Chapter Five

من 1 الى 40

5.1.1. Second – Order Systems

There are many applications of control systems work as a second order system such as servo system which is the main case study of our work.

The servo system shown in Fig.(5.1 (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.

By using the transfer function , Fig.(5.1 (a)) can be redrawn as in Fig.(5.1 (b)), which can be modified to that shown in Fig.(5.1 (c)).

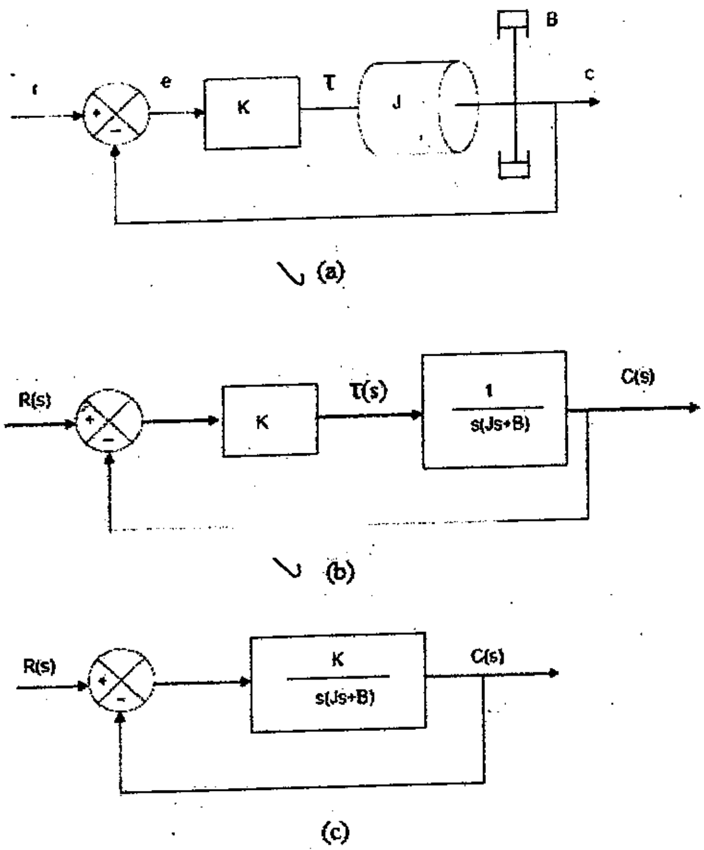


Fig.(5.1) (a) Servo system; (b) block diagram; (c) simplified block diagram.

The closed – loop transfer function is then obtained as:

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + Bs + K} = \frac{K/J}{s^2 + (B/J)\,s + (K/J)} \quad \ldots\ldots (5.1)$$

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).

5.1.1.1. Step Response of Second – Order System

The closed – Loop transfer function of the system shown in Fig.(5.1.(c)) is

$$\frac{C(s)}{R(s)} = \frac{K}{Js^2 + Bs + K} \quad \ldots\ldots (5.2)$$

Which can be rewritten as:

$$\frac{C(s)}{R(s)} = \frac{\frac{K}{J}}{\left[s + \frac{B}{2J} + \sqrt{\left(\frac{B}{2J}\right)^2 - \frac{k}{J}}\right]\left[s + \frac{B}{2J} - \sqrt{\left(\frac{B}{2J}\right)^2 - \frac{k}{J}}\right]} \quad \ldots\ldots (5.3)$$

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:

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

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

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

In terms of (ζ) and (ω_n), the system shown in Fig.(5.1(c)) can be modified to that shown in Fig.(5.2), and the closed – loop transfer function C(s) / R(s) given by equation (5.2) can be written:

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n\, s + \omega_n^2} \quad \ldots\ldots (5.4)$$

This form is called the standard form of the second – order system.

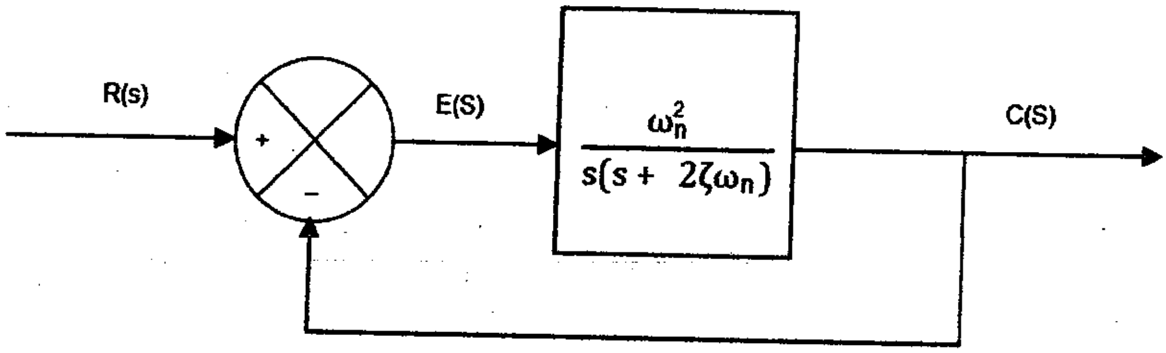


Fig.(5.2) Second–order system.

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 Fig.(5.2) 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 :

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{(s+\zeta\omega_n + j\omega_d)(s+\zeta\omega_n - j\omega_d)} \qquad \text{....... (5.5)}$$

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 :

$$C(s) = \frac{\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)\, s} \quad \text{.......(5.6)}$$

The inverse Laplace transform of equation (5.6) can be obtained easily if C(s) is written in the following form:

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

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

In appendix it was shown that

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

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

Hence the inverse Laplace transform of equation (5.6) is obtained as:

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

$$= 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin\left(\omega_d t + \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right), \quad \text{for } t \geq 0 \quad \text{....... (5.7)}$$

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, $R(s) = \frac{1}{s}$ and C(s) can be written:

$$C(s) = \frac{\omega_n^2}{(s + \omega_n)^2\, s} \quad \text{...... (5.8)}$$

   The inverse Laplace transform of equation (5.8) may be found as:

$$C(t) = 1 - e^{-\omega_n t}(1 + \omega_n t), \quad \text{for } t \geq 0 \quad \text{..... (5.9)}$$

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, $R(s) = \frac{1}{s}$ and C(s) can be written :

$$C(s) = \frac{\omega_n^2}{\left(s + \zeta\omega_n + \omega_n\sqrt{\zeta^2-1}\right)\left(s + \zeta\omega_n - \omega_n\sqrt{\zeta^2-1}\right)s} \quad \text{.... (5.10)}$$

The inverse Laplace transform of equation (5.10) is :

$$C(t)=1+\frac{1}{2\sqrt{\zeta^2-1}\,(\zeta+\sqrt{\zeta^2-1})}e^{-(\zeta+\sqrt{\zeta^2-1})\omega_n t}$$

$$-\frac{1}{2\sqrt{\zeta^2-1}\,(\zeta-\sqrt{\zeta^2-1})}e^{-(\zeta-\sqrt{\zeta^2-1})\omega_n t}$$

$$=1+\frac{\omega_n}{2\sqrt{\zeta^2-1}}\left(\frac{e^{-s_1 t}}{s_1}-\frac{e^{-s_2 t}}{s_2}\right), \quad \text{for } t\geq 0 \qquad \ldots\ldots (5.11)$$

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.

