Note 9

Closed-Loop Control

1. Introduction to Closed-Loop Control

1.1 Open-loop control vs. closed-loop control

Open-loop control systems. Those systems in which the output has no effect on the control action are called *open-loop control systems.* In other words, in an open-loop control system the output is neither measured nor fed back for comparison with the input. One practical example is a washing machine. Soaking, washing, and rinsing in the washer operate on a time basis. The machine does not measure the output signal, that is, the cleanliness of the clothes. Figure 1 (a) shows the schematic diagram of an open-loop control system.

Closed-loop control systems. Feedback control systems are often referred to as *closed-loop control systems.* In practice, the terms feedback control and closed-loop control are used interchangeably. In a closed-loop control system the actuating error signal, which is the difference between the input signal and the feedback signal (which may be the output signal itself or a function of the output signal and its derivatives and/or integrals), is fed to the controller so as to reduce the error and bring the output of the system to a desired value. The term closed-loop control always implies the use of feedback control action in order to reduce system error. Figure 1 (b) shows the schematic diagram of an closed-loop control system.
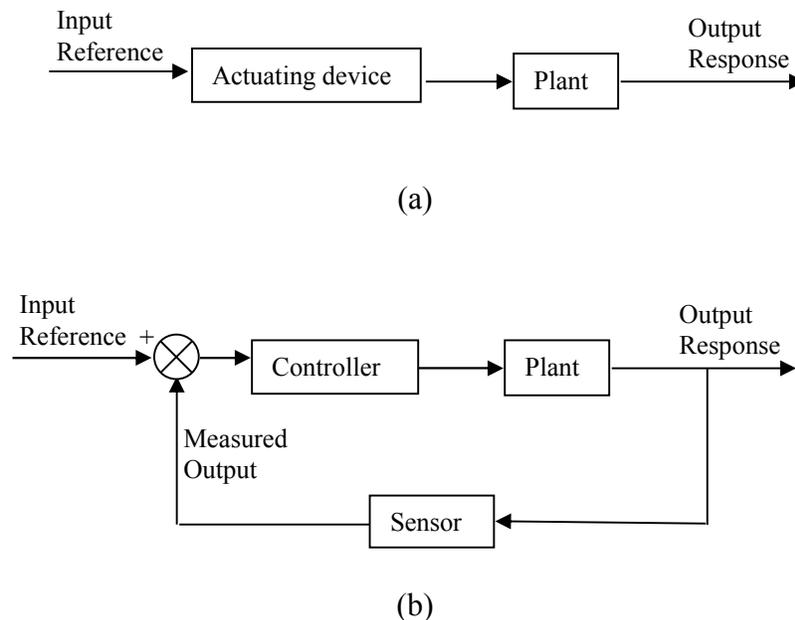

(a)

(b)

Figure 1: Schematic diagrams of (a) an open-loop control and (b) a closed-loop control system.

Closed-loop versus open-loop control systems. An advantage of the closed-loop control system is the fact that the use of feedback makes the system response relatively insensitive to external disturbances and internal variations in system parameters. It is thus possible to use relatively inaccurate and inexpensive components to obtain the accurate control of a given plant, whereas doing so is impossible in the open-loop case. From the stability point of view, the open-loop control system is easier to build because system stability is not a major problem. On the other hand, stability is a major problem in the closed-loop control system, which may tend to overcorrect errors that can cause oscillations of constant or changing amplitude.

1.2 Time response of control systems

The performance of a control system can be specified in terms of time response. The *time response* of a control system consists of two parts: the transient and steady-state responses. *Transient response* is the part where the output goes from the initial state to the final state and remains there. *Stead-state response* is the part of the output that proceeds the transient phase. Before reaching steady-state, the transient response of a control system often exhibits the dynamics of a second-order underdamped system. Thus, most system specifications are often given assuming that the system is second order. In addition, for the analysis of the high-order system, dominant pole techniques may be used to approximate the system with a second-order transfer function. The generalized form of the closed-loop second-order system is shown Figure 2; and the transfer function is given by

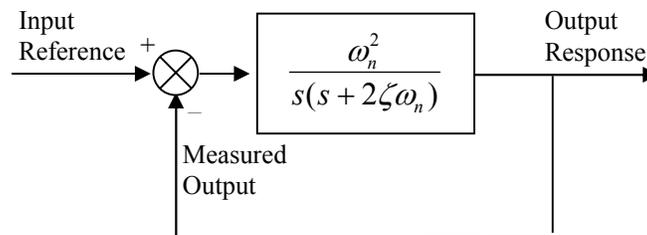$$G(S) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

Figure 2: Unity-feedback closed-loop second-order system.

