.[2] Letting $c(t) = 0.98$ in Eq. (4.6) and solving for time, t, we find the settling time to be

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

Block Diagram of a Closed-Loop System

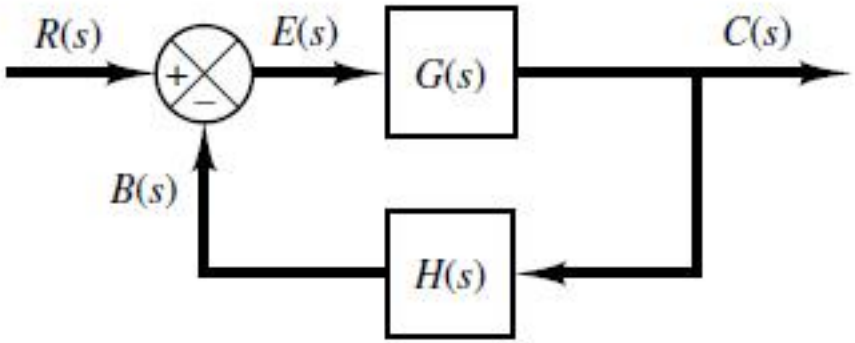
This Figure shows an example of a block diagram of a closed-loop system. The output C(s) is fed back to the summing point, where it is compared with the reference input R(s). The output of the block, C(s) in this case, is obtained by multiplying the transfer function G(s) by the input to the block, E(s). Any linear control system may be represented by a block diagram consisting of blocks, summing points, and branch points.

When the output is fed back to the summing point for comparison with the input, it is necessary to convert the form of the output signal to that of the input signal.

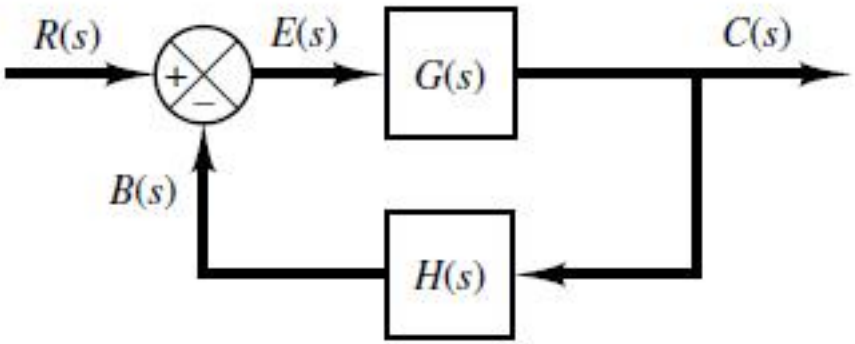


Block Diagram of a Closed-Loop temperature control System

For example, in a temperature control system, the output signal is usually the controlled temperature. The output signal, which has the dimension of temperature, must be converted to a force or position or voltage before it can be compared with the input signal. This conversion is accomplished by the feedback element whose transfer function is H(s), as shown in the Figure. The role of the feedback element is to modify the output before it is compared with the input. (The feedback element is a sensor that measures the output of the plant. The output of the sensor is compared with the system input and the actuating error signal is generated). In this example, the Feedback signal that is fed back to the summing point for comparison with the input is B(s) = H(s)C(s).



Open Loop TF
Feedforward TF



Open-Loop Transfer Function and Feedforward Transfer Function. Referring to Figure 2–4, the ratio of the feedback signal $B(s)$ to the actuating error signal $E(s)$ is called the *open-loop transfer function*. That is,

$$\text{Open-loop transfer function} = \frac{B(s)}{E(s)} = G(s)H(s)$$

The ratio of the output $C(s)$ to the actuating error signal $E(s)$ is called the *feedforward transfer function*, so that

$$\text{Feedforward transfer function} = \frac{C(s)}{E(s)} = G(s)$$

If the feedback transfer function $H(s)$ is unity, then the open-loop transfer function and the feedforward transfer function are the same.

Closed-Loop Transfer Function. For the system shown in Figure 2–4, the output $C(s)$ and input $R(s)$ are related as follows: since

$$C(s) = G(s)E(s)$$

$$E(s) = R(s) - B(s)$$

$$= R(s) - H(s)C(s)$$

eliminating $E(s)$ from these equations gives

$$C(s) = G(s)[R(s) - H(s)C(s)]$$

or

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \tag{2–3}$$

The transfer function relating $C(s)$ to $R(s)$ is called the *closed-loop transfer function*. It relates the closed-loop system dynamics to the dynamics of the feedforward elements and feedback elements.

From Equation (2–3), $C(s)$ is given by

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

Thus the output of the closed-loop system clearly depends on both the closed-loop transfer function and the nature of the input.

Obtaining Cascaded, Parallel, and Feedback (Closed-Loop) Transfer Functions with MATLAB. In control-systems analysis, we frequently need to calculate the cascaded transfer functions, parallel-connected transfer functions, and feedback-connected (closed-loop) transfer functions. MATLAB has convenient commands to obtain the cascaded, parallel, and feedback (closed-loop) transfer functions.

Suppose that there are two components $G_1(s)$ and $G_2(s)$ connected differently as shown in Figure 2–5 (a), (b), and (c), where

$$G_1(s) = \frac{\text{num1}}{\text{den1}}, \qquad G_2(s) = \frac{\text{num2}}{\text{den2}}$$

To obtain the transfer functions of the cascaded system, parallel system, or feedback (closed-loop) system, the following commands may be used:

```
[num, den] = series(num1,den1,num2,den2)
[num, den] = parallel(num1,den1,num2,den2)
[num, den] = feedback(num1,den1,num2,den2)
```

As an example, consider the case where

$$G_1(s) = \frac{10}{s^2 + 2s + 10} = \frac{\text{num1}}{\text{den1}}, \qquad G_2(s) = \frac{5}{s + 5} = \frac{\text{num2}}{\text{den2}}$$

MATLAB Program 2–1 gives $C(s)/R(s)$ = num/den for each arrangement of $G_1(s)$ and $G_2(s)$. Note that the command

```
printsys(num,den)
```

displays the num/den [that is, the transfer function $C(s)/R(s)$] of the system considered.

