C.4 Transient and Steady State Response Analysis

4.1 Introduction

Many applications of control theory are to servomechanisms which are systems using the feedback principle designed so that the output will follow the input. Hence there is a need for studying the time response of the system.

The time response of a system may be considered in two parts:
- Transient response: this part reduces to zero as $t \to \infty$
- Steady-state response: response of the system as $t \to \infty$

4.2 Response of the first order systems

Consider the output of a linear system in the form

$$Y(s) = G(s)U(s) \tag{4.1}$$

where
- $Y(s)$: Laplace transform of the output
- $G(s)$: transfer function of the system
- $U(s)$: Laplace transform of the input

Consider the first order system of the form $a\dot{y} + y = u$, its transfer function is

$$Y(s) = \frac{1}{as+1}U(s)$$

For a transient response analysis it is customary to use a reference unit step function $u(t)$ for which

$$U(s) = \frac{1}{s}$$

It then follows that

$$Y(s) = \frac{1}{(as+1)s} = \frac{1}{s} - \frac{1}{s+1/a} \tag{4.2}$$

On taking the inverse Laplace of equation (4.2), we obtain

$$y(t) = \underbrace{1}_{steady-state\ part} - \underbrace{e^{-t/a}}_{transient\ part} \qquad (t \ge 0) \tag{4.3}$$

Both of the input and the output to the system are shown in Fig. 4.1. The response has an exponential form. The constant a is called the *time constant* of the system.

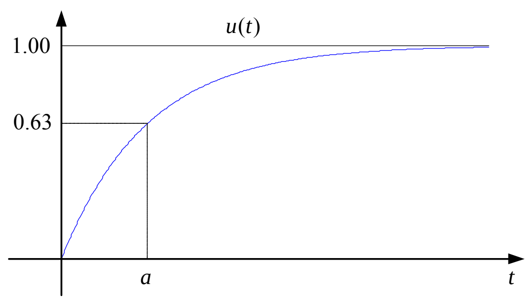

Fig. 4.1

Notice that when $t = a$, then $y(t) = y(a) = 1 - e^{-1} = 0.63$. The response is in two-parts, the transient part $e^{-t/a}$, which approaches zero as $t \to \infty$ and the steady-state part 1, which is the output when $t \to \infty$.

If the derivative of the input are involved in the differential equation of the system, that is, if $a\dot{y} + y = b\dot{u} + u$, then its transfer function is

$$Y(s) = \frac{bs+1}{as+1}U(s) = K\frac{s+z}{s+p}U(s) \tag{4.4}$$

where
- $K = b/a$
- $z = 1/b$: the zero of the system
- $p = 1/a$: the pole of the system

When $U(s) = 1/s$, Eq. (4.4) can be written as

$$Y(s) = \frac{K_1}{s} - \frac{K_2}{s+p}, \quad \text{where } K_1 = K\frac{z}{p} \text{ and } K_2 = K\frac{z-p}{p}$$

Hence,

$$y(t) = \underbrace{K_1}_{steady-state\ part} - \underbrace{K_2 e^{-pt}}_{transient\ part} \tag{4.5}$$

With the assumption that $z > p > 0$, this response is shown in Fig. 4.2.

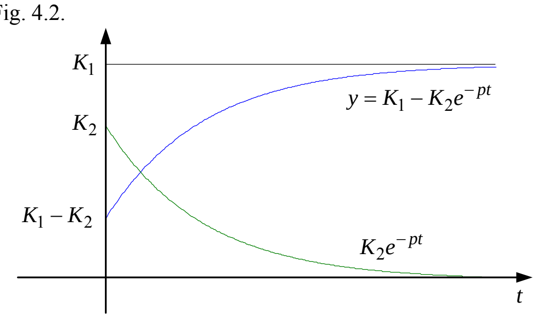

Fig. 4.2

We note that the responses to the systems (Fig. 4.1 and Fig. 4.2) have the same form, except for the constant terms K_1 and K_2. It appears that the role of the numerator of the transfer function is to determine these constants, that is, the size of $y(t)$, but its form is determined by the denominator.

4.3 Response of second order systems

An example of a second order system is a spring-dashpot arrangement shown in Fig. 4.3. Applying Newton's law, we find

$$M\ddot{y} = -\mu\dot{y} - ky + u(t)$$

where k is spring constant, μ is damping coefficient, y is the distance of the system from its position of equilibrium point, and it is assumed that $y(0) = \dot{y}(0) = 0$.

