BEE501- CONTROL SYSTEM

UNIT 1

SYSTEMS AND THEIR REPRESENTATION

Definition of Control System

A control system is a system of devices or set of devices, that manages commands, directs or regulates the behavior of other device(s) or system(s) to achieve desire results. In other words the **definition of control system** can be rewritten as **A control system is a system, which controls other system**. As the human civilization is being modernized day by day the demand of automation is increasing accordingly. Automation highly requires control of devices.

Requirement of Good Control System

Accuracy: Accuracy is the measurement tolerance of the instrument and defines the limits of the errors made when the instrument is used in normal operating conditions. Accuracy can be improved by using feedback elements. To increase accuracy of any control system error detector should be present in control system.

Sensitivity : The parameters of control system are always changing with change in surrounding conditions, internal disturbance or any other parameters. This change can be expressed in terms of sensitivity. Any control system should be insensitive to such parameters but sensitive to input signals only.

Noise : An undesired input signal is known as noise. A good control system should be able to reduce the noise effect for better performance.

Stability : It is an important characteristic of control system. For the bounded input signal, the output must be bounded and if input is zero then output must be zero then such a control system is said to be stable system.

Bandwidth : An operating frequency range decides the bandwidth of control system. Bandwidth should be large as possible for frequency response of good controlsystem.

Speed : It is the time taken by control system to achieve its stable output. A good control system possesses high speed. The transient period for such system is very small.

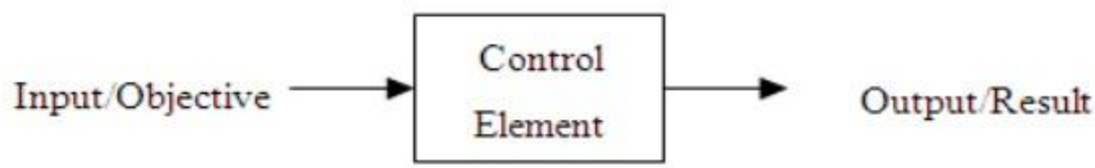
Oscillation : A small numbers of oscillation or constant oscillation of output tend to system to be stable.

Basic elements of control system:

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

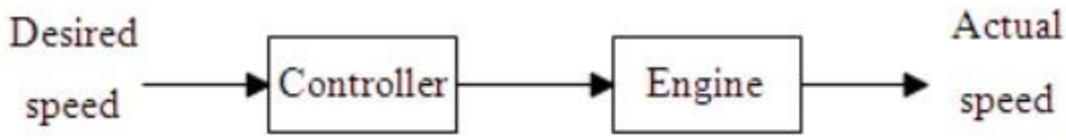
2) Closed-loop systems.



Basic Components of Control System

1. Open-loop systems:

The open-loop system is also called the non-feedback system. This is the simpler of the two systems. A simple example is illustrated by the speed control of an automobile as shown in Figure 1-2. In this open-loop system, there is no way to ensure the actual speed is close to the desired speed automatically. The actual speed might be way off the desired speed because of the wind speed and/or road conditions, such as uphill or downhill etc.



Basic Open Loop System

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

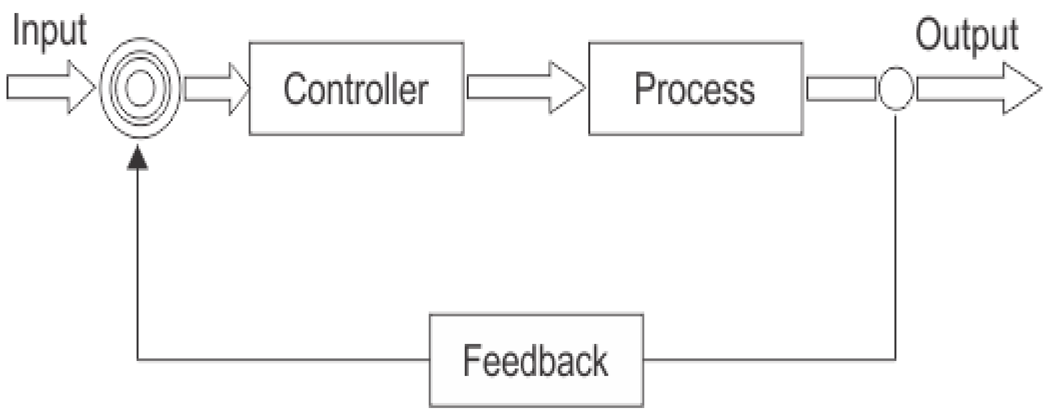
1. Simple in construction and design.
2. Economical.
3. Easy to maintain.
4. Generally stable.
5. Convenient to use as output is difficult to measure.

Disadvantages of Open Loop Control System

1. They are inaccurate.
2. They are unreliable.
3. Any change in output cannot be corrected automatically.

2. Closed-loop systems:

The closed-loop system is also called the feedback system. A simple closed-system is shown in Figure 1-3. It has a mechanism to ensure the actual speed is close to the desired speed automatically.



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

1. Closed loop control systems are more accurate even in the presence of non-linearity.
2. Highly accurate as any error arising is corrected due to presence of feedback signal.
3. Bandwidth range is large.
4. Facilitates automation.
5. The sensitivity of system may be made small to make system more stable.
6. This system is less affected by noise.

Disadvantages of Closed Loop Control System

1. They are costlier.
2. They are complicated to design.
3. Required more maintenance.
4. Feedback leads to oscillatory response.
5. Overall gain is reduced due to presence of feedback.
6. Stability is the major problem and more care is needed to design a stable closed loop system.

Comparison of Closed Loop And Open Loop Control System

Open Loop	Closed Loop
Any change in output has no effect on the input. Example : Feedback does not exists	Changes in output, affects the input which is possible by use of feedback
Output measurement is not required for operation of system	Output measurement is necessary
Feedback element is absent	Feedback element is present
Error detector is absent	Error detector is necessary
It is inaccurate and unreliable	Highly accurate and reliable
Highly sensitive to the disturbances	Less sensitive to the disturbances
Highly sensitive to the environmental changes	Less sensitive to the environmental changes
Bandwidth is small	Bandwidth is large
Simple to construct and cheap	Complicated to design and hence costly
Generally are stable in nature	Stability is the major consideration while designing
Highly affected by non-linearity	Reduced effect of nonlinearities

Electrical Analogies of Mechanical Systems

Two systems are said to be **analogous** to each other if the following two conditions are satisfied.

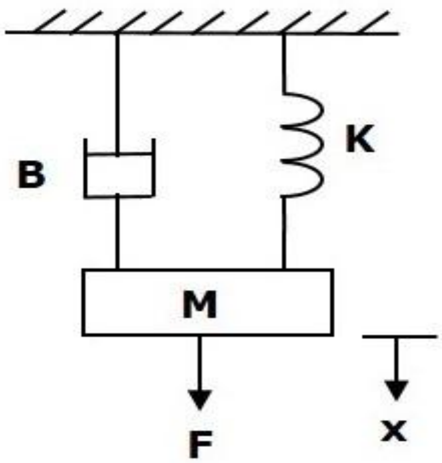
- The two systems are physically different
- Differential equation modelling of these two systems are same

Electrical systems and mechanical systems are two physically different systems. There are two types of electrical analogies of translational mechanical systems. Those are force voltage analogy and force current analogy.

Force Voltage Analogy

In force voltage analogy, the mathematical equations of **translational mechanical system** are compared with mesh equations of the electrical system.

Consider the following translational mechanical system as shown in the following figure.



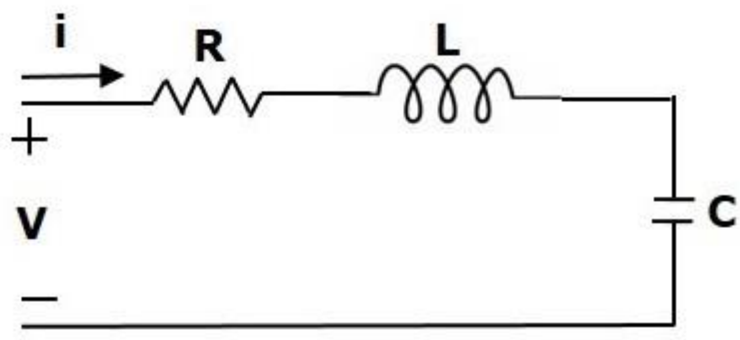
The **force balanced equation** for this system is

$$F = F_m + F_b + F_k$$

$$\Rightarrow F = M\frac{d^2x}{dt^2} + B\frac{dx}{dt} + Kx$$

(Equation 1)

Consider the following electrical system as shown in the following figure. This circuit consists of a resistor, an inductor and a capacitor. All these electrical elements are connected in a series. The input voltage applied to this circuit is VV volts and the current flowing through the circuit is ii Amps.



Mesh equation for this circuit is

$$V = Ri + L\frac{\mathrm{d}i}{\mathrm{d}t} + \frac{1}{c}\int i dt$$ **(Equation 2)**

Substitute, $i = \frac{\mathrm{d}q}{\mathrm{d}t}$ in Equation 2.

$$V = R\frac{\mathrm{d}q}{\mathrm{d}t} + L\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + \frac{q}{C}$$

$$\Rightarrow V = L\frac{\mathrm{d}^2q}{\mathrm{d}t^2} + R\frac{\mathrm{d}q}{\mathrm{d}t} + \left(\frac{1}{c}\right)q$$

(Equation 3)

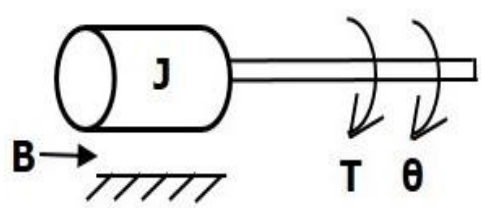
By comparing Equation 1 and Equation 3, we will get the analogous quantities of the translational mechanical system and electrical system. The following table shows these analogous quantities.

Rotational Mechanical System	Electrical System
Torque(T)	Voltage(V)

Moment of Inertia(J)	Inductance(L)
Rotational friction coefficient(B)	Resistance(R)
Torsional spring constant(K)	Reciprocal of Capacitance (1c)(1c)
Angular Displacement(θ)	Charge(q)
Angular Velocity(ω)	Current(i)

Torque Voltage Analogy:

In this analogy, the mathematical equations of **rotational mechanical system** are compared with mesh equations of the electrical system. Rotational mechanical system is shown in the following figure.



The torque balanced equation is

$$T = T_j + T_b + T_k$$

$$\Rightarrow T = J\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + B\frac{\mathrm{d}\theta}{\mathrm{d}t} + k\theta$$

(Equation 4)

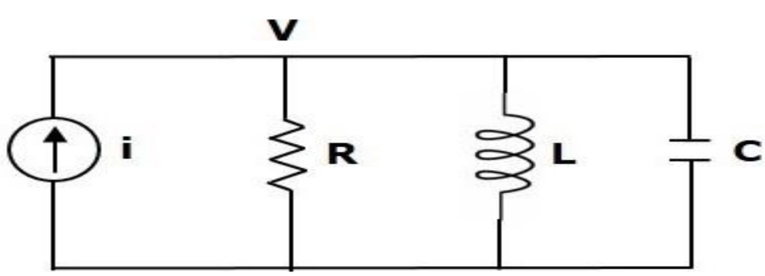
By comparing Equation 4 and Equation 3, we will get the analogous quantities of rotational mechanical system and electrical system. The following table shows these analogous quantities.

Rotational Mechanical System	Electrical System
Torque(T)	Voltage(V)
Moment of Inertia(J)	Inductance(L)
Rotational friction coefficient(B)	Resistance(R)
Torsional spring constant(K)	Reciprocal of Capacitance (1c)(1c)
Angular Displacement(θ)	Charge(q)
Angular Velocity(ω)	Current(i)

Force Current Analogy:

In force current analogy, the mathematical equations of the **translational mechanical system** are compared with the nodal equations of the electrical system.

Consider the following electrical system as shown in the following figure. This circuit consists of current source, resistor, inductor and capacitor. All these electrical elements are connected in parallel.



The nodal equation is

$$i = \frac{V}{R} + \frac{1}{L}\int V dt + C\frac{\mathrm{d}V}{\mathrm{d}t}$$ **(Equation 5)**

Substitute, $V = \frac{\mathrm{d}\Psi}{\mathrm{d}t}$ in Equation 5.

$$i = \frac{1}{R}\frac{\mathrm{d}\Psi}{\mathrm{d}t} + \left(\frac{1}{L}\right)\Psi + C\frac{\mathrm{d}^2\Psi}{\mathrm{d}t^2}$$

$$\Rightarrow i = C\frac{\mathrm{d}^2\Psi}{\mathrm{d}t^2} + \left(\frac{1}{R}\right)\frac{\mathrm{d}\Psi}{\mathrm{d}t} + \left(\frac{1}{L}\right)\Psi$$ **(Equation 6)**

By comparing Equation 1 and Equation 6, we will get the analogous quantities of the translational mechanical system and electrical system. The following table shows these analogous quantities.

Translational Mechanical System	Electrical System
Force(F)	Current(i)
Mass(M)	Capacitance(C)
Frictional coefficient(B)	Reciprocal of Resistance(1R)(1R)
Spring constant(K)	Reciprocal of Inductance(1L)(1L)
Displacement(x)	Magnetic Flux(ψ)
Velocity(v)	Voltage(V)

Similarly, there is a torque current analogy for rotational mechanical systems.

Torque Current Analogy:

In this analogy, the mathematical equations of the **rotational mechanical system** are compared with the nodal mesh equations of the electrical system.

By comparing Equation 4 and Equation 6, we will get the analogous quantities of rotational mechanical system and electrical system. The following table shows these analogous quantities.

Rotational Mechanical System	Electrical System
Torque(T)	Current(i)
Moment of inertia(J)	Capacitance(C)
Rotational friction coefficient(B)	Reciprocal of Resistance(1R)(1R)
Torsion spring constant(K)	Reciprocal of Inductance(1L)(1L)
Angular displacement(θ)	Magnetic flux(ψ)
Angular velocity(ω)	Voltage(V)

Transfer Function of Control System

A control system consists of an output as well as an input signal. The output is related to the input through a function call **transfer function**.