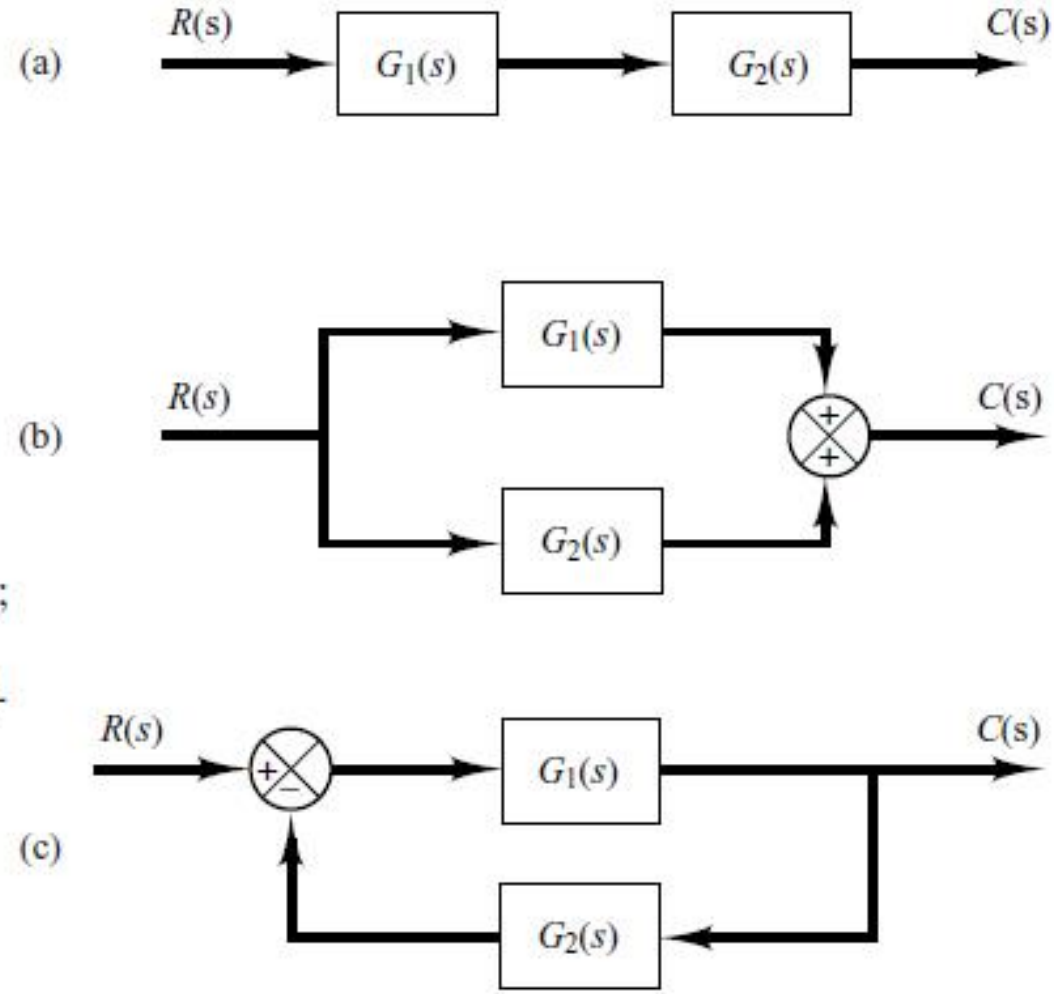
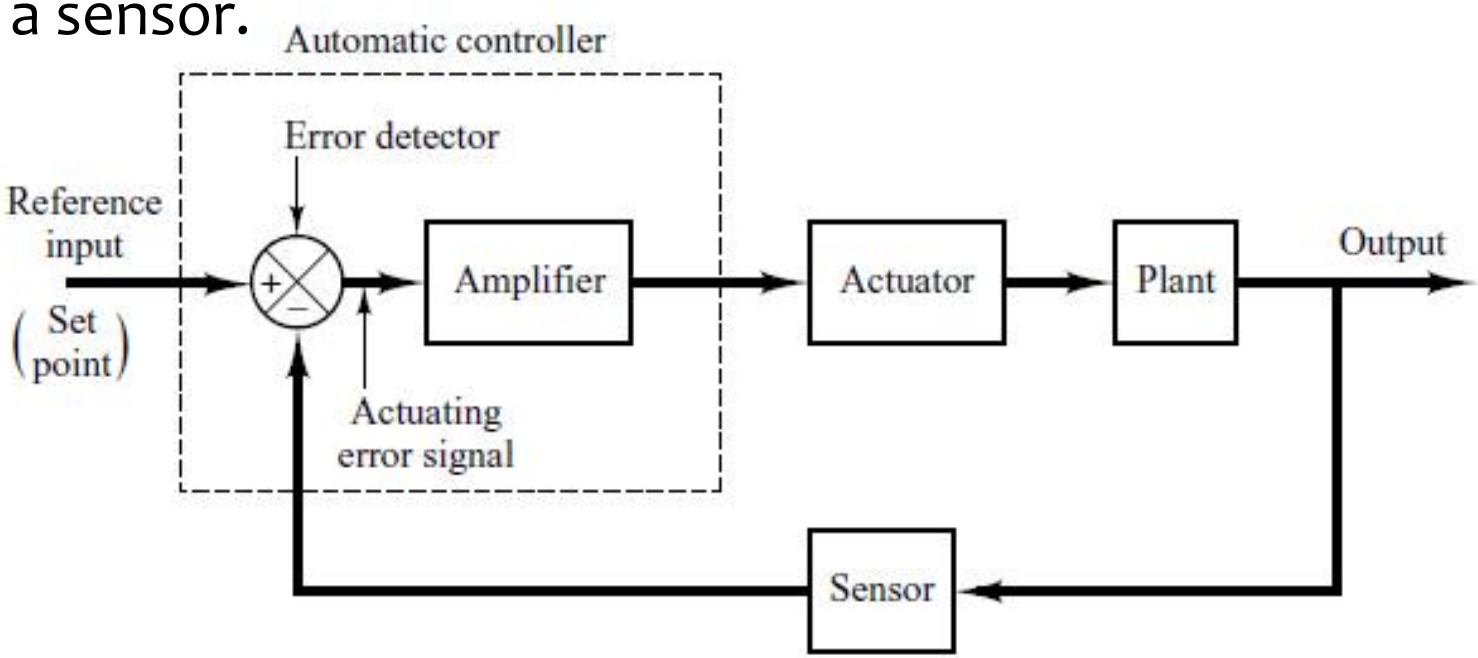


Figure 2–5
(a) Cascaded system;
(b) parallel system;
(c) feedback (closed-loop) system.

Automatic Controllers

An automatic controller compares the actual value of the plant output with the reference input (desired value), determines the deviation, and produces a control signal that will reduce the deviation to zero or to a small value. The manner in which the automatic controller produces the control signal is called the *control action*. Figure is a block diagram of an industrial control system which consists of an automatic controller, an actuator, a plant, and a sensor.



Classifications of Industrial Controllers

Most industrial controllers may be classified according to their control actions as:

1. Two-position or on–off controllers
2. Proportional controllers
3. Integral controllers
4. Proportional-plus-integral (PI) controllers
5. Proportional-plus-derivative (PD) controllers
6. Proportional - integral - derivative (PID) controllers

Most industrial controllers use electricity or pressurized fluid (oil or air as power sources. Consequently, controllers may be classified according to the kind of power employed in the operation

pneumatic controllers, hydraulic controllers, electronic controllers

Two-Position or On–Off Control Action

In a 2-position control system, the actuating element has only two fixed positions, say, on and off. Two-position or on–off control is simple and inexpensive, thus very widely used in industrial and domestic control systems.

Let the output signal from the controller be u(t) and the actuating error signal be e(t). In 2-position control, u(t) remains at either a max or min value, depending on whether the actuating error signal is positive or negative:

$$\begin{aligned} u(t) &= U_1, \qquad \text{for } e(t) > 0 \\ &= U_2, \qquad \text{for } e(t) < 0 \end{aligned}$$

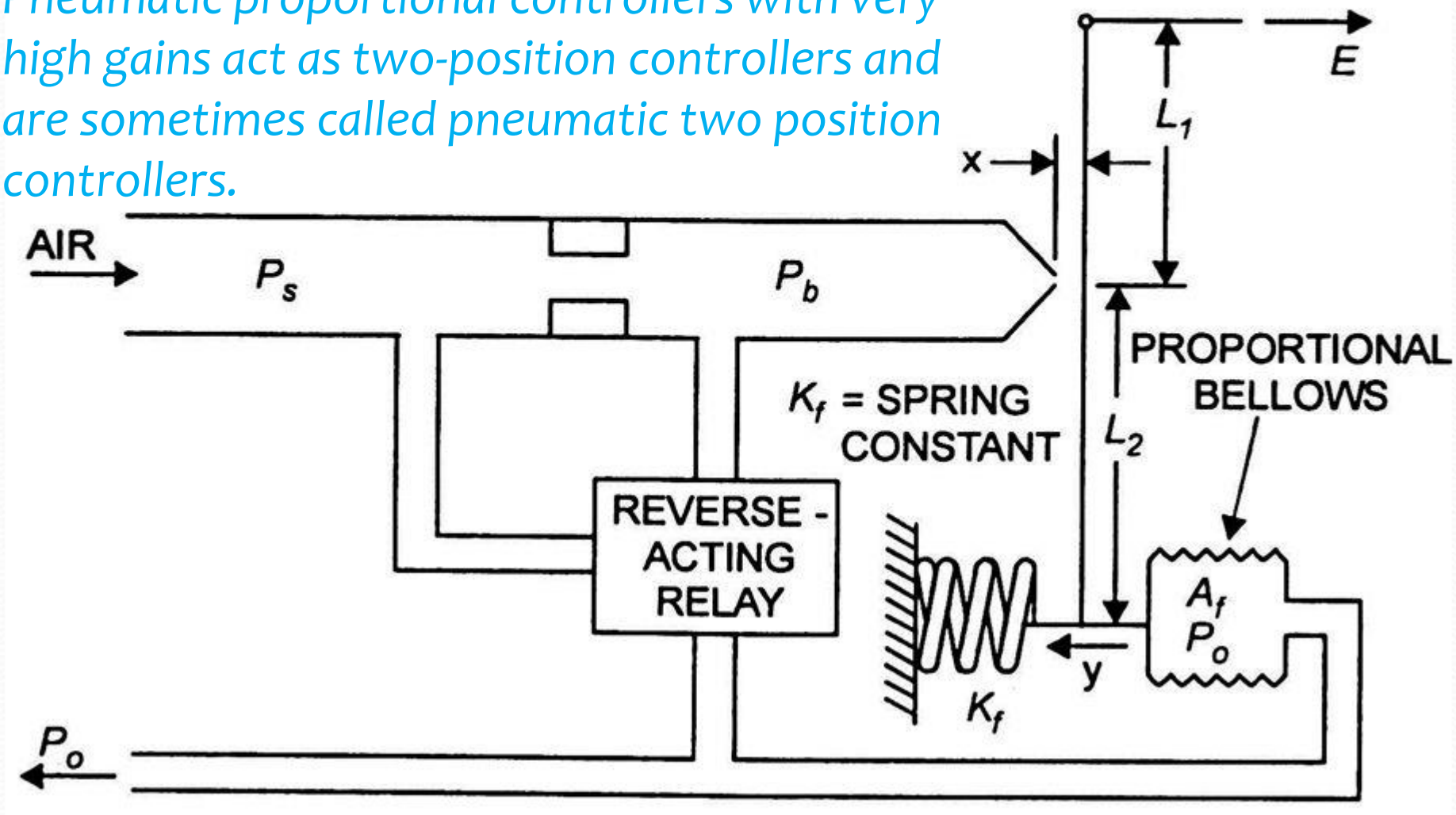
where U_1 and U_2 are constants.

