1.1 INTRODUCTION

A system is a combination of components connected to perform a required action. The control component of a system plays a major role in altering or maintaining the system output based on our desired characteristics. There are two types of control systems: manual control and automatic control. For example, in manual control, a man can switch ON or OFF the bore well motor to control the level of water in a tank. On the other hand, in automatic control, level switches and transducers are used to control the level of water in a tank. Control systems have naturally evolved in our ecosystem. In almost all living things, automatic control regulates the conditions necessary for life by tackling the disturbance through sensing and controlling functionalities. They operate complex systems and processes and achieve control with desired precision. The application of control systems facilitates automated manufacturing processes, accurate positioning and effective control of machine tools. They guide and control space vehicles, aircrafts, ships and high-speed ground transportation systems. modern automation of a plant involves components such as sensors, instruments, computers and application of techniques that involve data processing and control. It is essential to understand a system and its characteristics with the help of a model, before creating a control for it. The process of developing a model is known as modeling. Physical systems are modeled by applying notable laws that govern their behavior. For example, mechanical systems are described by Newton's laws and electrical systems are described by Ohm's law, Kirchhoff's laws, Faraday's laws and Lenz's law. These laws form the basis for the constitutive properties of the elements in a system.

BASIC ELEMENTS IN CONTROL SYSTEMS

In recent years, control systems have gained an increasingly importance in the development and advancement of the modern civilization and technology. Disregard the complexity of the system; it consists of an input (objective), the control system and its output (result). Practically our day-to-day activities are affected by some type of control systems. There are four basic elements of a typical motion control system. These are

- Controller
- Amplifier

- Actuator
- Feedback
- Error detector

The complexity of each of these elements will vary depending on the types of applications for which they are designed and built. A dynamical system manipulates entities such as energy, material, information, capital investment etc. It is characterized by relationships among certain variables that are chosen in its description. Usually inputs (causes) and outputs (effects) are important variables, which are connected by relations. Although a relationship is a function of time, the properties embedded in it may be time-invariant. A system may have only one input and one output. Such a system is termed a single-input-single-output (SISO) system. Some may be multiple-input-multiple-output (MIMO) systems. Large systems are characterized by several levels of organization, in a hierarchy. Figure 1 shows the schematic diagrams of systems indicating such features. The fields of systems, control and information processing are closely related to the science of cybernetics. Cybernetics attempts to understand the behavior of the system in nature. This understanding leads to the knowledge enabling us to improve the performance of natural or man-made processes.

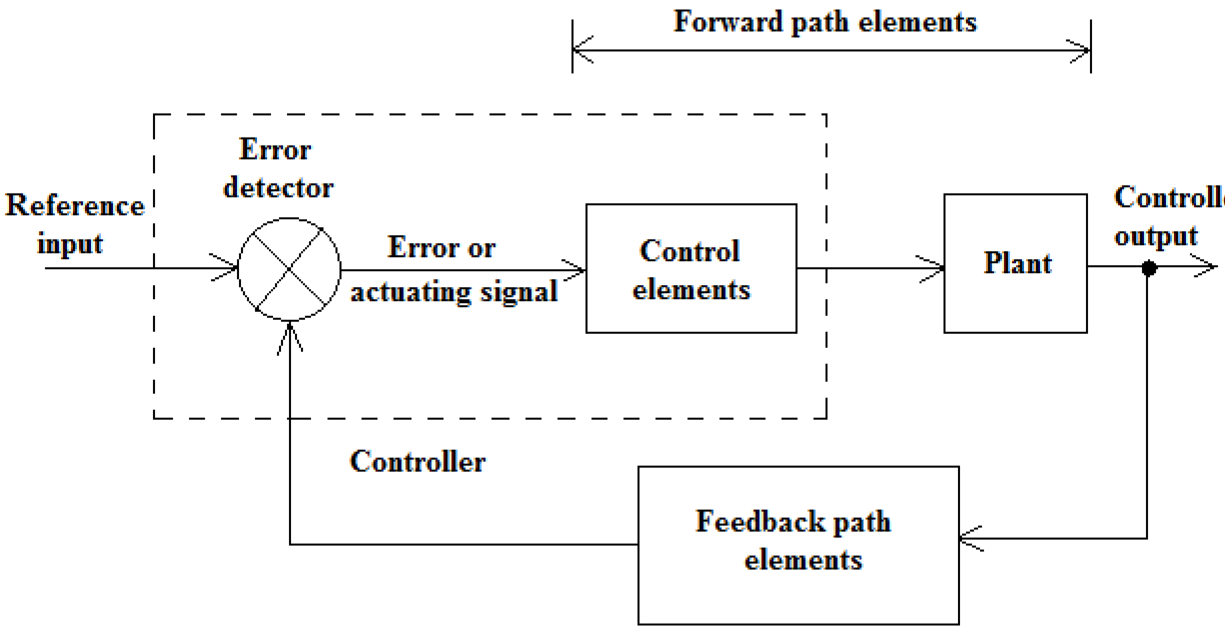


Figure 1.1.1 Basic Elements in Control Systems

[Source: "Control Systems Engineering" by I.J.Nagrath, M.Gopal, Page: 5]

1.2 OPEN AND CLOSED LOOP SYSTEMS

In recent years, control systems have gained an increasingly importance in the development and advancement of the modern civilization and technology. Figure shows the basic components of a control system. Disregard the complexity of the system; it consists of an input (objective), the control system and its output (result). Practically our day-to-day activities are affected by some type of control systems. There are two main branches of control systems:

1) Open-loop systems and 2) Closed-loop systems

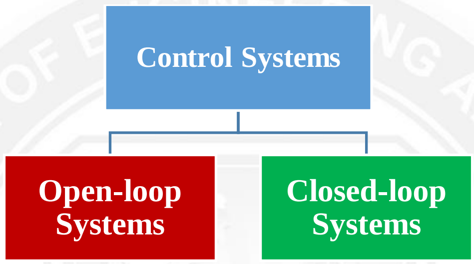


Figure 1.2.1 Classification of Control Systems

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.2]

OPEN LOOP SYSTEMS

A control system that cannot adjust itself to the changes is called open-loop control system. In general, manual control systems are open-loop systems. The block diagram of open-loop control system is shown in figure.

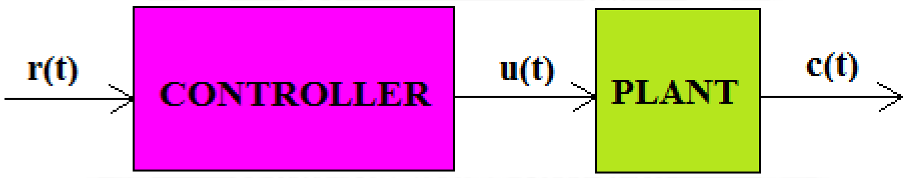


Figure 1.2.2 Block diagram of open loop system

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.2]

Here, r(t) is the input signal, u(t) is the control signal/actuating signal and c(t) is the output signal. In this system, the output remains unaltered for a constant input. In case of any discrepancy, the input should be manually changed by an operator. An open loop control system is suited when there is tolerance for fluctuation in the system and when the system parameter variation can be handled irrespective of the environmental conditions.

PRACTICAL EXAMPLES OF OPEN LOOP CONTROL SYSTEM

1. Electric Hand Drier-Hot air (output) comes out as long as you keep your hand under the machine, irrespective of how much your hand is dried.

2. Automatic Washing Machine-This machine runs according to the pre-set time irrespective of washing is completed or not.
3. Bread Toaster-This machine runs as per adjusted time irrespective of toasting is completed or not.
4. Automatic Tea/Coffee Maker-These machines also function for pre adjusted time only.
5. Timer Based Clothes Drier-This machine dries wet clothes for pre-adjusted time, it does not matter how much the clothes are dried.
6. Light Switch-Lamps glow whenever light switch is on irrespective of light is required or not.
7. Volume on Stereo System-Volume is adjusted manually irrespective of output volume level.

Advantages of Open Loop Control System

a) Simple in construction and design
b) Economical
c) Easy to maintain
d) Generally stable
e) Convenient to use as output is difficult to measure.

Disadvantages of Open Loop Control System

a) They are inaccurate
b) They are unreliable
c) Any change in output cannot be corrected automatically.

CLOSED LOOP SYSTEMS

Any system that can respond to the changes and make corrections by itself is known as closed loop control system. The only difference when compared to open loop system is the presence of feedback action. The block diagram of a closed loop system is shown in the figure. Here, r(t) is the input signal, e(t) is the error signal/actuating signal, u(t) or m(t) is the control signal/manipulated signal, b(t) is the feedback signal and c(t) is the controlled output. Here, the output of the machine is fed back to a comparator (error detector). The output signal is compared with the reference input, r(t) and the error signal, e(t) is sent to the controller. Based on the error, the controller adjusts the air conditioners input [control signal u(t)]. This process is continued till the error gets nullified.

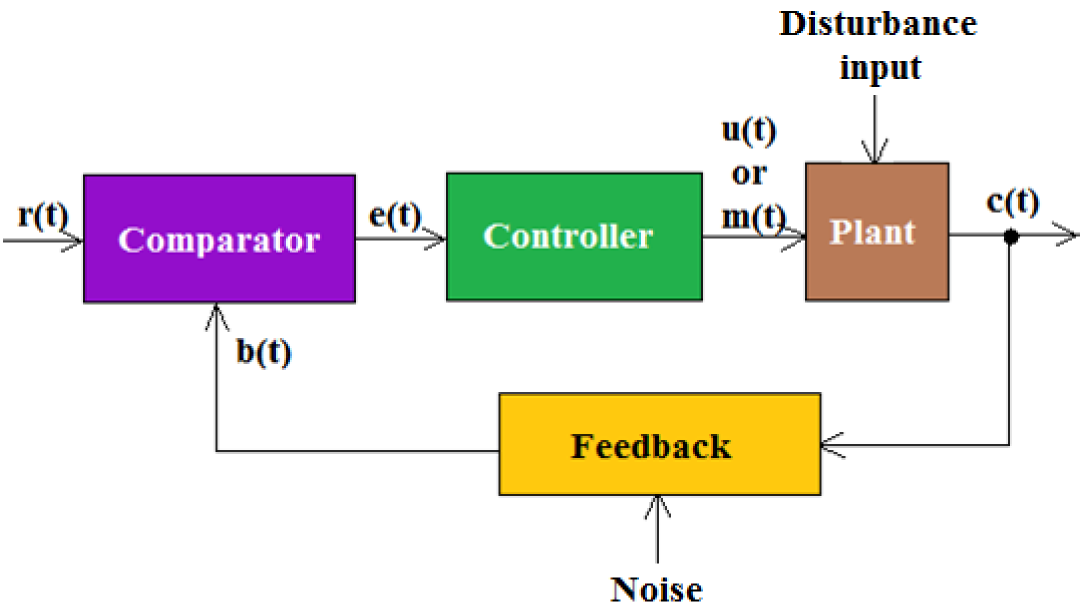


Figure 1.2.3 Block diagram of closed loop system

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.3]

Both the manual and automatic controls can be implemented in a closed loop system.

PRACTICAL EXAMPLES OF CLOSED LOOP CONTROL SYSTEM

1) Automatic Electric Iron-Heating elements are controlled by output temperature of the iron.
2) Servo Voltage Stabilizer-Voltage controller operates depending upon output voltage of the system.
3) Water Level Controller-Input water is controlled by water level of the reservoir.
4) Missile Launched and Auto Tracked by Radar-The direction of missile is controlled by comparing the target and position of the missile.
5) An Air Conditioner-An air conditioner functions depending upon the temperature of the room.
6) Cooling System in Car-It operates depending upon the temperature which it controls.

Advantages of Closed Loop Control System

a) Closed loop control systems are more accurate even in the presence of non-linearity.
b) Highly accurate as any error arising is corrected due to presence of feedback signal.

c) Bandwidth range is large.

d) Facilitates automation.

e) The sensitivity of system may be made small to make system more stable.

f) This system is less affected by noise.

Disadvantages of Closed Loop Control System

a) They are costlier.

b) They are complicated to design.

c) Required more maintenance.

d) Feedback leads to oscillatory response.

e) Overall gain is reduced due to presence of feedback.

f) Stability is the major problem and more care is needed to design a stable closed loop system.

S. No.	Open loop control system	Closed loop control system
1	Inaccurate	Accurate
2	Unreliable	Reliable
3	Stable	Unstable. It can be stabilized using the feedback or by reducing sensitivity
4	Bandwidth is small	Bandwidth is large
5	System is affected by noise	System is less affected by noise
6	Cheap	Costly
7	Simple in construction	Complex construction since a greater number of components are present
8	Requires less maintenance	Requires more maintenance
9	Overall gain is high	Overall high is reduced due to feedback

1.3 MECHANICAL TRANSLATIONAL AND ROTATIONAL SYSTEMS

The general classification of mechanical system is of two types namely translational and rotational systems.

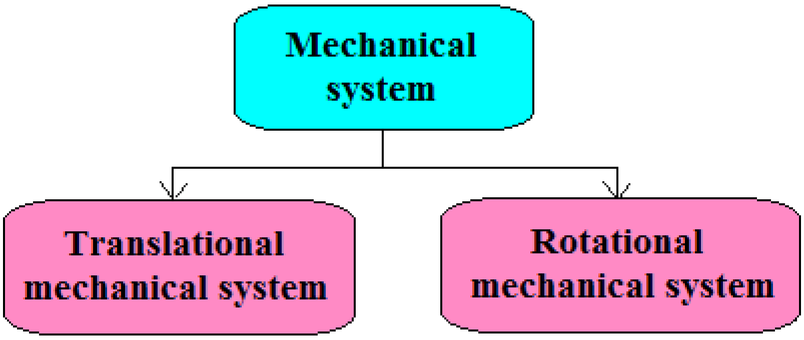


Figure 1.3.1 Classification of mechanical system

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.21]

MECHANICAL TRANSLATIONAL SYSTEMS

The model of mechanical translational systems can obtain by using three basic elements mass, spring and dashpot. When a force is applied to a translational mechanical system, it is opposed by opposing forces due to mass, friction and elasticity of the system. The force acting on a mechanical body is governed by Newton's second law of motion. For translational systems it states that the sum of forces acting on a body is zero.

Force balance equations of idealized elements:

Inertia force, $f_m(t)$

Consider an ideal mass element shown in figure, which has negligible friction and elasticity. Let a force be applied on it. The mass will offer an opposing force which is proportional to acceleration of a body.

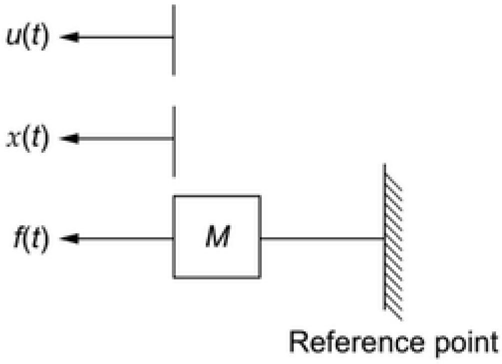


Figure 1.3.2 Mechanical translational element: Mass

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.21]

Let f(t) - applied force, f_m - opposing force due to mass,

$$f_m \propto \frac{d^2x}{dt^2}$$

By Newton's second law,

$$f = f_m = M\frac{d^2x}{dt^2}$$

Damper force, $f_b(t)$

Consider an ideal frictional element dash-pot shown in fig. which has negligible mass and elasticity. The dashpot's opposing force which is proportional to velocity of the body.

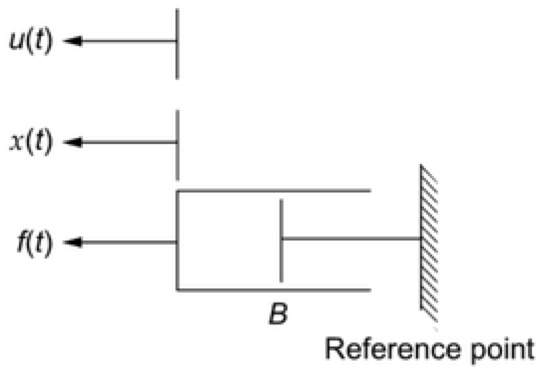


Figure 1.3.3 Mechanical translational element: Dashpot

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.23]

Let f = applied force, f_b = opposing force due to friction

$$f_b \propto \frac{dx}{dt}$$

By Newton's second law,

$$f = f_b = B\frac{dx}{dt}$$

Spring force, $f_k(t)$

Consider an ideal elastic element spring is shown in fig. This has negligible mass and friction.

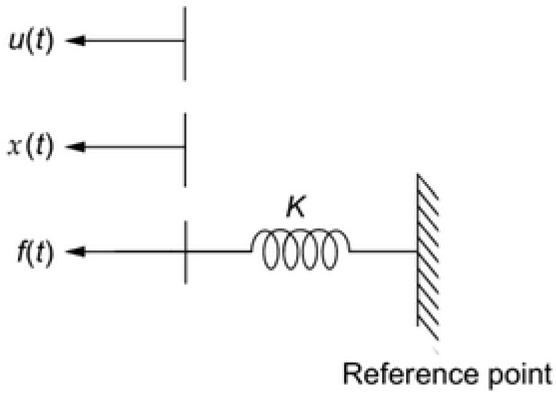


Figure 1.3.4 Mechanical translational element: Spring

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.24]

Let f = applied force, f_k = opposing force due to elasticity

$$f_k \propto x$$

By Newtons second law,

$$f = f_k = Kx$$

According to D'Alembert's principle, "The algebraic sum of the externally applied forces to any body is equal to the algebraic sum of the opposing forces restraining motion produced by the elements present in the body." A simple translational mechanical system and its free body diagram are shown in figures 1.3.5 (a) and (b) respectively.

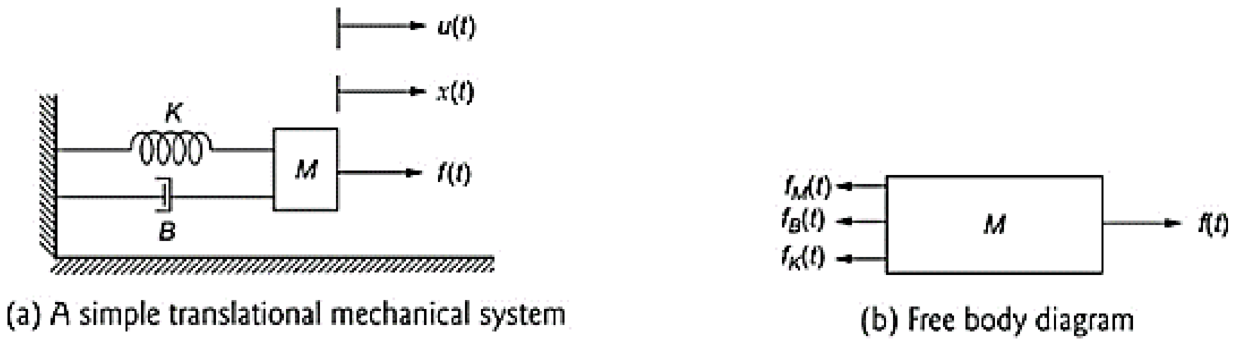


Figure 1.3.5 Mechanical translational system and its free body diagram

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.25]

$$f_m = M\frac{d^2x}{dt^2}$$

$$f_b = B\frac{dx}{dt}$$

$$f_k = Kx$$

$$f(t) = f_m + f_b + f_k = M\frac{d^2x}{dt^2} + B\frac{dx}{dt} + Kx$$

MECHANICAL ROTATIONAL SYSTEM

The modeling of a linear passive rotational mechanical system can be obtained by using three basic elements: inertia, rotational spring and rotational damper. The modeling of a rotational mechanical system is similar to that of a translational mechanical system except that the elements undergo a rotational instead of a translational movement. The opposing torques due to inertia, rotational spring and rotational damper act on a system when the system is subjected to a torque. Using D'Alembert's principle, for a linear passive rotational mechanical system, the sum of all the torques acting on a body is zero (i.e., the sum of applied torques is equal to the sum of the opposing torques on a body). Angular displacement, angular velocity and angular acceleration are the variables used to describe a linear passive rotational mechanical system. In rotational mechanical systems, the energy storage elements are inertia and rotational spring and the energy dissipating element is the rotational viscous damper. The analogous of the energy storage elements in ana electrical circuit are the inductors and the capacitors and the analogous of energy dissipating element in an electrical circuit is the resistor.

Torque balance equations of idealized elements:

Inertia Torque, $T_j(t)$

When a torque T(t) is applied to an inertia element J, it experiences an angular acceleration and it is shown in figure 1.3.6.

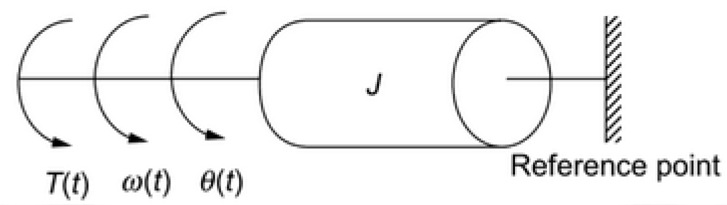


Figure 1.3.6 Mechanical rotational element: Inertia

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.38]

According to Newton's second law, the inertia torque is proportional to the angular acceleration.

$$T_j(t) \propto \frac{d^2\theta}{dt^2}$$

$$T_j(t) = J\frac{d^2\theta}{dt^2}$$

where J is the moment of inertia (kg-m^2/rad), θ(t) is the angular displacement (rad) and $T_j(t)$ is measured in Newton-meter (N-m).

Damping Torque, $T_b(t)$

When a torque, T(t) is applied to a damping element, B, it experiences an angular velocity and it is shown in figure 1.3.7.

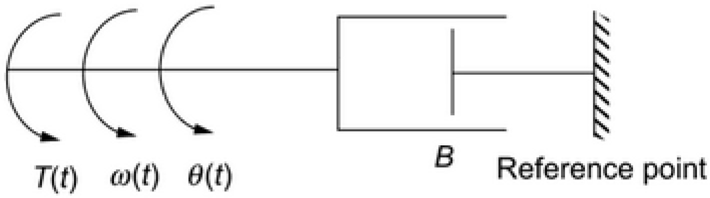


Figure 1.3.7 Mechanical rotational element: Dashpot

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.38]

The damping torque is proportional to the angular velocity. Therefore,

$$T_b(t) \propto \frac{d\theta}{dt}$$

$$T_b(t) = B\frac{d\theta}{dt}$$

where, B is the viscous friction coefficient (N-s/m), θ(t) is the angular displacement (rad). Damper element with two angular displacements and a single applied torque is shown in figure 1.3.8.

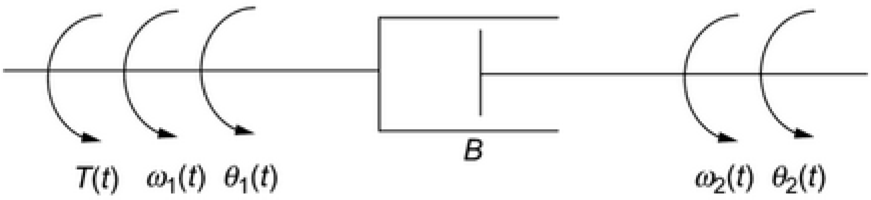


Figure 1.3.8 Mechanical rotational element: Dashpot

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.39]

$$T_b(t) = B\left(\frac{d\theta_1}{dt} - \frac{d\theta_2}{dt}\right)$$

Here, $T_b(t)$ is measured in Newton-meter.

Torsional/Rotational Spring Torque, $T_k(t)$

When a torque T(t) is applied to a spring element, K, it experiences ana angular displacement and it is shown in figure 1.3.9.

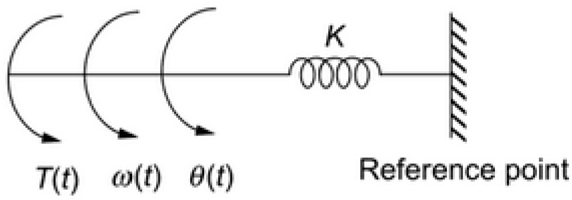


Figure 1.3.9 Mechanical rotational element: Dashpot

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.39]

According to Hooke's law, spring torque is proportional to the angular displacement.

$$T_k(t) \propto \theta$$

$$T_k(t) = K\theta$$

where, K is the spring constant (N-m/rad).

A spring element with two angular displacements is given in figure 1.3.10.

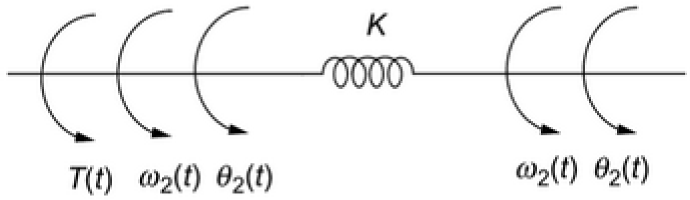


Figure 1.3.10 Mechanical rotational element: Dashpot

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.40]

$$T_k(t) = K(\theta_1 - \theta_2)$$

Here, $T_k(t)$ is measured in Newton-meter.

According to D'Alembert's principle, "The algebraic sum of the externally applied torques to any body is equal to the algebraic sum of the opposing torques restraining motion produced by the elements present in the body." A simple rotational mechanical system and its free body diagram are shown in figures 1.3.11 (a) and (b) respectively.

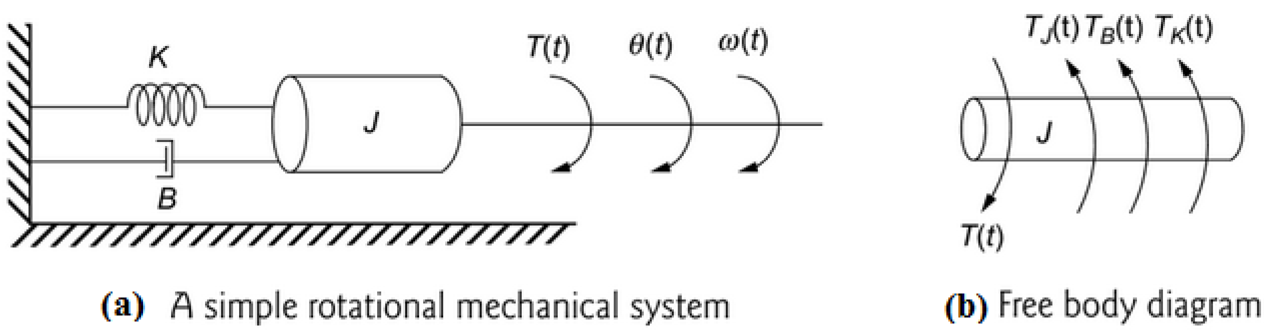


Figure 1.3.11 Mechanical rotational system and its free body diagram

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.40]

$$T_j = J\frac{d^2\theta}{dt^2}$$

$$T_b = B\frac{d\theta}{dt}$$

$$T_k = K\theta$$

$$T(t) = T_j + T_b + T_k = J\frac{d^2\theta}{dt^2} + B\frac{d\theta}{dt} + K\theta$$

Translational mechanical system	**Rotational mechanical system**
Force (F)	Torque (T)
Velocity (v)	Angular velocity (ω)
Displacement (x)	Angular displacement (θ)
Mass (M)	Moment of inertia (J)
Damping coefficient (B)	Rotational damping (B)
Spring constant (K)	Rotational spring constant (K)

1.4 ELECTRICAL ANALOGY OF MECHANICAL SYSTEMS

FORCE-VOLTAGE ANALOGY

Consider a simple translational mechanical system as shown in figure 1.4.1.

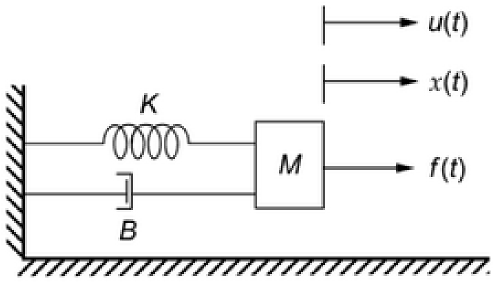


Figure 1.4.1 Translational mechanical system

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.51]

Using D' Alembert's principle, we have,

Sum of the applied forces = Sum of the opposing forces

$$f(t) = M\frac{du(t)}{dt} + B\,u(t) + K\int u(t)\,dt$$

Consider a series RLC circuit as shown in figure 1.4.2.

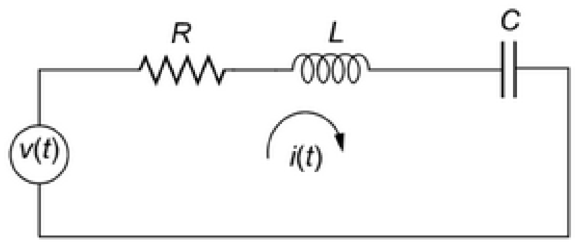


Figure 1.4.2 Series RLC circuit

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.52]

Using KVL, the integro-differential equations can be written as

$$v(t) = L\frac{di(t)}{dt} + R\,i(t) + \frac{1}{C}\int i(t)\,dt$$

Translational system	Electrical system
Force (f)	Voltage (v)
Velocity (u)	Current (i)
Displacement (x)	Charge (q)
Mass (M)	Inductance (L)
Damping coefficient (B)	Resistance (R)
Spring constant (K)	1/Capacitance (C)

FORCE-CURRENT ANALOGY

Consider a simple parallel RLC circuit as shown in figure 1.4.3.

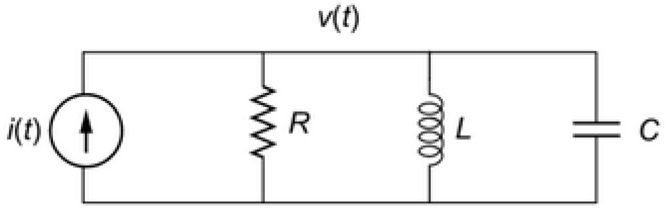


Figure 1.4.3 Parallel RLC circuit

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.52]

Using KCL, the integro-differential equations can be written as follows:

$$i(t) = C\frac{dv(t)}{dt} + G\,v(t) + \frac{1}{L}\int v(t)\,dt$$

where, conductance, G=1/R.

On comparing with the mechanical translational system equation, we get,

Translational System	Electrical System
Force (f)	Current (i)
Velocity (u)	Voltage (v)
Displacement (x)	Flux (Φ)
Mass (M)	Capacitance (C)
Damping coefficient (B)	Conductance (G)
Spring constant (K)	1/Inductance (L)

TORQUE-VOLTAGE ANALOGY

Consider a simple rotational mechanical system as shown in figure 1.4.4.

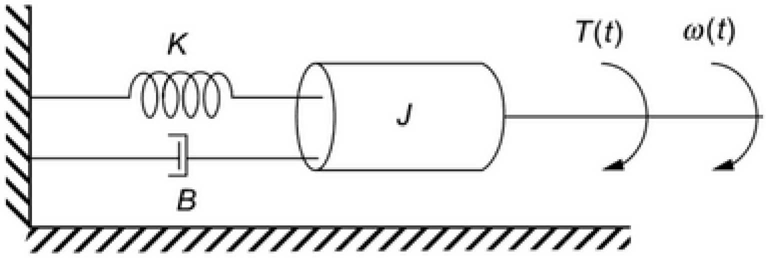


Figure 1.4.4 Rotational mechanical system

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.71]

Using D' Alembert's principle, we have,

Sum of the applied torques = Sum of the opposing torques

$$T(t) = J\frac{d\omega(t)}{dt} + B\,\omega(t) + K\int \omega(t)\,dt$$

Consider a series RLC circuit as shown in figure 1.4.5.

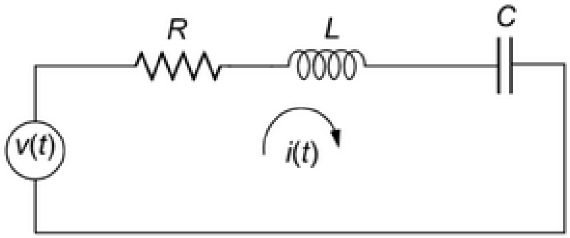


Figure 1.4.5 Series RLC circuit

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.72]

Using KVL, the integro-differential equations can be written as

$$v(t) = L\frac{di(t)}{dt} + R\,i(t) + \frac{1}{C}\int i(t)\,dt$$

On comparing with the mechanical rotational system equation, we get,

Rotational System	Electrical System
Torque (T)	Voltage (v)
Angular velocity (ω)	Current (i)
Angular displacement (θ)	Charge (q)
Moment of inertia (J)	Inductance (L)
Rotational damping (B)	Resistance (R)
Rotational spring constant (K)	1/Capacitance (C)

TORQUE-CURRENT ANALOGY

Consider a simple parallel RLC circuit as shown in figure 1.4.6.

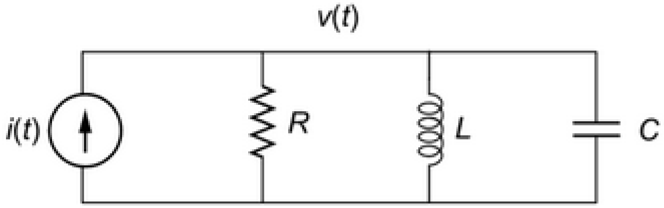


Figure 1.4.6 Parallel RLC circuit

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 1.73]

Using KCL, the integro-differential equations can be written as follows:

$$i(t) = C\frac{dv(t)}{dt} + G\,v(t) + \frac{1}{L}\int v(t)\,dt$$

where, conductance, G=1/R.

On comparing with the mechanical rotational system equation, we get,

Rotational Mechanical System	*T-I* Analogous
Torque (T)	Current (i)
Angular velocity (ω)	Voltage (v)
Angular displacement (θ)	Flux (Φ)
Moment of inertia (J)	Capacitance (C)
Rotational spring constant (K)	1/Inductance (L)
Rotational damping (B)	Conductance (G)

1.5 ELECTRICAL ANALOGY OF THERMAL SYSTEMS

There are two fundamental physical elements that make up thermal networks, thermal resistances and thermal capacitance. There are also three sources of heat, a power source, a temperature source, and fluid flow.

Example:

In practice temperature when we discuss temperature, we will use degree Celsius (ºC), while SI unit for temperature is to use Kelvins (0ºK = - 273.15ºC). Generally reference temperature (*T_1*) is taken and all temperatures are measured relative to this reference. Reference temperature is assumed to be constant.

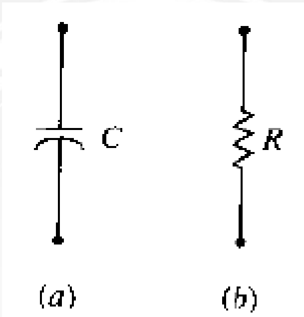


Figure 1.5.1 Network elements of thermal systems

[Source: "Linear Control System Analysis and Design" by John J. D'Azzo, Page: 76]

Thermal resistance

Consider the situation in which there is a wall, one side of which is at a temperature T_1, with the other side at temperature T_2, the wall has a thermal resistance of R_{12}.

Thermal capacitance

In addition to thermal resistance, objects can also have thermal capacitance (also called thermal mass). The thermal capacitance of an object is a measure of how much heat it can store. If an object has thermal capacitance its temperature will rise as heat flows into the object, and the temperature will lower as heat flows out. To understand this, envision a rock in the sun. During the day heat goes in to the rock from the sunlight, and the temperature of the rock increases as energy is stored in the rock as an increased temperature. At night energy is released, and the rock cools down. We represent a thermal capacitance in isolation in diagrams (and equations) as shown in Figure (in the drawing at the left the coil represents a power source and the stippled object is the thermal capacitance). In the thermal analogy, one end of the capacitor is always connected to the constant ambient temperature. The electrical model will always have one side of the capacitance connected to ground, or reference. Also, we could write the equation as = q but since T_1 is constant, it can be removed from the derivative. The thermal capacitance of an object is determined by its mass and specific heat.

$$C = mc_p$$

Where C is the thermal capacitance, m is the mass in kilograms, and c_p is the specific heat in J/(kg-ºK). it is always assumed that the capacitor is at a single uniform temperature, though this is obviously a simplification in many cases.

$$C\frac{dT_2}{dt} = q$$

Power source (or heat source)

A common part of a thermal model is a controlled power source that generates a predetermined amount of power, or heat, in a system. This power can either be constant or a function of time. In the electrical analogy, the power source is represented by a current source. An example of a power source is the quantity *q* in the diagrams for the thermal capacitance, above. In practice a power source is often an electrical heating element comprised of a coil of wire that is heated by a current flowing through it. Therefore, we use a diagram of a coil of wire to represent the power source. An ideal power source generates power that is independent of temperature.

Temperature source

An ideal temperature source maintains a given temperature independent of the amount of power required. Ambient temperature is considered to be reference temperature).

Mass Transfer (Fluid Flow)

If fluid with specific heat c_p J/kg-ºK) flows into a system with a flow rate of G kg/sec and a temperature of T_mºC above reference, and flows out at a temperature of T_{out}ºC below reference then the rate of heat flow into the system is given by

$$q_{in} = G\left\{\frac{kg}{sec}\right\}\cdot c_p\left\{\frac{J}{kg\ K}\right\}\cdot(T_{in}-T_{out})\{°C\} = Gc_p(T_{in}-T_{out})\{W\}$$

We can cancel the K and ºC since a temperature difference $(T_{in} – T_{out})$ is the same in Kelvin r Celsius. If you carefully observe this equation, it makes sense intuitively. Heat into s system goes up with mass flow rate into the system (increased mass flow, yields increased heat flow). Heat into a system also goes up with the specific heat of the mass (Higher specific heat indicates increased capacity to store heat). Finally, heat into system increases with an increased inflow temperature, or a decreased outflow temperature (if the temperature difference between inflow and outflow increases, more heat is being taken from the fluid). Note, the mass flow rate at the input and output must be equal to the mass (and thermal capacitance) of the system would be changing. This is not allowed for the systems being studied (time-invariant systems).