Definition of Transfer Function

The transfer function of a control system is defined as the ration of the Laplace transform of the output variable to Laplace transform of the input variable assuming all initial conditions to be zero.

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

For any control system there exists a reference input termed as excitation or cause which operates through a transfer operation termed as **transfer function** and produces an effect resulting in controlled output or response. Thus the cause and effect relationship between the output and input is related to each other through a **transfer function**.

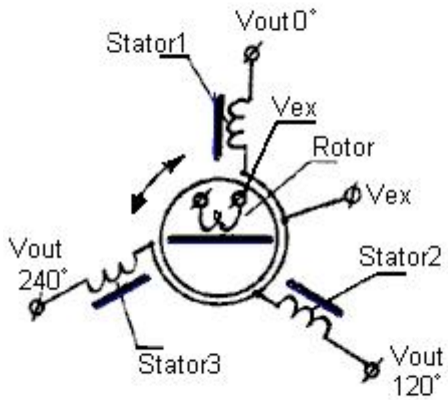
In Laplace Transform, if the input is represented by R(s) and output is represented by C(s), then the transfer function will be

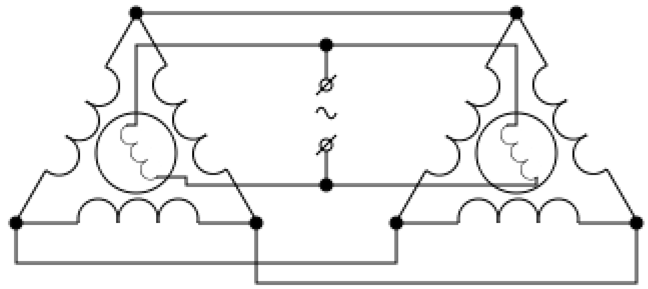
$$G(s) = \frac{C(s)}{R(s)} \Rightarrow R(s).G(s) = C(s)$$

That is, transfer function of the system multiplied by input function gives the output function of the system.

Synchro :

A **synchro** is, in effect, a transformer whose primary-to-secondary coupling may be varied by physically changing the relative orientation of the two windings. Synchros are often used for measuring the angle of a rotating machine such as an antenna platform. In its general physical construction, it is much like an electric motor. The primary winding of the transformer, fixed to the rotor, is excited by an alternating current, which by electromagnetic induction, causes currents to flow in three Y-connected secondary windings fixed at 120 degrees to each other on the stator. The relative magnitudes of secondary currents are measured and used to determine the angle of the rotor relative to the stator, or the currents can be used to directly drive a receiver synchro that will rotate in unison with the synchro transmitter. In the latter case, the whole device may be called a **selsyn** (a portmanteau of *self* and *synchronizing*).





There are two types of synchros systems: Torque systems and control systems.

In a torque system, synchros will provide a low-power mechanical output sufficient to position an indicating device, actuate a sensitive switch or move light loads without power amplification. In simpler terms, a torque synchros system is a system in which the transmitted signal does the usable work. In such a system, accuracy on the order of one degree is attainable.

Servo Motor:

Servo Motor is also called Control motors. They 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.The 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.The servo motor is widely used in radar and computers, robot, machine tool, tracking and guidance systems, processing controlling.

Classification of Servo Motor:

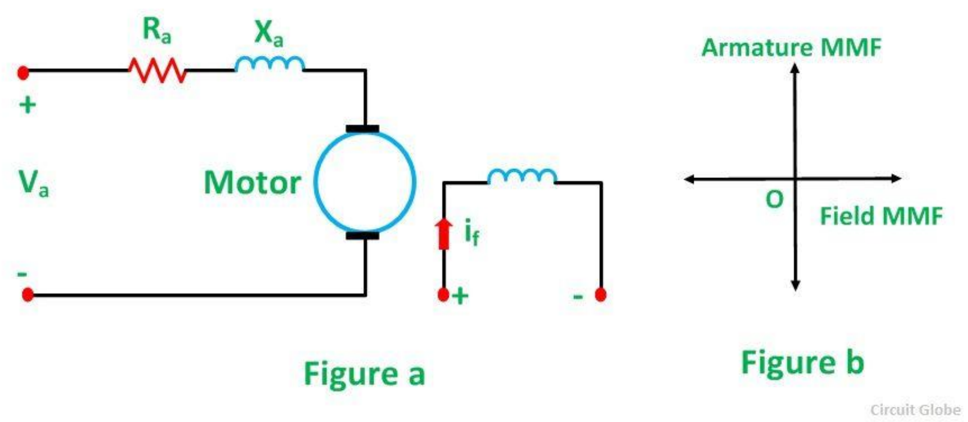
They are classified as AC and DC Servo Motor. The AC servomotor is further divided into two types.

Two Phase AC Servo Motor

Three Phase AC Servo Motor

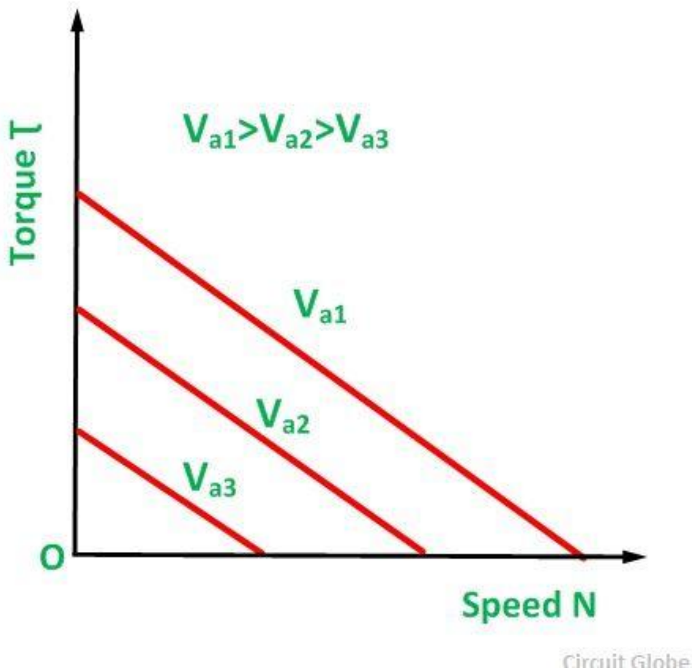
DC Servo Motor:

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.

The Torque-Speed Characteristics of the Motor is shown below.



As from the above characteristics, it is seen that the slope is negative. Thus, a negative slope provides viscous damping for the servo drive system.

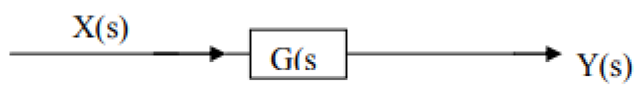
AC Servo Motor

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 ofservo motors can be controlled more precisely than those of standard DC motors, and they usually have three wires (power, ground & control).The 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.

BLOCK DIAGRAM

A control system may consist of a number of components. A block diagram of a system is a pictorial representation of the function performed by each component and of the flow of signals. Such a diagram depicts the inter-relationships which exists between the various components. A block diagram has the advantage of indicating more realistically the signal flows of the actual system. In a block diagram all system variables are linked to each other through functional blocks. The ―Functional Block‖ or simply ―Block‖ is a symbol for the mathematical operation on the input signal to the block which produces the output. The transfer functions of the components are usually entered in the corresponding blocks, which are connected by arrows to indicate the direction of flow of signals. Note that signal can pass only in the direction of arrows. Thus a block diagram of a control system explicitly shows a unilateral property.

Below Fig shows an element of the block diagram. The arrow head pointing towards the block indicates the input and the arrow head away from the block represents the output. Such arrows are entered as signals.



Block

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

The following figure shows a block having input X(s), output Y(s) and the transfer function G(s).

Transfer Function,

$$G(s) = \frac{Y(s)}{X(s)}$$

$$\Rightarrow Y(s) = G(s)X(s)$$

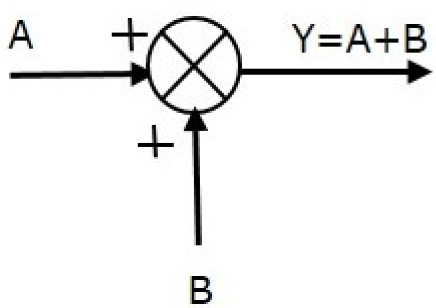
Output of the block is obtained by multiplying transfer function of the block with input.

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.**

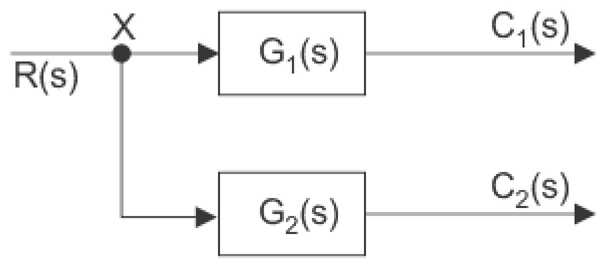
i.e.,Y = A + B.



Take-off Point

The take-off point is a point from which the same input signal can be passed through more than one branch. That means with the help of take-off point, we can apply the same input to one or more blocks, summing points.

In the following figure, the take-off point is used to connect the same input, R(s) to two more blocks.



Block Diagram Reduction Rules

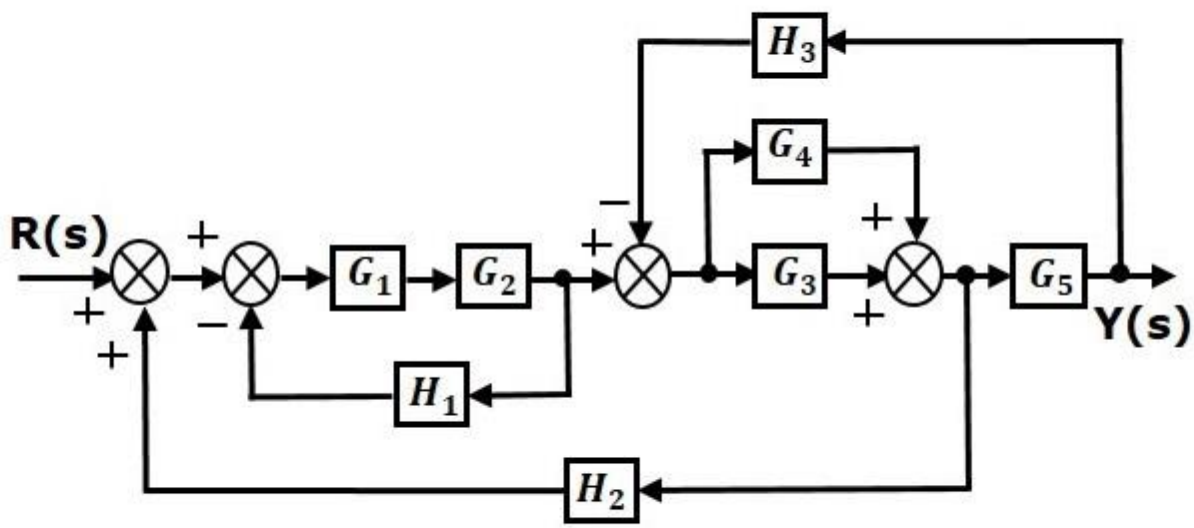
Follow these rules for simplifying (reducing) the block diagram, which is having many blocks, summing points and take-off points.

- **Rule 1 –** Check for the blocks connected in series and simplify.
- **Rule 2 –** Check for the blocks connected in parallel and simplify.
- **Rule 3 –** Check for the blocks connected in feedback loop and simplify.
- **Rule 4 –** If there is difficulty with take-off point while simplifying, shift it towards right.
- **Rule 5 –** If there is difficulty with summing point while simplifying, shift it towards left.
- **Rule 6 –** Repeat the above steps till you get the simplified form, i.e., single block.

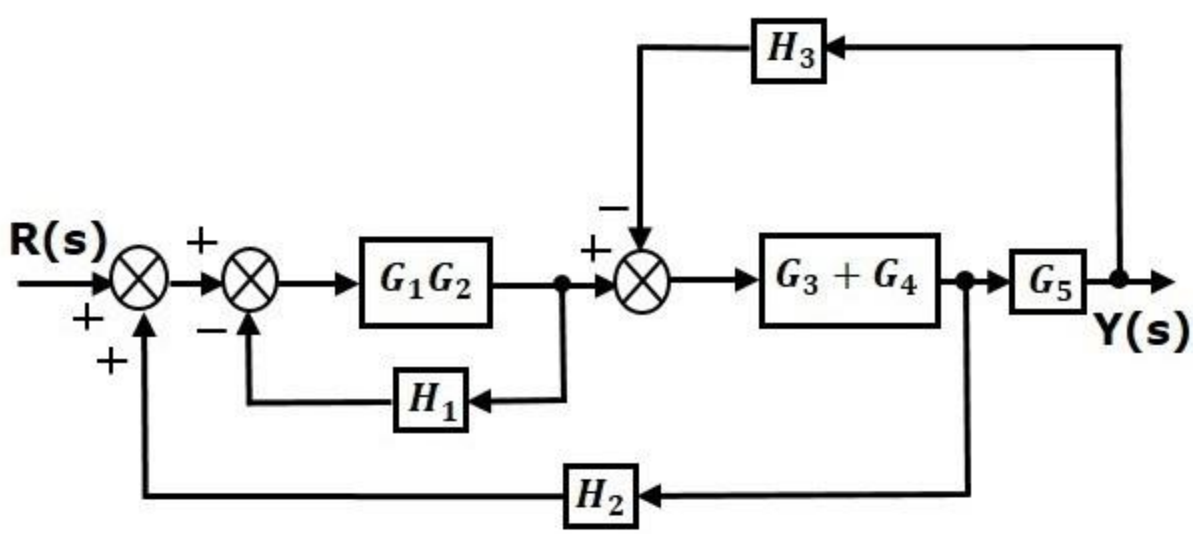
Note – The transfer function present in this single block is the transfer function of the overall block diagram.

