UNIT-1
STATE VARIABLE REPRESENTATIONS BY VARIOUS METHODS

The State Variables of a Dynamic System

The State Differential Equation

Signal-Flow Graph State Variables

The Transfer Function from the State Equation

Introduction

- In the previous chapter, we used Laplace transform to obtain the transfer function models representing linear, time-invariant, physical systems utilizing block diagrams to interconnect systems.

- In Chapter 3, we turn to an alternative method of system modeling using **time-domain methods**.

- In Chapter 3, we will consider physical systems described by an **nth-order ordinary differential equations**.

- Utilizing a set of variables known as **state variables**, we can obtain a set of first-order differential equations.

- The time-domain state variable model lends itself easily to computer solution and analysis.

Time-Varying Control System

- With the ready availability of digital computers, it is convenient to consider the time-domain formulation of the equations representing control systems.

- The time-domain is the mathematical domain that incorporates the response and description of a system in terms of time *t*.

- The time-domain techniques can be utilized for nonlinear, time-varying, and multivariable systems (a system with several input and output signals).

- A time-varying control system is a system for which one or more of the parameters of the system may vary as a function of time.

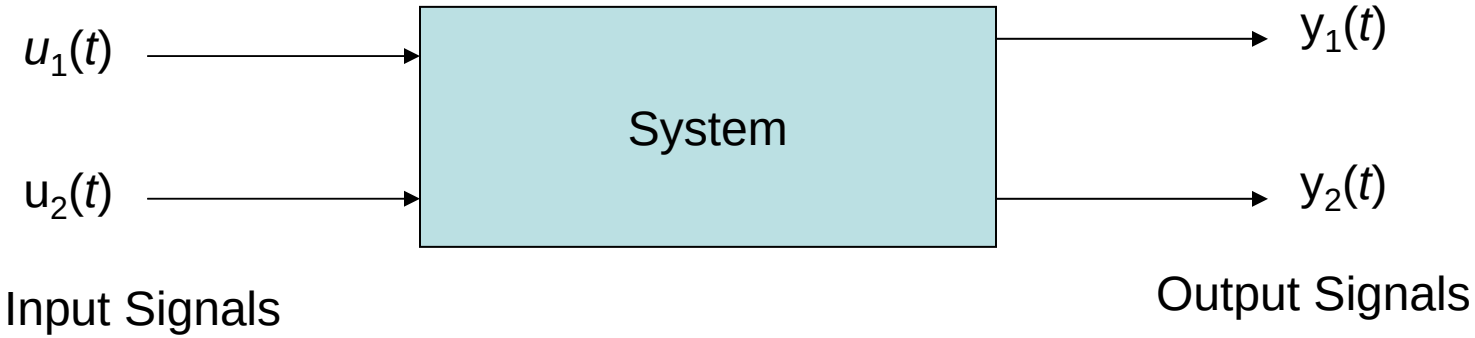
- For example, the mass of a missile varies as a function of time as the fuel is expended during flight

Terms

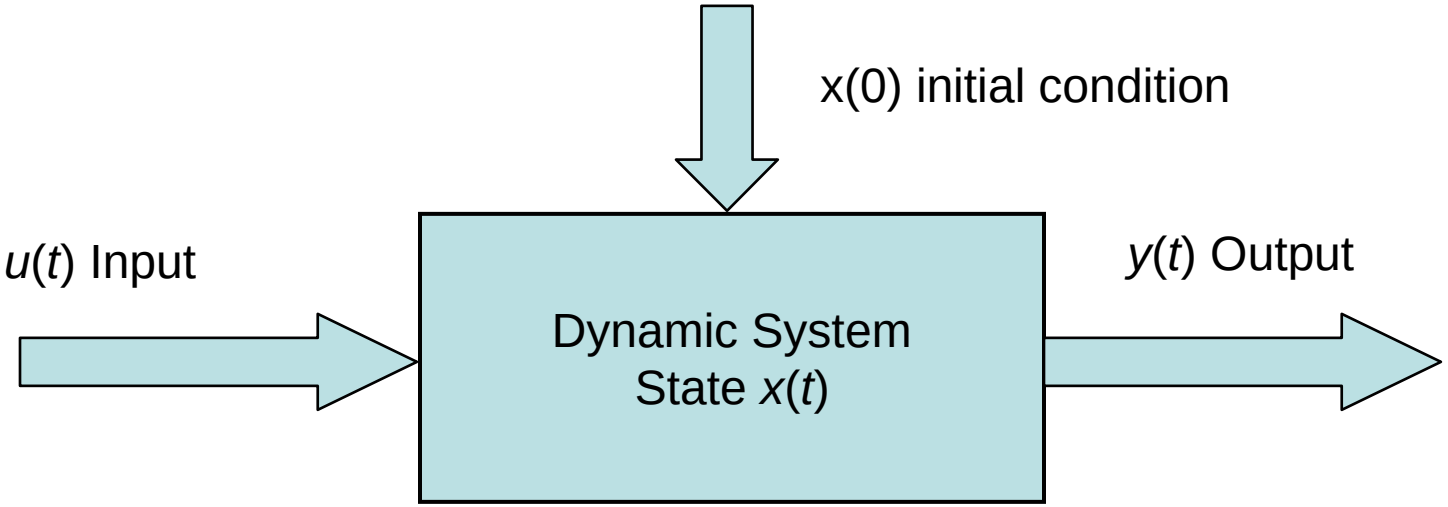
- **State:** The state of a dynamic system is the smallest set of variables (called state variables) so that the knowledge of these variables at $t = t_0$, together with the knowledge of the input for $t \geq t_0$, determines the behavior of the system for any time $t \geq t_0$.
- **State Variables:** 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.
- **State Vector:** If *n* state variables are needed to describe the behavior of a given system, then the *n* state variables can be considered the *n* components of a vector *x*. Such vector is called a state vector.
- **State Space:** The *n*-dimensional space whose coordinates axes consist of the x_1 axis, x_2 axis, .., x_n axis, where x_1, x_2, .., x_n are state variables, is called a state space.
- **State-Space Equations:** In state-space analysis, we are concerned with three types of variables that are involved in the modeling of dynamic system: input variables, output variables, and state variables.

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

The State Differential Equation

The state of a system is described by the set of first-order differential equations written in terms of the state variables (x_1, x_2, .., x_n)

$$\dot{x} = \frac{dx}{dt}$$

$$\dot{x}_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + b_{11}u_1 + \ldots + b_{1m}u_m$$

$$\dot{x}_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + b_{21}u_1 + \ldots + b_{2m}u_m$$

$$\dot{x}_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n + b_{n1}u_1 + \ldots + b_{nm}u_m$$

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ . \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ . & . & . \\ a_{n1} & a_{n2} & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ . \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11}....b_{1m} \\ ..... \\ b_{n1} \ldots . b_{nm} \end{bmatrix}\begin{bmatrix} u_1 \\ . \\ u_m \end{bmatrix}$$

A x B u

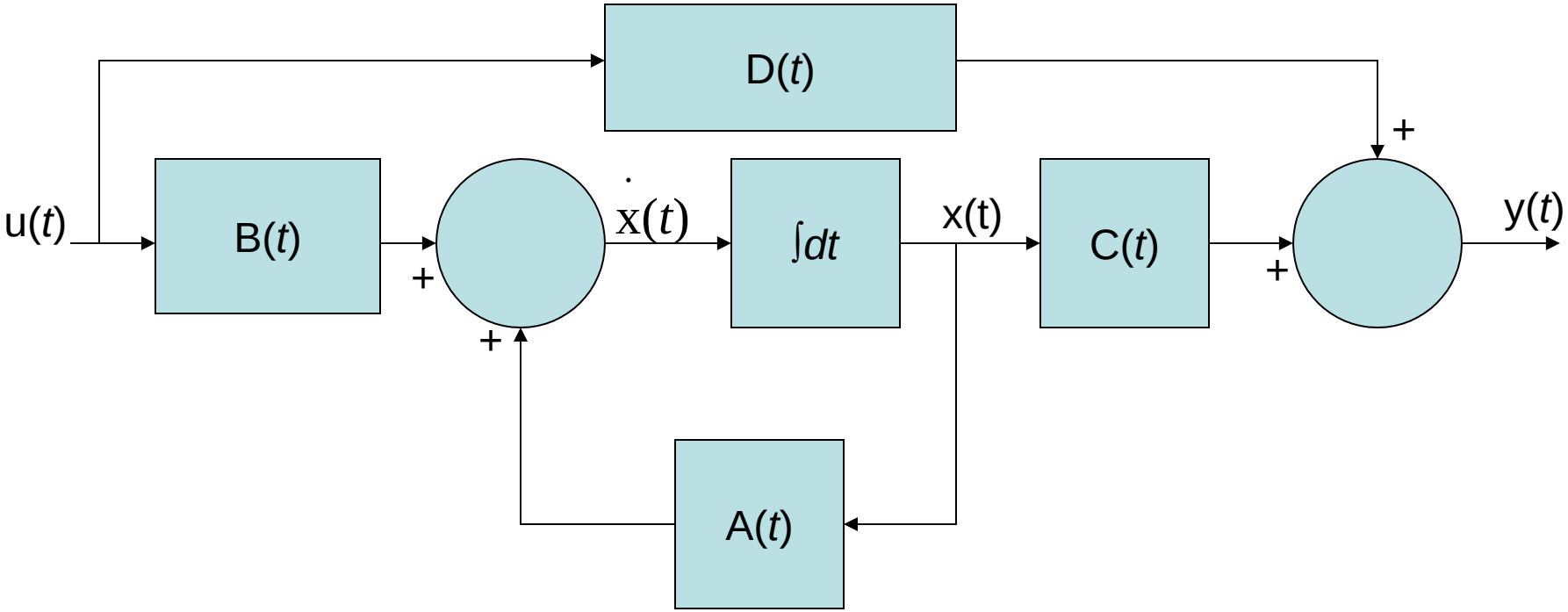
$$\dot{x} = Ax + Bu \text{ (State differential equation)}$$

$$y = Cx + Du \text{ (Output equation - output signals)}$$

A : State matrix; B : input matrix

C : Output matrix; D : direct transmission matrix

Block Diagram of the Linear, Continuous Time Control System



$$\dot{x}(t) = A(t)x(t) + B(t)\,u(t)$$

$$y(t) = C(t)\,x(t) + D(t)\,u(t)$$

Mass Grounded, *M* (kg)

Mechanical system described by the first-order differential equation

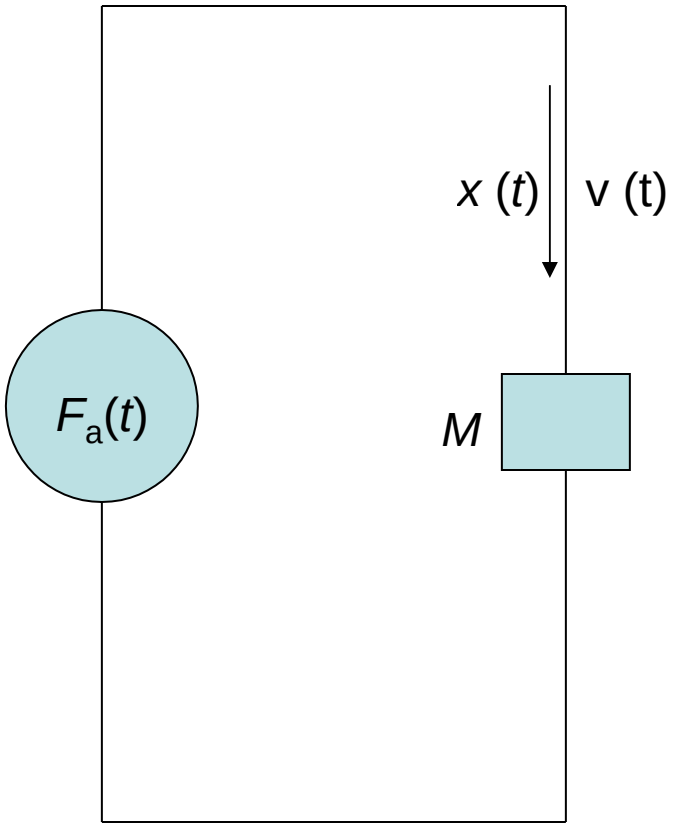
Appied torque $T_a(t)$ (N - m)

Linear velocity $v(t)$ (m/sec)

Linear position $x(t)$ (m)

$$F_a(t) = M\frac{dv}{dt} = M\frac{d^2x(t)}{dt^2}$$

$$v(t) = \frac{1}{M}\int_{t_0}^{t} F_a(t)dt$$



Mechanical Example: Mass-Spring Damper

A set of state variables sufficient to describe this system includes the position and the velocity of the mass, therefore, we will define a set of state variables as (x_1, x_2)

$$x_1(t) = y(t)$$

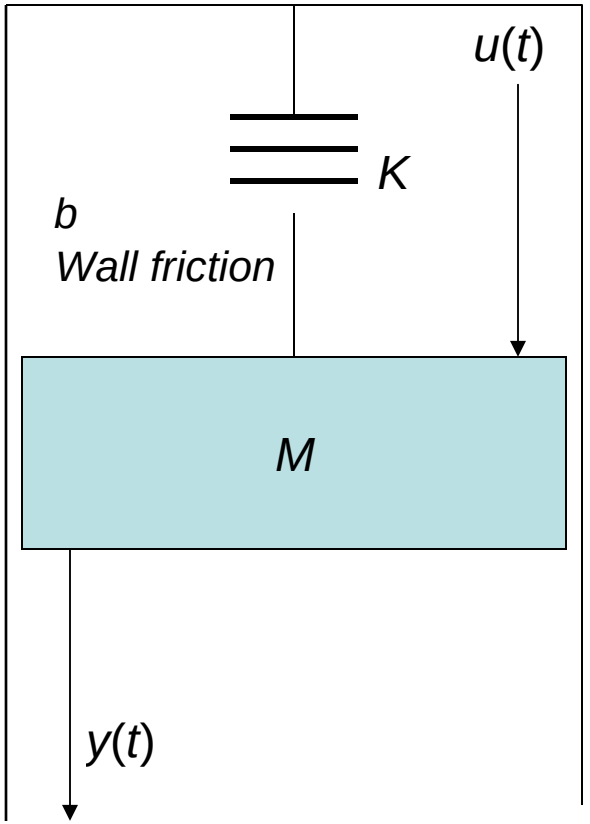
$$x_2(t) = \frac{dy(t)}{dt}$$

$$M\frac{d^2y}{dt^2} + b\frac{dy}{dt} + ky = u(t)$$

$$M\frac{dx_2}{dt} + bx_2 + kx_1 = u(t)$$

$$\frac{dx_1}{dt} = x_2;$$

$$\frac{dx_2}{dt} = -\frac{b}{m}x_2 - \frac{k}{M}x_1 + \frac{1}{M}u$$



k :Spring constant

Example 1: Consider the previous mechanical system. Assume that the system is linear. The external force $u(t)$ is the input to the system, and the displacement $y(t)$ of the mass is the output. The displacement y(t) is measured from the equilibrium position in the absence of the external force. This system is a single-input-single-output system.

$$m\ddot{y} + b\dot{y} + ky = u$$

This is a second order system. It means it involves two integrators.

Let us define two variables: $x_1(t)$ and $x_2(t)$

$$x_1(t) = y(t);\ x_2(t) = \dot{y}(t);\ \text{then } \dot{x}_1 = x_2$$

$$\dot{x}_2 = -\frac{k}{m}x_1 - \frac{b}{m}x_2 + \frac{1}{m}u$$

The output equation is : $y = x_1$

In a vector matrix form, we have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u \quad \text{(State Equation)}$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{(Output Equation)}$$

The state equation and the output equation are in the standard form :

$$\dot{x} = Ax + Bu;\ y = Cx + Du$$

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix},\ B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix},\ C = \begin{bmatrix} 1 & 0 \end{bmatrix},\ D = 0$$

Electrical and Mechanical Counterparts

Energy	Mechanical	Electrical
Kinetic	**Mass / Inertia** $0.5\ mv^2$ / $0.5\ j\omega^2$	**Inductor** $0.5\ Li^2$
Potential	**Gravity:** mgh **Spring:** $0.5\ kx^2$	**Capacitor** $0.5\ Cv^2$
Dissipative	**Damper / Friction** $0.5\ Bv^2$	Resistor Ri^2

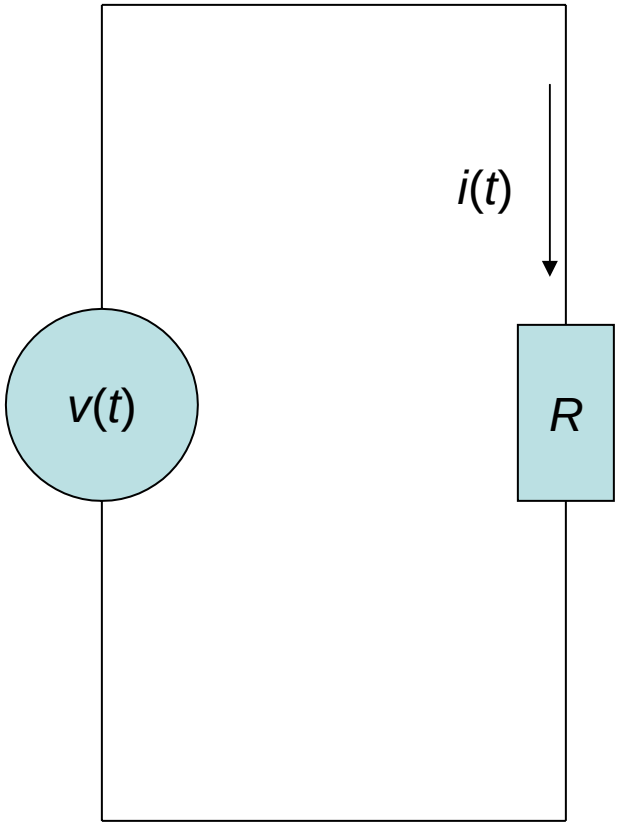
Resistance, *R* (ohm)

Appied voltage $v(t)$

Current $i(t)$

$$v(t) = Ri(t)$$

$$i(t) = \frac{1}{R}v(t)$$



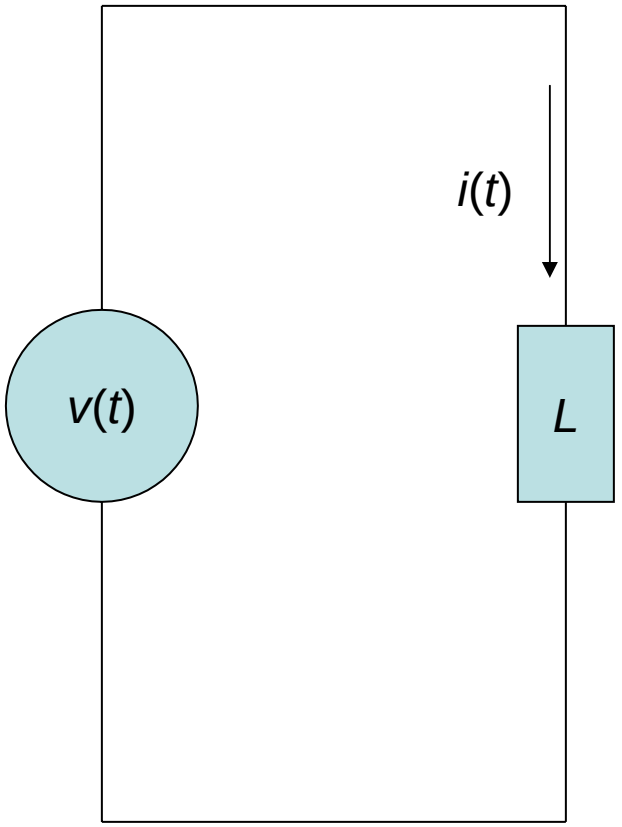
Inductance, L (H)

Appied voltage $v(t)$

Current $i(t)$

$$v(t) = L\frac{di(t)}{dt}$$

$$i(t) = \frac{1}{L}\int_{t_0}^{t} v(t)dt$$



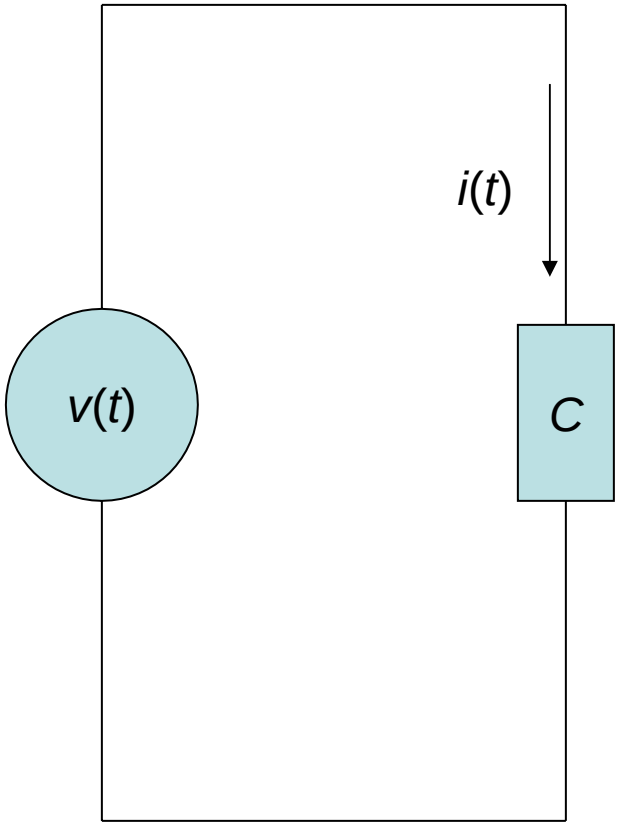
Capacitance, *C* (F)

Appied voltage $v(t)$

Current $i(t)$

$$v(t) = \frac{1}{C}\int_{t_0}^{t} i(t)dt$$

$$i(t) = C\,\frac{dv(t)}{dt}$$



Electrical Example: An *RLC* Circuit

$x_1 = v_C(t);\ x_2 = i_L(t)$

$\xi = (1/2)Li_L^2 + (1/2)Cv_c^2$

$x_1(t_0)$ and $x_2(t_0)$ is the total initial energy of the network

USE KCL at the junction

$$i_c = C\frac{dv_c}{dt} = +u(t) - i_L$$

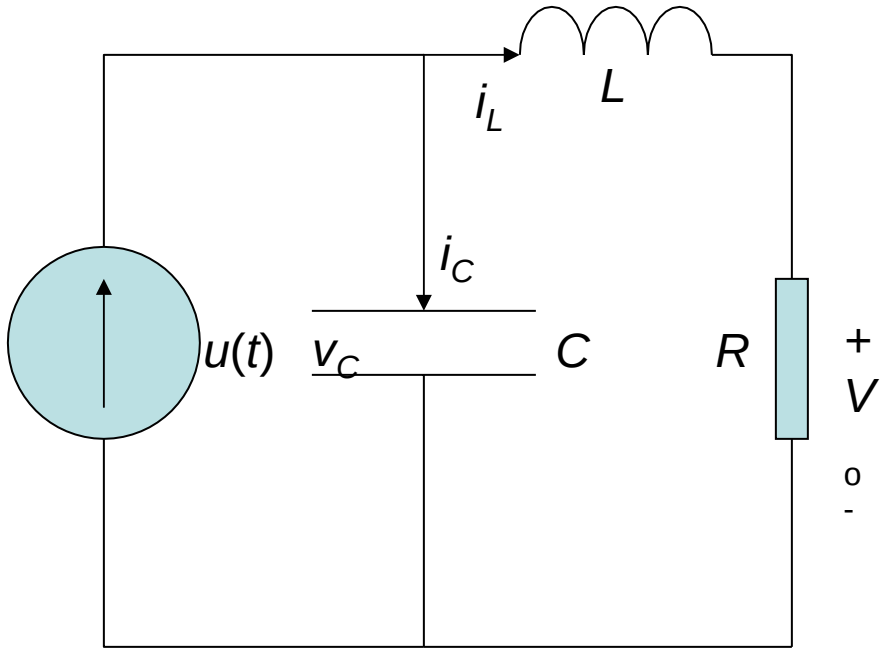
$$L\frac{di_L}{dt} = -Ri_L + v_c$$

The output of the system is represented by : $v_o = Ri_L(t)$

$$\frac{dx_1}{dt} = -\frac{1}{C}x_2 + \frac{1}{C}u(t)$$

$$\frac{dx_2}{dt} + \frac{1}{L}x_1 - \frac{R}{L}x_2$$

The output signal is then : $y_1(t) = v_o(t) = Rx_2$



Example 2: Use Equations from the *RLC* circuit

$$\dot{x} = \begin{bmatrix} 0 & -\frac{1}{C} \\ \frac{1}{L} & -\frac{R}{L} \end{bmatrix} x + \begin{bmatrix} \frac{1}{C} \\ 0 \end{bmatrix} u(t)$$

The output is

$$y = \begin{bmatrix} 0 & R \end{bmatrix} x$$

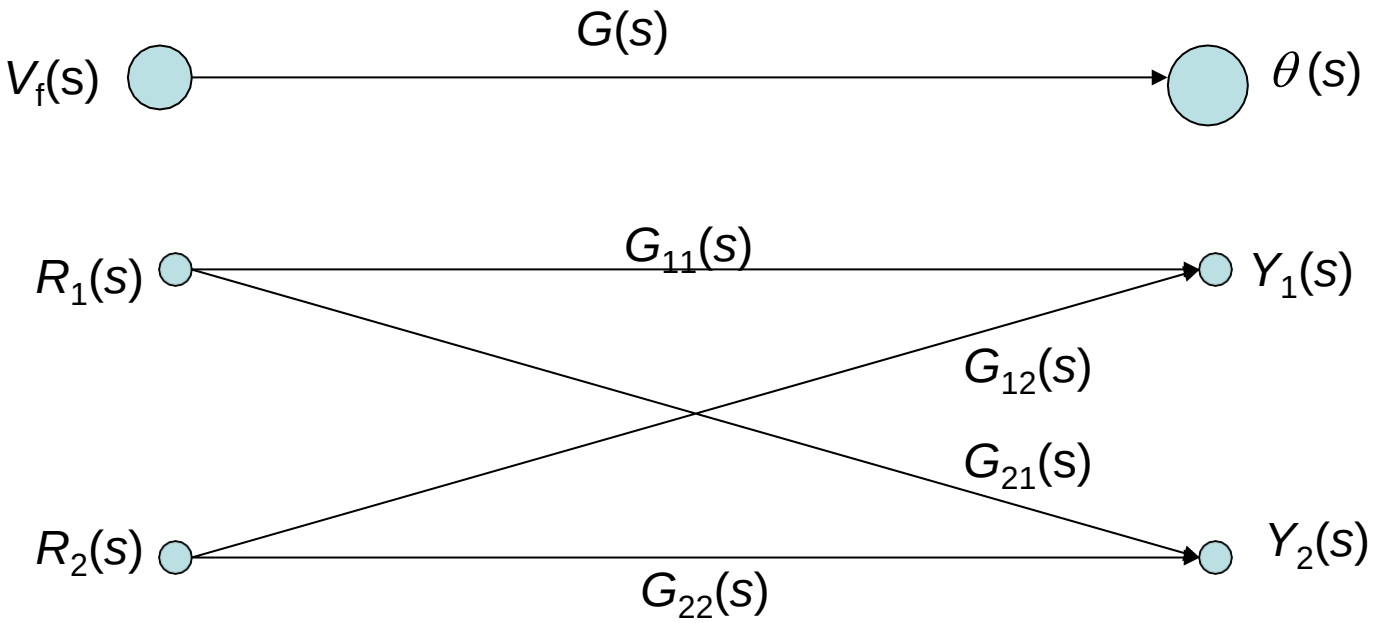
When $R = 3, L = 1, C = 1/2$, we have

$$\dot{x} = \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix} x + \begin{bmatrix} 2 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & 3 \end{bmatrix} x$$

Signal-Flow Graph Model

A signal-flow graph is a diagram consisting of nodes that are connected by several directed branches and is a graphical representation of a set of linear relations. Signal-flow graphs are important for feedback systems because feedback theory is concerned with the flow and processing of signals in system.