Energy balance

To develop a mathematical model of a thermal system we use the concept of an energy balance. The energy balance equation simply states that at any given location, or node, in a system, the heat into that node is equal to the heat out of the node plus any heat that is stored (heat is stored as increased temperature in thermal capacitances). The terms used in the equations is mentioned below:

Symbol	Quantity	U.S. customary units	Metric units
q	Rate of heat flow	Btu/minute	Joules/second
M	Mass	Pounds	Kilograms
S	Specific heat	Btu/(pounds)(°F)	Joules/(kilogram)(°C)
C	Thermal capacitance $C = MS$	Btu/°F	Joules/°C
K	Thermal conductance	Btu/(minute)(°F)	Joules/(second)(°C)
R	Thermal resistance	Degrees/ (Btu/minute)	Degrees/ (joule/second)
θ	Temperature	°F	°C
h	Heat energy	Btu	Joules

Additional heat stored in a body whose temperature is raised from θ_1 to θ_2 is given by

$$h = \frac{q}{D} = C(\theta_2 - \theta_1)$$

$$q = CD(\theta_2 - \theta_1)$$

Rate of heat flow through a body in terms of the two boundary temperatures θ_3 to θ_4

$$q = \frac{\theta_3 - \theta_4}{R}$$

The thermal resistance determines the rate of heat flow through the body. This is analogous to the resistance of a resistor in an electric circuit, which determines the current flow.

SIMPLE MERCURY THERMOMETER

Consider a thin glass-walled thermometer filled with mercury that has stabilized at a temperature θ_1. It is plunged into a bath of temperature θ_0 at t=0. In its simplest form, the thermometer can be considered to have a capacitance C that stores heat and a resistance R that limits the heat flow. The temperature at the center of the mercury is θ_m.The flow of heat into the thermometer is

$$q = \frac{\theta_0 - \theta_m}{R}$$

The heat entering the thermometer is stored in the thermal capacitance and is given by

$$h = \frac{q}{D} = C(\theta_m - \theta_1)$$

These equations can be combined to form

$$\frac{\theta_0 - \theta_m}{RD} = C(\theta_m - \theta_1)$$

Differentiating the above equation and rearranging the terms gives,

$$RC\, D\theta_m + \theta_m = \theta_0$$

The thermal network is drawn as in figure 1.5.2. Thus, the state equation is

$$\dot{x}_1 = -\frac{1}{RC}x_1 + \frac{1}{RC}u$$

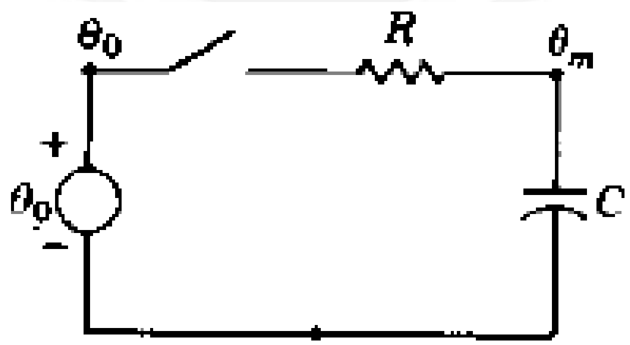


Figure 1.5.2 Network representation of a thermometer

[Source: "Linear Control System Analysis and Design" by John J. D'Azzo, Page: 77]

In general, *Heat in = Heat out + Heat stored*

$$q = \frac{T_r - T_a}{R_{ra}} + C\frac{dT_r}{dt}$$

1.6 TRANSFER FUNCTION

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

Open loop transfer function: G(s)

Loop transfer function: G(s)H(s)

Closed loop transfer function:

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

1.7 SERVOMOTOR

Servo Motor also called control motors, are used in feedback control systems as output actuators and does not use for continuous energy conversion. The principle of the Servomotor is similar to that of the other electromagnetic motor, but the construction and the operation are different. Their power rating varies from a fraction of a watt to a few hundred watts. Rotor inertia of the motors is low and have a high speed of response. The rotor of the Motor has the long length and smaller diameter. They operate at very low speed and sometimes even at the zero speed. Servo motor is widely used in radar and computers, robot, machine tool, tracking and guidance systems, processing controlling.

AC SERVOMOTORS

Servo motors are generally an assembly of four things: a DC motor, a gearing set, a control circuit and a position-sensor (usually a potentiometer). The position of servo motors can be controlled more precisely than those of standard DC motors, and they usually have three wires (power, ground & control). AC Servo Motors are divided into two types 2 and 3 Phase AC servomotor. Most of the AC servomotor are of the two-phase squirrel cage induction motor type. They are used for low power applications. The three phase squirrel cage induction motor is now utilized for the applications where high-power system is required. An AC servo motor is essentially a two-phase induction motor with modified constructional features to suit servo applications as shown in figure 1.7.1.

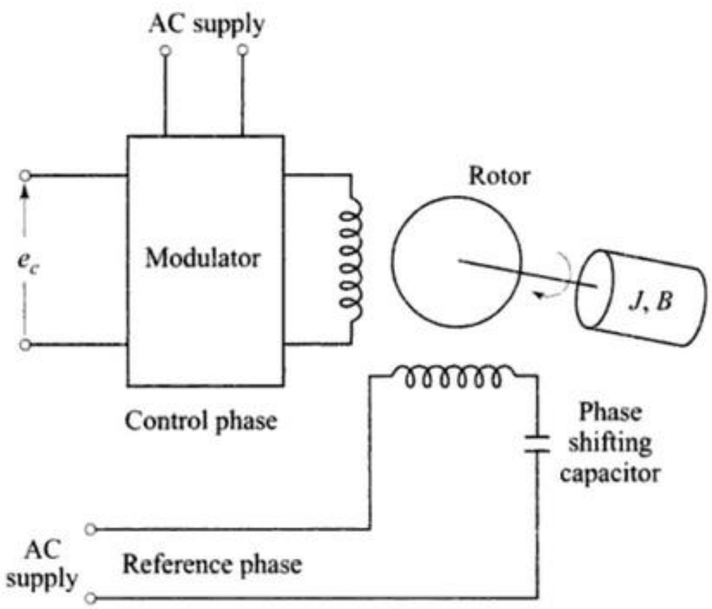


Figure 1.7.1 Symbolic representation of AC Servomotor

[Source: "Control Systems: Principles and Design" by M. Gopal, Page: 132]

It has two windings displaced by 90° on the stator. One winding, called as reference winding, is supplied with a constant sinusoidal voltage. The second winding, called control winding, is supplied with a variable control voltage which is displaced by -- 90° out of phase from the reference voltage.

The major differences between the normal induction motor and an AC servo motor are

1. The rotor winding of an ac servo motor has high resistance (R) compared to its inductive reactance (X) so that its X / *R* ratio is very low.
2. For a normal induction motor, X / *R* ratio is high so that the maximum torque is obtained in normal operating region which is around 5% of slip.

The torque speed characteristics of a normal induction motor and an ac servo motor are shown in figures 1.7.2 and 1.7.3.

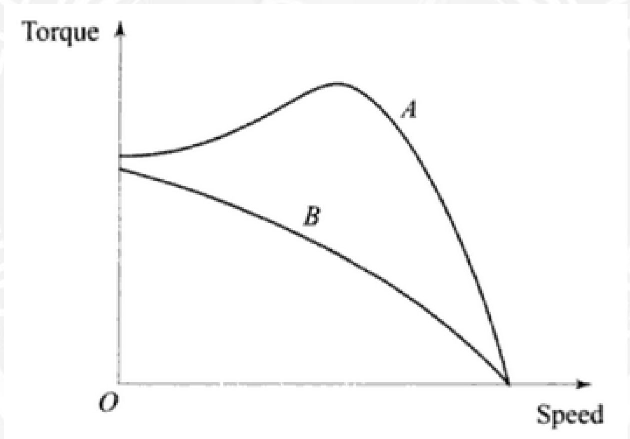


Figure 1.7.2 Torque speed characteristics of AC motors

[Source: "Control Systems: Principles and Design" by M. Gopal, Page: 131]

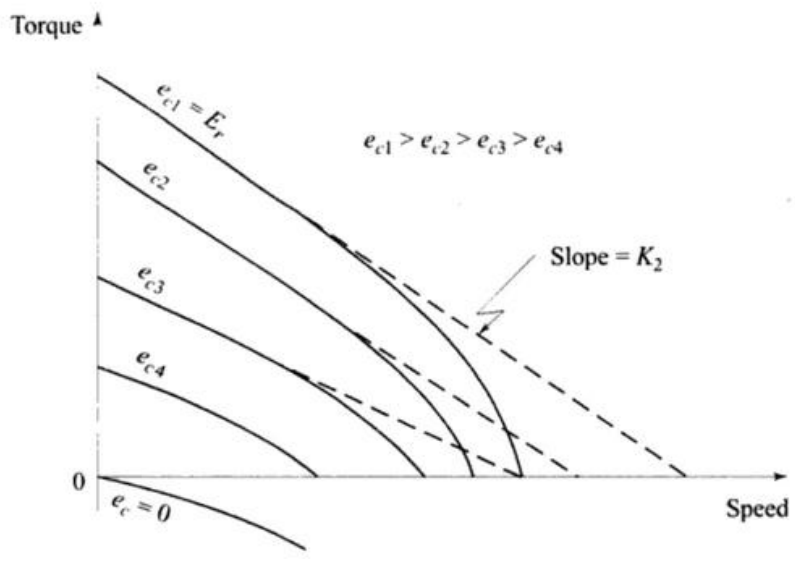


Figure 1.7.3 Torque speed characteristics of AC servomotor

[Source: "Control Systems: Principles and Design" by M. Gopal, Page: 133]

The torque-speed characteristic of a normal induction motor is highly nonlinear and has a positive slope for some portion of the curve. This is not desirable for control applications. as the positive slope makes the systems unstable. The torque speed characteristic of an ac servo motor is fairly linear and has negative slope throughout. The rotor construction is usually squirrel cage or drag cup type for an ac servo motor. The diameter is small compared to the length of the rotor which reduces inertia of the moving parts. Thus, it has good accelerating characteristic and good dynamic response. The supplies to the two windings of ac servo motor are not balanced as in the case of a normal induction motor. The control voltage varies both in magnitude and phase with respect to the constant reference vulture applied to the reference winding. The direction of rotation of the motor depends on the phase (± 90°) of the control voltage with respect to the reference voltage. For different rms values of control voltage the torque speed characteristics are shown in Figure. The torque varies approximately linearly with respect to speed and also controls voltage. The torque speed characteristics can be linearized at the operating point and the transfer function of the motor can be obtained.

DC SERVOMOTOR

A DC servo motor is used as an actuator to drive a load. It is usually a DC motor of low power rating. DC servo motors have a high ratio of starting torque to inertia and therefore they have a faster dynamic response. DC motors are constructed using rare earth permanent magnets which have high residual flux density and high coercively. As no field winding is used, the field copper losses am zero and hence, the overall efficiency of the motor is high. The speed torque characteristic of this motor is flat over a wide range, as the armature reaction is negligible. Moreover, speed in directly proportional to the armature voltage for a given torque. Armature of a DC servo motor is specially designed to have low inertia. DC Servo Motors are separately excited DC motor or permanent magnet DC motors. The figure (a) shows the connection of Separately Excited DC Servo motor and the figure (b) shows the armature MMF and the excitation field MMF in quadrature in a DC machine. This provides a fast torque response because torque and flux are decoupled. Therefore, a small change in the armature voltage or current brings a significant shift in the position or speed of the rotor. Most of the high-power servo motors are mainly DC.

(a) Armature controlled DC servo motor

The physical model of an armature controlled DC servo motor is given in

The armature winding has a resistance Ra and inductance La.

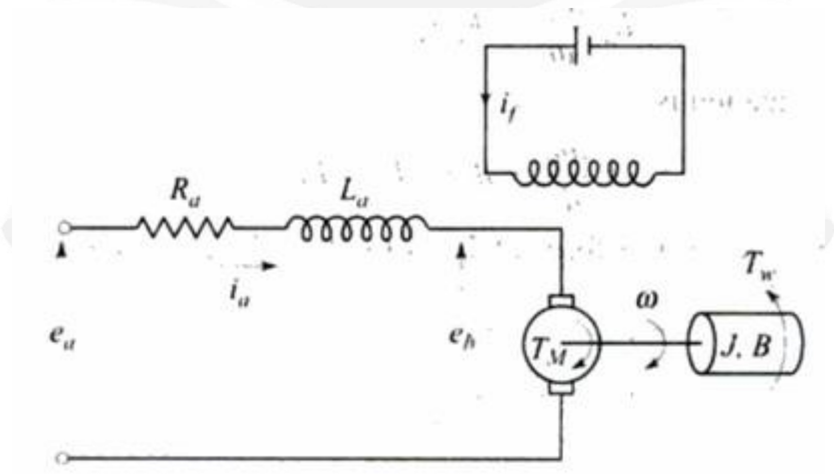


Figure 1.7.4 Armature controlled DC motor with load

[Source: "Control Systems: Principles and Design" by M. Gopal, Page: 117]

The field is produced either by a permanent magnet or the field winding is separately excited and supplied with constant voltage so that the field current, i_f is a constant. When the armature is supplied with a DC voltage of e_a volts, the armature rotates and produces a back emf, e_b. The armature current i_a depends on the difference of e_b and e_n. The armature has a permanent of inertia J, frictional coefficient B_0. The angular displacement of the motor is θ. The torque produced by the motor is given by

$$T_M = K_T i_a$$

$$e_b = K_b \omega$$

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

$$J \frac{d\omega}{dt} + B\omega + T_w = T_M$$

Taking Laplace transform,

$$T_M(s) = K_T I_a(s)$$

$$E_b(s) = K_b \omega(s)$$

$$L_a s I_a(s) + R_a I_a(s) + E_b(s) = E_a(s)$$

$$J s \omega(s) + B\omega(s) + T_w(s) = T_M(s)$$

where K_T is the motor torque constant. The back emf is proportional to the speed of the motor and hence

On solving,

$$\frac{\omega(s)}{E_a(s)} = \frac{K_T/R_a}{Js + B + K_T K_b/R_a}$$

$$\frac{\omega(s)}{E_a(s)} = \frac{K_m}{\tau_m s + 1}$$

where,

$$K_m = \frac{K_T}{R_a B + K_T K_b}$$

$$\tau_m = \frac{R_a J}{R_a B + K_T K_b}$$

K_m – motor gain constant, τ_m – motor time constant

(b) Field controlled DC servo motor

The schematic diagram of a field controlled DC servo motor is shown in figure 1.7.5.

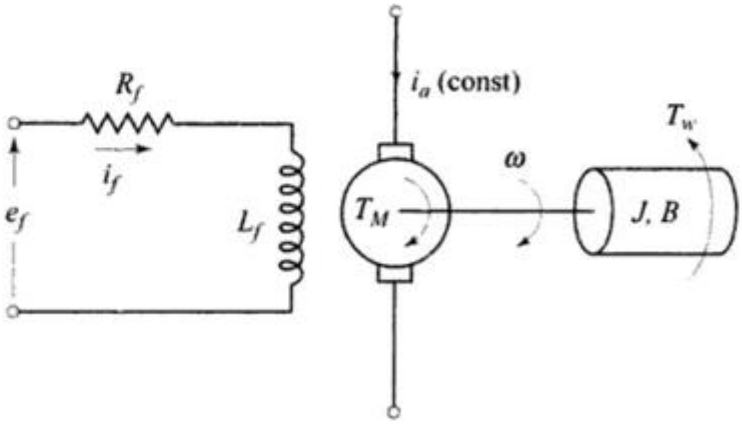


Figure 1.7.5 Field controlled DC servomotor

[Source: "Control Systems: Principles and Design" by M. Gopal, Page: 120]

$$T = K_{Tf} i_f$$

$$L_f \frac{di_f}{dt} + R_f i_f = e_f$$

$$J \frac{d\omega}{dt} + B\omega + T_w = T_M$$

Taking Laplace transform,

$$T(s) = K_{Tf} I_f(s)$$

$$L_f s I_f(s) + R_f I_f(s) = E_f(s)$$

$$Js^2\theta(s) + Bs\theta(s) + T_w(s) = T_M(s)$$

$$\frac{\theta(s)}{E_f(s)} = \frac{K_{Tf}}{s(Js+B)(R_f + sL_f)} = \frac{K_{Tf}/R_f B}{s(\frac{J}{B}s+1)(1+s\frac{L_f}{R_f})} = \frac{K_m}{s(\tau_m s+1)(1+s\tau_f)}$$

where, motor gain constant, $K_m = K_{Tf}/R_f B$

motor time constant, $\tau_m = \frac{J}{B}$

field time constant, $\tau_f = \frac{L_f}{R_f}$

1.8 BLOCK DIAGRAM REDUCTION TECHNIQUES

A system that can change its output in accordance with change in input is known as a closed loop system. This can be implemented by introducing a feedback path in an open-loop system and manipulating the input that is applied to the system. Such as closed-loop system can be represented by using a block diagram shown in figure 1.8.1.

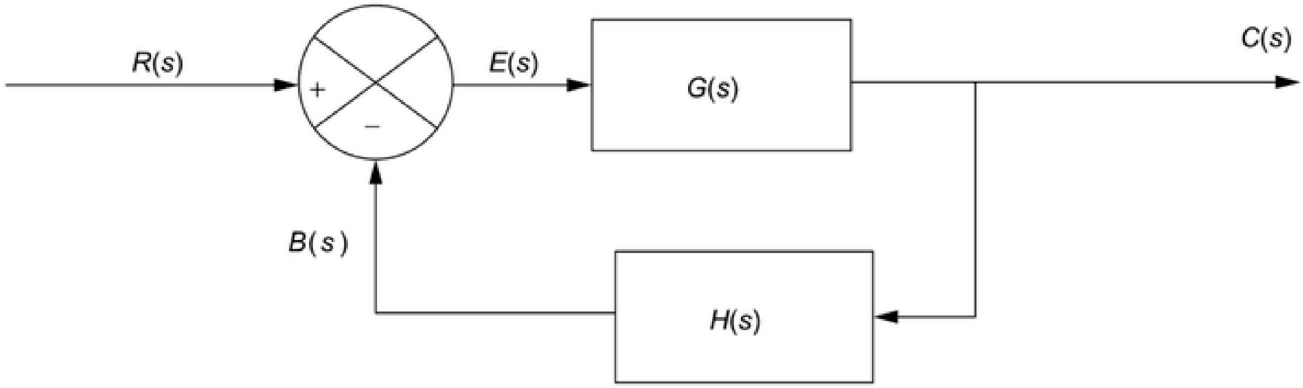


Figure 1.8.1 Simple block diagram representation

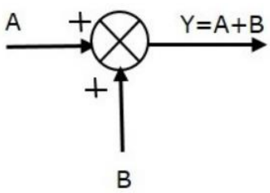
[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 3.2]

BLOCK

The transfer function of a component is represented by a block. Block has single input and single output.

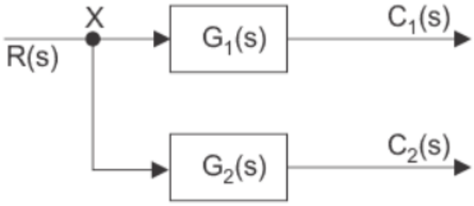
SUMMING POINT

The summing point is represented with a circle having cross (X) inside it. It has two or more inputs and single output. It produces the algebraic sum of the inputs. It also performs the summation or subtraction or combination of summation and subtraction of the inputs based on the polarity of the inputs. Let us see these three operations one by one. The following figure shows the summing point with two inputs (A, B) and one output (Y). Here, the inputs A and B have a positive sign. So, the summing point produces the output, Y as sum of A and B.



NODE

The node is a point from which the same input signal can be passed through more than one branch. That means with the help of node, we can apply the same input to one or more blocks, summing points. In the following figure, the node is used to connect the same input, R(s) to two more blocks.



The advantages of block diagram representation are:

(i) It facilitates easier representation of complex systems.

(ii) Calculation of transfer function by block diagram reduction techniques is easy.

(iii) Performance analysis of a complex system is simplified by determining its transfer function.

(iv) It facilitates easier access of individual elements in a system that is represented by a block diagram.

(v) It facilitates visualization of operation of the whole system by the flow of signals.

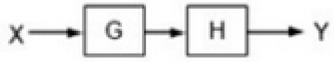
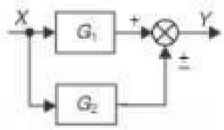
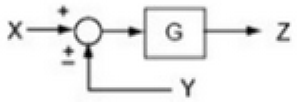
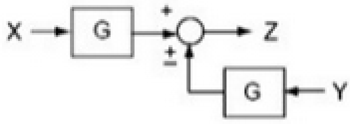
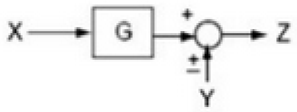
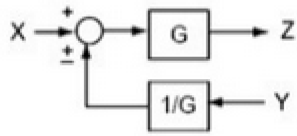
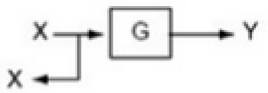
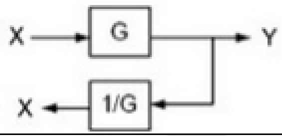
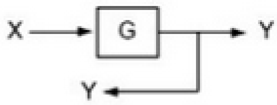
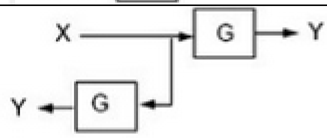
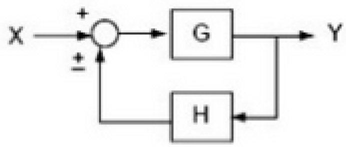
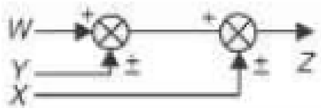
The disadvantages of block diagram representation are:

(i) It is difficult to determine the actual composition of individual elements in a system.

(ii) Representation of a system using block diagram is not unique.

(iii) The main source of signal flow cannot be represented definitely in a block diagram.

RULES FOR BLOCK DIAGRAM REDUCTION

Rule No.	Rule	Block diagram	Equivalent block diagram
1	Blocks in cascade		
2	Blocks in parallel		
3	Moving a summing point behind the block		
4	Moving a summing point ahead of the block		
5	Moving a branch point behind the block		
6	Moving a branch point ahead of the block		
7	Eliminating a feedback loop		
8	Interchanging the summing point		

1.9 SIGNAL FLOW GRAPH

The diagrammatic or pictorial representation of a set of simultaneous linear algebraic equations of a more complicated system is known as signal flow graph (SFG). It shows the flow of signals in the system. It is important to note that the flow of signals in SFG is only in one direction. To represent the set of algebraic equations using SFG, it is necessary that those algebraic equations are to be represented in the s-domain. The transfer function of the system which is represented by SFG can be obtained by using Mason's gain formula. The dependent and independent variables in the set of algebraic equations are represented by the nodes in the SFG. The branches are used to connect different nodes present in SFG. The connection between the different nodes is based on the relationship given in the algebraic equation. The arrow and the multiplication factor indicated on the branch of SFG represent the signal direction. The SFG and the block diagram representation of a system yield the same transfer function; but when a system is represented by SFG, the transfer function is obtained easily and quickly without using the SFG reduction techniques. The terminologies used in SFG rae explained with the help of SFG of a system as shown in figure 1.9.1.

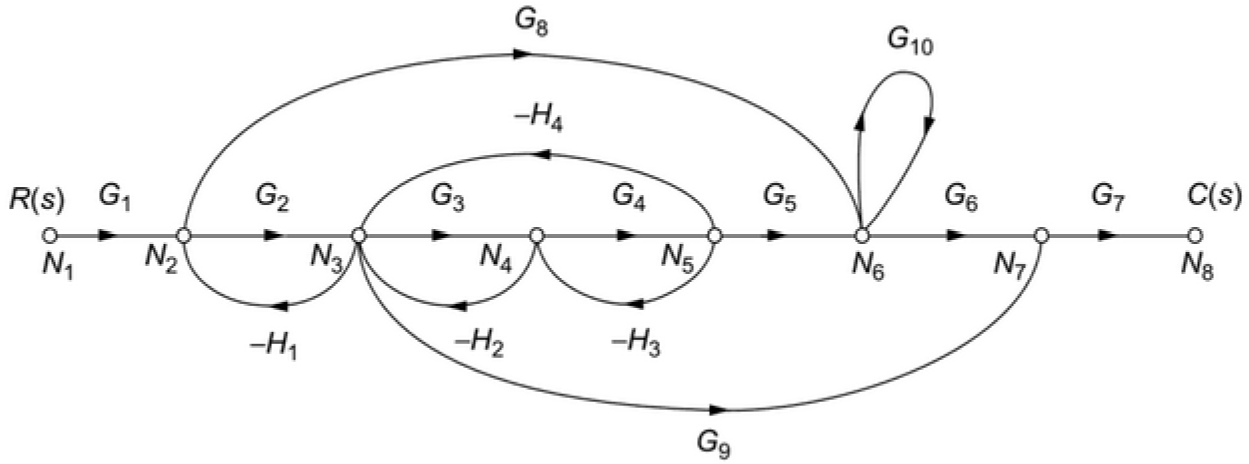


Figure 1.9.1 Signal flow graph of a system

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 4.1]

Node: The variables present in the set of algebraic equations are represented by a point called node.

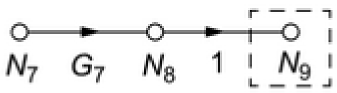


Figure 1.9.2 Node in signal flow graph

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 4.2]

Branch

The line segment joining the two nodes with a specific direction is known as a branch. The specific direction is indicated by an arrow in the branch.

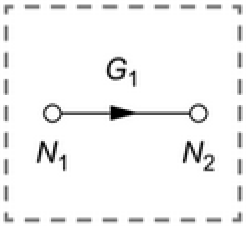


Figure 1.9.3 Branch in signal flow graph

[Source: "Control Systems Engineering" by S.Salivahanan, R.Rengaraj, G.R.Venkatakrishnan, Page: 4.2]

MASON'S GAIN FORMULA

A technique to reduce a signal flow graph to a single transfer function requires the application of one formula. The transfer function of a system represented by a signal flow graph is

$$T = \frac{C(s)}{R(s)} = \frac{\sum_{i=1}^{n} P_i \Delta_i}{\Delta}$$

where, k – number of forward path

P_i – ith forward path gain

Δ – 1- (sum of individual loop gains)+(sum of product of two non-touching loop gains)-(sum of product of three non-touching loop gains)+……

Δ_i – 1- (Δ of the loop non-touching the ith forward path)

Steps to determine the transfer function of a system using SFG Method

Step 1: Identify the number of forward paths.

Step 2: Identify the individual loops and find their respective loop gains.

Step 3: Identify the two non-touching loops and find the product of their gains.

Step 4: Identify the three non-touching loops and find the gain product and so on…

Step 5: Calculate the Δ value.

Step 6: Calculate the Δ_i value.

Step 7: Use Mason's gain formula to calculate the transfer function value, T.

Characteristics	Block Diagram	Signal flow graph
Time Consumption	More since the diagrams have to be redrawn repeatedly	Less since there is no necessary to redraw the diagrams
Technique applied	Block Diagram reduction technique	Mason's gain formula
Representation of elements	Blocks are used to represent the element.	Nodes are need to represent the elements
Representation of transfer function of each element	Represented inside the block of each element	Represented along the branches above the arrow ahead
Feedback paths	Present and hence the formula, (G/(1±GH)) is used to reduce the paths	Present, but there is no need for any formulae to reduce the paths
Self-loops	Absence of self-loops	Presence of self-loops
Summing points and takeoff points	Present in block diagram	Absence in SFG

2.1 TIME RESPONSE

The time response of the system is the output of the closed loop system as a function of time. It describes the behavior of a system and contains much information about it with respect to time response specification. Time response is formed by the transient response and the steady state response.

Time response = Transient response + Steady state response

Transient time response

Transient response (Natural response) describes the behavior of the system in its first short time until arrives the steady state value and this response will be our study focus. If the input is step function then the output or the response is called step time response and if the input is ramp, the response is called ramp time response, etc.

$$\mathbf{y(t) = y_{tr}(t) + y_{ss}(t)}$$

The transient response is defined as the part of the time response that goes to zero as time becomes very large. Thus $y_t(t)$ has the property

$$\mathbf{Lim\ y_t(t) = 0, \quad t \to \infty}$$

The time required to achieve the final value is called transient period. The transient response may be exponential or oscillatory in nature. Output response consists of the sum of forced response (form the input) and natural response (from the nature of the system).The transient response is the change in output response from the beginning of the response to the final state of the response and the steady state response is the output response as time is approaching infinity (or no more changes at the output).

Steady State Response

The steady state response is the part of the total response that remains after the transient has died out. For a position control system, the steady state response when compared to with the desired reference position gives an indication of the final accuracy of the system. If the steady state response of the output does not agree with the desired reference exactly, the system is said to have steady state error.

2.2 TIME DOMAIN SPECIFICATIONS

The desired performance characteristics of control systems are specified in terms of time domain specifications. Systems with energy storage elements cannot respond instantaneously and will exhibit transient responses, whenever they are subjected to inputs or disturbances. The desired performance characteristics of a system of any order may be specified in terms of the transient response to a unit step input signal. The response of a second order system for unit step input with various values of damping ratio is shown in figure 2.2.1.

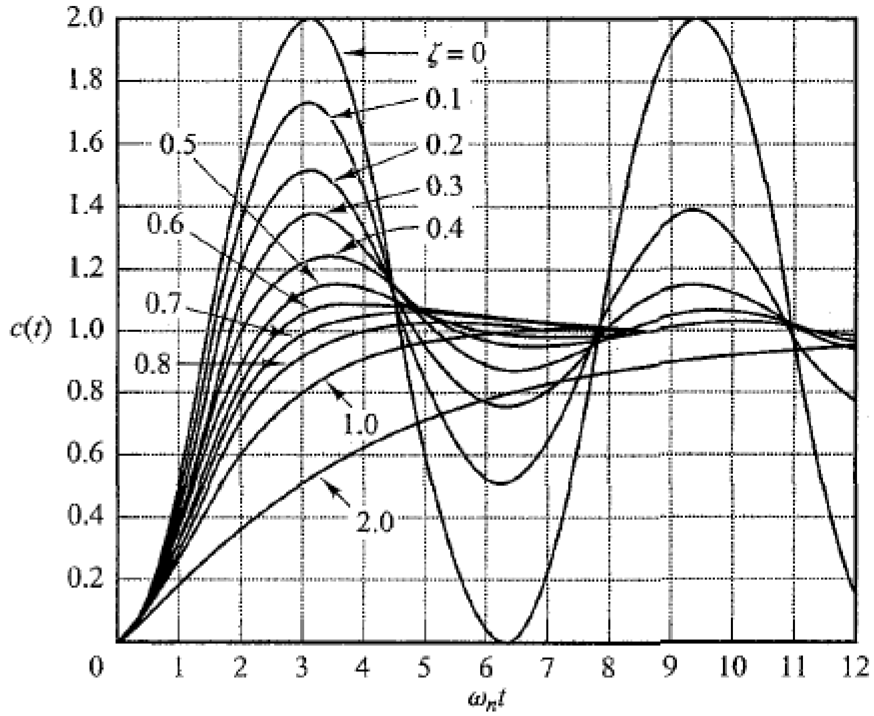


Figure 2.2.1 Time Response

[Source: "Modern Control Engineering" by Katsuhiko Ogata, Page: 229]

The transient response of a system to a unit step input depends on the initial conditions. Therefore, to compare the time response of various systems it is necessary to start with standard initial conditions. The most practical standard is to start with the system at rest and so output and all time derivatives before t=0 will be zero. The transient response of a practical control system often exhibits damped oscillation before reaching steady state. A typical damped oscillatory response of a system is shown in figure 2.2.2.

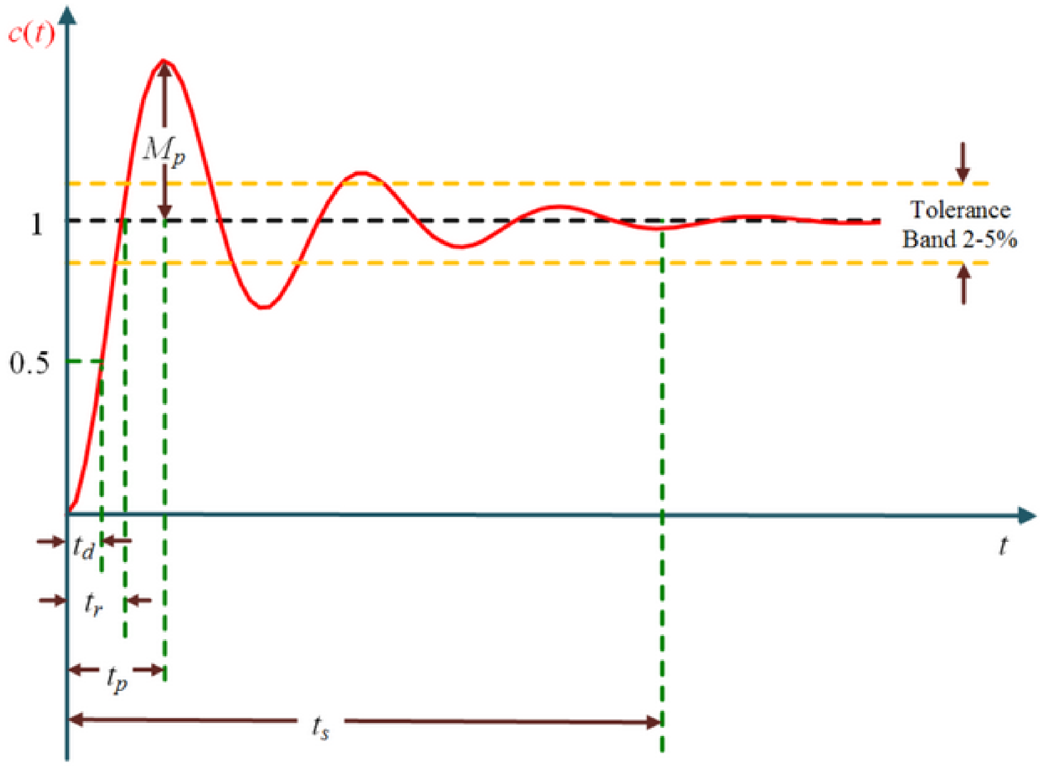


Figure 2.2.2 Transient and steady-state response analyses

[Source: "Modern Control Engineering" by Katsuhiko Ogata, Page: 230]

The transient response characteristics of a control system to a unit step input is specified in terms of the following time domain specifications:

1. Delay time, t_d: It is the time required for the response to reach 50% of the steady state value for the first time.

$$t_d = \frac{1 + 0.7\zeta}{\omega_n}$$

2. Rise time, t_r: It is the time required for the response to reach 100% of the steady state value for under damped systems. However, for over damped systems, it is taken as the time required for the response to rise from 10% to 90% of the steady state value. The unit step response of second order system for underdamped case is given by,

$$c(t) = 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{(1-\zeta^2)}}\sin(\omega_d t + \theta)$$

At $t = t_r$, $c(t) = c(t_r) = 1$

$$c(t_r) = 1 - \frac{e^{-\zeta\omega_n t_r}}{\sqrt{(1-\zeta^2)}}\sin(\omega_d t_r + \theta) = 1$$

$$\frac{-e^{-\zeta\omega_n t_r}}{\sqrt{(1-\zeta^2)}}\sin(\omega_d t_r+\theta)=0$$

Since $-e^{-\zeta\omega_n t_r}\neq 0$, the term, $\sin(\omega_d t_r+\theta)=0$,

When Φ = 0, π, 2π, 3π,…. sin Φ = 0

$$\omega_d t_r+\theta=\pi$$

$$\omega_d t_r=\pi-\theta$$

$$t_r=\frac{\pi-\theta}{\omega_d}$$

On constructing right angled triangle,

$$\tan\theta=\frac{\sqrt{(1-\zeta^2)}}{\zeta}$$

$$\theta=\tan^{-1}\frac{\sqrt{(1-\zeta^2)}}{\zeta}$$

$$Damped\ frequency,\ \ \omega_d=\omega_n\sqrt{(1-\zeta^2)}$$

$$t_r=\frac{\pi-\tan^{-1}\left(\frac{\sqrt{(1-\zeta^2)}}{\zeta}\right)}{\omega_n\sqrt{(1-\zeta^2)}}$$

3. Peak time, t_p: It is the time required for the response to reach the maximum or peak value of the response. To find the expression for peak time, t_p, differentiate c(t) with respect to 't' and equate to zero.

$$\frac{d}{dt}c(t)|_{t=t_p}=0$$

The unit step response of under damped second order system is given by

$$c(t)=1-\frac{e^{-\zeta\omega_n t}}{\sqrt{(1-\zeta^2)}}\sin(\omega_d t+\theta)$$

Differentiating c(t) with respect to 't',

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

Put $\omega_d=\omega_n\sqrt{(1-\zeta^2)}$,

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

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

$$= \frac{\omega_n e^{-\zeta\omega_n t}}{\sqrt{(1-\zeta^2)}}[\cos\theta \sin(\omega_d t + \theta) - \sin\theta \cos(\omega_d t + \theta)]$$

$$= \frac{\omega_n e^{-\zeta\omega_n t}}{\sqrt{(1-\zeta^2)}}[\sin(\omega_d t + \theta - \theta)]$$

$$= \frac{\omega_n e^{-\zeta\omega_n t}}{\sqrt{(1-\zeta^2)}}[\sin(\omega_d t)]$$

At t = t_p, $\frac{d}{dt}c(t) = 0$

$$\frac{\omega_n e^{-\zeta\omega_n t_p}}{\sqrt{(1-\zeta^2)}}\left[\sin(\omega_d t_p)\right] = 0$$

Since, $e^{-\zeta\omega_n t_p} \neq 0$, the term, $\left[\sin(\omega_d t_p)\right] = 0$

When Φ = 0, π, 2π, 3π,…. sin Φ = 0

$$\omega_d t_p = \pi$$

$$t_p = \frac{\pi}{\omega_d}$$

On substituting, we get,

$$t_p = \frac{\pi}{\omega_n\sqrt{(1-\zeta^2)}}$$

4. Peak overshoot, M_p: It is defined as the difference between the peak value of the response and the steady state value. Iris usually expressed in percent of the steady state value. If the time for the peak is t_p, percent peak overshoot is given by,

Maximum percent overshoot = $\frac{c(t_p)-c(\infty)}{c(\infty)}$

$$At\ t = \infty, c(t) = c(\infty) = 1 - \frac{e^{-\zeta\omega_n\infty}}{\sqrt{(1-\zeta^2)}}\sin(\omega_d t + \theta) = 1 - 0 = 1$$

$$At\ t = t_p, c(t) = c(t_p) = 1 - \frac{e^{-\zeta\omega_n t_p}}{\sqrt{(1-\zeta^2)}}\sin\left(\omega_d t_p + \theta\right)$$

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

$$= 1 - \frac{e^{-\zeta\frac{\pi}{\sqrt{(1-\zeta^2)}}}}{\sqrt{(1-\zeta^2)}}\sin(\pi + \theta)$$

$$= 1 - \frac{e^{-\zeta \frac{\pi}{\sqrt{(1-\zeta^2)}}}}{\sqrt{(1-\zeta^2)}} \sin(\theta) = 1 + \frac{e^{\frac{\zeta\pi}{\sqrt{(1-\zeta^2)}}}}{\sqrt{(1-\zeta^2)}} \sqrt{(1-\zeta^2)}$$

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

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

5. Settling time, t_s: It is the time taken by the response to reach and stay within a specified error. It is usually expressed as percentage of final value. The usual tolerable error is 2% and 5% of the final value.