Example

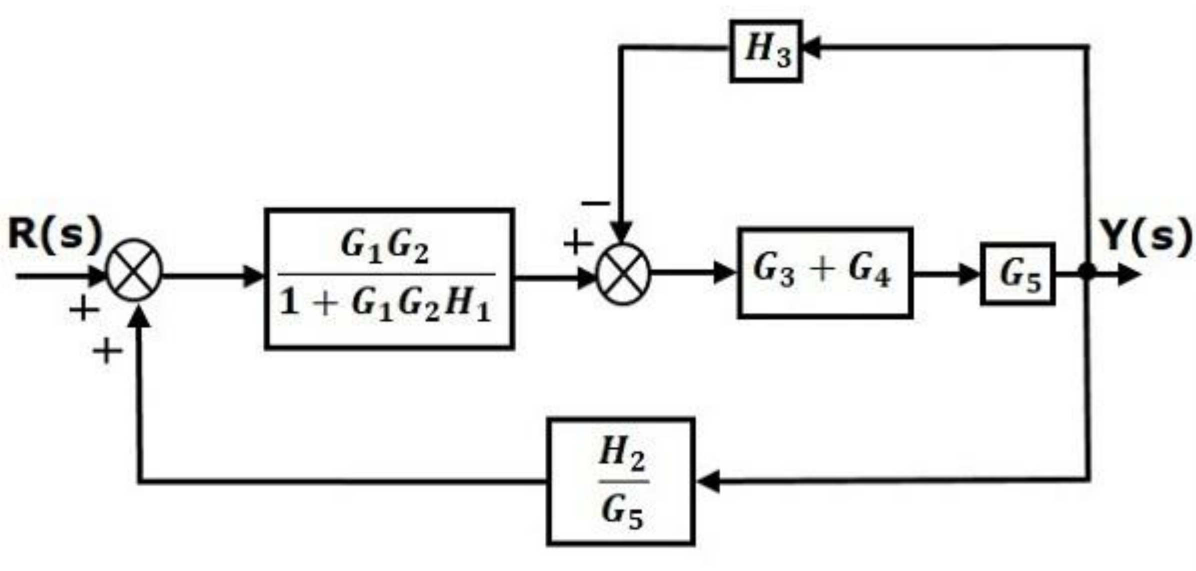
Consider the block diagram shown in the following figure. Let us simplify (reduce) this block diagram using the block diagram reduction rules.



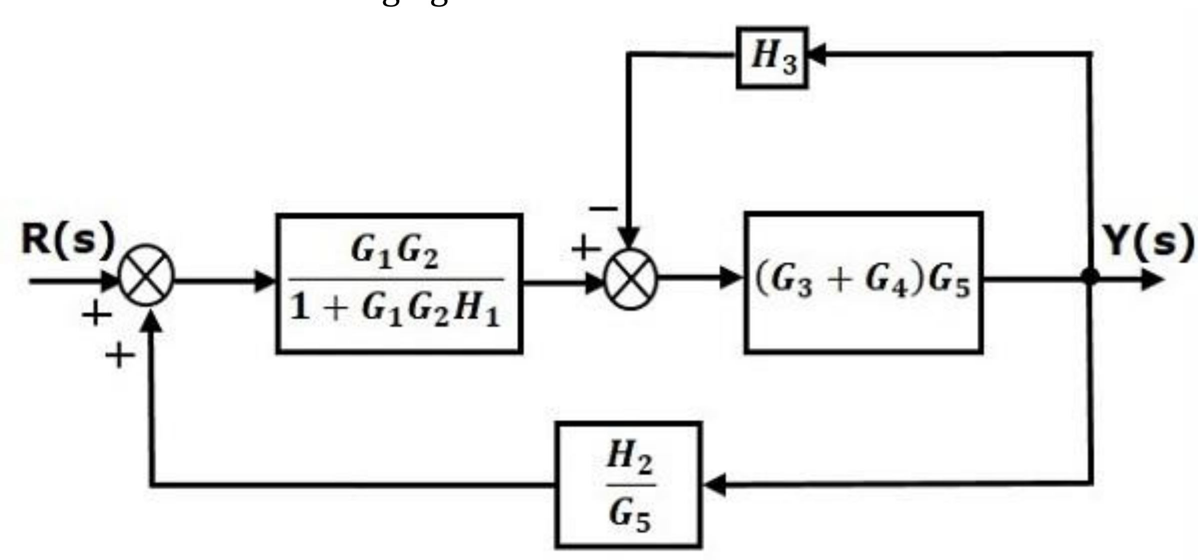
Step 1 – Use Rule 1 for blocks G1 and G2. Use Rule 2 for blocks G3 and G4. The modified block diagram is shown in the following figure.



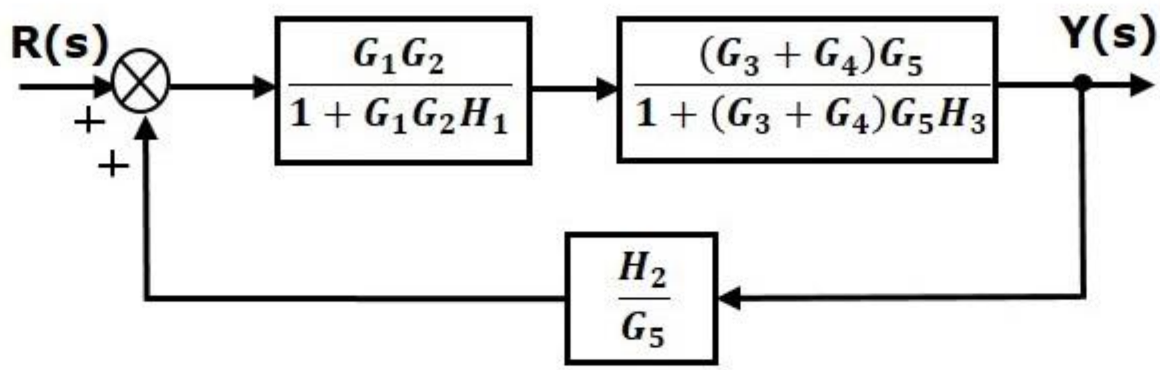
Step 2 – Use Rule 3 for blocks G1G2 and H1. Use Rule 4 for shifting take-off point after the block G5. The modified block diagram is shown in the following figure.



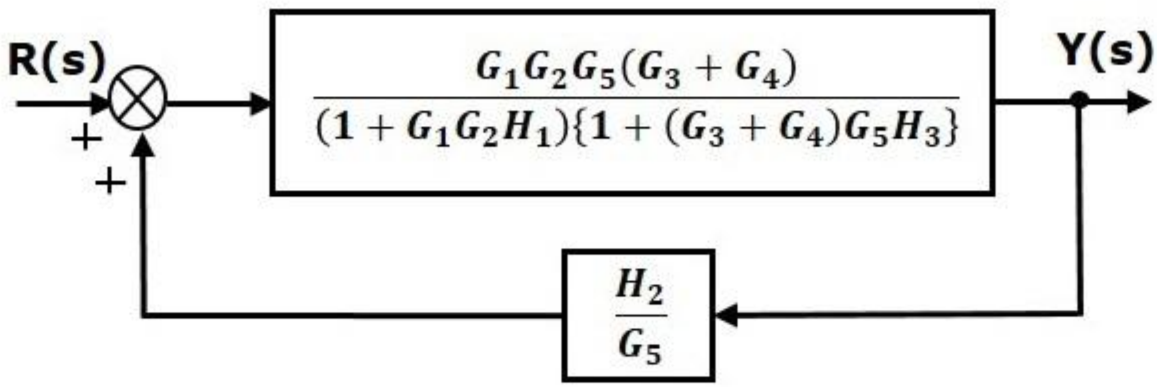
Step 3 – Use Rule 1 for blocks (G3+G4) and G5. The modified block diagram is shown in the following figure.



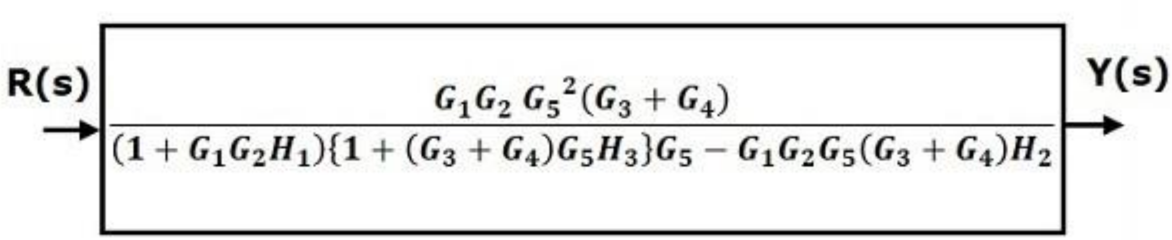
Step 4 – Use Rule 3 for blocks (G3+G4)G5 and H3. The modified block diagram is shown in the following figure.



Step 5 – Use Rule 1 for blocks connected in series. The modified block diagram is shown in the following figure.



Step 6 – Use Rule 3 for blocks connected in feedback loop. The modified block diagram is shown in the following figure. This is the simplified block diagram.



Therefore, the transfer function of the system is

Y(s)R(s)=G1G2G25(G3+G4)(1+G1G2H1){1+(G3+G4)G5H3}G5−G1G2G5(G3+G4)H2Y(s)R(s)=G1G2G52(G3+G4)(1+G1G2H1){1+(G3+G4)G5H3}G5−G1G2G5(G3+G4)H2

Note – Follow these steps in order to calculate the transfer function of the block diagram having multiple inputs.

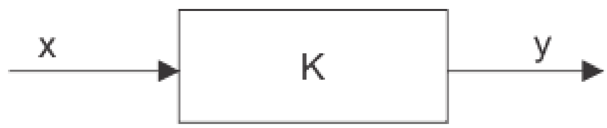
- **Step 1 –** Find the transfer function of block diagram by considering one input at a time and make the remaining inputs as zero.
- **Step 2 –** Repeat step 1 for remaining inputs.
- **Step 3 –** Get the overall transfer function by adding all those transfer functions.

Signal Flow Graph of Control System

The block diagram reduction process takes more time for complicated systems. Because, we have to draw the (partially simplified) block diagram after each step. So, to overcome this drawback, use signal flow graphs (representation).

In the next two chapters, we will discuss about the concepts related to signal flow graphs, i.e., how to represent signal flow graph from a given block diagram and calculation of transfer function just by using a gain formula without doing any reduction process.

Signal flow graph of control system is further simplification of block diagram of control system. Here, the blocks of transfer function, summing symbols and take off points are eliminated by branches and nodes.The transfer function is referred as transmittance in signal flow graph. Let us take an example of equation y = Kx. This equation can be represented with block diagram as below.



The same equation can be represented by signal flow graph, where x is input variable node, y is output variable node and a is the transmittance of the branch connecting directly these two nodes.

Key Definitions:

- Input Node: Node with only outgoing branches;
- Output Node: Node with incoming branches. Note: Any non-input node can be made an output node by adding a branch with gain= 1.
- Path: Collection of branches linked together in same direction.
- Forward Path: Path from input node to output node where node is visited more than once.
- Gain of Forward Path: Product of all gains of branches in the forward path.
- Loop: Path that originates and terminates at the same node. No other node is visited more than once.
- Loop Gain: Product of branch gains in a loop.
- Non-Touching: Two parts of a SFG are non-touching if they do not share at least one node

Mason's gain formula is

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

Where,

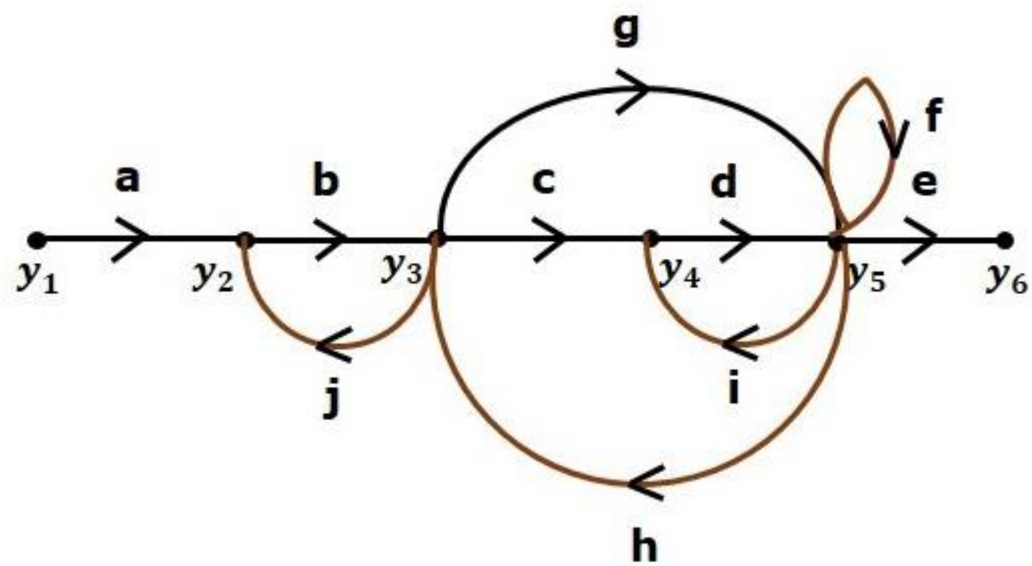
- **C(s)** is the output node
- **R(s)** is the input node
- **T** is the transfer function or gain between R(s)R(s) and C(s)C(s)
- **P_i** is the i^{th} forward path gain

Δ=1−(sum of all individual loop gains) +(sum of gain products of all possible two nontouching loops) −(sum of gain products of all possible three nontouching loops)+...

Δ_i is obtained from Δ by removing the loops which are touching the i^{th} forward path.

calculation of Transfer Function using Mason's Gain Formula

Let us consider the same signal flow graph for finding transfer function.



- Number of forward paths, N = 2.
- First forward path is - $y_1 \rightarrow y_2 \rightarrow y_3 \rightarrow y_4 \rightarrow y_5 \rightarrow y_6$
- First forward path gain, $p_1 = abcde$
- Second forward path is - $y_1 \rightarrow y_2 \rightarrow y_3 \rightarrow y_5 \rightarrow y_6$.
- Second forward path gain, $p_2 = abge$
- Number of individual loops, L = 5.
- Loops are - $y_2 \rightarrow y_3 \rightarrow y_2$, $y_3 \rightarrow y_5 \rightarrow y_3$, $y_3 \rightarrow y_4 \rightarrow y_5 \rightarrow y_3$, $y_4 \rightarrow y_5 \rightarrow y_4$ and $y_5 \rightarrow y_5$.