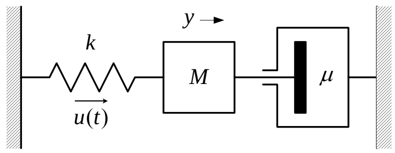

Fig. 4.3

Hence $u(t) = M\ddot{y} + \mu\dot{y} + k\,y$

On taking Laplace transforms, we obtain

$$Y(s) = \frac{1}{Ms^2 + \mu s + k}U(s) = \frac{K}{s^2 + a_1 s + a_2}U(s)$$

where $K = 1/M$, $a_1 = \mu/M$, $a_2 = k/M$. Applying a unit step input, we obtain

$$Y(s) = \frac{K}{s(s+p_1)(s+p_2)} \qquad (4.6)$$

where $p_{1,2} = \dfrac{a_1 \pm \sqrt{a_1^2 - 4a_2}}{2}$, p_1 and p_2 are the poles of the transfer function $G(s) = \dfrac{K}{s^2 + a_1 s + a_2}$, that is, the zeros of the denominator of G(s).

There are there cases to be considered:

Case 1: $a_1^2 > 4a_2 \rightarrow$ over-damped system

In this case p_1 and p_2 are both real and unequal. Eq. (4.6) can be written as

$$Y(s) = \frac{K_1}{s} + \frac{K_2}{s+p_1} + \frac{K_3}{s+p_2} \qquad (4.7)$$

where

$$K_1 = \frac{K}{p_1 p_2} = \frac{K}{a_2},\ K_2 = \frac{K}{p_1(p_1 - p_2)},\ K_3 = \frac{K}{p_2(p_2 - p_1)}$$

(notice that $K_1 + K_2 + K_3 = 0$). On taking Laplace transform of Eq.(4.7), we obtain

$$y(t) = K_1 + K_2 e^{-p_1 t} + K_3 e^{-p_2 t} \qquad (4.8)$$

The transient part of the solution is seen to be $K_2 e^{-p_1 t} + K_3 e^{-p_2 t}$.

Case 2: $a_1^2 = 4a_2 \rightarrow$ critically damped system

In this case, the poles are equal: $p_1 = p_2 = a_1/2 = p$, and

$$Y(s) = \frac{K}{s(s+p)^2} = \frac{K_1}{s} + \frac{K_2}{s+p} + \frac{K_3}{(s+p)^2} \qquad (4.9)$$

Hence $y(t) = K_1 + K_2 e^{-pt} + K_3\, t\, e^{-pt}$, where $K_1 = K/p^2$, $K_2 = -K/p^2$ and $K_3 = -K/p$ so that

$$y(t) = \frac{K}{p^2}(1 - e^{-pt} - p\,t\,e^{-pt}) \qquad (4.10)$$

Case 3: $a_1^2 < 4a_2 \rightarrow$ under-damped system

In this case, the poles p_1 and p_2 are complex conjugate having the form $p_{1,2} = \alpha \pm i\,\beta$ where $\alpha = a_1/2$ and $\beta = \frac{1}{2}\sqrt{4a_2 - a_1^2}$. Hence

$$Y(s) = \frac{K_1}{s} + \frac{K_2}{s+p_1} + \frac{K_3}{s+p_2},$$

where

$$K_1 = \frac{K}{\alpha^2 + \beta^2},\ K_2 = \frac{K(-\beta - i\alpha)}{2\beta(\alpha^2 + \beta^2)},\ K_3 = \frac{K(-\beta + i\alpha)}{2\beta(\alpha^2 + \beta^2)}$$

(Notice that K_2 and K_3 are complex conjugates)
It follows that

$$y(t) = K_1 + K_2 e^{-(\alpha + i\beta)t} + K_3 e^{-(\alpha - i\beta)t}$$

$$= K_1 + e^{-\alpha t}[(K_2 + K_3)\cos\beta t + (K_3 - K_2)i\sin\beta t]$$

(using the relation $e^{i\beta t} = \cos\beta t + i\sin\beta t$)

$$= \frac{K}{\alpha^2 + \beta^2} + \frac{K}{\alpha^2 + \beta^2}e^{-\alpha t}(-\cos\beta t - \frac{\alpha}{\beta}\sin\beta t \qquad (4.11)$$

$$= \frac{K}{\alpha^2 + \beta^2} - \frac{K}{\alpha^2 + \beta^2}\sqrt{1 + \frac{\alpha^2}{\beta^2}}\,e^{-\alpha t}\sin(\beta t + \varepsilon) \qquad (4.12)$$

where $\tan\varepsilon = \beta/\alpha$

Notice that when $t = 0$, $y(t) = 0$. The there cases discussed above are plotted in Fig. 4.4.