5.1.1.2 Definition 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 the 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 and as these specifications are shown graphically in Fig.(5.3).

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 over damped 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

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

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.

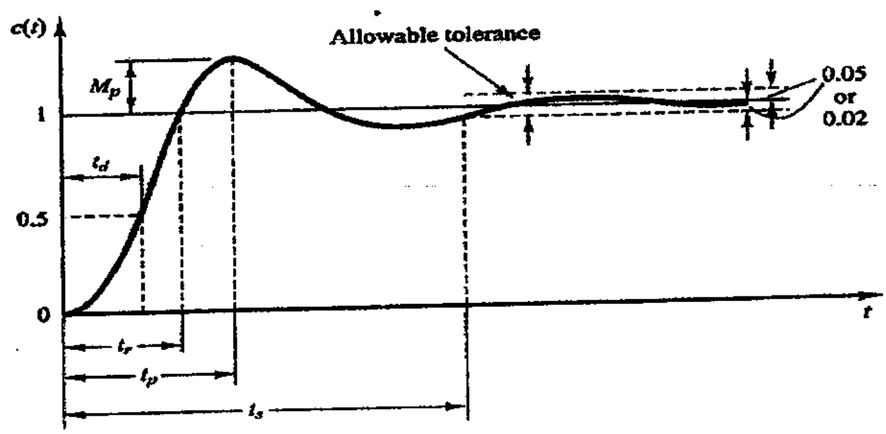


Fig.(5.3) Unit – step response curve showing t_d ,t_r ,t_p , M_p ,and t_s .

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.

5.2.1.2 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 rise time cannot be made smaller simultaneously. If one of them is made smaller, the other necessarily becomes larger.

5.2.1.2 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 (5.4). these values will be obtained in terms of (ζ) and (ω_n). The system is assumed to be underdamped.

(1) Rise time (t_r) : Referring to equation (5.7) , we obtain the rise time (t_r) by letting $c(t_r) = 1$.

$$C(t_r) = 1 = 1 - e^{-\zeta\omega_n t_r}\left(\cos\omega_d t_r + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin \omega_d t_r\right) \qquad \text{....... (5.12)}$$

Since $e^{-\zeta\omega_n t_r} \neq 0$, we obtain from equation (5.12) 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:

$$t_r = \frac{1}{\omega_d} \tan^{-1}\left(\frac{\omega_d}{-\sigma}\right) = \frac{\Pi - \beta}{\omega_d} \qquad \text{..... (5.13)}$$

Where β is defined in Fig.(5.4). Clearly, for a small value of t_r, ω_d must be large.

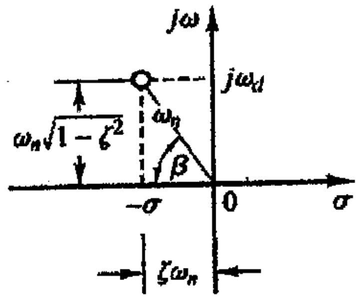


Fig.(5.4) Definition of the angle β .

(2) Peak time (t_p): Referring to equation (5.7), we may obtain the peak time by differentiating C (t) with respect to time and letting this derivative equal zero. Since

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

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

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

This last equation yields following the equation:

$$\sin \omega_d t_p = 0$$

or

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

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

$$t_p = \frac{\pi}{\omega_d} \qquad \ldots\ldots(5.14)$$

The peak time (t_p) corresponds to one – half cycle of the frequency of damped oscillation.

(3) 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 (5.7) as:

$$M_p = c(t_p) - 1$$

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

$$= e^{-(\sigma/\omega_d)\pi} = e^{-(\zeta/\sqrt{1-\zeta^2})\pi} \qquad \ldots\ldots(5.15)$$

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

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

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

(4) Settling time (t_s) : For an underdamped second – order system, it is convenient for comparing the responses of systems that the settling time (t_s) is defined as follows:

$$t_s = \frac{4}{\sigma} = \frac{4}{\zeta\omega_n} \qquad \text{(2\% criterion)} \qquad \ldots\ldots(5.16)$$

or

$$t_s = \frac{3}{\sigma} = \frac{3}{\zeta\omega_n} \qquad \text{(5\% criterion)} \qquad \text{.......(5.17)}$$

5.2.1 Higher – Order Systems

In this section, a transient response analysis of higher – order systems will be presented 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. Consider the system shown in Fig.(5.5). The closed – loop transfer function is:

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)\,H(s)} \qquad \text{....... (5.18)}$$

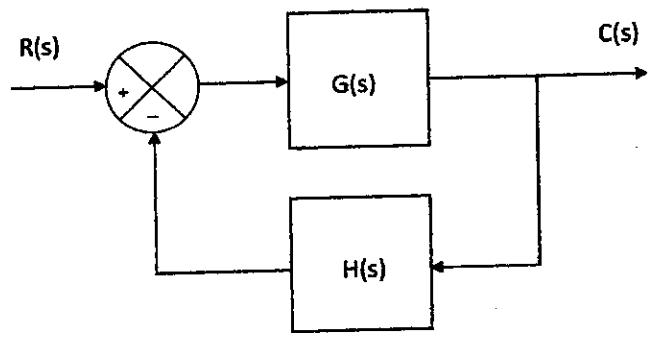


Fig.(5.5) Control system.

In general, G(s) and H(s) are given as ratios of polynomials in (s) , or

$$G(s) = \frac{P(s)}{q(s)} \qquad \text{and} \qquad H(s) = \frac{n(s)}{d(s)}$$

Where p (s), q (s) , n (s) , and d (s) are polynomials in (s). The closed– loop transfer function given by equation (5.18) may then be written:

$$\frac{C(s)}{R(s)} = \frac{P(s)\,d(s)}{q(s)\,d(s) + p(s)\,n(s)}$$

$$= \frac{b0s^{m} + b1s^{m-1} + \dots\dots + b_{m-1}\,s + bm}{a_0s^{n} + a_1s^{n-1} + \dots\dots\dots + a_{n-1}\,s + a_n} \qquad (m \le n)$$

The transient response of this system to any given input can be obtained by a computer simulation. 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:

$$\frac{C(s)}{R(s)}=\frac{k(s+z_1)(s+z_2)\ldots\ldots(s+z_m)}{(s+p_1)(s+p_2)\ldots\ldots(s+p_n)} \qquad \ldots\ldots (5.19)$$

Lets 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 (5.19) can be written :

$$C(s)=\frac{a}{s}+\sum_{i=1}^{n}\frac{a_i}{s+p_i} \qquad \ldots\ldots (5.20)$$

Where (a_i) is the residue of the pole at s = - p_i. (If the system involves multiple poles, the c (s) will have multiple – pole terms.)[The partial – fraction expansion of c(s) , as given by equation (), can .20 obtained easily with MATLAB. Use the residue command.]