The minimum value U_2 is usually either zero or $-U_1$.

PNEUMATIC CONTROLLERS

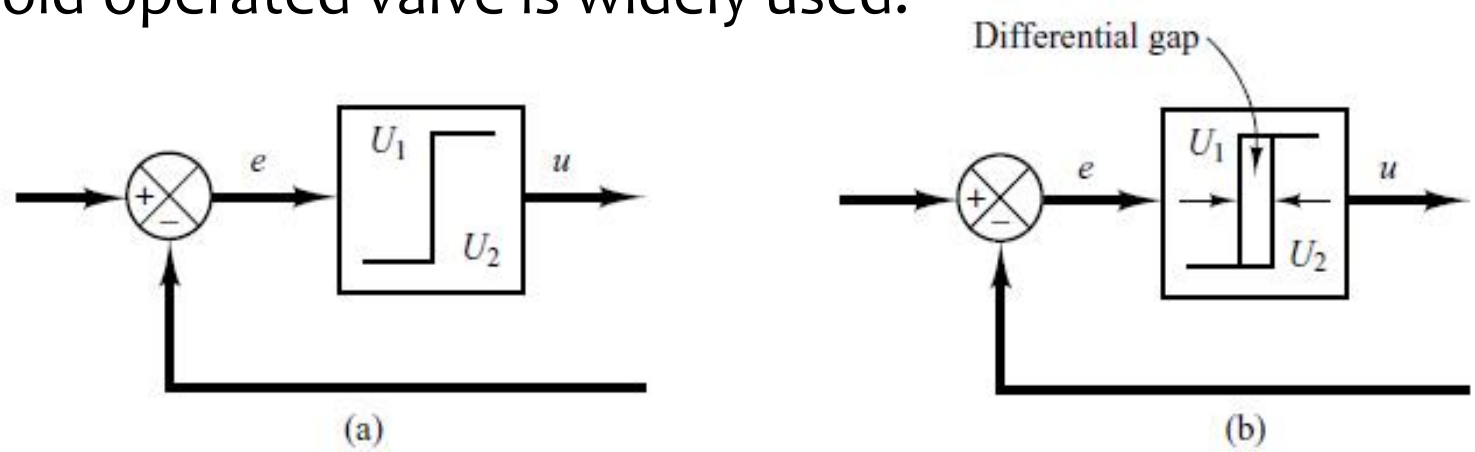
PROPORTIONAL MODE CONTROLLER

Pneumatic proportional controllers with very high gains act as two-position controllers and are sometimes called pneumatic two position controllers.

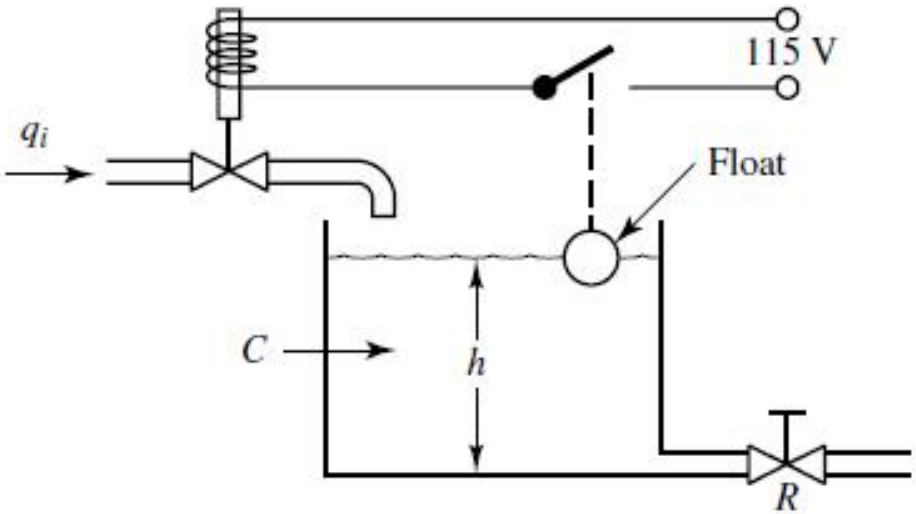


Two-Position or On–Off Controller

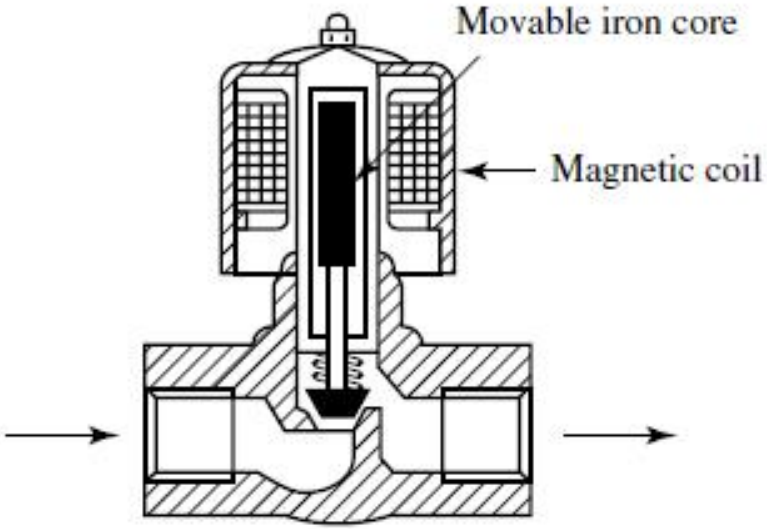
2-position controllers are generally electrical devices, and an electric solenoid-operated valve is widely used.



Figures show the block diagrams for on–off controllers. The range through which the actuating error signal must move before the switching occurs is called the *differential gap*. This gap causes the controller output u(t) to maintain its present value until the actuating error signal has moved slightly beyond the zero value. Sometimes, the differential gap is a result of unintentional friction and lost motion; quite often it is intentionally provided in order to prevent too-frequent on-off operation.



(a)

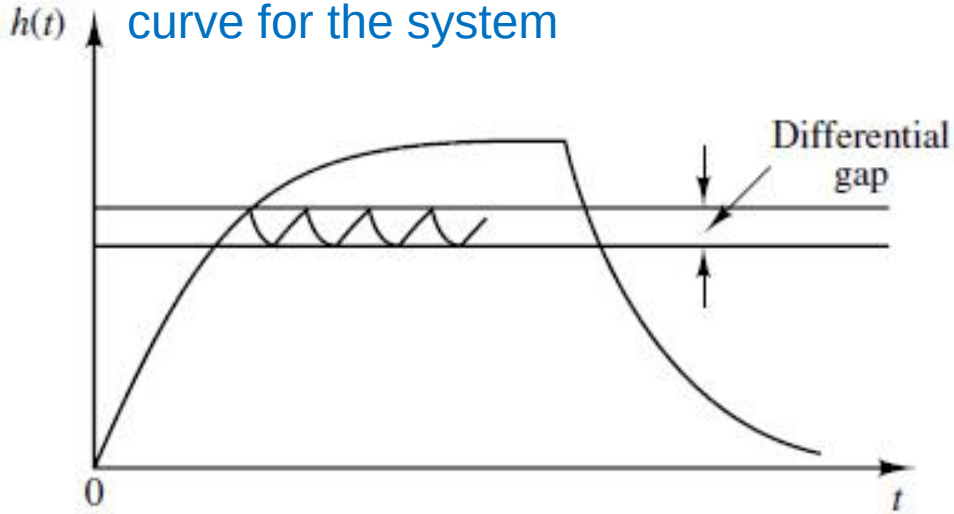


(b)

Liquid-level control system;

electromagnetic valve.

Level h(t)-versus-t curve for the system



Proportional Control Action. For a controller with proportional control action, the relationship between the output of the controller $u(t)$ and the actuating error signal $e(t)$ is

$$u(t) = K_p e(t)$$

or, in Laplace-transformed quantities,

$$\frac{U(s)}{E(s)} = K_p$$

where K_p is termed the proportional gain.

Whatever the actual mechanism may be and whatever the form of the operating power, the proportional controller is essentially an amplifier with an adjustable gain.

Integral Control Action. In a controller with integral control action, the value of the controller output $u(t)$ is changed at a rate proportional to the actuating error signal $e(t)$. That is,

$$\frac{du(t)}{dt} = K_i e(t)$$

or

$$u(t) = K_i \int_0^t e(t)\,dt$$

where K_i is an adjustable constant. The transfer function of the integral controller is

$$\frac{U(s)}{E(s)} = \frac{K_i}{s}$$

Proportional-Plus-Derivative Control Action. The control action of a proportional-plus-derivative controller is defined by

$$u(t) = K_p e(t) + K_p T_d \frac{de(t)}{dt}$$

and the transfer function is

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

where T_d is called the *derivative time*.