- Loop gains are - $l_1=bj$, $l_2=gh$, $l3=cd\ h$, $l4=di$ and $l5=f$.
- Number of two non-touching loops = 2.

- First non-touching loops pair is - $y_2 \to y_3 \to y_2$, $y_4 \to y_5 \to y_4$.
- Gain product of first non-touching loops pair, $l_1 l_4=bjdi$l1l4=bjdi
- Second non-touching loops pair is - $y_2 \to y_3 \to y_2$, $y_5 \to y_5$.
- Gain product of second non-touching loops pair is - $l_1 l_5=bjf$

Higher number of (more than two) non-touching loops are not present in this signal flow graph.

We know,

$\Delta=1-$(sum of all individual loop gains)+(sum of gain products of all possible two non touching loops)−(sum of gain products of all possible three non touching loops)+...

Substitute the values in the above equation,

$$\Delta=1-(bj+gh+cdh+di+f)+(bjdi+bjf)-(0)$$
$$\Rightarrow \Delta=1-(bj+gh+cdh+di+f)+bjdi+bjf$$

There is no loop which is non-touching to the first forward path.

So, $\Delta_1=1$.

Similarly, $\Delta_2=1$. Since, no loop which is non-touching to the second forward path.

Substitute, N = 2 in Mason’s gain formula

$$T=\frac{C(s)}{R(s)}=\frac{\Sigma_{i=1}^{2}P_i\Delta_i}{\Delta}$$

$$T=\frac{C(s)}{R(s)}=\frac{P_1\Delta_1+P_2\Delta_2}{\Delta}$$

Substitute all the necessary values in the above equation.

$$T=\frac{C(s)}{R(s)}=\frac{(abcde)1+(abge)1}{1-(bj+gh+cdh+di+f)+bjdi+bjf}$$

$$\Rightarrow T=\frac{C(s)}{R(s)}=\frac{(abcde)+(abge)}{1-(bj+gh+cdh+di+f)+bjdi+bjf}$$

Therefore, the transfer function is -

$$T=\frac{C(s)}{R(s)}=\frac{(abcde)+(abge)}{1-(bj+gh+cdh+di+f)+bjdi+bjf}$$

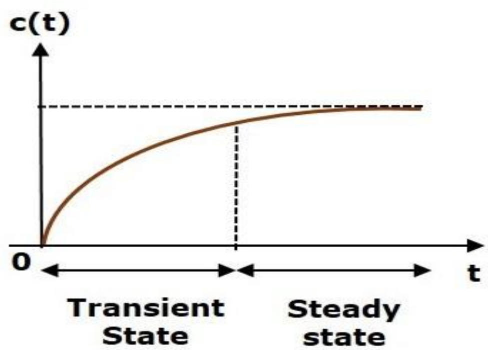
UNIT 2
TIME RESPONSE ANALYSIS

Time Response:

If the output of control system for an input varies with respect to time, then it is called the **time response** of the control system. The time response consists of two parts.

- Transient response
- Steady state response

The response of control system in time domain is shown in the following figure.



Here, both the transient and the steady states are indicated in the figure. The responses corresponding to these states are known as transient and steady state responses.

Mathematically, we can write the time response c(t) as

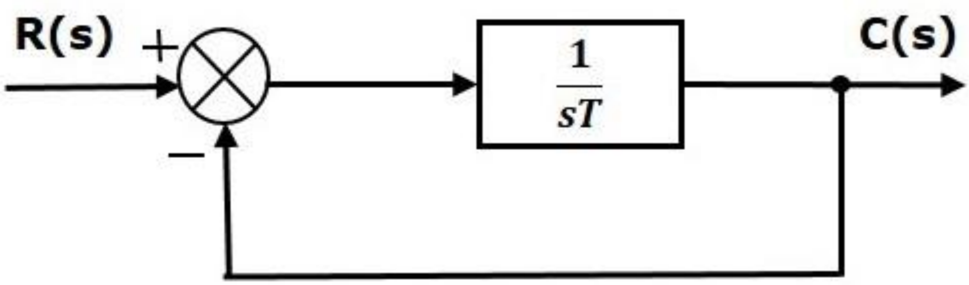
$$c(t)=ctr(t)+css(t)c(t)$$

Where,

- $c_{tr}(t)$ is the transient response
- $c_{ss}(t)$ is the steady state response
-

Response of the First Order System

Time response of the first order system. Consider the following block diagram of the closed loop control system. Here, an open loop transfer function, 1sT1sT is connected with a unity negative feedback.



We know that the transfer function of the closed loop control system has unity negative feedback as,

C(s)/R(s)=G(s)/1+G(s)

Substitute, G(s)=1sT in the above equation.

C(s)/R(s)=1sT1+1sT=1sT+1

The power of s is one in the denominator term. Hence, the above transfer function is of the first order and the system is said to be the **first order system**.We can re-write the above equation as

C(s)=(1sT+1)R(s)

Where,

- **C(s)** is the Laplace transform of the output signal c(t),
- **R(s)** is the Laplace transform of the input signal r(t), and
- **T** is the time constant.

Follow these steps to get the response (output) of the first order system in the time domain.

- Take the Laplace transform of the input signal r(t)r(t).
- Consider the equation, C(s)=(1sT+1)R(s)C(s)=(1sT+1)R(s)
- Substitute R(s)R(s) value in the above equation.
- Do partial fractions of C(s)C(s) if required.
- Apply inverse Laplace transform to C(s)C(s).

Step Response of First Order System

- Consider the **unit step signal** as an input to first order system.

$$\text{So, } r(t) = u(t)$$

Apply Laplace transform on both the sides.

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

Consider the equation, $C(s) = \left(\frac{1}{sT+1}\right) R(s)$

Substitute, $R(s) = \frac{1}{s}$ in the above equation.

$$C(s) = \left(\frac{1}{sT+1}\right)\left(\frac{1}{s}\right) = \frac{1}{s\,(sT+1)}$$

Do partial fractions of C(s).

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

$$\Rightarrow \frac{1}{s\,(sT+1)} = \frac{A\,(sT+1)+Bs}{s\,(sT+1)}$$

On both the sides, the denominator term is the same. So, they will get cancelled by each other. Hence, equate the numerator terms.

$$1 = A\,(sT+1) + Bs$$

By equating the constant terms on both the sides, you will get A = 1.

Substitute, A = 1 and equate the coefficient of the **s** terms on both the sides.

Substitute, A = 1 and equate the coefficient of the **s** terms on both the sides.

$$0 = T + B \Rightarrow B = -T$$

Substitute, A = 1 and B = −T in partial fraction expansion of $C(s)$.

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

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

Apply inverse Laplace transform on both the sides.

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

The **unit step response**, c(t) has both the transient and the steady state terms.

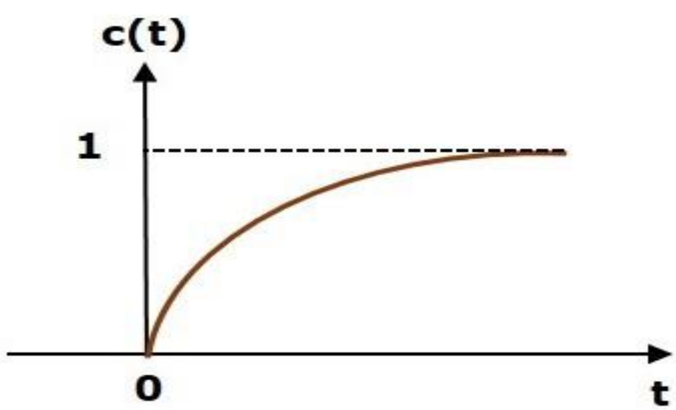
The transient term in the unit step response is -

$$c_{tr}(t) = -e^{-\left(\frac{t}{T}\right)} u(t)$$

The steady state term in the unit step response is -

$$c_{ss}(t) = u(t)$$

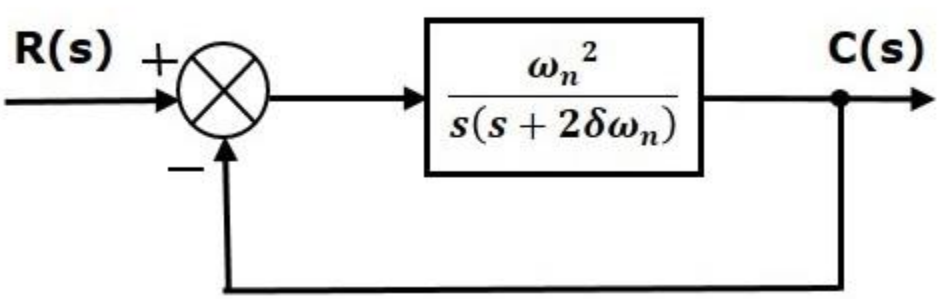
The following figure shows the unit step response.



- The value of the **unit step response, c(t)** is zero at t = 0 and for all negative values of t. It is gradually increasing from zero value and finally reaches to one in steady state. So, the steady state value depends on the magnitude of the input.

Response of Second Order System

Consider the following block diagram of closed loop control system. Here, an open loop transfer function, $\frac{\omega_n^2}{s(s+2\delta\omega_n)}$ is connected with a unity negative feedback.



We know that the transfer function of the closed loop control system having unity negative feedback as

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

Substitute, $G(s) = \frac{\omega_n^2}{s(s+2\delta\omega_n)}$ in the above equation.

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

The power of ‘s’ is two in the denominator term. Hence, the above transfer function is of the second order and the system is said to be the **second order system**.

The characteristic equation is –

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

The roots of characteristic equation are –

$$s = \frac{-2\omega\delta_n \pm \sqrt{(2\delta\omega_n)^2 - 4\omega_n^2}}{2} = \frac{-2(\delta\omega_n \pm \omega_n\sqrt{\delta^2 - 1})}{2}$$

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

- The two roots are imaginary when $\delta = 0$.
- The two roots are real and equal when $\delta = 1$.
- The two roots are real but not equal when $\delta > 1$.
- The two roots are complex conjugate when $0 < \delta < 1$.

We can write C(s)C(s) equation as,

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

.

Where,

- **C(s)** is the Laplace transform of the output signal, c(t)
- **R(s)** is the Laplace transform of the input signal, r(t)
- $\boldsymbol{\omega_n}$ is the natural frequency
- $\boldsymbol{\delta}$ is the damping ratio.

Follow these steps to get the response (output) of the second order system in the time domain.

- Take Laplace transform of the input signal, r(t)
- Consider the equation,

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

- Substitute R(s)R(s) value in the above equation.
- Do partial fractions of C(s)C(s) if required.
- Apply inverse Laplace transform to C(s)C(s).

Step Response of Second Order System

Consider the unit step signal as an input to the second order system.

Laplace transform of the unit step signal is,

R(s)=1sR(s)=1s

We know the transfer function of the second order closed loop control system is,

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

Case 1: δ = 0

Substitute, $\delta = 0$ in the transfer function.

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

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

Substitute, $R(s) = \frac{1}{s}$ in the above equation.

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

Apply inverse Laplace transform on both the sides.

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

So, the unit step response of the second order system when $/delta = 0$ will be a continuous time signal with constant amplitude and frequency.

Case 2: δ = 1

Substitute, $/delta = 1$ in the transfer function.

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

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

Substitute, $R(s) = \frac{1}{s}$ in the above equation.

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

Do partial fractions of $C(s)$.

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

After simplifying, you will get the values of A, B and C as $1,\ -1\, and\ -\omega_n$ respectively. Substitute these values in the above partial fraction expansion of $C(s)$.

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

Apply inverse Laplace transform on both the sides.

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

So, the unit step response of the second order system will try to reach the step input in steady state.

So, the unit step response of the second order system is having damped oscillations (decreasing amplitude) when 'δ' lies between zero and one.

Case 3: 0 < δ < 1

We can modify the denominator term of the transfer function as follows –

$$s^2 + 2\delta\omega_n s + \omega_n^2 = \left\{s^2 + 2(s)(\delta\omega_n) + (\delta\omega_n)^2\right\} + \omega_n^2 - (\delta\omega_n)^2$$

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

The transfer function becomes,

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

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

Substitute, $R(s) = \frac{1}{s}$ in the above equation.

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

Do partial fractions of $C(s)$.

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

After simplifying, you will get the values of A, B and C as $1,\ -1\ and\ -2\delta\omega_n$ respectively. Substitute these values in the above partial fraction expansion of C(s).

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

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

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

Substitute, $\omega_n\sqrt{1-\delta^2}$ as ω_d in the above equation.

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

Apply inverse Laplace transform on both the sides.

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

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

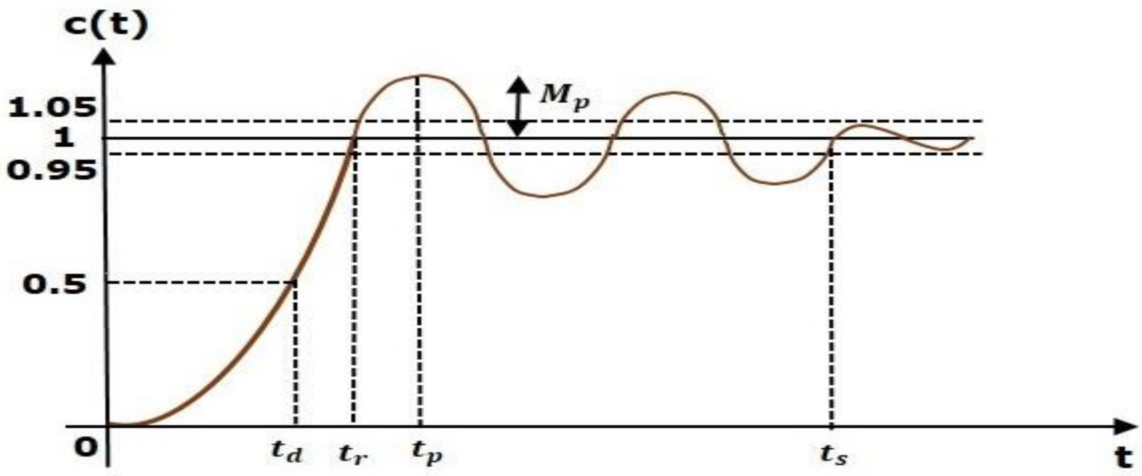
If $\sqrt{1-\delta^2} = \sin(\theta)$, then 'δ' will be cos(θ). Substitute these values in the above equation.