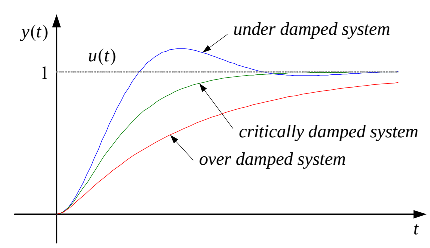

Fig. 4.4

From Fig. 4.4, we see that the importance of damping (note that $a_1 = \mu/M$, μ being the damping factor). We would expect that when the damping is 0 (that is, $a_1 = 0$) the system should oscillate indefinitely. Indeed when $a_1 = 0$, then

$\alpha = 0$, and $\beta = \sqrt{a_2}$

and since $\sin\varepsilon = 1$ and $\cos\varepsilon = 0$, then $\varepsilon = \pi/2$, Eq. (4.12) becomes

$$y(t) = \frac{K}{a_2}\left[1 - \sin\left(\sqrt{a_2}\,t + \frac{\pi}{2}\right)\right] = \frac{K}{a_2}\left[1 - \cos\sqrt{a_2}\,t\right]$$

This response of the undamped system is shown in Fig.4.5.

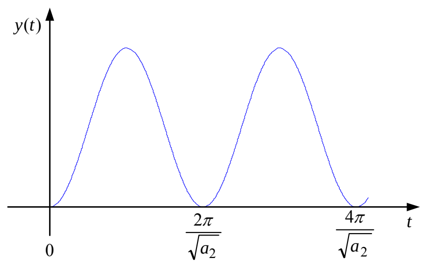

Fig. 4.5

There are two important constants associated with each second order system.

- The *undamped natural frequency* ω_n of the system is the frequency of the response shown in Fig. 4.5: $\omega_n = \sqrt{a_2}$
- The *damping ratio* ξ of the system is the ratio of the actual damping $\mu (= a_1 M)$ to the value of the damping μ_c, which results in the system being critically damped (that is, when $a_1 = 2\sqrt{a_2}$). Hence $\xi = \dfrac{\mu}{\mu_c} = \dfrac{a_1}{2\sqrt{a_2}}$.

We can write equation (4.12) in terms of these constants. We note that $a_1 = 2\omega_n \xi$ and $a_2 = \omega_n^2$. Hence

$$\sqrt{1+\alpha^2/\beta^2} = \sqrt{1+\frac{a_1^2}{4a_2 - a_1^2}} = \frac{2\sqrt{a_2}}{\sqrt{4a_2 - a_1^2}} = \frac{1}{\sqrt{1-\xi^2}}$$

Eq. (4.12) becomes

$$y(t) = \frac{K}{\omega_n^2}\left(1 - \frac{1}{\sqrt{1-\xi^2}} e^{-\omega_n \xi t} \sin(\omega t + \varepsilon)\right) \qquad (4.13)$$

where

$$\omega = \omega_n \sqrt{1-\xi^2} \quad \text{and} \quad \tan\varepsilon = \frac{\sqrt{1-\xi^2}}{\xi}.$$

It is conventional to choose $K/a_2 = 1$ and then plot graphs of the 'normalised' response $y(t)$ against ωt for various values of the damping ratio ξ. There typical graphs are shown in Fig. 4.6.

Some definitions

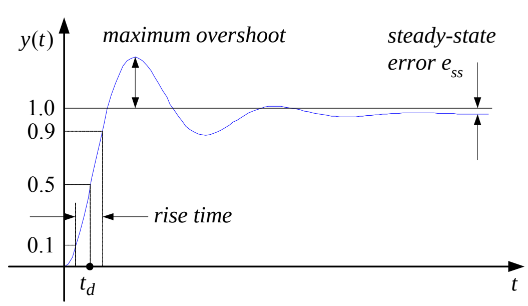

Fig. 4.7

(1) **Overshoot** defined as $\dfrac{\text{maximum overshoot}}{\text{final desired value}} \times 100\%$

(2) **Time delay** t_d, the time required for a system response to reach 50% of its final value.