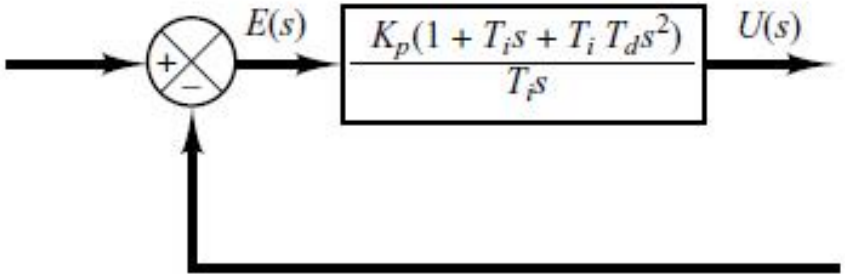
Proportional-Plus-Integral-Plus-Derivative Control Action. The combination of proportional control action, integral control action, and derivative control action is termed proportional-plus-integral-plus-derivative control action. It has the advantages of each of the three individual control actions. The equation of a controller with this combined action is given by

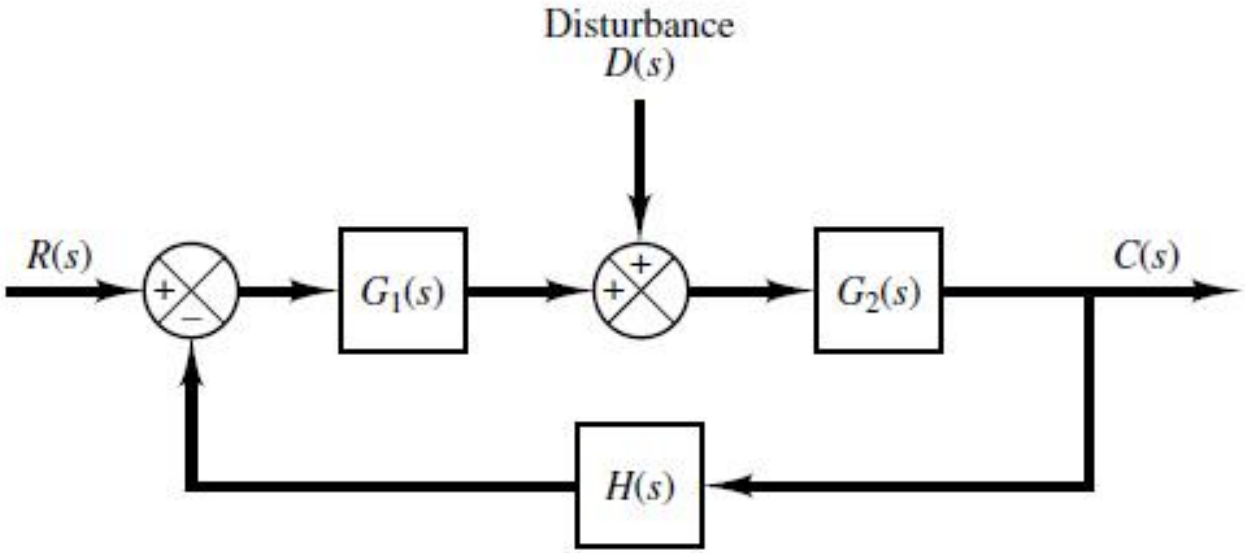
$$u(t) = K_p e(t) + \frac{K_p}{T_i}\int_0^t e(t)\,dt + K_p T_d \frac{de(t)}{dt}$$

or the transfer function is

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

where K_p is the proportional gain, T_i is the integral time, and T_d is the derivative time. The block diagram of a proportional-plus-integral-plus-derivative controller is shown in Figure 2–10.





Closed-Loop System Subjected to a Disturbance. Figure 2–11 shows a closed-loop system subjected to a disturbance. When two inputs (the reference input and disturbance) are present in a linear time-invariant system, each input can be treated independently of the other; and the outputs corresponding to each input alone can be added to give the complete output. The way each input is introduced into the system is shown at the summing point by either a plus or minus sign.

Consider the system shown in Figure 2–11. In examining the effect of the disturbance $D(s)$, we may assume that the reference input is zero; we may then calculate the response $C_D(s)$ to the disturbance only. This response can be found from

$$\frac{C_D(s)}{D(s)} = \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}$$

On the other hand, in considering the response to the reference input $R(s)$, we may assume that the disturbance is zero. Then the response $C_R(s)$ to the reference input $R(s)$ can be obtained from

$$\frac{C_R(s)}{R(s)} = \frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)}$$

The response to the simultaneous application of the reference input and disturbance can be obtained by adding the two individual responses. In other words, the response $C(s)$ due to the simultaneous application of the reference input $R(s)$ and disturbance $D(s)$ is given by

$$\begin{aligned} C(s) &= C_R(s) + C_D(s) \\ &= \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}\left[G_1(s)R(s) + D(s)\right] \end{aligned}$$

Consider now the case where $|G_1(s)H(s)| \gg 1$ and $|G_1(s)G_2(s)H(s)| \gg 1$. In this case, the closed-loop transfer function $C_D(s)/D(s)$ becomes almost zero, and the effect of the disturbance is suppressed. This is an advantage of the closed-loop system.

On the other hand, the closed-loop transfer function $C_R(s)/R(s)$ approaches $1/H(s)$ as the gain of $G_1(s)G_2(s)H(s)$ increases. This means that if $|G_1(s)G_2(s)H(s)| \gg 1$, then the closed-loop transfer function $C_R(s)/R(s)$ becomes independent of $G_1(s)$ and $G_2(s)$ and inversely proportional to $H(s)$, so that the variations of $G_1(s)$ and $G_2(s)$ do not affect the closed-loop transfer function $C_R(s)/R(s)$. This is another advantage of the closed-loop system. It can easily be seen that any closed-loop system with unity feedback, $H(s) = 1$, tends to equalize the input and output.

Procedures for Drawing a Block Diagram. To draw a block diagram for a system, first write the equations that describe the dynamic behavior of each component. Then take the Laplace transforms of these equations, assuming zero initial conditions, and represent each Laplace-transformed equation individually in block form. Finally, assemble the elements into a complete block diagram.

As an example, consider the *RC* circuit shown in Figure 2–12(a). The equations for this circuit are

$$i = \frac{e_i - e_o}{R} \tag{2–4}$$

$$e_o = \frac{\int i\,dt}{C} \tag{2–5}$$

The Laplace transforms of Equations (2–4) and (2–5), with zero initial condition, become

$$I(s) = \frac{E_i(s) - E_o(s)}{R} \tag{2–6}$$

$$E_o(s) = \frac{I(s)}{Cs} \tag{2–7}$$

Equation (2–6) represents a summing operation, and the corresponding diagram is shown in Figure 2–12(b). Equation (2–7) represents the block as shown in Figure 2–12(c). Assembling these two elements, we obtain the overall block diagram for the system as shown in Figure 2–12(d).

Familiar forms: cascade, parallel, feedback

Block Diagram Reduction. It is important to note that blocks can be connected in series only if the output of one block is not affected by the next following block. If there are any loading effects between the components, it is necessary to combine these components into a single block.

Any number of cascaded blocks representing nonloading components can be replaced by a single block, the transfer function of which is simply the product of the individual transfer functions.