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

So, the unit step response of the second order system is having damped oscillations (decreasing amplitude) when 'δ' lies between zero and one.

Time Domain Specifications

The time domain specifications of the second order system. The step response of the second order system for the underdamped case is shown in the following figure.



All the time domain specifications are represented in this figure. The response up to the settling time is known as transient response and the response after the settling time is known as steady state response.

Delay Time

It is the time required for the response to reach **half of its final value** from the zero instant. It is denoted by tdtd.

Consider the step response of the second order system for t ≥ 0, when 'δ' lies between zero and one.

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

The final value of the step response is one.

Therefore, at $t = t_d$, the value of the step response will be 0.5. Substitute, these values in the above equation.

$$c(t_d) = 0.5 = 1 - \left(\frac{e^{-\delta\omega_n t_d}}{\sqrt{1-\delta^2}}\right)\sin(\omega_d t_d + \theta)$$

$$\Rightarrow \left(\frac{e^{-\delta\omega_n t_d}}{\sqrt{1-\delta^2}}\right)\sin(\omega_d t_d + \theta) = 0.5$$

By using linear approximation, you will get the **delay time t_d** as

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

Rise Time

It is the time required for the response to rise from **0% to 100% of its final value**. This is applicable for the **under-damped systems**. For the over-damped systems, consider the duration from 10% to 90% of the final value. Rise time is denoted by $\mathbf{t_r}$.

At $t = t_1 = 0$, $c(t) = 0$.

We know that the final value of the step response is one.

Therefore, at $t=t_2$, the value of step response is one. Substitute, these values in the following equation.

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

$$c(t_2) = 1 = 1 - \left(\frac{e^{-\delta\omega_n t_2}}{\sqrt{1-\delta^2}}\right)\sin(\omega_d t_2 + \theta)$$

$$\Rightarrow \left(\frac{e^{-\delta\omega_n t_2}}{\sqrt{1-\delta^2}}\right)\sin(\omega_d t_2 + \theta) = 0$$

$$\Rightarrow \sin(\omega_d t_2 + \theta) = 0$$

$$\Rightarrow \omega_d t_2 + \theta = \pi$$

$$\Rightarrow t_2 = \frac{\pi - \theta}{\omega_d}$$

Substitute t_1 and t_2 values in the following equation of **rise time**,

$$t_r = t_2 - t_1$$

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

From above equation, we can conclude that the rise time t_r and the damped frequency ω_d are inversely proportional to each other.

3. ***Peak time tp:*** It is the time required for the response to reach the maximum or Peak value of the response.

$$t_p = \frac{\pi}{\omega_n \sqrt{1-\delta^2}}$$

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

$$M_p = 100\ e^{\frac{-\pi\delta}{\sqrt{1-\delta^2}}}\ \%$$

5.Settling time ts : It is the time required for the response to reach and remain within a specified tolerance limits (usually ± 2% or ± 5%) around the steady state value.

$$t_s \simeq \frac{e^{-\delta \omega_n t}}{\sqrt{1-\delta^2}}$$

6. ***Steady state error ess*** : It is the error between the desired output and the actual output as *t* ~ 00 or under steady state conditions. The desired output is given by the reference input *r (t)* and c(t).

$$e_{ss} = \underset{t \to \infty}{Lt}\ [r(t) - c(t)]$$

P, PI and PID Controllers

- The controller (an analogue/digital circuit, and software), is trying to keep the controlled variable such as temperature, liquid level, motor velocity, robot joint angle, at a certain value called the **set point (SP)**.
- A feedback control system does this by looking at the **error (E)** signal, which is the difference between where the controlled variable (called the **process variable (PV)**) is, and where it should be.
- Based upon the error signal, the controller decides the magnitude and the direction of the signal to the actuator.

The proportional (P), the integral (I), and the derivative (D), are all basic controllers.

Types of controllers: P, I, D, PI, PD, and PID controllers

- **Proportional Control**

With proportional control, the actuator *applies a corrective force that is proportional to the amount of error*:

$$Output_p = K_p \times E$$

$Output_p$ = system output due to proportional control

K_p = proportional constant for the system called **gain**

E = error, the difference between where the controlled variable should be and where it is. $E = SP - PV$.

UNIT 3

FREQUENCY RESPONSE ANALYSIS

Frequency response:

Frequency response is the quantitative measure of the output spectrum of a system or device in response to a stimulus, and is used to characterize the dynamics of the system. It is a measure of magnitude and phase of the output as a function of frequency, in comparison to the input.

Frequency domain specification:

(i)Resonant peak (Mr): Maximum value of M (jw) when w is varied from 0 to infinite. The magnitude of resonant peak gives the information about the relative stability of the system. A large value of resonant peak implies undesirable transient response.

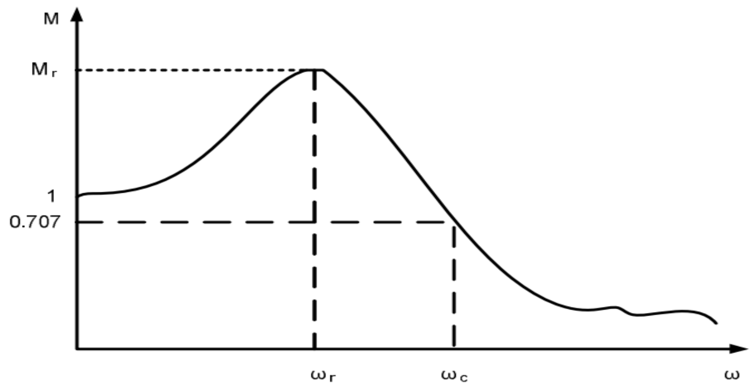
(ii) Resonant frequency (wr) : The frequency at which resonant peak (Mr) occurs. If resonant frequency is large, then the time response is fast.

(iii) Cut-off frequency (wc) : The frequency at which M (jw) has a value (1/2) ^1/2. It is the frequency at which the magnitude is 3dB below its zero frequency value.

(iv) Band-width (wb) : It is the range of frequencies in which the magnitude of a closed-loop the system is (1/2) ^1/2 times of Mr or the magnitude of the closed loop doesn't drop -3 dB.

(v) Cut-off rate: It is the slope of the log magnitude curve near cut-off frequency.

Different terms that reflect the frequency domain specification are elaborate below:



Response of Second-Order Control System

Consider the second-order system shown in Figure 3.10 where the system transfer function is written in the

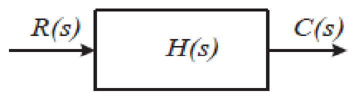


Figure 3.10: Second-order system

generalized form:

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

The system gain is 1, but any other value of the gain will change only the magnitude of the system response,

All coefficients of the polynomial in the denominator are positive. The special case when some of them have negative values will be discussed in the section presenting the problem of system stability.

The dynamic behaviour of the second-order system can be described in terms of two parameters: the natural frequency ωn, and the damping factor ζ.The order of a control system is determined by the power of s in the denominator of its transfer function. If the power of s in the denominator of transfer function of a control system is 2, then the system is said to be second-order control system. The general expression of transfer function of a second order control system is given as

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

Here, ζ and ω_n are damping ratio and natural frequency of the system respectively and we will learn about these two terms in detail later on. Therefore, the output of the system is given as

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

In the following we shall obtain the rise time, peak time, maximum overshoot and settling time of the step response of an under damped second-order system. The values will be obtained in terms of an under damped second-order system. The values will be obtained in terms of ζ and ω_n.

BODE PLOT:

A Bode plot is a standard format for plotting frequency response of LTI systems. Becoming familiar with this format is useful because:

1. It is a standard format, so using that format facilitates communication between engineers.
2. Many common system behaviours produce simple shapes (e.g. straight lines) on a Bode plot,
so it is easy to either look at a plot and recognize the system behaviour, or to sketch a plot
from what you know about the system behaviour.

The format is a log frequency scale on the horizontal axis and, on the vertical axis, phase in Degrees and magnitude in decibels. Thus, we begin with a review of decibels

1. Decibels

Definition: for voltages or other physical variables (current, velocity, pressure, etc.)

$$\text{decibels (dB)} = 20\log_{10}\frac{V_{\text{out}}}{V_{\text{in}}},$$

(Since power is propotional to voltage squared (or current, velocity, pressure, etc., squared) the definition can be rewritten in terms of power as

$$\text{decibels (dB)} = 20\log_{10}\frac{V_{\text{out}}}{V_{\text{in}}} = 10\log_{10}\left(\frac{V_{\text{out}}}{V_{\text{in}}}\right)^2 = 10\log_{10}\frac{P_{\text{out}}}{P_{\text{in}}}.$$

Common values

$$10\log_{10}2 = 20\log_{10}\sqrt{2} = 3\text{ dB}$$

$$10\log_{10}\frac{1}{2} = 20\log_{10}\frac{1}{\sqrt{2}} = -3\text{ dB}$$ "half power"

$$10\log_{10}10 = 20\log_{10}\sqrt{10} = 10\text{ dB}$$

$10\log_{10}100 = 20\log_{10}10 = 20$ dB, etc 10 dB for every factor of 10 in power

2. Bode plots

We are interested in the frequency response of an LTI system. The transfer function can be written like this:

$$H(s) = K\frac{(s-z_1)(s-z_2)\cdots}{(s-p_1)(s-p_2)(s-p_3)\cdots}$$

such that when we plug in $j\omega$ for s, we get

$$H(j\omega) = K\frac{(j\omega-z_1)(j\omega-z_2)\cdots}{(j\omega-p_1)(j\omega-p_2)(j\omega-p_3)\cdots}$$

Example 1

Obtain the Bode plot of the system given by the transfer function

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

We convert the transfer function in the following format by substituting $s = j\omega$

$$G(j\omega) = \frac{1}{2j\omega + 1}. \qquad (1)$$

We call $\omega = \frac{1}{2}$, the break point. So for

$\omega << \frac{1}{2}$, i.e., for small values of ω

$G(j\omega) \approx 1.$

Therefore taking the log magnitude of the transfer function for very small values of ω, we get

$$20\log\left|G(j\omega)\right| = 20\log(1) = 0.$$

Hence we see that below the break point the magnitude curve is approximately a constant.

For,

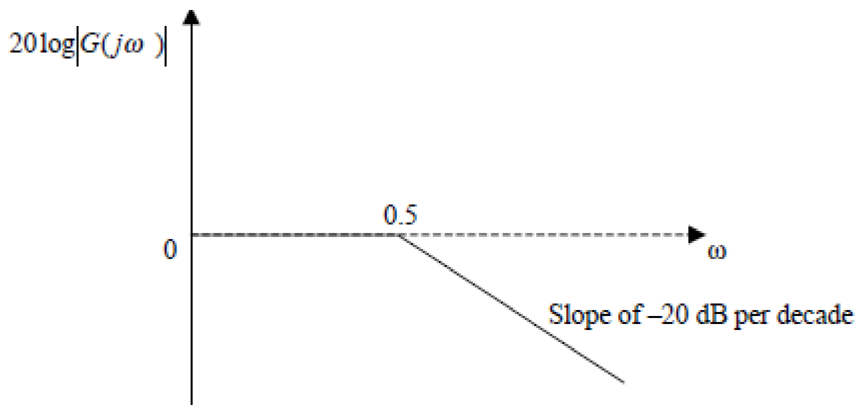
$\omega >> \frac{1}{2}$, i.e., for very large values of ω

$G(j\omega) \approx \frac{1}{2j\omega}.$

Similarly taking the log magnitude of the transfer function for very large values of ω, we have

$$20\log\left|G(j\omega)\right| = 20\log\left|\frac{1}{2j\omega}\right| = 20\log\left(\frac{1}{2\omega}\right) = 20\log(1) - 20\log(2\omega) = -20\log(2\omega).$$

So we see that, above the break point the magnitude curve is linear in nature with a slope of –20 dB per decade. The two asymptotes meet at the break point. The asymptotic bode magnitude plot is shown below.



The phase of the transfer function given by equation (1) is given by

$\phi = 0 - \tan^{-1}(2\omega) = -\tan^{-1}(2\omega).$

So for small values of ω, i.e., $\omega \simeq 0$, we get

$\phi \simeq 0.$

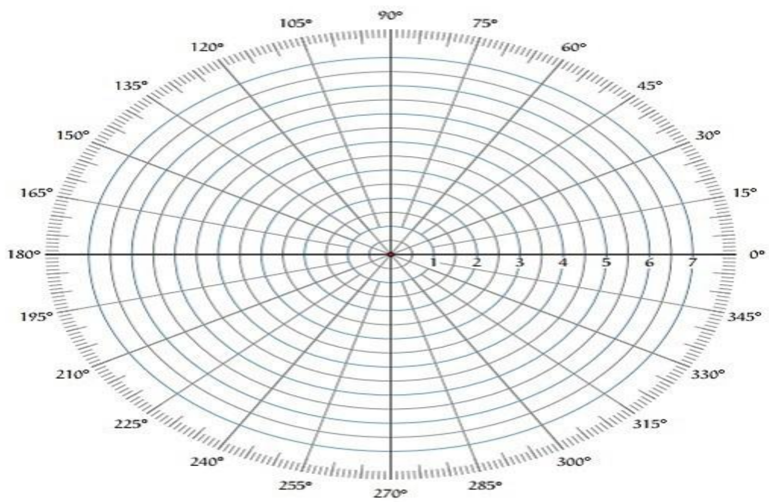
Control Systems - Polar Plots

We have two separate plots for both magnitude and phase as the function of frequency. Let us now discuss about polar plots. Polar plot is a plot which can be drawn between magnitude and phase. Here, the magnitudes are represented by normal values only.

The polar form of $G(j\omega)H(j\omega)$ is

$$G(j\omega)H(j\omega) = |G(j\omega)H(j\omega)| \angle G(j\omega)H(j\omega)$$

The **Polar plot** is a plot, which can be drawn between the magnitude and the phase angle of $G(j\omega)H(j\omega)$ by varying ω from zero to ∞. The polar graph sheet is shown in the following figure.