The response of second order system has two components. They are

a. Decaying exponential component, $\frac{e^{-\zeta\omega_n t}}{\sqrt{(1-\zeta^2)}}$

b. Sinusoidal component, $\sin(\omega_d t + \theta)$

In these terms, the decaying component term dampens or reduces the oscillations produced by sinusoidal component. Hence, the settling time is decided by the exponential component. The settling time can be found out by equating exponential component to percentage tolerance errors.

For 2% tolerance error band, at t = t_s, $\frac{e^{-\zeta\omega_n t_s}}{\sqrt{(1-\zeta^2)}} = 0.02$

For least values of ζ, $e^{-\zeta\omega_n t_s} = 0.02$

On taking natural logarithm on both sides, we get,

$$-\zeta\omega_n t_s = \ln(0.02) = -4$$

$$t_s = \frac{4}{\zeta\omega_n} = 4T$$

For 5% tolerance error band, at t = t_s, $\frac{e^{-\zeta\omega_n t_s}}{\sqrt{(1-\zeta^2)}} = 0.05$

For least values of ζ, $e^{-\zeta\omega_n t_s} = 0.05$

On taking natural logarithm on both sides, we get,

$$-\zeta\omega_n t_s = \ln(0.02) = -3$$

$$t_s = \frac{3}{\zeta\omega_n} = 3T$$

$$Settling\ time, t_s = \frac{4}{\zeta\omega_n}\ for\ 2\%\ error$$

$$Settling\ time, t_s = \frac{3}{\zeta\omega_n}\ for\ 5\%\ error$$

The performance of a system is usually evaluated in terms of the following qualities:

- How fast it is able to respond to the input?
- How fast it is reaching the desired output?
- What is the error between the desired output and the actual output, once the transients die down and steady state is achieved?
- Does it oscillate around the desired value?
- Is the output continuously increasing with time or is it bounded?
- These are the specifications to be given for the design of a controller for a given system.

2.3 TYPES OF TEST INPUT

The knowledge of input signal is required to predict the response of a system. In most of the systems, the input signals are not known ahead of time and also it is difficult to express the input signals mathematically by simple equations. The characteristics of actual input signals are a sudden shock, a sudden change, a constant velocity and a constant acceleration. Hence test signals which resembles these characteristics are used as input signals to predict the performance of the system. The commonly use test input signals are impulse, step, ramp, acceleration and sinusoidal signals.

Standard Input Signals

1. Step signal
2. Unit step signal
3. Ramp signal
4. Unit ramp signal
5. Parabolic signal
6. Unit parabolic signal
7. Impulse signal
8. Sinusoidal signal

STEP SIGNAL

The step signal is a signal whose value changes from zero to A at t=0 and remains constant at A for t > 0. The step signal resembles an actual steady input to a system. A special case of step signal is unit step in which A is unity.

RAMP SIGNAL

The ramp signal is a signal whose value increases linearly with time from an initial value of zero at t=0. The ramp signal resembles a constant velocity input to the system. A special case of ramp signal is unit ramp signal in which the value of A is unity.

PARABOLIC SIGNAL

In parabolic signal, the instantaneous value varies as square of the time from an initial value of zero at t=0. The sketch of the signal with respect to time resembles a parabola. The parabolic signal resembles a constant acceleration input to the system. A special case of parabolic signal is unit parabolic signal in which A is unity.

IMPULSE SIGNAL

A signal of very large magnitude which is available for very short duration is called impulse signal. Ideal impulse signal is a signal with infinite magnitude and zero duration but with an area of A. The unit impulse signal is a special case, in which A is unity. Since perfect impulse cannot be achieved in practice, it is usually approximated by a pulse of small width but with area, A. Mathematically an impulse signal is the derivative of a step signal. Laplace transform of the impulse function is unity.

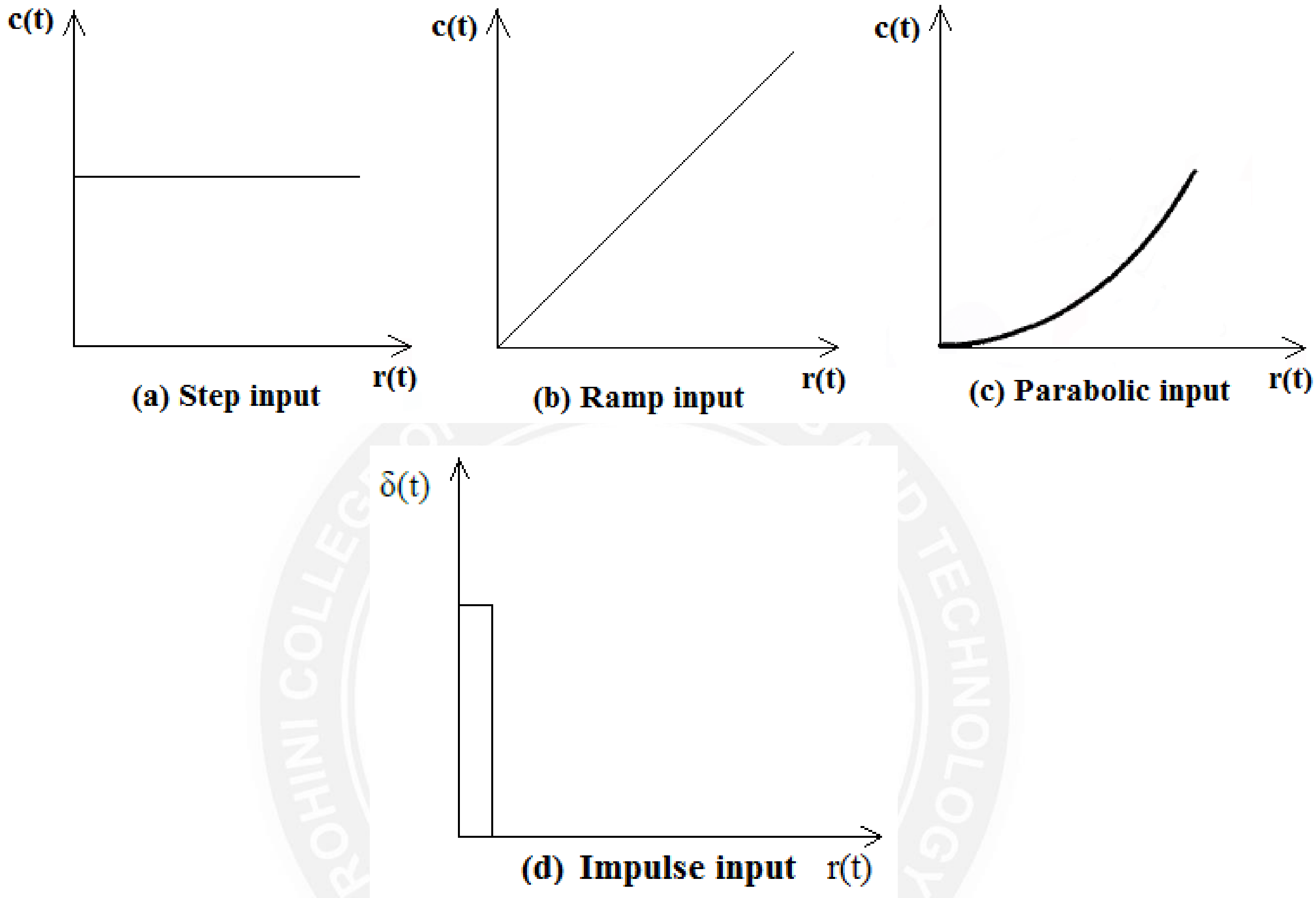
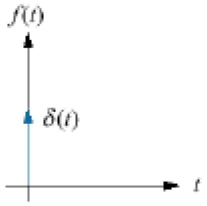
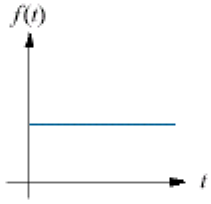
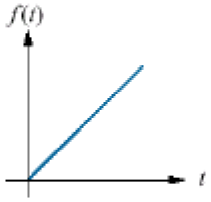
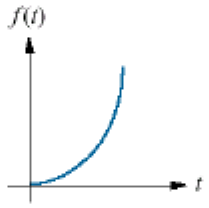
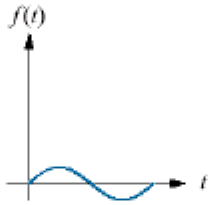


Figure 2.3.1 Standard test signals

[Source: "Control Systems Engineering" by I J Nagrath, M Gopal, Page: 196]

Input	r(t)	R(s)
Step input	A	A/s
Ramp input	At	A/s^2
Parabolic input	$At^2/2$	A/s^3
Impulse input	δ(t)	1

Input	Function	Description	Sketch	Use
Impulse	$\delta(t)$	$\delta(t) = \infty$ for $0- < t < 0+$ $= 0$ elsewhere $\int_{0-}^{0+} \delta(t)\,dt = 1$		Transient response Modeling
Step	$u(t)$	$u(t) = 1$ for $t > 0$ $= 0$ for $t < 0$		Transient response Steady-state error
Ramp	$tu(t)$	$tu(t) = t$ for $t \geq 0$ $= 0$ elsewhere		Steady-state error
Parabola	$\frac{1}{2}t^2u(t)$	$\frac{1}{2}t^2u(t) = \frac{1}{2}t^2$ for $t \geq 0$ $= 0$ elsewhere		Steady-state error
Sinusoid	$\sin \omega t$			Transient response Modeling Steady-state error

2.4 FIRST AND SECOND ORDER SYSTEM RESPONSE

Transfer Function

- It is the ratio of Laplace transform of output to Laplace transform of input with zero initial conditions.
- One of the types of modeling a system
- Using first principle, differential equation is obtained
- Laplace Transform is applied to the equation assuming zero initial conditions

Order of a system

- ✓ Order of a system is given by the order of the differential equation governing the system
- ✓ Alternatively, order can be obtained from the transfer function
- ✓ In the transfer function, the maximum power of s in the denominator polynomial gives the order of the system

Dynamic Order of Systems

- Order of the system is the order of the differential equation that governs the dynamic behaviour
- Working interpretation: Number of the dynamic elements / capacitances or holdup elements between a manipulated variable and a controlled variable
- Higher order system responses are usually very difficult to resolve from one another
- The response generally becomes sluggish as the order increases

SYSTEM RESPONSE

First-order system time response

- Transient
- Steady-state

Second-order system time response

- Transient
- Steady-state

FIRST ORDER SYSTEM

Response of First Order System for Unit Step Input

The standard form of closed loop transfer function of first order system is

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

If the input is unit step, then r(t) and R(s)=1/s

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

Applying partial fraction expansion,

$$C(s) = \frac{A}{s} + \frac{B}{1+sT}$$

On solving,

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

On taking inverse Laplace transform, the response in time domain is obtained as,

$$c(t) = 1 - e^{-\frac{t}{T}}$$

Hence, the input and output signal of the first order system is given by,

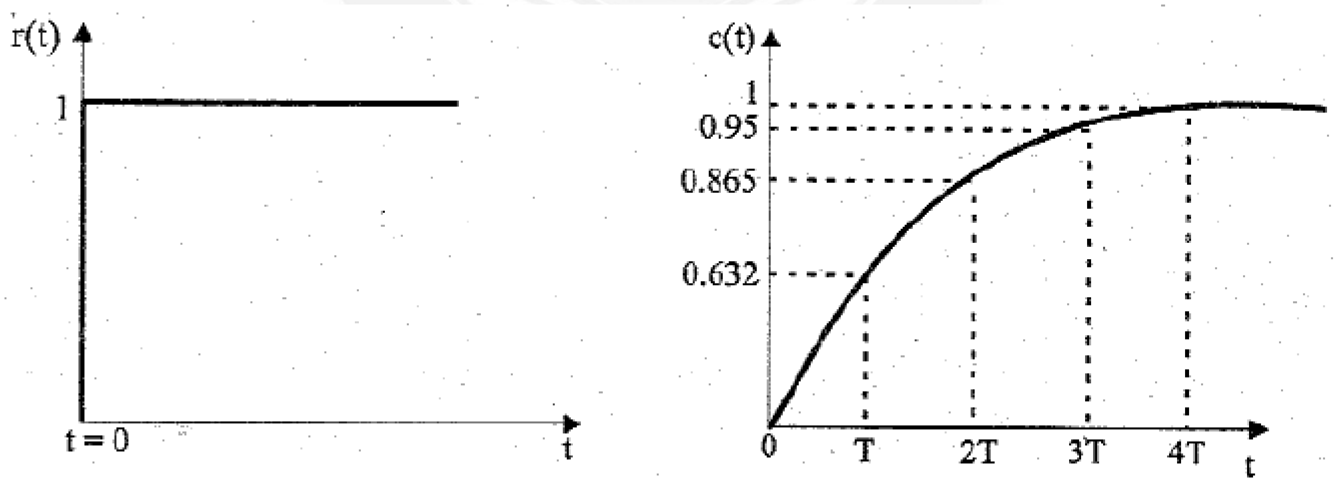


Figure 2.4.1 Response of first order system to unit step input

[Source: "Control Systems" by Nagoor Kani, Page: 2.20]

SECOND ORDER SYSTEM

LTI second-order system

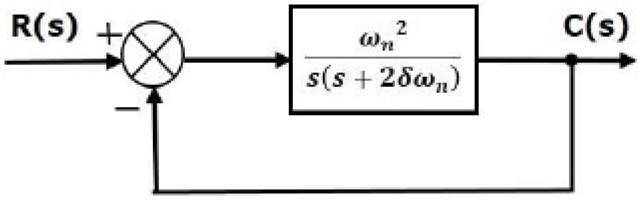


Figure 2.4.2 Closed loop for second order system

[Source: "Control Systems" by Nagoor Kani, Page: 2.20]

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

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

where, ζ is the damping ratio, ω_n is the natural frequency

DAMPING RATIO

It is the ratio of critical damping to actual damping.

CHARACTERISTIC EQUATION

$$s^2+2\zeta\omega_n s+\omega_n^2 = 0$$

$$s = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2-1}$$

The roots of characteristic equation are:

- The two roots are imaginary when $\zeta = 0$ (undamped system)
- The two roots are real and equal when $\zeta = 1$ (critically damped system)
- The two roots are real but not equal when $\zeta > 1$ (overdamped system)
- The two roots are complex conjugate when $0 < \zeta < 1$ (underdamped system)

Response of Second Order System for Unit Step Input

Consider the unit step signal as an input to the second order system. Laplace transform of the unit step signal is

R(s) = 1/s

Transfer function of the second order closed loop transfer function is

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

Case 1: Undamped system

When $\zeta = 0$,

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

For unit step input, R(s) = 1/s,

$$C(s) = \frac{\omega_n^2}{s^2 + \omega_n^2}\left(\frac{1}{s}\right) = \frac{\omega_n^2}{s(s^2 + \omega_n^2)}$$

Taking inverse Laplace transform,

$$c(t) = 1 - \cos\omega_n t$$

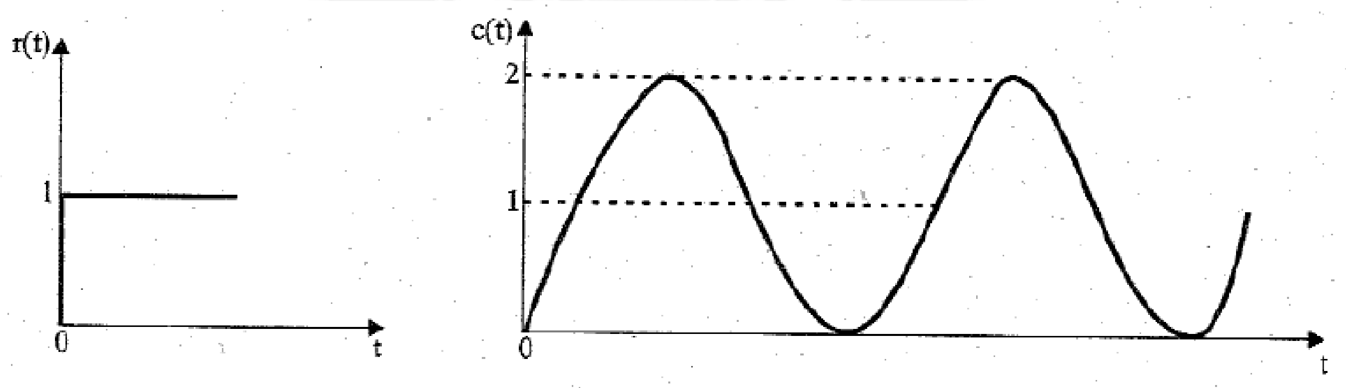


Figure 2.4.3 Response of undamped second order system to unit step input

[Source: "Control Systems" by Nagoor Kani, Page: 2.22]

Case 2: Underdamped system

When $0 < \zeta < 1$,

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

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

$$= (s + \zeta\omega_n)^2 + \omega_n^2(1 - \zeta^2)$$

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

For unit step input, R(s)=1/s,

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

By applying partial fraction,

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

On solving, we get,

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

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

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

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

$$- \frac{\zeta}{\sqrt{1 - \zeta^2}} \frac{\omega_n\sqrt{1 - \zeta^2}}{\left((s + \zeta\omega_n)^2 + \left(\omega_n\sqrt{1 - \zeta^2}\right)^2\right)}$$

On taking inverse Laplace transform,

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

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

We know, $\sin\theta = \sqrt{1-\zeta^2}, \cos\theta = \zeta$

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

$$c(t) = \left(1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}(\sin(\omega_d t + \theta))\right)$$

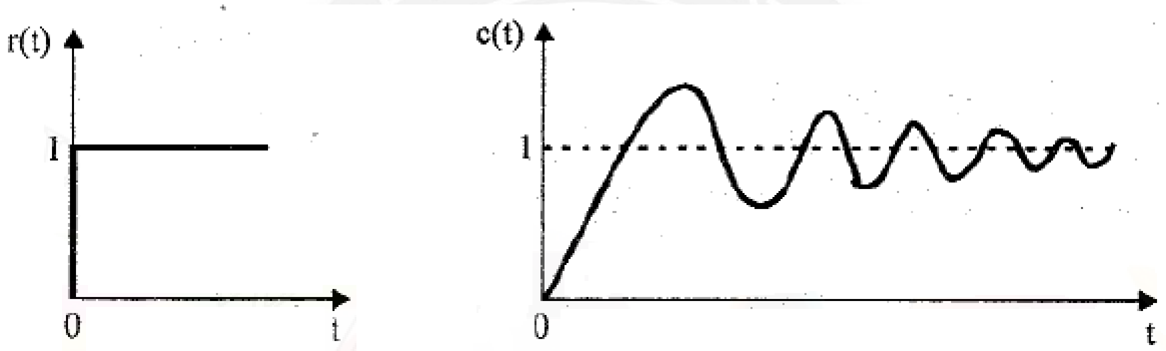


Figure 2.4.4 Response of underdamped second order system to unit step input

[Source: "Control Systems" by Nagoor Kani, Page: 2.24]

Case 3: Critically damped system

When $\zeta = 1$,

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

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

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

For a step input, R(s)=1/s

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

By applying partial fractions,

$$C(s) = \frac{A}{s} + \frac{B}{s + \omega_n} + \frac{C}{(s + \omega_n)^2}$$

On solving, we get

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

By taking inverse Laplace transform,

$$c(t) = 1 - e^{-\omega_n t} - \omega_n t e^{-\omega_n t}$$

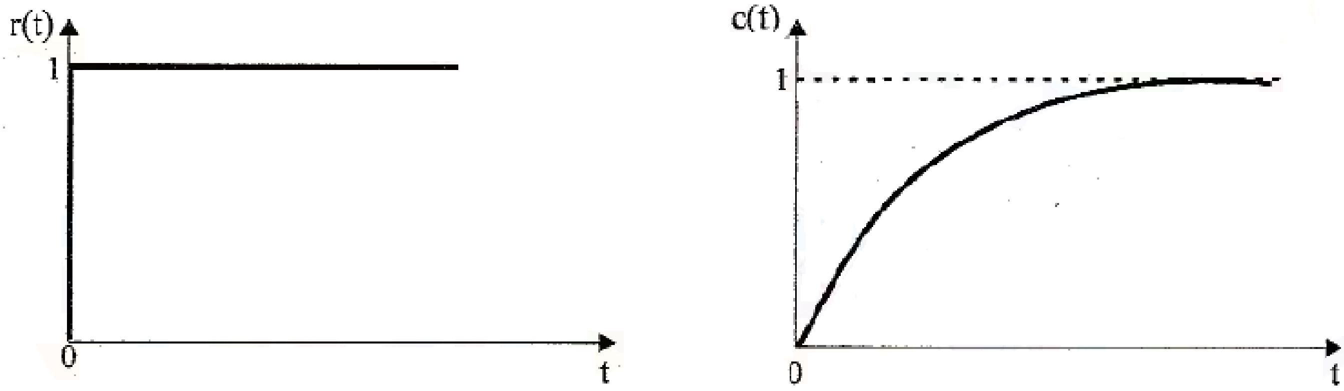


Figure 2.4.5 Response of critically damped second order system to unit step input

[Source: "Control Systems" by Nagoor Kani, Page: 2.25]

Case 4: Overdamped system

When $\zeta > 1$,

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

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = \{s^2 + 2\zeta\omega_n s + \omega_n^2 + \zeta^2\omega_n^2 - \zeta^2\omega_n^2\}$$

$$= (s + \zeta\omega_n)^2 - \omega_n^2(\zeta^2 - 1)$$

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

For unit step input, R(s)=1/s,

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

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

By applying partial fraction,

$$C(s) = \frac{A}{s} + \frac{B}{\left(s + \zeta\omega_n + \omega_n\sqrt{1 - \zeta^2}\right)} + \frac{C}{\left(s + \zeta\omega_n - \omega_n\sqrt{1 - \zeta^2}\right)}$$

By applying inverse Laplace transform,

$$c(t) = \left[1 + \left(\frac{1}{2\left(\zeta + \sqrt{\zeta^2 - 1}\right)\left(\sqrt{\zeta^2 - 1}\right)}\right)e^{-\left(\zeta\omega_n + \omega_n\sqrt{\zeta^2 - 1}\right)t} - \left(\frac{1}{2\left(\zeta - \sqrt{\zeta^2 - 1}\right)\left(\sqrt{\zeta^2 - 1}\right)}\right)e^{-\left(\zeta\omega_n - \omega_n\sqrt{\zeta^2 - 1}\right)t}\right]$$

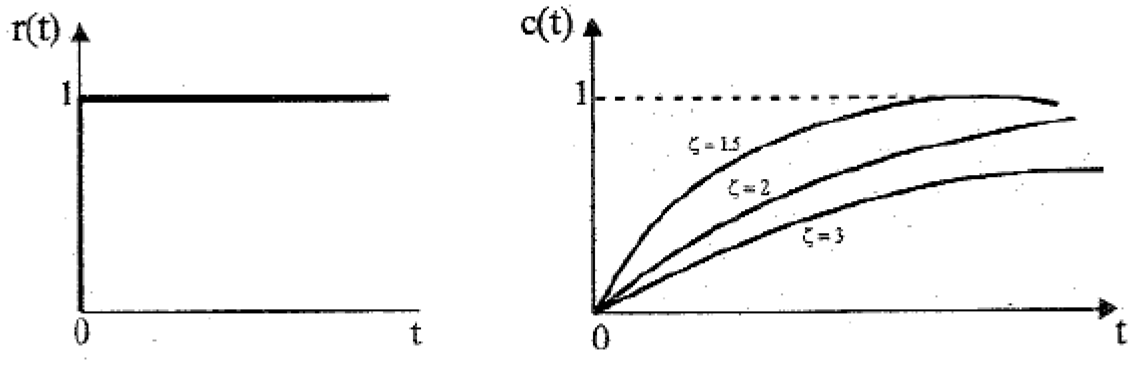


Figure 2.4.6 Response of over damped second order system to unit step input

[Source: "Control Systems" by Nagoor Kani, Page: 2.27]

2.5 ERROR COEFFICIENTS

There are two different types of error coefficient representation namely,

a) Static error constants

b) Generalized error coefficients

STATIC ERROR CONSTANTS

Positional error constant, $K_p = \lim_{s\to 0} G(s)H(s)$

Velocity error constant, $K_v = \lim_{s\to 0} sG(s)H(s)$

Acceleration error constant, $K_a = \lim_{s\to 0} s^2G(s)H(s)$

GENERALIZED ERROR COEFFICIENTS

$$C_o = \lim_{s\to 0} F(s)$$

$$C_1 = \lim_{s\to 0} \frac{dF(s)}{ds}$$

$$C_2 = \lim_{s\to 0} \frac{d^2F(s)}{ds^2}$$

where, $F(s) = \frac{1}{1+G(s)H(s)}$

Relation between static error constants and generalized error coefficients

$$C_o = \frac{1}{1+K_p}$$

$$C_1 = \frac{1}{K_v}$$

$$C_2 = \frac{1}{K_a}$$

2.6 STEADY STATE ERROR

The deviation of the output of control system from desired response during steady state is known as steady state error. It is represented as e_{ss}. We can find steady state error using the final value theorem as follows.

$$e_{ss} = \lim_{t \to \infty} e(t)$$

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

where, E(s) is the Laplace transform of the error signal, e(t)

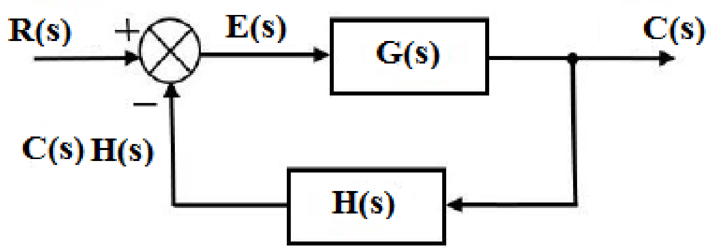


Figure 2.6.1 Closed loop control system

[Source: "Control Systems Engineering" by I J Nagrath, M Gopal, Page: 213]

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

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

$$E(s)\big(1 + G(s)H(s)\big) = R(s)$$

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

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

When a control system is excited with standard input signal, the steady state error may be zero, constant or infinity. Its value depends on the type number and input signal.

a) Type-0 system will have a constant steady state error when the input is step signal
b) Type-1 system will have a constant steady state error when the input is ramp signal
c) Type-2 system will have a constant steady state error when the input is parabolic signal

$$\text{For unit step input, } e_{ss} = \frac{1}{1+K_p}$$

$$\text{For unit ramp input, } e_{ss} = \frac{1}{K_v}$$

$$\text{For unit parabolic input, } e_{ss} = \frac{1}{K_a}$$

Static error constants for various type number of systems

Error constants	Type number of system			
	0	**1**	**2**	**3**
K_p	Constant	∞	∞	∞
K_v	0	Constant	∞	∞
K_a	0	0	Constant	∞

Steady state error for various types of input

Input signal	Type number of system			
	0	**1**	**2**	**3**
K_p	$\frac{1}{1+K_p}$	0	0	0
K_v	∞	$\frac{1}{K_v}$	0	0
K_a	∞	∞	$\frac{1}{K_a}$	0

2.7 ROOT LOCUS CONSTRUCTION

The root locus is a graphical representation in s-domain and it is symmetrical about the real axis. Because the open loop poles and zeros exist in the s-domain having the values either as real or as complex conjugate pairs.

Rules for Construction of Root Locus

The following rule structure is followed for constructing a root locus.

Rule 1 − Locate the open loop poles and zeros in the 's' plane.

Rule 2 − Find the number of root locus branches.

We know that the root locus branches start at the open loop poles and end at open loop zeros. So, the number of root locus branches **N** is equal to the number of finite open loop poles **P** or the number of finite open loop zeros **Z**, whichever is greater.

Mathematically, we can write the number of root locus branches **N** as

$$N = P \text{ if } P \geq Z$$

$$N = Z \text{ if } P < Z$$

Rule 3 − Identify and draw the **real axis root locus branches**.

If the angle of the open loop transfer function at a point is an odd multiple of 180^0, then that point is on the root locus. If odd number of the open loop poles and zeros exist to the left side of a point on the real axis, then that point is on the root locus branch. Therefore, the branch of points which satisfies this condition is the real axis of the root locus branch.

Rule 4 − Find the centroid and the angle of asymptotes.

- If $P=Z$

then all the root locus branches start at finite open loop poles and end at finite open loop zeros.

- If $P>Z$

then Z number of root locus branches start at finite open loop poles and end at finite open loop zeros and $P-Z$

number of root locus branches start at finite open loop poles and end at infinite open loop zeros.

- If $P<Z$

then P number of root locus branches start at finite open loop poles and end at finite open loop zeros and $Z-P$ number of root locus branches start at infinite open loop poles and end at finite open loop zeros.

So, some of the root locus branches approach infinity, when $P \neq Z$. Asymptotes give the direction of these root locus branches. The intersection point of asymptotes on the real axis is known as centroid.

We can calculate the **centroid α** by using this formula,

$$\alpha = \frac{\sum Real\ part\ of\ finite\ open\ loop\ poles - \sum Real\ part\ of\ finite\ open\ loop\ zeros}{P-Z}$$

Angle of asymptotes,

$$\theta = \frac{(2q+1)180^0}{P-Z}$$

where,

$$q=0,1,2,....,(P-Z)-1$$

Rule 5 − Find the intersection points of root locus branches with an imaginary axis.

We can calculate the point at which the root locus branch intersects the imaginary axis and the value of **K** at that point by using the Routh array method and special **case (ii)**.

- If all elements of any row of the Routh array are zero, then the root locus branch intersects the imaginary axis and vice-versa.
- Identify the row in such a way that if we make the first element as zero, then the elements of the entire row are zero. Find the value of **K** for this combination.
- Substitute this **K** value in the auxiliary equation. You will get the intersection point of the root locus branch with an imaginary axis.

Rule 6 − Find Break-away and Break-in points.

- If there exists a real axis root locus branch between two open loop poles, then there will be a break-away point in between these two open loop poles.
- If there exists a real axis root locus branch between two open loop zeros, then there will be a break-in point in between these two open loop zeros.

[Note − Break-away and break-in points exist only on the real axis root locus branches.]

Follow these steps to find break-away and break-in points.

1. Write K in terms of s from the characteristic equation $1+G(s)H(s)=0$

2. Differentiate *K* with respect to s and make it equal to zero. Substitute these values of *s* in the above equation.
3. The values of *s* for which the *K* value is positive are the break points.

Rule 7 − Find the angle of departure and the angle of arrival.

The Angle of departure and the angle of arrival can be calculated at complex conjugate open loop poles and complex conjugate open loop zeros respectively.

Angle of departure,

$$\phi_d = 180^0 - \phi$$

Angle of arrival,

$$\phi_a = 180^0 + \phi$$

where,

$$\phi = \sum \phi_P - \sum \phi_Z$$

2.8 ROUTH HURWITZ CRITERION

Consider a closed-loop transfer function

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

where the a_i's and b_i's are real constants and $m \le n$. An alternative to factoring the denominator polynomial, Routh's stability criterion, determines the number of closed-loop poles in the right-half s-plane.

Algorithm for applying Routh's stability criterion

The algorithm described below, like the stability criterion, requires the order of A(s) to be finite.

1. Factor out any roots at the origin to obtain the polynomial, and multiply by −1 if necessary, to obtain

$$a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n = 0$$
$$where, a_0 \neq 0 \; and \; a_n > 0$$

2. If the order of the resulting polynomial is at least two and any coefficient a_i is zero or negative, the polynomial has at least one root with nonnegative real part. To obtain the precise number of roots with nonnegative real part, proceed as follows. Arrange the coefficients of the polynomial, and values subsequently calculated from them as shown below:

s^n	a_0	a_2	a_4	a_6	$\cdots$
s^{n-1}	a_1	a_3	a_5	a_7	$\cdots$
s^{n-2}	b_1	b_2	b_3	b_4	$\cdots$
s^{n-3}	c_1	c_2	c_3	c_4	$\cdots$
s^{n-4}	d_1	d_2	d_3	d_4	$\cdots$
$\vdots$	$\vdots$	$\vdots$			
s^2	e_1	e_2			
s^1	f_1				
s^0	g_0				

The array is generated until all subsequent coefficients are zero. Similarly, cross multiply the coefficients of the two previous rows to obtain the c_i, d_i, etc. Until the nth row of the array has been completed. Missing coefficients are replaced by zeros. The resulting array is called the Routh array. The powers of s are not considered to be part of the array. We can think of them as labels. The column beginning with a0 is considered to be the first column of the array. The Routh array is seen to be triangular. It can be shown that multiplying a row by a positive number to simplify the calculation of the next row does not affect the outcome of the application of the Routh criterion.

where, the coefficients b_i are,

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

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

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

$$\vdots$$

3. Count the number of sign changes in the first column of the array. It can be shown that a necessary and sufficient condition for all roots of (2) to be located in the left-half plane is that all the a_i are positive and all of the coefficients in the first column be positive.

Example: Generic Cubic Polynomial

Consider the generic cubic polynomial:

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

where all the ai are positive. The Routh array is

s^3	a_0	a_2
s^2	a_1	a_3
s^1	$\frac{a_1 a_2 - a_0 a_3}{a_0}$	
s^0	a_3	

So, the condition that all roots have negative real parts is

$$a_1 a_2 > a_0 a_3$$

Example: A Quadratic Polynomial.

Next, we consider the fourth-order polynomial:

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

Here we illustrate the fact that multiplying a row by a positive constant does not change the result. One possible Routh array is given at left, and an alternative is given at right,

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

Also,

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

In this example, the sign changes twice in the first column so the polynomial equation A(s) = 0 has two roots with positive real parts.

Necessity of all coefficients being positive

In stating the algorithm above, we did not justify the stated conditions. Here we show that all coefficients being positive is necessary for all roots to be located in the left half-plane. It can be shown that any polynomial ins, all of whose coefficients are real, can be factored into a product of a maximal number linear and quadratic factors also having real coefficients. Clearly a linear factor (s+a) has nonnegative real root if a is positive. For both roots of a quadratic factor (s2+bs+c) to have negative real parts both b and c must be positive. (If c is negative, the square root ofb2−4cis real and the quadratic factor can be factored into two linear factors so the number of factors was not maximal.) It is easy to see that if all coefficients of the factors are positive, those of the original polynomial must be as well. To see that the condition is not sufficient, we can refer to several examples above.

Example: Determining Acceptable Gain Values

Consider a system whose closed-loop transfer function is

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

Characteristic equation

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

Routh array is

s^4	1	3	K
s^3	3	2	
s^2	7/3	K	
s^1	(14-9K)/7		
s^0	K		

For the system to be stable, the elements of the first column of the Routh array should be positive. Based on that condition, the s^1 row yields the condition that, for stability,

$$\frac{(14-9K)}{7} > 0$$

$$(14-9K) > 0$$

$$14 > 9K$$

$$\frac{14}{9} > K$$

The s^0 row yields the condition that, for stability,

$$K > 0$$

Hence, the system is stable when the value of K lies in the range of

$$0 < K < 14/9$$

Special Case: Zero First-Column Element.