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 (5.20) can be rewritten:

$$C(s)=\frac{a}{s}+\sum_{j=1}^{q}\frac{a_j}{s+p_j}+\sum_{k=1}^{r}\frac{b_k(s+\zeta_k\omega_k)+C_k\omega_k\sqrt{1-\zeta_k^2}}{s^2+2\zeta_k\omega_k s+\omega_k^2} \qquad (q+2r=n)$$

Where we assumed all closed – loop poles are distinct. [If the closed –loop poles involve multiple poles, C(s) must have multiple - pole terms.] Form 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:

$$C(t)=a+\sum_{j=1}^{q} a_j e^{-p_j t}+\sum_{k=1}^{r} b_k e^{-\zeta_k\omega_k t} \cos\omega_k\sqrt{1-\zeta_k^2}\,t$$

$$+\sum_{k=1}^{r} c_k e^{-\zeta_k\omega_k t} \sin\omega_k\sqrt{1-\zeta_k^2}\,t \qquad \text{for } t \geq 0 \text{ .... (5.21)}$$

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 (5.21) will approach zero as time (t) increases. The steady – state output is then $C(\infty) = a$. 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. The practical procedure for plotting time response curves of systems higher than second – order is through computer simulation.

5.3 PID Controller

PID (Proportional – Integral - Derivative) control is one of the earlier control strategies [160]. It has a simple control structure and it is easy to tune in real word. Therefore, it has a wide range of applications in industrial control. According to a survey for process control systems in 1989, more than 90% of the control loops were of the PID type [160]. PID control has been an active topic since decades ago. It is interesting to note that more than half of the industrial controllers in use today utilize PID or modified PID control schemes.

The usefulness of PID controls lies their general applicability to most control systems. In particular, when the mathematical model of the plant is not known and therefore analytical design methods cannot be used, PID controls prove to be most useful. In the field of process control systems, it is well known that the basic and modified PID control schemes have proved their usefulness providing satisfactory control, although in many given situations they may not provide optimal control [2].

A typical structure of PID control system is shown in Fig.(5.6), where it can be seen that in a PID controller, the error signal e(t) is used to generate the proportional, integral and derivative actions with the resulted signals weighted and summed to form the control signal u(t) applied to the plant model.

A mathematical description of the PID controller is:

$$u(t) = k_p \left[e(t) + \frac{1}{T_i}\int_0^t e(t)dt + T_d \frac{d\,e(t)}{dt}\right] \quad \text{......(5.22)}$$

where u(t) is the input signal, the error signal e(t) is defined as $e(t) = r(t) - y(t)$, and r(t) is the reference input signal.

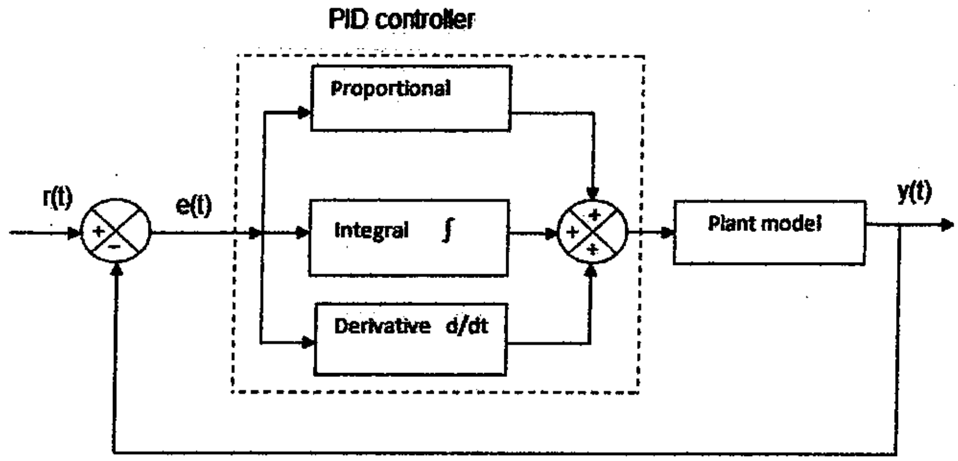


Fig.(5.6) A typical PID control structure.

5.3.1 Tuning Rules For PID Controllers

5.3.1.1 PID control of plants

Fig. (5.7) shows a PID control of a plant. If a mathematical model of the plant can be derived, then it is possible to apply various design techniques for determining parameters of the controller that will meet the transient and steady – state specifications of the closed – loop system. However, if the plant is so complicated that its mathematical model cannot be easily obtained, then an analytical approach to the design of a PID controller is not possible. Then we must resort to experimental approaches to the tuning of PID controllers.

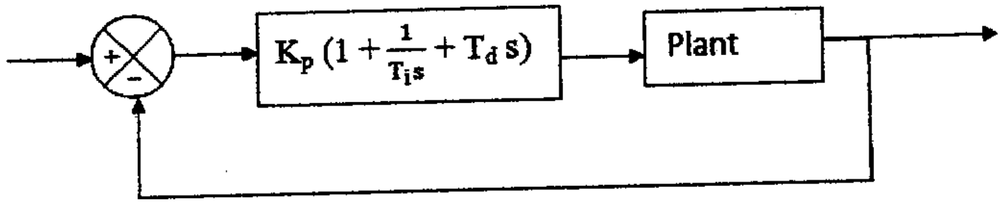


Fig.(5.7) PID control of a plant.

The process of selecting the controller parameters to meet given performance specifications is known as controller tuning. Ziegler and Nichols suggested rules for tuning PID controllers (meaning to set values K_p , T_i , and T_d) based on experimental step responses or based on the value of K_p that results in marginal stability when only proportional control action is used. Ziegler – Nichols rules, which are briefly presented in the following, are useful when mathematical models of plants are not known. (These rules can, of course, be applied to the design of systems with known mathematical models.)Such rules suggest a set of values of K_p , T_i , and T_d that will give a stable operation of the system. However, the resulting system may exhibit a large maximum overshoot in the step response, which is unacceptable. In such a case we need series of fine tuning until an acceptable result is obtained. In fact, the Ziegler – Nichols tuning rules give an educated guess for the parameter values and provide a starting point for fine tuning, rather than giving the final settings for K_p, T_i , and T_d in a single shot.

5.3.1.2 Zeigler – Nichols rules for tuning PID controllers

Zeigler and Nichols proposed rules for determining values of the proportional gain K_p , integral time T_i, and derivative time T_d based on the transient response characteristics of a given plant. Such determination of the parameters of PID controllers or tuning of PID controllers can be made by engineers on – site by experiments on the plant.