This graph sheet consists of concentric circles and radial lines. The **concentric circles** and the **radial lines** represent the magnitudes and phase anssgles respectively. These angles are represented by positive values in anti-clock wise direction. Similarly, we can represent angles with negative values in clockwise direction. For example, the angle 270^0 in anti-clock wise direction is equal to the angle -90^0 in clockwise direction.

Rules for Drawing Polar Plots

Follow these rules for plotting the polar plots.

- Substitute, $s = j\omega$ in the open loop transfer function.
- Write the expressions for magnitude and the phase of $G(j\omega)H(j\omega)$.
- Find the starting magnitude and the phase of $G(j\omega)H(j\omega)$ by substituting $\omega = 0$. So, the polar plot starts with this magnitude and the phase angle.
- Find the ending magnitude and the phase of $G(j\omega)H(j\omega)$ by substituting $\omega = \infty$. 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 $G(j\omega)H(j\omega)$ equal to zero and find the value(s) of ω.
- Check whether the polar plot intersects the imaginary axis, by making real term of $G(j\omega)H(j\omega)$ equal to zero and find the value(s) of ω.
- For drawing polar plot more clearly, find the magnitude and phase of $G(j\omega)H(j\omega)$ by considering the other value(s) of ω.

Example

Consider the open loop transfer function of a closed loop control system.

$$G(s)H(s) = \frac{5}{s(s+1)(s+2)}$$

Let us draw the polar plot for this control system using the above rules.

Step 1 – Substitute, $s = j\omega$ in the open loop transfer function.

$$G(j\omega)H(j\omega) = \frac{5}{j\omega(j\omega+1)(j\omega+2)}$$

The magnitude of the open loop transfer function is

$$M = \frac{5}{\omega(\sqrt{\omega^2+1})(\sqrt{\omega^2+4})}$$

The phase angle of the open loop transfer function isϕ=−900−tan−1ω−tan−1ω2ϕ=−900−tan−1ω−tan−1ω2

Step 2 − The following table shows the magnitude and the phase angle of the open loop transfer function at ω=0ω=0 rad/sec and ω=∞ω=∞ rad/sec.

Frequency (rad/sec)	Magnitude	Phase angle(degrees)
0	∞	-90 or 270
∞	0	-270 or 90

So, the polar plot starts at (∞,-90^0) and ends at (0,-270^0). The first and the second terms within the brackets indicate the magnitude and phase angle respectively.

Step 3 − Based on the starting and the ending polar co-ordinates, this polar plot will intersect the negative real axis. The phase angle corresponding to the negative real axis is -180^0 or 180^0. So, by equating the phase angle of the open loop transfer function to either -180^0 or 180^0, we will get the ωω value as 2–√2.

By substituting ω=2–√ω=2 in the magnitude of the open loop transfer function, we will get M=0.83M=0.83. Therefore, the polar plot intersects the negative real axis when ω=2–√ω=2 and the polar coordinate is (0.83,-180^0).

UNIT 4

STABILITY AND COMPENSATOR DESIGN

Stability Definition:

The stability of a system relates to its response to inputs or disturbances. A system which remains in a constant state unless affected by an external action and which returns to a constant state when the external action is removed can be considered to be stable.

Characteristics equation

The local behaviour of a system of differential equations,

$$\frac{dX_i}{dt} = f_i(X_1, \ldots, X_n) \qquad (i = 1, \ldots, n)$$

near an equilibrium point depends on the roots (eigenvalues) of the characteristic equation

$$|A - \lambda I| = 0 \qquad (4.1)$$

where $A = (a_{ij})$ is the matrix of first partial derivatives $\frac{\partial f_i}{\partial X_j}$ evaluated at the equilibrium point.

If the real parts of all the roots are negative, the system returns to equilibrium after a small perturbation. If the real parts of all the roots are positive, the system moves away from equilibrium (is locally unstable). If some roots have positive and some negative real parts, the behaviour of the system depends on how it is perturbed; it sometimes returns to equilibrium but for other displacements moves away. In biological systems we usually assume the perturbations to be unconstrained so that eventually the system will be displaced in a direction which allows the positive root to lead the system away from equilibrium. A single zero real part gives a neutral or passive equilibrium, but multiple zero roots can give unbounded solutions (unstable equilibrium). If a root is complex the system oscillates at a frequency given by the imaginary part while the amplitude behaves according to the real part of the root.

Routh Hurwitz Stability Criterion:

Any pole of the system lies on the right hand side of the origin of the s plane, it makes the system unstable. On the basis of this condition A. Hurwitz and E.J. Routh started investigating the necessary and sufficient conditions of stability of a system. We will discuss two criteria for stability of the system.

A first criterion is given by A. Hurwitz and this criterion is also known as **Hurwitz Criterion for stability** or **Routh Hurwitz Stability Criterion.**

Hurwitz Criterion

With the help of characteristic equation, we will make a number of Hurwitz determinants in order to find out the stability of the system. We define characteristic equation of the system as, now there are n determinants for nth order characteristic equation. Determinants from the coefficients of the characteristic equation. The step by step procedure for kth order characteristic equation is written below: Determinant one : The value of this determinant is given by |a1| where a1 is the coefficient of sn-1 in the characteristic equation.

Routh-Hurwitz Stability Criterion

The Routh-Hurwitz criterion is a method for determining whether a linear system is stable or not by examining the locations of the roots of the characteristic equation. The method determines only if there are roots that lie outside of the left half plane; it does not actually compute the roots. Consider the characteristic equation

$$1+GH(s)=0 \qquad \Rightarrow \boxed{D(s)=a_n s^n + a_{n-1}s^{n-1} + \ \ldots\ + a_1 s + a_0 = 0}$$

To determine whether this system is stable or not, check the following conditions:

1. Two necessary but not sufficient conditions that all the roots have negative real parts are

   a) All the polynomial coefficients must have the same sign.
   b) All the polynomial coefficients must be nonzero.

2. If condition (1) is satisfied, then compute the Routh-Hurwitz array as follows:

$$\begin{array}{c|ccccc}
s^n & a_n & a_{n-2} & a_{n-4} & a_{n-6} & \cdots \\
s^{n-1} & a_{n-1} & a_{n-3} & a_{n-5} & a_{n-7} & \cdots \\
s^{n-2} & b_1 & b_2 & b_3 & & \cdots \\
s^{n-3} & c_1 & c_2 & c_3 & & \cdots \\
s^{n-4} & & & \vdots & & \\
\vdots & & & \vdots & & \\
s^1 & & & \vdots & & \\
s^0 & & & \vdots & &
\end{array}$$

where the a_i are the polynomial coefficients, and the coefficients in the rest of the table are computed using the following pattern:

$$b_1 = \frac{-1}{a_{n-1}}\begin{vmatrix} a_n & a_{n-2} \\ a_{n-1} & a_{n-3} \end{vmatrix} = \frac{-1}{a_{n-1}}(a_n a_{n-3} - a_{n-2}a_{n-1}) \qquad b_2 = \frac{-1}{a_{n-1}}\begin{vmatrix} a_n & a_{n-4} \\ a_{n-1} & a_{n-5} \end{vmatrix}$$

$$b_3 = \frac{-1}{a_{n-1}}\begin{vmatrix} a_n & a_{n-6} \\ a_{n-1} & a_{n-7} \end{vmatrix} \ \ldots \qquad c_1 = \frac{-1}{b_1}\begin{vmatrix} a_{n-1} & a_{n-3} \\ b_1 & b_2 \end{vmatrix} \qquad c_2 = \frac{-1}{b_1}\begin{vmatrix} a_{n-1} & a_{n-5} \\ b_1 & b_3 \end{vmatrix} \qquad \ldots$$

3. The necessary condition that all roots have negative real parts is that all the elements of the first column of the array have the same sign. The number of changes of sign equals the number of roots with positive real parts.
4. Special Case 1: The first element of a row is zero, but some other elements in that row are nonzero. In this case, simply replace the zero element by "ε", complete the table development, and then interpret the results assuming that "ε" is a small number of the same sign as the element above it. The results must be interpreted in the limit as $\varepsilon \to 0$.
5. Special Case 2: All the elements of a particular row are zero. In this case, some of the roots of the polynomial are located symmetrically about the origin of the *s*-plane, e.g., a pair of purely imaginary roots. The zero row will always occur in a row associated with an odd power of *s*. The row just above the zero row holds the coefficients of the auxiliary polynomial. The roots of the auxiliary polynomial are the symmetrically placed roots. Be careful to remember that the coefficients in the array skip powers of *s* from one coefficient to the next.

Phase Margin

The Phase Margin is the amount of phase that needs to be added to a system such that the magnitude will be just unity while the phase is 1800. The figure below is showing 'theta' to be the phase margin

Often control engineers consider a system to be adequately stable if it has a phase margin of at least 300.

The root locus technique in control system was first introduced in the year 1948 by Evans. Any physical system is represented by a transfer function in the form of

$$G(s) = k \times \frac{numerator\ of\ s}{denomerator\ of\ s}$$

We can find poles and zeros from G(s). The location of poles and zeros are crucial keeping view stability, relative stability, transient response and error analysis. When the system put to service stray inductance and capacitance get into the system, thus changes the location of poles and zeros. In root locus technique in control system we will evaluate the position of the roots, their locus of movement and associated information. These information will be used to comment upon the system performance.

Advantages of Root Locus Technique

1. Root locus technique in control system is easy to implement as compared to other methods.
2. With the help of root locus we can easily predict the performance of the whole system.
3. Root locus provides the better way to indicate the parameters.

Now there are various terms related to root locus technique that we will use frequently in this article.

1. **Characteristic Equation Related to Root Locus Technique:** 1 + G(s)H(s) = 0 is known as characteristic equation. Now on

differentiating the characteristic equation and on equating dk/ds equals to zero, we can get break away points.

2. **Break away Points:** Suppose two root loci which start from pole and moves in opposite direction collide with each other such that after collision they start moving in different directions in the symmetrical way. Or the break away points at which multiple roots of the characteristic equation 1 + G(s)H(s)= 0 occur. The value of K is maximum at the points where the branches of root loci break away. Break away points may be real, imaginary or complex.
3. **Break in Point :** Condition of break in to be there on the plot is written below : Root locus must be present between two adjacent zeros on the real axis.
4. **Centre of Gravity :** It is also known centroid and is defined as the point on the plot from where all the asymptotes start. Mathematically, it is calculated by the difference of summation of poles and zeros in the transfer function when divided by the difference of total number of poles and total number of zeros. Centre of gravity is always real and it is denoted by σ_A.

$$\sigma_A = \frac{(Sum\ of\ real\ parts\ of\ poles) - (Sum\ of\ real\ parts\ of\ zeros)}{N - M}$$

   Where, N is number of poles and M is number of zeros.

5. **Asymptotes of Root Loci :** Asymptote originates from the center of gravity or centroid and goes to infinity at definite some angle. Asymptotes provide direction to the root locus when they depart break away points.

6. **Angle of Asymptotes:** Asymptotes makes some angle with the real axis and this angle can be calculated from the given formula, Where, p = 0, 1, 2 ..(N-M-1)N is the total number of poles M is the total number of zeros.

7. **Angle of Arrival or Departure:** We calculate angle of departure when there exists complex poles in the system. Angle of departure can be

calculated as 180-{(sum of angles to a complex pole from the other poles)-(sum of angle to a complex pole from the zeros)}.

$$Angle\ of\ asymptotes\ = \frac{(2p+1)\times 180}{N-M}$$

Where, p = 0, 1, 2 .... (N-M-1)
N is the total number of poles
M is the total number of zeros.

8. **Intersection of Root Locus with the Imaginary Axis:** In order to find out the point of intersection root locus with imaginary axis, we have to use Routh Hurwitz criterion. First, we find the auxiliary equation then the corresponding value of K will give the value of the point of intersection.
9. **Gain Margin:** We define gain margin as a by which the design value of the gain factor can be multiplied before the system becomes unstable. Mathematically it is given by the formula.

$$Gain\ margin = \frac{Value\ of\ K\ at\ the\ imaginary\ axes\ cross\ over}{Design\ value\ of\ K}$$

10. **Phase Margin :** Phase margin can be calculated from the given formula.

$$Phase\ margin = 180 + \angle(G(jw)H(jw))$$

11. **Symmetry of Root Locus:** Root locus is symmetric about the x axis or the real axis.
    - Magnitude Criteria : At any points on the root locus we can apply magnitude criteria as,

    $$|G(s)H(s)| = 1$$

    Using this formula we can calculate the value of K at any desired point.

        - Using Root Locus Plot : The value of K at any s on the root locus is given by

$$K = \frac{product\ of\ all\ of\ the\ vector\ lengths\ drawn\ from\ the\ poles\ of\ G(s)H(s)\ to\ s}{product\ of\ all\ of\ the\ vector\ lengths\ drawn\ from\ the\ zeros\ of\ G(s)H(s)\ to\ s}$$

Procedure to Plot Root Locus

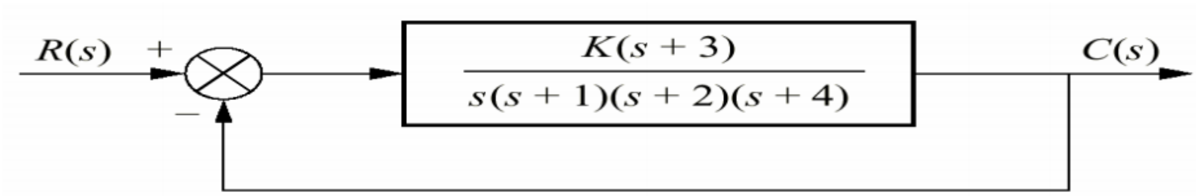
Keeping all these points in mind we are able to draw the **root locus plot** for any kind of system. Now let us discuss the procedure of making a root locus.

1. Find out all the roots and poles from the open loop transfer function and then plot them on the complex plane.
2. All the root loci starts from the poles where k = 0 and terminates at the zeros where K tends to infinity. The number of branches terminating at infinity equals to the difference between the number of poles & number of zeros of G(s)H(s).
3. Find the region of existence of the root loci from the method described above after finding the values of M and N.
4. Calculate break away points and break in points if any.
5. Plot the asymptotes and centroid point on the complex plane for the root loci by calculating the slope of the asymptotes.
6. Now calculate angle of departure and the intersection of root loci with imaginary axis.
7. Now determine the value of K by using any one method that I have described above.