Read Examples : 2.8 - 2.11

Mason's Gain Formula for Signal Flow Graphs

In many applications, we wish to determine the relationship between an input and output variable of the signal flow diagram. The transmittance between an input node and output node is the overall gain between these two nodes.

$$P = \frac{1}{\Delta}\sum_k P_k \Delta_k$$

P_k = path gain of k_{th} forward path

Δ = determinant of graph

= 1- (sum of all individual loop gain) +

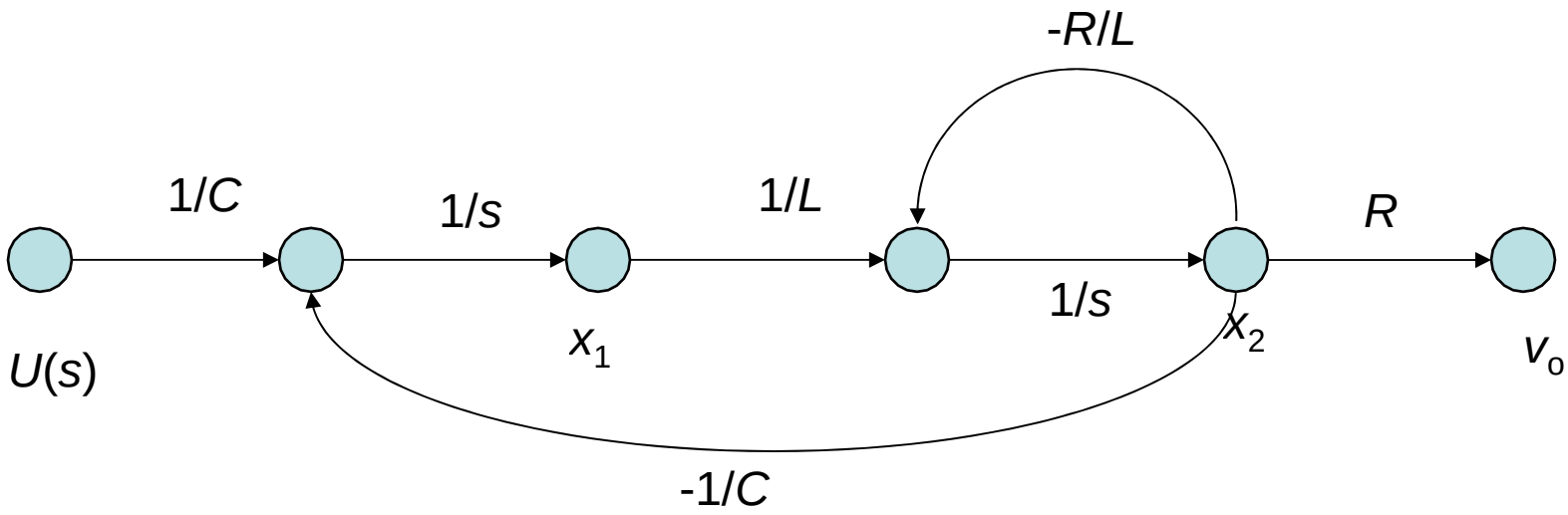
(sum of gain of all possible combinations of two nontouching loops)

- (sum of gain products of all possible combinations of these nontouching loops) + ..

$$= 1 - \sum_a L_a + \sum_{b,c} L_b L_c - \sum_{d,e,f} L_d L_c L_f$$

Δ_k = cofactor of the kth forward path determinant of the graph with the loops touching the kth forward path removed, that is, the cofactor Δ_k is obtained from Δ by removing the loops that touch path P_k.

Signal-Flow Graph State Models



$$G(s) = \frac{V_o(s)}{U(s)} = \frac{\square\alpha}{s^2 + \beta s + \gamma}$$

$$\dot{x}_1 = -\frac{1}{C} x_2 + \frac{1}{C} u(t)$$

$$\dot{x}_2 = \frac{1}{L} x_1 - \frac{R}{L} x_2 \; ; v_o = R x_2$$

$$\frac{Vo(s)}{U(s)} = \frac{R/LCs^2}{1+(R/Ls)+(1/LCs^2)}; = \frac{R/LC}{s^2 + (R/L)s + (1/LC)}$$

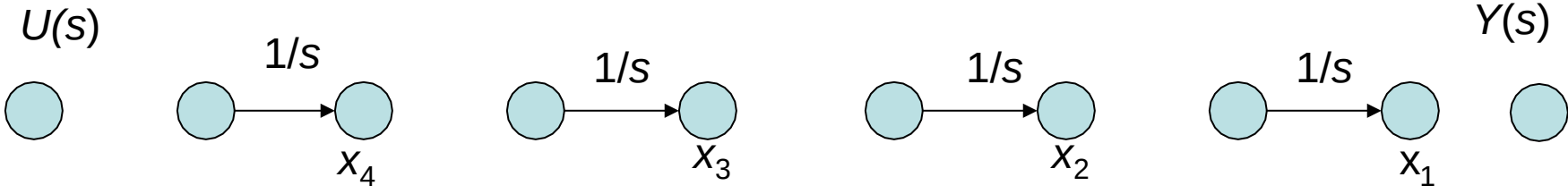
$$G(s) = \frac{Y(s)}{U(s)} = \frac{s^m + b_{m-1}s^{m-1} + ... + b_1 s + b_o}{s^n + a_{n-1}s^{n-1} + ... + a_1 s + a_o}$$

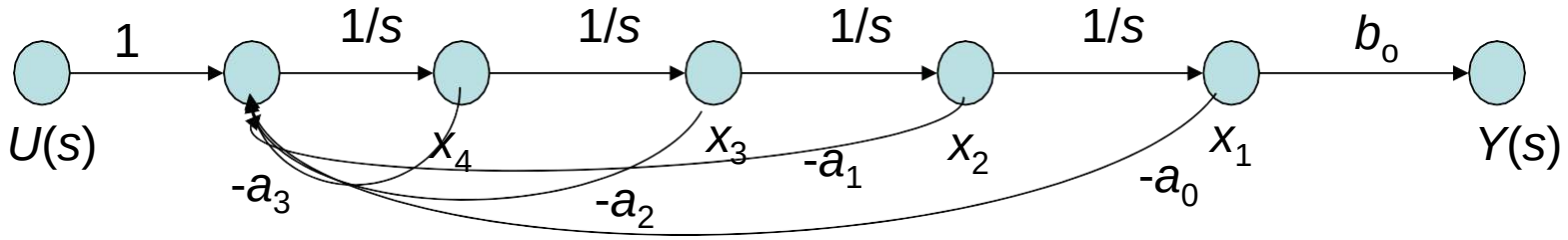
$$G(s) = \frac{Y(s)}{U(s)} = \frac{s^{-(n-m)} + b_{m-1}s^{-(n-m+1)} + ... + b_1 s^{-(n-1)} + b_o s^{-n}}{1 + a_{n-1}s^{-1} + ... + a_1 s^{-(n-1)} + a_o s^{-n}}$$

$$G(s) = \frac{Y(s)}{U(s)} = \frac{\sum_k P_k \Delta k}{\Delta}$$

$$G(s) = \frac{\sum_k P_k}{1 - \sum_{q=1}^{N} Lq} = \frac{\text{Some of the forward - path factors}}{\text{1- sum of the feedback loop factor}}$$

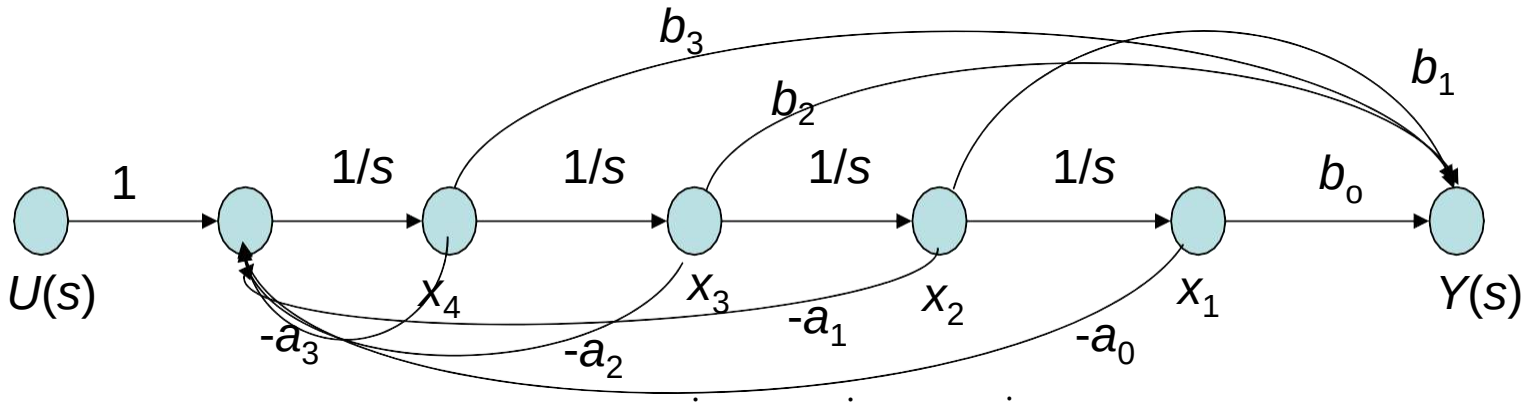
Phase Variable Format: Let us initially consider the fourth-order transfer function. Four state variables (x_1, x_2, x_3, x_4); Number of integrators equal the order of the system.





$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_0}{s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0}$$

$$= \frac{b_0 s^{-4}}{1 + a_3 s^{-1} + a_2 s^{-2} + a_1 s^{-3} + a_0 s^{-4}}$$



$$G(s)=\frac{b_3 s^3 + b_2 s^2 + b_1 s + b_0}{s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0}$$

$$= \frac{b_3 s^{-1} + b_2 s^{-2} + b_1 s^{-3} + b_0 s^{-4}}{1 + a_3 s^{-1} + a_2 s^{-2} + a_1 s^{-3} + a_0 s^{-4}}$$

$$x_1 = x_2 ;\ x_2 = x_3 ;\ x_3 = x_4$$

$$\dot{x}_4 = -a_0 x_1 - a_1 x_2 - a_2 x_3 - a_3 x_4 + u$$

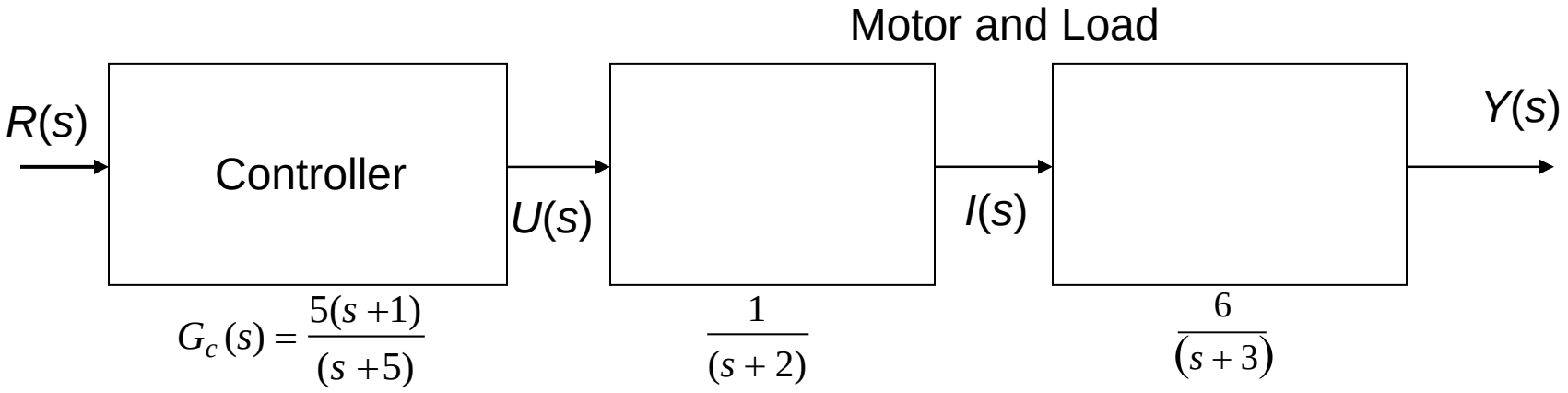
$$y(t) = b_0 x_1 + b_1 x_2 + b_2 x_3 + b_3 x_4$$

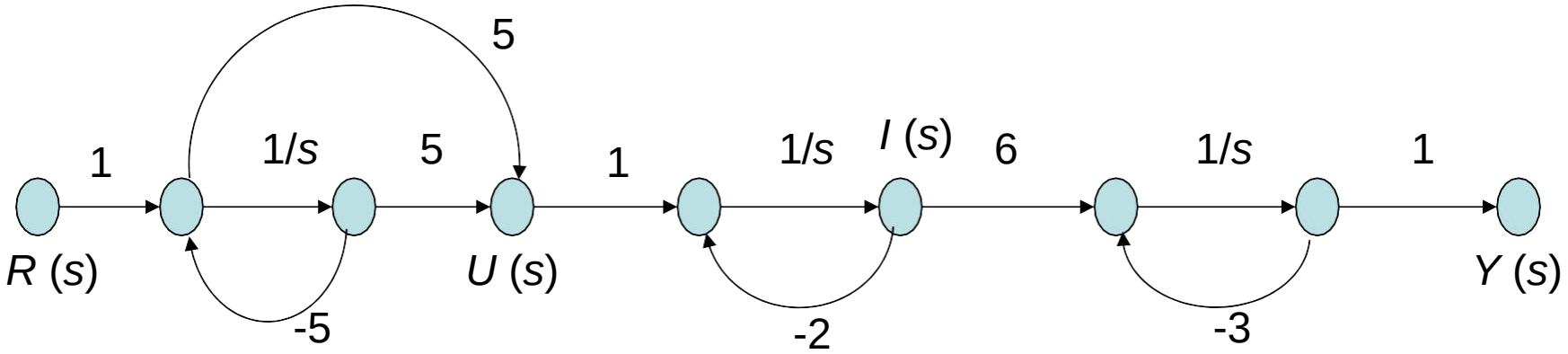
$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u(t)$$

Read Example 3.1 of the textbook

$$y(t) = \mathrm{C}x = \begin{bmatrix} b_0 & b_1 & b_2 & b_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

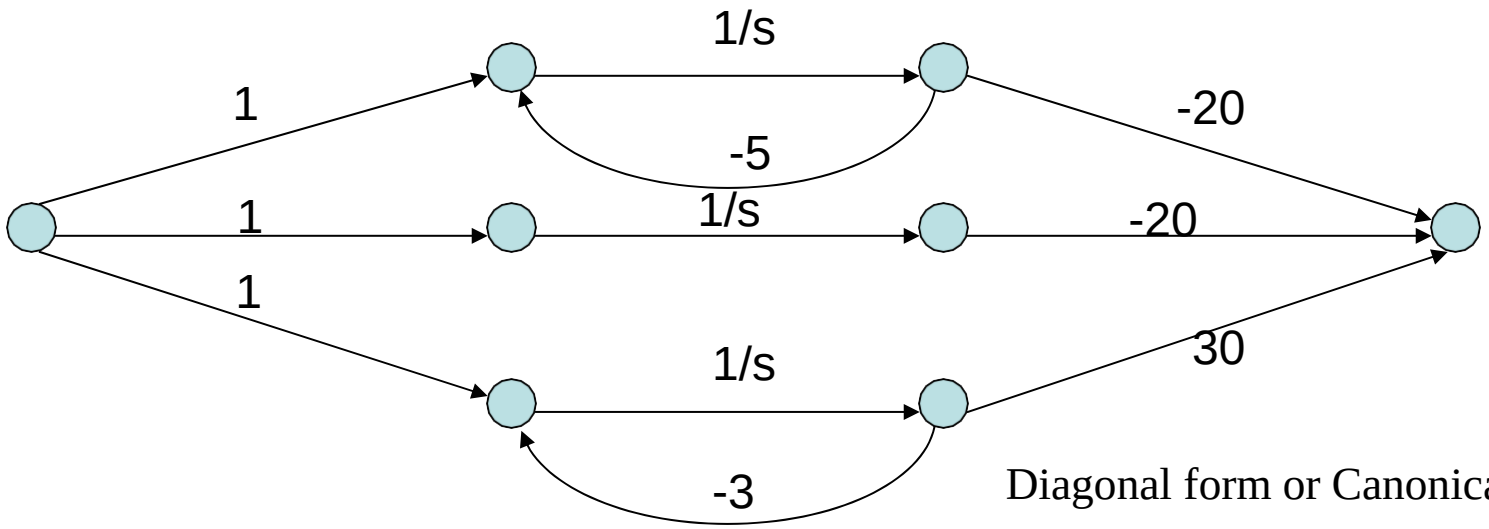
Alternative Signal-Flow Graph State Models





$$\dot{x} = \begin{bmatrix} -3 & 6 & 0 \\ 0 & -2 & -20 \\ 0 & 0 & -5 \end{bmatrix} x + \begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix} r(t) \qquad y = \begin{bmatrix} 10 & 0 \end{bmatrix} x$$

The State Variable Differential Equations



Diagonal form or Canonical form

$$\frac{Y(s)}{R(s)} = T(s) = \frac{30(s+1)}{(s+5)(s+2)(s+3)} = \frac{q(s)}{(s-s_1)(s-s_2)(s-s_3)}$$

$$\frac{Y(s)}{R(s)} = T(s) = \frac{k_1}{(s+5)} + \frac{k_2}{(s+2)} + \frac{k_3}{(s+3)}$$

$k_1 = -20$, $k_2 = -10$, and $k_3 = 30$

$$\dot{x} = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} r(t); \; y(t) = \begin{bmatrix} -20 & -10 & 30 \end{bmatrix} x$$

The State Variable Differential Equations

$$\dot{\mathbf{x}} = \begin{bmatrix} -3 & 6 & 0 \\ 0 & -2 & -5 \\ 0 & 0 & -5 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix} r(t)$$

$$\frac{Y(s)}{R(s)} = T(s) = \frac{30(s+1)}{(s+5)(s+2)(s+3)} = \frac{q(s)}{(s-s_1)(s-s_2)(s-s_3)}$$

$$\frac{Y(s)}{R(s)} = T(s) = \frac{k_1}{(s+5)} + \frac{k_2}{(s+2)} + \frac{k_3}{(s+3)}$$

$$k_1 = -20,\ k_2 = -10,\ \text{and}\ k_3 = 30$$

$$\dot{\mathbf{x}} = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} r(t)$$

$$y(t) = \begin{bmatrix} -20 & -10 & 30 \end{bmatrix} \mathbf{x}$$

The Transfer Function from the State Equation

Given the transfer function $G(s)$, we may obtain the state variable equations using the signal-flow graph model. Recall the two basic equations

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

$$y = \mathrm{C}\mathrm{x}$$

y is the single output and

u is the single input.

Take the Laplace transform

$$s\mathrm{X}(s) = \mathrm{A}\,\mathrm{X}(s) + \mathrm{B}U(s)$$

$$Y(s) = \mathrm{C}\mathrm{X}(s)$$

$$\left(s\mathrm{I} - \mathrm{A}\right)\mathrm{X}(s) = \mathrm{B}U(s)$$

$$\text{Since } \left[s\mathrm{I} - \mathrm{A}\right]^{-1} = \Phi(s)$$

$$\mathrm{X}(s) = \Phi(s)\,\mathrm{B}U(s)$$

$$Y(s) = \mathrm{C}\,\Phi(s)\,\mathrm{B}U(s)$$

$$G(s) = \frac{Y(s)}{U(s)} = \mathrm{C}\,\Phi(s)\,\mathrm{B}$$

Exercises: E3.2 (DGD)

A robot-arm drive system for one joint can be represented by the differential equation,

$$\frac{dv(t)}{dt} = -k_1 v(t) - k_2 y(t) + k_3 i(t)$$

where $v(t)$ = velocity, $y(t)$ = position, and $i(t)$ is the control-motor current. Put the equations in state variable form and set up the matrix form for $k_1=k_2=1$

$$v = \frac{dy}{dt}$$

$$\frac{dv}{dt} = -k_1 v(t) - k_2 y(t) + k_3 i(t)$$

$$\frac{d}{dt}\begin{pmatrix} y \\ v \end{pmatrix} = \begin{bmatrix} 0 & 1 \\ -k_2 & -k_1 \end{bmatrix}\begin{pmatrix} y \\ v \end{pmatrix} + \begin{bmatrix} 0 \\ k_3 \end{bmatrix} i$$

Define $u = i$, and let $k_1 = k_2 = 1$

.

$$\dot{x} = Ax + Bu;\ A = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ k_3 \end{bmatrix},\ x = \begin{pmatrix} y \\ v \end{pmatrix}$$

E3.3: A system can be represented by the state vector differential equation of equation (3.16) of the textbook. Find the **characteristic roots** of the system (DGD).

$$\dot{x} = Ax + Bu \qquad A = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$$

$$\text{Det}(\lambda I - A) = \text{Det}\left(\begin{bmatrix} \lambda & -1 \\ 1 & (\lambda+1) \end{bmatrix}\right)$$

$$= \lambda(\lambda+1)+1 = \lambda^2 + \lambda + 1 = 0$$

$$\lambda_1 = -\frac{1}{2} + j\frac{\sqrt{3}}{2}; \lambda_2 = -\frac{1}{2} - j\frac{\sqrt{3}}{2}$$

E3.7: Consider the spring and mass shown in Figure 3.3 where M = 1 kg, k = 100 N/m, and b = 20 N/m/sec. (a) Find the state vector differential equation. (b) Find the roots of the characteristic equation for this system (DGD).