There are two methods called Ziegler – Nichols tuning rules: the first method and the second method.

1.First method : In the first method , we obtain experimentally the response of the plant to a unit – step input , as shown in Fig.(5.8).

If the plant involves neither integrator(s) nor dominant complex–conjugate poles, then such a unit – step response curve may look S–shaped, as shown in Fig.(5.9). This method applies if the response to a step input exhibits an S–shaped curve. Such step – response curves may be generated experimentally or from a dynamic simulation of the plant.

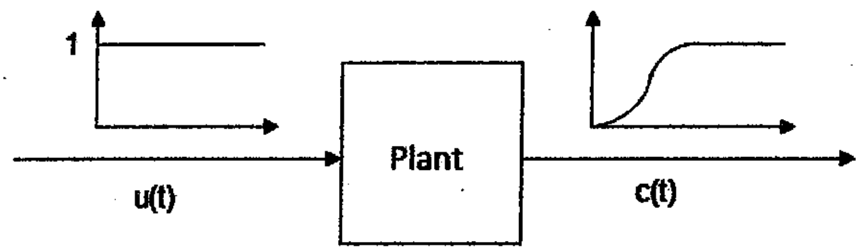


Fig.(5.8) Unit–step response of a plant.

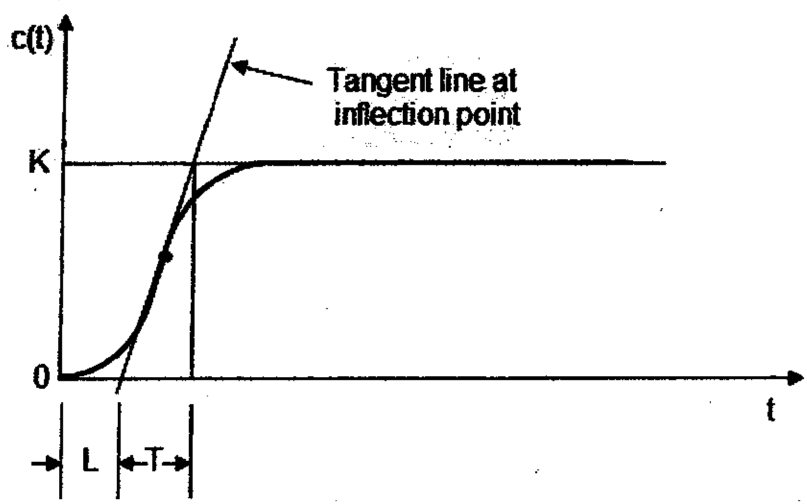


Fig.(5.9) S–shaped response curve.

The S–shaped curve may be characterized by two constants, delay time (L) and time constant (T). The delay time and time constant are determined by drawing a tangent line at the inflection point of the S–shaped curve and determining the intersections of the tangent line with the time axis and line $c(t) = K$, as shown in

Fig.(5.9). The transfer function C(s) / U(s) may then be approximated by a first – order system with a transport lag as follows:

$$\frac{C(s)}{U(s)} = \frac{ke^{-LS}}{Ts+1}$$

Ziegler and Nichols suggested to set the values of K_p , T_i , and T_d according to the formula shown in Table 5.1.

Table 5.1 Zeigler – Nichols Tuning Rule Based on Step Response of Plant (First Method)

Type of Controller	K_p	T_i	T_d
P	$\frac{T}{L}$	∞	0
PI	$0.9\frac{T}{L}$	$\frac{L}{0.3}$	0
PID	$1.2\frac{T}{L}$	2L	0.5L

Notice that the PID controller tuned by the first method of Ziegler – Nichols rules gives:

$$G(s) = K_p\,(1 + \frac{1}{T_i s} + T_d\, s)$$

$$= 1.2\frac{T}{L}\,(1 + \frac{1}{2LS} + 0.5\ LS)$$

$$= 0.6\ T\,\frac{(S+\frac{1}{L})^2}{S}$$

Thus, the PID controller has a pole at the origin and double zeros at s = -1/L.

2. **Second Method:** In the second method, we first set $T_i = \infty$ and $T_d = 0$. Using the proportional control action only (as shown in Fig.(5.10)), increase K_p from (0) to a critical value K_{cr} at which the output first exhibits sustained oscillations. (If the output does not exhibit sustained oscillations for whatever value k_p may take, then

this method does not apply.) Thus, the critical gain K_{cr} and the corresponding period P_{cr} are experimentally determined (see Fig.(5.11)).

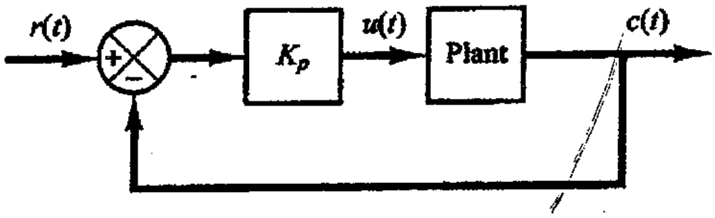


Fig.(5.10) Closed–loop system with a proportional controller.

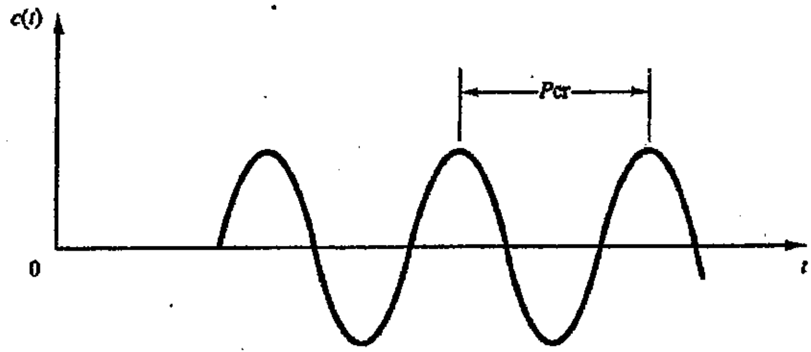


Fig.(5.11) Sustained oscillation with period P_{cr} .

Ziegler and Nichols suggested that we set the values of the parameters K_p , T_i , and T_d according to the formula shown in Table 5.2.

Table 5.2. Zeigler – Nichols Tuning Rule Based on Critical Gain K_{cr} and Critical Period P_{cr} (Second Method).

Type of Controller	K_p	T_i	T_d
P	$0.5\ K_{cr}$	∞	0
PI	$0.45\ K_{cr}$	$\frac{1}{1.2}\ P_{cr}$	0
PID	$0.6\ K_{cr}$	$0.5\ P_{cr}$	$0.125\ P_{cr}$

Notice that the PID controller tuned by the second method of Ziegler – Nichols rules gives:

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

$$= 0.6\ K_{cr}\left(1 + \frac{1}{0.5\ P_{cr} s} + 0.125\ P_{cr} s\right)$$

$$= 0.075\ K_{cr}\ P_{cr}\ \frac{\left(S + \frac{4}{P_{cr}}\right)^2}{S}$$

Thus, the PID controller has a pole at the origin and double zeros at $s = -4/P_{cr}$.