By following above procedure you can easily draw the **root locus plot** for any open loop transfer function.

8. Calculate the gain margin.
9. Calculate the phase margin.
10. You can easily comment on the stability of the system by using Routh array.

Sketch the root locus for the system shown in figure.



Solution : Begin by calculating the asymptotes :

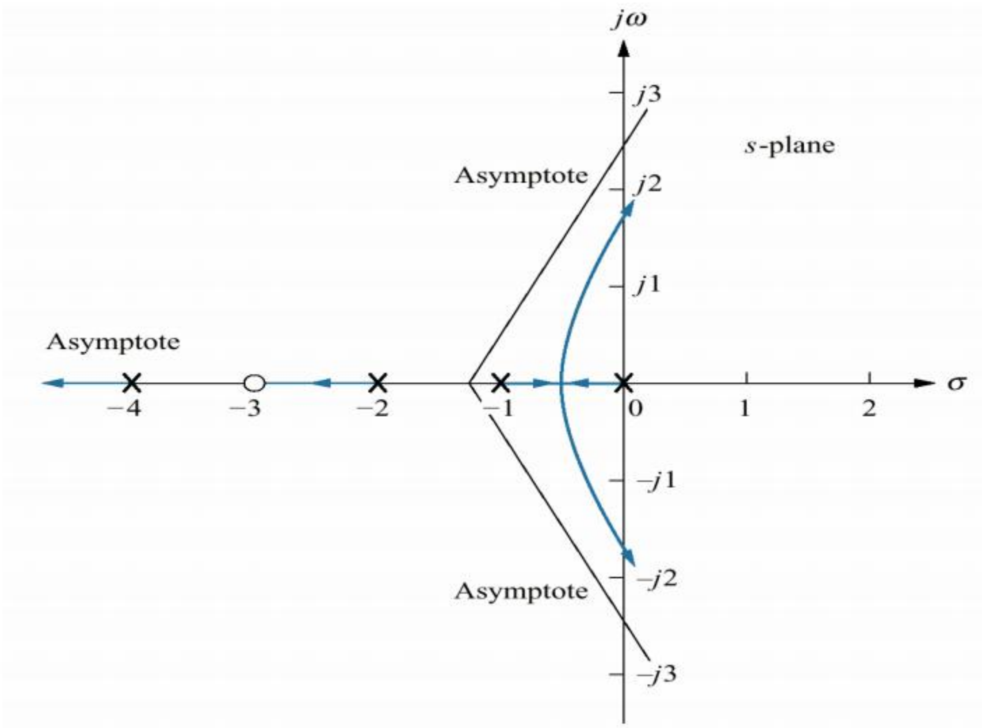
$$\sigma_a = \frac{\sum \textit{finite}\text{ poles } - \sum \text{finite zeros}}{(\text{number of finite poles}) - (\text{number of finite zeros})} = \frac{(-1-2-4)-(-3)}{4-1} = -\frac{4}{3}$$

$$\theta_a = \frac{(2k+1)\pi}{(\text{number of finite poles}) - (\text{number of finite zeros})} = \pi/3 \quad \text{for k} = 0$$

$$= \pi \quad \text{for k} = 1$$

$$= 5\pi/3 \quad \text{for k} = 2$$

If the values for k continued to increase, the angels would begin repeat. The number of lines obtained equals the difference between the number of finite poles and the number of finite zeros. Rule 4 states that the root locus begins at the open loop poles and ends at the open loop zeros. For the example, there are more open loop poles than open loop zeros. Thus, there must be zeros at infinity. The asymptotes tell us how we get to these zeros at infinity.



Compensation

A closed-loop system is usually an unstable system.Hence, because it is unstable, there must be some kind of compensators than can compensate the stability of a closed loop system.

Compensators are used to alter the output response of a system in order to accommodate to the set of desired criteria. This is achieved by introducing additional poles and/or zeros to the system transfer function.

The introduction of additional zeros and/or poles will speed-up the response of the system.

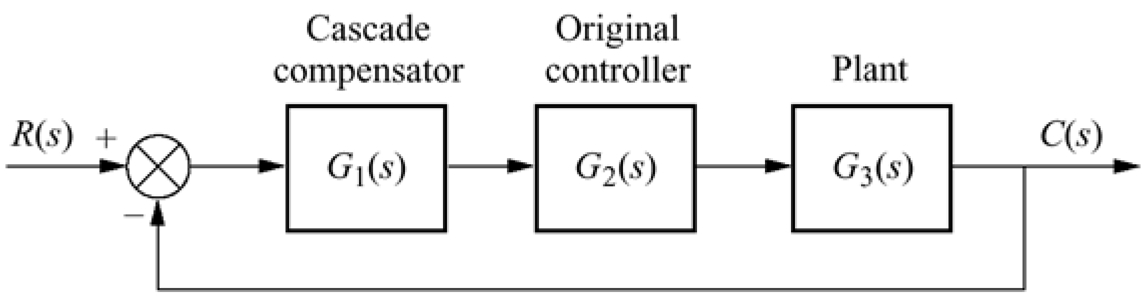
• When we introduce additional poles/zeros, we areactually improving the transient response of the system,as well as reducing the steady-state errors.

• Additional poles will eventually improve the steady-state characteristics, while additional zeros will improve the transient response.

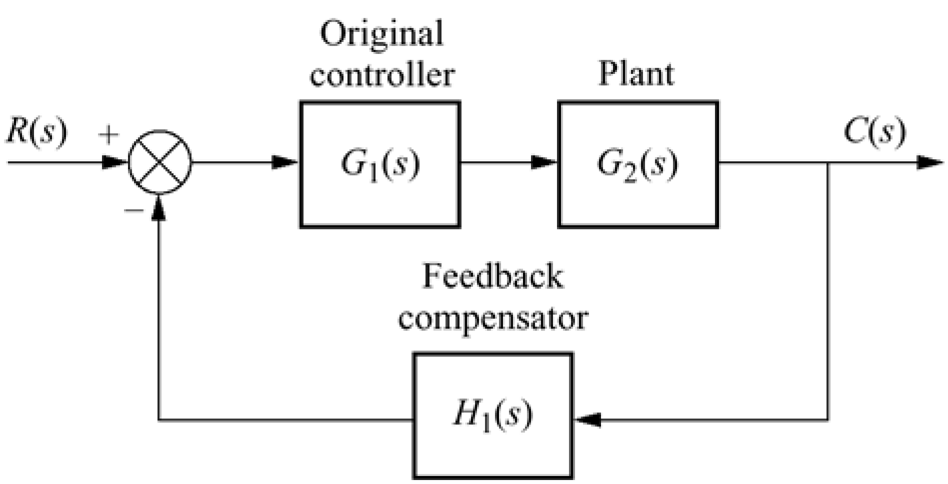
• Recall that poles are also called integrators in s-domain while zeros are called differentiators.

Compensation Configuration

Two configuration of compensation are commonly used:
– (a) cascade compensation.
– (b) feedback compensation.



(a)

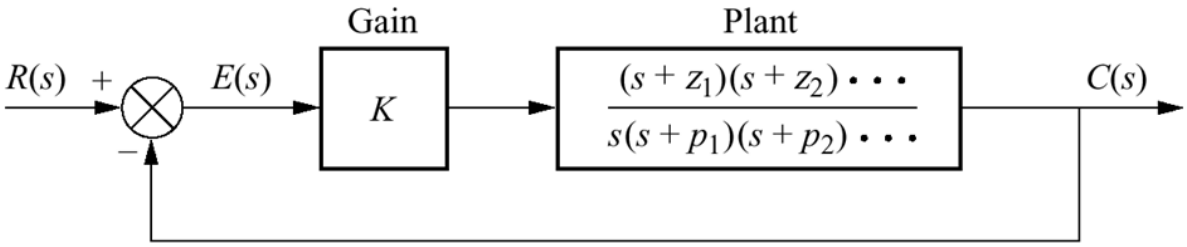


(b)

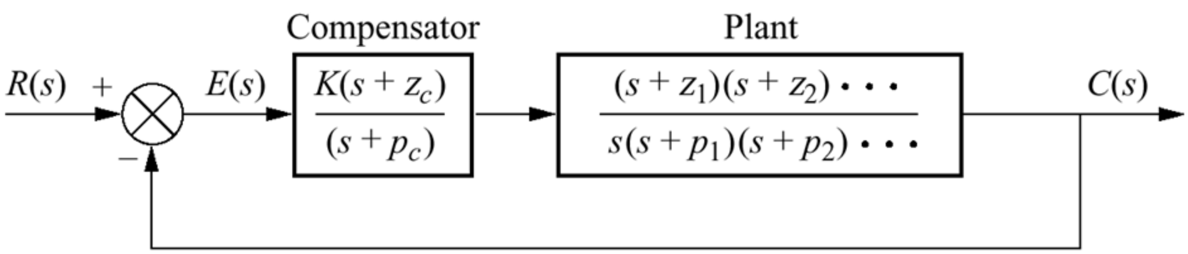
Design of Lag Compensation

We note that the design of ideal integral compensation involves the use of pure integrator. Now, this time, we will not use pure integrator. We will use poles and zeros that are close to the origin, but not necessarily on the origin.

• Let us consider a type 1 uncompensated system which is shown below:

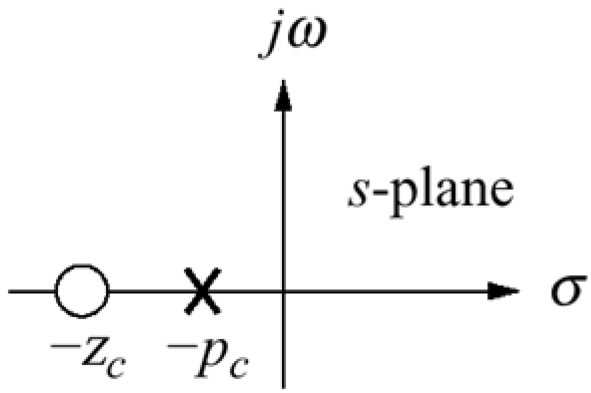


The above system can be improved by adding a compensator transfer function in the feed-forward section of the loop:



• The pole-zero plot of this compensator is shown below:

$$G_c(s) = \frac{(s + z_c)}{(s + p_c)}$$



UNIT V

STATE VARIABLE ANALYSIS

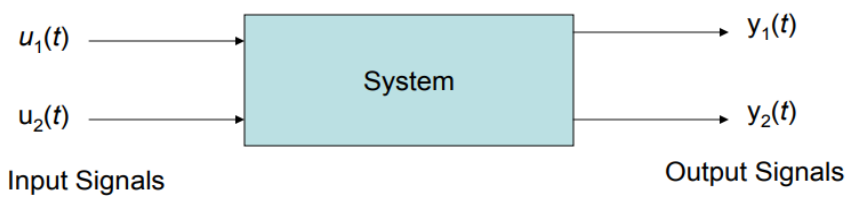
Concept of state variables:

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 behavior in the absence of any external forces affecting the system.The state variables of a dynamic system are the variables making up the smallest set of variables that determine the state of the dynamic system.

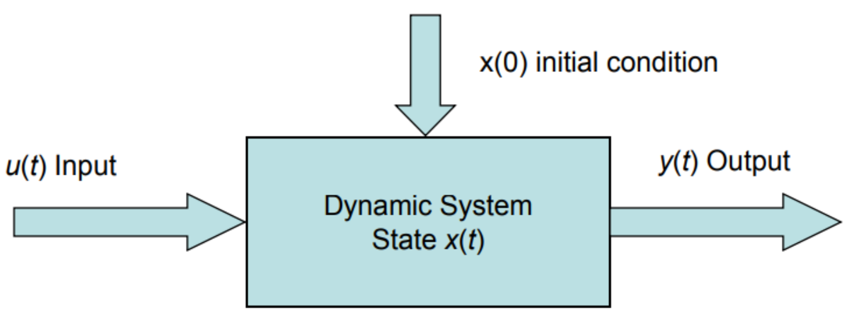
The State Variables of a Dynamic System

• The state of a system is a set of variables such that the knowledge of these variables and the input functions will, with the equations describing the dynamics, provide the future state and output of the system.

• For a dynamic system, the state of a system is described in terms of a set of state variables.



State Variables of a Dynamic System:



The state variables describe the future response of a system, given the present state, the excitation inputs, and the equations describing the dynamics.

State models for linear and time invariant Systems:

State model:

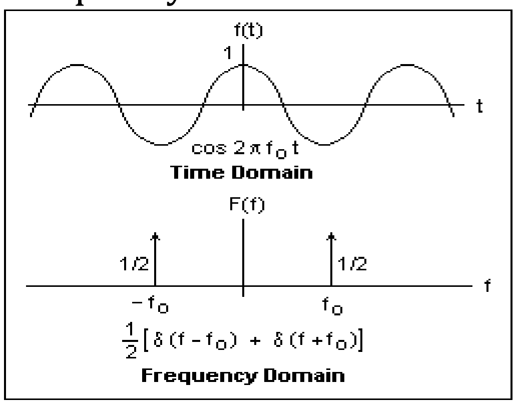
In control engineering, a state-space representation is a mathematical model of a physical system as a set of input, output and state variables related by first-order differential equations. "State space" refers to the Euclidean space in which the variables on the axes are the state variables.

Linear Time Invariant Systems:

Linear Time Invariant Systems (LTI systems) are a class of systems used in signals and systems that are both linear and time invariant. Linear systems are systems whose outputs for a linear combination of inputs are the same as a linear combination of individual responses to those inputs. Time invariant systems are systems where the output does not depend on when an input was applied. These properties make LTI systems easy to represent and understand graphically.

LTI systems are superior to simple state machines for representation because they have more memory. LTI systems, unlike state machines, have memory of past states and have to ability predicted the future. LTI systems are used to predict long-term behavior in a system. So, they are often used to model systems like power plants. Another important application of LTI systems is electrical circuits. These circuits, made up of inductors, transistors, and resistors, are the basis upon which modern technology is built.