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -100x_1 - 20x_2 + u$$

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -100 & -20 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$\text{Det}\,(\lambda \text{I} - \text{A}) = \text{Det} \begin{bmatrix} \lambda & -1 \\ 100 & \lambda + 20 \end{bmatrix} = \lambda^2 + 20\lambda + 100$$

$$= (\lambda + 10)^2 = 0; \lambda_1 = \lambda_2 = -10$$

E3.8: The manual, low-altitude hovering task above a moving land deck of a small ship is very demanding, in particular, in adverse weather and sea conditions. The hovering condition is represented by the **A** matrix (DGD)

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -5 & -2 \end{bmatrix}$$

$$\mathrm{Det}(\lambda I - A) = \mathrm{Det}\begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ 0 & 5 & \lambda = 2 \end{bmatrix} =$$

$$\lambda\left(\lambda^2 + 2\lambda + 5\right) = 0$$

$$\lambda_1 = 0; \lambda_2 = -1 + j2; \lambda_3 = -1 - j2$$

E3.9: See the textbook (DGD)

$$\dot{x}_1 = \frac{1}{2}x_2 - x_1$$

$$\dot{x}_2 = -x_1 - x_2$$

$$\dot{x} = \begin{bmatrix} -1 & 1/2 \\ 1 & -3/2 \end{bmatrix} x,\ y = \begin{bmatrix} 1 & -3/2 \end{bmatrix} x$$

$$s^2 + \frac{5}{2}s + 1 = (s+2)\left(s + \frac{1}{2}\right) = 0$$

$$\dot{z} = \begin{bmatrix} -2 & 0 \\ 0 & -1/2 \end{bmatrix} z;\ y = \begin{bmatrix} -0.35 & -1.79 \end{bmatrix} z$$

P3.1 (DGD-ELG4152):

Apply KVL

$$v(t) = Ri(t) + L\frac{di}{dt} + +v_c$$

$$v_c = \frac{1}{C}\int i dt$$

(a) Select the state variables as $x_1 = i$ and $x_2 = v_c$

(b) The state equations are :

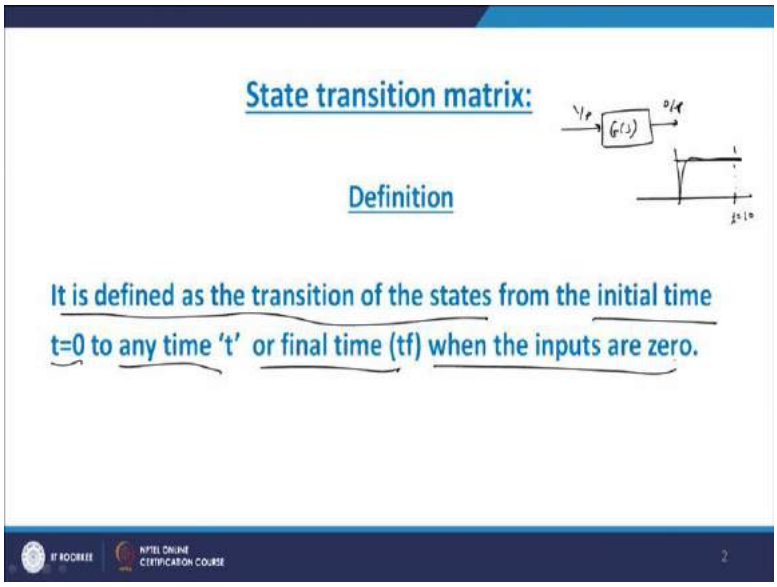
$$\dot{x}_1 = \frac{1}{L}v - \frac{R}{L}x_1 - \frac{1}{L}x_2$$

$$\dot{x}_2 = \frac{1}{C}x_1$$

(c) $$\dot{x} = \begin{bmatrix} -R/L & -1/L \\ 1/C & 0 \end{bmatrix} x + \begin{bmatrix} 1/L \\ 0 \end{bmatrix} u$$

State Transition Matrix

We start with a State Transition Matrix. So, what is meant by State Transition Matrix?

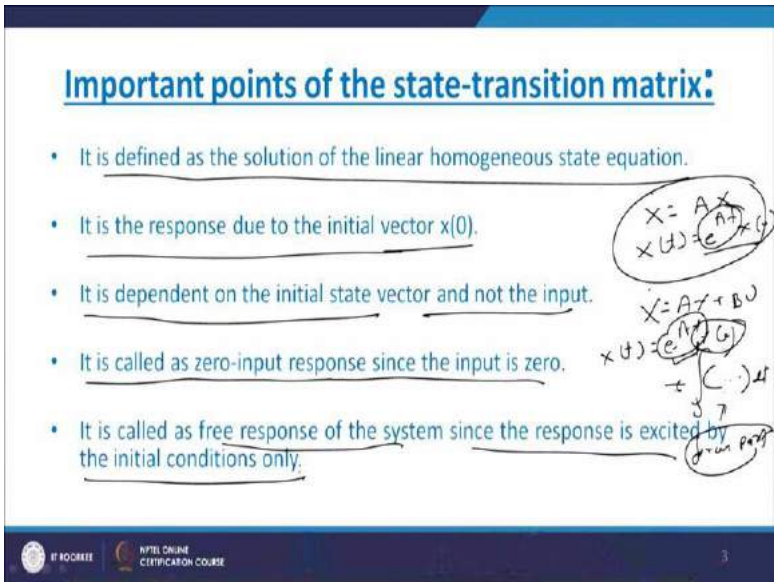


If you see a classical system the classical system here is a involvement of a plants say G of s we are we are applying input and we are getting a output. So, impede input is applied at t equals to 0, and we are expecting that the output as certain time interval say output say at say t equals to say 10 second. Now, this is the step input and now is a output. So, here output we have got at t equals to 10 second.

But, in case of a state space system along with this output there are states which are the important parameter in the system, that has been also traveled that is from t equals to 0 to t equals to f the state has been transferred. So, this transferring of the states we can we can say that state transition matrix.

Now, how will you define the state transition matrix? It is defined as the transition of states from initial time t naught to any time t or a final time tf when the inputs are zero,

that is this trans state transition matrix has not concerned with the any type of input it is concerned with the some initial state that the t 0 to tf. Now, you see what are the significant, means say what are the importance of these state transition matrix.



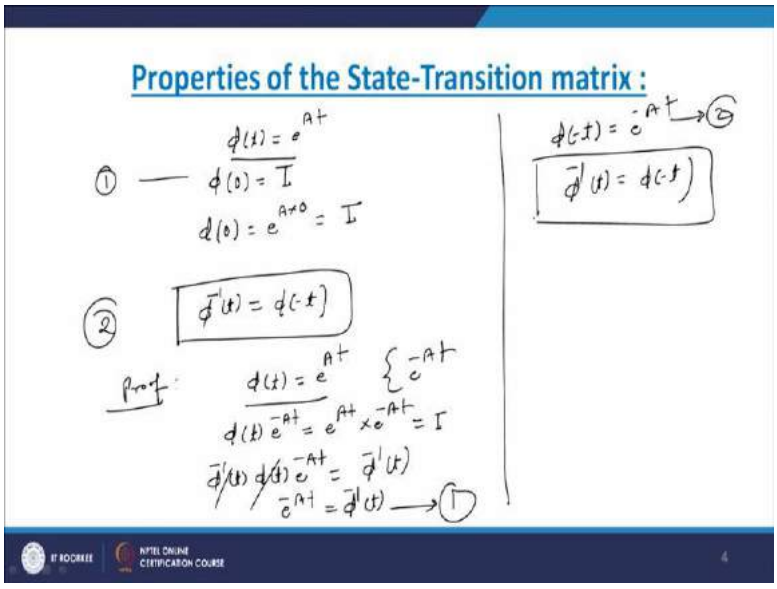
The first point it is defined as the solution of linear homogeneous state equations. We have seen last time, the X dot equal to A X X of t equal to e raised to A t X naught. So, this is e raised to A t is called state transition matrix and this is a homogeneous state equations.

It is response due to initial vector X naught it is response due to initial vector X naught here this is what, but if you consider a first system like this X dot equals to AX plus BU. So, here we will get X of t equal to e raised to A t X naught plus 0 to t some values will get here coefficient will get and we written d of t.

So, in this particular case this is u of t it has been taken care by X naught only whereas, if you take this part this portion that is a force portion force part in this particular part there is there is no involvement of X naught. So, therefore, what we are saying here it is dependent on the initial state vector not the input. The input is concern with the force part only, this part state transition matrix is concerned with only X naught. So, it is called as zero input response since input is 0, and whereas, if you take the force part this is called as zero state response. Here in this particular force part there is no initial state has been involved, but as far as now, we are discussing about the state transition matrix it is

concern with the zero input response because this portion e raised to A t into X naught this portion, there is no input.

It is called as a free response of the since the response is excited by the initial condition only, that is due to initial conditions the state has been transferred and this is called state transition matrix.



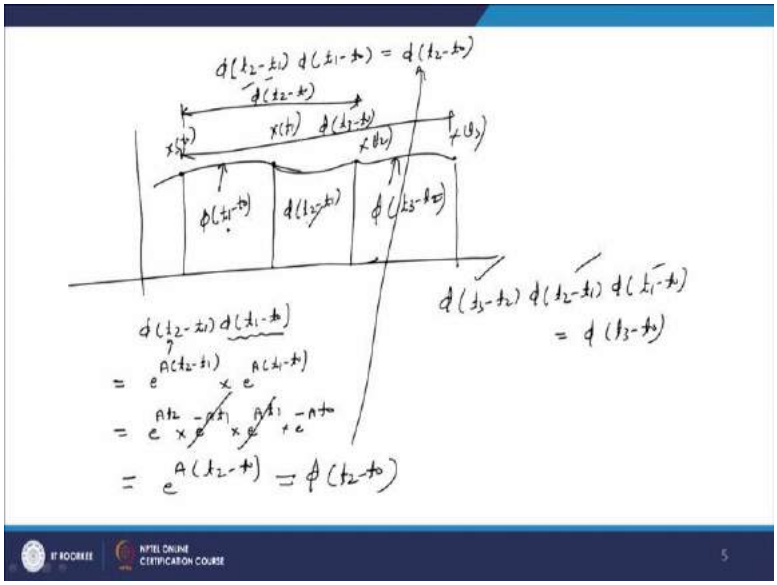
Now, we have to see a few properties of this state transition matrix. Say phi of t equal to e raised to A t that is a state transition matrix, and now we have to pre prove this phi of 0 equal to I this is a first part first property. So, how do you get I, identity matrix? So, in this particular equations, that is phi of 0 we replace we will get identity matrix.

Now, the second properties, we have to prove phi of minus t equal to phi of minus t that is phi inverse of minus t equal to phi of minus t. Now, how will you prove this one? So, now, we start with the same equation phi of t equal to e raised to A t. Now, through this equation we wow, post multiply by e raised to minus t, that is what we have to do to this equation we have to multi post multiply by e raised to minus A t.

So, after post multiplication what we will get? We will get phi of t to e raised to minus at equal to e raised to A t into e raised to minus at this is post multiplications, that is phi of t multiplied by e raised to minus at here e raised to A t multiplied by e raised to minus at we will get a anti matrix. So, now, our main purpose is to get phi inverse of t, equal to

phi of minus t. Now, what we will do? We will pre multiply both side by phi inverse of t. So, what we will get? So, if you multiply this phi inverse of t into phi of t to e of e raised to minus at, so phi phi inverse of t, that is I. So, this is cancelled; what we will get e raised to minus at equals to phi inverse of t.

So, e raised to minus at equals to phi inverse of t now, this if the phi of minus t equal to e raised to minus at just like phi of t equals to e raised to plus at, but if you say phi of minus t equal to e raised to minus A t. Now, if you see this equation say 1, you can say this equation 2, if you compare these equations, so what we will get? We will get phi of phi inverse t equals 2 phi of minus t, this is a second property.



Now, we see the another property, then another property is phi of t 2 minus t 1, phi of t 1 minus t 0 that is equal to phi of t 2 minus t 0, phi of t 2 minus t 1, phi of t 1 minus 0 equal to phi of t 2 minus t 0.

Now, this we can express through a diagram. Now, here on point this one, is a another point here. Now, here X of t 0 and now here is X of t 1. Now, the state has been transferred from X of t 0 to t 1. So, that has been represented by phi of t 1 minus t 0 then same state. Now, you see take X of t 2, t 2. Now, here also the state has been transferred form X of t 1 to X of t 2 so it has been represented by phi of t 2 minus t 1.

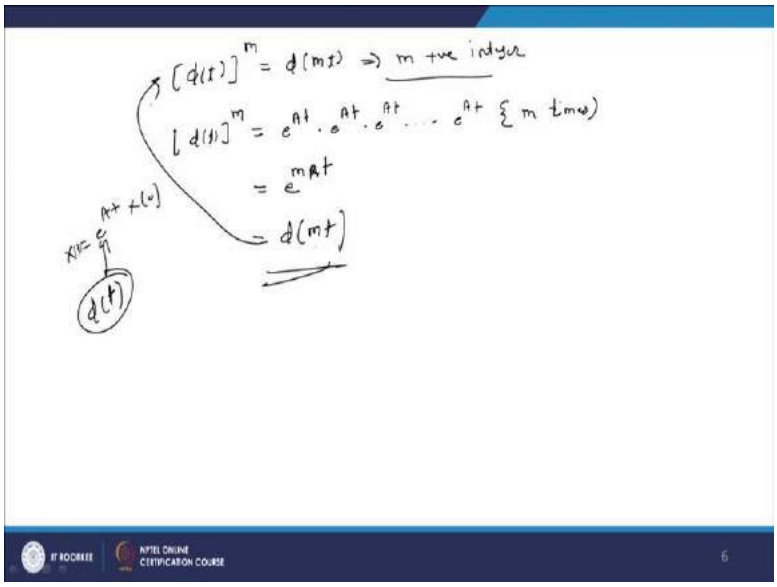
Now, you just see this phi of t 2 minus t 1 phi of t 1 minus t 0 equals to phi t 2 minus t 0 that is this particular portion, phi of t 2 minus t 0. Now, we have to prove this part mathematically. How will you prove this? So, here phi of t 2 minus t 1 phi of t 1 minus t 0, now, you have to solve this. So, this is also a transition matrix state transition matrix. So, it has been represented by e raised to a t 2 minus t 1 you see this portion this part.

And about this phi of t 1 minus t 0 it has been written by A of t 1 minus t 0. Now, we solve this one. So, we can write down as e of A t 2 to A of minus A into t 1 that is manipulations e of A into t 1 multiplied by minus A into t naught. Now, you see here e of minus A t 1, minus A t 1 this will be cancelled. So, what we will get? E of a t 2 minus t 0 and that is nothing, but phi of t 2 minus t 0. Now, we have prove this part.

So, your transition for state from t 0 to t 1, t 1 to t 2, but it may possible that the state at X equal to t 2 may transferred to X equals to say t 3 that means, X equal to t 3 that is it can be represented by a phi of t 3 minus t 2. Now, how do you express this through state transition matrices? That is phi of t 3 minus t 2, phi of t 2 minus t 1, phi of t 1 minus t naught, that is equal to phi of t 3 minus t naught that is you see here phi of t 3 minus t 2 this one, phi of t 2 minus t 1, then phi of t 1 minus t 0 and this is a complete part is the involvement of phi of t 3 minus t naught.

So, similarly more state elements we can add it and we can express these in terms of state transition matrix.

(Refer Slide Time: 12:48)

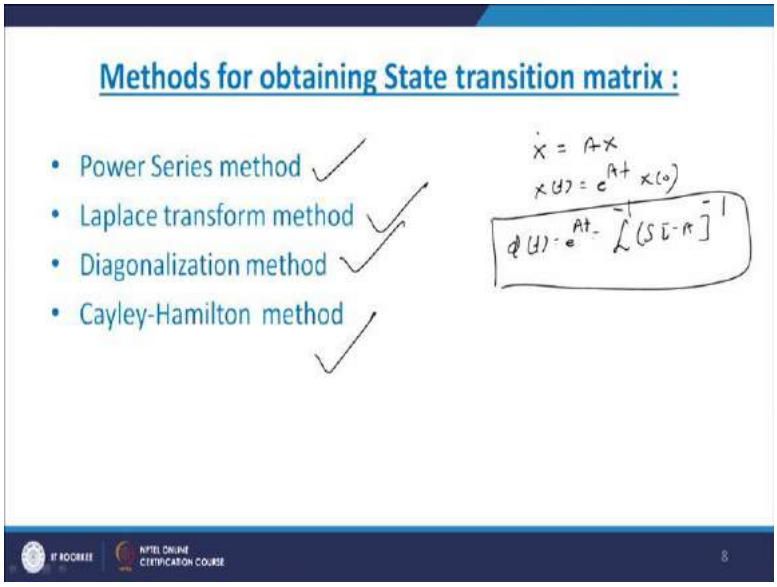


Now, the next properties say phi of t say raised to m equal to phi of m into t. So, in this case m is any positive integer. Now, this we have to prove. So, now, this phi of t raised to m. So, phi of t is nothing, but e raised to A t let us say m equals to 2 that means, e raised to A t multiplied by e raised to A t. Now, here m is there that means, e raised to A t has been written in terms of m times. Therefore, this equation we can write down as e raised to A t multiplied by e raised to A t e raised to A t up to e raised to A t.

And now up to how much terms we have to write? This we have to write m times, see here if phi of t raised to m e raised to A t e raised to A t m times. So, this we can write down as e how much time is a m into A into t, m into A into t and now we seen that this X t equals to e raised to t X naught. So, this is written as phi of t. So, here this can be written as phi of m into t here m term has been added therefore, this results has been proved.

So, these are the properties of the state transition matrix. Sometimes when we are doing the analysis of the system and many states are involved this type of properties can be useful. Now, the point has come this e raised to A t is there so that means, whenever required state transition matrix we have to calculate e raised to A t. Now, how to calculate it? So, if you see the literature there are various methods are available. So, these methods are useful to calculate the state transition matrix then we see the various methods for determination of state transition matrix.

(Refer Slide Time: 15:27)

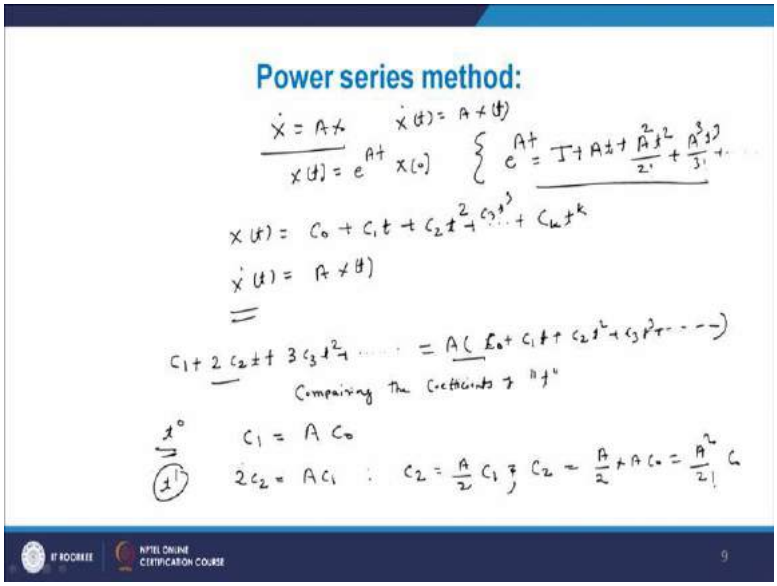


So, first one is Power Series method, Laplace transform method, Diagonalization method and Cayley-Hamilton method. So, these 4 methods can useful to get the state transition matrix. If you take a any example then apply this 4 method you will get the same results but now which method to be applied. So, this depends upon the type of systems. If you see the power series method we will see it, but you will find that as a order increases calculation is quite difficult and sometimes if we do not know whether we getting we will get result or not. This Laplace transform method is based on a s domain and Laplace is very important tool. Therefore, we will find that the Laplace transform method is quite useful.

We are already seen this approach earlier to get the state equation, that is X dot equals to AX X of t equal to e raised to A t into X naught and now these phi of t equal to e raised to A t equals to Laplace inverse of SI minus A inverse. By means of this we can calculate this state transition matrix using Laplace transform method that is you have take sI minus A inverse.

Now, the diagonalization method, this method is based on the concept of converting the given system matrix into the diagonal form and in this case there is need to go for the model matrix. So, if you see the calculation wise this method is quite cumbersome, but Laplace transfer is quite better.

Now, the next is Cayley-Hamilton theorem this is a very important theorem and it has wide applications. So, we will find that this method is also useful in getting the state transition matrix. Now, we will see through, but this is the method we already seen only, so method which we have to see power and diagonalization and Cayley-Hamilton theorem. Now, we start with power series method.



Now, we start with power series method. So, here we start with X dot equal to AX or this we can write down as X dot t equal to A into X of t. And here X of t state equation equal to e raised to A t into X naught and here this e raised to A t that is equal to I, A into t A square into t square factorial 2 plus A cube t cube factorial 3 and this processor. Now, we have to prove this. Why? The state transition matrix e raised to A t equals to this way.

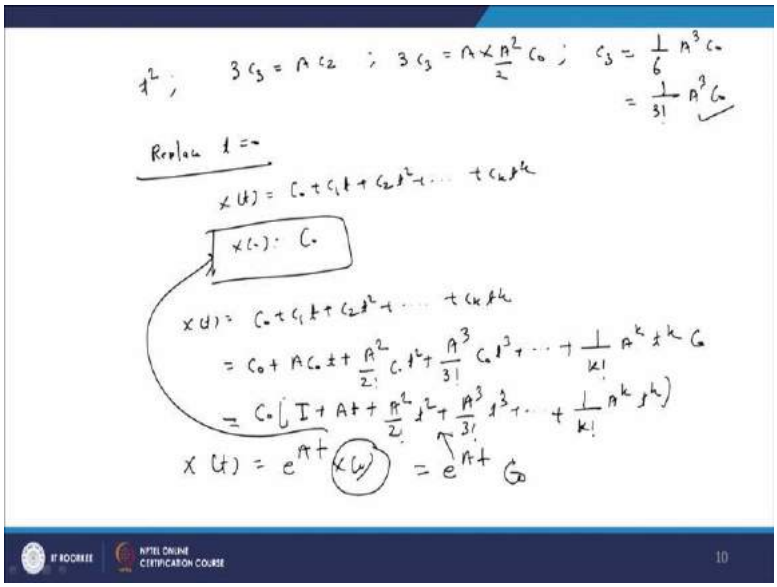
Now, we start doing it. So, first of all what we will do we will assume that say X of t equal to C 0 plus C 1 t plus C 2 t square plus say C k into t raised to k X of t. Now, we are concerned with the X dot t. So, here X dot t equal to A into X of t, now if you see X dot t so now, differentiate these equations. So, if you differentiate X dot t this term of C 0 will not present. So, terms of C 1 C 2 up to C k will be present therefore, we write the equation as this is C 1 plus this depression of these 2 into C 2, 2 say 2 into C 2 2 into t plus 3 C 3 into t square and this process repeated that is s C through C 3 is there.

So, here C 3 t cube means 3, 3 C 3 into t square equal to A into X of t. So, we can write down as A into X 0 C 0 C 1 t C 2 t square like this or we can write C 3 t cube. So, these terms we have got and now our main purpose is to get this result. So, what we can do? Means in what way we imposes. So, we can see that we can compare the various terms of t and some somewhere we can achieve the result.

So, please see it. So, now, what we will do comparing the coefficients of t. Now, you compare the coefficient of t. So, now, see here, for if see that the term t raised to 0. So,

you will find that t raised to 0 means 1. So, here for this it is C 1 and here is A into C naught. Similarly we compare that terms t raised to 1. So, here that is t. So, we can write down as see here t raised to 1, 2 C 2, so 2 C 2. Now, here we find A into C 1 into t's. Just see here A 1 C 1 into t, so we can write down as A into C 1. So, from this we can write down as C 2 equal to a by 2 into C 1 C 2 equals to a by 2 into C 1. But we know C 1; what is C 1? The C 1 is A into C naught. So, we replace the value of C 1 in this equation. So, we will get C 2 equal to A by 2 into A into C naught equal to A square by 2. So, we can write down as a factorial 2 into C naught.

So, why are your factorial, we will see later on. Now, here we have compare the coefficient of t 1. Now, similarly we have to compare the coefficient of t 2.



So, if you compare the coefficient of t 2 we will get 3 C 3 into A into C 2, so here 3 C 3 into A. Now, 3 C 3 into a what is C 2? C 2 we are already got A square factorial 2 into C naught. So, we replace these values here. So, you will can write A square by 2 into C naught.

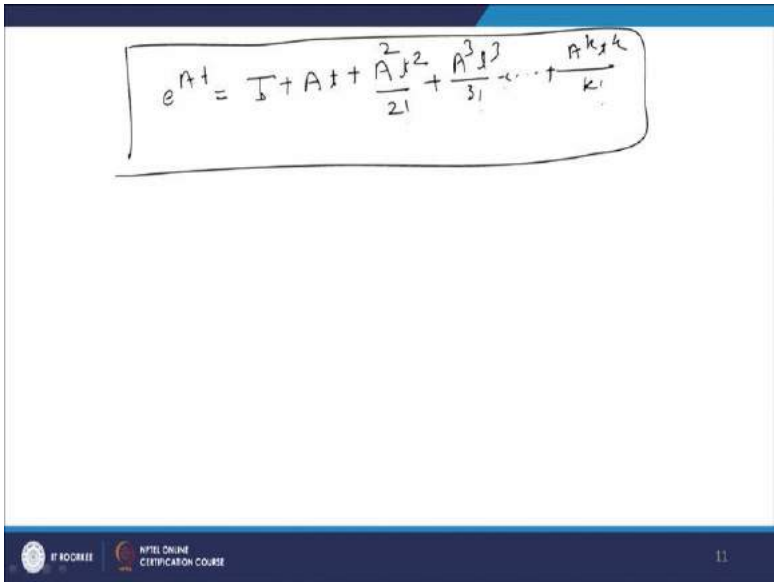
Now, we have got A square by 2 C naught, but you know if we know also C naught. So, C naught is, so you will find that that C naught is no value. So, we can keep like this only. So, we can write down as this C 3 equal to 3 we will take this side and we will get a A cube into C naught, so here we can write down this as 6 into A 3 into C naught and if you write this as this as a factorial 3 that is 3 into 2, 6 A cube into C naught.

Now, similarly we have to push it further. Suppose if you replace, if you replace t equals to 0 in equation of X of t that is X of t equals to C 0, C 1 t C 2 t square plus C k into t k. So, you will get X of 0 equal to C naught. So, now, we have got the value of X naught C 0, C 1, C 2, C 3 assume that we got up to C k. Now, we replace all this element in the given equations and now we will see what results we will get that is X of t equal to C 0, C 1 t plus C 2 t square up to C k into t k.

So, we have got all values of C 0, C 1, C 2, C into C k into t k. So, you can also find this process A equals C 0, C 1, C 2 into C 3 into t cube this portion and we have got all values and now we try to replace. So, if you replace it C 0, what is the value of C 1? C 1 you will find you get C 1 equals to A into C naught, C 2 equals to A square factorial into C naught and C 3 equal to 1 by factorial 3 A cube into C naught.

So, we replace these values and we will find C 0 into A into C 0 into t plus C 2 A square factorial 2, C 0 into t square plus A cube factorial 3 C 0 into t cube and this process repeated and we will find that this is 3 our terms are up to k. So, we have 1 by a factorial k into A raised to 3. So, k, so we a raised to k into t raised to k.