(3) **Rise time**, the time required for the system response to rise from 10% to 90% of its final value.

(4) **Settling time**, the time required for the eventual settling down of the system response to be within (normally) 5% of its final value.

(5) **Steady-state error** e_{ss}, the difference between the steady state response and the input.

In fact, one can often improve one of the parameters but at the expense of the other. For example, the overshoot can be decreased at the expense of the time delay.

In general, the quality of a system may be judged by various standards. Since the purpose of a servomechanism is to make the output follow the input signal, we may define expressions which will describe dynamics accuracy in the transient state. Such expression are based on the *system error*, which in its simplest form is the difference between the input and the output and is denoted by $e(t)$, that is, $e(t) = y(t) - u(t)$, where $y(t)$ is the actual output and $u(t)$ is the desired output ($u(t)$ is the input).

The expression called the *performance index* can take on various forms, typically among them are:

(1) integral of error squared (IES) $\int_0^\infty e^2(t)dt$

(2) integral of absolute error (IAS) $\int_0^\infty |e|(t)dt$

(3) integral of time multiplied absolute error criterion (ITAE) $\int_0^\infty t|e|(t)dt$

Having chosen an appropriate performance index, the system which minimizes the integral is called *optimal*. The object of modern control theory is to design a system so that it is optimal with respect to a performance index and will be discussed in the part II of this course.

4.4 Response of higher order systems

We can write the transfer function of an n^{th} - order system in the form

$$G(s) = \frac{K(s^m + b_1 s^{m-1} + \ldots + b_m)}{s^n + a_1 s^{n-1} + \ldots + a_n} \qquad (4.14)$$

Example 4.1

With reference to Fig. 2.11, calculate the close loop transfer function $G(s)$ given the transfer functions $A(s) = \dfrac{1}{s+3}$ and $B(s) = 2/s$

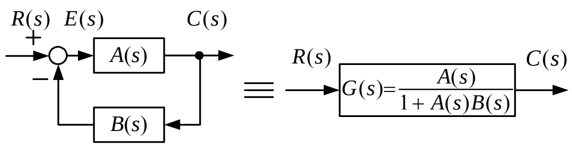

Fig. 2.11

We obtain

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

The response of the system having the transfer function (4.14) to a unit step input can be written in the form

$$Y(s)=\frac{K(s+z_1)(s+z_2)\cdots(s+z_m)}{s(s+p_1)(s+p_2)\cdots(s+p_n)} \qquad (4.15)$$

where

- $z_1,z_2,\cdots,z_m$: the zeros of the numerator
- $p_1,p_2,\cdots,p_n$: the zeros of the denominator

We first assume that $n\ge m$ in equation (4.14); we then have two cases to consider:

Case 1: $p_1,p_2,\cdots,p_n$ are all distinct numbers. The partial fraction expansion of equation (4.15) has the form

$$Y(s)=\frac{K_1}{s}+\frac{K_2}{s+p_1}+\cdots+\frac{K_{n+1}}{s+p_n} \qquad (4.16)$$

$K_1,K_2,\cdots,K_{n+1}$ are called the *residues* of the expansions. The response has the form

$$y(t)=K_1+K_2e^{-p_1t}+\ldots+K_{n+1}e^{-p_nt}$$

Case 2: $p_1,p_2,\cdots,p_n$ are not distinct any more. Here at least one of the roots, say p_1, is of multiplicity r, that is

$$Y(s)=\frac{K(s+z_1)(s+z_2)\cdots(s+z_m)}{s(s+p_1)^r(s+p_2)\cdots(s+p_n)} \qquad (4.17)$$

The partial fraction expansion of equation (4.17) has the form

$$Y(s)=\frac{K_1}{s}+\frac{K_{21}}{s+p_1}+\cdots+\frac{K_{2r}}{(s+p_1)^r}+\cdots+\frac{K_{n-r+2}}{s+p_{n-r+1}} \qquad (4.18)$$