The dynamics behavior of the second-order system is described in terms of two parameters: the natural frequency of ω_n and the damping ratio of ζ. Figure 3 shows is a typical step response of such a second-order system, and the specifications are given in the following.
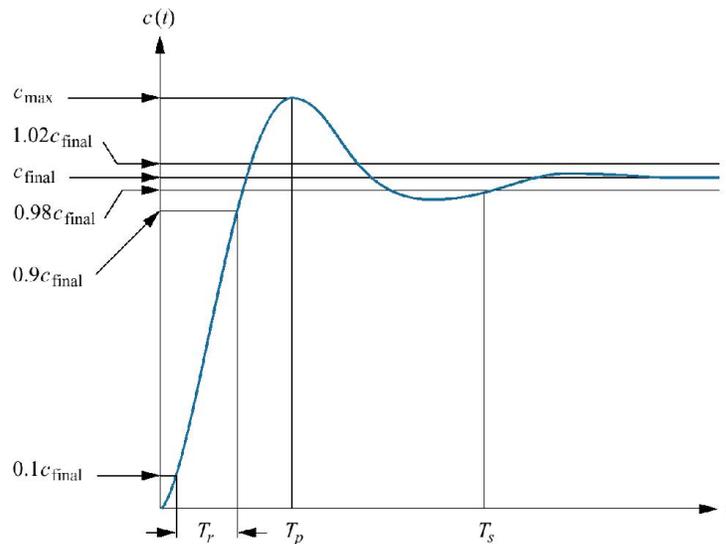

Figure 3: Typical step response of a control system.

Peak time, T_P. The peak time is the time required for the response to reach the first peak, which is given by

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

Percent overshoot, %*OS*. The percent overshoot is defined as the amount that the waveform at the peak time overshoots the steady-state value, which is expressed as a percentage of the steady-state value.

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

Settling time, T_s. The settling time is the time required for the amplitude of the sinusoid to decay to 2% of the steady-state value.

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

In most control design, the interest is specifically in the final value, or steady-state output value. This is known as *steady-state accuracy*. Ideally, in the steady-state phase, the system tracks the input reference and the steady-state error is zero. The s*teady-state error* is the difference between the input and the output for a prescribed test input (typically a step) as $t \to \infty$. Consider the following feedback system
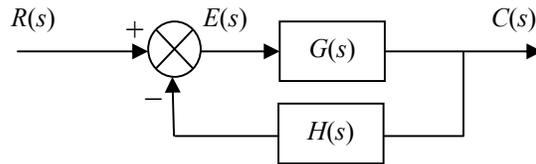
Recall the closed-loop transfer function of the system, $T(s)$, is

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

The system output can then be expressed as

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

The error, $E(s)$, between the system input, $R(s)$, and the measured output, $C(s)$ $H(s)$, is

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

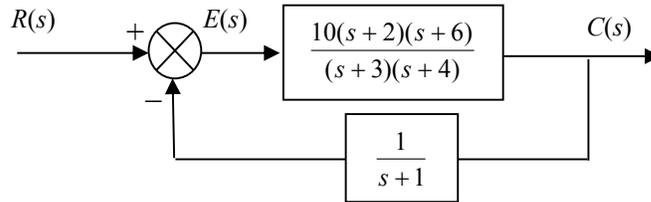Using the final value theorem, the steady-state error is

$$e(\infty) = \lim_{s \to 0} \frac{sR(s)}{1 + G(s)H(s)}$$

For a step input, i.e., $r(t) = u(t)$, its Laplace transform is $R(s) = \frac{1}{s}$, the steady-state error is

$$e(\infty) = \lim_{s \to 0} \frac{s(1/s)}{1 + G(s)} = \frac{1}{1 + \lim_{s \to 0} G(s)H(s)}.$$

Example 1

For the feedback system shown in the following figure, find the steady-state error if the input is $5u(t)$.