If the first term in a row is zero, but the remaining terms are not, the zero is replaced by a small, positive value of ϵ and the calculation continues as described above. Here's an example:

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

Routh array is

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

Special Case: Zero Row

If all the coefficients in a row are zero, a pair of roots of equal magnitude and opposite sign is indicated. These could be two real roots with equal magnitudes and opposite signs or two conjugate imaginary roots. The zero row is replaced by taking the coefficients of dP(s)/ds, where P(s), called the auxiliary polynomial, is obtained from the values in the row above the zero row. The pair of roots can be found by solving dP(s)/ds= 0. Note that the auxiliary polynomial always has even degree. It can be shown that an auxiliary polynomial of degree 2n has n pairs of roots of equal magnitude and opposite sign.

2.9 TIME RESPONSE ANALYSIS

- Two types of inputs can be applied to a control system
- Command Input or Reference Input $y_r(t)$
- Disturbance Input w(t)

(External disturbances w(t) are typically uncontrolled variations in the load on a control system). In systems controlling mechanical motions, load disturbances may represent forces. In voltage regulating systems, variations in electrical load area major source of disturbances.

In general, the closed loop transfer function of a system is denoted as M(s).

$$M(s) = \frac{b_0 s^M + b_1 s^{M-1} + b_2 s^{M-2} + \dots\dots\dots + b_{M-1} s + b_M}{a_0 s^N + a_1 s^{N-1} + a_2 s^{N-2} + \dots\dots\dots + a_{N-1} s + a_N}$$

3.1 FREQUENCY RESPONSE

The response of a system for the sinusoidal input is called sinusoidal response. The ratio of sinusoidal response to sinusoidal input is called sinusoidal transfer function of the system and in general, it is denoted by, T(jω). The sinusoidal transfer function is the frequency domain representation of the system and so it is also called frequency domain transfer function.

The frequency domain transfer function T(jω) is a complex function of ω. Hence, it can be separated into magnitude function and phase function. Now, the magnitude and phase functions will be real functions of ω and they are called frequency response.

The frequency response can be evaluated for open loop system and closed loop system. The frequency domain transfer function of open loop and closed loop systems can be obtained from the s-domain transfer function by replacing 's' by jω as shown:

Open loop transfer function: $\boldsymbol{G(j\omega) = |G(j\omega)|\angle G(j\omega)}$

Loop transfer function: $\boldsymbol{G(j\omega)H(j\omega) = |G(j\omega)H(j\omega)|\angle G(j\omega)H(j\omega)}$

Closed loop transfer function: $\boldsymbol{M(j\omega) = |M(j\omega)|\angle M(j\omega)}$

The advantages of frequency response analysis are the following:

1. The absolute and relative stability of the closed loop system can be estimated from the knowledge of their open loop frequency response.
2. The practical testing of systems can be easily carried with available sinusoidal signal generators and precise measurement equipments.
3. The transfer function of complicated systems can be determined experimentally by frequency response tests.
4. The design and parameter adjustment of the open loop transfer function of a system for specified closed loop performance is carried out more easily in frequency domain.
5. When the system is designed by the use of frequency response analysis, the effects of noise disturbance and parameter variations are relatively easy to visualize and incorporate corrective measures.
6. The frequency response analysis and designs can be extended to certain non-linear control systems.

The frequency response of a system is a frequency dependent function which expresses how a sinusoidal signal of a given frequency on the system input is transferred through the system. Time-varying signals at least periodical signals – which excite systems, as the reference (set point) signal or a disturbance in a control system or measurement signals which are inputs signals to signal filters, can be regarded as consisting of a sum of frequency components. Each frequency component is a sinusoidal signal having certain amplitude and a certain frequency. (The Fourier series expansion or the Fourier transform can be used to express these frequency components quantitatively.) The frequency response expresses how each of these frequency components is transferred through the system. Some components may be amplified, others may be attenuated, and there will be some phase lag through the system. The frequency response is an important tool for analysis and design of signal filters (as low pass filters and high pass filters), and for analysis, and to some extent, design, of control systems. Both signal filtering and control systems applications are described (briefly) later in this chapter. The definition of the frequency response – which will be given in the next section – applies only to linear models, but this linear model may very well be the local linear model about some operating point of a non-linear model. The frequency response can be found experimentally or from a transfer function model. It can be presented graphically or as a mathematical function.

FREQUENCY DOMAIN SPECIFICATIONS

The performance and characteristics of a system in frequency domain are measured in terms of frequency domain specifications. The requirements of a system to be designed are usually specified in terms of these specifications.

The frequency domain specifications are,

a) Resonant peak, M_r
b) Resonant frequency, ω_r
c) Bandwidth, ω_b
d) Cut-off rate
e) Gain margin, K_g
f) Phase margin, γ

FREQUENCY DOMAIN SPECIFICATIONS OF SECOND ORDER SYSTEM

Resonant peak, M_r

The maximum value of the magnitude of closed loop transfer function is called the resonant peak, M_r. A large resonant peak corresponds to a large overshoot in transient response. Consider the closed loop transfer function of second order system,

$$\frac{C(s)}{R(s)} = M(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

The sinusoidal transfer function M(jω) is obtained by letting s=jω.

$$M(j\omega) = \frac{\omega_n^2}{(j\omega)^2 + 2\zeta\omega_n(j\omega) + \omega_n^2}$$

$$= \frac{\omega_n^2}{-\omega^2 + 2\zeta\omega_n(j\omega) + \omega_n^2}$$

$$= \frac{\omega_n^2}{\omega_n^2\left(-\frac{\omega^2}{\omega_n^2} + 2j\zeta\frac{\omega}{\omega_n} + 1\right)}$$

$$= \frac{1}{1 - \left(\frac{\omega}{\omega_n}\right)^2 + 2j\zeta\frac{\omega}{\omega_n}}$$

Let normalized frequency, $u = \left(\frac{\omega}{\omega_n}\right)$,

$$M(j\omega) = \frac{1}{1 - u^2 + 2j\zeta u}$$

Let, M – Magnitude of closed loop transfer function

α – Phase of closed loop transfer function

$$M = |M(j\omega)| = [(1-u^2)^2 + 4\zeta^2u^2]^{-\frac{1}{2}}$$

$$\alpha = \angle M(j\omega) = -\tan^{-1}\frac{2\zeta u}{1-u^2}$$

The resonant peak is the maximum value of M. The condition for maximum value of M can be obtained by differentiating the equation of M with respect to u and letting (dM/du=0) when ($u=u_r$) with normalized frequency, $u_r = \frac{\omega_r}{\omega}$.

On differentiating 'M' with respect to 'u', we get,

$$\frac{dM}{du} = \frac{d}{du}[1 - u^2 + 2j\zeta u]^{-\frac{1}{2}}$$

$$= -\frac{1}{2}[1-u^2+2j\zeta u]^{-\frac{3}{2}}[2(1-u^2)(-2u)+8\zeta^2 u]$$

$$= -\frac{[-4u(1-u^2)+8\zeta^2 u]}{2[(1-u^2)^2+4\zeta^2u^2]^{\frac{3}{2}}}$$

$$= -\frac{[4u(1-u^2)-8\zeta^2 u]}{2[(1-u^2)^2+4\zeta^2u^2]^{\frac{3}{2}}}$$

Replacing u by u_r and equating dM/du to zero,

$$\frac{[4u_r(1-{u_r}^2)-8\zeta^2 u_r]}{2[(1-{u_r}^2)^2+4\zeta^2{u_r}^2]^{\frac{3}{2}}} = 0$$

$$4u_r(1-{u_r}^2)-8\zeta^2 u_r = 0$$

$$4u_r-4{u_r}^3-8\zeta^2 u_r = 0$$

$$4u_r-4{u_r}^3 = 8\zeta^2 u_r$$

$$4{u_r}^3 = 4u_r-8\zeta^2 u_r$$

$${u_r}^2 = 1-2\zeta^2$$

$$u_r = \sqrt{1-2\zeta^2}$$

Therefore, the resonant peak occurs when $u_r = \sqrt{1-2\zeta^2}$

On substituting for M with M=M_r and u=u_r,

$$M_r = \frac{1}{[(1-{u_r}^2)^2+4\zeta^2{u_r}^2]^{\frac{1}{2}}} = \frac{1}{\left[\left(1-(1-2\zeta^2)\right)^2+4\zeta^2(1-2\zeta^2)\right]^{\frac{1}{2}}}$$

$$= \frac{1}{[4\zeta^4+4\zeta^2-8\zeta^4]^{\frac{1}{2}}} = \frac{1}{[4\zeta^2-4\zeta^4]^{\frac{1}{2}}} = \frac{1}{[4\zeta^2(1-\zeta^2)]^{\frac{1}{2}}} = \frac{1}{2\zeta\sqrt{1-\zeta^2}}$$

$$M_r = \frac{1}{2\zeta\sqrt{1-\zeta^2}}$$

Resonant frequency, ω_r

The frequency at which the resonant peak occurs is called resonant frequency, ω_r. This is related to the frequency of oscillation in the step response and thus it is indicative of the speed of transient response.

Normalized resonant frequency,

$$u_r = \frac{\omega_r}{\omega_n} = \sqrt{1-2\zeta^2}$$

$$\omega_r = \omega_n\sqrt{1-2\zeta^2}$$

Bandwidth, ω_b

The bandwidth is the range of frequencies for which the system normalized gain is more than -3db. The frequency at which the gain is -3db is called cut-off frequency. Bandwidth is usually defined for closed loop system and it transmits the signals whose frequencies are less than the cut-off frequency. The bandwidth is a measure of the ability of a feedback system to reproduce the input signal, noise rejection characteristics and rise time. A large bandwidth corresponds to a small rise time or fast response.

Let, normalized bandwidth,

$$u_b = \frac{\omega_b}{\omega_n}$$

When $u=u_b$, the magnitude M, of the closed loop system is $1/\sqrt{2}$ or (-3db)

On substituting for M with $u=u_b$ and equating it to $1/\sqrt{2}$

$$M = \frac{1}{[(1-{u_b}^2)^2 + 4\zeta^2{u_b}^2]^{\frac{1}{2}}} = \frac{1}{\sqrt{2}}$$

On squaring and cross multiplying, we get,

$$(1-{u_b}^2)^2 + 4\zeta^2{u_b}^2 = 2$$

$$1 + {u_b}^4 - 2{u_b}^2 + 4\zeta^2{u_b}^2 = 2$$

$${u_b}^4 - 2{u_b}^2(1-2\zeta^2) - 1 = 0$$

Let $x = {u_b}^2$,

$$x^2 - 2x(1-2\zeta^2) - 1 = 0$$

Hence,

$$x = \frac{2(1-2\zeta^2) \pm \sqrt{4(1-2\zeta^2)^2+4}}{2} = \frac{2(1-2\zeta^2) \pm 2\sqrt{(1+4\zeta^4-4\zeta^2)+1}}{2}$$

Let us take only the positive sign,

$$x = 1 - 2\zeta^2 + \sqrt{(2+4\zeta^4-4\zeta^2)}$$

But, $u_b = \sqrt{x}$

$$u_b = \left[1 - 2\zeta^2 + \sqrt{(2+4\zeta^4-4\zeta^2)}\right]^{\frac{1}{2}}$$

Also, $u_b = \frac{\omega_b}{\omega_n}$

$$\omega_b = \omega_n \left[1 - 2\zeta^2 + \sqrt{(2 + 4\zeta^4 - 4\zeta^2)}\right]^{\frac{1}{2}}$$

Cut-off rate

The slope of the log-magnitude curve near the cut-off frequency is called cut-off rate. The cut-off rate indicates the ability of the system to distinguish the signal from noise.

Gain margin, K_g

The gain margin, K_g is defined as the value of gain, to be added to system, in order to bring the system to the verge of instability. The gain margin is given by the reciprocal of the magnitude of open loop transfer function at phase crossover frequency.

The frequency at which the phase of open loop transfer function is 180° is called the phase crossover frequency, ω_{pc}.

$$K_g = \frac{1}{|G(j\omega_{pc})|}$$

$$K_g\, in\ db = 20 \log K_g = 20 \log \frac{1}{|G(j\omega_{pc})|}$$

The gain margin in db is given by the negative of the db magnitude of G(jω) at phase crossover frequency. The gain margin indicates the additional gain that can be provided to system without affecting the stability of the system.

[Note: The gain margin of second order system is infinite].

Phase margin, γ

The phase margin, γ is defined as the additional phase lag to be added at the gain crossover frequency in order to bring the system to the verge of instability.

The gain crossover frequency, ω_{gc} is the frequency at which the magnitude of the open loop transfer function is unity (or it is the frequency at which the db magnitude is zero).

The phase margin is obtained by adding 180° to the phase angle, ϕ of the open loop transfer function at the gain crossover frequency. The phase margin indicates the additional phase lag that can be provided to the system without affecting stability.

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

Put s=jω,

$$G(j\omega) = \frac{\omega_n^2}{j\omega(j\omega + 2\zeta\omega_n)} = \frac{\omega_n^2}{\omega_n(j\frac{\omega}{\omega_n})\omega_n(j\frac{\omega}{\omega_n} + 2\zeta)} = \frac{1}{(j\frac{\omega}{\omega_n})(j\frac{\omega}{\omega_n} + 2\zeta)}$$

Let normalized frequency, $u = \frac{\omega}{\omega_n}$

$$G(j\omega) = \frac{1}{(ju)(ju + 2\zeta)}$$

Magnitude of G(jω),

$$|G(j\omega)| = \frac{1}{(u)\sqrt{(u^2 + 4\zeta^2)}} = \frac{1}{\sqrt{(u^4 + 4u^2\zeta^2)}}$$

Phase of G(jω),

$$\angle G(j\omega) = -90^o - \tan^{-1}\frac{u}{2\zeta}$$

At the gain crossover frequency ω_{gc}, the magnitude is unity.

Hence, at $u=u_{gc}$,

$$|G(j\omega_{gc})| = \frac{1}{\sqrt{({u_{gc}}^4 + 4{u_{gc}}^2\zeta^2)}} = 1$$

$$({u_{gc}}^4 + 4{u_{gc}}^2\zeta^2) = 1$$

$$({u_{gc}}^4 + 4{u_{gc}}^2\zeta^2) - 1 = 0$$

Let $x= u_{gc}^2$

$$x^2 + 4x\zeta^2 - 1 = 0$$

$$x = \frac{-4\zeta^2 \pm \sqrt{16\zeta^4 + 4}}{2} = -2\zeta^2 \pm \sqrt{4\zeta^4 + 1}$$

Let us take only the positive sign,

$$x = -2\zeta^2 + \sqrt{4\zeta^4 + 1}$$

Hence,

$$u_{gc} = \left[-2\zeta^2 + \sqrt{4\zeta^4 + 1}\right]^{\frac{1}{2}}$$

Phase margin,

$$\gamma = 180^o + \phi_{gc}$$

$$\gamma = 180^o + \angle G(j\omega_{gc}) = 180^o + \left(-90^o - \tan^{-1}\frac{u_{gc}}{2\zeta}\right)$$

$$\gamma = 90^o - \tan^{-1}\frac{\left[-2\zeta^2 + \sqrt{4\zeta^4 + 1}\right]^{\frac{1}{2}}}{2\zeta}$$

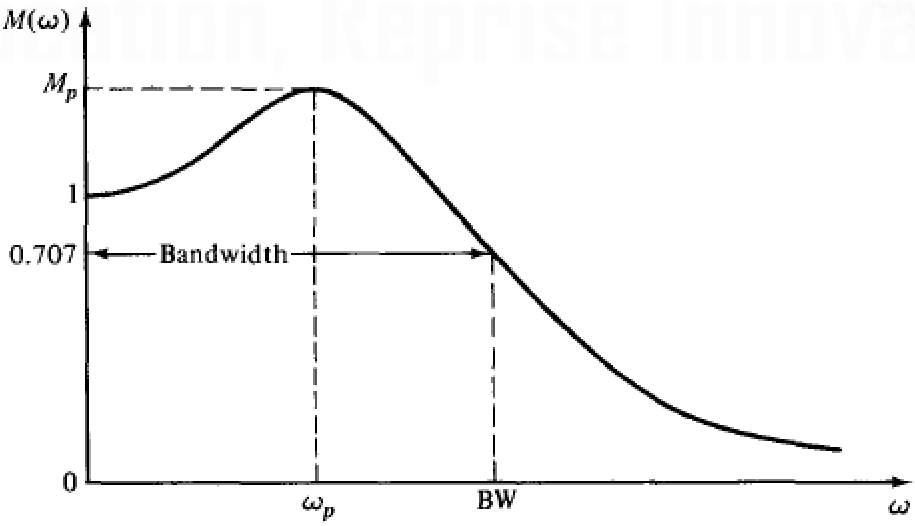


Figure 3.1.1 Typical magnification curve of a feedback control system

[Source: "Automatic Control Systems" by Benjamin C. Kuo, Page: 463]

3.2 BODE PLOT

The Bode plot is a frequency response plot of the sinusoidal transfer function of a system. One is a plot of the magnitude of a sinusoidal transfer function versus log ω. The other is a plot of the phase angle of a sinusoidal transfer function versus log ω. The main advantage of bode plot is that multiplication of magnitudes can be converted into addition. Also, a simple method for sketching an approximate log-magnitude curve is available. A Bode plot is a (semilog) plot of the transfer function magnitude and phase angle as a function of frequency.

The gain magnitude is many times expressed in terms of decibels (dB) = 20 $\log_{10}A$.

Semilog sheet

Two sets of axes: gain on top, phase below (identical)

Logarithmic frequency axes

Gain axis is logarithmic – either explicitly or as units of decibels(dB)

Phase axis is linear with units of degrees

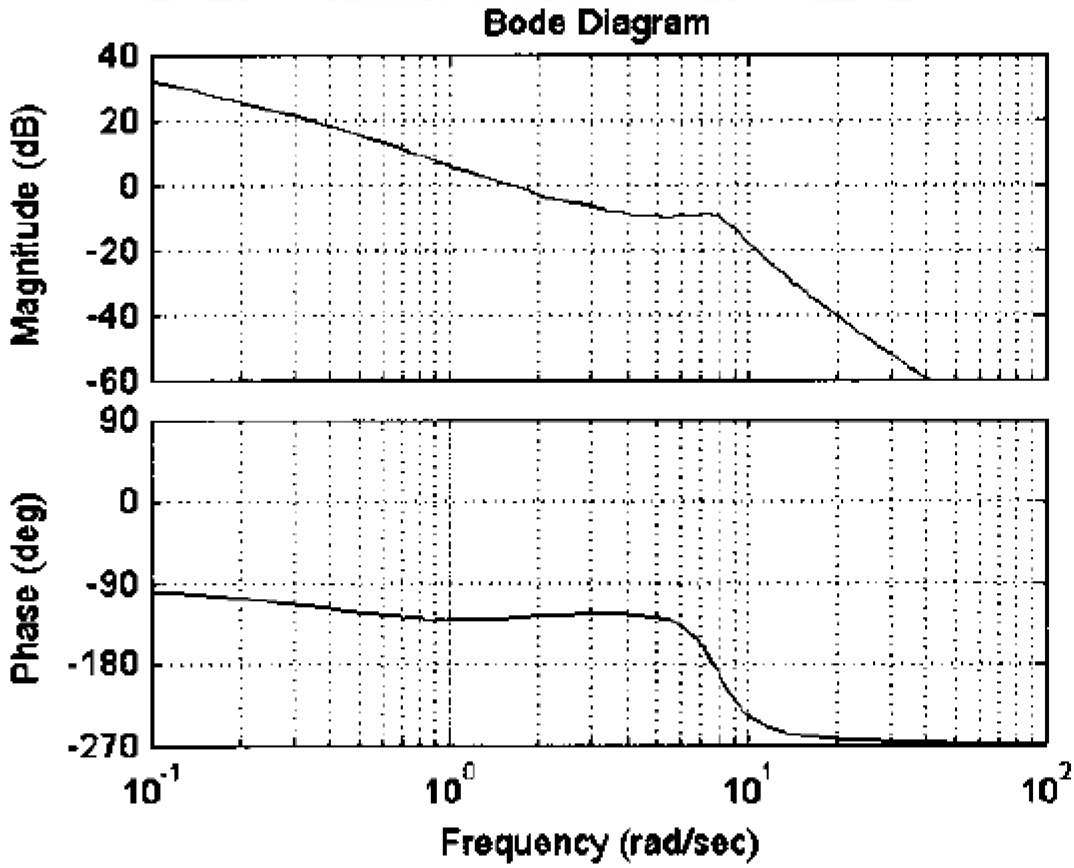


Figure 3.2.1 Magnitude and phase plots of Bode plot

[Source: "Linear Control System Analysis and Design with MATLAB" by John J D'Azzo, Constantine, Stuart, Page: 318]

BASIC FACTORS OF G(jω)

The basic factors that are very frequently occur in a typical transfer function G(jω) are,

1. Constant gain, K
2. Integral and derivative factors $(j\omega)^{\mp 1}$
3. First-order factors $(1 + j\omega T)^{\mp 1}$
4. Quadratic factors $\left(1 + 2\zeta\left(j\frac{\omega}{\omega_n}\right) + \left(j\frac{\omega}{\omega_n}\right)^2\right)^{\mp 1}$

Constant Gain, K

Let G(s)=K,

$$G(j\omega) = K = K\angle 0^o$$

$$A = |G(j\omega)| in\ db = 20\log K$$

$$\phi = \angle G(j\omega) = 0^o$$

The magnitude plot for a constant gain K is a horizontal straight line at the magnitude of 20log K db. The phase plot is a straight line at 0°.

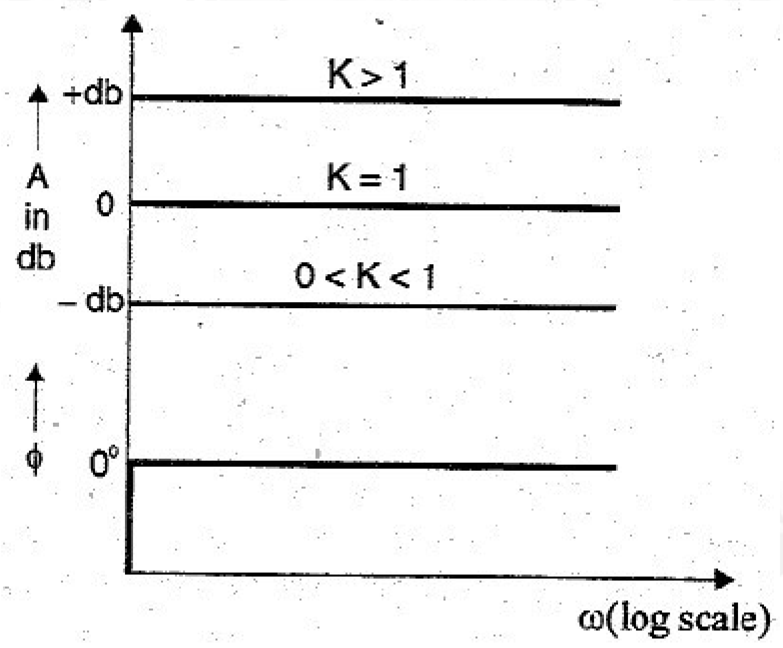


Figure 3.2.2 Bode plot of constant gain, K

[Source: "Control Systems" by A Nagoor Kani, Page: 3.10]

Integral Factor

Let G(s)=K/s,

$$G(j\omega) = \frac{K}{j\omega} = \frac{K}{\omega}\angle - 90^o$$

$$A = |G(j\omega)| in\ db = 20\log\left(\frac{K}{\omega}\right)$$

$$\phi = \angle G(j\omega) = -90^o$$

The magnitude plot of the integral factor is a straight line with the slope of -20db/dec and passing through zero db when ω=K. The phase plot is a straight line at -90°.

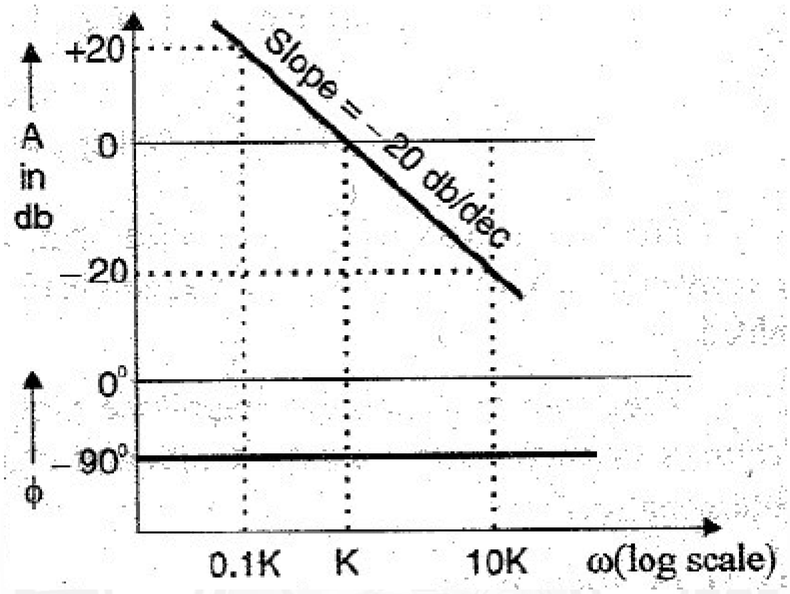


Figure 3.2.3 Bode plot of integral factor, K/jω

[Source: "Control Systems" by A Nagoor Kani, Page: 3.11]

Derivative factor

Let G(s)=Ks,

$$G(j\omega) = Kj\omega = K\omega\angle 90^o$$

$$A = |G(j\omega)| in\ db = 20\log(K\omega)$$

$$\phi = \angle G(j\omega) = +90^o$$

The magnitude plot of the integral factor is a straight line with the slope of 20db/dec and passing through zero db when ω=K. The phase plot is a straight line at +90°.

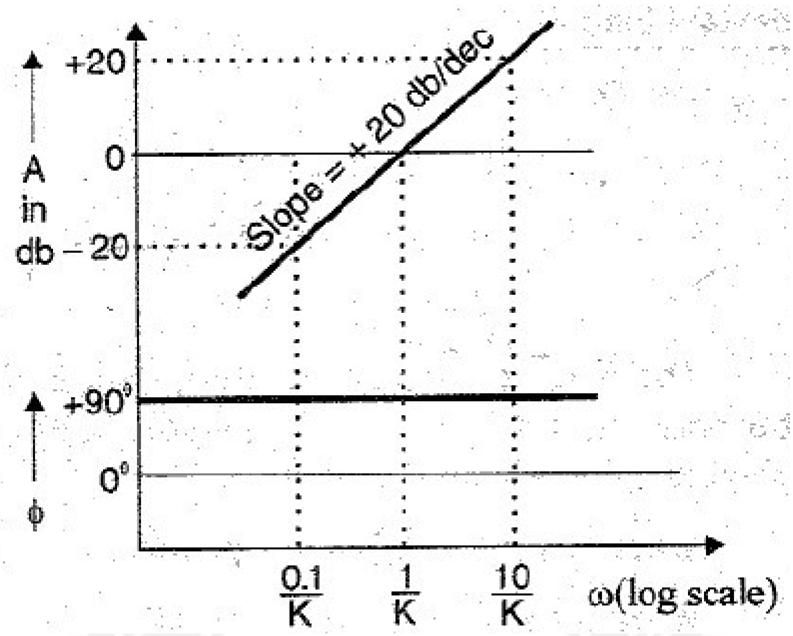


Figure 3.2.4 Bode plot of derivative factor, K x jω

[Source: "Control Systems" by A Nagoor Kani, Page: 3.11]

First order factor in denominator

Let $G(s) = \frac{1}{1+sT}$

$$G(j\omega) = \frac{1}{1 + j\omega T} = \frac{1}{\sqrt{1 + \omega^2 T^2}} \angle - \tan^{-1} \omega T$$

$$A = |G(j\omega)| in\ db = 20 \log \left(\frac{1}{\sqrt{1 + \omega^2 T^2}} \right)$$

$$\phi = \angle G(j\omega) = \angle - \tan^{-1} \omega T$$

The magnitude plot of the first order factor can be approximated by two straight lines, one is a straight line at zero db for the frequency range, $0<\omega<1/T$, and the other is a straight line with slope -20db/dec for the frequency range, $1/T<\omega<\infty$. The corner frequency is $\omega_c=1/T$ and the loss in db at the corner frequency is -3db. The phase angle of the first order factor varies from 0^o to -90^o as ω is varied from zero to infinity. The phase plot is a curve passing through -45^o at ω_c.

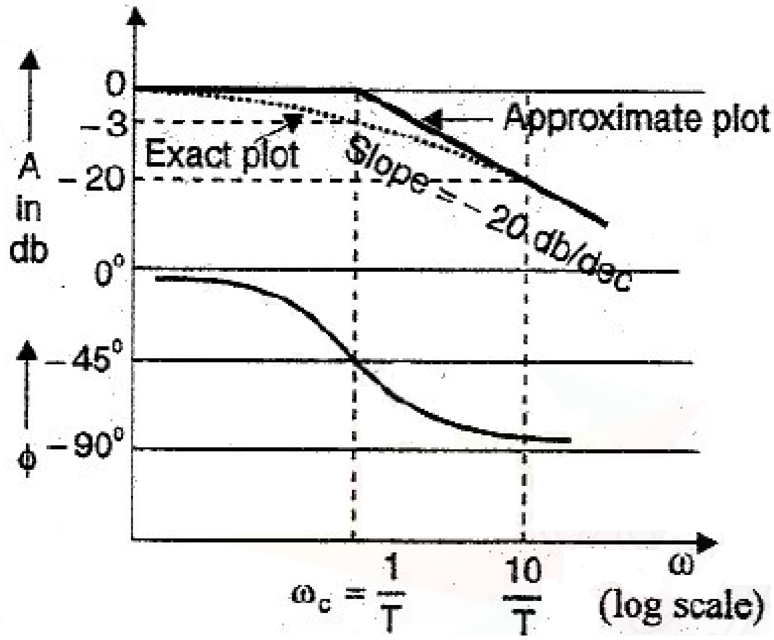


Figure 3.2.5 Bode plot of first order factor in denominator, 1/(1+jωT)

[Source: "Control Systems" by A Nagoor Kani, Page: 3.13]

First order factor in numerator

Let $G(s) = 1 + sT$

$$G(j\omega) = 1 + j\omega T = \sqrt{1 + \omega^2 T^2} \angle \tan^{-1} \omega T$$

$$A = |G(j\omega)| in\ db = 20 \log\left(\sqrt{1 + \omega^2 T^2}\right)$$

$$\phi = \angle G(j\omega) = \angle \tan^{-1} \omega T$$

The magnitude plot of the first order factor can be approximated by two straight lines, one is a straight line at zero db for the frequency range, $0<\omega<1/T$, and the other is a straight line with slope 20db/dec for the frequency range, $1/T<\omega<\infty$. The corner frequency is $\omega_c=1/T$ and the loss in db at the corner frequency is +3db. The phase angle of the first order factor varies from 0^o to $+90^o$ as ω is varied from zero to infinity. The phase plot is a curve passing through $+45^o$ at ω_c.

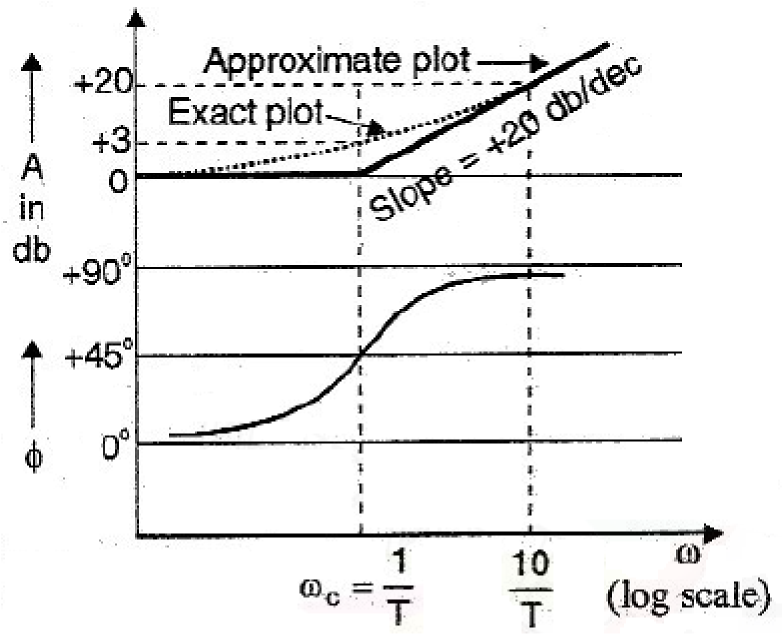


Figure 3.2.6 Bode plot of first order factor in numerator, (1+jωT)

[Source: "Control Systems" by A Nagoor Kani, Page: 3.14]

Quadratic factor in denominator

Second order closed loop transfer function is given by

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \frac{1}{\left(\frac{s}{\omega_n}\right)^2 + 2\zeta\frac{s}{\omega_n} + 1}$$

$$G(j\omega) = \frac{1}{\left(\frac{j\omega}{\omega_n}\right)^2 + 2\zeta\frac{j\omega}{\omega_n} + 1} = \frac{1}{-\left(\frac{\omega}{\omega_n}\right)^2 + 2\zeta\frac{j\omega}{\omega_n} + 1}$$

$$G(j\omega) = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{{\omega_n}^2}\right) + 4\zeta^2\frac{\omega^2}{{\omega_n}^2}}} \angle -\tan^{-1}\left(\frac{2\zeta\frac{\omega}{\omega_n}}{1 - \frac{\omega^2}{{\omega_n}^2}}\right)$$

At low frequencies when $\omega \ll \omega_n$, the magnitude is,

$$A = -20\log\sqrt{1 - \frac{\omega^2}{{\omega_n}^2}(2 - 4\zeta^2) + \frac{\omega^4}{{\omega_n}^4}} \cong -20\log 1 = 0$$

At high frequencies when $\omega \gg \omega_n$, the magnitude is,

$$A = -20\log\sqrt{1 - \frac{\omega^2}{{\omega_n}^2}(2 - 4\zeta^2) + \frac{\omega^4}{{\omega_n}^4}}$$

$$A \cong -20\log\sqrt{\frac{\omega^4}{{\omega_n}^4}} = -\,20\log\frac{\omega^2}{{\omega_n}^2} = -20\log\left(\frac{\omega}{\omega_n}\right)^2$$

$$\phi = \angle G(j\omega) = -\tan^{-1}\left(\frac{2\zeta\frac{\omega}{\omega_n}}{1-\frac{\omega^2}{{\omega_n}^2}}\right)$$

The magnitude plot of the quadratic factor in the denominator can be approximated by two straight lines, one is a straight line at zero db for the frequency range, $0<\omega<\omega_n$, and the other is a straight line with slope -40db/dec for the frequency range, $\omega_n<\omega<\infty$. The frequency at which the two asymptotes meet is called the corner frequency. For the quadratic factor, the frequency, ω_n is the corner frequency, ω_c. The phase angle of the quadratic factor varies from 0^o to -180^o as ω is varied from zero to infinity. The phase plot is a curve passing through -90^o at ω_c. At the corner frequency, phase angle is -90^o and independent of ζ, but at all other frequency it depends on ζ.