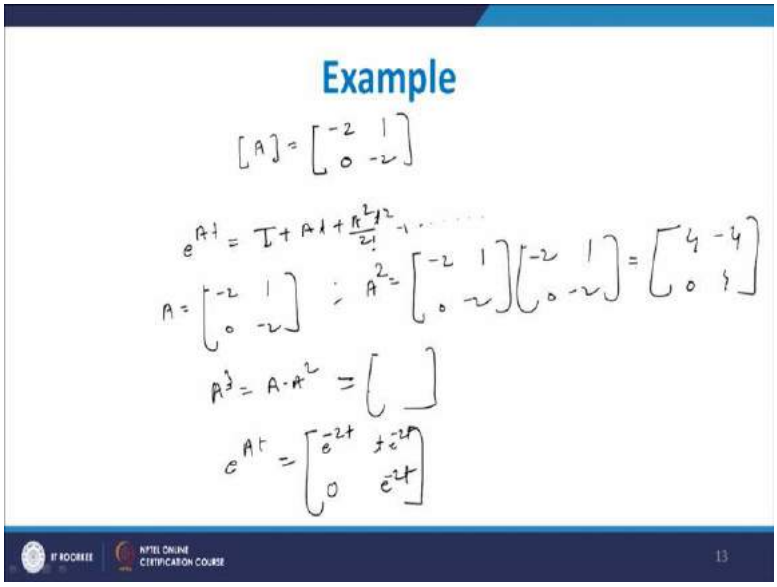
Now, we will find that in all these cases in all these cases the term C 0 is common. So, we will take C 0. So, we can write down I plus A into t A square factorial 2 into t square, A cube factorial 3 into t cube plus 1 by factorial k a raised to k into t k. So now, we have got this one. Now, what is the X of t? So, X of t is equal to e raised to A t into X naught. And what is a X naught? This X naught we are already calculated as C naught and therefore, this is e raised to A t into C naught and therefore, if you compare this equation this is the same equation.



Now, here e raised to A t is nothing but this part that means, we can write down e raised to A t equal to I plus A into t plus A square t square factorial 2 plus A cube t cube factorial 3 and like this A k t k factorial k. Now, this is the direct formula. So, by means of this formula we can easily get the state transition matrix.

So, in that case what is the input is available to you input is available to is A, A matrix. So, you replace the A matrix and solve it we can get the result.

(Refer Slide Time: 28:06)



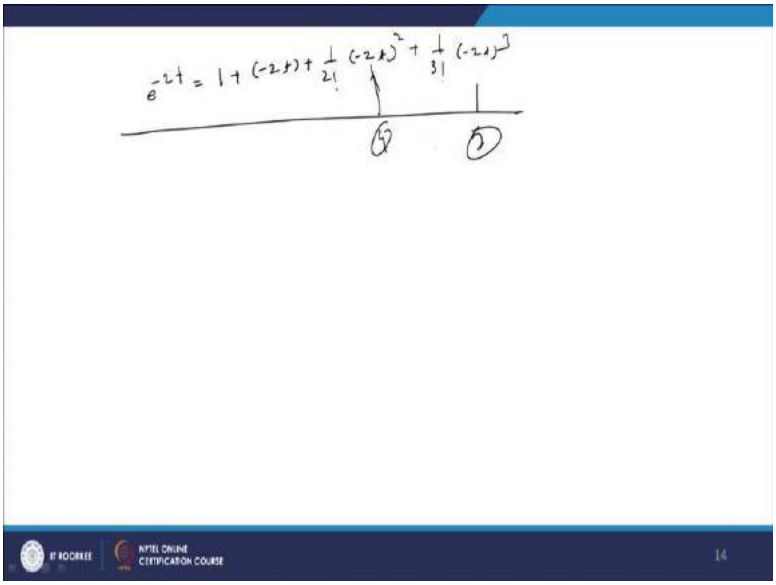
Now, we try to solve one example. Suppose if you are taking a matrix equal to minus 2, 1, 0, minus 2. And we want to determined we have to determine the state transition matrix e raised to A t by means of power series approach. So, here e raise to A t equal to same formula we write is I, A into t plus A square t square factorial 2 and like this.

Now, here you find that we will get A, A square, A cube it means that we have to calculate a many multiplications, multiple multiplication has to be solved and we will find out do this calc to do these calculations or to carry out this calculation sometimes it is quite cumbersomes.

Just we see here suppose if you take A is like this same thing minus 2, 1, 0, minus 2. Now, we have to determine A square, so A square is minus 2 1 0 minus 2 here minus 2 1 0 minus 2 and if you solve this one you will get 4 minus 4 0 into 4 and we have got A square, similarly you can get A cube A cube equals to A into A square.

So, we have to solve you get some values and finally, you have to replace in these particular equations. So, you will find that there are some terms. So, everywhere these terms are added and finally, if you get e raised to A t as e raised to minus 2 t t into e raised to minus 2 t 0 e raised to minus 2 t. That means what you do? You have to determine A square, A cube, A 4 like this they replace these values.

(Refer Slide Time: 30:24)



And from this same expression taking different elements we will get e raised to minus 2 t 2 t, that is e raised to minus 2 t portion is nothing but 1 plus minus 2 t plus 1 plus factorial 2 minus 2 t square plus factorial 3 minus 2 t whole cube. This is e raised to minus 2 t. Similarly we will get t e raised to minus 2 two t 0 e raised to minus 2 t. So, this is very simple calculations you can easily do it.

But here the main intention of this example is that sometimes it may not possible that you can get a similar type of results that is instead of minus 2 sometimes we get 4, it may get 7. In that case to make these elements or make adjust this element by means of power series exam power series expansion is quite difficult. Therefore, as we increase the order we will not recommend this method to get the state transition matrix because calculation is quite difficult. That means, in that is case it is better to go for the Laplace transform approach or there are other approaches that is called Cayley-Hamilton theorem or diagonalization that also helpful, but this approach sometimes is quite difficult.

State-Space Representations of Transfer Function Systems

1 State-Space Representation in Canonical Forms

We here consider a system de ned by

$$y^{(n)} + a_1 y^{(n-1)} + \quad + a_{n-1} y + a_n y = b_0 u^{(n)} + b_1 u^{(n-1)} + \quad + b_{n-1} u + b_n u \; ; \qquad (1)$$

where u is the control input and y is the output. We can write this equation as

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \quad + b_{n-1} s + b_n}{s^n + a_1 s^{n-1} + \quad + a_{n-1} s + a_n} : \qquad (2)$$

Later, we shall present state-space representation of the system de ned by (1) and (2) in controllable canonical form, observable canonical form, and diagonal canonical form.

1.1 Controllable Canonical Form

We consider the following state-space representation, being called a controllable canon-ical form, as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u \qquad (3)$$

$$y = (b_n - a_n b_0) \; (b_{n-1} - a_{n-1} b_0) \; \cdots \; (b_1 - a_1 b_0) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + b_0 u \qquad (4)$$

Note that the controllable canonical form is important in dicsussing the pole-placement approach to the control system design.

1.2 Observable Canonical Form

We consider the following state-space representation, being called an observable canon-ical form, as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} b_n - a_n b_0 \\ b_{n-1} - a_{n-1} b_0 \\ b_{n-2} - a_{n-2} b_0 \\ \vdots \\ b_1 - a_1 b_0 \end{bmatrix} u \tag{5}$$

$$y = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u \tag{6}$$

1.3 Diagonal Canonical Form

We here consider the transfer function system given by (2). We have the case where the dominator polynomial involves only distinct roots. For the distinct root case, we can write (2) in the form of

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \cdots + b_{n-1} s + b_n}{(s+p_1)(s+p_2)\cdots(s+p_n)} \tag{7}$$

$$= b_0 + \frac{c_1}{s+p_1} + \frac{c_2}{s+p_2} + \cdots + \frac{c_n}{s+p_n}: \tag{8}$$

The diagonal canonical form of the state-space representation of this system is given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} -p_1 & & & \\ & -p_2 & & \\ & & \ddots & \\ & & & -p_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} u \tag{9}$$

$$y = \begin{bmatrix} c_1 & c_2 & \cdots & c_{n-1} & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + b_0 u \tag{10}$$

2 Numerical Examples

Example 1: Obtain the transfer function of the system de ned by the following state-space equations:

$$\begin{bmatrix} \dot{x}_2 \\ \dot{x}_1 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -6 & -11 & -6 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u ; \tag{11}$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \\ x_3 \end{bmatrix} : \tag{12}$$

Solution: From (11) and (12), we determine the following parameters: $b_0 = 0$, $b_1 = 0$, $b_2 = 0$, $b_3 = 1$, $a_1 = 6$, $a_2 = 11$, $a_3 = 6$. Thus, the resulting transfer function is

$$G(s) = \frac{Y(s)}{U(s)} = \frac{1}{s^3 + 6s^2 + 11s + 6} :$$

Example 2: Find the state-space representation of the following transfer function system (13) in the diagonal canonical form.

$$G(s) = \frac{2s + 3}{s^2 + 5s + 6} : \tag{13}$$

Solution: Partial fraction expansion of (13) is

$$\frac{2s + 3}{s^2 + 5s + 6} = \frac{A}{s + 2} + \frac{B}{s + 3} :$$

Hence, we get A = 1 and B = 3. We now have two distinct poles. For this, we can write the transfer function (13) in the following form:

$$\begin{bmatrix} \dot{x}_2 \\ \dot{x}_1 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \tag{14}$$

$$y = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \tag{15}$$

Example 3: Obtain the state-space representation of the transfer function system (16) in the controllable canonical form.

$$G(s) = \frac{s^2 + 3s + 3}{s^2 + 2s + 1} \tag{16}$$

Solution: From the transfer function (16), we obtain the following parameters: $b_0 = 1$, $b_1 = 3$, $b_2 = 3$, $a_1 = 2$, and $a_2 = 1$. The resulting state-space model in controllable canonical form is obtained as

$$\begin{matrix} \dot{x}_1 \\ \dot{x}_2 \end{matrix} = \begin{matrix} 0 & 1 \\ 1 & 2 \end{matrix} \begin{matrix} x_1 \\ x_2 \end{matrix} + \begin{matrix} 0 \\ 1 \end{matrix} u ; \qquad (17)$$

$$y = \begin{matrix} 2 & 1 \end{matrix} \begin{matrix} x_1 \\ x_2 \end{matrix} + u : \qquad (18)$$

Example 4: Consider the following state equations:

$$\dot{x}_1 = x_2(t) ; \qquad (19)$$

$$\dot{x}_2 = x_3(t) ; \qquad (20)$$

$$\dot{x}_3 = -6x_1(t) - 11x_2(t) - 6x_3(t) + 6u(t) ; \qquad (21)$$

$$y = x_1(t) ; \qquad (22)$$

and determine the controllable canonical form.

Solution: Using the state equations (19), (20), (21), and (22), we write the following high order di erential equation:

$$\frac{d^3}{dt^3} y(t) + 6\frac{d^2}{dt^2} y(t) + 11\frac{d}{dt} y(t) + 6y(t) = 6u(t) :$$

The state variables $x(t) = y$, $x_2(t) = \dot{y}$, and $x_3 = \ddot{y}$. Hence, we get

$$\dot{x}_1 = x_2(t) ;$$

$$\dot{x}_2 = x_3(t) ;$$

$$\dot{x}_3 = \frac{d^3}{dt^3} y(t) = -6y(t) - 11\dot{y}(t) - 6\ddot{y}(t) + 6u(t)$$

$$= -6x_3(t) - 11x_2(t) - 6x_1(t) + 6u(t) :$$

In matrix form, we have

$$\dot{x}(t) = \begin{matrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -6 & -11 & -6 \end{matrix} x(t) + \begin{matrix} 0 \\ 0 \\ 6 \end{matrix} u(t) ;$$

$$y(t) = 1 \quad 0 \quad 0\ x(t) :$$

Example 5: Consider the following state equations

$$\dot{x}_1(t) = x_1(t) + x_3(t) + 4u(t) ; \tag{23}$$

$$\dot{x}_2(t) = 3x_1(t) + 2u(t) ; \tag{24}$$

$$\dot{x}_3(t) = 5x_1(t) + x_2(t) + u(t) ; \tag{25}$$

$$y(t) = x_1(t) ; \tag{26}$$

and determine the observable canonical form.

Solution: Using the state equations (23), (24), (25), and (26), we write the following high order di erential equation:

$$\frac{d^3}{dt^3}y(t) + \frac{d^2}{dt^2}y(t) + 5\frac{d}{dt}y(t) + 3y(t) = 4\frac{d^2}{dt^2}u(t) + \frac{d}{dt}u(t) + 2u(t) :$$

We introduce $x_1(t) = y(t)$ in the equation, and collect all terms without di erentiation on the right hand side, we get

$$\frac{d^3}{dt^3}x_1(t) + \frac{d^2}{dt^2}x_1(t) + \frac{d}{dt}x_1(t) \quad \frac{d^2}{dt^2}u(t) \quad \frac{d}{dt}u(t) = 3x_1(t) + 2u(t) ;$$

i.e.,

$$\frac{d}{dt}\left[\frac{d^2}{dt^2}x_1(t) + \frac{d}{dt}x_1(t) + x_1(t) \quad \frac{d}{dt}u(t) \quad u(t)\right] = 3x_1(t) + 2u(t) :$$

Now introduce the expression within the paranthesis as a new state variable

$$x_2(t) = \frac{d^2}{dt^2}x_1(t) + \frac{d}{dt}x_1(t) + 5x_1(t) \quad 4u(t) \quad u(t) ;$$

i.e.,

$$\dot{x}_2(t) = \quad 3x_1(t) + 2u(t) : \tag{27}$$

Repeating this precedure yields

$$\frac{d}{dt}\left[\frac{d}{dt}x_1(t) + x_1(t) \quad 4u(t)\right] = x_2(t) \quad 5x_1(t) + u(t) ; \tag{28}$$

and we introduce

$$x_3(t) = \frac{d}{dt}x_1(t) + x_1(t) \quad 4u(t) ;$$

i.e.,

$$\dot{x}_1(t) = x_3(t) \quad x_1(t) + 4u(t) : \tag{29}$$

From (27), (28), and (29), we de ne the state-space form of

$$\dot{x}(t) = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 2 \\ 4 \end{bmatrix} u(t) ;$$

$$y(t) = \begin{bmatrix} 4 & 5 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 5 \end{bmatrix} u(t) :$$

$$y(t) = \begin{matrix} 1 & 0 & 0 \end{matrix} x(t) :$$

3 References

1. Katsuhiko Ogata, Modern Control Engineering, 4th Ed., Prentice Hall Inc., New Jersey, 2002.
2. KTH Reglerteknik, Reglerteknik Allman Kurs, Del 2, 2007.

Module 8: Controllability, Observability and Stability of Discrete Time Systems
Lecture Note 1

Controllability and observability are two important properties of state models which are to be studied prior to designing a controller.

Controllability deals with the possibility of forcing the system to a particular state by appli-cation of a control input. If a state is uncontrollable then no input will be able to control that state. On the other hand whether or not the initial states can be observed from the output is determined using observability property. Thus if a state is not observable then the controller will not be able to determine its behavior from the system output and hence not be able to use that state to stabilize the system.

1 Controllability

Before going to any details, we would first formally define controllability. Consider a dynamical system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{1}$$

where $A \in R^{N \times N}$, $B \in R^{N \times M}$, $C \in R^{P \times N}$, $D \in R^{P \times M}$.

Definition 1. Complete State Controllability: The state equation (1) (or the pair (A, B)) is said to be completely state controllable or simply state controllable if for any initial state $x(0)$ and any final state $x(N)$, there exists an input sequence $u(k)$, $k = 0, 1, 2, , N$, which transfers $x(0)$ to $x(N)$ for some finite N . Otherwise the state equation (1) is state uncontrollable.

Definition 2. Complete Output Controllability: The system given in equation (1) is said to be completely output controllable or simply output controllable if any final output can be reached from any initial state $x(0)$ by applying an unconstrained input sequence $y(N)$ $k = 0, 1, 2, , u(k)$, N , for some finite N . Otherwise (1) is not output controllable.

1.1 Theorems on controllability

State Controllability:

1. The state equation (1) or the pair (A, B) is state controllable if and only if the n × nm state controllability matrix

$$U_C = \begin{bmatrix} B & AB & A^2B & ..... & A^{N-1}B \end{bmatrix}$$

has rank n, i.e., full row rank.

2. The state equation (1) is controllable if the n × n controllability grammian matrix

$$W_C = \sum_{I=0}^{N-1} A^I B B^T (A^I)^T = \sum_{I=0}^{N-1} A^{N-1-I} B B^T (A^{N-1-I})^T$$

is non-singular for any nonzero finite N .

3. If the system has a single input and the state model is in controllable canonical form then the system is controllable.

4. When A has distinct eigenvalues and in Jordan/Diagonal canonical form, the state model is controllable if and only if all the rows of B are nonzero.

5. When A has multiple order eigenvalues and in Jordan canonical form, then the state model is controllable if and only if

   i. each Jordan block corresponds to one distinct eigenvalue and

   ii. the elements of B that correspond to last row of each Jordan block are not all zero.

Output Controllability: The system in equation (1) is completely output controllable if and only if the p × (n + 1)m output controllability matrix

$$U_{OC} = \begin{bmatrix} D & CB & CAB & CA^2B & ..... & CA^{N-1}B \end{bmatrix}$$

has rank p, i.e., full row rank.

1.2 Controllability to the origin and Reachability

There exist three different definitions of state controllability in the literature:

1. Input transfers any state to any state. This definition is adopted in this course.

2. Input transfers any state to zero state. This is called controllability to the origin.

3. Input transfers zero state to any state. This is referred as controllability from the origin or reachability.

Above three definitions are equivalent for continuous time system. For discrete time systems definitions (1) and (3) are equivalent but not the second one.

Example: Consider the system $x(k+1) = Ax(k) + Bu(k)$, $y(k) = Cx(k)$. where

$$A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } C = [0\ 1]$$

Show if the system is controllable. Find the transfer function $\frac{Y}{U}(z)$. Can you see any connection between controllability and the transfer function?

SOLUTION: The controllability matrix is given by

$$U_C = [B\ AB] = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$$

Its determinant $|U_C| = 0 \Rightarrow U_C$ has a rank 1 which is less than the order of the matrix, i.e., 2. Thus the system is not controllable. The transfer function

$$G(z) = \frac{Y(z)}{U(z)} = C(zI - A)^{-1} B = [0\,1] \begin{bmatrix} z+2 & -1 \\ -1 & z+2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{z+1}$$

Although state model is of order 2, the transfer function has order 1. The eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = -3$. This implies that the transfer function is associated with pole-zero cancellation for the pole at -3. Since one of the dynamic modes is cancelled, the system became uncontrollable.

2 Observability

Definition 3. The state model (1) (or the pair (A, C)) is said to be observable if any initial state x(0) can be uniquely determined from the knowldge of output y(k) and input sequence u(k), for k = 0, 1, 2, , N , where N is some finite time. Otherwise the state model (1) is unobservable.

2.1 Theorems on observability

1. The state model (1) or the pair (A, C) is observable if the np × n observability matrix

$$U = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{bmatrix}$$

has rank n, i.e., full column rank.

2. The state model (1) is observable if the n × n observability grammian matrix

$$W_O = \sum_{I=0}^{N-1} (A^I)^T C^T CA^I = \sum_{I=0}^{N-1} (A^{N-1-I})^T C^T CA^{N-1-I}$$

is non-singular for any nonzero finite N .

3. If the state model is in observable canonical form then the system is observable.

4. When A has distinct eigenvalues and in Jordan/Diagonal canonical form, the state model is observable if and only if none of the columns of C contain zeros.

5. When A has multiple order eigenvalues and in Jordan canonical form, then the state model is observable if and only if

   i. each Jordan block corresponds to one distinct eigenvalue and

   ii. the elements of C that correspond to first column of each Jordan block are not all zero.

2.2 Theorem of Duality

The pair (A, B) is controllable if and only if the pair (A^T , B^T) is observable.

Exercise: Prove the theorem of duality.

3 Loss of controllability or observability due to pole-zero cancellation

We have already seen through an example that a system becomes uncontrollable when one of the modes is cancelled. Let us take another example.

Example:

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k) \\ y(k) &= [1 \;\; 1]x(k) \end{aligned}$$

The controllability matrix

$$U_C = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix}$$

implies that the state model is controllable. On the other hand, the observability matrix

$$U_O = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

has a rank 1 which implies that the state model is unobservable. Now, if we take a different set of state variables so that, $x_1(k) = y(k)$, then the state variable model will be:

$$\begin{aligned} x_1(k+1) &= y(k+1) \\ x_1(k+2) &= y(k+2) = -y(k) - 2y(k+1) + u(k+1) + u(k) \end{aligned}$$

Lets us take $x_2(k) = y(k+1) - u(k)$. The new state variable model is:

$$\begin{aligned} x_1(k+1) &= x_2(k) + u(k) \\ x_2(k+1) &= -x_1(k) - 2x_2(k) - u(k) \end{aligned}$$

which implies

$$A = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad C = [1\ 0]$$

The controllability matrix

$$U_C = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

implies that the state model is uncontrollable. The observability matrix

$$U_O = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

implies that the state model is observable. The system difference equation will result in a transfer function which would involve pole-zero cancellation. Whenever there is a pole zero cancellation, the state space model will be either uncontrollable or unobservable or both.

4 Controllability/Observability after sampling

Question: If a continuous time system is undergone a sampling process will its controllability or observability property be maintained?

The answer to the question depends on the sampling period T and the location of the eigenvalues of A.

- Loss of controllability and/or observability occurs only in presence of oscillatory modes of the system.
- A sufficient condition for the discrete model with sampling period T to be controllable is that whenever $Re[\lambda_I - \lambda_J] = 0$, $|Im[\lambda_I - \lambda_J]| = 2\pi m/T$ for $m = 1, 2, 3, ...$
- The above is also a necessary condition for a single input case.

Note: If a continuous time system is not controllable or observable, then its discrete time version, with any sampling period, is not controllable or observable.

UNIT 2

Phase plane analysis for linear systems

In this lecture we are going to talk about systems of two linear ODE of the first order in the form

$$\begin{aligned} x' &= a_{11}x + a_{12}y, \\ y' &= a_{22}x + a_{22}y. \end{aligned} \tag{1}$$

Here I chose the notations $(x(t), y(t))^T$ for the unknown functions, and the independent variable is, as before, denoted by t, meaning "time." Note that *planar* system (1) is a particular case of the general nonlinear system

$$\begin{aligned} x' &= f(x, y), \\ y' &= g(x, y), \end{aligned} \tag{2}$$

where $f, g \in C^1(U ; R)$, $U \subseteq R^2$, are given functions of two variables. The common point between (1) and (2) is that their right-hand sides do not depend explicitly on t. They, of course, depend on t through the variables x(t) and y(t), but t itself is absent. Such systems are called *autonomous* (cf. autonomous first order ordinary differential equations).

Assume that system (2) (or (1)) has a solution $x = x(t; x_0)$, $y = y(t; y_0)$, where $(x_0, y_0)^T$ are the initial conditions. We can consider this solution as a parametrically defined curve: for each time moment t we have two numbers $(x, y) \in R^2$, which can be represented as a point on the plane xy. If we change t, the point position will change, but since $x(t; x_0)$ and $y(t; x_0)$ are differentiable, then the change will be small, and we actually obtain a smooth curve. Moreover, by increasing or decreasing t we move on xy plane along this curve. Such curve with the direction of time increase on it is called an *orbit*, or a *trajectory* of system (2) (or system (1)). Our task here is to analyze the structure of orbits of system (1) on the plane xy, which is called the *phase plane*. Since some of the properties of the orbits of (1) hold in the general case (2), I will start with the more general system.

- If $(\hat{x}, \hat{y})$ are such that $f(\hat{x}, \hat{y}) = 0$ and $g(\hat{x}, \hat{y}) = 0$, then $x = \hat{x}, y = \hat{y}$ is a solution to (2), and the corresponding orbit is simply a point on the phase plane with coordinates $(\hat{x}, \hat{y})$. For the linear system (1) point $(\hat{x}, \hat{y})$ has to be a solution to

$$\begin{aligned} 0 &= a_{11}x + a_{12}y, \\ 0 &= a_{22}x + a_{22}y, \end{aligned}$$

  i.e., a solution to a homogeneous system of two linear algebraic equations with the matrix

$$A = \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}.$$

  This system always has solution $(\hat{x}, \hat{y}) = (0, 0)$. This solution is unique if and only if $\det A = 0$. If $\det A = 0$ then we have infinitely many solutions. In order not to complicate the following discussion I will assume that $\det A = 0$.

  Such points $(\hat{x}, \hat{y})$ are called *equilibrium points*, or *rest points*, or *stationary points*, or *critical points* of system (2). Hence the assumption for (1) that $\det A = 0$ is equivalent to saying that system (1) has only one equilibrium at the origin.

- If $x = x(t)$, $y = y(t)$ is a solution to (2), then $\tilde{x} = x(t + c)$, $\tilde{y} = y(t + c)$ is a also a solution to (2) for any constant c.

Proof. Note that if x(t), y(t) is a solution, then for the first equation in (2) it means that

$$\frac{dx}{dt}(t) = f\big(x(t), y(t)\big),$$

and since this is true for any t, it is true for t + c:

$$\frac{dx}{dt}(t + c) = f\big(x(t + c), y(t + c)\big),$$

which, due to the chain rule, can be rewritten as

$$\frac{dx}{d(t + c)}(t + c) = f\big(x(t + c), y(t + c)\big),$$

or, using new variable $\tau = t + c$,

$$\frac{dx}{d\tau}(\tau) = f\big(x(\tau), y(\tau)\big).$$

But since $x(\tau) = \tilde{x}(t)$, $y(\tau) = \tilde{y}(t)$, this exactly means that $\tilde{x}(t) = x(t + c)$, $\tilde{y}(t) = y(t + c)$ is a solution to (2).