Note that it the system has a known mathematical model (such as the transfer function), then we can use the root – locus method to find the critical gain K_{cr} and the frequency of the sustained oscillations ω_{cr}, where $2\pi / \omega_{cr} = P_{cr}$. These values can be found from the crossing points of the root –locus branches with the $j\omega$ axis. (Obviously, if the root – locus branches do not cross the jw axis, this method does not apply.)

Ziegler – Nichols tuning rules (and other tuning rules presented in the literature) have been widely used to tune PID controllers in process control systems where the plant dynamics are not precisely known. Over many years, such tuning rules proved to be very useful. Zeigler – Nichols tuning rules can, of course, be applied to plants whose dynamics are known. (If the plant dynamics are known, many analytical and graphical approaches to the design of PID controllers are available, in addition to Ziegler – Nichols tuning rules.)

Chapter Five

5.1 Modifications of PID Control Schemes

Consider the basic PID control system shown in Fig.(5.12(a)), where the system is subjected to disturbances and noises. Fig.(5.12(b)) is a modified block diagram of the same system.

In the basic PID control system such as the one shown in Fig.(5.12(b)), if the reference input is a step function, then, because of the presence of the derivative term in the control action, the manipulated variable u(t) will involve an impulse function (delta function). In an actual PID controller, instead of the pure derivative term $T_d s$, we employ :

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

Where the value of (γ) is somewhere around 0.1. Therefore, when the reference input is a step function, the manipulated variable u(t) will not involve an impulse function, but will involve a sharp pulse function. Such a phenomenon is called set – point kick.

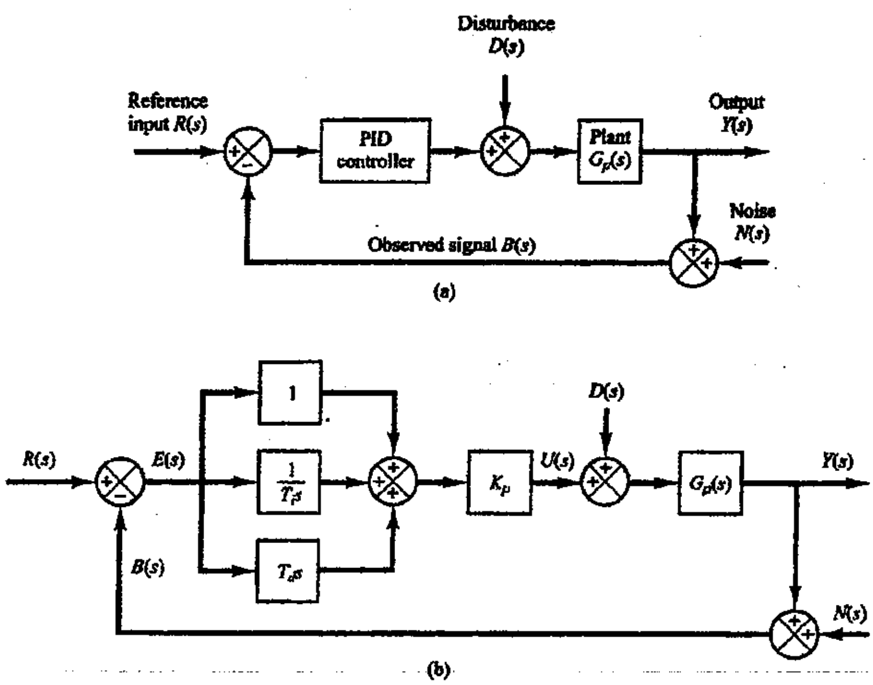


Fig.(5.12) (a) PID controlled system; (b) Equivalent block diagram.

5.1.1. PI – D Control

To avoid the set – point kick phenomenon, we may wish to operate the derivative action only in the feedback path so that differentiation occurs only on the feedback signal and not on the reference signal. The control scheme arranged in this way is called the PI – D control. Fig.(5.13) shows a PI – D controlled system.

From Fig.(5.13), it can be seen that the manipulated signal U(s) is given by:

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

Notice that in the absence of the disturbances and noises, the closed – loop transfer function of the basic PID control system [shown in Fig.(5.12 (b))] and the PI–D control system (shown in Fig.(5.13)) are given, respectively, by:

$$\frac{Y(s)}{R(s)} = \left(1 + \frac{1}{T_i s} + T_d s\right) \frac{K_p\, G_p(s)}{1 + \left(1 + \frac{1}{T_i s} + T_d s\right) K_p\, G_p(s)}$$

and

$$\frac{Y(s)}{R(s)} = \left(1 + \frac{1}{T_i s}\right) \frac{K_p\, G_p(s)}{1 + \left(1 + \frac{1}{T_i s} + T_d s\right) K_p\, G_p(s)}$$

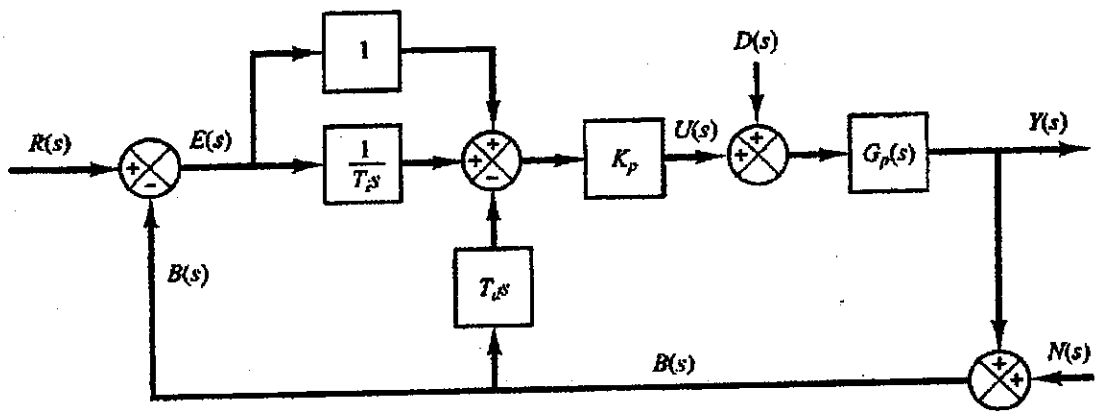


Fig.(5.13) PI–D controlled system.

It is important to point out that in the absence of the reference input and noises, the closed – loop transfer function between the disturbance D(s) and the output Y(s) in either case is the same and is given by:

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

5.1.1. I – PD Control

Consider again the case where the reference input is a step function. Both PID control and PI – D control involve a step function in the manipulated signal. Such a step change in the manipulated signal may not be desirable in many occasions. Therefore, it may be advantageous to move the proportional action and derivative action to the feedback path so that these actions affect the feedback signal only. Fig.(5.14) shows such a control scheme. It is called the I – PD control. The manipulated signal is given by:

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

Notice that the reference input R(s) appears only in the integral control part. Thus, in I – PD control, it is imperative to have the integral control action for proper operation of the control system.