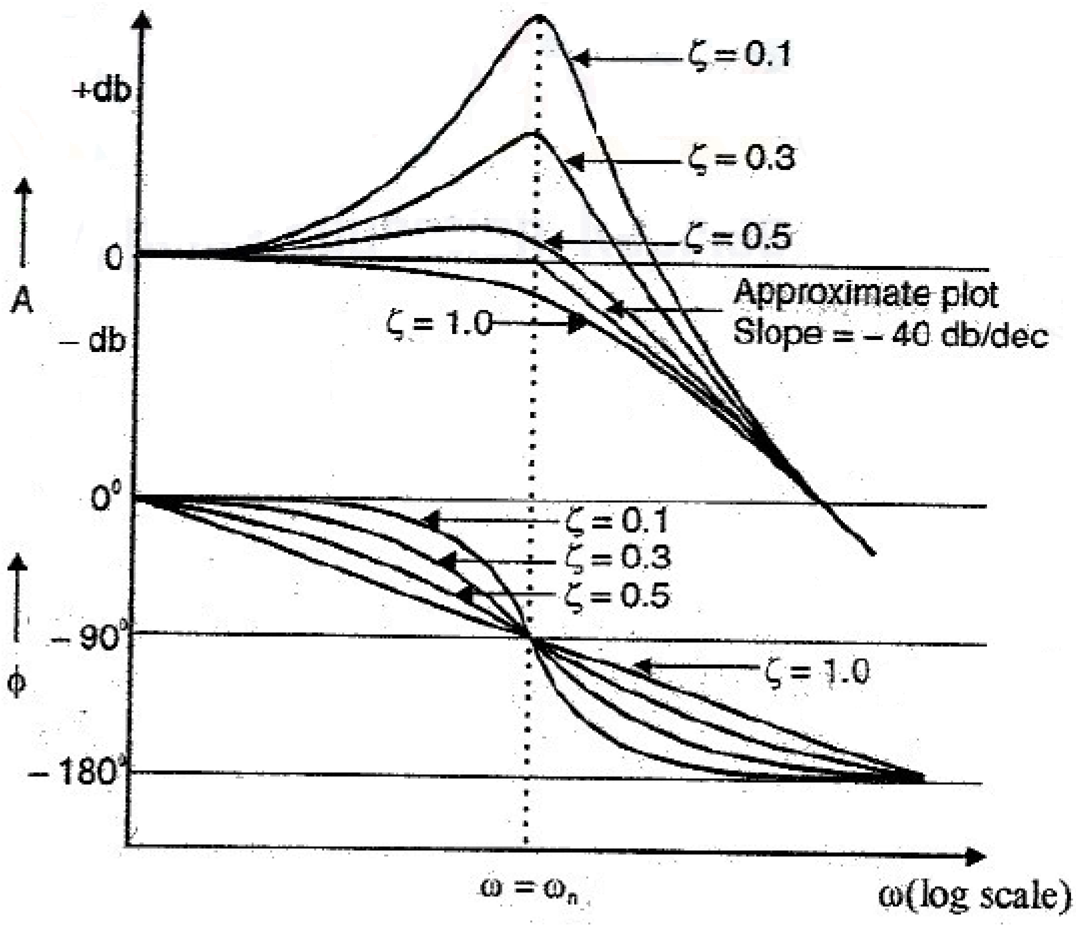


Figure 3.2.7 Bode plot of quadratic factor in denominator

[Source: "Control Systems" by A Nagoor Kani, Page: 3.15]

Quadratic factor in the numerator

$$G(s) = \frac{s^2 + 2\zeta\omega_n s + \omega_n^2}{\omega_n^2} = \left(\frac{s}{\omega_n}\right)^2 + 2\zeta\frac{s}{\omega_n} + 1$$

$$G(j\omega) = \left(\frac{j\omega}{\omega_n}\right)^2 + 2\zeta\frac{j\omega}{\omega_n} + 1 = -\left(\frac{\omega}{\omega_n}\right)^2 + 2\zeta\frac{j\omega}{\omega_n} + 1$$

$$G(j\omega) = \sqrt{\left(1 - \frac{\omega^2}{{\omega_n}^2}\right) + 4\zeta^2\frac{\omega^2}{{\omega_n}^2}} \angle \tan^{-1}\left(\frac{2\zeta\frac{\omega}{\omega_n}}{1 - \frac{\omega^2}{{\omega_n}^2}}\right)$$

At low frequencies when $\omega << \omega_n$, the magnitude is,

$$A = 20\log\sqrt{1 - \frac{\omega^2}{{\omega_n}^2}(2 - 4\zeta^2) + \frac{\omega^4}{{\omega_n}^4}} \cong 20\log 1 = 0$$

At high frequencies when $\omega >> \omega_n$, the magnitude is,

$$A = 20\log\sqrt{1 - \frac{\omega^2}{{\omega_n}^2}(2 - 4\zeta^2) + \frac{\omega^4}{{\omega_n}^4}}$$

$$A \cong 20\log\sqrt{\frac{\omega^4}{{\omega_n}^4}} = 20\log\frac{\omega^2}{{\omega_n}^2} = 20\log\left(\frac{\omega}{\omega_n}\right)^2$$

$$\phi = \angle G(j\omega) = \tan^{-1}\left(\frac{2\zeta\frac{\omega}{\omega_n}}{1 - \frac{\omega^2}{{\omega_n}^2}}\right)$$

The magnitude plot of the quadratic factor in the denominator can be approximated by two straight lines, one is a straight line at zero db for the frequency range, $0<\omega<\omega_n$, and the other is a straight line with slope +40db/dec for the frequency range, $\omega_n<\omega<\infty$. The frequency at which the two asymptotes meet is called the corner frequency. For the quadratic factor, the frequency, ω_n is the corner frequency, ω_c. The phase angle of the quadratic factor varies from 0^o to $+180^o$ as ω is varied from zero to infinity. The phase plot is a curve passing through $+90^o$ at ω_c. At the corner frequency, phase angle is $+90^o$ and independent of ζ, but at all other frequency it depends on ζ.

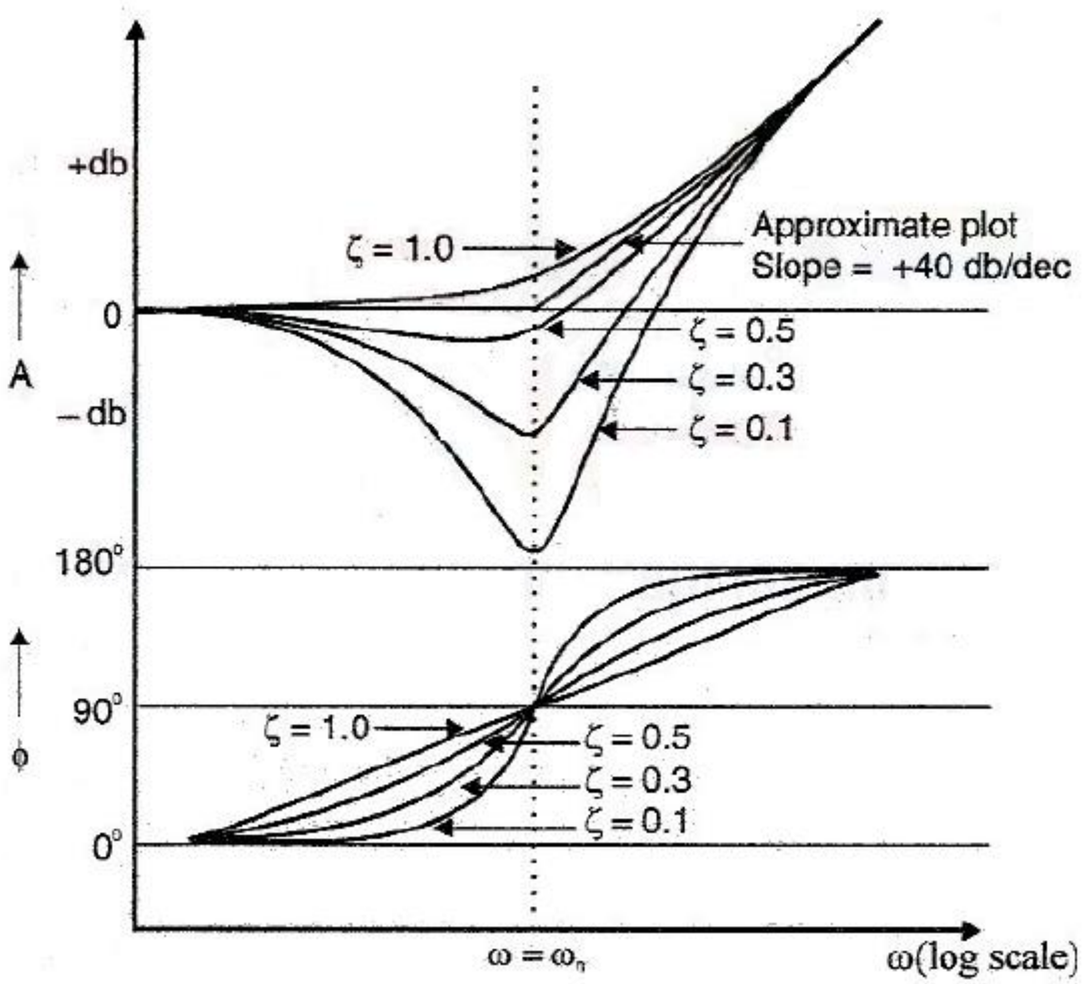


Figure 3.2.8 Bode plot of quadratic factor in numerator

[Source: "Control Systems" by A Nagoor Kani, Page: 3.16]

Derivative factor: magnitude

$$M = 20\log\left|j\omega\right| = 20\log\omega \text{ dB}$$

$$\angle j\omega = 90^0$$

$$\Delta M = 20\log\omega_2 - 20\log\omega_1 = 20\log\frac{\omega_2}{\omega_1} \text{ dB/decade}$$

$$\Delta M = 20\log 10 = 20 \text{ dB/decade}$$

$$\Delta M = 20\log 2 \approx 6 \text{ dB/octave}$$

Integral factor: magnitude

$$M = 20\log\left|\frac{1}{j\omega}\right| = -20\log\omega \text{ dB}$$

$$\angle j\omega = 270^0$$

$$\Delta M = -20\log\omega_2 + 20\log\omega_1 = -20\log\frac{\omega_2}{\omega_1} \text{ dB/decade}$$

$$\Delta M = -20\log 10 = -20 \text{ dB/decade}$$

$$\Delta M = 20\log 2 \approx -6 \text{ dB/octave}$$

First-order derivative factor: magnitude

$$M = 20\log|1 + j\omega\tau| = 20\log(\sqrt{1+[\omega\tau]}^2 \text{ dB}$$

For $\omega \ll \omega_c$, $M \approx 0$ dB

For $\omega \gg \omega_c$,

$$M \approx 20\log\frac{\omega}{\omega_c} \text{ dB}$$

Here, $\omega_c = 1/\tau$ = corner frequency

For $\omega > \omega_c$

$$\Delta M = 20\log\omega_2 - 20\log\omega_1 = 20\log\frac{\omega_2}{\omega_1}$$

$$\Delta M = 20\log 10 = 20 \text{ dB/decade}$$

$$\Delta M = 20\log 2 \approx 6 \text{ dB/octave}$$

First-order derivative factor: phase

$$\theta = \angle 1 + j\omega\tau = \arctan(\omega\tau)$$

$$\theta \approx 0 \qquad ; w < \frac{w_c}{10}$$

$$\theta = 45^0\left(1 + \log\frac{\omega}{\omega_c}\right) \qquad ; \frac{w_c}{10} < w < 10w_c$$

$$\theta \approx 90 \qquad ; w > 10w_c$$

First-order integral factor: magnitude

$$M = 20\log\left|\frac{1}{1+j\omega\tau}\right| = 20\log\left(\frac{1}{\sqrt{1+[\omega\tau]^2}}\right)\text{dB}$$

$$M \approx 0, \qquad w << w_c$$

$$M \approx -20\log\frac{\omega}{\omega_c} dB, \quad w >> w_c$$

$$\Delta M = -20\log\omega_2 + 20\log\omega_1 = -20\log\frac{\omega_2}{\omega_1}\text{dB/decade}$$

$$\Delta M = -20\log 2 \approx -6 \text{ dB/octave}$$

First-order integral factor: phase

$$\theta = 360, \ \omega < \omega_c/10$$

$$\theta = 360 - 45^0\left(1+\log\frac{\omega}{\omega_c}\right), \ \omega_c/10 < \omega < 10\ \omega_c$$

$$\theta = 360 - 45^0\left(1+\log\frac{\omega}{\omega_c}\right)$$

$$\theta = 270, \ \omega > 10\ \omega_c$$

Second-order derivative factor: magnitude

$$M = 20\log|\omega_n^2 - \omega^2 + j2\zeta\omega\omega_n|$$

$$= 20\log\left(\omega_n^2\sqrt{\left(1-\frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}\right)$$

$$\mathbf{M \approx 40\log w_n, w << w_n}$$

$$\mathbf{M = 20\log\ (2zw_n{}^2), w = w_n}$$

$$\mathbf{M = 40\ \log w, w >> w_n}$$

For $w >> w_n$

$$\Delta M = 40\log\omega_2 - 40\log\omega_1 = 40\log\frac{\omega_2}{\omega_1}\ \text{dB/decade}$$

$$\Delta M = 40\log 10 = 40\ \text{dB/decade}$$

$$\Delta M = 40\log 2 \approx 12\ \text{dB/octave}$$

Second-order derivative factor: phase

$$\theta = \angle|\omega_n^2 - \omega^2 + j2\zeta\omega\omega_n| = \arctan\left(\frac{2\zeta\frac{\omega}{\omega_n}}{1-\frac{\omega^2}{\omega_n^2}}\right)$$

$$\theta = 0^0,\quad w < \frac{w_n}{10}$$

$$\theta = 90^0,\quad w = w_n$$

$$\theta = 180^0,\quad w > 10w_n$$

Second-order integral factor

$$M = 20\log\left|\frac{1}{\omega_n^2 - \omega^2 + j2\zeta\omega\omega_n}\right| dB = 20\log\left(\frac{1}{\omega_n^2\sqrt{\left(1-\frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}}\right) dB$$

$M \approx -40\log\omega_n,\ \omega \ll \omega_n$

$M = -20\log(2\zeta\omega_n^2),\ \omega = \omega_n$

$M = -40\log\omega,\ \omega \gg \omega_n$

$$\Delta M = -40\log\omega_2 + 40\log\omega_1 = -40\log\frac{\omega_2}{\omega_1}\ dB/decade$$

PROCEDURE FOR MAGNITUDE PLOT OF BODE PLOT

Step 1: Convert the transfer function into Bode form or time constant form.

Step 2: List the corner frequencies in the increasing order and prepare a table as shown

Term	Corner frequency rad/sec	Slope db/dec	Change in Slope db/dec

In the above table, enter K or $K/(j\omega)^n$ or $K(j\omega)^n$ as the first term and the other terms in the increasing order of corner frequencies. Then enter the corner frequency, slope contributed by each term and change in slope at every corner frequency.

Step 3: Choose an arbitrary frequency ω_l which is lesser than the lowest corner frequency. Calculate the db magnitude of K or $K/(j\omega)^n$ or $K(j\omega)^n$ at ω_l and at the lowest corner frequency.

Step 4: Then calculate the gain (db magnitude) at every corner frequency one by one by using the formula,

$$\text{Gain at } \omega_y = \text{change in gain from } \omega_x \text{ to } \omega_y + \text{Gain at } \omega_x$$

$$A_y = (\text{Slope from } \omega_x \text{ to } \omega_y \times \log(\omega_y/\omega_x) + \text{Gain at } \omega_x$$

Step 5: Choose an arbitrary frequency ω_h which is greater than the highest corner frequency. Calculate the gain at ω_h by using the formula in step 4.

Step 6: In a semilog graph sheet mark the required range of frequency on x-axis (log scale) and the range of db magnitude on y-axis (ordinary scale) after choosing proper units.

Step 7: Mark all the points obtained in steps 3, 4, 5 on the graph and join the points by straight lines. Mark the slope at every part of the graph.

PROCEDURE FOR PHASE PLOT OF BODE PLOT

The phase plot is an exact plot obtained with exact phase angles of $G(j\omega)$ computed for various values of ω and is then tabulated. The choice of frequencies are preferably the frequencies chosen for magnitude plot. Usually the magnitude plot and phase plot are drawn in a single semilog sheet on a common frequency scale. Take another y-axis in the graph where the magnitude plot is drawn and, in this y-axis, mark the desired range of phase angles after choosing proper units. From the tabulated values of ω and phase angles, mark all the points on the graph. Join the points by a smooth curve.

DETERMINATION OF GAIN MARGIN AND PHASE MARGIN FROM BODE PLOT

The gain margin in db is given by the negative of db magnitude of $G(j\omega)$ at the phase crossover frequency, ω_{pc}. The ω_{pc} is the frequency at which phase of $G(j\omega)$ is 180^o. if the db magnitude of $G(j\omega)$ at ω_{pc} is negative then gain margin is positive and vice versa.

Let Φ_{gc} be the phase angle of $G(j\omega)$ at gain cross over frequency, ω_{gc}. The ω_{gc} is the frequency at which the db magnitude of $G(j\omega)$ is zero. Now the phase margin, γ is given by, $\gamma = 180^o + \Phi_{gc}$. If Φ_{gc} is less negative than -180°, then phase margin is positive and vice versa. The positive and negative gain margins and phase margins are illustrated in figure 3.2.9.

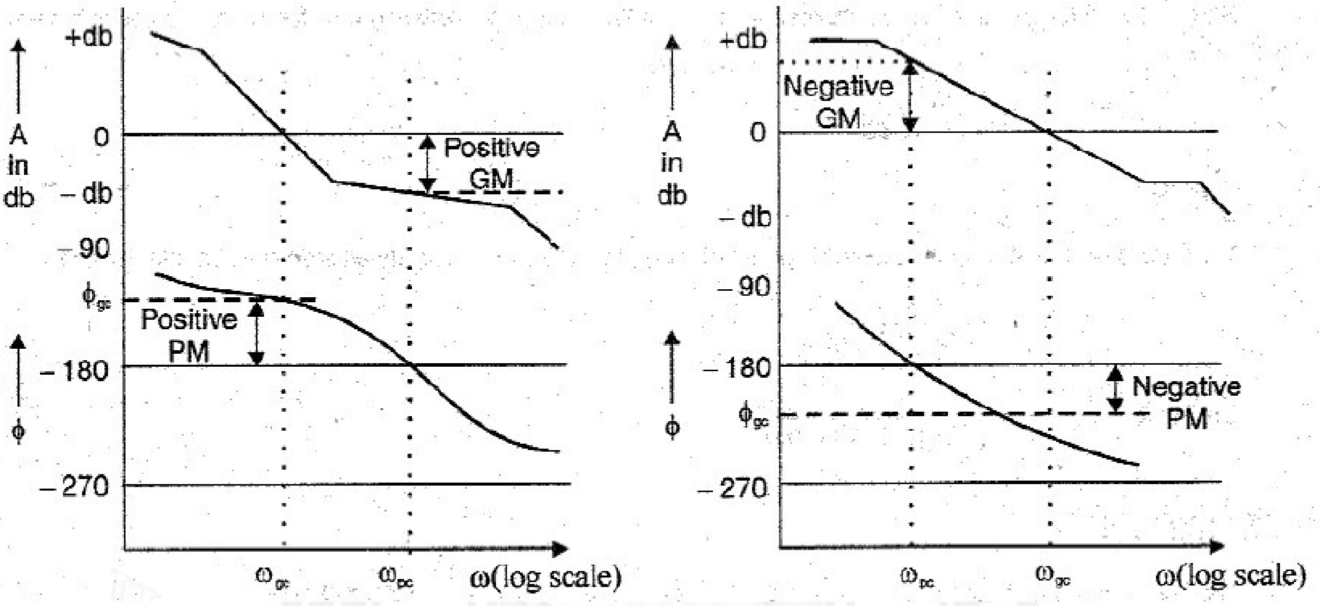


Figure 3.2.9 Gain margin and Phase margin in Bode plot

[Source: "Control Systems" by A Nagoor Kani, Page: 3.20]

3.3 POLAR PLOT

The polar plot of a sinusoidal transfer function G(jω) is a plot of the magnitude G(jω) versus the phase angle of G(jω) on polar coordinates as ω is varied from zero to infinity. Thus, the polar plot is the locus of vectors |G(jω)| < as ω is varied from zero to infinity. The polar plot is also called Nyquist plot. It is a graphical method of determining stability of feedback control systems by using the polar plot of their open-loop transfer functions. Polar plot is a plot to be drawn between magnitude and phase. Polar plot is a plot of magnitude of G(jω) versus the phase of G(jω) in polar co-ordinates. But the magnitudes are presented with normal values only. The Polar plot is a plot, which can be drawn between the magnitude and the phase angle of G(jω) H(jω) by differentiating g ω from zero to ∞. The polar graph sheet is described in below mentioned image. This graph sheet includes various concentric circles and radial lines. The concentric circles and the radial lines are considered as the magnitudes and phase angles.

- Angles are highlighted with positive values in anti-clock wise direction.
- Mark angles with negative values in clockwise direction.

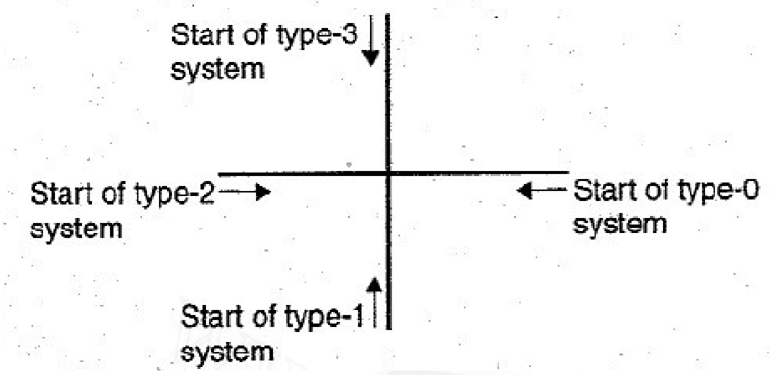
The polar plot is usually plotted on a polar graph sheet. The polar graph sheet has concentric circles and radial lines. The circles represent the magnitude and the radial lines represent the phase angles. Each point on the polar graph has a magnitude and phase angle. The magnitude of a point is given by the value of the circle passing through that point and the phase angle is given by the radial line passing through that point. In polar graph sheet a positive phase angle is measured in anticlockwise from the reference axis (0°) and a negative angle is measured clockwise from the reference axis (0°). In order to plot the polar plot, magnitude and phase of G(jω) are computed for various values of ω and tabulated. Usually the choice of frequencies are corner frequencies and frequencies around corner frequencies. Choose proper scale for the magnitude circles. Fix all the points on polar graph sheet and join the points by smooth curve, write the frequency corresponding to each point of the plot. Alternatively, if G(jω) can be expressed in rectangular coordinates as,

$$G(j\omega) = G_R(j\omega) + jG_i(j\omega)$$

where, $G_R(j\omega)$ = Real part of G(jω), $G_i(j\omega)$ = Imaginary part of G(jω)

Then the polar plot can be plotted in ordinary graph sheet between $G_R(j\omega)$ and $G_i(j\omega)$ by varying ω from 0 to infinity. In order to plot the polar plot on ordinary graph sheet, the magnitude and phase if G(jω) are computed for various values of ω. Then convert the polar coordinates to rectangular coordinates using P → R conversion (polar to rectangular conversion) in the calculator. Sketch the polar plot using rectangular coordinates. For minimum phase transfer function with only poles, type number of the system determines the quadrant at which the polar plot starts and the order of the system determines quadrant at which the polar plot ends. The minimum phase systems are systems with all poles and zeros on left half of s-plane. The start and end of polar plot of all pole minimum phase system are shown in figures respectively. Some typical sketches of polar plot are shown in table. The change in shape of polar plot can be predicted due to addition of a pole or zero.

1. When a pole is added to s system, the polar plot end point will shift by -90°.
2. When a zero is added to s system, the polar plot end point will shift by +90°.



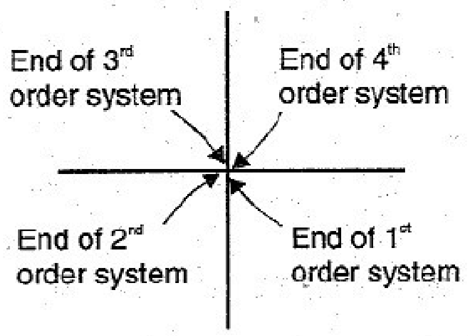


Figure 3.3.1 Start and end of polar plot of all pole minimum phase system

[Source: "Control Systems" by A Nagoor Kani, Page: 3.38]

RULES FOR DRAWING POLAR PLOT

- ✓ Substitute, s=jω in the open loop transfer function.
- ✓ Write the expressions for magnitude and the phase of G(jω) H(jω).
- ✓ Find the starting magnitude and the phase of G(jω) H(jω) by substituting ω=0. So, the polar plot starts with this magnitude and the phase angle.
- ✓ Find the ending magnitude and the phase of G(jω) H(jω) by substituting ω=∞. So, the polar plot ends with this magnitude and the phase angle.
- ✓ Check whether the polar plot intersects the real axis, by making the imaginary term of G(jω) H(jω) equal to zero and find the value(s) of ω.

- ✓ Check whether the polar plot intersects the imaginary axis, by making real term of G(jω) H(jω) equal to zero and find the value(s) of ω.
- ✓ For drawing polar plot more clearly, find the magnitude and phase of G(jω) H(jω) by considering the other value(s) of ω.

DETERMINATION OF GAIN MARGIN AND PHASE MARGIN FROM POLAR PLOT

The gain margin is defined as the inverse of the magnitude of G(jω) at phase crossover frequency. The phase crossover frequency is the frequency at which the phase of G(jω) is 180º. Let the polar plot cut the 180º axis at point B and the magnitude circle passing through the point B be G_B. Now the gain margin, $K_g = 1/ G_B$. If the point B lies within unity circle, the gain margin is positive otherwise negative. If the polar plot is drawn in ordinary graph sheet using rectangular coordinates then the point B is the cutting point of G(jω) locus with negative real axis and $K_g = 1/|G_B|$ where G_B is the magnitude corresponding to point B). The phase margin is defined as, phase margin, $\gamma = 180^o + \Phi_{gc}$ is the phase angle of G(jω) at gain crossover frequency. The gain crossover frequency is the frequency at which the magnitude of G(jω) is unity. Let the polar plot cut the unity circle at point A as shown in figures. Now the phase margin, γ is given by ∠AOP, i.e., ∠AOP is below -180º axis then the phase margin is positive and if it is above -180º axis then the phase margin is negative.

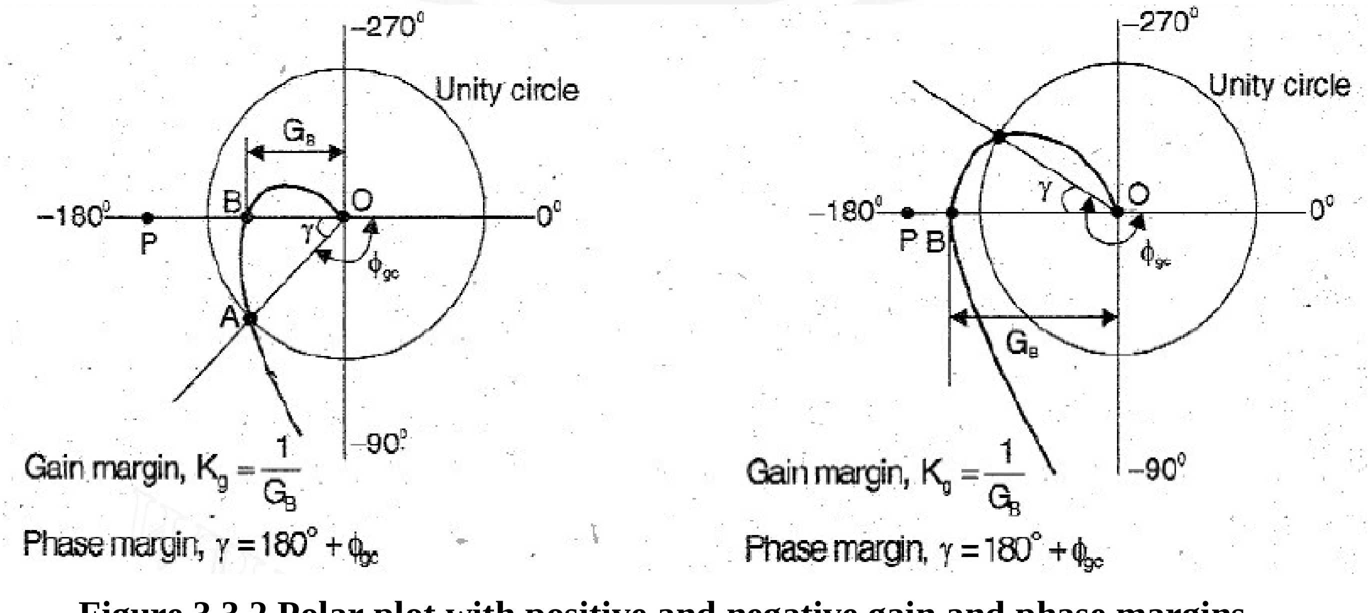


Figure 3.3.2 Polar plot with positive and negative gain and phase margins

[Source: "Control Systems" by A Nagoor Kani, Page: 3.41]

GAIN ADJUSTMENT USING POLAR PLOT

To determine K for specified GM

Draw G(jω) locus with K = 1. Let it cut the -180^o axis at point B corresponding to gain of G_B. Let the specified gain margin be x db. For this gain margin, the G(jω) locus will cut -180^o at point A whose magnitude is G_A.

$$20 \log \frac{1}{G_A} = x$$

$$\log \frac{1}{G_A} = \frac{x}{20}$$

$$\frac{1}{G_A} = 10^{\frac{x}{20}}$$

$$G_A = \frac{1}{10^{\frac{x}{20}}}$$

Now the value of K is given by, K = G_A/G_B

If, K>1, then the system gain should be increased.

If K<1, then the system gain should be reduced.

To determine K for specified PM

Draw G(jω) locus with K = 1. Let it cut the unity circle at point B. (The gain at point B is G_B and equal to unity). Let the specified phase margin be x^o. For a phase margin of x^o, let Φ_{gcx} be the phase angle of G(jω) at gain crossover frequency.

$$x^o = 180^o + \phi_{gcx}$$

$$\phi_{gcx} = x^o - 180^o$$

In the polar plot, the radial line corresponding to will cut the locus of G(jω) with K = 1 at point A and the magnitude corresponding to that point be G_A.

Now, K = G_B/G_A = 1/G_A (since G_B = 1)

3.4 DETERMINATION OF CLOSED LOOP RESPONSE FROM OPEN LOOP RESPONSE

M and N circles

Peak magnitude

$$M_r = 20\log\left|\frac{C(j\omega)}{R(j\omega)}\right| dB$$

where, 3 dB is considered good.

M-CIRCLES

$$M(j\omega) = \frac{G(j\omega)}{1+G(j\omega)}$$

$$G(j\omega) = X + jY$$

$$M(j\omega) = \frac{X+jY}{1+X+jY} = \frac{\sqrt{X^2+Y^2}\angle\tan^{-1}\left(\frac{Y}{X}\right)}{\sqrt{(1+X)^2+Y^2}\angle\tan^{-1}\left(\frac{Y}{1+X}\right)}$$

$$= \frac{\sqrt{X^2+Y^2}}{\sqrt{(1+X)^2+Y^2}}\angle\tan^{-1}\left(\frac{Y}{X}\right) - \tan^{-1}\left(\frac{Y}{1+X}\right)$$

Let, M = Magnitude of M(jω)

$$|M(j\omega)| = \frac{\sqrt{X^2+Y^2}}{\sqrt{(1+X)^2+Y^2}}$$

$$M^2(1+X)^2 + M^2Y^2 = X^2 + Y^2$$

$$X^2(1-M^2) + (1-M^2)Y^2 - 2M^2X = M^2$$

$$X^2 + Y^2 - 2\frac{M^2}{(1-M^2)}X = \frac{M^2}{(1-M^2)}$$

Adding $\left(\frac{M^2}{(1-M^2)}\right)^2$ on both sides, we get,

$$\left(X - \frac{M^2}{(1-M^2)}\right)^2 + Y^2 = \left(\frac{M}{(1-M^2)}\right)^2$$

The above equation represents a family of circles with its

centre at $\left(\frac{M^2}{(1-M^2)}, 0\right)$ and radius $\frac{M}{(1-M^2)}$

Family of M-circles corresponding to the closed loop magnitudes, M of a unit feedback system is given by the figure 3.4.1.

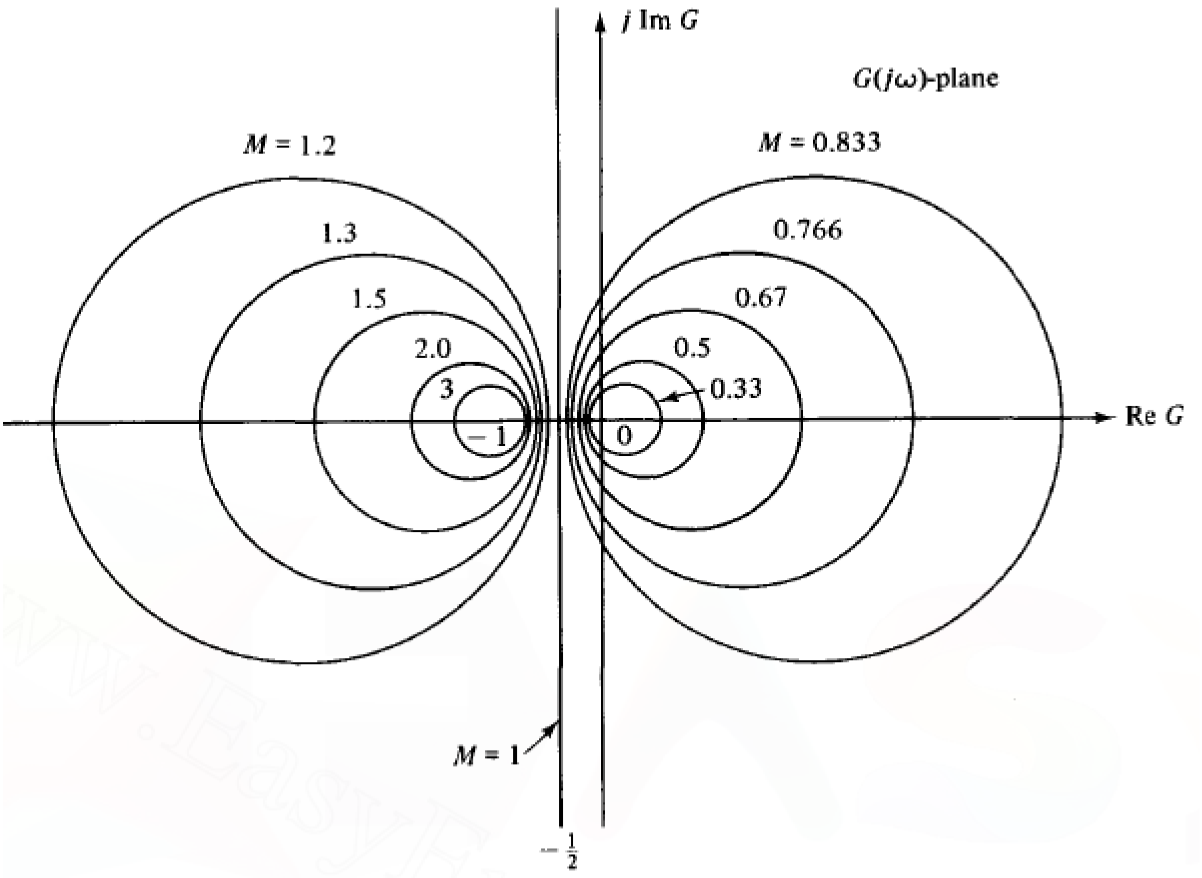


Figure 3.4.1 Constant M-circles in the polar co-ordinates

[Source: "Automatic Control Systems" by Benjamin C. Kuo, Page: 487]

N-CIRCLES

$$\angle M(j\omega) = \alpha = \frac{\angle G(j\omega)}{\angle(1 + G(j\omega))}$$

$$\alpha = \tan^{-1}\frac{Y}{X} - \tan^{-1}\frac{Y}{1+X}$$

$$\tan\alpha = N = \tan\left(\tan^{-1}\frac{Y}{X} - \tan^{-1}\frac{Y}{1+X}\right)$$

We know,

$$\tan(A-B) = \frac{\tan A - \tan B}{1 + \tan A \tan B}$$

$$\mathrm{N} = \left(\frac{Y}{X^2 + X + Y^2}\right)$$

$$\left(X + \frac{1}{2}\right)^2 + \left(Y - \frac{1}{2N}\right)^2 = \frac{1}{4} + \left(\frac{1}{2N}\right)^2$$

The above equation represents the family of circles with its

Centre at $\left(-\frac{1}{2}, \frac{1}{2N}\right)$ and radius $\sqrt{\frac{1}{4} + \left(\frac{1}{2N}\right)^2}$

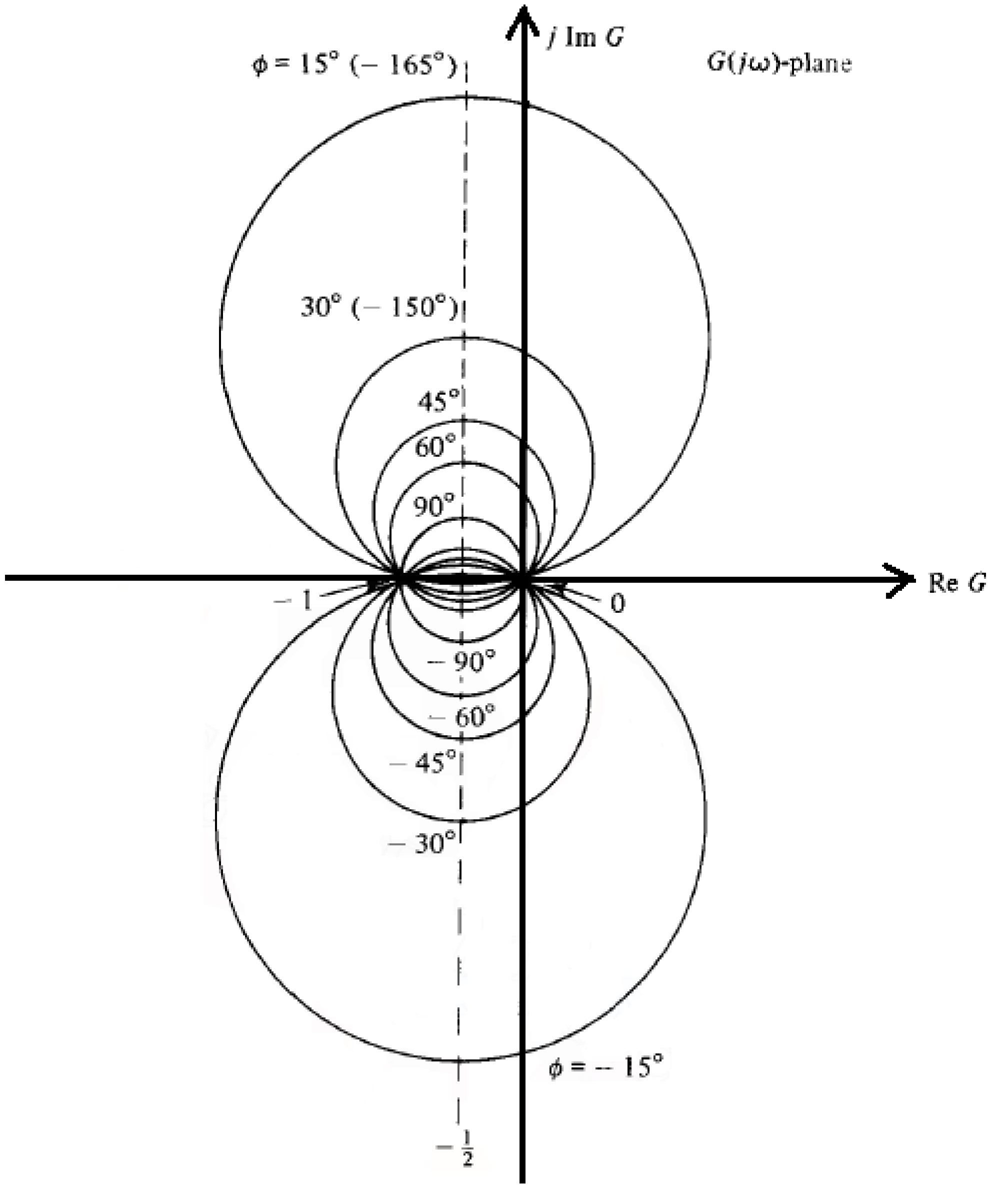


Figure 3.4.2 Constant N-circles in the polar co-ordinates

[Source: "Automatic Control Systems" by Benjamin C. Kuo, Page: 490]

3.5 CORRELATION BETWEEN FREQUENCY DOMAIN AND TIME DOMAIN SPECIFICATIONS

For a second order system,

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

Put $s=j\omega$

$$\frac{C(j\omega)}{R(j\omega)} = \frac{\omega_n^2}{-\omega^2 + 2\zeta\omega_n j\omega + \omega_n^2}$$

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

Let $u = \frac{\omega}{\omega_n}$, then

$$\frac{C(j\omega)}{R(j\omega)} = \frac{1}{(1-u^2) + 2\zeta ju}$$

We know,

$$M(j\omega) = |M(j\omega)|\angle M(j\omega)$$

$$|M(j\omega)| = \frac{1}{\sqrt{(1-u^2)^2 + (2\zeta u)^2}}$$

$$\theta = -\tan^{-1}\left(\frac{2\zeta u}{1-u^2}\right)$$

Now,

$$M_r = \frac{1}{2\zeta\sqrt{1-\zeta^2}}$$

$$\omega_r = \omega_n\sqrt{1-2\zeta^2}$$

$$\omega_b = \omega_n\sqrt{1-2\zeta^2+\sqrt{4\zeta^4-4\zeta^2+2}}$$

$$PM = -180^o + \phi$$

where,

$$\phi = \tan^{-1}\left(\frac{2\zeta}{\sqrt{\sqrt{4\zeta^2+1}-2\zeta^2}}\right)$$

3.6 NYQUIST STABILITY CRITERION

Nyquist criterion is a graphical method of determining stability of feedback control systems by using the Nyquist plot of their open-loop transfer functions.