This simple and very important fact means that if x(t), y(t) is the solution to (2) with the initial condition $x(t_0) = x_0$, $y(t_0) = y_0$, then $x(t + t_0)$, $y(t + t_0)$ is the solution to system (2) with the initial condition $x(0) = x_0$, $y(0) = y_0$. From the geometric point of view it means that we can use different parametrizations to define the same curve on the phase plane.

For the linear system we can prove this fact explicitly. Recall that any solution to (1) is given by $e^{At}v$ for some vector $v \in R^2$. Now consider $e^{A(t+c)}v = e^{At}u$, where $u = e^{Ac}v$, which is clearly a solution to the linear system.

- *Orbits do not intersect*. Suppose contrary: there are two orbits x(t), y(t) and $\tilde{x}(t)$, $\tilde{y}(t)$ that pass through the same point (x_0, y_0) for different time moments t_1 and t_2: i.e.,

$$(x_0, y_0) = x(t_1), y(t_1) = \tilde{x}(t_2), \tilde{y}(t_2).$$

Since $\tilde{x}(t), \tilde{y}(t)$ is a solution, then, according to the previous property, $\tilde{x}(t + (t_2 - t_1)), \tilde{y}(t + (t_2 - t_1))$ is also a solution, which corresponds to the same orbit, but with a different time parametrization. On the other hand, the value of this solution at the point t_1 coincides with the value of $x(t), y(t)$ at the same point, which according to the uniqueness and existence theorem means that $\tilde{x}(t+(t_2-t_1)), \tilde{y}(t+(t_2-t_1))$ and $x(t), y(t)$ coincide, which yields that the existence of a common point of these orbits coincide, hence no intersections.

This property, as well as the previous one, is not true for non-autonomous systems.

We obtained that the phase plane consists of orbits, which cannot intersect. It is impossible to depict all the orbits, but it is usually enough to draw only a few to get a general idea of the behavior of the solutions of system (2). In particular, it is always advisable first to plot equilibria. Several key orbits on the phase plane representing the general picture are called the *phase portrait*. It is usually quite difficult to draw the phase portrait of the general nonlinear system (2). For system (1), especially assuming that $\det A = 0$, this problem can be solved completely, as I will show next.

Phase portraits of linear system (1)

There are only a few types of the phase portraits possible for system (1). Let me start with a very simple one:

$$\dot{x} = \lambda_1 x,$$
$$\dot{y} = \lambda_2 y.$$

This means that the matrix of the system has the diagonal form

$$A = \begin{matrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{matrix},$$

i.e., it has real eigenvalues λ_1, λ_2 with the eigenvectors $(1, 0)^T$ and $(0, 1)^T$ respectively. The equations are decoupled and the general solution to this system is given by

$$\begin{matrix} y(t) \\ x(t) \end{matrix} = C_1 \begin{matrix} 0 \\ 1 \end{matrix} e^{\lambda_1 t} + C_2 \begin{matrix} 1 \\ 0 \end{matrix} e^{\lambda_2 t}.$$

Note that this is a fancy way to write that

$$x(t) = C_1 e^{\lambda_1 t}, \quad y(t) = C_2 e^{\lambda_2 t}.$$

Now to figure out the behavior of the orbits (i.e., their qualitative form and the directions on them), we can argue as follows: Consider initially only special orbits corresponding to $C_1 = 0, \pm 1$ and $C_2 = 0, \pm 1$. First of all there is always the equilibrium (0, 0), which means that the corresponding orbit is a point at the origin. Next, take $C_1 = +1$ and $C_2 = 0$, which implies that $x(t) = e^{\lambda_1 t}$ and $y(t) = 0$. To get further information we need to specify the signs of λ_1 and λ_2.

Case $\lambda_1 > 0 > \lambda_2$. I repeat that I consider the case $C_1 = 1$ and $C_2 = 0$. This corresponds to the line, whose direction is given by the first eigenvector $v_1 = (1, 0)^T$. If $t \to \infty$ then $x(t) \to \infty$ for $\lambda_1 > 0$, hence the orbit constitutes the half line ($x > 0$, $y = 0$) with the direction from the origin to infinity (since also $x(t) \to 0$ if $t \to -\infty$). Similarly, taking $C_1 = -1$, $C_2 = 0$ we will find that $x(t) \to -\infty$ if $t \to \infty$ and $x(t) \to 0$ if $t \to -\infty$ on the half line ($x < 0$, $y = 0$). Hence we fully described the orbit structure on the line corresponding to the direction of v_1 (see the figure below): There are three orbits there, two half-lines separated by the equilibrium at the origin, and on both half lines the direction is *from* the origin (the corresponding eigenvalue is positive). Now take $C_1 = 0$, $C_2 = \pm 1$. In this case we find ourselves on the direction corresponding to $v_2 = (0, 1)^T$, i.e., on y-axis. We again have three orbits there, but the direction is reversed, we approach the origin along these orbits because λ_2 is negative (see the figure below).

What about the case $C_1 = 0$ and $C_2 = 0$? For both $t \to \infty$ and $t \to -\infty$ one of the coordinate will approach infinity. Moreover, our intuition tells us that close orbits should behave similarly, therefore we do not have much choice as to obtain the orbit structure shown in the figure above. An equilibrium point for which we have two real eigenvalues, one is negative and one is positive, is called *saddle*.

I actually was quite vague about why the orbits not on the axes have this particular shape. Here is a proof. We have, again, that

$$x(t) = C_1 e^{\lambda_1 t}, \quad y(t) = C_2 e^{\lambda_2 t},$$

or

$$\frac{x}{C_1} = e^{\lambda_1 t}, \qquad \frac{y}{C_2} = e^{\lambda_2 t}.$$

Raise the first equality to the power λ_2 and the second equality to the power λ_1. We find, by eliminating t, that

$$\frac{x^{\lambda_2}}{C_1^{\lambda_2}} = \frac{y^{\lambda_1}}{C_2^{\lambda_1}} \Rightarrow y = Ax^{\frac{\lambda_2}{\lambda_1}},$$

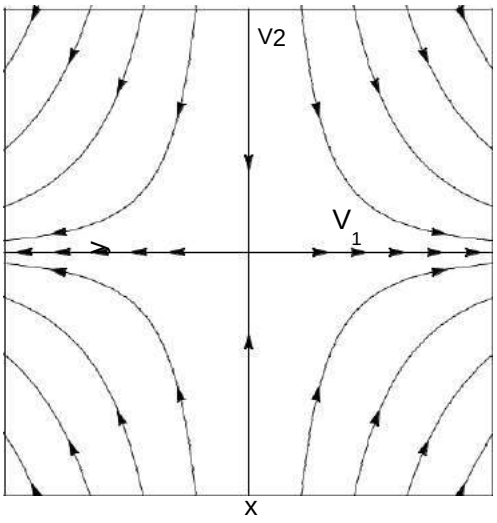


Figure 1: Saddle. The case $\lambda_1 > 0 > \lambda_2$. The first eigenvector corresponds to x-axis, and the second one corresponds to y-axis

where A is a new constant depending on C_1 and C_2. Since the eigenvalues have opposite signs, we find that orbits corresponds to "hyperbolas"

$$y = Ax^{-\gamma}, \qquad \gamma > 0,$$

which we can see in the figure. By eliminating t we lost the information on the direction along the orbits. But since we already know the directions along the axes, we can restore it for our curves by continuity.

Q: Can you plot a phase portrait of the system with the diagonal matrix $A = \mathrm{diag}(\lambda_1, \lambda_2)$ such that $\lambda_1 < 0 < \lambda_2$? What will be the difference with respect to the figure above?

Case $\lambda_2 > \lambda_1 > 0$. Formally, we have exactly the same general solution

$$\begin{matrix} y(t) \\ x(t) \end{matrix} = C_1 \begin{matrix} 0 \\ 1 \end{matrix} e^{\lambda_1 t} + C_2 \begin{matrix} 1 \\ 0 \end{matrix} e^{\lambda_2 t},$$

but note that opposite to the previous case here for any choice of C_1 and C_2 $(x, y) \to (0, 0)$ if $t \to -\infty$, hence geometrically all the orbits represent curves coming out from the origin and approaching infinity as $t \to \infty$. This is true in particular for the *characteristic directions* corresponding to the eigenvectors v_1 and v_2 (see the figure below, left panel).

The subtle question is however how exactly orbits approach the origin for $t \to -\infty$. For this we recall that the equation for the curves on the plane is given by

$$y = Ax^{\frac{\lambda_2}{\lambda_1}},$$

and since we assumed that $\lambda_2 > \lambda_1$ then we have

$$y = Ax^{\gamma}, \quad \gamma > 1,$$

which corresponds to "parabolas," i.e., to the curves that are tangent to the x-axis at the origin. Another way to see why x-axis is more important in this case is to note that when $t \to -\infty$, then $e^{\lambda_2 t}$ is much much smaller than $e^{\lambda_1 t}$, hence it is the first eigenvector that plays the most important role. For $t \to \infty$ the situation is opposite, since $e^{\lambda_2 t}$ is much much bigger than $e^{\lambda_1 t}$ and the second eigenvector shows the slope of the orbits.

An equilibrium point for which we have two real eigenvalues of the same sign is called *node*. To obtain more intuition about how exactly the orbits approach the origin consider

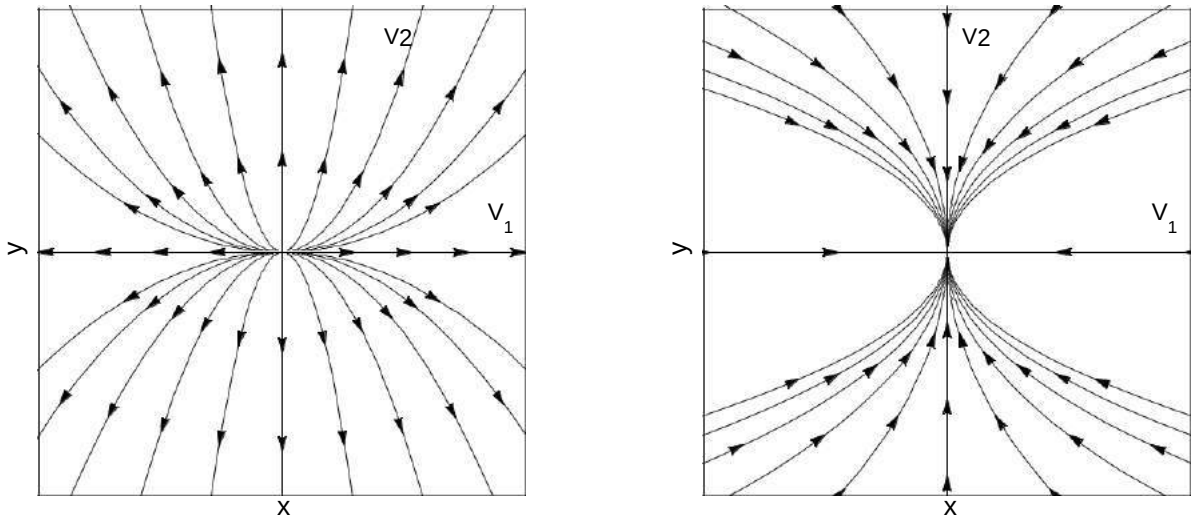


Figure 2: Left: Node. The case $\lambda_2 > \lambda_1 > 0$. The first eigenvector corresponds to x-axis, and the second one corresponds to y-axis. Right: Node. The case $\lambda_1 < \lambda_2 < 0$. The first eigenvector corresponds to x-axis, and the second one corresponds to y-axis

Case $\lambda_1 < \lambda_2 < 0$. Here the general solution shows that the direction on any orbits is from $\pm\infty$ to the origin as t goes from $-\infty$ to ∞. Now however when $t \to \infty$ $e^{\lambda_2 t}$ is much much bigger than $e^{\lambda_1 t}$ and hence the orbits behave as the second eigenvector v_2, whereas for $t \to -\infty$ the first eigenvector becomes dominant, and therefore the orbits far from the origin are parallel to the first eigenvector (see the right panel of the figure above). The same conclusion can be seen from the equation for the phase curves

$$y = Ax^{\gamma}, \qquad 0 < \gamma < 1.$$

It is a good exercise to consider two remaining cases $\lambda_2 < \lambda_1 < 0$ and $\lambda_1 > \lambda_2 > 0$.

Case $\lambda_1 = \lambda_2 < 0$. To make our discussion full, consider also the case of equal negative eigenvalues for the diagonal matrix $A = \mathrm{diag}(\lambda_1, \lambda_2)$. Since the eigenvalues are negative, the direction on the orbits is *to* the origin (see the left panel in the figure below). *Q:* Do you know what to change on the figure to present a phase portrait for $\lambda_1 = \lambda_2 > 0$?

Up till now we discussed only diagonal matrix $A = \mathrm{diag}(\lambda_1, \lambda_2)$, where λ_1 and λ_2 are real numbers. It turns out that not much changes if we consider a general matrix A with two real eigenvalues. Consider two examples.

Example 1. Consider the system

$$\dot{x} = x + 3y,$$
$$\dot{y} = x - y,$$

which means that we have matrix

$$A = \begin{matrix} 1 & 3 \\ 1 & -1 \end{matrix}$$

with the eigenvalues $\lambda_1 = 2$ and $\lambda_2 = -2$ (cf. our very first example of the saddle point) and the corresponding eigenvectors $v_1 = (3, 1)^T$ and $v_2 = (-1, 1)^T$. Hence the general solution to our problem is given by

$$y(t) \qquad {}_1 \qquad {}_1$$

$$x(t) = C_1 \begin{matrix} 3 \\ 1 \end{matrix} e^{2t} + C_2 \begin{matrix} -1 \\ 1 \end{matrix} e^{-2t}.$$

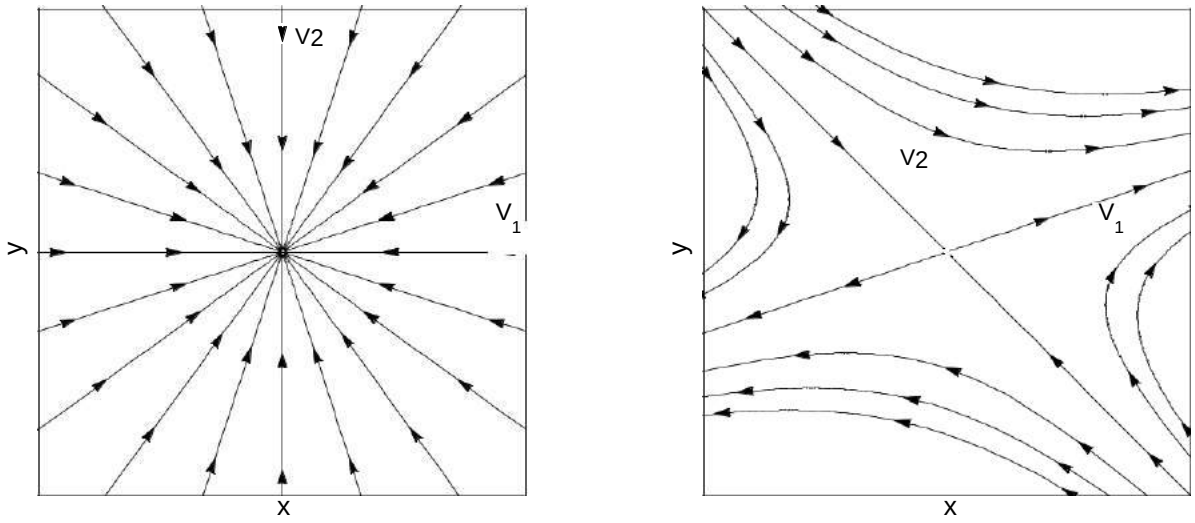


Figure 3: Left: Node. The case $\lambda_1 = \lambda_2 < 0$. The first eigenvector corresponds to x-axis, and the second one corresponds to y-axis. Right: Saddle. The case $\lambda_1 > 0 > \lambda_2$. Here the eigenvector directions do not coincides with the axes

The analysis of the general solution shows that on the characteristic direction corresponding to v_1 the orbits point *from* the origin (because $\lambda_1 = 2 > 0$) and on the line with the direction v_2 the orbits point to the origin (because $\lambda_2 = -2 < 0$). The rest of the orbits do not approach the origin either for $t \to \infty$ or for $t \to -\infty$, and can be plotted by continuity (see the right panel in the figure above).

So what is exactly different from the first example in this lecture, when we considered the diagonal matrix? Not much actually. You can see that the resulting picture in this example can be obtained from the figure of the first considered case by some stretching and rotation (without any tear!). It is true in general: If the eigenvalues of a general matrix A_2 are real and have opposite sign, then the origin is a saddle. To plot it you first need to plot two lines with the directions corresponding to the eigenvectors, put arrows to the origin on the line which corresponds to the negative eigenvalue, and put the arrows from the origin on the line corresponding to the positive eigenvalue. The rest of the orbits are plotted by continuity and remembering that the orbits cannot intersect.

Example 2. Consider

$$\dot{x} = -y,$$
$$\dot{y} = 8x - 6y,$$

hence our eigenvalues are -2, -4 with the eigenvectors $(1, 2)^T$, $(1, 4)^T$. Because both eigenvalues are negative, we know that all the orbits approach the origin when $t \to \infty$. Again, the only subtle thing here is to decide along which direction the orbits tend to the origin. Since we have that $\lambda_2 < \lambda_1 < 0$ then for $t \to \infty$ λ_1 is more important, hence the orbits will be parallel to v_1 when the orbit is close to the origin. For $t \to -\infty$, far from the origin, λ_2 becomes dominant, and therefore the orbits will be parallel to v_2 (see the figure, left panel).

Therefore, for any matrix with two real eigenvalues λ_1 and λ_2 with two distinct eigenvectors, we have the equilibrium point at the origin, which is called node. This point *attracts* orbits (in the sense that the direction on the orbits points to the origin) if eigenvalues are negative and *repels* them if eigenvalues are positive. To determine the direction along which the orbits approach the origin, you need to look for the eigenvector that corresponds to the eigenvalue that is closer to zero. Contrary, to see the behavior of the orbits far from the origin, we need to look for the direction of the eigenvector corresponding to the eigenvalue that is further from zero.

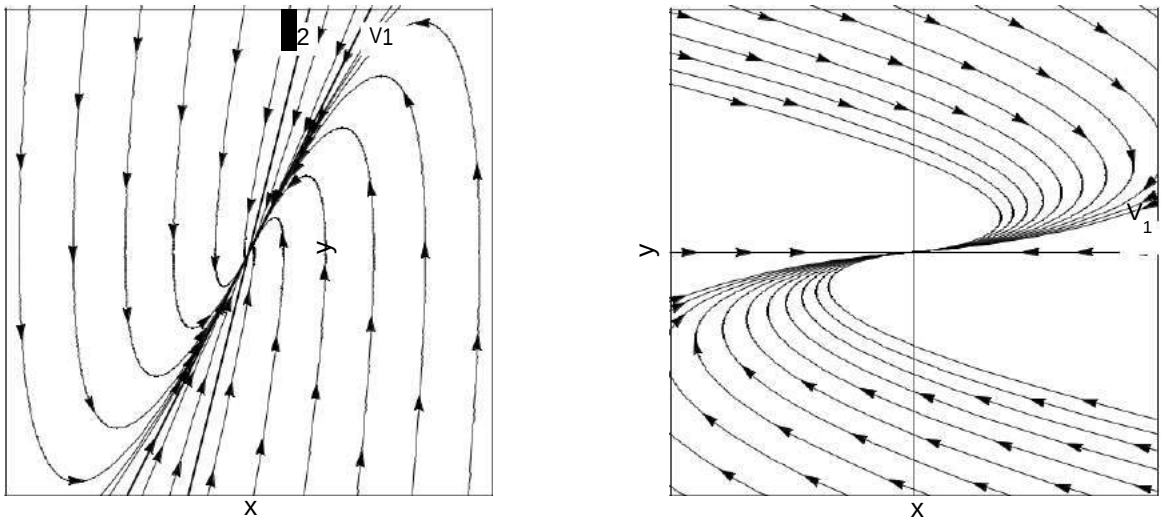


Figure 4: Left: Node. The case $\lambda_2 < \lambda_1 < 0$. Here the eigenvector directions do not coincides with the axes. Right: The case of equal eigenvalues with only one eigenvector

To conclude the discussion of the case when the matrix has real eigenvalues, recall that it is possible to have equal eigenvalues with only one eigenvector. Consider such a case with

$$A = \begin{matrix} 0 & \lambda \\ \lambda & 1 \end{matrix},$$

which has eigenvalue λ multiplicity 2 with the eigenvector $v = (1, 0)^T$. In this case we have only one characteristic direction, and the orbits approach the origin along it. Moreover, the direction on the orbits from the origin if $\lambda > 0$ and to the origin is $\lambda < 0$ (see figure, right panel for an example).

Now we switch to the case when the eigenvalues are not real.

Consider the system

$$\begin{aligned} x' &= ax - by, \\ y' &= bx + ay, \end{aligned}$$

for some constants a and b. Hence we have the matrix

$$A = \begin{matrix} a & -b \\ b & a \end{matrix}.$$

We can analyze this case exactly as before, by finding the eigenvalues and eigenvectors (for example, the eigenvalues are $\lambda_{1,2} = a \pm ib$). However, there is a simple way to figure out the behavior of the orbits. Multiply the second equation by i and add them, after some simplifications we will find that for $z = x + iy$ it is true that:

$$z' = \lambda z, \ \lambda = a + ib.$$

Now I use the polar form of a complex number $z = \rho e^{i\theta}$, from which (fill in the details)

$$\rho' = a\rho, \quad \theta = b,$$

from which we find

$$\rho(t) = C_1 e^{at}, \qquad \theta(t) = bt + C_2,$$

for the polar coordinates ρ and θ. To see how exactly the orbits look like, assume that a > 0 and b > 0. Therefore, we find that ρ → ∞ as t → ∞ and ρ → 0 as t → −∞. For b positive it means that the polar angle changes in the positive direction, which is counterclockwise. Considering superposition of these two movements, we find that all the orbits (safe for the equilibrium point at the origin) are spiral, which the direction on them *from* the origin. An example for the same matrix with a < 0 and b > 0 is given on the right panel in the same figure. An equilibrium with such structure of orbits is called *focus* or *spiral*.

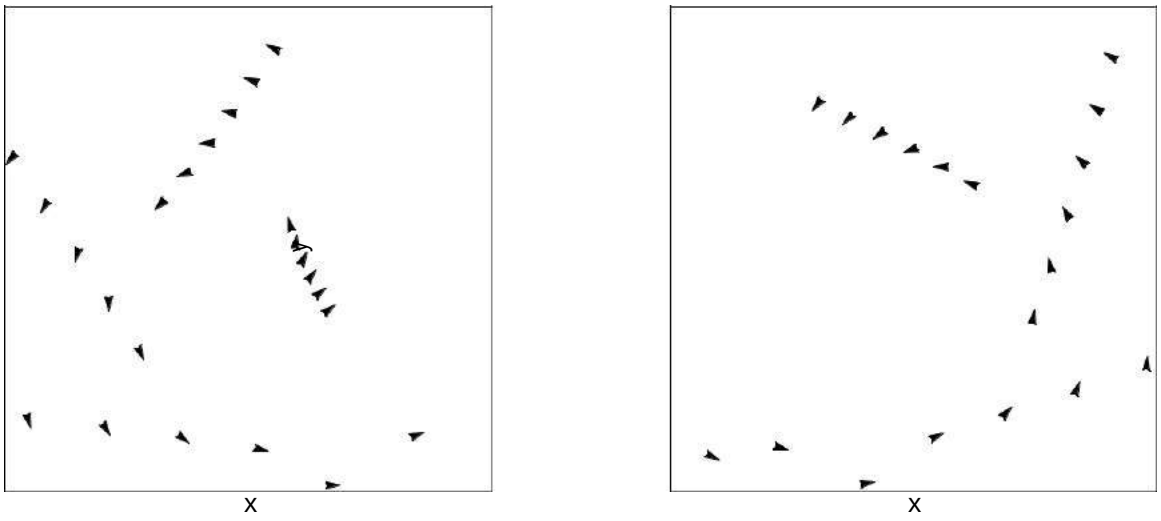


Figure 5: Left: Focus. a > 0, b > 0. Right: Focus. a < 0, b > 0

The general case is when the matrix has two complex conjugate eigenvalues $\lambda_{1,2} = a \pm b$. By the sign of a we know the direction on the orbits: If a > 0 then the direction is from the origin, and if a < 0 then the direction towards the origin. However, the subtle thing here is to determine whether the rotation occurs clockwise or counterclockwise. To make sure that you find the correct direction, it is useful to take any point (x, y) = (0, 0) and find the direction at this point (this direction is given be the vector $(a_{11}x + a_{12}y, a_{21}x + a_{22}y)$). If you know whether the origin attracts or repels the orbits and one point with the precise direction, it becomes clear what is the whole picture on the phase plane.

As an example, and to make the discussion complete, consider the linear system with the matrix

$$A = \begin{matrix} 1 & 0 \\ 0 & -3 \end{matrix},$$

with the eigenvalues $\lambda_{1,2} = \pm i\sqrt{3}$. Since a = 0 we find that the orbits neither approach nor leave the origin. Actually, it can be shown that all the orbits in this case are ellipses. To infer the direction of rotation, pick a point, e.g., (x, y) = (1, 0). At this point we find the vector $(0, 1)^T$, which points in the counterclockwise direction, hence the orbits look like ellipses with the counterclockwise directions on them (see the figure). An equilibrium with such phase portrait is called *center*.