The closed–loop transfer function Y(s)/R(s) in the absence of the disturbance input and noise input is given by :

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

It is noted that in the absence of the reference input and noise signals, the closed–loop transfer function between the disturbance input and the output is given by:

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

This expression is the same as that for PID control or PI–D control.

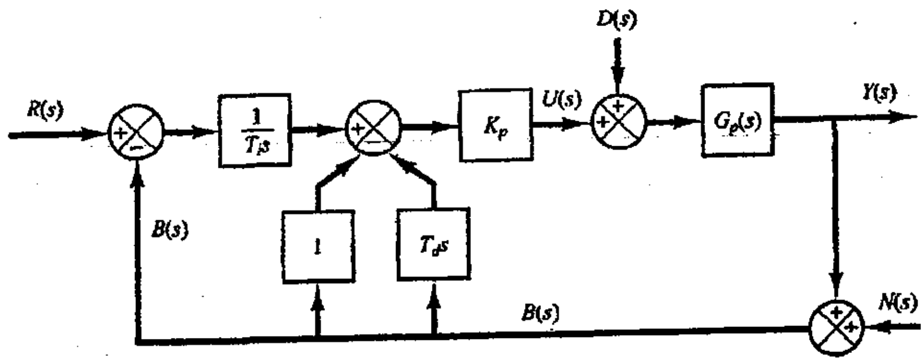


Fig.(5.14) I–PD controlled system.

5.1.3. Dynamic Characteristics of the PI Controller, PD Controller, and PID Controller

The PI controller is characterized by the transfer function

$$G_c(s) = K_p(1 + \frac{1}{T_i s})$$

The PI controller is a lag compensator. It possesses a zero at $s = -1/T_i$ and a pole at $S = 0$. Thus, the characteristic of the PI controller is infinite gain at zero frequency. This improves the steady–state characteristics. However, inclusion of the PI control action in the system increases the type number of the compensated system by (1), and this causes the compensated system to be less stable or even makes the system unstable. Therefore, the values of K_p and T_i must be chosen carefully to ensure a proper transient response. By properly designing the PI controller, it is possible to make the transient response to a step input exhibit relatively small or no overshoot. The speed of response, however, becomes much slower. This is because the PI controller, being a low –pass filter, attenuates the high – frequency components of the signal. Based on the preceding discussions, we can summarize the advantages and disadvantages of a properly designed PI controller as [161].

1. Improving damping and reducing maximum overshoot.
2. Increasing rise time.
3. Improving gain margin and phase margin.
4. Filtering out high–frequency noise.

5. May make the system unstable.
6. Integral control of the system can eliminate the steady–state error in the response to the step input [2].

The PD controller is a simplified version of the lead compensator. The PD controller has the transfer function $G_c(s)$, where:

$$G_c(s) = K_p (1 + T_d s)$$

The value of K_p is usually determined to satisfy the steady – state requirement. The corner frequency $1/T_d$ is chosen such that the phase lead occurs in the neighborhood of the gain crossover frequency. Although the phase margin can be increased, the magnitude of the compensator continues to increase for the frequency region $1/T_d < \omega$. (Thus, the PD controller is a high – pass filter.) Such a continued increase of the magnitude is undesirable, since it amplifies high – frequency noises that may be present in the system. Lead compensation can provide a sufficient phase lead, while the increase of the magnitude for the high–frequency region is very much smaller than that for PD control. Therefore, lead compensation is preferred over PD control. The PD control, as in the case of the lead compensator, improves the transient – response characteristics, improves system stability, and increases the system band width, which implies fast rise time. PD controller can affect the performance of a control system in the following ways [161].

1. Improving damping and reducing maximum overshoot.
2. Reducing rise time and settling time.
3. Possibly accentuating noise at higher frequencies.
4. Derivative control action, when added to a proportional controller, provides a means of obtaining a controller with high sensivity [2].
5. Derivative control is essentially anticipatory, measure the instantaneous error velocity, and predicts the large overshoot a head of time and produces an appropriate counteraction before too large an overshoot occurs [2].

The PID controller is a combination of the PI and PD controllers. It is a lag – lead compensator. Note that the PI control action and PD control action occur in different frequency regions. The PI control action occurs at the low – frequency region and PD control action occurs at the high – frequency region. The PID control may be used when the system requires improvements in both transient and steady–state performances. Notice that the PID control, when designed properly, captures the advantages of both the PD and the PI controls [161].

The PID controller may be written in the following form:

$$G_c(s) = K_P + \frac{K_I}{s} + K_D\, s$$

Where,

K_P = Proportional gain .

K_I = Integral gain .

K_D= Derivative gain .

The effect of K_P , K_I , and K_D can be summarized as in the table 5.3.

Table 5.3 Effect of increasing the PID gains K_P , K_I ,and K_D on the step response.

PID Gain	Percent overshoot	Settling time	Rise time	Steady–state error
K_P	Increases	Minimal impact (Small change)	Decreases	Decreases
K_I	Increases	Increases	Decreases	Eliminate (Zero steady – state error)
K_D	Decreases	Decreases	Small change	Small change (No impact)

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–[illegible], 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.)

هذه المحاضرات ملخصة من ogata

تكملة فصل (5)

عماد عبد السلام حيوت
دكتوراه هندسة ميكانيك

* Routh's stability Criterion

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

where, $m \leqslant n$.

The procedure in Routh's stability criterion is as follows:

(1) write the polynomial in (S) in the following form:

$$a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n = 0 \quad --- (*)$$

where the ~~coeffeff~~ coefficients are real quantities. and $(a_n \neq 0)$. [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.

(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				

للأمانة العلمية
هذه المحاضرات ملخصة من ogata
وليست مؤلفة، د. عماد عبد السلام

$$b_1 = \frac{a_1 a_2 - a_0 a_3}{a_1}$$

$$b_2 = \frac{a_1 a_4 - a_0 a_5}{a_1}$$

$$b_3 = \frac{a_1 a_6 - a_0 a_7}{a_1}$$

$$\vdots$$

$$c_1 = \frac{b_1 a_3 - a_1 b_2}{b_1}$$

$$c_2 = \frac{b_1 a_5 - a_1 b_3}{b_1}$$

$$c_3 = \frac{b_1 a_7 - a_1 b_4}{b_1}$$

$$\vdots$$

$$d_1 = \frac{c_1 b_2 - b_1 c_2}{c_1}$$

$$d_2 = \frac{c_1 b_3 - b_1 c_3}{c_1}$$

$$\vdots$$

* Routh's stability criterion states that the number of roots of eq(*) with positive real parts is equal to the number of changes in sign of the coefficients of the first column of the array.