Feedback transfer function

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

Poles and zeros of the open loop transfer function

$$G(s)H(s) = \frac{K(s - z_1)(s - z_2) \ldots (s - z_m)}{(s - p_1)(s - p_2) \ldots (s - p_n)}$$

$$1 + G(s)H(s) = \frac{(s - p_1)(s - p_2) \ldots (s - p_n) + K(s - z_1)(s - z_2) \ldots (s - z_m)}{(s - p_1)(s - p_2) \ldots (s - p_n)}$$

Number of closed loop poles – Number of zeros of 1+GH = Number of open loop poles

$$1 + G(s)H(s) = \frac{(s - z_{c1})(s - z_{c2}) \ldots (s - z_{cm})}{(s - p_1)(s - p_2) \ldots (s - p_n)}$$

where, z_{c1}, z_{c2}, ….., z_{cm} = zeros of 1+G(s)H(s)

These are also poles of the closed loop transfer function

$$Magnitude, |1 + G(s)H(s)| = \frac{|(s - z_{c1})||(s - z_{c2})| \ldots |(s - z_{cm})|}{|(s - p_1)||(s - p_2)| \ldots |(s - p_n)|}$$

$$Angle, \angle 1 + G(s)H(s) = \frac{\angle(s - z_{c1})\angle(s - z_{c2}) \ldots \angle(s - z_{cm})}{\angle(s - p_1)\angle(s - p_2) \ldots \angle(s - p_n)}$$

The s-plane to 1+GH plane mapping phase angle of the 1+G(s)H(s) vector, corresponding to a point on the s-plane is the difference between the sum of the phase of all vectors drawn from zeros of 1+GH (closed loop poles) and open loops on the s plane. If this point s is moved along a closed contour enclosing any or all of the above zeros and poles, only the phase of the vector of each of the enclosed zeros or open-loop poles will change by 3600. The direction will be in the same sense of the contour enclosing zeros and in the opposite sense for the contour enclosing open-loop poles. A stability test for time invariant linear systems can also be derived in the frequency domain. It is known as Nyquist stability criterion. It is based on the complex analysis result known as *Cauchy's principle of argument*. Note that the system transfer function is a complex function. By applying Cauchy's principle of argument to the *open-loop system* transfer function, we will get information about stability of the closed-loop system transfer function and arrive at the Nyquist stability criterion (Nyquist, 1932). The importance of Nyquist stability lies in the fact that it can also be used to determine the relative degree of system stability by producing the so-called phase and gain stability margins. These stability margins are needed for frequency domain controller design techniques. Only the essence of the Nyquist stability criterion is presented and the phase and gain stability margins are defined. The Nyquist method is used for studying the stability of linear systems with pure time delay.

For a SISO feedback system the closed-loop transfer function is given by,

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

where, G(s) represents the system and H(s) is the feedback element. Since the system poles are determined as those values at which its transfer function becomes infinity, it follows that the closed-loop system poles are obtained by solving the following equation.

$$1 + G(s)H(s) = 0 = \Delta(s)$$

which, in fact, represents the *system characteristic equation*.

Principles of Argument

When a closed contour in the s-plane encloses a certain number of poles and zeros of 1+G(s)H(s) in the clockwise direction, the number of encirclements of the origin by the corresponding contour in the G(s)H(s) plane will encircle the point (-1,0) a number of times given by the difference between the number of its zeros and poles of 1+G(s)H(s) it enclosed on the s-plane. Let F(s) be an analytic function in a closed region of the complex plane given in figure 3.6.1 except at a finite number of points (namely, the poles of F(s)). It is also assumed that F(s) is analytic at everypoint on the contour. Then, *as s travels around the contour in the s - plane in the clockwise direction, the function encircles the origin in the (Re{F(s)}, Im{F(s)}) - plane in the same direction times (see figure 4.3.1), with given by,*

$$N = Z - P$$

where Z and P stand for the number of zeros and poles (including their multiplicities) of the function F(s) inside the contour.

$$arg\{F(s)\} = (Z - P)2\pi = 2\pi N$$

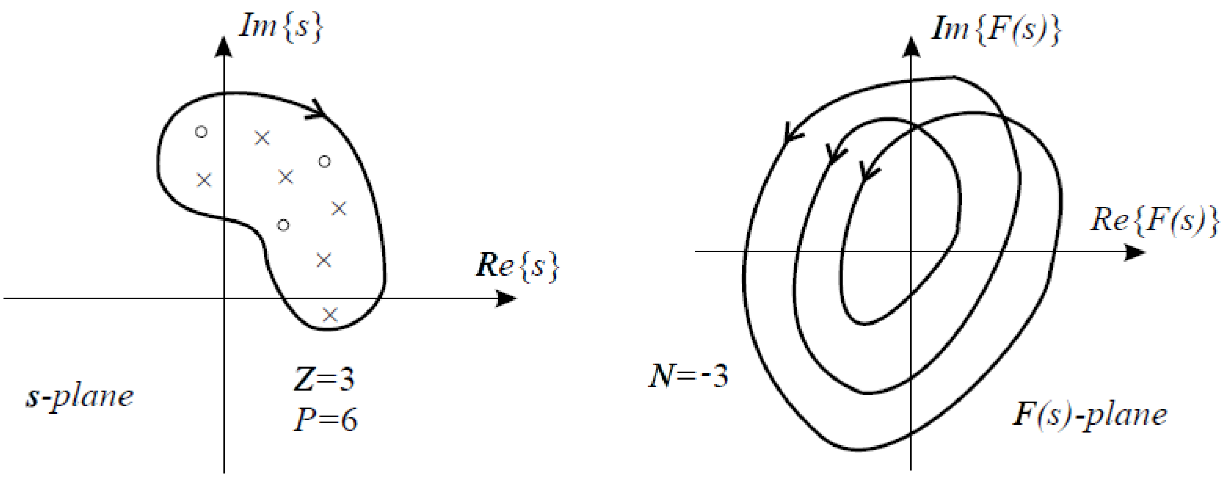


Figure 3.6.1 s-plane and F(s) plane contours

[Source: "Control Systems" by A Nagoor Kani, Page: 4.27]

Contour in the s-plane

The Nyquist plot is a polar plot of the function D(s) = 1+G(s)H(s) when 's' travels around the contour given in figure 3.6.2.

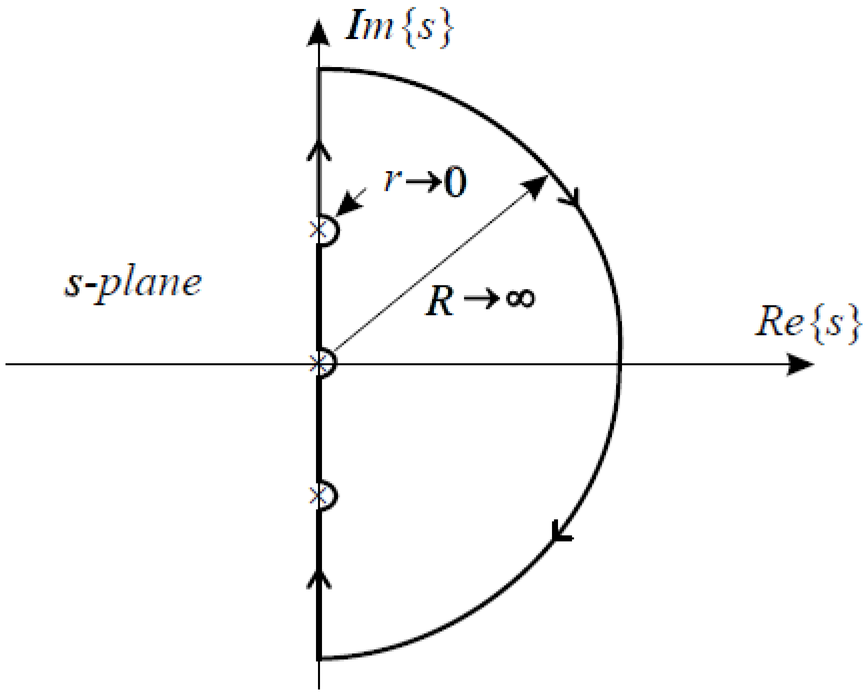


Figure 3.6.2 Nyquist contour when the poles are on imaginary axis and at origin

[Source: "Control Systems" by A Nagoor Kani, Page: 4.33]

Phase and Gain Stability Margins

Two important notions can be derived from the Nyquist diagram: *phase and gain stability margins*. The phase and gain stability margins are presented in figure 3.6.3.

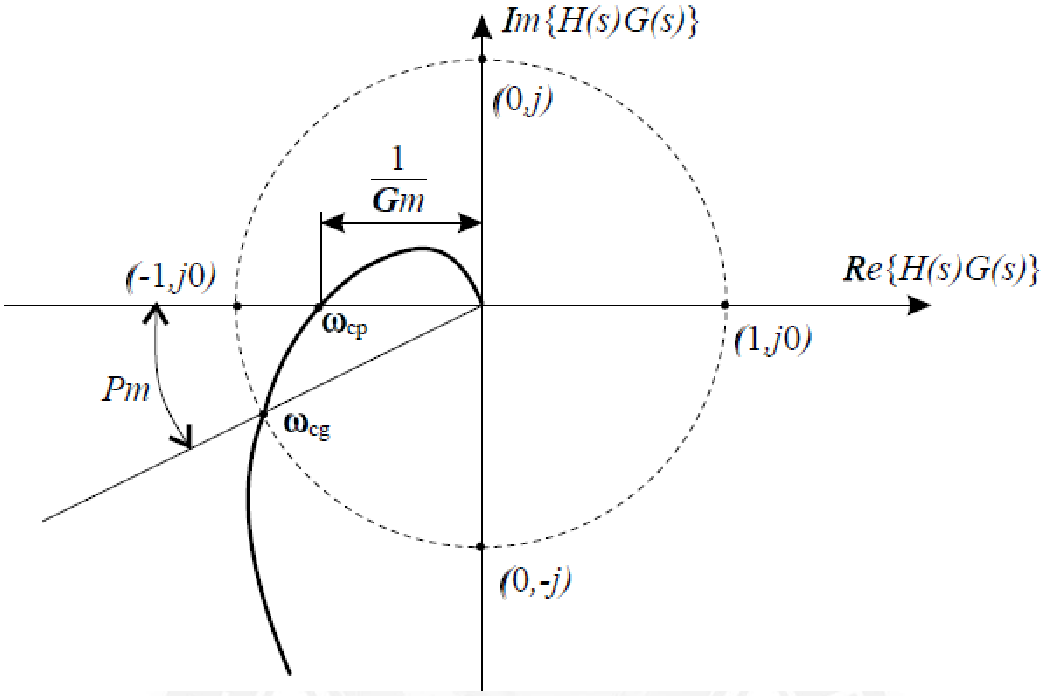


Figure 3.6.3 Gain and Phase margin

[Source: "Control Systems" by A Nagoor Kani, Page: 4.33]

They give the degree of relative stability; in other words, they tell how far the given system is from the instability region. Their formal definitions are given by

$$PM = 180^o + arg\{G(j\omega_{gc})H(j\omega_{gc})\}$$

$$GM(dB) = 20\log\frac{1}{|G(j\omega_{pc})H(j\omega_{pc})|}, (dB)$$

where, ω_{gc} and ω_{pc} stand for gain and phase crossover frequency respectively.

$$|G(j\omega_{gc})H(j\omega_{gc})| = 1 \Rightarrow \omega_{gc}$$

$$arg\{G(j\omega_{pc})H(j\omega_{pc})\} = 180^o \Rightarrow \omega_{pc}$$

PROCEDURE FOR INVESTIGATING STABILITY USING NYQUIST CRITERION

The following procedure can be followed to investigate the stability of closed loop system from the knowledge of open loop system, using Nyquist stability criterion.

1. Choose a Nyquist contour as shown in figure, which encloses the entire right half s-plane except the singular points. The Nyquist contour encloses all the right half s-plane poles and zeros of G(s)H(s). [The poles on imaginary axis are singular points and so they are avoided by taking a detour around it as shown in figures.
2. The Nyquist contour should be mapped in the G(s)H(s)-plane using the function G(s)H(s) to determine the encirclement -1 + j0 point in the G(s)H(s)-plane. The

Nyquist contour of the figure can be divided into four sections $C_1.C_2.C_3$ and C_4. The mapping of the four sections in the G(s)H(s)-plane can be carried sectionwise and then combined together to get entire G(s)H(s)-contour.

3. In section C_1, the value of ω varies from 0 to + infinite. The mapping of section C_1 is obtained by letting s = jω in G(s)H(s) and varying ω from 0 to + infinite.

   The locus of G(jω)H(jω) as ω is varied from 0 to + infinite will be the G(s)H(s)-contour in G(s)H(s)-plane corresponding to section C_1 in s-plane. This locus is the plot of G(jω)H(jω). There are three ways of mapping this section of G(s)H(s)-contour, they are,

   (i) Calculate the values of G(jω)H(jω) for various values of ω and sketch the actual locus of G(jω)H(jω).

   (or)

   (ii) Separate the real part and imaginary part of G(jω)H(jω). Equate the imaginary part to zero, to find the frequency at which the G(jω)H(jω) locus crosses real axis (to find phase crossover frequency). Substitute this frequency on real part and find the crossing point of the locus on real axis. Sketch the approximate locus of G(jω)H(jω) from the knowledge of type number and order of the system (or from the value of G(jω)H(jω) at ω = 0 and ω = infinite).

   (or)

   (iii) Separate the magnitude and phase of G(jω)H(jω). Equate the phase of G(jω)H(jω) to -180º and solve for ω. This value of ω is the phase crossover frequency and the magnitude at this frequency is the crossing point on real axis. Sketch the approximate root locus as mentioned in method (ii).

4. The section C_2 of Nyquist contour has a semicircle of infinite radius. Therefore, every point on section C_2 has infinite magnitude but the argument varies from $+\pi/2$ to $-\pi/2$. Consider the loop transfer function in time constant form and with y number of poles at origin, as shown below. Let G(s)H(s) has m zeros & n poles including poles at origin. For practical systems, n>m. From the above two equations we can conclude that the section C_2 of Nyquist contour in s-plane is mapped as circles/circular are around origin with radius tending to zero in the G(s)H(s)-plane.

5. In section C3, the value of ω varies from $-\infty$ to 0. The mapping of section C3 is obtained by letting $s=j\omega$ in G(s)H(s) and varying ω from $-\infty$ to 0. The locus of $G(j\omega)H(j\omega)$ as ω is varied from $-\infty$ to 0 will be the G(s)H(s)-contour in G(s)H(s)-plane corresponding to section C3 in s-plane. This locus is the inverse polar plot of $G(j\omega)H(j\omega)$. The inverse polar plot is given by the mirror image of polar plot with respect to real axis.
6. The section C4 of Nyquist contour has a semicircle of zero radius. Therefore, every point on semicircle has zero magnitude but the argument varies from $-\pi/2$ to $\pi/2$. Hence the mapping of section C4 from s-plane to G(s)H(s)-plane can be obtained by letting in G(s)H(s) and varying θ from $-\pi/2$ to $\pi/2$.

PERFORMANCE CRITERIA

For ordinary random inputs (i.e. inputs such that the error E is a stationary random function of time t), it is usual to adopt the mean -square- error as the performance criterion. This is the analogue of integral- square-error for simple transient inputs.

4.1 CONCEPT OF STATE VARIABLES

State space analysis is an excellent method for the design and analysis of control systems. The conventional and old method for the design and analysis of control systems is the transfer function method. The transfer function method for design and analysis had many drawbacks.

Drawbacks of transfer function model analysis:

a. Transfer function is defined under zero initial conditions
b. Transfer function is applicable to linear time invariant systems
c. Transfer function analysis is restricted to single input and single output systems
d. Does not provide information regarding the internal state of the system

Advantages of state variable analysis:

- It can be applied to linear system
- It can be applied to non-linear system
- It can be applied to time varying system
- It can be applied to time invariant system
- It can be applied to multiple input multiple output system
- Its gives idea about the internal state of the system

A state variable is one of the set of variables that are used to describe the mathematical "state" of a dynamical system. Intuitively, the state of a system describes enough about the system to determine its future behaviour in the absence of any external forces affecting the system. The state variable analysis can be applied for any type of systems. In this method of analysis, it is not necessary that the state variables represent physical quantities of the system, but variables that do not represent physical quantities and those that are neither measurable nor observable may be chosen as state variables.

STATE SPACE FORMULATION

State:

The state of a dynamic system is the minimal set of variables called state variables such that the knowledge of these variables at time t = t_0 (initial condition), together with the knowledge of input for $t \geq t_0$, completely determines the behaviour of the system for any time $t > t_0$. (or) A set of variables which describes the system at any time instant are called state variables. In the state variable formulation of a system, in general, a system consists of m-inputs, p-outputs and n-state variables. The state space representation of the system may be visualized as shown in figure 4.1.1.

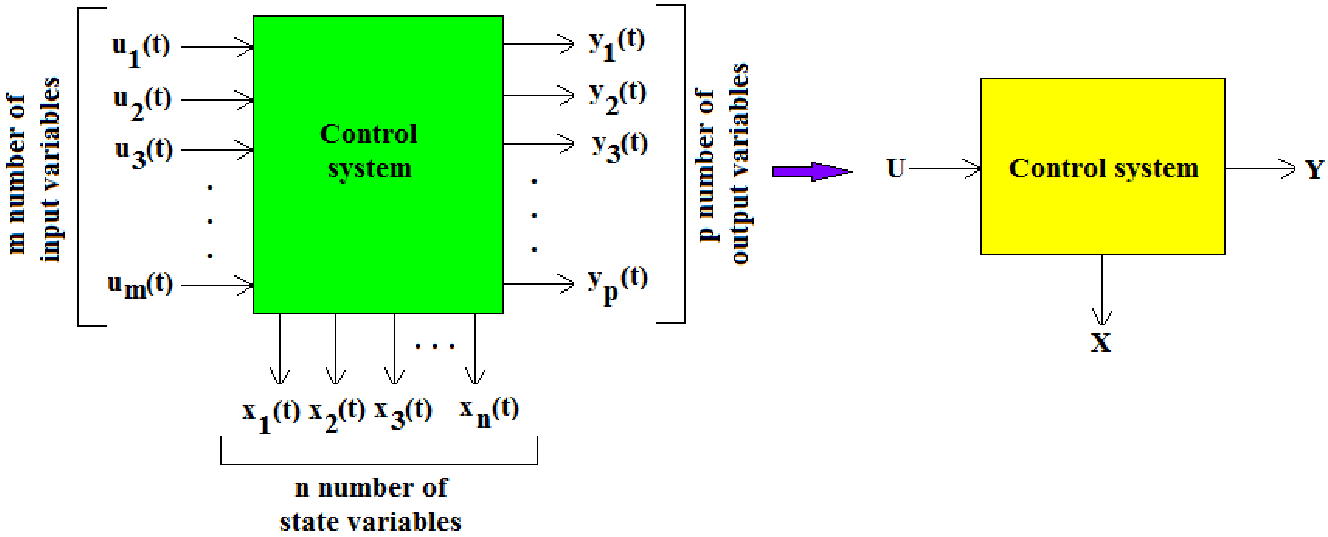


Figure 4.1.1 State space representation of a system

[Source: "Control Systems" by A. Nagoor Kani, Page: 5.2]

Let us consider a multi input & multi output (MIMO) system is having

m inputs: $u_1(t), u_2(t), \ldots\ldots, u_m(t)$

p number of outputs: $y_1(t), y_2(t), \ldots\ldots, y_p(t)$

n number of state variables: $x_1(t), x_2(t), \ldots\ldots, x_n(t)$

The different variables may be represented by the vectors (column matrix) as shown below:

Input vector

$$U(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_m(t) \end{bmatrix}$$

Output vector

$$Y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_p(t) \end{bmatrix}$$

State variable vector

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

State vector:

If n state variables are needed to completely describe the behaviour of a given system, then these n state variables can be considered the n components of a vector X. Such a vector is called a state vector.

State space:

The n-dimensional space whose co-ordinate axes consists of the x_1 axis, x_2 axis……….. x_n axis, where x_1, x_2, ..., x_n are state variables is called a state space.

4.2 STATE MODELS FOR LINEAR AND TIME INVARIANT SYSTEMS

State model is given by state and output equation

State equation:

$$\dot{X}(t) = AX(t) + BU(t)$$

Output equation:

$$Y(t) = CX(t) + DU(t)$$

where,

A is state matrix of size (n x n)

B is the input matrix of size (n x m)

C is the output matrix of size (p x n)

D is the direct transmission matrix of size (p x m)

X(t) is the state vector of size (n x 1)

Y(t) is the output vector of size (p x 1)

U(t) is the input vector of size (m x 1)

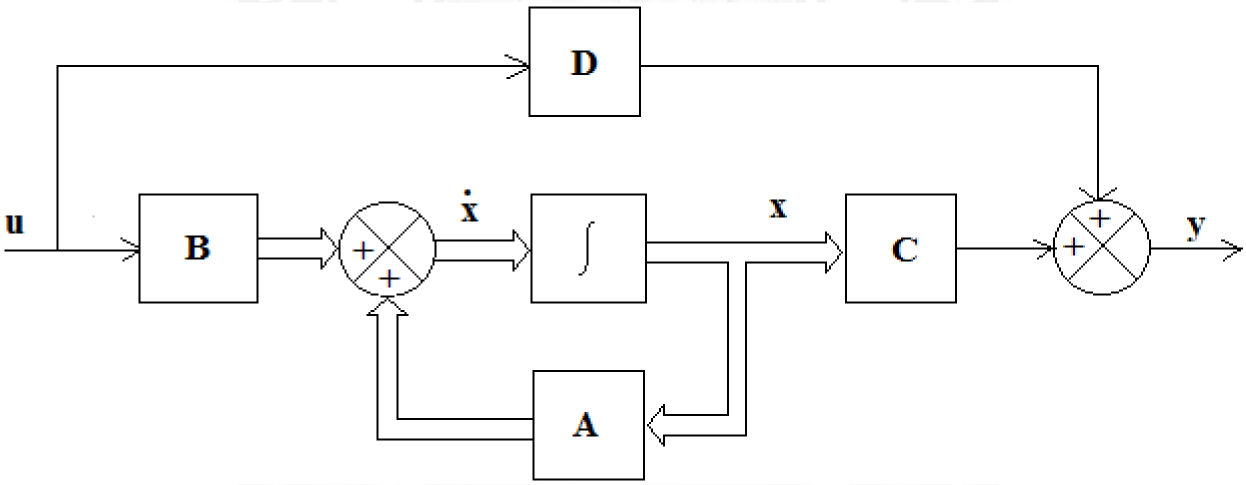


Figure 4.2.1 State space model diagram

[Source: "Modern Control Engineering" by Katsuhiko Ogata, Page: 828]

STATE SPACE REPRESENTATION USING PHYSICAL VARIABLES

In state-space modelling of systems, the choice of state variables is arbitrary. One of the possible choices of state variables. The physical variables of electrical systems are current or voltage in the R, L and C elements. The physical variables of mechanical systems are displacement, velocity and acceleration. The advantages of choosing the physical variables (or quantities) of the system as state variables are the following,

1. *The state variables can be utilized for the purpose of feedback*
2. *The implementation of design with state variable feedback becomes straight forward*
3. *The solution of state equation gives time variation of variables which have direct relevance to the physical system.*

The drawback in choosing the physical quantities as state variables is that the solution of state equation may be a difficult task. In state space modelling using physical variables, the state equations are obtained from a basic model of the system which is developed using the fundamental elements of the system.

Electrical System

The basic model of an electrical system can be obtained by using the fundamental elements Resistor, Capacitor and Inductor. Using these elements, the electrical network or equivalent circuit of the system is drawn. Then the differential equations governing the electrical systems can be formed by writing Kirchhoff's Current Law equations by choosing various nodes in the network or Kirchhoff's Voltage Law by choosing various closed path in the network. A minimal number of state variables are chosen for obtaining the state model of the system. The best choice of state variables in electrical system are currents and voltages in energy storage elements. The energy storage elements are inductance and capacitance. The physical variables in the differential equations are replaced by state variables and the equations are rearranged as first order differential equations. These set of first order equations constitutes the state equation of the system. The inputs to the system are voltage sources or current sources. The outputs in electrical system are usually voltages or currents in energy dissipating elements. The resistance is energy dissipating element in electrical network. In general, the output variables can be any voltage or current in the network.

Mechanical Translational System

The basic model of mechanical translational system can be obtained by using three basic elements; mass, spring and dash-pot. When a force is applied to a mechanical translational system, it is opposed by opposing forces due to mass, friction and elasticity of the system. The forces acting on a body are governed by Newton's second law of motion. The differential equations governing the system are obtaining by writing force balance equations at various nodes in the system. A node is a meeting points of elements.

Guidelines to form the state model of mechanical translational systems

1. For each node in the system one differential equation can be framed by equating the sum of applied forces to the sum of opposing forces. Generally, the nodes are mass elements of the system, but in some cases the nodes may be without mass element.
2. Assign a displacement to each node and draw a free body diagram for each node. The free body diagram is obtained by drawing each mass of node separately and them marking all the forces acting on it.
3. In the free body diagram, the opposing forces due to mass, spring and dash –pot are always act in a direction opposite to applied force. The displacement, velocity and acceleration will be in the direction of applied force or in the direction opposite to that of opposing force.
4. For each free body diagram write one differential equation by equating the sum of applied forces to the sum of opposing forces.
5. Choose a minimum number of state variables. The choice of state variables are displacement, velocity or acceleration.
6. The physical variables in differential equations are replaced by state variables and the equations are rearranged as first order differential equations. These set of first order equations constitute the state equation of the system.
7. The inputs are the applied forces and the outputs are the displacement, velocity or acceleration of the desired nodes.

Mechanical Rotational System

The basic model of mechanical rotational system can be obtained by using three basic elements moment of inertia of mass, rotational dash-pot and rotational spring. When a torque is applied to a mechanical rotational system, it is opposed by opposing torques due to moment of inertia, friction and elasticity of the system. The torque acting on a body are governed by Newton's second law of motion. The differential equations governing the system are obtained by writing torque balance equations at various nodes in the system. A node is a meeting point of elements.

Guidelines to form the state model of mechanical rotational systems

1. For each node in the system one differential equation can be framed by equating the sum of applied torques to the sum of opposing torques. Generally, the nodes are mass elements of the system, but in some cases the nodes may be without mass element.
2. Assign an angular displacement to each node and draw a free body diagram for each node. The free body diagram is obtained by drawing each mass of node separately and them marking all the torques acting on it.
3. In the free body diagram, the opposing torques due to mass of inertia, spring and dashpot are always act in a direction opposite to applied force. The angular displacement, velocity and acceleration will be in the direction of applied torque or in the direction opposite to that of opposing torque.
4. For each free body diagram write one differential equation by equating the sum of applied torques to the sum of opposing torques.
5. Choose a minimum number of state variables. The choice of state variables are angular displacement, velocity or acceleration.
6. The physical variables in differential equations are replaced by state variables and the equations are rearranged as first order differential equations. These set of first order equations constitute the state equation of the system.
7. The inputs are the applied torques and the outputs are the angular displacement, velocity or acceleration of the desired nodes.

STATE SPACE REPRESENTATION USING PHASE VARIABLES

The phase variables are defined as those particular state variables which are obtained from one of the system variables and its derivatives. There are three methods of modelling a system using phase variables. They are,

METHOD 1

Consider the following n[th] order linear differential equation relating the output y(t) to the input u(t) of a system,

$$\dot{y}^{n} + a_1\dot{y}^{n-1} + a_2\dot{y}^{n-2} + \cdots + a_{n-2}\ddot{y} + a_{n-1}\dot{y} + a_n y = bu$$

By choosing the output, y and their derivatives as state variables, we get,

$$x_1 = y$$

$$x_2 = \dot{y}$$

$$x_3 = \ddot{y}$$

$$\vdots$$

$$x_n = \dot{y}^{n-1}$$

$$\dot{x_n} = \dot{y}^{n}$$

$$\dot{x_n} + a_1x_n + a_2x_{n-1} + \cdots + a_{n-2}x_3 + a_{n-1}x_2 + a_nx_1 = bu$$

$$\dot{x_n} = -a_1x_n - a_2x_{n-1} - \cdots - a_{n-2}x_3 - a_{n-1}x_2 - a_nx_1 + bu$$

The state equations of the system are

$$\dot{x_1} = x_2$$

$$\dot{x_2} = x_3$$

$$\dot{x_3} = x_4$$

$$\vdots$$

$$\dot{x}_{n-1} = x_n$$

$$\dot{x_n} = -a_1x_n - a_2x_{n-1} - \cdots - a_{n-2}x_3 - a_{n-1}x_2 - a_nx_1 + bu$$

On arranging the above equations in the matrix form, we get,

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x_n} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 1 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & -a_{n-3} & \vdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix} u$$

$$\dot{X} = AX + BU$$

This form of matrix A is known as ***Bush form (or) Companion form***.

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

$$Y = CX$$

METHOD 2

Consider the following n[th] order linear differential equation relating the output y(t) to the input u(t) of a system,

$$\dot{y}^n + a_1\dot{y}^{n-1} + a_2\dot{y}^{n-2} + \cdots + a_{n-2}\ddot{y} + a_{n-1}\dot{y} + a_n y = bu$$

Let n = m = 3

$$\dddot{y} + a_1\ddot{y} + a_2\dot{y} + a_3 y = b_0\dddot{u} + b_1\ddot{u} + b_2\dot{u} + b_3 u$$

On taking Laplace transform with zero initial conditions, we get,

$$s^3Y(s) + a_1s^2Y(s) + a_2sY(s) + a_3Y(s)$$
$$= b_0s^3U(s) + b_1s^2U(s) + b_2sU(s) + b_3U(s)$$

$$[s^3 + a_1s^2 + a_2s + a_3]Y(s) = [b_0s^3 + b_1s^2 + b_2s + b_3]U(s)$$

$$\frac{Y(s)}{U(s)} = \frac{[b_0s^3 + b_1s^2 + b_2s + b_3]}{[s^3 + a_1s^2 + a_2s + a_3]} = \frac{s^3[b_0 + \frac{b_1}{s} + \frac{b_2}{s^2} + \frac{b_3}{s^3}]}{s^3[1 + \frac{a_1}{s} + \frac{a_2}{s^2} + \frac{a_3}{s^3}]} = \frac{[b_0 + \frac{b_1}{s} + \frac{b_2}{s^2} + \frac{b_3}{s^3}]}{1 - [-\frac{a_1}{s} - \frac{a_2}{s^2} - \frac{a_3}{s^3}]}$$

From Mason's gain formula, the transfer function of the system is given by,

$$T(s) = \frac{1}{\Delta}\sum_K P_K\,\Delta_K$$

where, P_K – path gain of Kth forward path

Δ = 1 – (sum of loop gain of all individual loops) + (sum of gain products of all possible combinations of two non-touching loops) – ……..

Δ_K = Δ for that part of the graph which is not touching Kth forward path

The transfer function of the system with four forward paths and three feedback loops (touching each other) is given by,

$$T(s) = \frac{P_1 + P_2 + P_3 + P_4}{1 - (P_{11} + P_{12} + P_{13})}$$

By comparing the above equations,

$$P_1 = b_0;\ P_2 = \frac{b_1}{s};\ P_3 = \frac{b_2}{s^2};\ P_4 = \frac{b_3}{s^3};\ P_{11} = -\frac{a_1}{s};\ P_{12} = -\frac{a_2}{s^2};\ P_{13} = -\frac{a_3}{s^3}$$

On arranging the above equations in the matrix form, we get,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} -a_1 & 1 & 0 & 0 & \vdots & 0 \\ -a_2 & 0 & 1 & 0 & \vdots & 0 \\ -a_3 & 0 & 0 & 1 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_{n-1} & 0 & 0 & 0 & \vdots & 1 \\ -a_n & 0 & 0 & 0 & \vdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 - a_1 b_0 \\ b_2 - a_2 b_0 \\ b_3 - a_3 b_0 \\ \vdots \\ b_{n-1} - a_{n-1} b_0 \\ b_n - a_n b_0 \end{bmatrix} u$$

$$\dot{X} = AX + BU$$

$$y = [1 \quad 0 \quad 0 \quad \cdots \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u$$

$$Y = CX + DU$$

METHOD 3

Consider the following n[th] order linear differential equation relating the output y(t) to the input u(t) of a system,

$$\dot{y}^n + a_1 \dot{y}^{n-1} + a_2 \dot{y}^{n-2} + \cdots + a_{n-2} \ddot{y} + a_{n-1} \dot{y} + a_n y = bu$$

Let n = m = 3

$$\dddot{y} + a_1 \ddot{y} + a_2 \dot{y} + a_3 y = b_0 \ddot{u} + b_1 \ddot{u} + b_2 \dot{u} + b_3 u$$

On taking Laplace transform with zero initial conditions, we get,

$$s^3 Y(s) + a_1 s^2 Y(s) + a_2 s Y(s) + a_3 Y(s)$$
$$= b_0 s^3 U(s) + b_1 s^2 U(s) + b_2 s U(s) + b_3 U(s)$$

$$[s^3 + a_1 s^2 + a_2 s + a_3] Y(s) = [b_0 s^3 + b_1 s^2 + b_2 s + b_3] U(s)$$

$$\frac{Y(s)}{U(s)} = \frac{[b_0 s^3 + b_1 s^2 + b_2 s + b_3]}{[s^3 + a_1 s^2 + a_2 s + a_3]}$$

Let,

$$\frac{Y(s)}{U(s)} = \frac{Y(s)}{X_1(s)} \cdot \frac{X_1(s)}{U(s)}$$

$$\frac{X_1(s)}{U(s)} = \frac{1}{[s^3 + a_1 s^2 + a_2 s + a_3]}$$

$$\frac{Y(s)}{X_1(s)} = [b_0 s^3 + b_1 s^2 + b_2 s + b_3]$$

State Equation

On cross multiplying the equation, we get,

$$X_1(s)[s^3 + a_1 s^2 + a_2 s + a_3] = U(s)$$

$$s^3 X_1(s) + a_1 s^2 X_1(s) + a_2 s X_1(s) + a_3 X_1(s) = U(s)$$

$$\dddot{x}_1 + a_1 \ddot{x}_1 + a_2 \dot{x}_1 + a_3 x_1 = u$$

Let the state variable be x_1, x_2, x_3.