Time invariant systems are systems where the output for a particular input does not change depending on when that input was applied. A time invariant systems that takes in signal and produces output will also, when excited by signal , produce the time shifted output .Thus, the entirety of an LTI system can be described by a single function called its impulse response. This function exists in the time domain of the system. For an arbitrary input, the output of an LTI system is the convolution of the input signal with the system's impulse response. Conversely, the LTI system can also be described by its transfer function. The transfer function is the Laplace transform of the impulse response. This transformation changes the function from the time domain to the frequency domain. This transformation is important because it turns differential equations into algebraic equations, and turns convolution into multiplication. In the frequency domain, the output is the product of the transfer function with the transformed input. The shift from time to frequency is illustrated in the following image.



Shifting from the time to the frequency domain[1]

In addition to linear and time invariant, LTI systems are also memory systems, invertible, causal, real, and stable. That means they have memory, they can be inverted, they depend only on current and past events, they have fully real inputs and outputs, and they produce bounded output for bounded input.

Solution of state and output equation in controllable canonical form:

State-Space Canonical Forms For any given system, there are essentially an infinite number of possible state space models that will give the identical input/output dynamics. Thus, it is desirable to have certain standardized state space model structures: these are the so-called canonical forms. Given a system transfer function, it is possible to obtain each of the canonical models. And, given any particular canonical form it is possible to transform it to another form. Consider the system defined by

$$\overset{(n)}{y} + a_1 \overset{(n-1)}{y} + \ldots + a_{n-1}\dot{y} + a_n y = b_0 \overset{(n)}{u} + b_1 \overset{(n-1)}{u} + \ldots + b_{n-1}\dot{u} + b_n u$$

where u is the input, y is the output and $\overset{(n)}{y}$ represents the n^{th} derivative of y with respect to time. Taking the Laplace transform of both sides we get:

$$Y(s)\left(s^n + a_1 s^{n-1} + \ldots + a_{n-1}s + a_n\right) = U(s)\left(b_0 s^n + b_1 s^{n-1} + \ldots + b_{n-1}s + b_n\right)$$

which yields the transfer function:

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \ldots + b_{n-1}s + b_n}{s^n + a_1 s^{n-1} + \ldots + a_{n-1}s + a_n} \tag{1}$$

Given the a system having transfer function as defined in (1) above, we will define the controllable canonical and observable canonical forms.

Controllable Canonical Form:

The controllable canonical form arranges the coefficients of the transfer function denominator across one row of the A matrix:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \ldots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} b_n - a_n b_0 & b_{n-1} - a_{n-1} b_0 & \ldots & b_1 - a_1 b_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u$$

The controllable canonical from is useful for the pole placement controller design technique.

Example1.1. Consider the system given by

$$\frac{U(s)}{Y(s)} = \frac{s+3}{s^2+3s+2}$$

Obtain a state space representation in controllable canonical form.

By inspection, $n = 2$ (the highest exponent of s), therefore $a_1 = 3$, $a_2 = 2$, $b_0 = 0$, $b_1 = 1$ and $b_2 = 3$. Therefore, we can simply write the state space model as follows:

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y &= \begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

Observable Canonical Form:

The observable canonical form is defined in terms of the transfer function coefficients of (1) as follows:

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} &= \begin{bmatrix} 0 & 0 & \dots & 0 & -a_n \\ 1 & 0 & \dots & 0 & -a_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} b_n - a_n b_0 \\ b_{n-1} - a_{n-1} b_0 \\ \vdots \\ b_1 - a_1 b_0 \end{bmatrix} u \\ y &= \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u \end{aligned}$$

Note the relationship between the observable and controllable forms:

$$\begin{aligned} A_{obs} &= A_{cont}^T \\ B_{obs} &= C_{cont}^T \\ C_{obs} &= B_{cont}^T \\ D_{obs} &= D_{cont} \end{aligned}$$

Example1.2. Given the system transfer function of example 1.1, find the observable canonical form state space model. Recall that by inspection, we have n = 2 (the highest exponent of s), and therefore a1 = 3, a2 = 2, b0 = 0, b1 = 1 and b2 = 3. Thus, we can write the observable canonical form model as follows:

$$
\begin{aligned}
\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix} &= \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} u \\
y &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\end{aligned}
$$

Diagonal Canonical Form

The diagonal canonical form is a state space model in which the poles of the transfer function are arranged diagonally in the A matrix. Given the system transfer function having a denominator polynomial that can be factored into distinct (p1 6= p2 6= . . . 6= pn) roots as follows:

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \ldots + b_{n-1}s + b_n}{(s+p_1)(s+p_2)\ldots(s+p_n)}$$

The denominator polynomial can be rewritten by partial fraction expansion as follows:

$$b_0 + \frac{c_1}{s+p_1} + \frac{c_2}{s+p_2} + \ldots + \frac{c_n}{s+p_n}$$

Then the diagonal canonical form state space model can be written as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} -p_1 & & & 0 \\ & -p_2 & & \\ & & \ddots & \\ 0 & & & -p_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} c_1 & c_2 & \ldots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + b_0 u$$

Example1.3.

Given the system transfer function of example 1.1, find the diagonal canonical form state space model. 3 The transfer function of the system can be re-written with the denominator factored as follows:

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

therefore $p_1 = 1$ and $p_2 = 2$. Trivially, the partial fraction expansion of the denominator gives $c_1 = 2$ and $c_2 = 1$ and the diagonal from model can be written as:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Jordan Form:

The Jordan form is a type of diagonal form canonical model in which the poles of the transfer function are arranged diagonally in the A matrix. Consider the case in which the denominator polynomial of the transfer function involves multiple repeated roots:

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \ldots + b_{n-1}s + b_n}{(s+p_1)^3(s+p_4)(s+p_5)\ldots(s+p_n)}$$

The denominator polynomial can be rewritten by partial fraction expansion as follows:

$$b_0 + \frac{c_1}{(s+p_1)^3} + \frac{c_1}{(s+p_1)^2} + \frac{c_1}{s+p_1} + \frac{c_2}{s+p_2} + \ldots + \frac{c_n}{s+p_n}$$

Then the Jordan canonical form state space model can be written as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \left[\begin{array}{ccc|ccc} -p_1 & 1 & 0 & 0 & \ldots & 0 \\ 0 & -p_1 & 1 & \vdots & & \vdots \\ 0 & 0 & -p_1 & 0 & \ldots & 0 \\ \hline 0 & \ldots & 0 & -p_4 & & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \ldots & 0 & 0 & & -p_n \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} c_1 & c_2 & \ldots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + b_0 u$$

Concepts of controllability and observability:

Controllability: In the world of control engineering, there are a slew of systems available that need to be controlled. The task of a control engineer is to design controller and compensator units to interact with these pre-existing systems. However, some systems simply cannot be controlled (or, more often, cannot be controlled in specific ways). The concept of controllability refers to the ability of a controller to arbitrarily alter the functionality of the system plant. Complete state controllability (or simply controllability if no other context is given) describes the ability of an external input to move the internal state of a system from any initial state to any other final state in a finite time interval

We will start off with the definitions of the term controllability, and the related term reachability

Controllability

A system with internal state vector x is called controllable if and only if the system states can be changed by changing the system input.

Reachability

A particular state x1 is called reachable if there exists an input that transfers the state of the system from the initial state x0 to x1 in some finite time interval [t0, t).

Controllability Matrix

For LTI (linear time-invariant) systems, a system is reachable if and only if its controllability matrix, ζ, has a full row rank of p, where p is the dimension of the matrix A, and p × q is the dimension of matrix B.

$$\zeta = [B \quad AB \quad A^2B \quad \cdots \quad A^{p-1}B] \in R^{p \times pq}$$

A system is controllable or "Controllable to the origin" when any state x_1 can be driven to the zero state *x = 0* in a finite number of steps.

A system is controllable when the rank of the system matrix A is *p*, and the rank of the controllability matrix is equal to:

$$Rank(\zeta) = Rank(A^{-1}\zeta) = p$$

Observability:

The state-variables of a system might not be able to be measured for any of the following reasons:

The location of the particular state variable might not be physically accessible (a capacitor or a spring, for instance).

There are no appropriate instruments to measure the state variable, or the state-variable might be measured in units for which there does not exist any measurement device.

The state-variable is a derived "dummy" variable that has no physical meaning.

If things cannot be directly observed, for any of the reasons above, it can be necessary to calculate or estimate the values of the internal state variables, using only the input/output relation of the system, and the output history of the system from the starting time. In other words, we must ask whether or not it is possible to determine what the inside of the system (the internal system states) is like, by only observing the outside performance of the system (input and output)? We can provide the following formal definition of mathematical observability

Observability

A system with an initial state, {\displaystyle x(t_{0})} $x(t_0)$ is observable if and only if the value of the initial state can be determined from the system output y(t) that has been observed through the time interval {\displaystyle t_{0}<t<t_{f}} $t_0 < t < t_f$. If the initial state cannot be so determined, the system is unobservable.

Complete Observability

A system is said to be completely observable if all the possible initial states of the system can be observed. Systems that fail these criteria are said to be unobservable.

Detectability:

A system is Detectable if all states that cannot be observed decay to zero asymptotically.

Constructability:

A system is constructible if the present state of the system can be determined from the present and past outputs and inputs to the system. If a system is observable, then it is also constructible. The relationship does not work the other way around.

Constructability

A state x is unconstructable at a time t1 if for every finite time t < t1 the zero input response of the system is zero for all time t.

A system is completely state constructible at time t1 if the only state x that is unconstructable at t0 is x = 0.

If a system is observable at an initial time t0, then it is constructible at some time t > t0, if it is constructible at t1.

Observability Matrix

The observability of the system is dependent only on the system states and the system output, so we can simplify our state equations to remove the input terms:

$$x'(t) = Ax(t)$$
$$y(t) = Cx(t)$$

Therefore, we can show that the observability of the system is dependent only on the coefficient matrices A and C. We can show precisely how to determine whether a system is observable, using only these two matrices. If we have the observability matrix Q:

$$Q = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{p-1} \end{bmatrix}$$

Example 5.1:

Consider the following system with measurements

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$
$$y(k) = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

Observability matrix :

The observability matrix for this second-order system is given by

$$\mathcal{O} = \begin{bmatrix} \mathbf{C}_d \\ \mathbf{C}_d\mathbf{A}_d \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 7 & 10 \end{bmatrix}$$

Since the rows of the matrix Q are linearly independent, then $\mathrm{rank}\mathcal{O} = 2 = n,$ i.e. the system under consideration is observable.

Another way to test the completeness of the rank of square matrices is to find their determinants. In this case

$$\det\mathcal{O} = -4 \neq 0 \;\Leftrightarrow\; \text{full rank} \;=\; n = 2$$

Example 5.2:

Consider a case of an unobservable system, which can be obtained by slightly modifying Example 5.1. The corresponding system and measurement matrices are given by

$$\mathbf{A}_d = \begin{bmatrix} 1 & -2 \\ -3 & -4 \end{bmatrix}, \quad \mathbf{C}_d = [1 \quad 2]$$

The observability matrix is

$$\mathcal{O} = \begin{bmatrix} 1 & 2 \\ -5 & -10 \end{bmatrix}$$

so that $\text{rank}\mathcal{O} = 1 < 2$, and the system is unobservable.

Effect of state feedback:

In state feedback, the value of the state vector is fed back to the input of the system. We define a new input, r, and define the following relationship:

$$u(t) = r(t) + Kx(t)$$

K is a constant matrix that is external to the system, and therefore can be modified to adjust the locations of the poles of the system. This technique can only work if the system is controllable.

Closed-Loop System

If we have an external feedback element K, the system is said to be a closed-loop system. Without this feedback element, the system is said to be an open-loop system. Using the relationship we've outlined above between r and u, we can write the equations for the closed-loop system:

$$x' = Ax + B(r + Kx)$$
$$x' = (A + BK)x + Br$$

Now, our closed-loop state equation appears to have the same form as our open loop state equation, except that the sum (A + BK) replaces the matrix A. We can define the closed-loop state matrix as:

$$A_{cl} = (A_{ol} + BK)$$

Acl is the closed-loop state matrix, and Aol is the open-loop state matrix. By altering K, we can change the Eigen values of this matrix, and therefore change the locations of the poles of the system. If the system is controllable, we can find the characteristic equation of this system as:

$$\alpha(s) = |sI - A_{cl}| = |sI - (A_{ol} + BK)|$$

Computing the determinant is not a trivial task, the determinant of that matrix can be very complicated, especially for larger systems. However, if we transform the system into controllable canonical form, the calculations become much easier. Another alternative to compute K is by Ackerman's Formula.

Ackerman's Formula

Consider a linear feedback system with no reference input:

$$u(t) = -Kx(t)$$

where *K* is a vector of gain elements. Systems of this form are typically referred to as **regulators**. Notice that this system is a simplified version of the one we introduced above, except that we are ignoring the reference input. Substituting this into the state equation gives us:

$$x' = Ax - BKx$$

Ackerman's Formula gives us a way to select these gain values *K* in order to control the location's of the system poles. Using Ackerman's formula, if the system is controllable, we can select arbitrary poles for our regulator system.

$$K = [0 \quad 0 \quad \cdots \quad 1]\zeta^{-1}a(z) \qquad \text{[Ackerman's Formula]}$$

where *a(z)* is the desired characteristic equation of the system and ζ is the controllability matrix of the original system.

The gain *K* can be computed in MATLAB using Ackerman's formula with the following command:

Effects of the Addition of an Observer to State Feedback

In the pole placement design process, it is assumed that the actual state is available for feedback. In practice, the actual state may not be measurable, so it is necessary to design a state observer. Therefore the design process involves a two stage process. First stage includes determination of the feedback gain matrix to yield the desired characteristic equation and the second stage involves the determination of the observer gain matrix to yield the desired observer characteristic equation. The closed loop poles of the observed-state feedback control system consist of the poles due to the poleplacement design and the poles due to the observer design. If the order of the plant is n, then the observer is also nth order and the resulting characteristic equation for the entire closed-loop system becomes the order of 2n. The desired closed-loop poles to be generated by state feedback are chosen in such a way that the system satisfies the performance requirements. The poles of the observer are usually chosen so that the observer response is much faster than system response. A rule of thumb is to choose an observer response at least two to five times faster than system response. The maximum speed of the observer is limited only by the noise and sensitivity problem involved in the control system. Since the observer poles are placed left of the desired closed loop poles in the pole placement process, the closed loop poles will dominate the response.