$$R(s) \quad + \quad E(s) \quad \frac{10(s+2)(s+6)}{(s+3)(s+4)} \quad C(s)$$

$$\frac{1}{s+1}$$

1.3 Basic feedback control types

Figure 4 shows the three basic feedback control actions: (a) proportional, (2) integral, and (c) derivative control actions. Proportional control action is generated based on the current error, the integral control action is generated based on the past error, and the derivate control action is generated based on the anticipated future error. Integral of the error can be interpreted as the past information about it. Derivative of the error can be interpreted as a measure of future error. Assume that the error signal entering the controller has a trapezoidal form, the control actions generated by the proportional, integral, and derivative actions are shown in Figure 5.
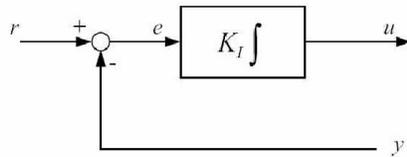
Proportional control (P)
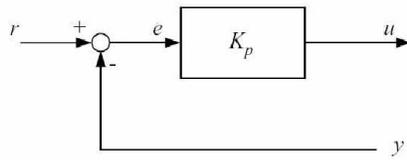
$$u(t) = K_p e(t)$$

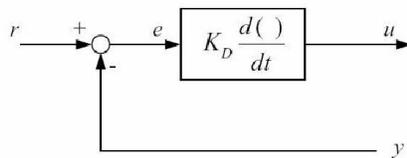$$D(s) = \frac{u(s)}{e(s)} = K_p$$

Integral control (I)

$$u(t) = K_I \int_0^t e(\tau)d\tau$$

$$D(s) = \frac{u(s)}{e(s)} = \frac{K_I}{s}$$

Derivative control (D)

$$u(t) = K_D \frac{d}{dt}\left(e(t)\right)$$

$$D(s) = \frac{u(s)}{e(s)} = K_D s$$

Figure 4: Basic feedback control actions: proportional, integral, and derivative.

Based on the aforementioned three feedback control action, various controllers can be constituted, such as *proportional (P) controller*, *proportional-integral (PI) controller*, *proportional-derivative (PD) controller*, and *proportional-integral-derivative (PID) controller*, to achieve desired control performance. As an example, Figure 6 shows the block diagram of a PID controller. The control algorithm can be expressed in both continuous (analog) time domain (which can be implemented using op-amps) and in discrete (digital) time domain (which can be implemented using digital computers). At any given time, *t*, the control signal, *u(t)* is determined as a function

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau)d\tau + K_D \dot{e}(t)$$

which shows that the control signal is a function of the error between the commanded and measured output signals, $e(t)$ at time t, as well as the derivative of the error signal, $\dot{e}(t)$, and the integral of the error signal since the control loop is enable $(t = 0)$, $\int_0^t e(\tau)d\tau$.
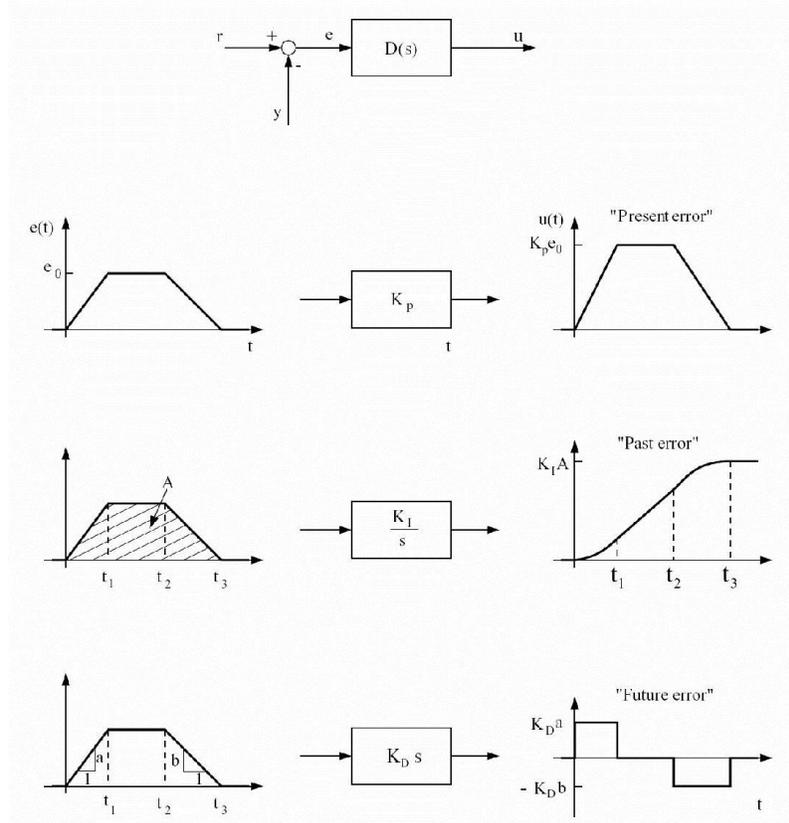