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

$$x_3 = \ddot{x}_1 = \dot{x}_2$$

$$\dot{x}_3 = \dddot{x}_1$$

On substituting the state variables, we get,

$$\dot{x}_3 + a_1 x_3 + a_2 x_2 + a_3 x_1 = u$$

The state equations are

$$\dot{x}_1 = x_2 \, ; \, \dot{x}_2 = x_3$$

$$\dot{x}_3 = -a_1 x_3 - a_2 x_2 - a_3 x_1 + u$$

Output Equation

On cross multiplying the equation, we get,

$$Y(s) = [b_0 s^3 X_1(s) + b_1 s^2 X_1(s) + b_2 s X_1(s) + b_3 X_1(s)]$$

Taking inverse Laplace transform, we get,

$$y = b_0 \dddot{x}_1 + b_1 \ddot{x}_1 + b_2 \dot{x}_1 + b_3 x_1$$

On substituting the state variables, we get,

$$y = b_0 \dot{x}_3 + b_1 x_3 + b_2 x_2 + b_3 x_1$$

Substituting $\dot{x}_3 = -a_1 x_3 - a_2 x_2 - a_3 x_1 + u$, we get,

$$y = b_0(-a_1 x_3 - a_2 x_2 - a_3 x_1 + u) + b_1 x_3 + b_2 x_2 + b_3 x_1$$

$$y = (b_3 - a_3 b_0) x_1 + (b_2 - a_2 b_0) x_2 + (b_1 - a_1 b_0) x_3 + b_0 u$$

Framing the state and output equation in matrix form, we get,

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x_n} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 1 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & -a_{n-3} & \vdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u$$

$$\dot{X} = AX + BU$$

This form of matrix A is known as ***Bush form (or) Companion form***.

$$y = [b_n - a_n b_0 \quad b_{n-1} - a_{n-1} b_0 \quad \cdots \quad \cdots \quad b_2 - a_2 b_0 \quad b_1 - a_1 b_0] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u$$

$$Y = CX$$

STATE SPACE REPRESENTATION USING CANONICAL VARIABLES

In canonical form (or normal form) of state model, the system matrix A will be a diagonal matrix. The elements on the diagonal are the poles of the transfer function of the system. By partial fraction expansion, the transfer function Y(s)/U(s) of the n^{th} order system can be expressed as,

$$\frac{Y(s)}{U(s)} = b_0 + \frac{C_1}{s+\lambda_1} + \frac{C_2}{s+\lambda_2} + \cdots + \frac{C_n}{s+\lambda_n}$$

where, C_1, C_2, C_3, ….., C_n are residues and λ_1, λ_2, λ_3, ….., λ_n are roots of denominator polynomial (or poles of the system).

$$\frac{Y(s)}{U(s)} = b_0 + \frac{C_1/s}{1+\lambda_1/s} + \frac{C_2/s}{1+\lambda_2/s} + \cdots + \frac{C_n/s}{1+\lambda_n/s}$$

$$Y(s) = b_0 U(s) + \frac{C_1/s}{1+\lambda_1/s}U(s) + \frac{C_2/s}{1+\lambda_2/s}U(s) + \cdots + \frac{C_n/s}{1+\lambda_n/s}U(s)$$

The state equation can be framed as,

$$\dot{x}_1 = -\lambda_1 x_1 + u$$

$$\dot{x}_2 = -\lambda_2 x_2 + u$$

$$\vdots$$

$$\dot{x}_n = -\lambda_n x_n + u$$

The output equation can be framed as,

$$y = C_1 x_1 + C_2 x_2 + \cdots + C_n x_n + b_0 u$$

The canonical form of state model in the matrix form is given by,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} -\lambda_1 & 0 & 0 & 0 & \vdots & 0 \\ 0 & -\lambda_2 & 0 & 0 & \vdots & 0 \\ 0 & 0 & -\lambda_3 & 0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & -\lambda_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} C_1 & C_2 & C_3 & \cdots & C_{n-1} & C_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u$$

JORDAN CANONICAL FORM

$$A = J = \begin{bmatrix} -\lambda_1 & 0 & 0 & 0 & \vdots & 0 \\ 0 & -\lambda_1 & 0 & 0 & \vdots & 0 \\ 0 & 0 & -\lambda_1 & 0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & -\lambda_n \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}$$

$$\dot{Z} = JZ + \tilde{B}U$$

$$Y = \tilde{C}Z + DU$$

where,

$$J = M^{-1}AM; \quad \tilde{B} = M^{-1}B; \quad \tilde{C} = CM$$

4.3 SOLUTION OF STATE AND OUTPUT EQUATION IN CONTROLLABLE CANONICAL FORM

Consider the state equation of a linear time invariant system as,

$$\dot{X}(t) = AX(t) + BU(t)$$

The matrices A and B are constant matrices. This state equation can be of two types,

1. Homogeneous
2. Non-homogeneous

HOMOGENEOUS EQUATION

If A is a constant matrix and input control forces are zero then the equation takes the form

$$\dot{X}(t) = AX(t)$$

Such an equation is called homogeneous equation. The obvious equation is considered if input is zero. In such systems, the driving force is provided by the initial conditions of the system to produce the output. For example, consider a series RC circuit in which a capacitor is initially charged to V volts. The current is the output. Now there is no input control force, i.e., external voltage applied to the system. But the initial voltage on the capacitor drives the current through the system and capacitor starts discharging through the resistance, R. such a system works on the initial conditions without any input applied to it is called homogeneous system.

NON-HOMOGENEOUS EQUATION

If A is a constant matrix and matrix U(t) is non-zero vector i.e. the input control forces are applied to the system then the equation takes normal form as,

$$\dot{X}(t) = AX(t) + BU(t)$$

Such an equation is called non-homogeneous equation. Most of the practical systems require inputs to dive them. Such systems arc nonhomogeneous linear systems. The solution of the state equation is obtained by considering basic method of finding the solution of homogeneous equation.

STATE TRANSITION MATRIX

Properties of State Transition Matrix

1. $\phi(0) = e^{A\times 0} = I(unit\ matrix)$
2. $\phi(t) = e^{At} = (e^{-At})^{-1} = [\phi(-t)]^{-1}$

   or $\phi^{-1}(t) = \phi(-t)$
3. $\phi(t_1 + t_2) = e^{A(t_1+t_2)} = e^{At_1}e^{At_2} = \phi(t_1)\phi(t_2)$

Computation of State transition matrix

The state transition matrix, e^{At} can be computed by any one of the following two methods:

Method 1: Computation of e^{At} using matrix exponential

If the system matrix 'A' is an (n×n) square matrix, then each of these exponentials is an (n×n) square matrix of time functions, and one of the consequences of a theorem developed in linear algebra, known as the Cayley-Hamilton theorem, shows that such a matrix may be expressed as an $(n-1)^{st}$ degree polynomial in the matrix A.

That is,

$$e^{At} = I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots + \frac{1}{i!}A^it^i$$

where, e^{At} – State transition matrix of order n x n

A – System matrix of order n x n

I – Unit matrix of order n x n

Method 2: Computation of e^{At} using Laplace transform

The theorem also states that the equation remains an equality if I is replaced by unity and A is replaced by any one of the scalar roots sI of the nth-degree scalar equation, det(sI-A) = 0. The expression det(sI-A) indicates the determinant of the matrix (sI-A). This determinant is an nth-degree polynomial ins. Let us assume that then roots are all different. This equation is called the characteristic equation of the matrix a, and the values of s which are the roots of the equation are known as the eigen values of A.

Consider the state equation without input vector,

$$\dot{X}(t) = AX(t)$$

On taking Laplace transform, we get,

$$sX(s) - X(0) = AX(s)$$

$$sX(s) - AX(s) = X(0)$$

$$sIX(s) - AX(s) = X(0)$$

$$(sI - A)X(s) = X(0)$$

Pre-multiplying both sides by $(sI-A)^{-1}$,

$$X(s) = (sI - A)^{-1}X(0)$$

On taking inverse Laplace transform,

$$x(t) = L^{-1}[(sI - A)^{-1}]x(0)$$

On comparing with solution of state equation,

$$e^{At} = L^{-1}[(sI - A)^{-1}]$$

Also,

$$e^{At} = \phi(t)$$

where,

$$\phi(s) = (sI - A)^{-1}$$

which is the resolvent matrix.

Consider the state equation with input vector,

$$\dot{X}(t) = AX(t) + BU(t)$$

On taking Laplace transform, we get,

$$sX(s) - X(0) = AX(s) + BU(s)$$

$$sIX(s) - AX(s) = X(0) + BU(s)$$

$$(sI - A)X(s) = X(0) + BU(s)$$

Pre-multiplying both sides by $(sI-A)^{-1}$,

$$X(s) = (sI - A)^{-1}X(0) + (sI - A)^{-1}BU(s)$$

$$X(s) = \phi(s)X(0) + \phi(s)BU(s)$$

On taking inverse Laplace transform,

$$x(t) = \phi(t)x(0) + L^{-1}[\phi(s)BU(s)]$$

Solution of output equation by Laplace Transform

$$Y(s) = CX(s) + DU(s)$$

$$y(t) = L^{-1}[CX(s) + DU(s)]$$

CONTROLLABLE CANONICAL FORM (CCF)

Probably the most straightforward method for converting from the transfer function of a system to a state space model is to generate a model in "controllable canonical form."

Consider a system defined by,

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}\dot{y} + a_n y = b_0 u^{(n)} + b_1 u^{(n-1)} + \cdots + b_{n-1}\dot{u} + b_n u$$

where u is the control input and y is the output. It can be written as,

$$\frac{Y(s)}{U(s)} = \frac{[b_0 s^n + b_1 s^{n-1} + \cdots + b_{n-1}s + b_n]}{[s^n + a_1 s^{n-1} + a_{n-1}s + a_n]}$$

Controllable canonical form of this system is given by,

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x_n} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 1 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & -a_{n-3} & \vdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} b_n - a_n b_0 & b_{n-1} - a_{n-1} b_0 & \cdots & \cdots & b_2 - a_2 b_0 & b_1 - a_1 b_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u$$

OBSERVABLE CANONICAL FORM

The observable canonical form of the state-space representation of this system is given by

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x_n} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \vdots & 0 & -a_n \\ 1 & 0 & 0 & \vdots & 0 & -a_{n-1} \\ 0 & 1 & 0 & \vdots & 0 & -a_{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & 0 & -a_2 \\ 0 & 0 & 0 & \vdots & 1 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} b_n - a_n b_0 \\ b_{n-1} - a_{n-1} b_0 \\ b_{n-2} - a_{n-2} b_0 \\ \vdots \\ b_2 - a_2 b_0 \\ b_1 - a_1 b_0 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u$$

DIAGONAL CANONICAL FORM

There are cases where the dominator polynomial involves only distinct roots. For the distinct root case, we can write the equation in the form of

$$\frac{Y(s)}{U(s)} = \frac{[b_0 s^n + b_1 s^{n-1} + \cdots + b_{n-1} s + b_n]}{(s + p_1)(s + p_2)\ldots(s + p_n)} = b_0 + \frac{c_1}{s + p_1} + \frac{c_2}{s + p_2} + \cdots + \frac{c_n}{s + p_n}$$

The diagonal canonical form of the state-space representation of this system is given by

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x_n} \end{bmatrix} = \begin{bmatrix} -p_1 & 0 & 0 & 0 & \vdots & 0 \\ 0 & -p_2 & 0 & 0 & \vdots & 0 \\ 0 & 0 & -p_3 & 0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & -p_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u$$

4.4 CONCEPTS OF CONTROLLABILITY AND OBSERVABILITY

CONCEPT OF CONTROLLABILITY

A system is said to be completely controllable, if it is possible to transfer the system state from any initial state x(t0) to any desired state x(t) in specified finite time by a control vector u(t).

If any of the state variable is independent of the control u(t), there would be no way of driving this particular state variable to desired state in finite time by means of control effort. Therefore, this particular state is said to be uncontrollable. As long as there is at least one uncontrollable state, the system is said to be not completely controllable or 'uncontrollable'.

Consider a single input, linear time invariant system:

$$\dot{X}(t) = AX(t) + BU(t)$$

Let the initial system state be x(0) and the final state be $x(t_f)$. The system is controllable if it is possible to construct a control signal, which in finite time interval $0 < t <= t_f$, will transfer the system state from x(0) to $x(t_f)$. The above equation is completely controllable if and only if the rank of the composite matrix is n.

$$Q_C = [B \vdots AB \vdots \cdots \vdots A^{n-1}B]$$

Since only matrices A and B are involved, we may say that the pair (A;B) is controllable if rank of Q_C is n.

CONCEPT OF OBSERVABILITY

A system is said to be completely observable, if every state $x(t_0)$ can be completely identified by measurement of outputs y(t) over a finite time interval. Given a LTI system that is described by the dynamic equations, the state $x(t_0)$ is said to be observable if given any input u(t), there exists a finite time $t_f >= t_0$ such that knowledge of u(t) for $t_0 <= t < t_f$, matrices A,B,C, & D and the output y(t); for $t_0 <= t < t_f$ are sufficient to determine $x(t_0)$. The necessary and sufficient condition for the system to be completely observable it is necessary and sufficient that the following n x n_p observability matrix has rank of n.

$$Q_O = [C^T \quad A^TC^T \quad (A^2)^TC^T \quad \cdots \quad (A^{n-1})^TC^T]$$

5.1 CHARACTERISTIC EQUATION

The characteristic equation is nothing more than setting the denominator of the closed-loop transfer function to zero. In control theory, there are two main methods of analyzing feedback systems: the transfer function (or frequency domain) method and the state space method. The characteristic equation is the equation which is solved to find a matrix's eigenvalues, also called the characteristic polynomial. Characteristic equation is used to solve linear differential equations. Characteristic equations of auxiliary differential equations are used to solve a partial differential equation.

The properties of transfer function are given below:

- The ratio of Laplace transform of output to Laplace transform of input assuming all initial conditions to be zero.
- The transfer function of a system does not depend on the inputs to the system.
- The system poles and zeros can be determined from its transfer function.

Closed loop transfer function:

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

Characteristic Equation:

$$1 \pm G(s)H(s) = 0$$

5.2 EFFECT OF LAG, LEAD AND LAG-LEAD COMPENSATION ON FREQUENCY RESPONSE

Every control system which has been designed for a specific application should meet certain performance specification. There are always some constraints which are imposed on the control system design in addition to the performance specification. The choice of a plant is not only dependent on the performance specification but also on the size, weight & cost. Although the designer of the control system is free to choose a new plant, it is generally not advised due to the cost & other constraints. Under this circumstance, it is possible to introduce some kind of corrective sub-systems in order to force the chosen plant to meet the given specification. We refer to these sub-systems as compensator whose job is to compensate for the deficiency in the performance of the plant.

Necessary of Compensation

1. In order to obtain the desired performance of the system, we use compensating networks. Compensating networks are applied to the system in the form of feed forward path gain adjustment.
2. Compensate a unstable system to make it stable.
3. A compensating network is used to minimize overshoot.
4. These compensating networks increase the steady state accuracy of the system. An important point to be noted here is that the increase in the steady state accuracy brings instability to the system.
5. Compensating networks also introduces poles and zeros in the system thereby causes changes in the transfer function of the system. Due to this, performance specifications of the system change.

EFFECT OF LAG COMPENSATION ON FREQUENCY RESPONSE

The Lag Compensator is an electrical network which produces a sinusoidal output having the phase lag when a sinusoidal input is applied. The lag compensator circuit in the 's' domain is shown in the following figure.

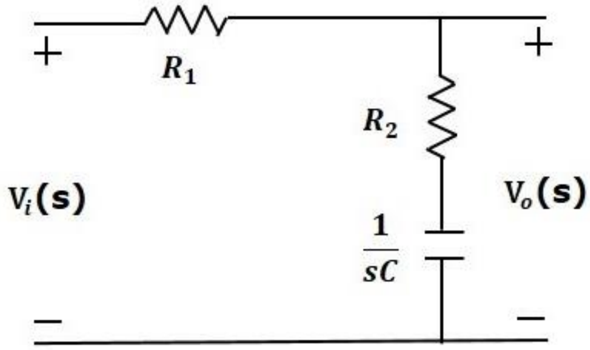


Figure 5.2.1 Electrical lag compensator

[Source: "Control Systems" by A Nagoor Kani, Page: 4.65]

Here, the capacitor is in series with the resistor R_2 and the output is measured across this combination. The transfer function of this lag compensator is

$$\frac{V_o(s)}{V_i(s)} = \frac{1}{\beta}\left(\frac{s+\frac{1}{\tau}}{s+\frac{1}{\beta\tau}}\right)$$

$$\tau = R_2 C$$

$$\beta = \frac{R_1 + R_2}{R_2}$$

$$\beta > 1$$

$$\text{Pole, } s = -\frac{1}{\beta\tau}$$

$$\text{Zero, } s = -\frac{1}{\tau}$$

Let s = jω,

$$\frac{V_o(j\omega)}{V_i(j\omega)} = \frac{1}{\beta}\left(\frac{j\omega+\frac{1}{\tau}}{j\omega+\frac{1}{\beta\tau}}\right)$$

Phase angle,

$$\phi = \tan^{-1}\omega\tau - \tan^{-1}\beta\omega\tau$$

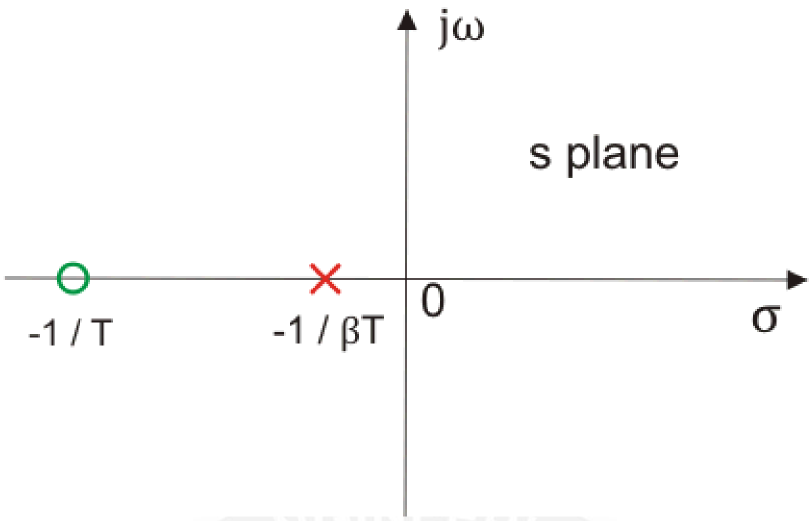


Figure 5.2.2 Pole-zero plot of lag compensator

[Source: "Control Systems" by A Nagoor Kani, Page: 4.65]

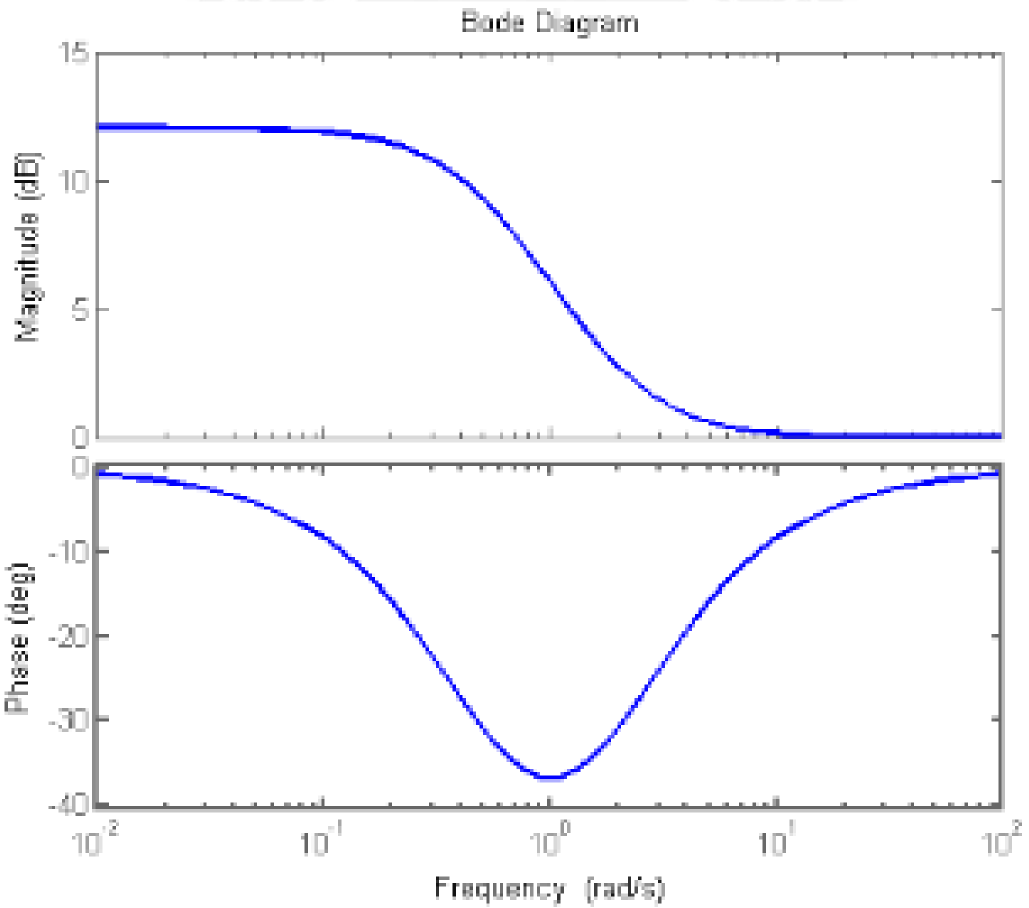


Figure 5.2.3 Bode plot of lag compensator

[Source: "Control Systems" by A Nagoor Kani, Page: 4.67]

The phase of the output sinusoidal signal is equal to the sum of the phase angles of input sinusoidal signal and the transfer function. So, in order to produce the phase lag at the output of this compensator, the phase angle of the transfer function should be negative. This will happen when $\beta > 1$.

Effect of Phase Lag Compensation

1. Gain crossover frequency increases.
2. Bandwidth decreases.
3. Phase margin will be increase.
4. Response will be slower before due to decreasing bandwidth, the rise time and the settling time become larger.

Advantages of Phase Lag Compensation

1. Phase lag network allows low frequencies and high frequencies are attenuated.
2. Due to the presence of phase lag compensation the steady state accuracy increases.

Disadvantages of Phase Lag Compensation

1. Due to the presence of phase lag compensation the speed of the system decreases.

EFFECT OF LEAD COMPENSATION ON FREQUENCY RESPONSE

The lead compensator is an electrical network which produces a sinusoidal output having phase lead when a sinusoidal input is applied. The lead compensator circuit in the 's' domain is shown in the following figure. Lead compensator are used to improve the transient response of a system.

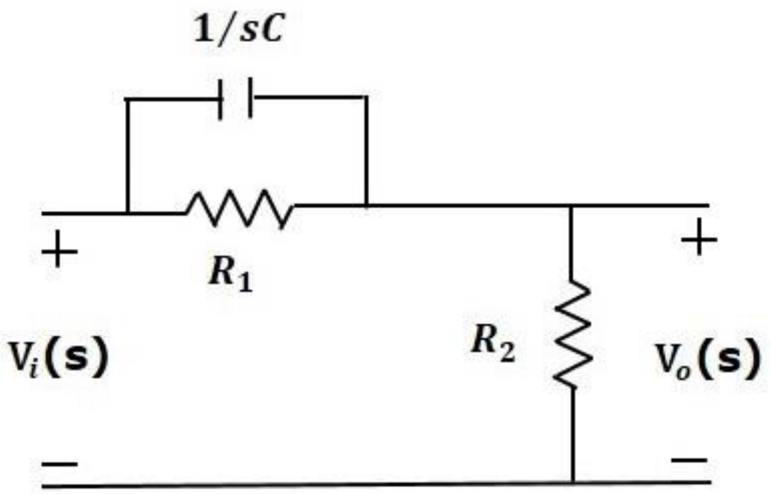


Figure 5.2.4 Electrical lead compensator

[Source: "Control Systems" by A Nagoor Kani, Page: 4.70]

Taking i_2=0 and applying Laplace Transform, we get,

$$\frac{V_o(s)}{V_i(s)} = \frac{R_2(R_1Cs + 1)}{R_1 + R_2 + R_2R_1Cs}$$

$$\frac{V_o(s)}{V_i(s)} = \alpha\left(\frac{\tau s + 1}{\alpha\tau s + 1}\right)$$

$$\tau = R_1C$$

$$\alpha = \frac{R_2}{R_1 + R_2}$$

$$\alpha < 1$$

Pole, $s = -\frac{1}{\alpha\tau}$

Zero, $s = -\frac{1}{\tau}$

Let s = jω,

$$\frac{V_o(j\omega)}{V_i(j\omega)} = \alpha\left(\frac{\tau j\omega + 1}{\alpha\tau j\omega + 1}\right)$$

Phase angle,

$$\phi = \tan^{-1}\omega\tau - \tan^{-1}\alpha\omega\tau$$

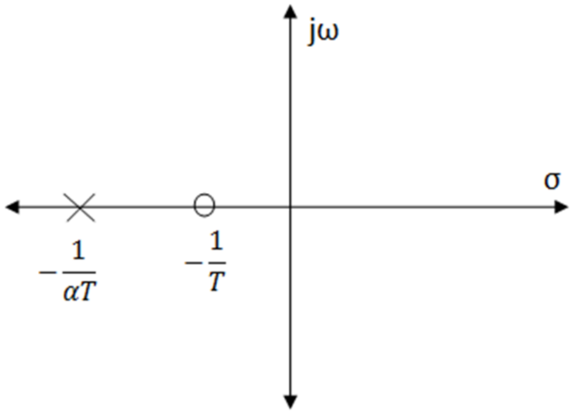


Figure 5.2.5 Pole-zero plot of lead compensator

[Source: "Control Systems" by A Nagoor Kani, Page: 4.69]

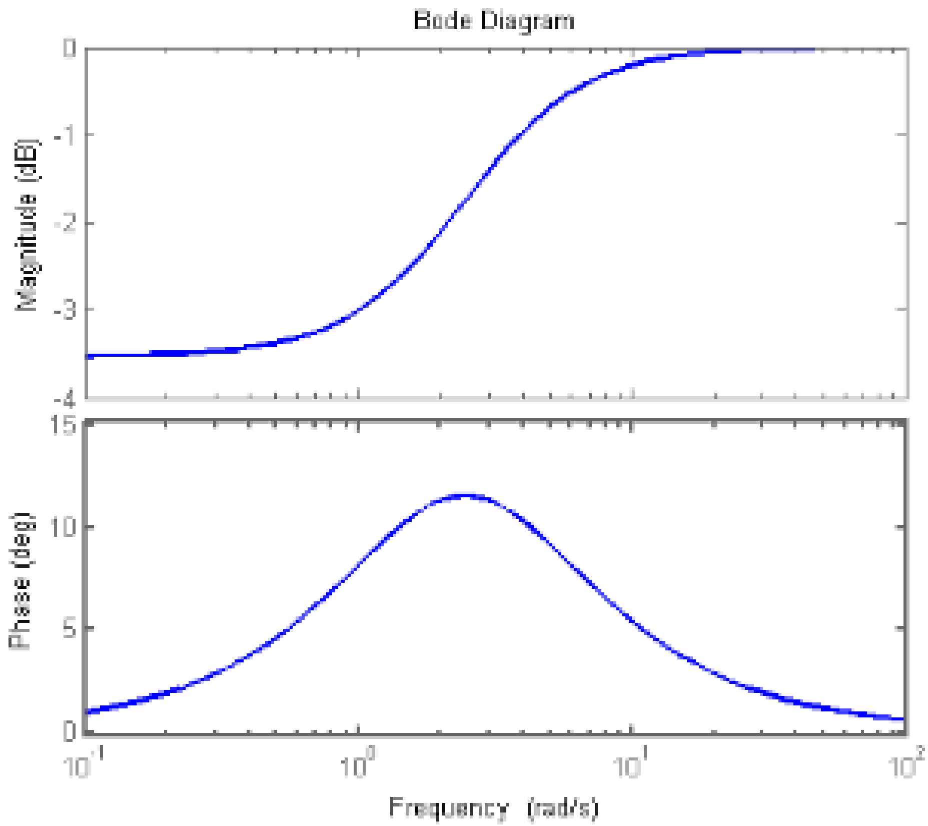


Figure 5.2.6 Bode plot of lead compensator

[Source: "Control Systems" by A Nagoor Kani, Page: 4.71]

The phase of the output sinusoidal signal is equal to the sum of the phase angles of input sinusoidal signal and the transfer function. So, in order to produce the phase lead at the output of this compensator, the phase angle of the transfer function should be positive. This will happen when $0<\alpha<1$. Therefore, zero will be nearer to origin in pole-zero configuration of the lead compensator.

Bode plot of lead compensator

Maximum phase lead occurs at

$$\omega_m = \frac{1}{\tau\sqrt{\alpha}}$$

Let Φ_m = maximum phase lead

$$\sin\phi_m = \frac{1-\alpha}{1+\alpha}$$

$$\alpha = \frac{1-\sin\phi_m}{1+\sin\phi_m}$$

Magnitude at maximum phase lead

$$|G_c(j\omega)| = \frac{1}{\sqrt{\alpha}}$$

Effect of Phase Lead Compensation

1. The velocity constant K_v increases.
2. The slope of the magnitude plot reduces at the gain crossover frequency so that relative stability improves and error decrease due to error is directly proportional to the slope.
3. Phase margin increases.
4. Response becomes faster.

Advantages of Phase Lead Compensation

1. Due to the presence of phase lead network the speed of the system increases because it shifts gain crossover frequency to a higher value.
2. Due to the presence of phase lead compensation maximum overshoot of the system decreases.

Disadvantages of Phase Lead Compensation

1. Steady state error is not improved.

EFFECT OF LAG-LEAD COMPENSATION ON FREQUENCY RESPONSE

Lag-Lead compensator is an electrical network which produces phase lag at one frequency region and phase lead at other frequency region. It is a combination of both the lag and the lead compensators. The lag-lead compensator circuit in the ‘s’ domain is shown in the following figure.

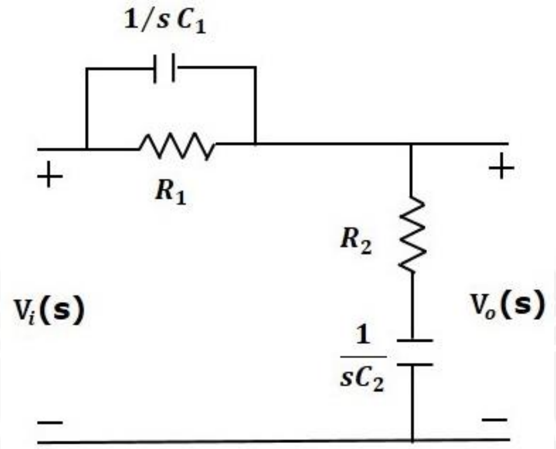


Figure 5.2.7 Electrical lag-lead compensator

[Source: “Control Systems” by A Nagoor Kani, Page: 4.73]

This circuit looks like both the compensators are cascaded. So, the transfer function of this circuit will be the product of transfer functions of the lead and the lag compensators.

$$\frac{V_o(s)}{V_i(s)} = \beta\left(\frac{\tau_1 s + 1}{\beta\tau_1 s + 1}\right)\frac{1}{\alpha}\left(\frac{s + \frac{1}{\tau_2}}{s + \frac{1}{\alpha\tau_2}}\right)$$

We know, αβ=1

$$\frac{V_o(s)}{V_i(s)} = \left(\frac{s + \frac{1}{\tau_1}}{s + \frac{1}{\beta\tau_1}}\right)\left(\frac{s + \frac{1}{\tau_2}}{s + \frac{1}{\alpha\tau_2}}\right)$$

where,

$$\tau_1 = R_1 C_1$$

$$\tau_2 = R_2 C_2$$

Advantages of Phase Lag Lead Compensation

1. Due to the presence of phase lag-lead network the speed of the system increases because it shifts gain crossover frequency to a higher value.
2. Due to the presence of phase lag-lead network accuracy is improved.

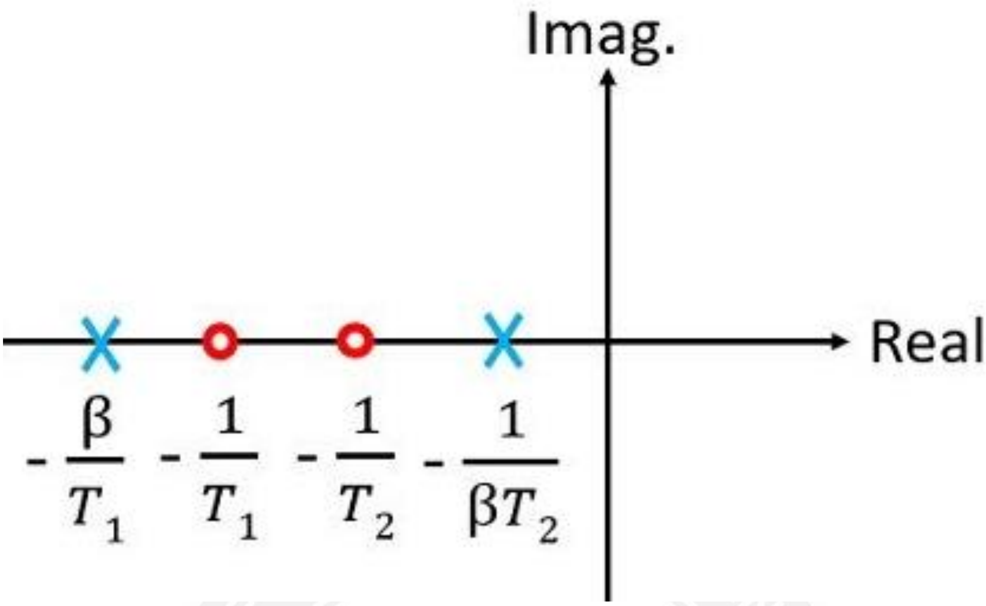


Figure 5.2.8 Pole-zero plot of lag-lead compensator

[Source: "Control Systems" by A Nagoor Kani, Page: 4.73]

5.3 DESIGN OF LAG, LEAD AND LAG LEAD COMPENSATOR USING BODE PLOTS

DESIGN PROCEDURE OF LAG COMPENSATOR USING BODE PLOTS

1. Determine the compensator gain K to meet the steady state error requirement.
2. Draw the Bode plots of KG(s).
3. From the Bode plots, find the frequency ω_g at which the phase of KG(s) is

$$\angle KG(\omega_g) = PM - 180^o + 5^o \sim 10^o$$

4. Calculate β to make ω_g the gain crossover frequency,

$$20\log\beta = 20\log K + 20\log|G(j\omega_g)|$$

5. Choose T to be much greater than 1/ ω_g, for example, T=10/ ω_g.
6. Verify the results using MATLAB.

DESIGN PROCEDURE OF LEAD COMPENSATOR USING BODE PLOTS

1. Draw the Bode plot for the uncompensated system and obtain the current phase margin available.
2. Calculate the phase margin required to meet the damping coefficient or percent overshoot requirement. Don't forget to add some extra phase margin to compensate for imperfections in the controller design (approximately 10 degrees of phase is good).

$$PM = \tan^{-1}\left(\frac{2\zeta}{\sqrt{-2\zeta^2 + \sqrt{1+4\zeta^4}}}\right)$$

$$\zeta \cong \frac{PM}{100}$$

3. Calculate the value of alpha from the following equation. Use the phase margin obtained in Step (4) as the maximum phase value:

$$\alpha = \frac{1-\sin\phi_{max}}{1+\sin\phi_{max}}$$

4. Calculate the gain corresponding to the maximum phase frequency using the equation below. We are going to look for the new phase margin frequency that we want to design for by looking for places where this gain is present on the Bode plot.

$$|G(j\omega_{max})| = \frac{1}{\sqrt{\beta}}$$

5. Find the new maximum phase margin frequency by looking for the point where the uncompensated system's magnitude curve is the negative value of the gain calculated in Step (4).
6. Select the break frequencies, T and beta*T using the maximum frequency equation given below:

$$\omega_{max} = \frac{1}{T\sqrt{\beta}}$$

7. Reset the system gain to adjust for the compensator's gain.
8. Check that the bandwidth still meets design requirements. Simulate the system and repeat the design as necessary.

DESIGN PROCEDURE OF LAG-LEAD COMPENSATOR USING BODE PLOTS

The lag-lead compensator is the analog to the PID controller. The lag-lead compensator can meet multiple design requirements: the lag component reduces high frequency gain, stabilizes the system and meets steady state requirements, while the lead component is used to meet transient response design requirements.

The general equation for this kind of compensator is given below:

$$G_{lag-lead}(s) = G_{lead}(s)G_{lag}(s) = \left(\frac{s+\frac{1}{T_1}}{s+\frac{\gamma}{T_1}}\right)\left(\frac{s+\frac{1}{T_2}}{s+\frac{1}{\gamma T_2}}\right), \gamma > 1$$

(1) Calculate the required bandwidth to meet the transient performance requirement (usually expressed in terms of the settling time, rise time or peak time). Use the equation provided above.