Summary

In the previous section we found that it is possible to have the phase portrait around the origin that belongs to one of the following types:

- saddle (two real eigenvalues of opposite sign);
- node (two real eigenvalues of the same sign);

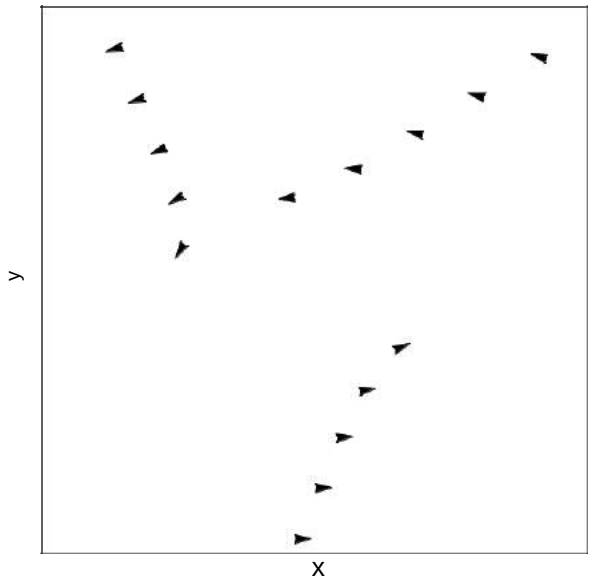


Figure 6: Center, the eigenvalues are purely imaginary

- degenerate node (one eigenvalue of multiplicity two with only one eigenvector);
- focus (two complex conjugate eigenvalues, moreover, Re $\lambda = 0$);
- center (two imaginary eigenvalues, Re $\lambda = 0$).

However, the analysis was mostly based on three types of matrices:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} a & b \\ b & -a \end{pmatrix}.$$

To state why it is enough to consider only these three matrices, I will need the following

Definition 3. *Matrices* $A, B \in M_n(\mathbb{R})$ *are called similar is there exists an invertible matrix* S*, such that*

$$A = S^{-1}BS.$$

Similar matrices share a lot of properties. For example, they have the same eigenvalues (can you prove it?). The main fact is the next theorem, which I state without proof.

Theorem 4. *Any* 2×2 *matrix is similar to one of the three matrices above.*

As an important corollary we obtain that for any matrix $A \in M_2(\mathbb{R})$ such that $\det A = 0$, the only possible phase portraits are given in the previous section.

Stability of the origin

Having at our disposable all the possible phase portraits of linear planar systems of ODE makes it very intuitively clear what it means to have the origin stable.

Definition 5. *The origin of the linear system*

$$\dot{x} = a_{11}x + a_{12}y,$$
$$\dot{y} = a_{22}x + a_{22}y,$$

with $\det A = 0$ *is called*

- *Lyapunov stable, if any orbit, starting close enough to the origin, stays close to the origin for all positive* t*;*
- *asymptotically stable, if any orbit, starting close enough to the origin, is 1) Lyapunov stable, and 2) tends to the origin as* $t \to \infty$*;*
- *unstable, if there exits an orbit starting close enough to the origin that leaves a small neighborhood of the origin for some positive* t*.*

Using this definition we find that

- saddles are always unstable, since it is possible to find orbits close to the origin that eventually leave any neighborhood of the origin;
- nodes can be either asymptotically stable (this requires that both eigenvalues are negative) or unstable (both eigenvalues are positive);
- foci can be either asymptotically stable (if $\mathrm{Re}\,\lambda_{1,2} < 0$) or unstable ($\mathrm{Re}\,\lambda_{1,2} > 0$);
- center is Lyapunov stable, but not asymptotically stable, since the orbits do not approach the origin.

Putting everything together, we obtain a very important fact that says that

Theorem 6. *The origin of the linear planar system with the matrix* A *such that* $\det A = 0$*, is stable if for all the eigenvalues* $\mathrm{Re}\,\lambda \le 0$ *and unstable otherwise. Moreover, it is asymptotically stable if* $\mathrm{Re}\,\lambda_{1,2} < 0$*.*

Since there is only one equilibrium in the linear system, it is often said that the system is stable or asymptotically stable, meaning that the origin is stable or asymptotically stable.

It is convenient to summarize all the information about the linear planar systems using two parameters: trace and determinant of matrix A. For an arbitrary matrix A the characteristic polynomial has the form

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = \lambda^2 - \mathrm{tr}\,A\,\lambda + \det A,$$

where I used the notation $\mathrm{tr}\,A := a_{11} + a_{22}$ for the *trace* of matrix A, which is given by the sum of the elements on the main diagonal. Therefore, the eigenvalues can be found as

$$\lambda_{1,2} = \frac{\mathrm{tr}\,A \pm \sqrt{(\mathrm{tr}\,A)^2 - 4\det A}}{2}.$$

For example, if $\mathrm{tr}\,A > 0$, $\det A > 0$, and $= (\mathrm{tr}\,A)^2 - 4\det A > 0$ then we obtain that there are two real eigenvalues of the same sign, which means that in this parameter region we have unstable node.

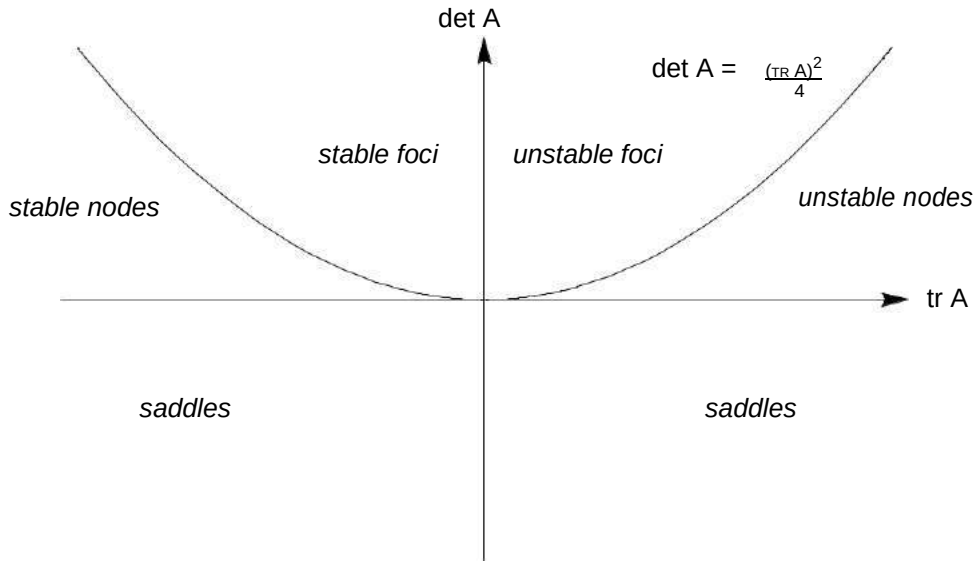
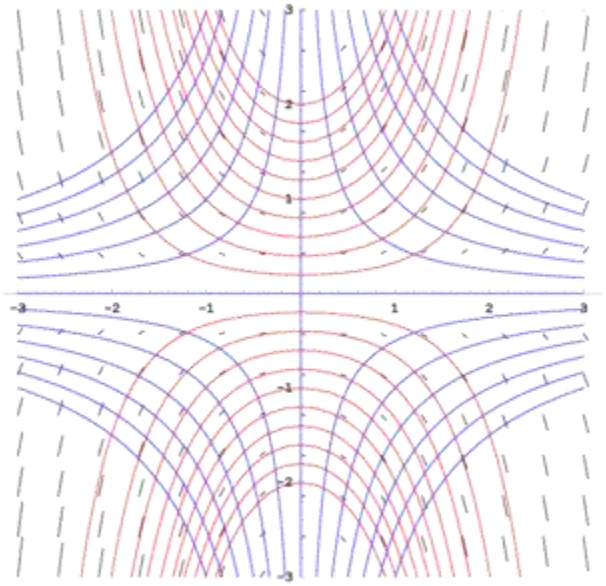


Figure 7: Parametric portrait of linear planar systems. There are six domains here, with the boundaries det A = 0, tr A = 0 and det A = $(\text{tr } A)^2/4$. Note that on the line tr A = 0 when det A > 0 we have centers, and if det A < 0 we still have saddles. On the line det A = $(\text{tr } A)^2/4$ we have two equal eigenvalues, and hence degenerate nodes

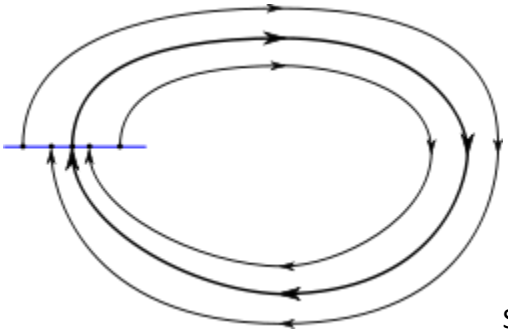
Method of isoclines:- **Isoclines** are often used as a graphical **method** of solving ordinary differential equations. In an equation of the form y' = f(x,y), the **isoclines** are lines in the (x, y) plane obtained by setting f(x,y) equal to a constant.



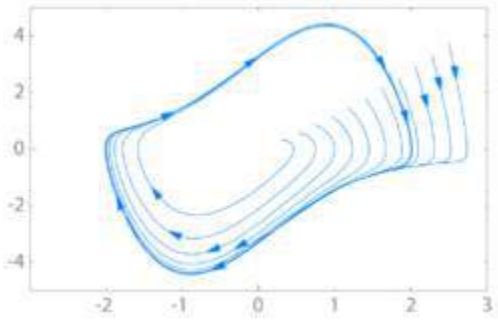
Given a family of curves, assumed to be differentiable, an **isocline** for that family is formed by the set of points at which some member of the family attains a given slope. The word comes from the Greek words ἴσος (isos), meaning "same", and the κλίνειν, meaning "make to slope". Generally, an isocline will itself have the shape of a curve or the union of a small number of curves.

Isoclines are often used as a graphical method of solving ordinary differential equations. In an equation of the form $y' = f(x,y)$, the isoclines are lines in the (x, y) plane obtained by setting $f(x,y)$ equal to a constant. This gives a series of lines (for different constants) along which the solution curves have the same gradient. By calculating this gradient for each isocline, the slope field can be visualised; making it relatively easy to sketch approximate solution curves; as in fig.

Limit cycle:- In mathematics, in the study of dynamical systems with two-dimensional phase space, a **limit cycle** is a closed trajectory in phase space having the property that at least one other trajectory spirals into it either as time approaches infinity or as time approaches negative infinity.



Stable limit cycle

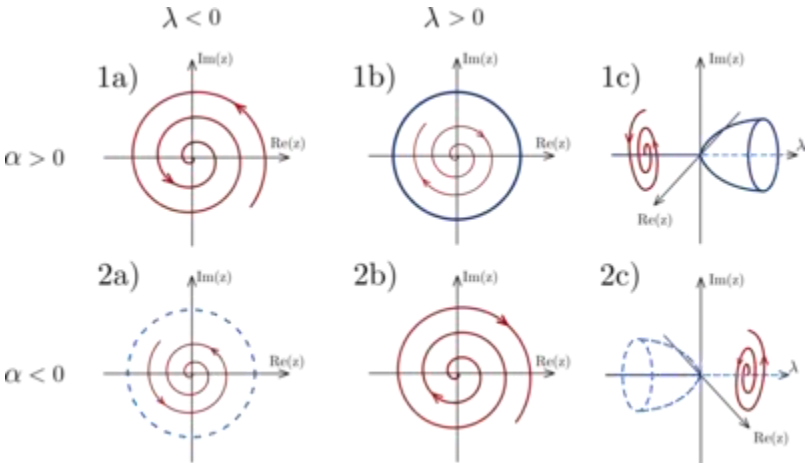


Stable limit cycle with van der pall oscillator

In mathematics, in the study of dynamical systems with two-dimensional phase space, a **limit cycle** is a closed trajectory in phase space having the property that at least one other trajectory spirals into it either as time approaches infinity or as time approaches negative infinity. Such behavior is exhibited in some nonlinear systems. Limit cycles have been used to model the behavior of a great many real world oscillatory systems. The study of limit cycles was initiated by Henri Poincaré.

Stable, Unstable and semi stable limit cycles:- In the case where all the neighbouring trajectories approach the limit cycle as time approaches infinity, it is called a *stable* or *attractive* limit cycle (ω-limit cycle). If instead all neighbouring trajectories approach it as time approaches negative infinity, then it is an *unstable* limit cycle (α-limit cycle). If there is a neighbouring trajectory which spirals into the limit cycle as time approaches infinity, and another one which spirals into it as time approaches negative infinity, then it is a *semi-stable* limit cycle. There are also limit cycles which are neither stable, unstable nor semi-stable: for instance, a neighboring trajectory may approach the limit cycle from the outside, but the inside of the limit cycle is approached by a family of other cycles (which wouldn't be limit cycles).

Stable limit cycles are examples of attractors. They imply self-sustained oscillations: the closed trajectory describes perfect periodic behavior of the system, and any small perturbation from this closed trajectory causes the system to return to it, making the system stick to the limit cycle.



Examples of limit cycles branching from fixpoints near Hopf bifurcation. Trajectories in red, stable structures in dark blue, unstable structures in light blue. The parameter choice determines the occurrence and stability of limit cycles.

Singular points:- **Singular Point**. A **singular point** of an algebraic curve is a **point** where the curve has "nasty" behavior such as a cusp or a **point** of self-intersection (when the underlying field is taken as the reals). More formally, a **point** on a curve is **singular** if the and partial derivatives of are both zero at the **point** .

In geometry, a **singular point** on a curve is one where the curve is not given by a smooth embedding of a parameter. The precise definition of a singular point depends on the type of curve being studied.

Algebraic curves in the plane may be defined as the set of points (x, y) satisfying an equation of the form $f(x, y)=0$, where f is a polynomial function $f:\mathbf{R}^2\to\mathbf{R}$. If f is expanded as

f = a+bx+by+cx+……..

b0=df/dx

b1=df/dy

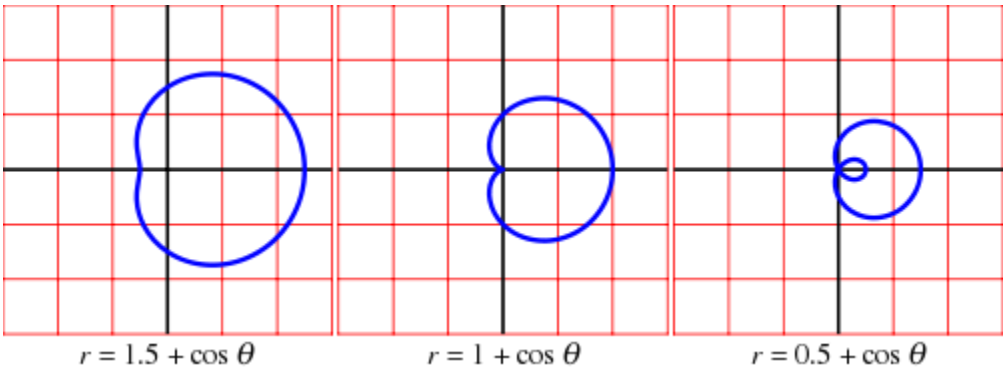
If the origin (0, 0) is on the curve then $a_0=0$. If $b_1\neq 0$ then the implicit function theorem guarantees there is a smooth function h so that the curve has the form $y=h(x)$ near the origin. Similarly, if $b_0\neq 0$ then there is a smooth function k so that the curve has the form $x=k(y)$ near the origin. In either case, there is a smooth map from **R** to the plane which defines the curve in the neighborhood of the origin. Note that at the origin.

f (x,y) = df/dx = df/dy = 0

Regular points:- Assume the curve passes through the origin and write $y=mx$. Then f can be written

f = (b0+mb1)x + (c0+2mc1)+…….

If b_0+mb_1 is not 0 then $f=0$ has a solution of multiplicity 1 at $x=0$ and the origin is a point of single contact with line $y=mx$. If $b_0+mb_1=0$ then $f=0$ has a solution of multiplicity 2 or higher and the line $y=mx$, or $b_0x+b_1y=0$, is tangent to the curve. In this case, if $c_0+2mc_1+c_2m^2$ is not 0 then the curve has a point of double contact with $y=mx$. If the coefficient of x^2, $c_0+2mc_1+c_2m^2$, is 0 but the coefficient of x^3 is not then the origin is a point of inflection of the curve. If the coefficients of x^2 and x^3 are both 0 then the origin is called *point of undulation* of the curve. This analysis can be applied to any point on the curve by translating the coordinate axes so that the origin is at the given point.



the left curve acquires an acnode at the origin, which is an isolated point in the plane. The central curve, the cardioid, has a cusp at the origin. The right curve has a crunode at the origin and the curve crosses itself to form a loop.

Some of the possible singularities are:-

An isolated point, acnode

Two point crossing, crunode

A cusp, spinode

A tacnode

A rhamhoid cusp

Then amongst singular points, an important distinction is made between a **regular singular point**, where the growth of solutions is bounded (in any small sector) by an algebraic function, and an **irregular singular point**, where the full solution set requires functions with higher growth rates. This distinction occurs, for example, between the hypergeometric equation, with three regular singular points, and the Bessel equation which is in a sense a limiting case, but where the analytic properties are substantially different.

Stability of non linear system:- The behavior of a nonlinear system is described in mathematics by a **nonlinear system of equations**, which is a set of simultaneous equations in which the unknowns (or the unknown functions in the case of differential equations) appear as variables of a polynomial of degree higher than one or in the argument of a function which is not a polynomial of degree one. In other words, in a nonlinear system of equations, the equation(s) to be solved cannot be written as a linear combination of the unknown variables or functions that appear in them. Systems can be defined as nonlinear, regardless of whether known linear functions appear in the equations. In particular, a differential equation is *linear* if it is linear in terms of the unknown function and its derivatives, even if nonlinear in terms of the other variables appearing in it.

As nonlinear dynamical equations are difficult to solve, nonlinear systems are commonly approximated by linear equations (linearization). This works well up to some accuracy and some range for the input values, but some interesting phenomena such as solitons, chaos, and singularities are hidden by linearization. It follows that some aspects of the dynamic behavior of a nonlinear system can appear to be counterintuitive, unpredictable or even chaotic. Although such chaotic behavior may resemble random behavior, it is in fact not random. For example, some aspects of the weather are seen to be chaotic, where simple changes in one part of the system produce complex effects throughout. This nonlinearity is one of the reasons why accurate long-term forecasts are impossible with current technology.

stability theory addresses the stability of solutions of differential equations and of trajectories of dynamical systems under small perturbations of initial conditions. The heat equation, for example, is a stable partial differential equation because small perturbations of initial data lead to small variations in temperature at a later time as a result of the maximum principle. In partial differential equations one may measure the distances between functions using Lp norms or the sup norm, while in differential geometry one may measure the distance between spaces using the Gromov–Hausdorff distance.

In dynamical systems, an orbit is called *Lyapunov stable* if the forward orbit of any point is in a small enough neighborhood or it stays in a small (but perhaps, larger) neighborhood. Various criteria have been developed to prove stability or instability of an orbit. Under favorable circumstances, the question may be reduced to a well-studied problem involving eigenvalues of matrices. A more general method involves Lyapunov functions. In practice, any one of a number of different stability criteria are applied.

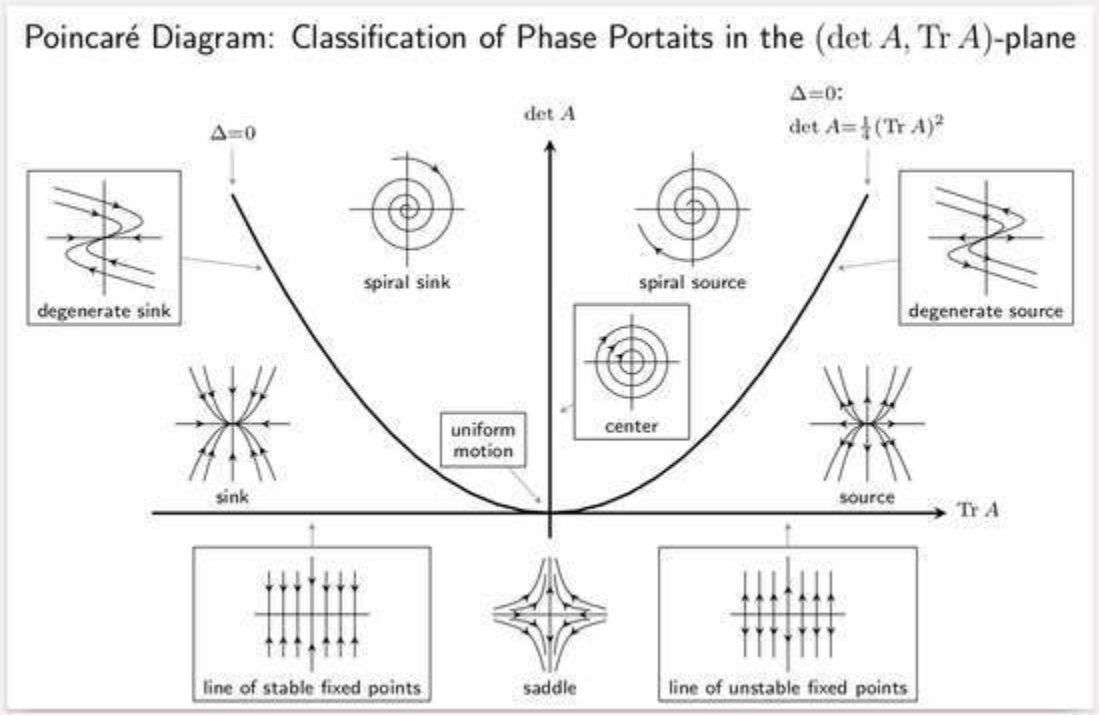
Many parts of the qualitative theory of differential equations and dynamical systems deal with asymptotic properties of solutions and the trajectories—what happens with the system after a long period of time. The simplest kind of behavior is exhibited by equilibrium points, or fixed points, and by periodic orbits. If a particular orbit is well understood, it is natural to ask next whether a small change in the initial condition will lead to similar behavior.

Stability means that the trajectories do not change too much under small perturbations. The opposite situation, where a nearby orbit is getting repelled from the given orbit, is also of interest. In general, perturbing the initial state in some directions results in the trajectory asymptotically approaching the given one and in other directions to the trajectory getting away from it. There may also be directions for which the behavior of the perturbed orbit is more complicated (neither converging nor escaping completely), and then stability theory does not give sufficient information about the dynamics.

The simplest kind of an orbit is a fixed point, or an equilibrium. If a mechanical system is in a stable equilibrium state then a small push will result in a localized motion, for example, small oscillations as in the case of a pendulum. In a system with damping, a stable equilibrium state is moreover asymptotically stable. On the other hand, for an unstable equilibrium, such as a ball resting on a top of a hill, certain

small pushes will result in a motion with a large amplitude that may or may not converge to the original state.

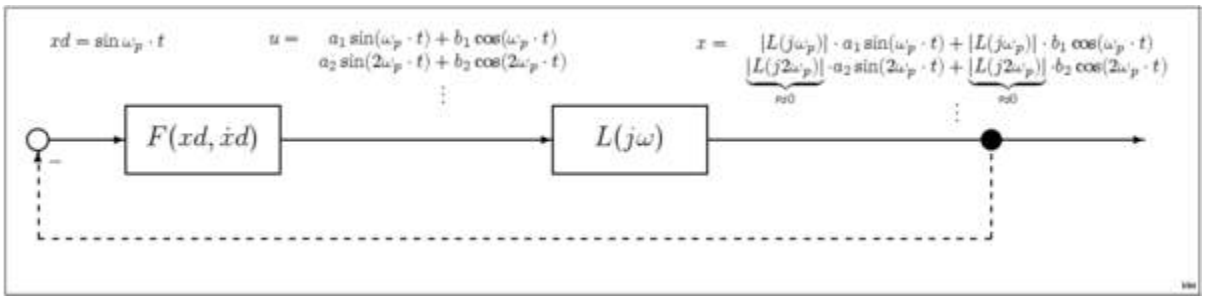
There are useful tests of stability for the case of a linear system. Stability of a nonlinear system can often be inferred from the stability of its linearization.



UNIT 3

Describing function:-

In control systems theory, the **describing function** (DF) method, developed by NikolayMitrofanovichKrylov and NikolayBogoliubov in the 1930s, and extended by Ralph Kochenburgeris an approximate procedure for analyzing certain nonlinear control problems. It is based on quasi-linearization, which is the approximation of the non-linear system under investigation by a linear time-invariant (LTI) transfer function that depends on the amplitude of the input waveform. By definition, a transfer function of a true LTI system cannot depend on the amplitude of the input function because an LTI system is linear. Thus, this dependence on amplitude generates a family of linear systems that are combined in an attempt to capture salient features of the non-linear system behavior. The describing function is one of the few widely applicable methods for designing nonlinear systems, and is very widely used as a standard mathematical tool for analyzing limit cycles in closed-loop controllers, such as industrial process controls, servomechanisms, and electronic oscillators.



Non linear system in the state of harmonic balance

Consider feedback around a discontinuous (but piecewise continuous) nonlinearity (e.g., an amplifier with saturation, or an element with deadband effects) cascaded with a slow stable linear system. The continuous region in which the feedback is presented to the nonlinearity depends on the amplitude of the output of the linear system. As the linear system's output amplitude decays, the nonlinearity may move into a different continuous region. This switching from one continuous region to another can generate periodic oscillations. The describing function method attempts to predict characteristics of those oscillations (e.g., their fundamental frequency) by assuming that the slow system acts like a low-pass or bandpass filter that concentrates all energy around a single frequency. Even if the output waveform has several modes, the method can still provide intuition about properties like frequency and possibly amplitude; in this case, the describing function method can be thought of as describing the sliding mode of the feedback system.