Since $\mathscr{L}^{-1}\left\{\frac{K}{(s+p)^j}\right\}=\frac{K}{(j-1)!}t^{j-1}e^{-pt}$, $(j=1,2,\cdots,r)$, the response has the form

$$y(t)=K_1+K_{21}e^{-p_1t}+K_{22}e^{-p_1t}+\ldots+\frac{K_{2r}}{(r-1)!}t^{r-1}e^{-p_1t}+$$
$$K_3e^{-p_2t}+\cdots+K_{n-r+2}e^{-p_{n-r+1}t} \qquad (4.19)$$

We now consider that $n<m$ in equation (4.14); which is the case when the system is improper; that is, it can happen when we consider idealized and physically non-realisable systems, such as resistanceless circuits. We then divide the numerator until we obtain a proper fraction so that when applying a unit step input, we can write $Y(s)$ as

$$Y(s)=K(c_1+c_2s+\ldots+cs^{n-m-1})+\frac{K_1(s^n+d_1s^{n-1}+\ldots+d_n)}{s(s^n+a_1s^{n-1}+\ldots+a_n)} \qquad (4.20)$$

where c_s,d_s,K and K_1 are all constants.

The inverse Laplace transform of the first right term of (4.20) involves the impulse function and various derivatives of it. The second term of (4.20) is treated as in Case 1 or Case 2 above.

Example 4.2

Find the response of the system having a transfer function

$$G(s)=\frac{5(s^2+5s+6)}{s^3+6s^2+10s+8}$$

to a unit step input.

In this case,

$$Y(s)=\frac{5(s^2+5s+6)}{s^3+6s^2+10s+8}=\frac{5(s+2)(s+3)}{s(s+4)[s+(1+i)][s+(1-i)]}$$

The partial fraction expansion as

$$Y(s)=\frac{K_1}{s}+\frac{K_2}{s+4}+\frac{K_3}{s+(1+i)}+\frac{K_4}{s+(1-i)}$$

where $K_1=\frac{15}{4}$, $K_2=-\frac{1}{4}$, $K_3=\frac{-7+i}{4}$, $K_4=\frac{-7-i}{4}$.

Hence

$$y(t)=\frac{15}{4}-\frac{1}{4}e^{-4t}+\frac{1}{4}e^{-t}(-14\cos t+2\sin t)$$
$$=\frac{15}{4}-\frac{1}{4}e^{-4t}+\frac{\sqrt{2000}}{4}e^{-t}\sin(t+352)$$

4.5 Steady state error

Consider a unity feedback system as in Fig. 4.8

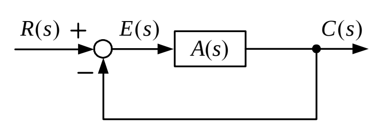

Fig. 4.8

where

- $r(t)$: reference input
- $c(t)$: system output
- $e(t)$: error

We define the *error function* as

$$e(t)=r(t)-c(t) \qquad (4.21)$$

hence, $e_{ss} = \lim_{t\to\infty} e(t)$. Since $E(s) = R(s) - A(s)E(s)$, it follows that $E(s) = \dfrac{R(s)}{1+A(s)}$ and by the final value theorem

$$e_{ss} = \lim_{s\to 0} sE(s) = \lim_{s\to 0} \frac{sR(s)}{1+A(s)} \qquad (4.22)$$

We now define three error coefficients which indicate the steady state error when the system is subjected to three different standard reference inputs $r(s)$.

(1) **step input**, $r(t) = ku(t)$ (k is a constant)

$$e_{ss} = \lim_{s\to 0} \frac{s\,k/s}{1+A(s)} = \frac{k}{1+\lim_{s\to 0} A(s)}$$

$\lim_{s\to 0} A(s) = K_p$, called the *position error constant*, then

$$e_{ss} = \frac{k}{1+K_p} \quad \text{or} \quad K_p = \frac{k - e_{ss}}{e_{ss}} \qquad (4.23)$$

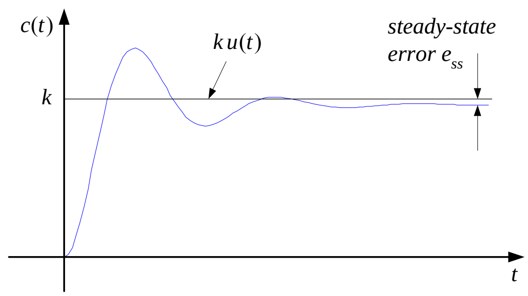

Fig. 4.9

(2) **Ram input**, $r(t) = ktu(t)$ (k is a constant)

In this case, $R(s) = \dfrac{k}{s^2}$, so that $e_{ss} = \dfrac{k}{K_v}$ or $K_v = \dfrac{k}{e_{ss}}$, where $K_v = \lim_{s\to 0} sA(s)$ is called the *velocity error constant*.