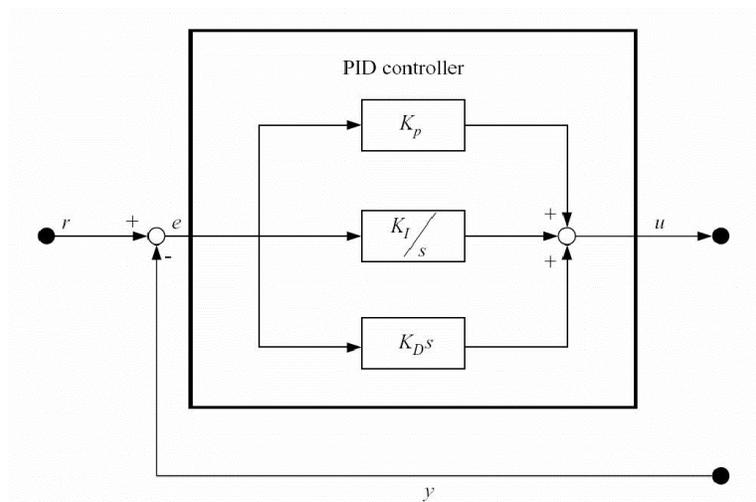Figure 5: Illustration of the input-output behavior of the basic feedback control actions.

Figure 6: Block diagram of a PID controller.

Design of controller via root locus that is introduced in ME 431 is one of the methods that we can use. Design via root locus requires that the designer has the mathematic model of the plant. However, if the plant is complicated, a mathematical model cannot be easily be derived, and analytical approach to the design of a controller is not possible. In such cases, experimental methods need to be used. Ziegler and Nichols developed such a method for tuning the controller parameters based on experiments, called *ultimate sensitivity method*. In this method, the criteria for adjusting the parameters are based on evaluating the amplitude and frequency of oscillations of the system at the limit of stability. To use the method, the proportional gain is increased until the system becomes marginally stable, as shown in Figure 7. The corresponding gain is defined as K_u (called the ultimate gain) and the period of oscillation is P_u (called the ultimate period). P_u should be measured when the amplitude of oscillation is as same as possible. The tuning parameters are selected as shown in the following Table 1.
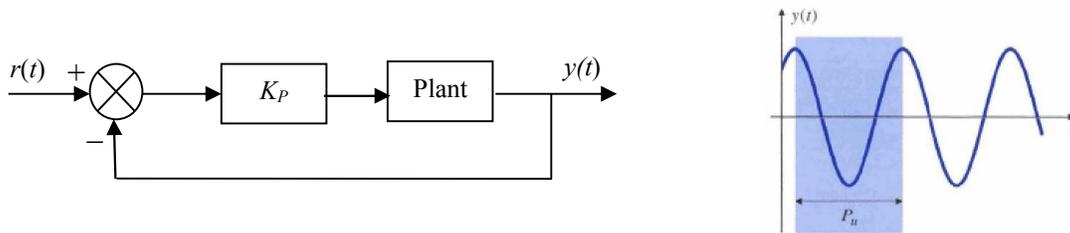

Figure 7: Determination of ultimate gain and period.

Table 1: Ziegler-Nichols tuning for the controller $D_c(s) = K_P \left(1 + \dfrac{1}{T_I s} + T_D s \right)$.

Type of Controller	Optimum Gain
Proportional	$k_p = 0.5 K_u$
PI	$\begin{cases} k_p = 0.45 K_u, \\ T_I = \frac{P_u}{1.2} \end{cases}$
PID	$\begin{cases} k_p = 0.6 K_u, \\ T_I = \frac{1}{2} P_u, \\ T_D = \frac{1}{8} P_u \end{cases}$

An example of using the *ultimate sensitivity method* is illustrated in class.