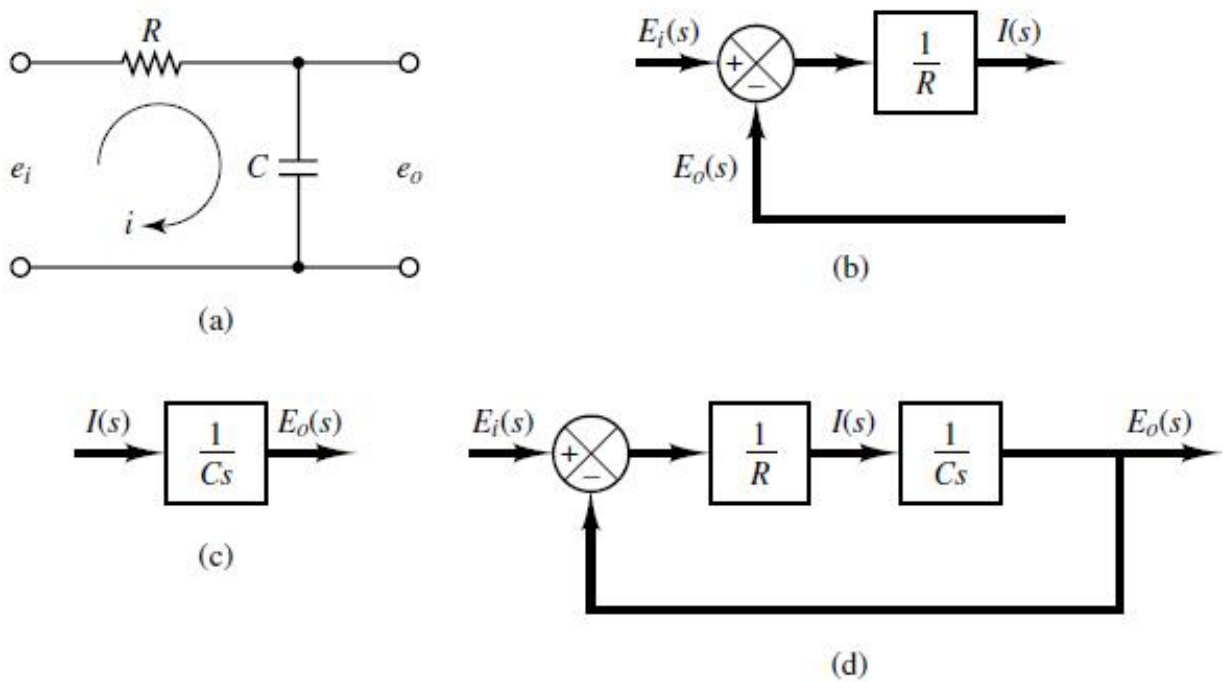
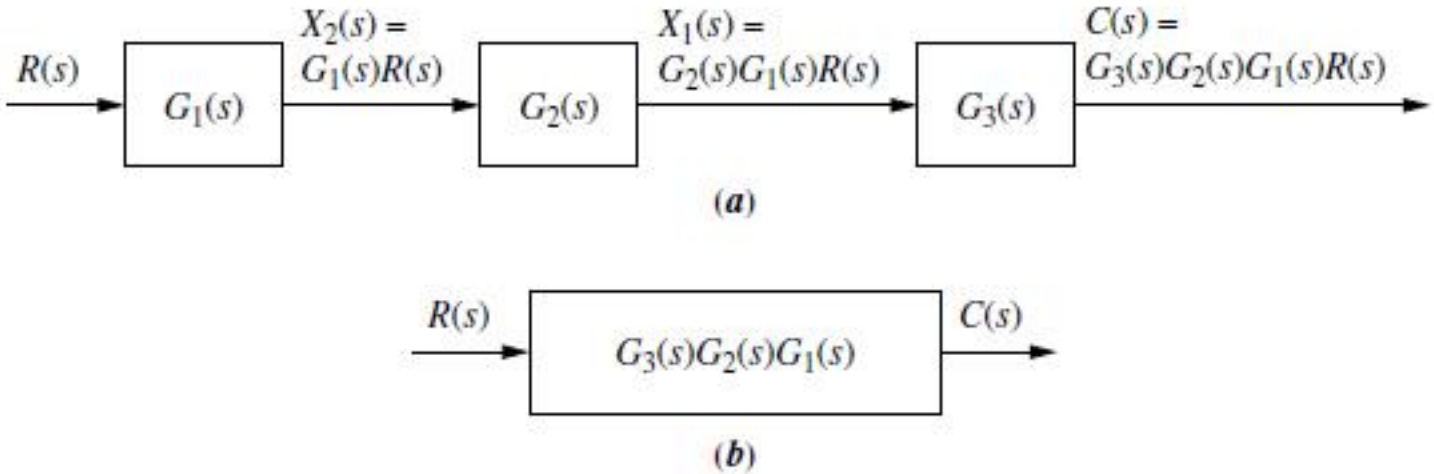


Figure 2–12
(a) *RC* circuit;
(b) block diagram representing Equation (2–6);
(c) block diagram representing Equation (2–7);
(d) block diagram of the *RC* circuit.

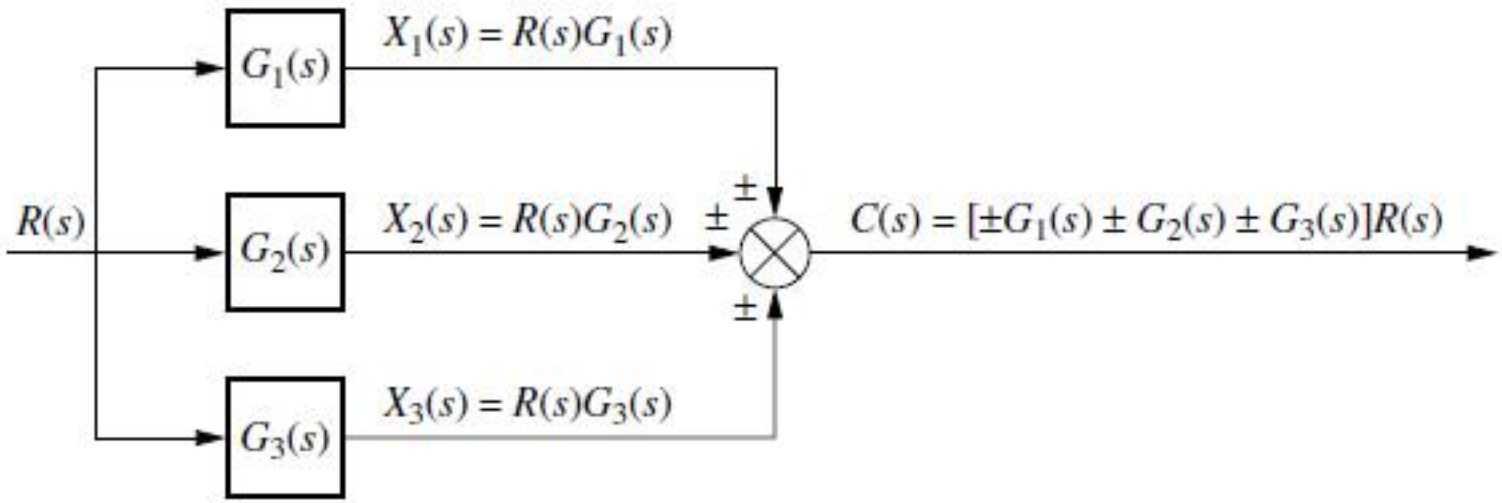
1. Cascade Form



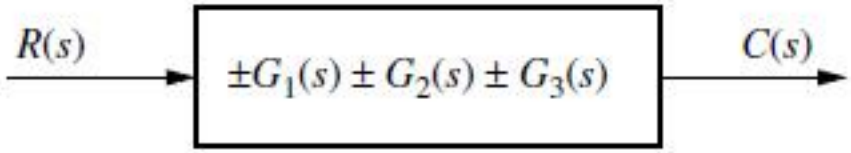
transform from Figure 5.3(a), or

$$\boxed{G_e(s) = G_3(s)G_2(s)G_1(s)}$$

2. Parallel Form



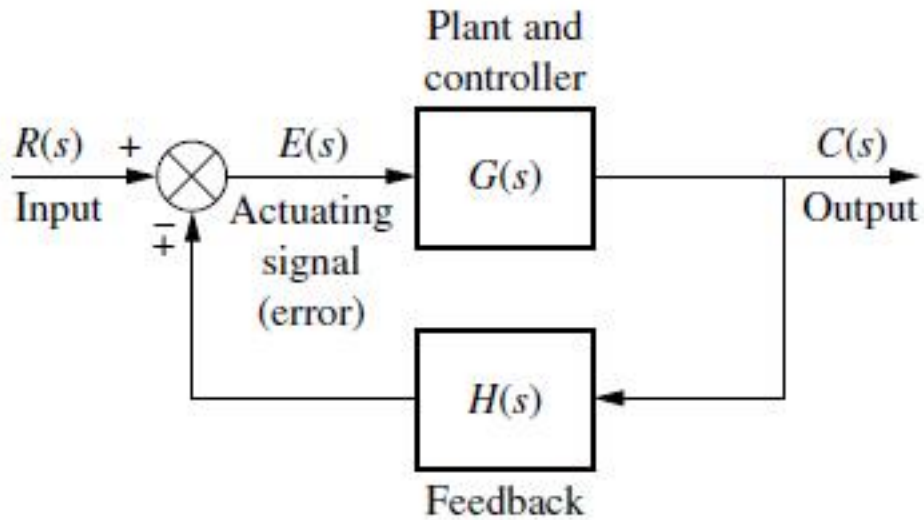
(a)



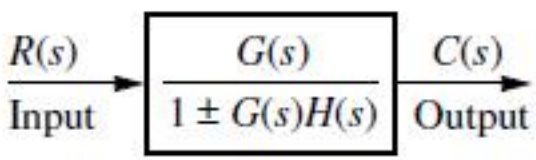
(b)

$$G_e(s) = \pm G_1(s) \pm G_2(s) \pm G_3(s)$$

3. Feedback Form



(b)



(c)

$$G_e(s) = \frac{G(s)}{1 \pm G(s)H(s)}$$

[1] The system is said to have *negative feedback* if the sign at the summing junction is negative and *positive feedback* if the sign is positive.