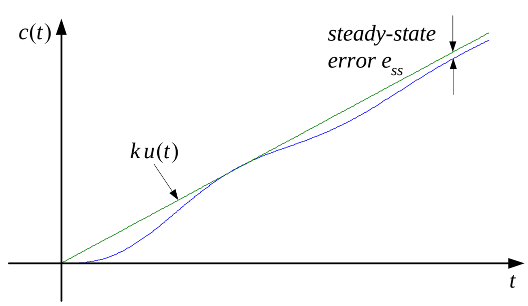

Fig. 4.10

(3) **Parabolic input**, $r(t) = \dfrac{1}{2}kt^2 u(t)$ (k is a constant)

In this case, $R(s) = \dfrac{k}{s^3}$, so that $e_{ss} = \dfrac{k}{K_a}$ or $K_a = \dfrac{k}{e_{ss}}$, where

$K_a = \lim_{s\to 0} s^2 A(s)$ is called the *acceleration error constant*.

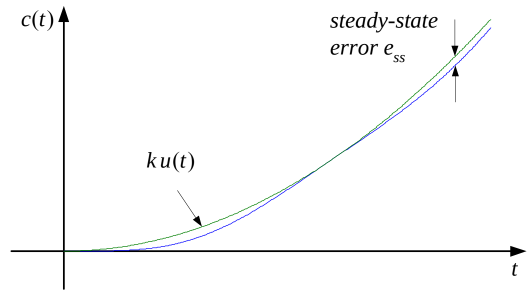

Fig. 4.11

Example 4.3

Find the (static) error coefficients for the system having a open loop transfer function $A(s) = \dfrac{8}{s(4s+2)}$

$$K_p = \lim_{s\to 0} A(s) = \infty$$
$$K_v = \lim_{s\to 0} sA(s) = 4$$
$$K_a = \lim_{s\to 0} s^2 A(s) = 0$$

From the definition of the error coefficients, it is seen that e_{ss} depends on the number of poles at $s = 0$ of the transfer function. This leads to the following classification. A transfer function is said to be of type N if it has N poles at the origin. Thus if

$$A(s) = \frac{K(s-z_1)\cdots(s-z_m)}{s^j (s-p_1)\cdots(s-p_n)} \qquad (4.24)$$

At $s = 0$, $A(s) = \lim_{s\to 0} \dfrac{K_1}{s^j}$ where $K_1 = \dfrac{K(-z_1)\cdots(-z_m)}{(-p_1)\cdots(-p_n)}$ (4.25)

K_1 is called the *gain* of the transfer function. Hence the steady state error e_{ss} depends on j and $r(t)$ as summarized in Table 4.1

Table 4.1

j	System	e_{ss}		
		$r(t)=ku(t)$	$r(t)=ktu(t)$	$r(t)=½kt^2u(t)$
0	Type 1	Finite	∞	∞
1	Type 2	0	finite	∞
2	Type 3	0	0	finite

4.6 Feedback Control

Consider a negative feedback system in Fig. 4.12

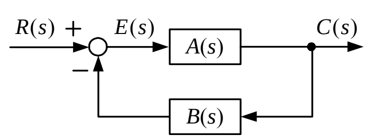

Fig. 4.12

The *close loop* transfer function is related to the *feed-forward* transfer function $A(s)$ and *feedback* transfer function $B(s)$ by

$$G(s) = \frac{A(s)}{1 + A(s)B(s)} \quad (4.26)$$

We consider a simple example of a first order system for which $A(s) = \frac{K}{as+1}$ and $B(s) = c$, so that

$$G(s) = \frac{K}{as + Kc + 1} = \frac{K/a}{s + \frac{Kc+1}{a}}$$

On taking Laplace inverse transform, we obtain the impulse response of the system, where

(1) $c = 0$ (response of open loop system): $g(t) = \frac{K}{a}e^{-t/a}$

(2) $c \neq 0$: $g(t) = \frac{K}{a}e^{-\frac{Kc+1}{a}t} = \frac{K}{a}e^{-\frac{t}{\alpha}}$, where $\alpha = \frac{a}{Kc+1}$

a and α are respectively the time-constants of the open loop and closed loop systems. a is always positive, but α can be either positive or negative.