2. Design Via State Space

4.1 State-Space Representation

Consider a mechanical system given in the following, with $M = 1$ kg, $K = 100$ N/m and $f_v = 12$ N-s/m
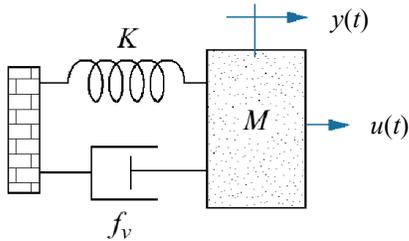
Linear combination. A linear combination of n variables, x_i, for $i = 1$ to n, is given by the following sum, S:

$$S = K_n x_n + K_{n-1} x_{n-1} + \cdots + K_1 x_1$$

Linear independence. A set of variable is said to be linearly independence if none of the variables can be written as a linear combination of the others.

System variable. Any variable that responds to an input or initial conditions in a system.

State variables. The smallest set of linearly independent system variables.

State vector. A vector whose elements are the state variables.

State equations. A set of n simultaneous, first-order differential equation with n variables, where the n variables to be solved are the state variables.

Output equation. The algebraic equation that expresses the output variables of a system as linear combinations of the state variables and inputs.

A system is represented in state space by the following equations:

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

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

where
$\mathbf{x}$ = state vector
$\dot{\mathbf{x}}$ = derivative of the state vector with respect to time
$\mathbf{y}$ = output vector
$\mathbf{u}$ = input or control vector
$\mathbf{A}$ = system matrix
$\mathbf{B}$ = input matrix
$\mathbf{C}$ = output matrix
$\mathbf{D}$ = feedforward matrix

Consider the differential equation

$$\frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_1\frac{dy}{dt} + a_0 y = b_0 u \qquad (3.48)$$

A convenient way to choose state variables is to choose the output, $y(t)$, and its $(n-1)$ derivatives as the state variables. This choice is called the *phase-variable choice*. Choosing the state variables, x_i, we get

$$x_1 = y \qquad (3.49a)$$

$$x_2 = \frac{dy}{dt} \qquad (3.49b)$$

$$x_3 = \frac{d^2 y}{dt^2} \qquad (3.49c)$$

$$(3.49d)$$

$$x_n = \frac{d^{n-1} y}{dt^{n-1}}$$

and differentiating both sides yields

$$\dot{x}_1 = \frac{dy}{dt} \qquad (3.50a)$$

$$\dot{x}_2 = \frac{d^2 y}{dt^2} \qquad (3.50b)$$

$$\dot{x}_3 = \frac{d^3 y}{dt^3} \qquad (3.50c)$$

$$(3.50d)$$

$$\dot{x}_n = \frac{d^n y}{dt^n}$$

where the dot above the x signifies differentiation with respect to time.

Substituting the definitions of Eq. (3.49) into Eq. (3.50), the state equations are evaluated as

$$\dot{x}_1 = x_2 \qquad (3.51a)$$

$$\dot{x}_2 = x_3 \qquad (3.51b)$$

$$(3.51c)$$

$$x_{n-1} = x_n$$

$$\dot{x}_n = -a_0 x_1 - a_1 x_2 \cdots - a_{n-1} x_n + b_0 u \qquad (3.51d)$$

where Eq. (3.51d) was obtained from Eq. (3.48) by solving for $d^n y/dt^n$ and using Eq. (3.49). In vector-matrix form, Eq. (3.51) become

$$
\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} =
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\
\vdots & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\
-a_0 & -a_1 & -a_2 & -a_3 & -a_4 & -a_5 & \cdots & -a_{n-1}
\end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}
+
\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ b_0 \end{bmatrix} u
$$

(3.52)