Moving Blocks to Create Familiar Forms

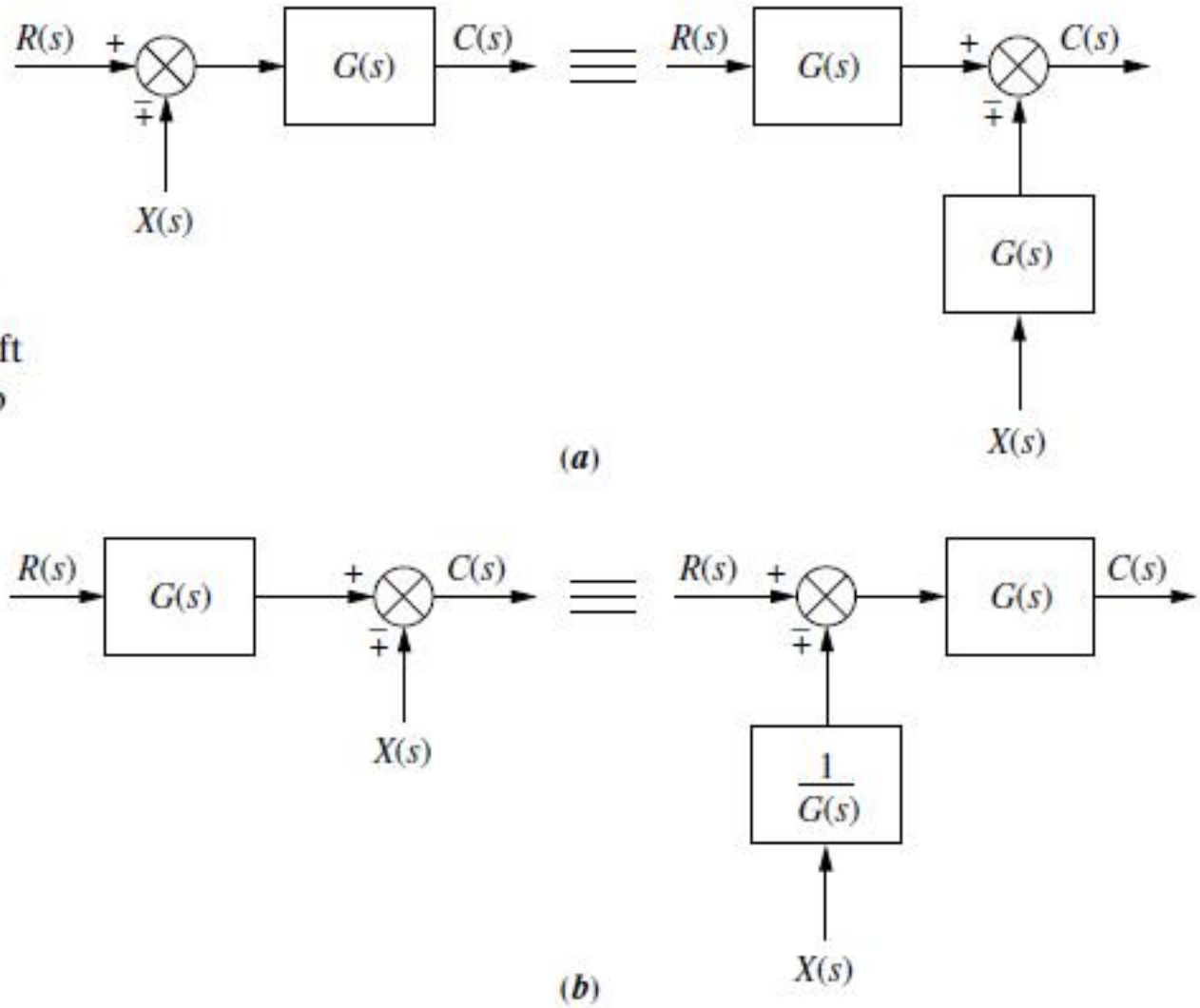


FIGURE 5.7 Block diagram algebra for summing junctions—equivalent forms for moving a block **a.** to the left past a summing junction; **b.** to the right past a summing junction

Moving Blocks to Create Familiar Forms

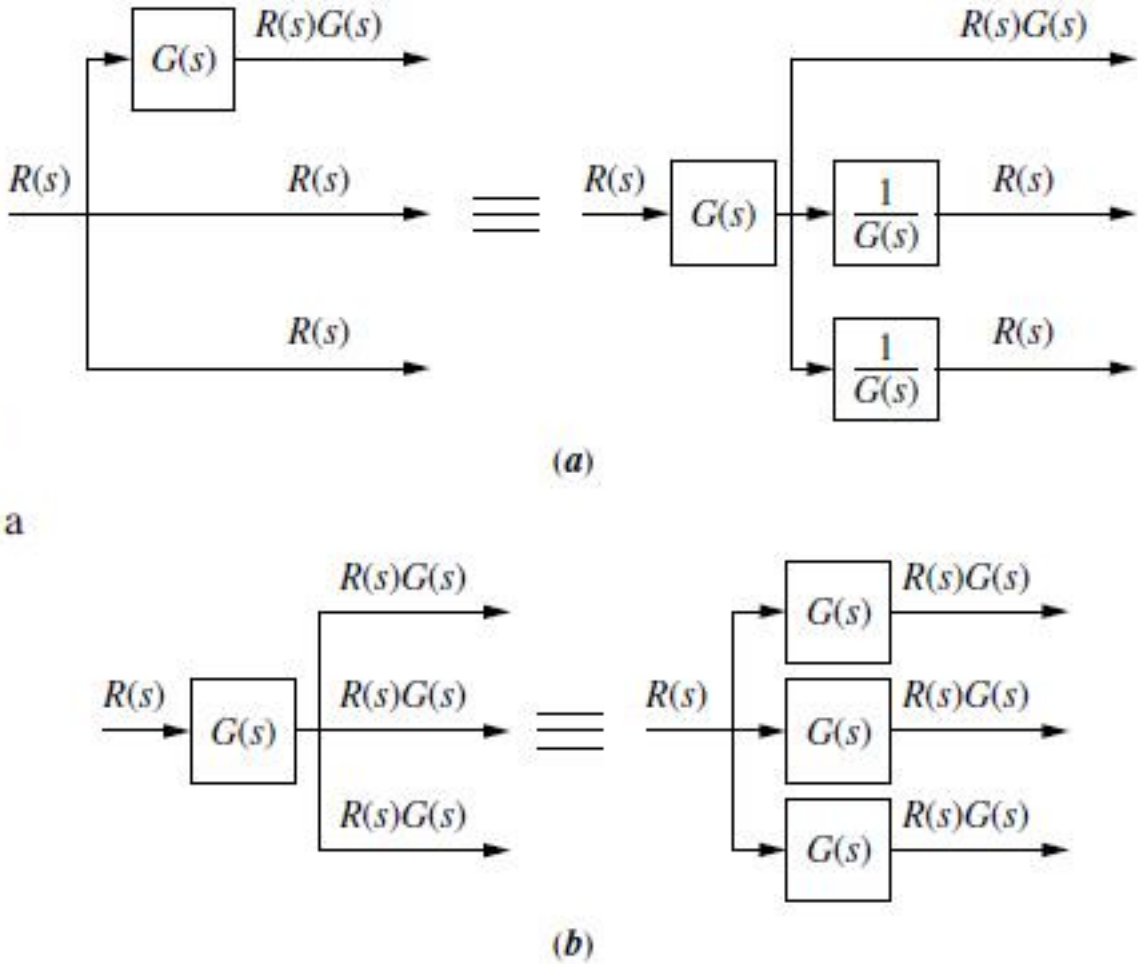
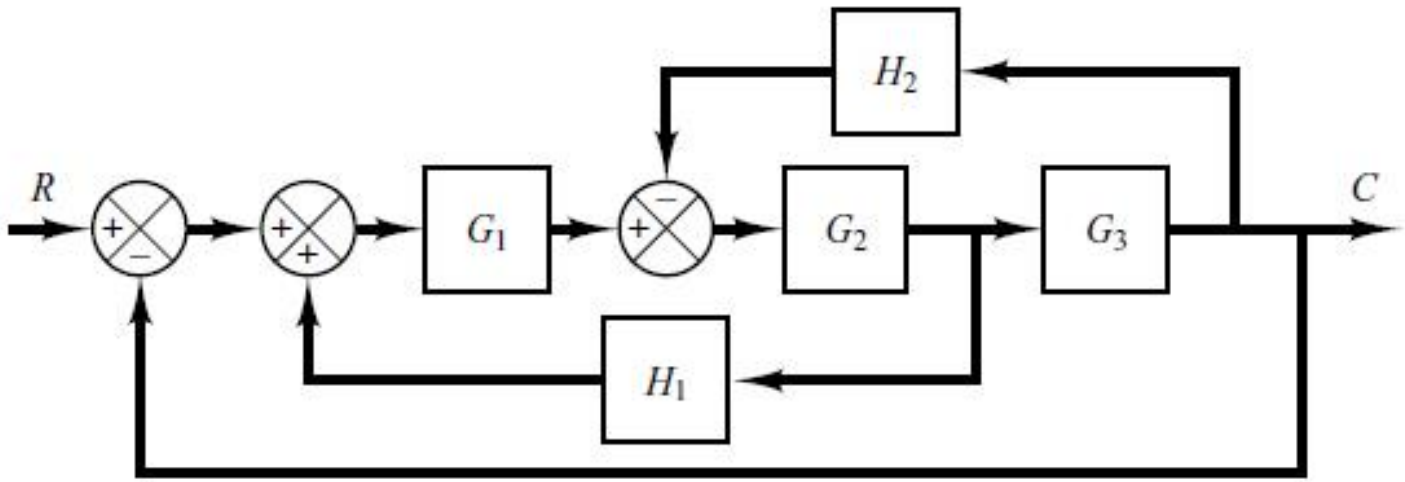