Fig. 4.13 shows how the time responses vary with different values of Kc.

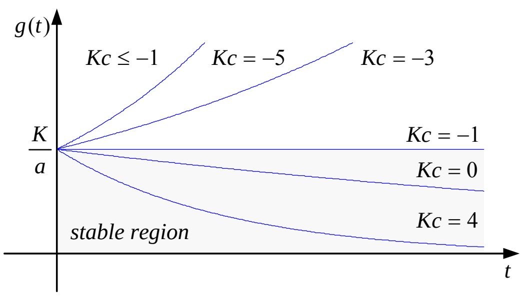

Fig. 4.13

If the impulse response does not decay to zero as t increase, the system is *unstable*. From the Fig. 4.13, the instability region is defined by $Kc \leq -1$.

In many applications, the control system consists basically of a plant having the transfer function $A(s)$ and a *controller* having a transfer function $B(s)$, as in Fig. 4.14.

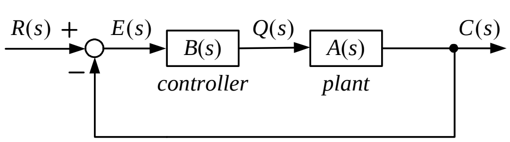

Fig. 4.14

With the closed loop transfer function

$$G(s) = \frac{A(s)B(s)}{1 + A(s)B(s)} \quad (4.27)$$

The controllers can be of various types.

(1) **The on-off Controller**
The action of such a controller is very simple.

$$q(t) = \begin{cases} Q_1 & \text{if } e(t) > 0 \\ Q_2 & \text{if } e(t) < 0 \end{cases}$$

where $q(t)$ is output signal from the controller
Q_1, Q_2 are some constants

The on-off controller is obviously a nonlinear device and it cannot be described by a transfer function.

(2) **Proportional Controller**
For this control action

$q(t) = K_p e(t)$

where K_p is a constant, called the *controller gain*. The transfer function of this controller is

$$B(s) = \frac{Q(s)}{E(s)} = K_p \quad (4.28)$$

(3) **Integral Controller**

In this case $q(t) = K\int_0^t e(t)dt$, hence

$$B(s) = K/s \quad (4.29)$$

(4) **Derivative Controller**

In this case $q(t) = K\frac{de}{dt}$, hence

$$B(s) = Ks \quad (4.30)$$

(5) **Proportional-Derivative Controller (PD)**

In this case $q(t) = K_p e(t) + K_1\frac{de}{dt}$, hence

$$B(s) = K_p\left(1 + \frac{K_1}{K_p}s\right) = K_p(1 + Ks) \quad (4.30)$$

(6) **Proportional-Integral Controller (PI)**

In this case $q(t) = K_p e(t) + K_1\int_0^t e(t)dt$, hence

$$B(s) = K_p\left(1 + \frac{K_1}{K_p}\frac{1}{s}\right) = K_p\left(1 + \frac{K}{s}\right) \quad (4.31)$$

(7) **Proportional-Derivative-Integral Controller (PID)**

In this case $q(t) = K_p e(t) + K_1\frac{de}{dt} + K_2\int_0^t e(t)dt$, hence

$$B(s) = K_p\left(1 + \frac{K_1}{K_p}s + \frac{K_2}{K_p}\frac{1}{s}\right) = K_p(1 + k_1 s + k_2/s)$$

Example 4.4

Design a controller for a plant having the transfer function $A(s) = 1/(s+2)$ so that the resulting closed loop system has a zero steady state error to a reference ramp input.

For zero steady state error to a ramp input, the system must be of type 2. Hence if we choose an integral controller with $B(s) = K/s$ then the transfer function of the closed loop system including the plant and the controller is

$$\frac{A(s)B(s)}{1 + A(s)B(s)} = \frac{K}{s^3 + 2s^2 + K}$$

The response of this control system depends on the roots of the denominator polynomial $s^3 + 2s^2 + K = 0$.

If we use PI controller, $B(s) = K_p(1 + K/s)$ the system is of type 2 and response of the system depends on the roots of $s^3 + 2s^2 + K_p s + KK_p = 0$