Other types of describing functions that have been used are DFs for level inputs and for Gaussian noise inputs. Although not a complete description of the system, the DFs often suffice to answer specific questions about control and stability. DF methods are best for analyzing systems with relatively weak nonlinearities. In addition the higher order sinusoidal input describing functions (HOSIDF), describe the response of a class of nonlinear systems at harmonics of the input frequency of a sinusoidal input. The HOSIDFs are an extension of the SIDF for systems where the nonlinearities are significant in the response.

Although the describing function method can produce reasonably accurate results for a wide class of systems, it can fail badly for others. For example, the method can fail if the system emphasizes higher harmonics of the nonlinearity. Such examples have been presented by Tzypkin for bang–bang systems. A fairly similar example is a closed-loop oscillator consisting of a non-inverting Schmitt trigger followed by an *inverting* integrator that feeds back its output to the Schmitt trigger's input. The output of the Schmitt trigger is going to be a square waveform, while that of the integrator (following it) is going to have a triangle waveform with peaks coinciding with the transitions in the square wave. Each of these two oscillator stages lags the signal exactly by 90 degrees (relative to its input). If one were to perform DF analysis on this circuit, the triangle wave at the Schmitt trigger's input would be replaced by its fundamental (sine wave), which passing through the trigger would cause a phase shift of less than 90 degrees (because the sine wave would trigger it sooner than the triangle wave does) so the system would appear not to oscillate in the same (simple) way.

Limitations of Describing function:-

1. The describing-function method of analysis fails to predict the existence of a limit cycle when applied to a relay control system for a nuclear reactor.
2. This is demonstrated by a direct solution of the system differential equations and by experiment on the Universities' Research Reactor.
3. It involves complex computations.
4. This process waste a lot of time.
5. It causes various types of non linearities in the system.
6. Produces various errors.

Use of describing function for stability:-

The describing function is an approximate procedure for analyzing certain nonlinear control problems in control engineering. To start, let us first recall the basic definition of a linear control system. Linear control systems are those where the principle of superposition (if the two inputs are applied simultaneously, then the output will be the sum of two outputs) is applicable.

Analysis of different nonlinear controls system is very difficult due to their nonlinear behavior. We cannot use conventional analysis methods such as the Nyquist stability criterion or pole-zero method in order to analyze these nonlinear systems, as these methods are restricted to linear systems.

Describing function for Ideal Relay:-

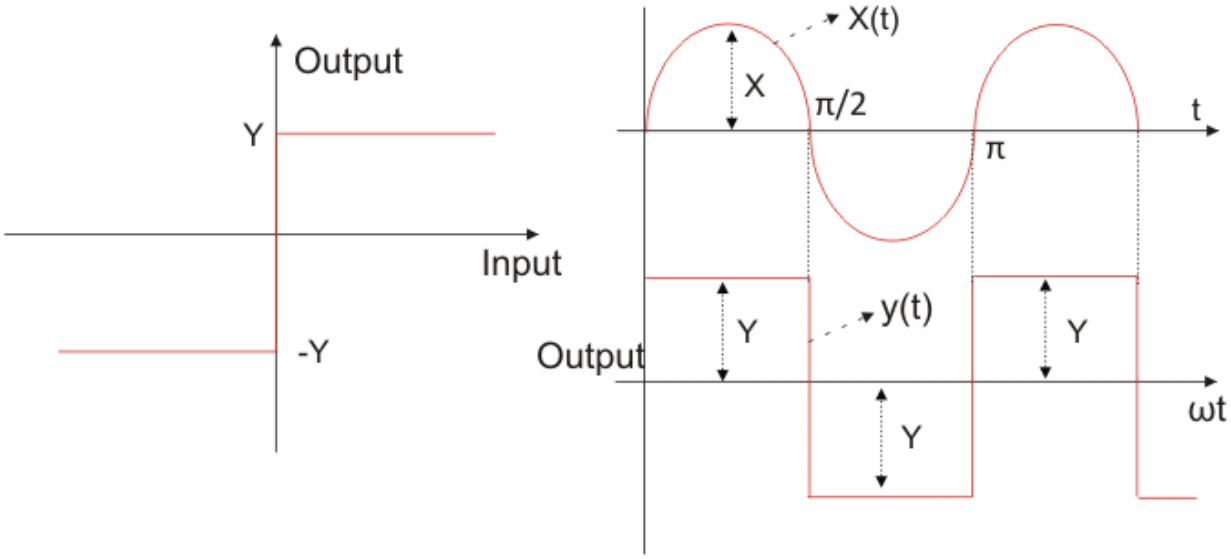
We have the characteristic curve for ideal relay as shown in the given figure.

Let us take input function as:

$$X(t) = X\sin(\omega t).$$

Now from the curve we can define the output as

$$Y(t) = Y \ for \ 0 \leq \omega t \leq \pi$$
$$Y(t) = -Y \ for \ \pi \leq \omega t \leq 2\pi$$



The output periodic function has odd symmetry :

$$y(\omega t) = -y(-\omega t)$$

Let us first calculate Fourier series constant A_1.

$$A_1 = \frac{1}{\pi}\int_0^{2\pi} y(t)\cos\omega t \ d(\omega t)$$

On substituting the value of the output in the above equation and integrating the function from 0 to 2π we have the value of the constant A_1 as zero.

Similarly we can calculate the value of Fourier constant B_1 for the given output and the value of B_1 can be calculated as

$$B_1 = \frac{1}{\pi}\int_0^{2\pi} y(t)\sin\omega t \ d(\omega t)$$
$$= \frac{2}{\pi}\int_0^{\pi} y(t)\sin\omega t \ d(\omega t)$$

On substituting the value of the output in the above equation y(t) = Y we have the value of the constant B_1

$$B_1 = \frac{2Y}{\pi}\int_0^{\pi} y(t)\sin\omega t\, d(\omega t) = \frac{4Y}{\pi}$$

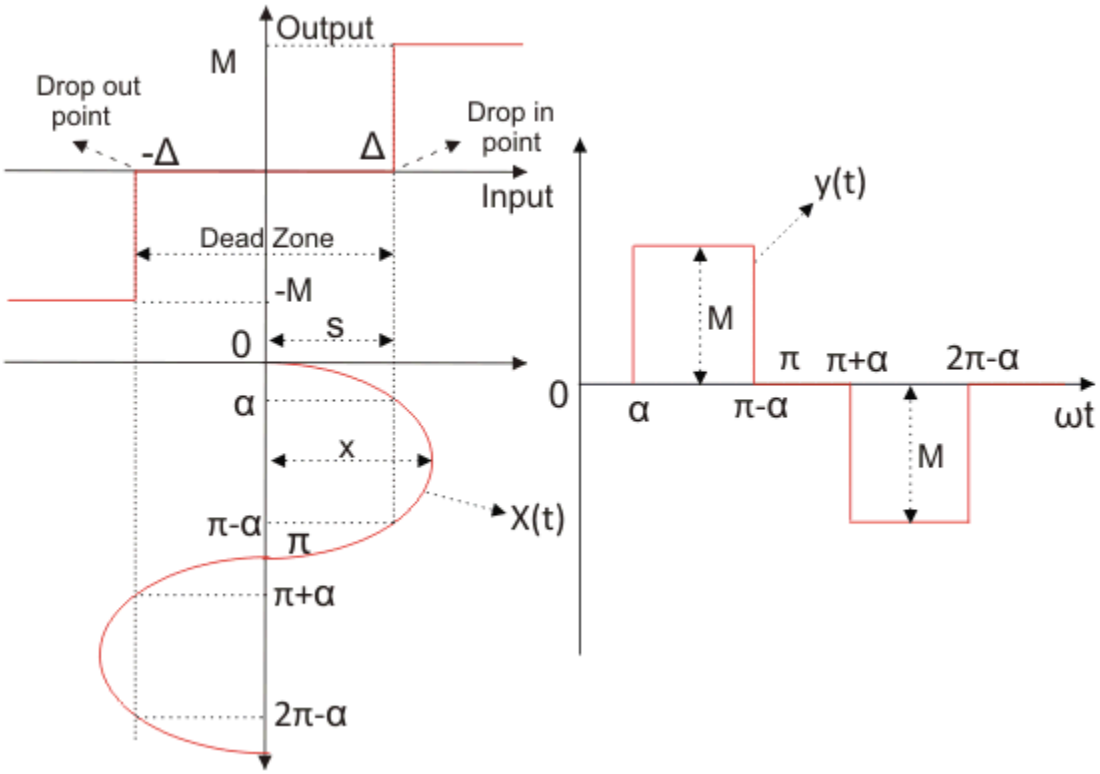
And the phase angle for the describing function can be calculated as

$$\angle \tan^{-1}\left(\frac{A_1}{B_1}\right) = \angle 0^\circ$$

Thus the describing function for an ideal relay is

$$N = \frac{Y_1}{X}\angle 0^\circ = \frac{4Y}{\pi X}\angle 0^\circ$$

<u>Describing function for Relay (Relay with Dead Zone):-</u>



We have the characteristic curve for real realy as shown in the given figure. If X is less than dead zone Δ, then the relay produces no output; the first harmonic component of Fourier series is of course zero and describing function is also zero. If X > Δ the relay produces the output.

Let us take input function as:

$$X(t) = X\sin(\omega t).$$

Now from the curve we can define the output as:

$$Y(t) = 0 \; for \; 0 \leq \omega t \leq \alpha$$
$$Y(t) = M \; for \; \alpha \leq \omega t \leq (\pi - \alpha)$$
$$Y(t) = 0 \; for \; (\pi - \alpha) \leq \omega t \leq \pi$$
$$Y(t) = -M \; for \; (\pi + \alpha) \leq \omega t \leq (2\pi - \alpha)$$
$$Y(t) = 0 \; for \; (2\pi - \alpha) \leq \omega t \leq 2\pi$$

$$Where, \; X\sin\alpha = \Delta$$
$$\Rightarrow \alpha = \sin^{-1}\left(\frac{\Delta}{X}\right)$$

The output periodic function has odd symmetry :

$$y(\omega t) = -y(-\omega t)$$

Let us first calculate Fourier series constant A_1.

$$A_1 = \frac{1}{\pi}\int_0^{2\pi} y(t)\cos\omega t \; d(\omega t)$$

On substituting the value of the output in the above equation and integrating the function from 0 to 2π we have the value of the constant A_1 as zero.

Similarly we can calculate the value of Fourier constant B for the given output and the value of B can be calculated as

$$B_1 = \frac{1}{\pi}\int_0^{2\pi} y\sin\omega t \; d(\omega t)$$

Due to the symmetry of y, the coefficient B_1 can be calculated as follows,

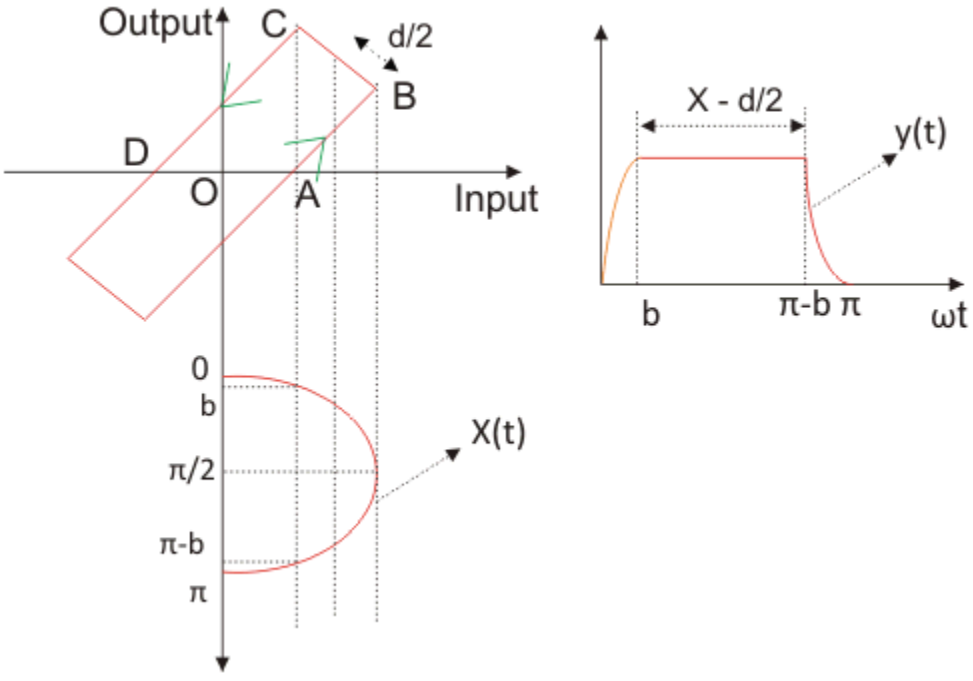
$$B_1 = \frac{4}{\pi}\int_0^{\pi/2} y \sin\omega t \; d(\omega t)$$
$$= \frac{4M}{\pi}\int_\alpha^{\pi/2} y \sin\omega t \; d(\omega t)$$
$$= \frac{4M}{\pi}\cos\alpha$$

Therefore, the describing function is

$$N = \frac{4M}{\pi X}\cos\alpha \; \angle 0^\circ$$
$$= \frac{4M}{\pi X}\sqrt{1-\left(\frac{\Delta}{X}\right)^2} \; \angle 0^\circ \quad \left[Since \; \sin\alpha = \frac{\Delta}{X}\right]$$

Describing function for Backlash Non Linearity:-

We have the characteristic curve for backlash as shown in the given figure. Let us take input function as



$$X(t) = X\sin(\omega t).$$

Now from the curve we can define the output as

$$Y(t) = K\left[X\sin\left(\omega t - \frac{d}{2}\right)\right] \quad for \;\; 0 \le \omega t \le \frac{\pi}{2}$$
$$Y(t) = K\left[X - \frac{d}{2}\right] \quad for \;\; \frac{\pi}{2} \le \omega t \le (\pi - b)$$
$$Y(t) = K\left[X\sin\left(\omega t + \frac{d}{2}\right)\right] \quad for \;\; (\pi - b) \le \omega t \le \pi$$

Let us first calculate Fourier series constant A_1.

$$A_1 = \frac{1}{\pi}\int_0^{2\pi} y(t)\cos\omega t \; d(\omega t)$$

On substituting the value of the output in the above equation and integrating the function from zero to 2π we have the value of the constant A_1 as

$$A_1 = \frac{4KX}{\pi}\left[\frac{(d/2)^2}{X^2} - \frac{(d/2)}{X}\right]$$

Similarly we can calculate the value of Fourier constant B for the given output and the value of B_1 can be calculated as

$$B_1 = \frac{1}{2\pi}\int_0^{2\pi} y(t)\sin\omega t \; d(\omega t)$$

On substituting the value of the output in the above equation and integrating the function from zero to pi we have the value of the constant B_1 as

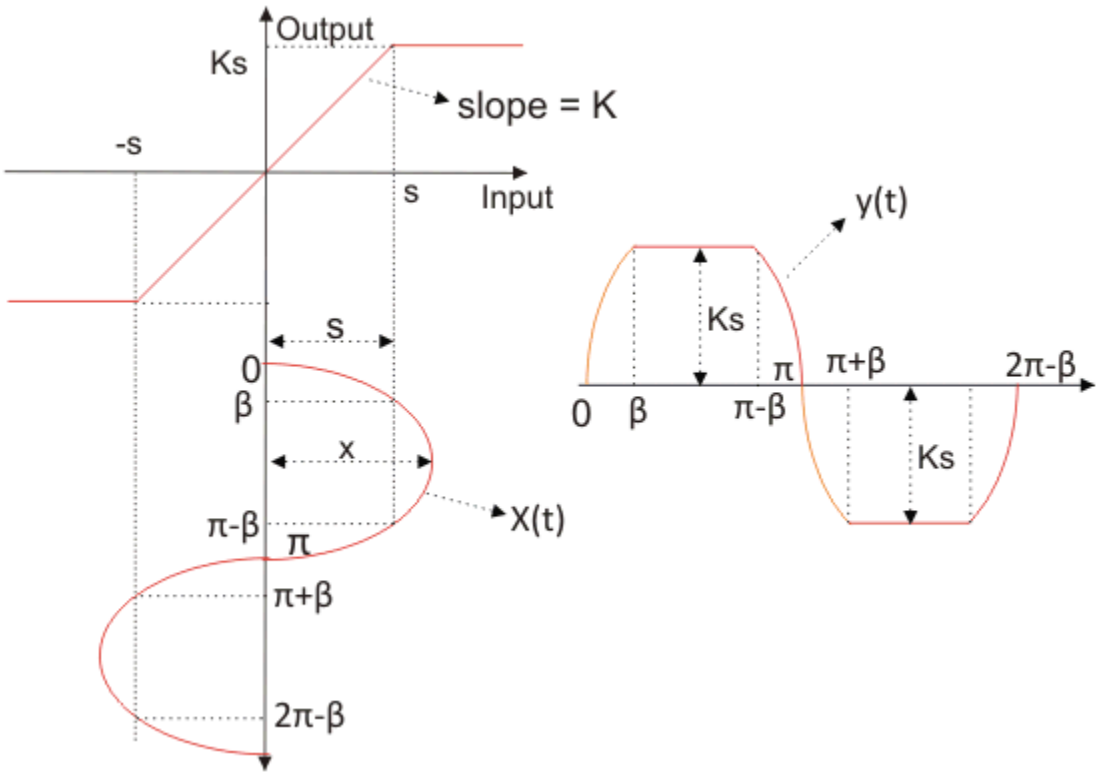
$$B_1 = \frac{KX}{\pi}\left[\frac{\pi}{2} + b + \frac{d(X-d)}{X^2}\sqrt{\frac{2X}{d} - 1}\right]$$

We can easily calculate the describing function of backlash from below equation

$$N = \frac{\sqrt{A_1^2 + B_1^2}}{X}\angle\tan^{-1}\left(\frac{A_1}{B_1}\right)$$

Describing function for Saturation Non linearity:-

We have the characteristic curve for saturation as shown in the given figure.



Let us take input function as

$$X(t) = X\sin(\omega t).$$

Now from the curve we can define the output as:

$$Y(t) = KX\sin\omega t \; for \; 0 \leq \omega t \leq \beta$$
$$Y(t) = Ks \; for \; \beta \leq \omega t \leq (\pi - \beta)$$
$$Y(t) = KX\sin\omega t \; for \; (\pi - \beta) \leq \omega t \leq \pi$$

Let us first calculate Fourier series constant A_1.

$$A_1 = \frac{1}{\pi}\int_0^{2\pi} y(t)\cos\omega t \; d(\omega t)$$

On substituting the value of the output in the above equation and integrating the function from 0 to 2π we have the value of the constant A1 as zero.

Similarly we can calculate the value of Fourier constant B_1 for the given output and the value of B_1 can be calculated as,

$$B_1 = \frac{1}{\pi}\int_0^{2\pi} y(t)\sin\omega t\ d(\omega t)$$
$$= \frac{4}{\pi}\int_0^{\pi/2} y(t)\sin\omega t\ d(\omega t)$$
$$= \frac{1}{\pi}\left[\int_0^{\beta} KX\sin^2\omega t\ d(\omega t) + \int_{\beta}^{\pi/2} Ks \times \sin\omega t\ d(\omega t)\right]$$
$$= \frac{4K}{\pi}\left[\frac{X\beta}{2} - \frac{X}{4}\sin 2\beta + s\cos\beta\right]$$
$$= \frac{2KX}{\pi}\left[\beta + 2\frac{S}{X}\cos\beta - \sin\beta\cos\beta\right]$$

$$Now\ KX\sin\omega t = Ks\ when\ \omega t = \beta$$
$$So,\ \sin = \frac{Ks}{KX}$$
$$\Rightarrow \beta = \sin^{-1}\frac{s}{X}$$
$$\therefore B_1 = \frac{2KX}{\pi}\left[\sin^{-1}\left(\frac{s}{X}\right) + 2\frac{S}{X}\cos\sin^{-1}\left(\frac{s}{X}\right) - \sin\sin^{-1}\left(\frac{s}{X}\right).\cos\sin^{-1}\left(\frac{s}{X}\right)\right]$$
$$= \frac{2KX}{\pi}\left[\sin^{-1}\left(\frac{s}{X}\right) + \frac{s}{X}\sqrt{1-\left(\frac{s}{X}\right)^2}\right]$$

The phase angle for the describing function can be calculated as

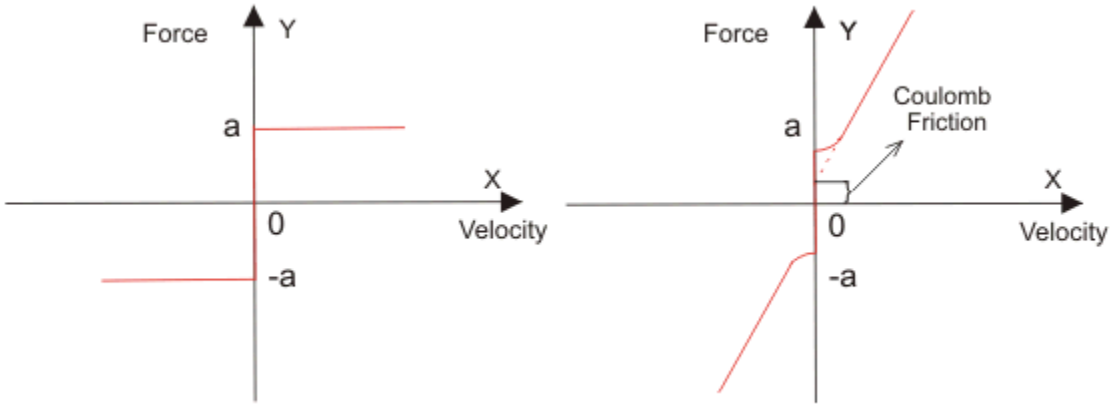
$$\angle\tan^{-1}\left(\frac{A_1}{B_1}\right) = \angle 0^\circ$$

Thus the describing function for saturation is

$$N = \frac{B_1}{X}\angle 0^\circ$$
$$= \frac{2K}{\pi}\left[\sin^{-1}\left(\frac{s}{X}\right) + \frac{s}{X}\sqrt{1-\left(\frac{s}{X}\right)^2}\right]\angle 0^\circ$$

Friction Non linearity:-

Anything which opposes the relative motion of the body is called friction. It is a kind of non linearity present in the system. The common example in an electric motor in which we find coulomb friction drag due to the rubbing contact between the brushes and the commutator.



Friction may be of three types and they are written below:

1. Static Friction : In simple words, the static friction acts on the body when the body is at rest.
2. Dynamic Friction : Dynamic friction acts on the body when there is a relative motion between the surface and the body.
3. Limiting Friction : It is defined as the maximum value of limiting friction that acts on the body when it is at rest.

Dynamic friction can also be classified as (a) Sliding friction (b) Rolling friction. Sliding friction acts when two bodies slides over each other while rolling acts when the bodies rolls over another body.
In mechanical system we have two types of friction namely (a) Viscous friction (b) Static friction.

Liapunov Function and 2nd method:-

In the theory of ordinary differential equations (ODEs), **Lyapunov functions** are scalar functions that may be used to prove the stability of an equilibrium of an ODE. Named after the Russian mathematician AleksandrMikhailovichLyapunov, Lyapunov functions (also called the Lyapunov's second method for stability) are important to stability theory of dynamical systemsand control theory. A similar concept appears in the theory of general state space Markov chains, usually under the name Foster–Lyapunov functions.

For certain classes of ODEs, the existence of Lyapunov functions is a necessary and sufficient condition for stability. Whereas there is no general technique for constructing Lyapunov functions for ODEs, in many specific cases the construction of Lyapunov functions is known. For instance, quadratic functions suffice for systems with one state; the solution of a particular linear matrix inequality provides Lyapunov functions for linear systems; and conservation laws can often be used to construct Lyapunov functions for physical systems

Various types of stability may be discussed for the solutions of differential equations or difference equations describing dynamical systems. The most important type is that concerning the stability of solutions near to a point of equilibrium. This may be discussed by the theory of **AleksandrLyapunov.** In simple terms, if the solutions that start out near an equilibrium point stay near forever, then is **Lyapunov stable**. More strongly, if is Lyapunov stable and all solutions that start out near converge to then is asymptotically stable. The notion of **exponential stability** guarantees a minimal rate of decay, i.e., an estimate of how quickly the solutions converge. The idea of Lyapunov stability can be extended to infinite-dimensional manifolds, where it is known as structural stability, which concerns the behavior of different but "nearby" solutions to differential equations. Input-to-state stability (ISS) applies Lyapunov notions to systems with inputs.

Lyapunov, in his original 1892 work, proposed two methods for demonstrating stability. The first method developed the solution in a series which was then proved convergent within limits. The second method, which is now referred to as the Lyapunov stability criterion or the Direct Method, makes use of a *Lyapunov function V(x)* which has an analogy to the potential function of classical dynamics. It is introduced as follows for a system, x. = f(x)having a point of equilibrium at x=0, then

V(x)=0, if and only if x=0

V(x)>0, if and only if x is not equal to 0

Then *V(x)* is called a Lyapunov function and the system is stable in the sense of Lyapunov (Note that is required; otherwise for example would "prove" that is locally stable). An additional condition called "properness" or "radial unboundedness" is required in order to conclude global stability. Global asymptotic stability (GAS) follows similarly.

It is easier to visualize this method of analysis by thinking of a physical system (e.g. vibrating spring and mass) and considering the energy of such a system. If the system loses energy over time and the energy is never restored then eventually the system must grind to a stop and reach some final resting state. This final state is called the attractor. However, finding a function that gives the precise energy of a physical system can be difficult, and for abstract mathematical systems, economic systems or biological systems, the concept of energy may not be applicable.

Lyapunov's realization was that stability can be proven without requiring knowledge of the true physical energy, provided a Lyapunov function can be found to satisfy the above constraints.