FIGURE 5.8 Block diagram algebra for pickoff points—equivalent forms for moving a block **a.** to the left past a pickoff point; **b.** to the right past a pickoff point

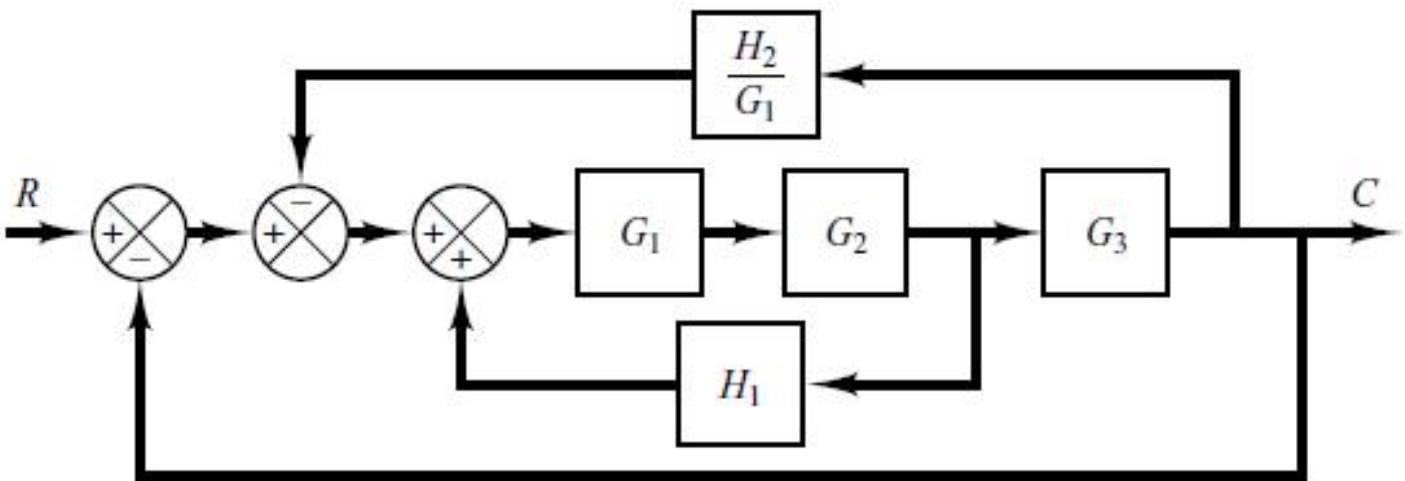
Consider the system shown in Figure 2–13(a). Simplify this diagram.

By moving the summing point of the negative feedback loop containing H_2 outside the positive feedback loop containing H_1, we obtain Figure 2–13(b). Eliminating the positive feedback loop, we have Figure 2–13(c). The elimination of the loop containing H_2/G_1 gives Figure 2–13(d). Finally, eliminating the feedback loop results in Figure 2–13(e).

(a)



(b)



Block diagram simplification: as the block diagram is simplified, the transfer functions in new blocks become more complex: new poles and zeros generated.

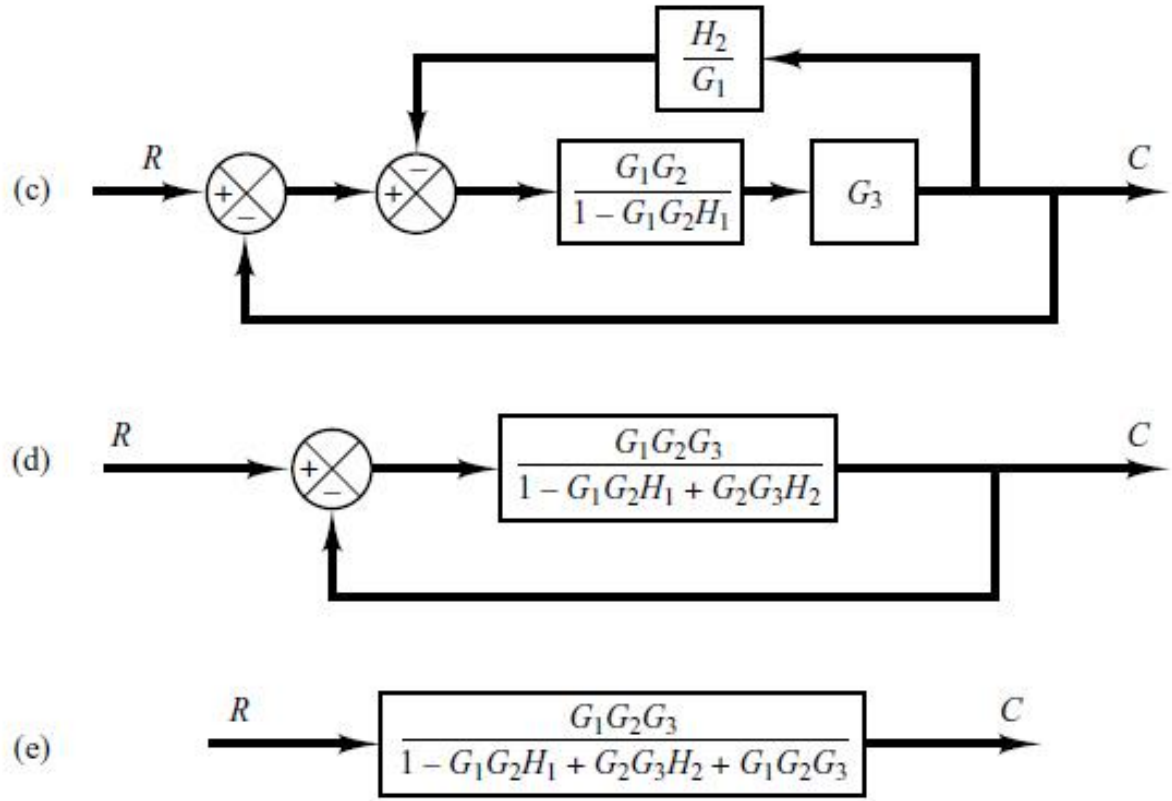


Figure 2–13
(a) Multiple-loop system;
(b)–(e) successive reductions of the block diagram shown in (a).

Notice that the numerator of the closed-loop transfer function $C(s)/R(s)$ is the product of the transfer functions of the feedforward path. The denominator of $C(s)/R(s)$ is equal to

$$1 + \sum \text{(product of the transfer functions around each loop)}$$

$$= 1 + (-G_1G_2H_1 + G_2G_3H_2 + G_1G_2G_3)$$

$$= 1 - G_1G_2H_1 + G_2G_3H_2 + G_1G_2G_3$$

Linearize the nonlinear equation

$$z = xy$$

in the region $5 \le x \le 7, 10 \le y \le 12$. 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 $5 \le x \le 7, 10 \le y \le 12$, 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%.

Modeling in State Space

Modern Control Theory. The modern trend in engineering systems is toward greater complexity, due mainly to the requirements of complex tasks and good accuracy. Complex systems may have multiple inputs and multiple outputs and may be time varying. Because of the necessity of meeting increasingly stringent requirements on the performance of control systems, the increase in system complexity, and easy access to large scale computers, modern control theory, which is a new approach to the analysis and design of complex control systems, has been developed since around 1960.This new approach is based on the concept of state. The concept of state by itself is not new, since it has been in existence for a long time in the field of classical dynamics and other fields.

Modeling in State Space

Modern Control Theory Versus Conventional Control Theory.

Modern control theory is contrasted with conventional control theory in that the former is applicable to multiple-input, multiple-output systems, which may be linear or nonlinear,
time invariant or time varying, while the latter is applicable only to linear time invariant single-input, single-output systems.
Also, modern control theory is essentially time-domain approach and frequency domain approach (in certain cases such as H-infinity control), while conventional control theory is a complex frequency-domain approach.
Before we proceed further, we must define state, state variables, state vector, and state space.

Modeling in State Space

State. The state of a dynamic system is the smallest set of variables (called *state variables*) such that knowledge of these variables at $t=t_0$, together with knowledge of the input for $t >= t_0$, completely determines the behavior of the system for any time $t >= t_0$. Note that the concept of state is by no means limited to physical systems. It is applicable to biological systems, economic systems, social systems, and others.

State Variables. The state variables of a dynamic system are the variables making up the smallest set of variables that determine the state of the dynamic system. If at least *n* variables x1, x2 ,… , xn are needed to completely describe the behavior of a dynamic system (so that once the input is given for t >= t0 and the initial state at t=t0 is specified, the future state of the system is completely determined), then such n variables are a set of state variables. Note that state variables need not be physically measurable or observable quantities. Variables that are neither measurable nor observable can be chosen too as state variables. This is an advantage of the state-space methods. Practically, however, it is convenient to choose easily measurable quantities for the state variables, if this is possible at all, because optimal control laws will require the feedback of all state variables with suitable weighting.