(2) Set the DC gain of the uncompensated system to meet the steady state requirements (this requires use of the Final Value Theorem).

(3) Draw the Bode plot for the uncompensated system and obtain the current phase margin available.

(4) Calculate the phase margin required to meet the damping coefficient or percent overshoot requirement. Don't forget to add some extra phase margin to compensate

for imperfections in the controller design (approximately 10 degrees of phase is good).

(5) Select a new phase margin frequency that is slightly less than the bandwidth.

(6) At this new phase margin frequency, calculate the phase lead required to obtained the phase margin from Step (4). Add some additional phase to adjust for the lag compensator's effects, if you have not already done so in Step (4).

(7) Design the lag compensator. Choose the higher breakpoint frequency as the phase margin frequency divided by 10. A plot of the interaction between beta and the phase margin is used to select beta, but for our purposes, I think that spacing the pole and zero of the lag compensator apart by a factor of 10 is sufficient for our design purposes.

(8) We can calculate gamma as the inverse of beta, so now we are ready to design the lead compensator. Use the equation given below to find T.

(9) Check that the bandwidth still meets design requirements. Simulate the system and repeat the design as necessary.

5.4 EFFECTS OF P, PI, PID MODES OF FEEDBACK CONTROL

PROPORTIONAL CONTROLLER (P-Controller)

The proportional controller is a device that produces a control signal, u(t) proportional to the input error signal, e(t)

$$u(t) \propto e(t)$$

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

where, K_p = Proportional gain or constant

On taking Laplace transform of equation, we get,

$$U(s) = K_p E(s)$$

Transfer function,

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

The equation gives the output of the P-controller for the input E(s) and it is the transfer function of P-controller. The block diagram of the P-controller is shown in the figure 5.4.1.

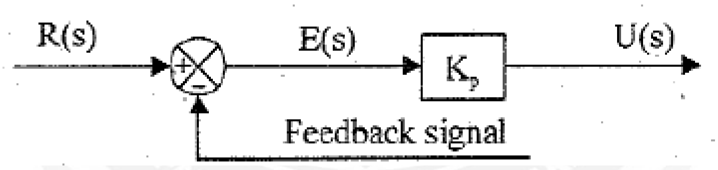


Figure 5.4.1 Block diagram of proportional controller

[Source: "Control Systems" by Nagoor Kani, Page: 2.79]

From the equation, we can conclude that the proportional controller amplifies the error signal by an amount K_p. Also the introduction of the controller on the system increases the loop gain by an amount K_p. The increase in loop gain improves the steady state tracking accuracy, disturbance signal rejection and the relative stability and also makes the system less sensitive to parameter variations. But increasing the gain to very large values mau lead to instability of the system. The drawback in P-controller is that it leads to a constant steady state error.

Example of Electronic P-controller

The proportional controller can be realized by an amplifier with adjustable gain. Either the non- inverting operational amplifier or the inverting operational amplifier followed by sign changer will work as a proportional controller. The op-amp proportional controller is shown in the figures 5.4.2.

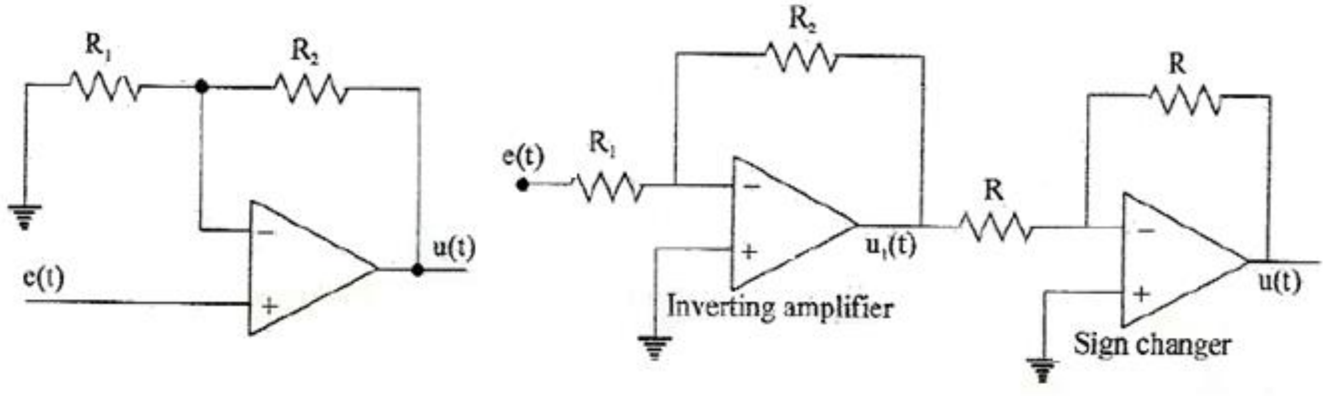


Figure 5.4

5.4.2 P-controller using non-inverting and inverting amplifier

[Source: "Control Systems" by Nagoor Kani, Page: 2.80]

By deriving the transfer function of the controller shown in figures and comparing with the transfer function of P-controller defined by equation, it can be shown that they work as P-controllers.

Analysis of P-controller

In figure 2.8.2, the input e(t) is applied to positive input. By symmetry of op-amp the voltage of negative input is also e(t). Also, we assume an ideal op-amp so that input current is zero. Based on the above assumptions the equivalent circuit of the controller is shown in figure 5.4.3.

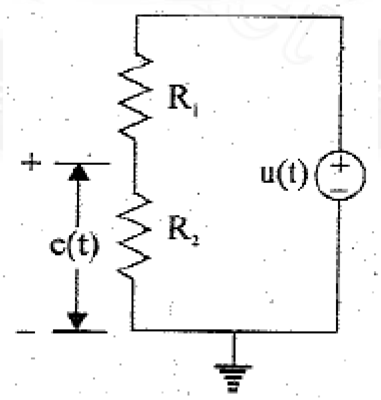


Figure 5.4.3 Equivalent circuit of P-controller

[Source: "Control Systems" by Nagoor Kani, Page: 2.80]

By voltage division rule,

$$e(t) = \frac{R_1}{R_1 + R_2} u(t)$$

On taking Laplace transform of equation we get,

$$\frac{U(s)}{E(s)} = \frac{R_1 + R_2}{R_1}$$

The equation is the transfer function of op-amp P-controller. On comparing, we get,

$$K_p = \frac{R_1 + R_2}{R_1}$$

Therefore, by adjusting the values of R_1 and R_2 the value of gain, K_p can be varied.

Analysis of P-controller

The assumption made in op-amp circuit analysis are,

1. The voltages at both inputs are equal
2. The input current is zero

Based on the above assumptions, the equivalent circuit of op-amp amplifier and sign changer are shown in figure 5.4.4.

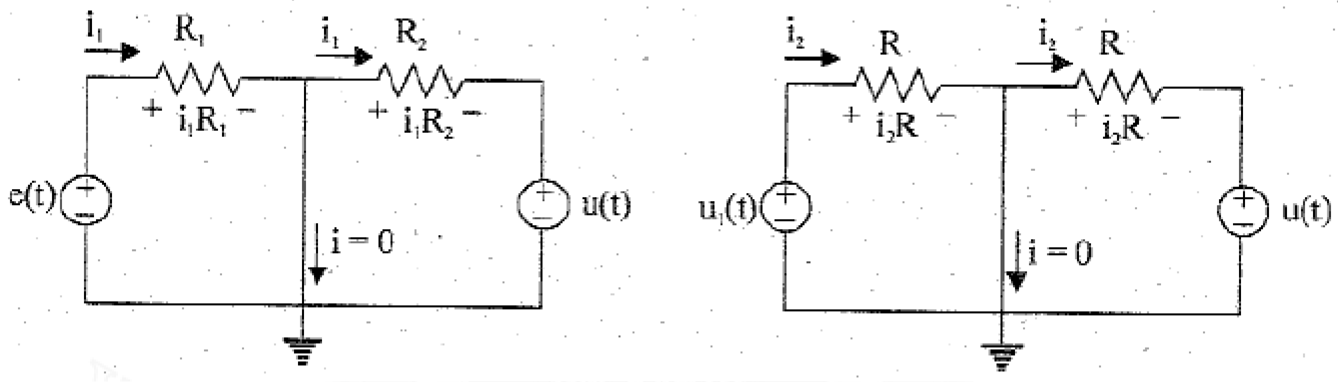


Fig 5.4.4 Equivalent circuit of amplifier and sign changer

[Source: "Control Systems" by Nagoor Kani, Page: 2.81]

From the circuit,

$$e(t) = i_1 R_1$$

$$u_1(t) = -i_1 R_2$$

Substitute for i_1,

$$u_1(t) = -\frac{e(t)}{R_1} R_2$$

Also, from the circuit,

$$u(t) = -i_2 R$$

$$u_1(t) = i_2 R$$

Substitute for i_2,

$$u_1(t) = -u(t)$$

On equating the equations we get,

$$u(t) = \frac{e(t)}{R_1} R_2$$

On taking Laplace transform of equation we get,

$$\frac{U(s)}{E(s)} = \frac{R_2}{R_1}$$

The equation is the transfer function of op-amp P-controller. On the comparing equations, Proportional gain,

$$K_p = \frac{R_2}{R_1}$$

Therefore, by adjusting the values of R_1 and R_2 the value of gain K_p can be varied.

INTEGRAL CONTROLLER (I-Controller)

The integral controller is a device that produces a control signal u(t) which is proportional to integral of the input error signal, e(t).

In I-controller

$$u(t) \propto \int e(t)dt$$

$$u(t) = K_i \int e(t)dt$$

where K_i = integral gain or constant

On taking Laplace transform of equation with zero initial conditions we get,

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

The equation gives the output of the I-controller for the input E(s) and equation is the transfer function of the I-controller, the block diagram of I-controller is shown in the figure 5.4.5.

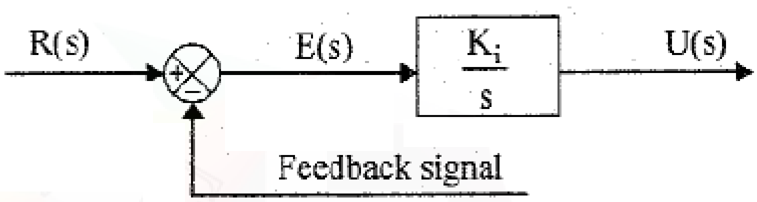


Figure 5.4.5 Block diagram of integral controller

[Source: "Control Systems" by Nagoor Kani, Page: 2.82]

The integral controller removes or reduces the steady error without the need for manual reset. Hence the I-controller is sometimes called automatic reset. The drawback in integral controller is that it may lead to oscillatory response of increasing or decreasing amplitude which is undesirable and the system may become unstable.

Example of electronic I-controller

The integral controller can be realized by an integrator using op-amp followed by a sign changer as shown in figure 2.8.6.

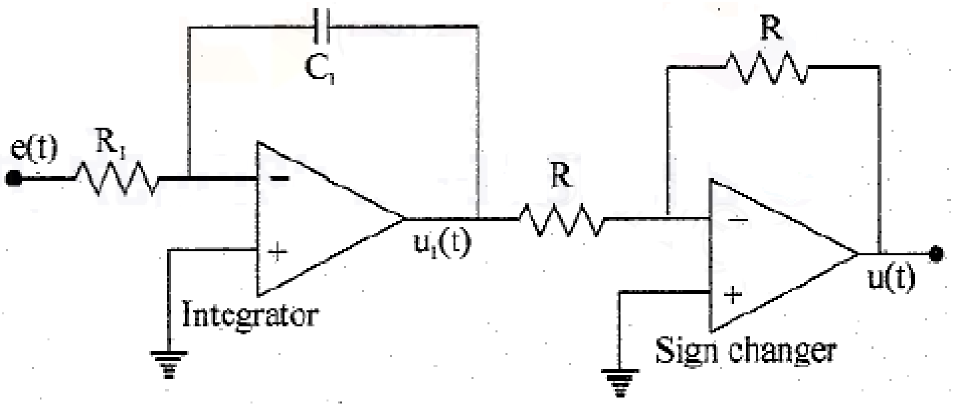


Figure 5.4.6 I-controller using inverting amplifier

[Source: "Control Systems" by Nagoor Kani, Page: 2.82]

By deriving the transfer function of the controller shown in figure and comparing with the transfer function of I-controller defined by equation.

Analysis of I-controller

The assumptions made in op-amp circuit analysis are,

1. The voltages of both inputs are equal
2. The input current is zero.

Based on the above assumptions, the equivalent circuit of op-amp amplifier and sign changer are shown in figure 5.4.7.

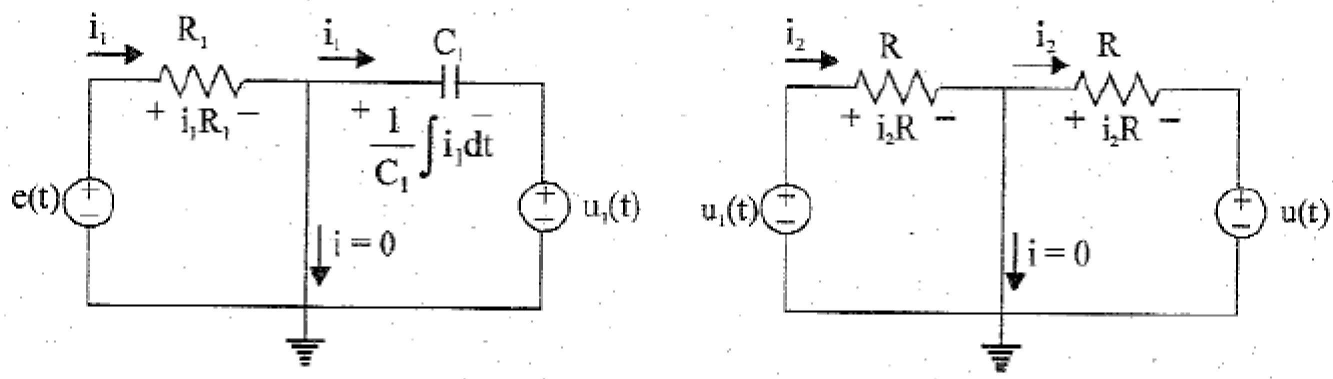


Figure 5.4.7 Equivalent circuit of amplifier and sign changer

[Source: "Control Systems" by Nagoor Kani, Page: 2.83]

From the circuit,

$$e(t) = i_1 R_1$$

$$u_1(t) = -\frac{1}{C_1}\int i_1\, dt$$

Substitute for i_1,

$$u_1(t) = -\frac{1}{R_1C_1}\int e(t)dt$$

Also, from the circuit,

$$u(t) = -i_2R$$

$$u_1(t) = i_2R$$

Substitute for i_2,

$$u_1(t) = -u(t)$$

On equating equations we get

$$u(t) = \frac{1}{R_1C_1}\int e(t)dt$$

On taking Laplace transform of equation we get,

$$\frac{U(s)}{E(s)} = \frac{1}{sR_1C_1}$$

The equation is the transfer function of op-amp P-controller. On the comparing equations, Integral gain,

$$K_i = \frac{1}{R_1C_1}$$

Therefore, by adjusting the values of R_1 and C_1 the value of gain K_i can be varied.

PROPORTIONAL PLUS INTEGRAL CONTROLLER (PI-CONTROLLER)

The proportional plus integral controller (PI controller) produces and output signal consisting of two terms: *one proportional to error signal and the other proportional to the integral of error signal.*

In PI controller,

$$u(t) \propto \left[e(t) + \int e(t)dt\right]$$

$$u(t) = K_p e(t) + \frac{K_p}{T_i}\int e(t)dt$$

On taking Laplace transform of equation with zero initial conditions, we get,

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

The equation gives the output of the PI-controller for the input E(s) and it is the transfer function of PI-controller. The block diagram of PI-controller is shown in figure 5.4.8.

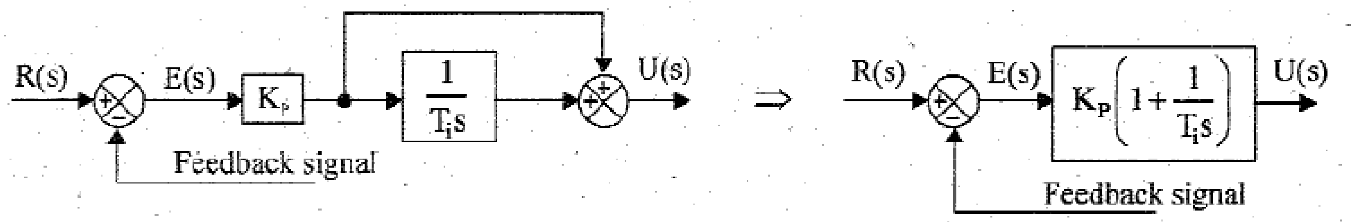


Figure 5.4.8 Block diagram of PI controller

[Source: "Control Systems" by Nagoor Kani, Page: 2.84]

The advantages of both P-controller and I –controller is combined in PI-controller. The proportional action increases the loop gain and makes the system less sensitive to variations of system parameters. The integral action eliminates or reduces the steady state error. The integral control action is adjusted by varying the integral time. The change in value of K_p affects both the proportional and integral parts of control action. The inverse of the integral time T_i is called the reset rate.

Example of Electronic PI-controller

The PI controller can be realized by an op-amp differentiator with gain followed by a sign changer as shown in figure 5.4.9.

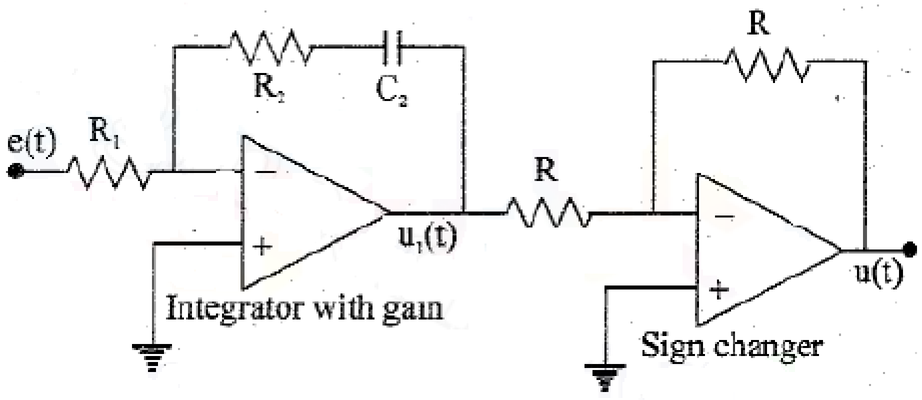


Figure 5.4.9 PI-controller using inverting amplifier

[Source: "Control Systems" by Nagoor Kani, Page: 2.84]

By deriving the transfer function of the controller shown in figure and comparing with the transfer function of PI-controller defined by equation, it can be proved that the circuit shown in figure will work as PI-controller.

Analysis of PI-controller

The assumptions made in op-amp circuit analysis are,

1. The voltages of both inputs are equal
2. The input current is zero.

Based on the above assumptions, the equivalent circuit of op-amp amplifier and sign changer are shown in figure 5.4.10.

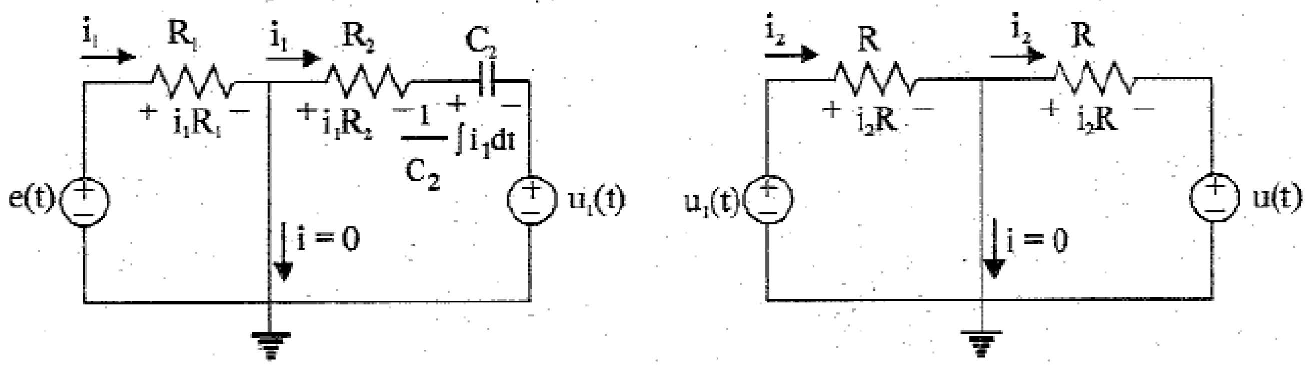


Figure 5.4.10 Equivalent circuit of amplifier and sign changer

[Source: "Control Systems" by Nagoor Kani, Page: 2.85]

From the circuit,

$$e(t) = i_1 R_1$$

$$u_1(t) = -i_1 R_2 - \frac{1}{C_2}\int i_1\, dt$$

Substitute for i_1,

$$u_1(t) = -\frac{e(t)}{R_1}R_2 - \frac{1}{R_1C_2}\int e(t)dt$$

Also, from the circuit,

$$u(t) = -i_2R$$

$$u_1(t) = i_2R$$

Substitute for i_2,

$$u_1(t) = -u(t)$$

On equating equations we get

$$u(t) = \frac{e(t)}{R_1}R_2 + \frac{1}{R_1C_2}\int e(t)dt$$

On taking Laplace transform of equation we get,

$$\frac{U(s)}{E(s)} = \frac{R_2}{R_1}\left(1 + \frac{1}{sR_2C_2}\right)$$

The equation is the transfer function of op-amp P-controller. On the comparing equations,

Proportional gain,

$$K_p = \frac{R_2}{R_1}$$

Integral time,

$$T_i = R_2C_2$$

By varying the values of R_1 and R_2, the value of gain K_p and T_i can be adjusted.

PROPORTIONAL PLUS DERIVATIVE CONTROLLER (PD-CONTROLLER)

The PD controller produces and output signal consisting of two terms: *one proportional to error signal, the other one proportional to derivatives of error signal.*

In PD controller,

$$u(t) \propto \left[e(t) + \frac{d}{dt}e(t)\right]$$

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

On taking Laplace transform of equation with zero initial conditions, we get,

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

The equation gives the output of the PD-controller for the input E(s) and it is the transfer function of PD-controller. The block diagram of PD-controller is shown in figure 5.4.11.

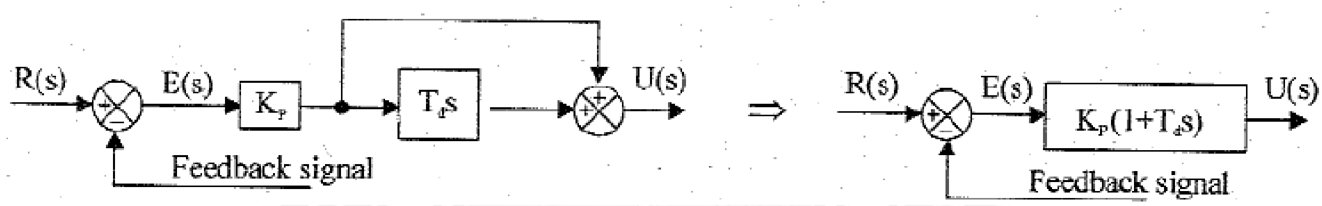


Figure 5.4.11 Block diagram of PD controller

[Source: "Control Systems" by Nagoor Kani, Page: 2.86]

The derivative control acts on a rate of change of error and not on the actual error signal. The derivative control action is effective only during transient periods and so it does not produce corrective measures for any constant error. Hence the derivative controller is never used alone, but it is employed in association with proportional and integral controllers. The derivative controller does not affect the steady-state error directly but anticipates the error, initiates an early corrective action and tends to increase the stability of the system. While derivative control action has an advantage of being anticipatory it has the disadvantage that it amplifies noise signals and may cause a saturation effect in the actuator. The derivative control action is adjusting by varying the derivative time. The change in the value of K_p affects both the proportional and derivative parts of control action. The derivative control is also called rate control.

Example of Electronic PD-controller

The PD controller can be realized by an op-amp amplifier with integral and derivative action followed by a sign changer as shown in figure 5.4.12.

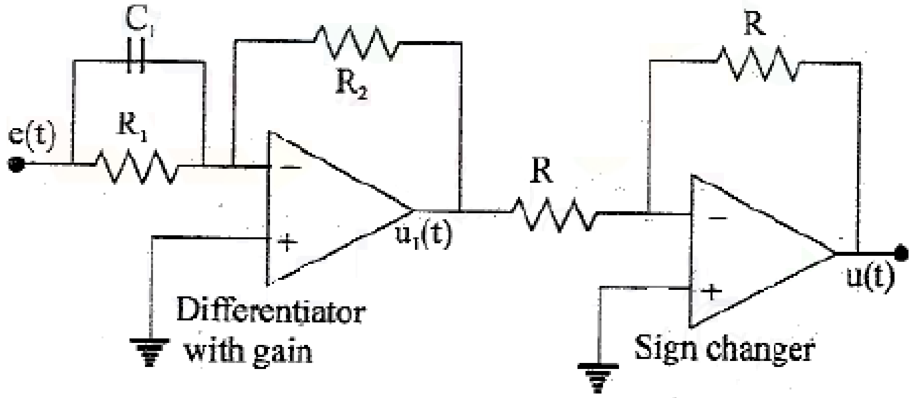


Figure 5.4.12 PD-controller using inverting amplifier

[Source: "Control Systems" by Nagoor Kani, Page: 2.86]

By deriving the transfer function of the controller shown in figure and comparing with the transfer function of PD-controller defined by equation, it can be proved that the circuit shown in figure will work as PD-controller.

Analysis of PD-controller

The assumptions made in op-amp circuit analysis are,

1. The voltages of both inputs are equal
2. The input current is zero.

Based on the above assumptions, the equivalent circuit of op-amp amplifier and sign changer are shown in figure 5.4.13.

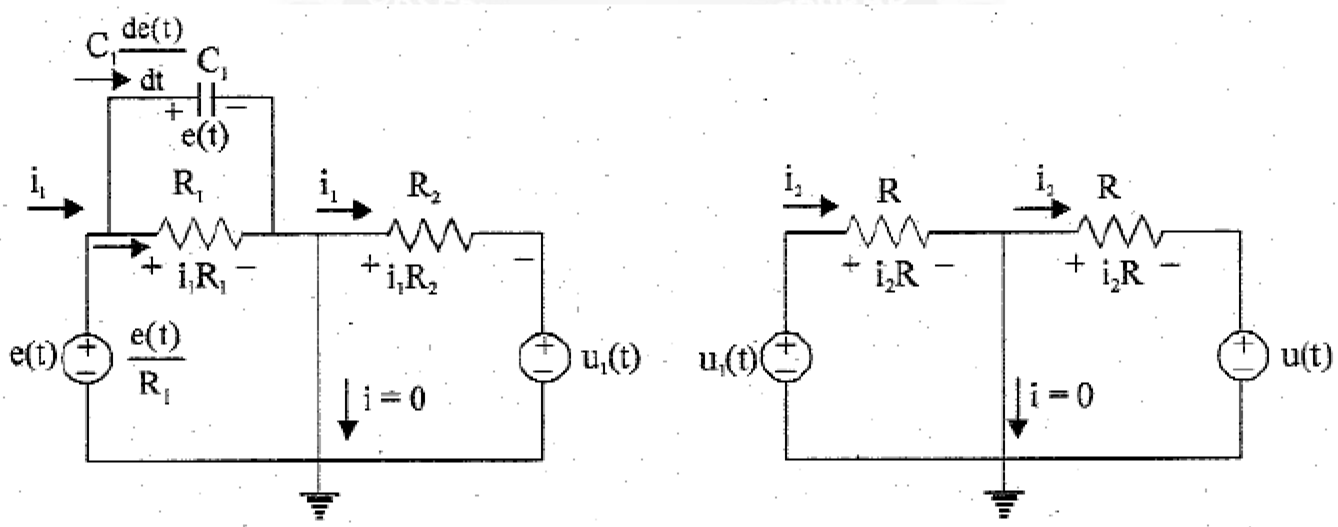


Figure 5.4.13 Equivalent circuit of amplifier and sign changer

[Source: "Control Systems" by Nagoor Kani, Page: 2.87]

From the circuit,

$$i_1 = \frac{e(t)}{R_1} + C_1 \frac{de(t)}{dt}$$

$$u_1(t) = -i_1 R_2$$

Substitute for i_1,

$$u_1(t) = -\frac{e(t)}{R_1} R_2 - R_2 C_1 \frac{d}{dt} e(t)$$

Also, from the circuit,

$$u(t) = -i_2 R$$

$$u_1(t) = i_2 R$$

Substitute for i_2,

$$u_1(t) = -u(t)$$

On equating the equations, we get,

$$u(t) = \frac{e(t)}{R_1} R_2 + R_2 C_1 \frac{d}{dt} e(t)$$

On taking Laplace transform of equation we get,

$$\frac{U(s)}{E(s)} = \frac{R_2}{R_1}(1 + sR_1 C_1)$$

The equation is the transfer function of op-amp P-controller. On the comparing equations, Proportional gain,

$$K_p = \frac{R_2}{R_1}$$

Derivative time,

$$T_d = R_1 C_1$$

By varying the values of R_1 and R_2, the value of K_p and T_d are adjusted.

PROPORTIONAL PLUS INTEGRAL PLUS DERIVATIVE (PID) CONTROLLER

The PID controller produces and output signal consisting of two terms: *one proportional to error signal, another one proportional to the integral of error signal and the third one proportional to derivatives of error signal.*

$$u(t) \propto \left[e(t) + \int e(t)dt + \frac{d}{dt}e(t)\right]$$

$$u(t) = K_p e(t) + \frac{K_p}{T_i}\int e(t)dt + K_p T_d \frac{d}{dt}e(t)$$

On taking Laplace transform of equation with zero initial conditions, we get,

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

The equation gives the output of the PID-controller for the input E(s) and it is the transfer function of PID-controller. The block diagram of PID-controller is shown in figure 5.4.14.

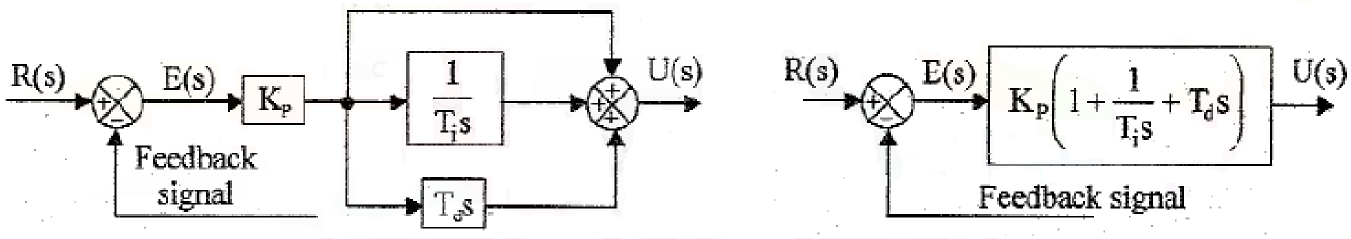


Figure 5.4.14 Block diagram of PID controller

[Source: "Control Systems" by Nagoor Kani, Page: 2.88]

The combination of proportional control action, integral control action and derivative control action is called PID-control action. This combined action has the advantages of the each of the three individual control actions. The proportional controller stabilizes the gain but produces a steady state error. The integral controller reduces or eliminates the steady state error. The derivative controller reduces the rate of change of error.

Example of Electronic PID-controller

The PID controller can be realized by an op-amp amplifier with integral and derivative action followed by a sign changer as shown in figure 5.4.15.

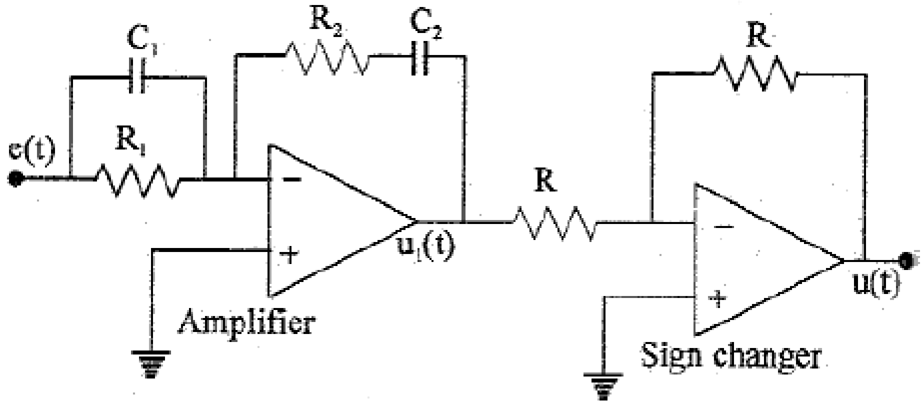


Figure 5.4.15 PID-controller using inverting amplifier

[Source: "Control Systems" by Nagoor Kani, Page: 2.88]

By deriving the transfer function of the controller shown in figure and comparing with the transfer function of PID-controller defined by equation, it can be proved that the circuit shown in figure will work as PID-controller.

Analysis of PID-controller

The assumptions made in op-amp circuit analysis are,

1. The voltages of both inputs are equal
2. The input current is zero.

Based on the above assumptions, the equivalent circuit of op-amp amplifier and sign changer are shown in figure 5.4.16.

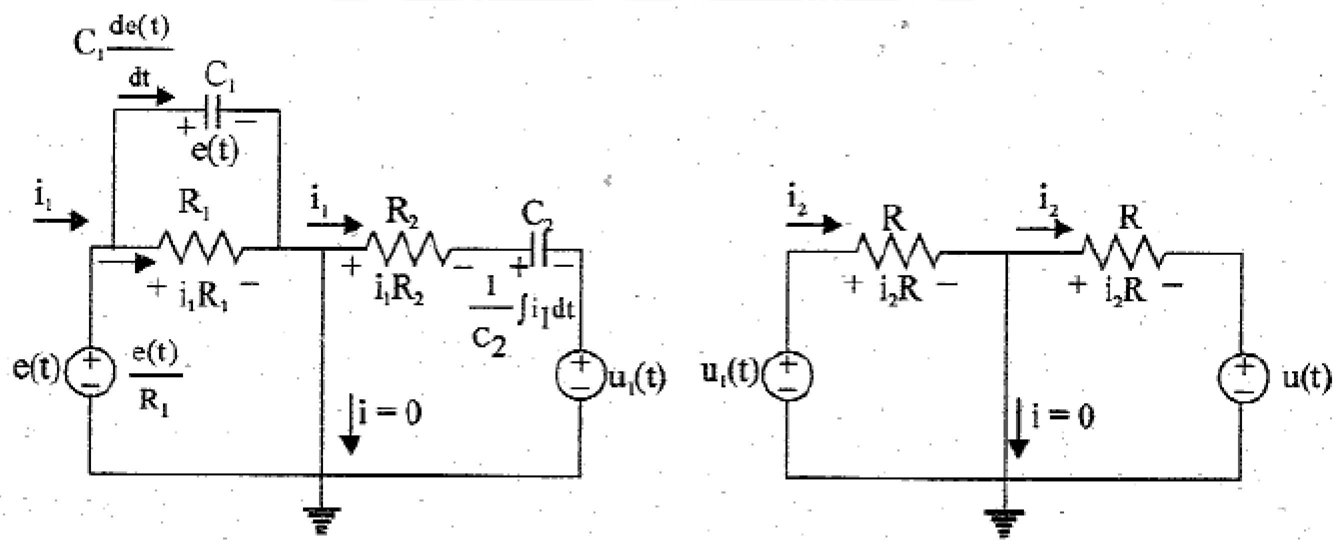


Figure 5.4.16 Equivalent circuit of amplifier and sign changer

[Source: "Control Systems" by Nagoor Kani, Page: 2.89]

From the circuit,

$$i_1 = \frac{e(t)}{R_1} + C_1 \frac{de(t)}{dt}$$

On taking Laplace transform of equation with zero initial conditions, we get,

$$I_1(s) = \left(\frac{1}{R_1} + C_1 s\right) E(s)$$

Also, from the circuit,

$$u_1(t) = -i_1 R_2 - \frac{1}{C_2} \int i_1 \, dt$$

On taking Laplace transform of equation with zero initial conditions, we get,

$$U_1(s) = -I_1(s) R_2 - \frac{1}{sC_2} I_1(s)$$

Substitute for i_1, from equations

$$U_1(s) = -\left(\frac{R_2}{R_1} + \frac{C_1}{C_2} + \frac{1}{R_1 C_2 s} + R_2 C_1 s\right) E(s)$$

Also, from the circuit,

$$u(t) = -i_2 R$$

$$u_1(t) = i_2 R$$

Substitute for i_2,

$$u_1(t) = -u(t)$$

On equating the equations, we get,

$$\frac{U(s)}{E(s)} = \frac{R_2}{R_1}\left(1 + \frac{R_1 C_1 + R_2 C_2}{R_2 C_2} + \frac{1}{R_2 C_2 s} + R_1 C_1 s\right)$$

The equation is the transfer function of op-amp PID-controller. On the comparing, we get,

Proportional gain, $K_p = \frac{R_2}{R_1}$

Derivative time, $T_d = R_1 C_1$

Integral time, $T_i = R_2 C_2$

Also, $\frac{R_1 C_1 + R_2 C_2}{R_2 C_2} = 1$

By varying the values of R_1 and R_2, the value of K_p, T_d and T_i are adjusted.