.

UNIT 4

Optimal Control System:- **Optimal control theory** is a branch of applied mathematics that deals with finding a control law for a dynamical system over a period of time such that an objective function is optimized. It has numerous applications in both science and engineering. For example, the dynamical system might be a spacecraft with controls corresponding to rocket thrusters, and the objective might be to reach the moon with minimum fuel expenditure. Or the dynamical system could be a nation's economy, with the objective to minimize unemployment, the controls in this case could be fiscal and monetary policy.

Optimal control is an extension of the calculus of variations, and is a mathematical optimization method for deriving control policies. The method is largely due to the work of Lev Pontryagin and Richard Bellman in the 1950s, after contributions to calculus of variations by Edward J. McShane. Optimal control can be seen as a control strategy in control theory.

Optimal control deals with the problem of finding a control law for a given system such that a certain optimality criterion is achieved. A control problem includes a cost functional that is a function of state and control variables. An **optimal control** is a set of differential equations describing the paths of the control variables that minimize the cost function. The optimal control can be derived using Pontryagin's maximum principle (a necessary condition also known as Pontryagin's minimum principle or simply Pontryagin's Principle), or by solving the Hamilton–Jacobi–Bellman equation (a sufficient condition).

Variation Calculas:-

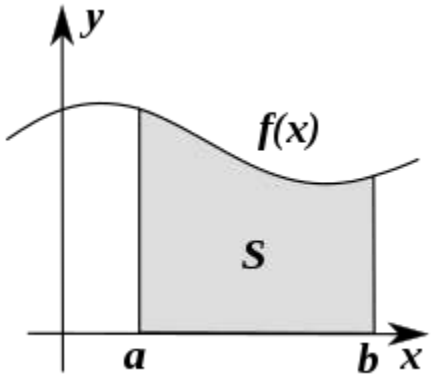
The **calculus of variations** is a field of mathematical analysis that uses variations, which are small changes in functions and functionals, to find maxima and minima of functionalsmappings from a set of functions to the real numbers. Functionals are often expressed as definite integrals involving functions and their derivatives. Functions that maximize or minimize functionals may be found using the Euler–Lagrange equation of the calculus of variations.

A simple example of such a problem is to find the curve of shortest length connecting two points. If there are no constraints, the solution is a straight line between the points. However, if the curve is constrained to lie on a surface in space, then the solution is less obvious, and possibly many solutions may exist. Such solutions are known as geodesics. A related problem is posed by Fermat's principle: light follows the path of shortest optical length connecting two points, where the optical length depends upon the material of the medium. One corresponding concept in mechanics is the principle of least/stationary action.

Many important problems involve functions of several variables. Solutions of boundary value problems for the Laplace equation satisfy the Dirichlet principle. Plateau's problem requires finding a surface of minimal area that spans a given contour in space, a solution can often be found by dipping a frame in a solution of soap suds. Although such experiments are relatively easy to perform, their mathematical interpretation is far from simple: there may be more than one locally minimizing surface, and they may have non-trivial topology.

Functionals of a single function:-

The arc length functional has as its domain the vector space of rectifiable curves and outputs a real scalar. This is an example of a non-linear functional.



The Riemann integral is a linear functional on the vector space of Riemann-integrable functions from a to b. .

This article is mainly concerned with the second concept, which arose in the early 18th century as part of the calculus of variations. The first concept, which is more modern and abstract, is discussed in detail in a separate article, under the name linear form. The third concept is detailed in the article on higher-order functions.

Eular Lagrange Equation:-

In the calculus of variations, the **Euler equation**is a second-order partial differential equation whose solutions are the functions for which a given functional is stationary. It was developed by Swiss mathematician Leonhard Euler and French mathematician Joseph-Louis Lagrange in the 1750s.

Because a differentiable functional is stationary at its local extrema, the Euler–Lagrange equation is useful for solving optimization problems in which, given some functional, one seeks the function minimizing or maximizing it. This is analogous to Fermat's theorem in calculus, stating that at any point where a differentiable function attains a local extremum its derivative is zero.

In Lagrangian mechanics, according to Hamilton's principle of stationary action, the evolution of a physical system is described by the solutions to the Euler equation for the action of the system. In this context Euler equations are usually called **Lagrange equations**. In classical mechanics, it is equivalent to Newton's laws of motion, but it has the advantage that it takes the same form in any system

of generalized coordinates, and it is better suited to generalizations. In classical field theory there is an analogous equation to calculate the dynamics of a field.

The Euler–Lagrange equation was developed in the 1750s by Euler and Lagrange in connection with their studies of the tautochrone problem. This is the problem of determining a curve on which a weighted particle will fall to a fixed point in a fixed amount of time, independent of the starting point.

Lagrange solved this problem in 1755 and sent the solution to Euler. Both further developed Lagrange's method and applied it to mechanics, which led to the formulation of Lagrangian mechanics. Their correspondence ultimately led to the calculus of variations, a term coined by Euler himself in 1766.

The Euler–Lagrange equation is an equation satisfied by a function ***q*** of a real argument *t*, which is a stationary point of the functional.

L(t,q(t), dq/dt)-dL/dt(t,q(t), dq/dt) = 0, for i=1……n

$$l = \int_A^B (dx^2 + dy^2)^{1/2} = \int_a^b [1 + y'^2(x)]^{1/2}\, dx,$$

$$y' \equiv dy/dx$$

.

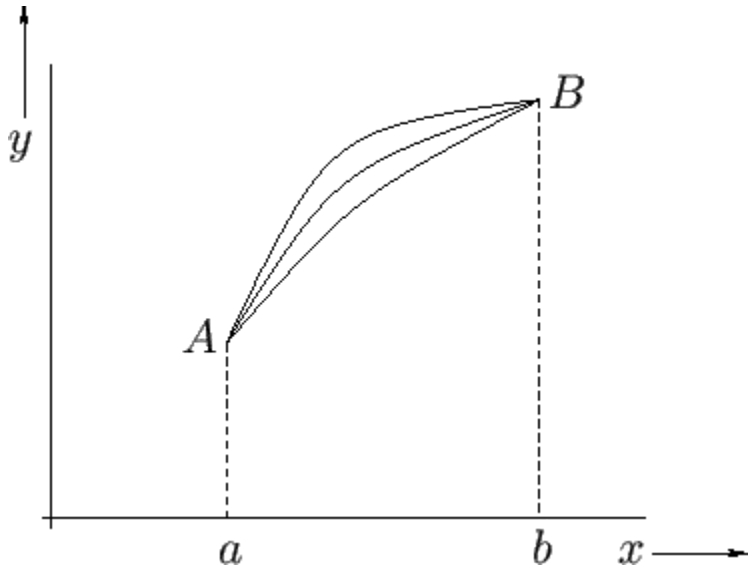


Figure: Different paths between points A and B .

In order to find the shortest path between points A and B , we need to minimize the functional l with respect to small variations in the function $y(x)$, subject to the constraint that the end points, A and B , remain fixed.

Consider a general functional of the form

$$\int_a^b \left(\frac{\partial F}{\partial y}\,\delta y + \frac{\partial F}{\partial y'}\,\delta y' \right)\,dx = 0.\, \delta I = \int_a^b \left(\frac{\partial F}{\partial y}\,\delta y + \frac{\partial F}{\partial y'}\,\delta y' \right) dx,$$

$$I = \int_a^b F(y, y', x)\,dx,$$

Integrating the second term in the integrand of the previous equation by parts, we get

$$\int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'} \right) \right] \delta y\,dx + \left[\frac{\partial F}{\partial y'}\,\delta y \right]_a^b = 0.$$

$$\int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'} \right) \right] \delta y\,dx = 0.$$

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y} = 0.$$

$$\frac{\partial F}{\partial y'} = \text{const.}$$

$$y' \frac{d}{dx}\left(\frac{\partial F}{\partial y'} \right) - y' \frac{\partial F}{\partial y} = 0.$$

$$\frac{d}{dx}\left(y' \frac{\partial F}{\partial y'} \right) = y' \frac{d}{dx}\left(\frac{\partial F}{\partial y'} \right) + y'' \frac{\partial F}{\partial y'}.$$

$$\frac{d}{dx}\left(y' \frac{\partial F}{\partial y'} \right) = y' \frac{\partial F}{\partial y} + y'' \frac{\partial F}{\partial y'}.$$

$$\frac{d}{dx}\left(y' \frac{\partial F}{\partial y'} \right) = \frac{dF}{dx},$$

$$y' \frac{\partial F}{\partial y'} - F = \text{const.}$$

$$\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1+y'^2}} = c,$$

$$y' = \frac{c}{\sqrt{1-c^2}} = \text{const.}$$

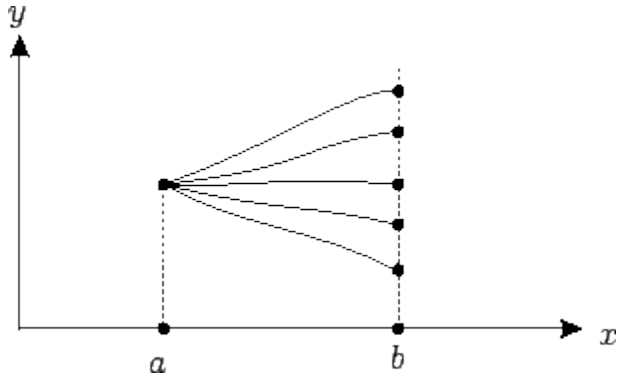
$$y' = \text{constant}$$

Of course, is the equation of a straight-line. Thus, the shortest distance between two fixed points in a plane is indeed a straight-line.

Variable End point problem:-

So far we have been considering the Basic Calculus of Variations Problem, in which the curves have both their endpoints fixed by the boundary conditions . Accordingly, the class of admissible perturbations is restricted to those vanishing at the endpoints. This fact, reflected in , was explicitly used in the derivation of the Euler-Lagrange equation . Indeed, the first-order necessary condition which serves as the basis for the Euler-Lagrange equation--need only hold for admissible perturbations.

If we change the boundary conditions for the curves of interest, then the class of admissible perturbations will also change, and in general the necessary condition for optimality will be different.



$$L_z(b, y(b), y'(b))\eta(b) = 0$$

$$\delta J|_y(\eta) = \int_a^b \Big(L_y(x, y(x), y'(x)) - \frac{d}{dx} L_z(x, y(x), y'(x)) \Big) \eta(x) dx + L_z(b, y(b), y'(b))$$

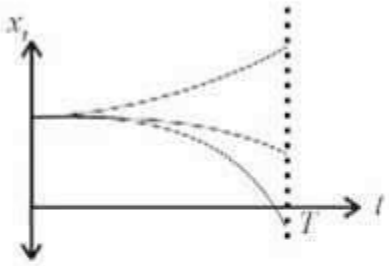
$$L_z(b, y(b), y'(b)) = 0.$$

We can think of as replacing the boundary condition $y(b) = y_1$. Recall that we want to have two boundary conditions to uniquely specify an extremal. Comparing with the Basic Calculus of Variations Problem, here we have only one endpoint fixed a priori, but on the other hand we have a richer perturbation family which allows us to obtain one extra condition.

Transversality Conditions:-

Transversality condition. In optimal control theory, a **transversality condition** is a boundary **condition** for the terminal values of the costate variables. They are one of the necessary **conditions** for optimality in finite-horizon optimal control problems without an endpoint constraint on the state variable.

4-2



We will now confirm this intuition by deriving the transversality condition for this particular problem and at the same time giving a more formal presentation of Pontryagin's maximum principle.

The objective function is

$V = \int_0^T F(t,x,z)\ dt$

now, setting up an equation as a Lagrangian with the state equation constraint, we have

$L = \int_0^T \left[F(t,x,z) + \lambda_t \left(f(t,x,z) - \dot{x}_t \right) \right] dt$

We put the constraint inside the integral because it must hold at every point in time. Note that the shadow price variable, λ_t is actually not a single variable, but is instead defined at every point in time in the interval 0 to T. Since the state equation must be satisfied at each point in

A necessary condition for optimality in variational problems with variable end-points. The arbitrary constants on which the solution of the Euler equation depends are determined by means of the tranversality condition. The transversality condition is a necessary condition for the vanishing of the first variation of a functional.

For the simplest problem in variational calculus with variable end-points.

$$J(x) = \int_{t_1}^{t_2} F(t, x, \dot{x})\, dt,$$

in which the point

$$(t_1, x(t_1), t_2, x(t_2)) = (t_1, x_1, t_2, x_2)$$

is not fixed but can belong to a certain manifold, the transversality condition can be written in the form of the relation

$$[(F - \dot{x}F_{\dot{x}})\, dt + F_{\dot{x}}\, dx]_1^2 = 0,$$

which must be satisfied for any values of the tangent differentials dt_1, dx_1, dt_2, dx_2 of the boundary condition manifold.

If the left- and right-hand end-points of the extremal can be displaced along prescribed curves $x = \phi_1(t)$ and $x = \phi_2(t)$, then since

$$dx_1 = \dot{\phi}_1(t)\, dt_1, \qquad dx_2 = \dot{\phi}_2(t)\, dt_2$$

and the variations of dt_1 and dt_2 are independent, (1) implies

$$\left.\begin{aligned} F(t_1, x_1, \dot{x}_1) - [\dot{\phi}_1(t_1) - \dot{x}_1]\, F_{\dot{x}}(t_1, x_1, \dot{x}_1) = 0, \\ F(t_2, x_2, \dot{x}_2) + [\dot{\phi}_2(t_2) - \dot{x}_2]\, F_{\dot{x}}(t_2, x_2, \dot{x}_2) = 0\,. \end{aligned}\right\}$$

If the equations of the curves along which the left- and right-hand end-points are displaced are given in implicit form, $\omega_1(t, x) = 0$ and $\omega_2(t, x) = 0$, then the transversality condition (1) can be written in the form

$$\left.\begin{aligned} \frac{F - \dot{x}F_{\dot{x}}}{\omega_{1t}} = \frac{F_{\dot{x}}}{\omega_{1x}} \qquad & \text{at the left-hand end-point,} \\ \frac{F - \dot{x}F_{\dot{x}}}{\omega_{2t}} = \frac{F_{\dot{x}}}{\omega_{2x}} \qquad & \text{at the right-hand end-point.} \end{aligned}\right\}$$

If there are no constraints on one of the end-points, then at this end-point, by virtue of the independence of the respective tangent differentials dt and dx, the transversality condition takes the form

$$F = 0, \qquad F_{\dot{x}} = 0\,.$$

The above relations are called Transversality conditions.

Limitations of calculas of variation:-

1. To find minimizers.
2. Necessary conditions which have to satisfy minimizers.
3. Find solutions (externals) which satisfy the necessary conditions.
4. Sufficients conditions which guarantee that such solutions are minimizers.
5. Qualitative properties of minimizers, like regularities properties.
6. How depend minimizers on parameters.
7. Dtability of extremals depending on parameters.

Pontriyagin'sMinimun Principle:-

Pontriyagin's minimum principle is used in optimal control theory to find the best possible control for taking a dynamical system from one state to another, specially in the presence of constraints for the state or input controls. It states that it is necessary for any optimal control along with the optimal state trajectory to solve the so called Hamitonian system, which is a two point boundry value problem, plus a maximum condition of the Hamitonian. These necessary conditions becomes sufficient under certain convexity conditions on the objective and constraints functions.

The minimum principle was formulated by the Russian Mathematician Lev Pontryagin and his students, and its initial application was to maximization of the terminal speed of rocket. The result was derived using ideas from the classical calculus of variation.

ECE 680 FALL 2013

Pontryagin's Minimum Principle[1]

In this handout, we provide a derivation of the minimum principle of Pontryagin, which is a generalization of the Euler-Lagrange equations that also includes problems with constraints on the control inputs. Only a special case of the minimum principle is stated. However, this special case covers a large class of control problems. We consider the problem of minimizing the performance index given by

$$J = \Phi(x(t_F)) + \int_{T_0}^{T_F} F(x(t), u(t))dt \tag{1}$$

subject to

$$\dot{x} = f(x, u), \quad x(t_0) = x_0, \quad \text{and} \quad x(t_F) = x_F . \tag{2}$$

We consider two cases: fixed final state and free final state.

To proceed we define the Hamiltonian function $H(x, u, p)$ as

$$H(x, u, p) = H = F + p^T f ,$$

where the costate vector p will be determined in the analysis to follow.

We discuss the case when the final time t_F is fixed. We follow the development of Luenberger [1, Section 11.1]. We "adjoin" to J additional terms that sum up to zero. We note that because the state trajectories must satisfy the equation, $\dot{x} = f(x, u)$, we have

$$f(x, u) - \dot{x} = 0. \tag{3}$$

We introduce the modified objective performance index,

$$\tilde{J} = J + \int_{T_0}^{T_F} p(t)^T (f(x, u) - \dot{x})\, dt. \tag{4}$$

[1] handout prepared by S. H. Zak

Because for any state trajectory, (3) is satisfied, for any choice of $p(t)$ the value of $\tilde{J}$ is the same as that of J. Therefore, we can express J as

$$\begin{aligned}\tilde{J} &= \Phi(x(t_F)) + \int_{t_0}^{t_F} F(x(t), u(t))\,dt + \int_{t_0}^{t_F} p(t)^T (f(x,u) - \dot{x})\,dt \\ &= \Phi(x(t_F)) + \int_{t_0}^{t_F} (H(x,u,p) - p^T\dot{x})\,dt. \qquad (5)\end{aligned}$$

Let $u(t)$ be a nominal control strategy; it determines a corresponding state trajectory $x(t)$. If we apply another control strategy, say $v(t)$, that is "close" to $u(t)$, then $v(t)$ will produce a state trajectory close to the nominal trajectory. This new state trajectory is just a perturbed version of $x(t)$ and it can be represented as

$$x(t) + \delta x(t).$$

The change in the state trajectory yields a corresponding change in the modified performance index. We represent this change as $\delta\tilde{J}$; it has the form,

$$\delta\tilde{J} = \Phi(x(t_F) + \delta x(t_F)) - \Phi(x(t_F)) + \int_{t_0}^{t_F} (H(x+\delta x, v, p) - H(x,u,p) - p^T\delta\dot{x})\,dt. \qquad (6)$$

We evaluate the integral above. Integrating by parts gives

$$\int_{t_0}^{t_F} p^T\delta\dot{x}\,dt = p(t_F)^T\delta x(t_F) - p(t_0)^T\delta x(t_0) - \int_{t_0}^{t_F} \dot{p}^T\delta x\,dt. \qquad (7)$$

Note that $\delta x(t_0) = 0$ because a change in the control strategy does not change the initial state. Taking into account the above, we represent (6) as

$$\begin{aligned}\delta\tilde{J} &= \Phi(x(t_F) + \delta x(t_F)) - \Phi(x(t_F)) - p(t_F)^T\delta x(t_F) \\ &\quad + \int_{t_0}^{t_F} (H(x+\delta x, v, p) - H(x,u,p) + \dot{p}^T\delta x)\,dt.\end{aligned}$$

We next replace $(\Phi(x(t_F) + \delta x(t_F)) - \Phi(x(t_F)))$ with its first-order approximation, and add and subtract the term $H(x, v, p)$ under the integral to obtain

$$\begin{aligned}\delta\tilde{J} &= \nabla_x\Phi|_{T=T_F} - p(t_F)^T\delta x(t_F) \\ &\quad + \int_{t_0}^{t_F} (H(x+\delta x, v, p) - H(x,v,p) + H(x,v,p) - H(x,u,p) + \dot{p}^T\delta x)\,dt.\end{aligned}$$

Replacing $(H(x + \delta x, v, p) - H(x, v, p))$ with its first-order approximation,

$$H(x + \delta x, v, p) - H(x, v, p) = \frac{\partial H}{\partial x}\,\delta x,$$

gives

$$\tilde{o}J = \nabla x\Phi|_{T=T_F} - p(t_F)^T\,\delta x(t_F) + \int_{T_0}^{T_F} \left(\frac{\partial H}{\partial x} + \dot{p}^T\right)\delta x + H(x, v, p) - H(x, u, p)\,dt.$$

Selecting in the above p as the solution to the differential equation

$$\frac{\partial H}{\partial x} + \dot{p}^T = 0^T$$

with the final condition

$$p(t_F) = \nabla x\Phi|_{T=T_F},$$

reduces $\tilde{\delta}J$ to

$$\tilde{\delta}J = \int_{T_0}^{T_F} (H(x, v, p) - H(x, u, p))\,dt.$$

If the original control $u(t)$ is optimal, then we should have $\tilde{\delta}J \geq 0$, that is, $H(x, v, p) \geq H(x, u, p)$. We summarize our development in the following theorem.

Theorem 1 Necessary conditions for $u \in U$ to minimize (1) subject to (2) are:

$$\dot{p} = -\left(\frac{\partial H}{\partial x}\right)^T,$$

where $H = H(x, u, p) = F(x, u) + p^T f(x, u)$,

$$H(x^*, u^*, p^*) = \min_{u \in U} H(x, u, p).$$

If the final state, $x(t_F)$, is free, then in addition to the above conditions it is required that the following end-point condition is satisfied,

$$p(t_F) = \nabla x\Phi|_{T=T_F}.$$

The equation,

$$\dot{p} = -\left(\frac{\partial H}{\partial x}\right)^T \tag{8}$$

is called in the literature the adjoint or costate equation.

We illustrate the above theorem with the following well-known example.

Example 1 We consider a controlled object modeled by

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

with the performance index

$$J = \int_0^{t_F} dt.$$

The control is required to satisfy

$$|u(t)| \leq 1$$

for all $t \in [0, t_F]$. This constraint means that the control must have magnitude no greater than 1. Our objective is to find admissible control that minimizes J and transfers the system from a given initial state x_0 to the origin.

We begin by finding the Hamiltonian function for the problem,

$$\begin{aligned} H &= 1 + p^T(Ax + bu) \\ &= 1 + \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ &= 1 + p_1x_2 + p_2u. \end{aligned}$$

The costate equations are

$$\begin{bmatrix} \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = -\begin{bmatrix} \frac{\partial H}{\partial x_1} \\ \frac{\partial H}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 0 \\ -p_1 \end{bmatrix}.$$

Solving the costate equations yields

$$p_1 = d_1 \quad \text{and} \quad p_2 = -d_1t + d_2,$$

where d_1 and d_2 are integration constants. We now find an admissible control minimizing the Hamiltonian,

$$\begin{aligned} \arg_U \min H &= \arg_U \min(1 + p_1 x_2 + p_2 u) \\ &= \arg_U \min(p_2 u) \end{aligned}$$

$$u t \quad \text{if} \quad p_2 <$$

$$u(t) = -1 \quad \text{if } p_2 > 0.$$

Hence,

$$u^*(t) = -\text{sign}(p^*_2) = -\text{sign}(-d_1 t + d_2),$$

where the signum function is defined as

$$\text{sign}(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{if } z = 0 \\ -1 & \text{if } z < 0. \end{cases}$$

Thus, the optimal control law is piecewise constant taking the values 1 or -1. This control law has at most two intervals of constancy because the argument of the sign function is a linear function, $-d_1 t + d_2$, that changes its sign at most once. This type of control is called a bang-bang control because it switches back and forth between its extreme values. System trajectories for $u = 1$ and $u = -1$ are families of parabolas. Segments of the two parabolas through the origin form the switching curve,

$$x_1 = -\frac{1}{2} x^2_2 \, \text{sign}(x_2).$$

This means that if an initial state is above the switching curve, then $u = -1$ is used until the switching curve is reached. Then, $u = 1$ is used to reach the origin. For an initial state below the switching curve, the control $u = 1$ is used first to reach the switching curve, and then the control is switched to $u = -1$. We implement the above control action as

$$u = -\text{sign}(v) = \begin{cases} 1 & \text{if } v < 0 \\ 0 & \text{if } v = 0 \\ -1 & \text{if } v > 0, \end{cases}$$

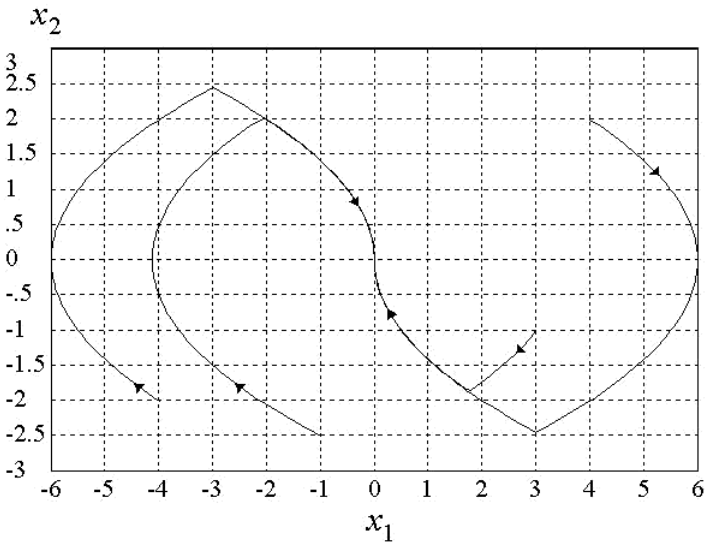


Figure 1: A phase-plane portrait of the time optimal closed-loop system of Example 1.

where $v = v(x_1, x_2) = x_1 + \frac{1}{2} x_2^2 \text{sign}(x_2) = 0$ is the equation describing the switching curve. We can use the above equation to synthesize a closed-loop system such that starting at an arbitrary initial state in the state plane, the trajectory will always be moving in an optimal fashion towards the origin. Once the origin is reached, the trajectory will stay there. A phase portrait of the closed-loop time optimal system is given in Figure 1.

References

[1] D. G. Luenberger. Introduction to Dynamic Systems: Theory, Models, and Applications. John Wiley & Sons, New York, 1979.