Example: let us apply Routh's stability criterion to the following third-order polynomial:

$$a_0 s^3 + a_1 s^2 + a_2 s + a_3 = 0$$

Solution:

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		

The condition that all roots have negative real parts is given by:

$$a_1 a_2 > a_0 a_3$$

example: Consider the following polynomial:

$$s^4 + 2s^3 + 3s^2 + 4s + 5 = 0$$

s^4	1	3	5
s^3	2	4	0
s^2	1	5	
s^1	-6		
s^0	5		

s^4	1	3	5
s^3	1	2	0 → بالقسم على 2
s^2	1	5	
s^1	-3		
s^0	5		

عدد الجذور التي تحت المتقوس غير مستقره = 2
وهو يعدد تغيرات إشارة من موجب
إلى سالب ومن سالب
إلى موجب

Special cases — حالات خاصة

Example: $s^3 + 2s^2 + s + 2 = 0$

s^3	1	1
s^2	2	2
s^1	$0 \approx \epsilon$	
s^0	2	

ملاحظات:

1. إذا كانت الإشارة قبل وبعد (ϵ) متشابهة فإن المعادلة تحتوي على زوج من الجذور الخيالية.
2. إذا كانت الإشارة قبل وبعد (ϵ) مختلفة فإن المعادلة تحتوي على تغير واحد بالإشارة.

Example: $s^3 - 3s + 2 = (s-1)^2 (s+2) = 0$

s^3	1	-3
s^2	$0 \approx \epsilon$	2
s^1	$-3 - \frac{2}{\epsilon}$	
s^0	2	

One sign change ($s^2 \to s^1$)

One sign change ($s^1 \to s^0$)

There are two sign changes.

example:

$$s^5 + 2s^4 + 24s^3 + 48s^2 - 25s - 50 = 0$$

solution:

s^5	1	24	−25
s^4	2	48	−50 ← Auxiliary polynomial $p(s)$
s^3	0	0	

$$p(s) = 2s^4 + 48s^2 - 50$$

* هناك اما جذور حقيقية متعاكسة بالإشارة .
او جذور خيالية متعاكسة (conjugate)

$$\frac{dp(s)}{ds} = 8s^3 + 96s$$

s^5	1	24	−25
s^4	2	48	−50
s^3	8	96 ← coefficients of $dp(s)/ds$	
s^2	24	−5	
s^1	112.7	0	
s^0	−50		

$$2s^4 + 48s^2 - 50 = 0$$

$$s^2 = 1 \quad , \quad s^2 = -25$$

$$s = \pm 1 \quad , \quad s = \pm j5$$

$$(s+1)(s-1)(s+j5)(s-j5)(s+2) = 0$$

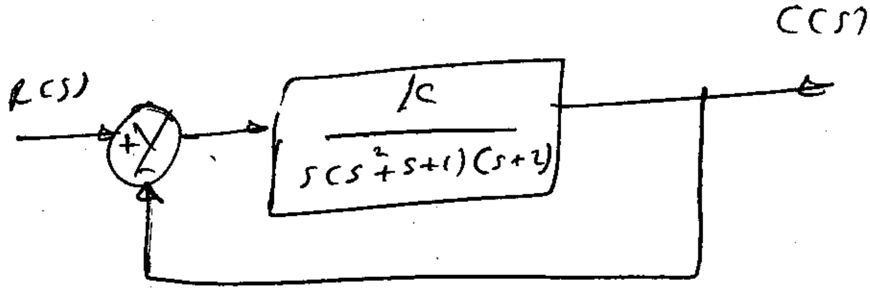
One sign change (one root with a positive real part).

Example:

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

$$s^4 + 3s^3 + 3s^2 + 2s + K = 0$$

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		



$$2 - \frac{9}{7}K > 0$$

$$2 > \frac{9}{7}K$$

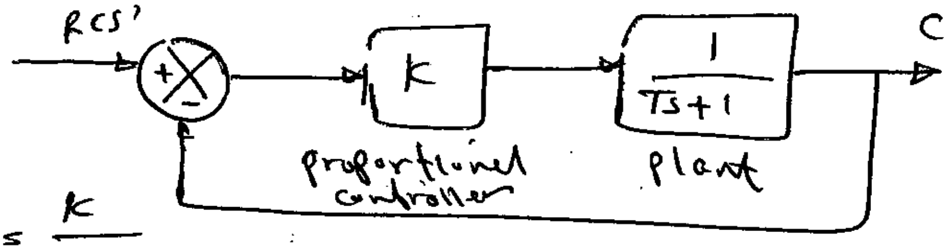
$$\boxed{\frac{14}{9} > K > 0}$$ the range of stability

* Effect of Integral and Derevative control Actions on System Performance.

* Integral Control Action

The 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.

* proportional control of systems.



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

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

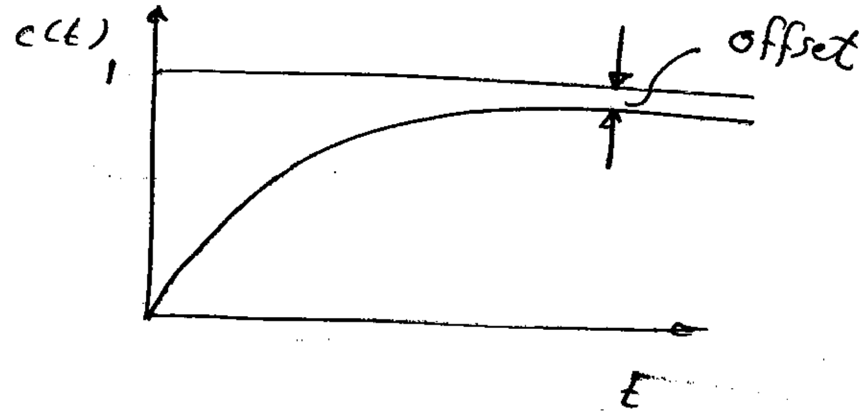
$$E(s) = \frac{1}{1+G(s)} R(s) = \frac{1}{1+\frac{K}{Ts+1}} R(s)$$

unit-step $\left(R(s) = \frac{1}{s}\right)$

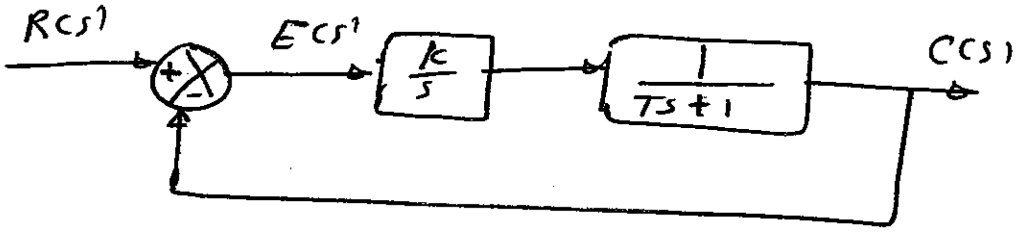
$$E(s) = \frac{Ts+1}{Ts+1+K} \frac{1}{s}$$

The steady-state error is :

$$e_{ss} = \lim_{t \to \infty} e(t) = \lim_{s \to 0} s E(s) = \lim \frac{Ts+1}{Ts+1+K} = \frac{1}{K+1}$$



* Integral control of systems.