Modeling in State Space

State Vector. If n state variables are needed to completely describe the behavior of a given system, then these n state variables can be considered the n components of a vector **x**. Such a vector is called a *state vector*. A state vector is thus a vector that determines uniquely the system state **x**(t) for any time t t_0, once the state at $t=t_0$ is given and the input u(t) for t t_0 is specified.

State Space. The n-dimensional space whose coordinate axes consist of the x_1 axis, x_2 axis, … , x_n axis, where x_1, x_2, … , x_n are state variables, is called a *state space*.

Any state can be represented by a point in the state space.

Modeling in State Space

State-Space Equations. In state-space analysis we are concerned with three types of variables that are involved in the modeling of dynamic systems: input variables, output variables, and state variables. The state-space representation for a given system is not unique, except that the number of state variables is the same for any of the different state-space representations of the same system.

The dynamic system must involve elements that memorize the values of the input for $t >= t_1$. Since integrators in a continuous-time control system serve as memory devices, the outputs of such integrators can be considered as the variables that define the internal state of the dynamic system. Thus the outputs of integrators serve as state variables. The number of state variables to completely define the dynamics of the system is equal to the number of integrators involved in the system.

Modeling in State Space

Assume that a MIMO system involves n integrators. Assume also that there are r inputs $u_1(t)$, $u_2(t)$, … , $u_r(t)$ and m outputs $y_1(t)$, $y_2(t)$, … , $y_m(t)$. Define n outputs of the integrators as state variables: $x_1(t)$, $x_2(t)$, … , $x_n(t)$ The system may be described by

$$\dot{x}_1(t) = f_1(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t)$$
$$\dot{x}_2(t) = f_2(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t)$$
$$\vdots$$
$$\dot{x}_n(t) = f_n(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t)$$

The outputs y1(t), y2(t), … , ym(t) of the system may be given by:

$$y_1(t) = g_1(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t)$$
$$y_2(t) = g_2(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t)$$
$$\vdots$$
$$y_m(t) = g_m(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t)$$

If we define:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \cdot \\ \cdot \\ \cdot \\ x_n(t) \end{bmatrix}, \qquad \mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t) \\ f_2(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t) \\ \cdot \\ \cdot \\ \cdot \\ f_n(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t) \end{bmatrix},$$

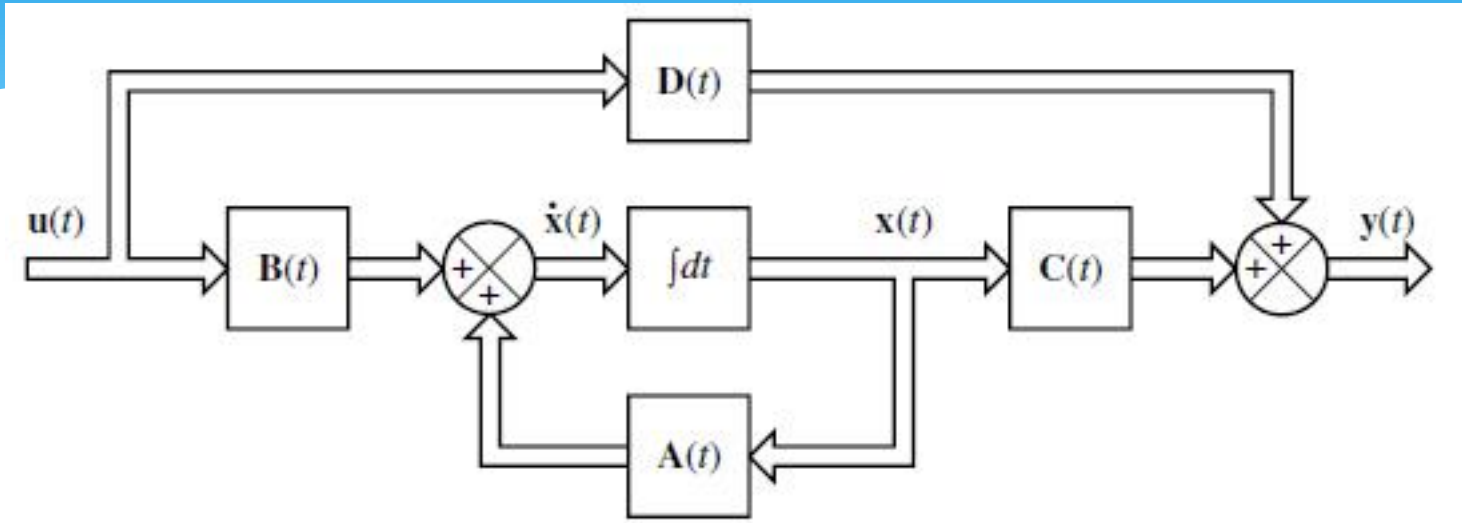
$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \cdot \\ \cdot \\ \cdot \\ y_m(t) \end{bmatrix}, \qquad \mathbf{g}(\mathbf{x}, \mathbf{u}, t) = \begin{bmatrix} g_1(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t) \\ g_2(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t) \\ \cdot \\ \cdot \\ \cdot \\ g_m(x_1, x_2, \ldots, x_n; u_1, u_2, \ldots, u_r; t) \end{bmatrix}, \qquad \mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \cdot \\ \cdot \\ \cdot \\ u_r(t) \end{bmatrix}$$

then Equations (2–8) and (2–9) become:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$$ → (the state equation)

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}, \mathbf{u}, t)$$ →(the output equation)

Block diagram of the linear, continuous time control system represented in State Space



If vector functions **f** and/or **g** involve time t explicitly, then the system is called a time - varying system.

If Equations (2–10) and (2–11) are linearized about the operating state, then we have the following linearized state equation and output equation:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t)$$

where **A**(t) is called the state matrix, **B**(t) the input matrix, **C**(t) the output matrix, and **D**(t) the direct transmission matrix.

mechanical system represented in State Space

EXAMPLE 2–2 Consider the mechanical system shown in Figure 2–15. We assume that the system is linear. The external force $u(t)$ is the input to the system, and the displacement $y(t)$ of the mass is the output. The displacement $y(t)$ is measured from the equilibrium position in the absence of the external force. This system is a single-input, single-output system.

From the diagram, the system equation is

$$m\ddot{y} + b\dot{y} + ky = u \tag{2–16}$$

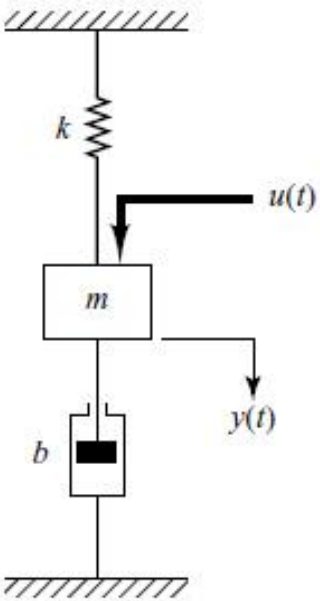


Figure 2–15
Mechanical system.

This system is of second order. This means that the system involves two integrators. Let us define state variables $x_1(t)$ and $x_2(t)$ as

$$x_1(t) = y(t)$$
$$x_2(t) = \dot{y}(t)$$

Then we obtain

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = \frac{1}{m}(-ky - b\dot{y}) + \frac{1}{m}u$$

or

$$\dot{x}_1 = x_2 \tag{2–17}$$
$$\dot{x}_2 = -\frac{k}{m}x_1 - \frac{b}{m}x_2 + \frac{1}{m}u \tag{2–18}$$

The output equation is

$$y = x_1 \tag{2–19}$$

mechanical system represented in State Space

In a vector-matrix form, Equations (2–17) and (2–18) can be written as

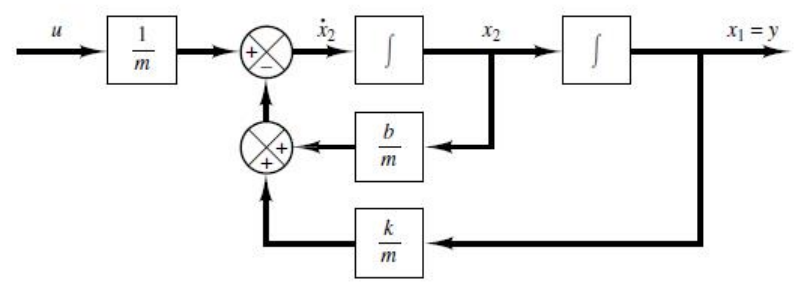
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\dfrac{k}{m} & -\dfrac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{1}{m} \end{bmatrix} u \tag{2–20}$$

The output equation, Equation (2–19), can be written as

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{2–21}$$

Equation (2–20) is a state equation and Equation (2–21) is an output equation for the system. They are in the standard form:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$$
$$y = \mathbf{C}\mathbf{x} + Du$$



where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\dfrac{k}{m} & -\dfrac{b}{m} \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 \\ \dfrac{1}{m} \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad D = 0$$

Figure 2–16 is a block diagram for the system. Notice that the outputs of the integrators are state variables.