Equation (3.52) is the phase-variable form of the state equations. This form is easily recognized by the unique pattern of 1's and 0's and the negative of the coefficients of the differential equation written in reverse order in the last row of the system matrix.
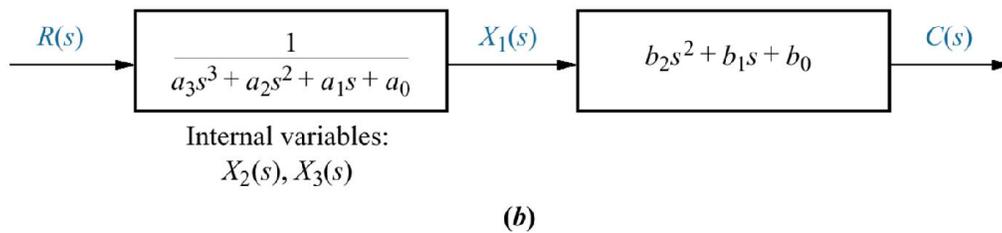
Finally, since the solution to the differential equation is $y(t)$, or x_1, the output equation is

$$
y = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}
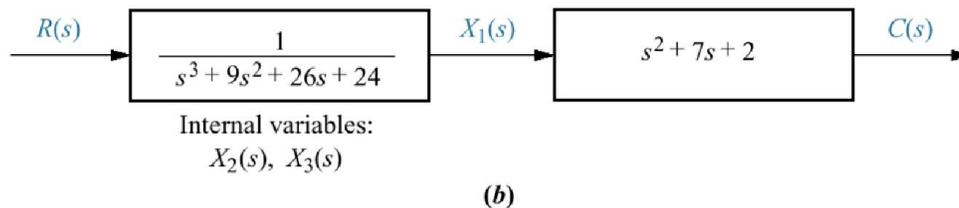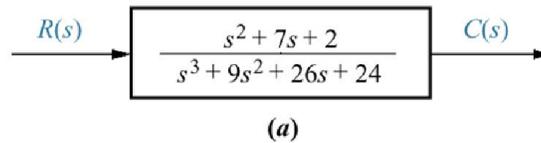\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}
$$

(3.53)

Example 2. Find the state-space representation of the transfer function shown in following figure.
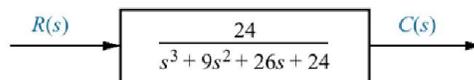
If a transfer function has a polynomial in s in the numerator that is of order less than the polynomial in the denominator, as shown below, the numerator and the denominator can be handled separately.

$$R(s) \rightarrow \boxed{\dfrac{b_2 s^2 + b_1 s + b_0}{a_3 s^3 + a_2 s^2 + a_1 s + a_0}} \rightarrow C(s)$$

(a)

$$R(s) \rightarrow \boxed{\dfrac{1}{a_3 s^3 + a_2 s^2 + a_1 s + a_0}} \xrightarrow{X_1(s)} \boxed{b_2 s^2 + b_1 s + b_0} \rightarrow C(s)$$

Internal variables:
$X_2(s), X_3(s)$

(b)
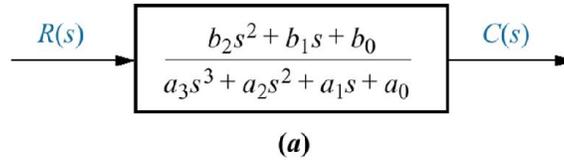
Example 3. Find the state-space representation of the transfer function shown in following figure.

$$R(s) \rightarrow \boxed{\dfrac{s^2 + 7s + 2}{s^3 + 9s^2 + 26s + 24}} \rightarrow C(s)$$

(a)

$$R(s) \rightarrow \boxed{\dfrac{1}{s^3 + 9s^2 + 26s + 24}} \xrightarrow{X_1(s)} \boxed{s^2 + 7s + 2} \rightarrow C(s)$$

Internal variables:
$X_2(s), \ X_3(s)$

(b)

4.2 Control Design Via State-Space

This section shows how to introduce additional parameters into a system so that we can control the location of all closed-loop poles. An *n*th-order feedback control system has a *n*th-order closed-loop characteristic equation of the form

$$s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 = 0$$

Since the coefficient of the highest power of *s* is unity, there are *n* coefficients whose values determine the system closed-loop pole locations. Thus, if we can introduce *n* adjustable parameters into the system and relate them to the coefficients in the above equation, all of the poles of the closed-loop system can be set to any desired location.