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

$$\frac{E(s)}{R(s)} = \frac{R(s)-C(s)}{R(s)} = \frac{s(Ts+1)}{s(Ts+1)+K}$$

$$e_{ss} = \lim_{s\to 0} sE(s)$$

$$= \lim_{s\to 0} \frac{s^2(Ts+1)}{Ts^2+s+K}\,\frac{1}{s}$$

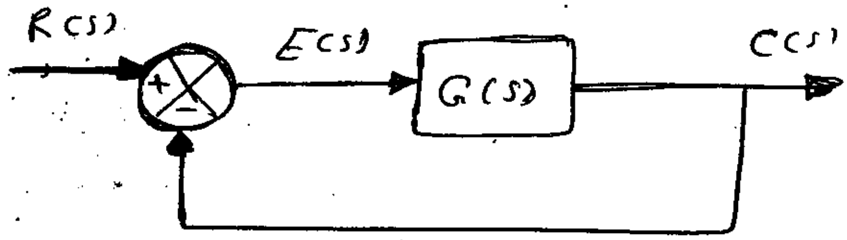
$$= 0$$

* Derevative Control Action.

• Derevative 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. Derevative control thus anticipates the actuating error, initiates an early corrective action, and tends to increas the Stability of the system.

* 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. (may be as PD or PID).

* Steady-State Errors in Unity-Feedback Control Systems.



$$G(s) = \frac{K(T_a s + 1)(T_b s+1) \cdots (T_m s + 1)}{s^N (T_1 s+1)(T_2 s+1) \cdots (T_p s+1)}$$

N : type of the system.

* steady-state errors.

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

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

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

$$e_{ss} = \lim_{t \to \infty} e(t) = \lim_{s \to 0} s E(s) = \lim_{s \to 0} \frac{s R(s)}{1+G(s)}$$

* Static position Error Constant K_p

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

$$= \frac{1}{1+G(0)}$$

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

$$e_{ss} = \frac{1}{1+K_p}$$

For a type 0 system,

$$K_p = \lim_{s \to 0} \frac{K(T_a s+1)(T_b s+1)\cdots}{(T_1 s+1)(T_2 s+1)} = K$$

for $N=1$

$$K_p = \lim_{s \to 0} \frac{K(T_a s+1)(T_b s+1)\cdots}{s^N(T_1 s+1)(T_2 s+1)\cdots} = \infty \text{ , for } N \geq$$

$$e_{ss} = \frac{1}{1+K} \quad \text{for } N=0$$

$$e_{ss} = 0 \quad \text{for } N \geq 1$$

* Static Velocity Error Constant K_v.

$$e_{ss} = \lim_{s \to 0} \frac{s}{1+G(s)} \frac{1}{s^2} \qquad \text{(unit-ramp input)}$$

$$= \lim_{s \to 0} \frac{1}{s\,G(s)}$$

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

$$e_{ss} = \frac{1}{K_v}$$

For $N = 0$ $\qquad K_v = \lim_{s \to 0} \frac{s\,K\,(T_a s+1)(T_b s+1)\ldots}{(T_1 s+1)(T_2 s+1)\ldots} = 0$

$N = 1 \qquad K_v = K$

$N \geq 2 \qquad K_v = \infty$

$e_{ss} = \frac{1}{K_v} = \infty \qquad N = 0$

$e_{ss} = \frac{1}{K_v} = \frac{1}{K} \qquad N = 1$

$e_{ss} = \frac{1}{K_v} = 0 \qquad N \geq 2$

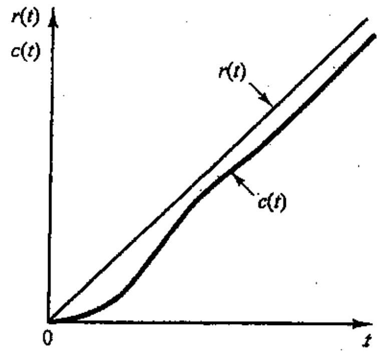


Figure 5–50
Response of a type 1 unity-feedback system to a ramp input.

For a type 2 or higher system,

$$K_v = \lim_{s\to 0} \frac{sK(T_a s + 1)(T_b s + 1)\cdots}{s^N(T_1 s + 1)(T_2 s + 1)\cdots} = \infty, \qquad \text{for } N \geq 2$$

The steady-state error e_{ss} for the unit-ramp input can be summarized as follows:

$$e_{ss} = \frac{1}{K_v} = \infty, \qquad \text{for type 0 systems}$$

$$e_{ss} = \frac{1}{K_v} = \frac{1}{K}, \qquad \text{for type 1 systems}$$

$$e_{ss} = \frac{1}{K_v} = 0, \qquad \text{for type 2 or higher systems}$$

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

$$r(t) = \frac{t^2}{2}, \qquad \text{for } t \geq 0$$
$$= 0, \qquad \text{for } t < 0$$

is given by

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

The static acceleration error constant K_a is defined by the equation

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

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,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{(T_1 s + 1)(T_2 s + 1) \cdots} = 0$$

For a type 1 system,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{s(T_1 s + 1)(T_2 s + 1) \cdots} = 0$$

For a type 2 system,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{s^2(T_1 s + 1)(T_2 s + 1) \cdots} = K$$

For a type 3 or higher system,

$$K_a = \lim_{s \to 0} \frac{s^2 K(T_a s + 1)(T_b s + 1) \cdots}{s^N(T_1 s + 1)(T_2 s + 1) \cdots} = \infty, \qquad \text{for } N \geq 3$$

Thus, the steady-state error for the unit parabolic input is

$$e_{ss} = \infty, \qquad \text{for type 0 and type 1 systems}$$

$$e_{ss} = \frac{1}{K}, \qquad \text{for type 2 systems}$$

$$e_{ss} = 0, \qquad \text{for type 3 or higher systems}$$

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.

(39)

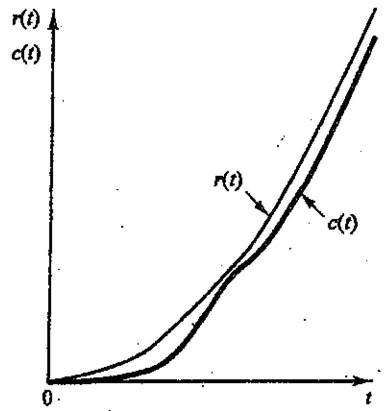


Figure 5–51
Response of a type 2 unity-feedback system to a parabolic input.

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.