The idea behind the introduction of *n* adjustable parameters into the system will be illustrated in class.
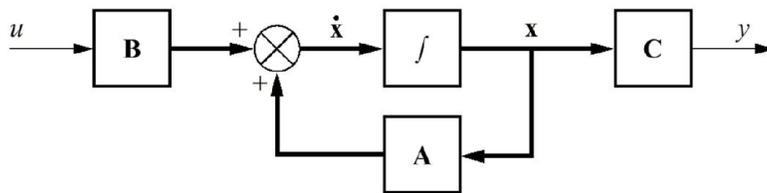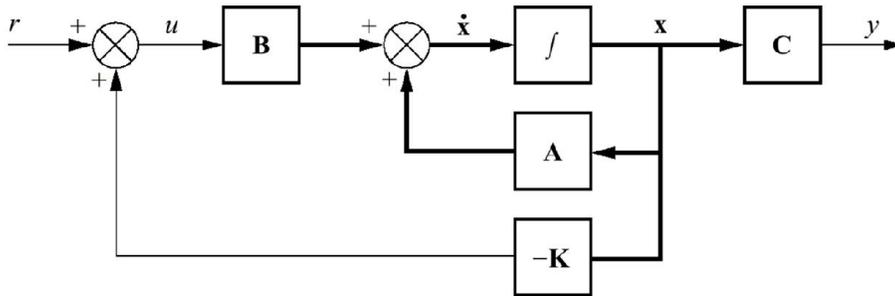
(a) State-space representation of a plant
(b) Plant with state-variable feedback

Design Problem:

Given the plant

$$G(s) = \frac{20(s+5)}{s(s+1)(s+4)}$$

Design the feedback gains to yield 9.5% overshoot and a settling time of 0.74 second.

3. Analog Implementation of Controllers

A controller can be realized by op-amps. Figure 8 shows such a circuit in general, and its transfer function is given by

$$\frac{V_o(z)}{V_i(z)} = -\frac{Z_2(s)}{Z_1(s)}$$

By judicious choice of $Z_1(s)$ and $Z_2(s)$, this circuit can be used as a building block to implement the controllers, such as PID controllers. Table 2 summarized the realization of common controllers.
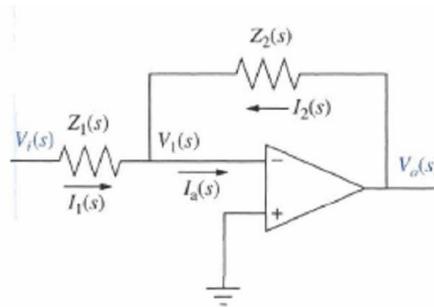

Figure 8: Op-amps for controller realization.

Table 2: Implementation of controllers using op-amps.

Function	$Z_1(s)$	$Z_2(s)$	$G_c(s) = -\dfrac{Z_2(s)}{Z_1(s)}$
Gain	R_1	R_2	$-\dfrac{R_2}{R_1}$
Integration	R	C	$-\dfrac{1}{RC}\dfrac{1}{s}$
Differentiation	C	R	$-RCs$
PI controller	R_1	$R_2 \quad C$	$-\dfrac{R_2}{R_1}\dfrac{\left(s+\frac{1}{R_2 C}\right)}{s}$
PD controller	C / R_1	R_2	$-R_2 C\left(s+\dfrac{1}{R_1 C}\right)$
PID controller	C_1 / R_1	$R_2 \quad C_2$	$-\left[\left(\dfrac{R_2}{R_1}+\dfrac{C_1}{C_2}\right)+R_2 C_1 s+\dfrac{\frac{1}{R_1 C_2}}{s}\right]$

4. Digital Implementation of Controllers

The basic components of a digital controller are shown in Figure 9:

(1) A CPU for implementing the control algorithms.
(2) Discrete state input and output devices (e.g., witches and lamps)
(3) A/D converter to convert the sensor signals from analog to digital signals.
(4) D/A converter to convert the control decisions made by the control algorithm in the CPU to the analog signal form so that it can be commanded to actuation system.
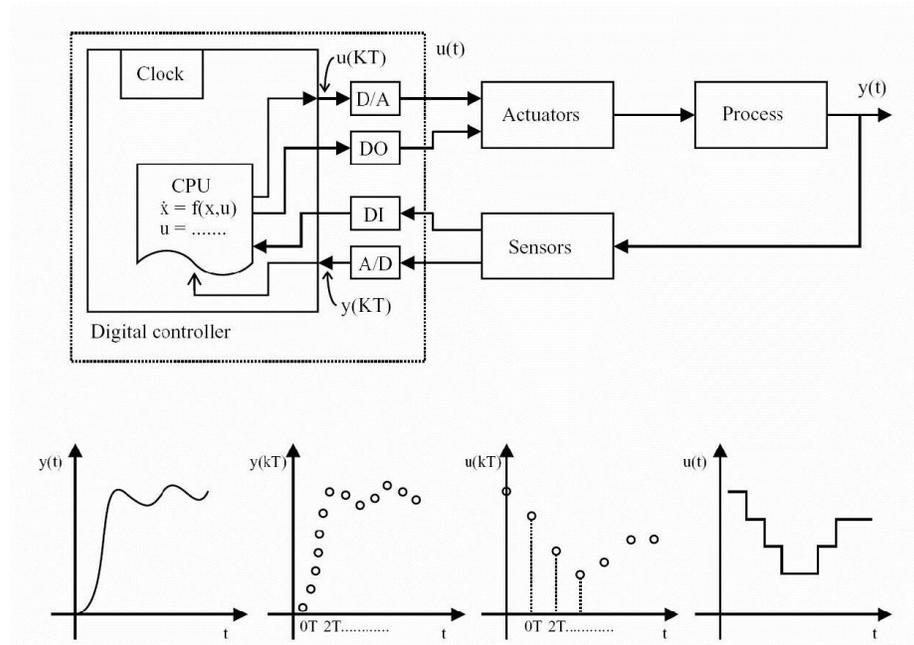

Figure 9: Digital closed-loop control system and the nature of the signal involved.

In Figure 9, it is shown that the signals travel from sensors to the control computer in analog form. Similarly, the control signals from the controller to the amplifier/actuator travel in analogy form. The conversion of signal from analogy to digital form occurs at the control computer end. Similarly, the conversion of digital signal to analogy signal occurs at the control computer end and travels to the actuators in analog form. Recent trends in the computer-controlled systems are such that the analog to digital, and digital to analog conversion occurs at the sensor and actuator point. Such sensors and actuators are marketed as "*smart sensors*" and "*smart actuators*". In this approach, the signal travels from sensor point to control computer, and from control computer to the actuator point in digital form. Especially, the use of fiber-optic transmission medium provides very high signal transmission speed with high noise immunity. It also simplifies the interface problems between the computers, sensors, and amplifiers. In either case, digital input and output (DI/DO), A/D, and D/A operations are needed in a computer control system as interface between the digital world of computers and the analogy world of real systems.

The exact location of the digital and analog interface functions may vary from application to application.

Digital control systems can be modeled adequately by the discrete equivalent to the differential equation, namely the **difference equation**. For example, the general second-order difference equation

$$a_2 y(kT) + a_1 y(kT - T) + a_0 y(kT - 2T) = b_2 x(kT) + b_1 x(kT - T) + b_0 x(kT - 2T)$$

where y is the system output and x is the system input.

In analog or continuous control systems, we used Laplace transforms in our analysis. In a similar manner, in digital control systems we need to use a new transformation in order to simplify our analysis, which is called the **Z-transform**. The z-transform is defined by

$$z\{f(KT)\} = F(z) = \sum_{k=0}^{\infty} f(kT)z^{-k}$$

Based on the above definition, we have the equation, called real translation theorem.

$$z\{f(KT - nT)\} = z^{-n}F(z)$$

Or

$$z^{-1}\{z^{-n}F(z)\} = f(KT - nT)$$

The Tustin transformation is used to transform the continuous compensator, $G_c(s)$, to the digital compensator, $G_c(z)$. The Tustin transformation is given by

$$s = \frac{2(z-1)}{T(z+1)}$$

and its inverse by

$$z = \frac{1 + \frac{T}{2}s}{1 - \frac{T}{2}s}$$

As the sampling interval, T, gets smaller (high sampling rate), the digital compensator's output yields a closer match to the analog compensator. If the sampling rate is not high enough, there is a discrepancy at higher frequencies between the digital and analog frequency responses.

Example 4

Find the difference equations for the digital implementation of a controller with a transfer function of $G_c(s) = 1.4\dfrac{s+6}{s}$, given that the sampling period is 0.07 s and 0.035 s, respectively.