DIFFERENTIAL EQUATIONS

M.A. (Previous)

Directorate of Distance Education
Maharshi Dayanand University
ROHTAK – 124 001

Maharshi Dayanand University
ROHTAK – 124 001

Developed & Produced by EXCEL BOOKS PVT. LTD., A-45 Naraina, Phase 1, New Delhi-110028

Contents

M.A. (Previous)
DIFFERENTIAL EQUATIONS

Paper-V **M.Marks: 100**
Time: 3 Hrs.

Note: *Question paper will consist of three sections. Section-I consisting of one question with ten parts of 2 marks each covering whole of the syllabus shall be compulsory. From Section-II, 10 questions to be set selecting two questions from each unit. The candidate will be required to attempt any seven questions each of the five marks. Section-III, five questions to be set, one from each unit. The candidate will be required to attempt any three questions each of fifteen marks.*

Unit-I

Preliminaries: Initial value problem and the equivalent integral equation, mth order equation in d-dimensions as a first order system, concepts of local existence, existence in the large and uniquness of solutions with examples.

Basic Theorems: Ascoli-Arzela Theorem, A theorem on convergene of solutions of a family of initial value problems.

Picard-Lindel of theorem: Peano's existence theorem and corollary. Maximal intervals of existence. Extension theorem and corollaries. Kamke's convergence theorem. Kneser' theorem (statement only).

Unit-II

Dependence on initial conditions and parameters: Preliminaries, Continuity, Differentiability, Higher Order Differentiability.

Differential Inequalities and Uniqueness: Gronwall's inequality, Maximal and Minimal solutions. Differential inequalities. A theorem of Wintner, Uniqueness Theorems, Nagumo's and Osgood's criteria.

Egres points and Lyapunov functions. Successive approximations.

Unit-III

Linear Differential Equations: Linear Systems, Variation of constants, reduction to smaller systems. Basic inequalities, constant coefficients, Floquet theory, Adjoint systems, Higher order equations.

Unit-IV

Poincare-Bendixson Theory: Autonomous sytems, Umlanfsatz, Index of stationary point. Poincare-Bendixson theorem. Stability of periodic solutions, rotation points, foci, nodes and saddle points.

Use of Implicit function and fixed point theorems: Period solutions, Linear equations, Non-linear problems.

Second order Boundary value problems – Linear Problems, Non-linear problems, Aprori bounds.

Unit-V

Linear second order equations: Preliminaries, Basic facts, Theorems of Sturm. Sturm-Liouville Boundary Value Problems, Number of zeros, Non-oscillatory equations and principal solutions. Non-oscillation theorems.

1

DIFFERENTIAL AND INTEGRAL EQUATIONS

The subject of differential equations is large, diverse, powerful, useful, and full of surprises. Differential equations can be studied on their own-just because they are intrinsically interesting. Or, they may be studied by a physicist, engineer, biologist, economist, physician, or political scientist because they can model (quantitatively explain) many physical or abstract systems. Just what is a differential equation? A differential equation having y as the dependent variable (unknown function) and t as the independent variable has the form

$$F\left(t, y\frac{dy}{dt}, \frac{d^2y}{dt^2}, ..., \frac{d^ny}{dt^n}\right) = 0$$

for some positive integer n. (If n is 0, the equation is an algebraic or transcendental equation, rather than a differential equation). Here is the same idea in words :

Definition. A **differential equation** is an equation that relates in a non-trivial manner an unknown function and one or more of the derivatives or differentials of that unknown function with respect to one or more independent variables.

The phrase "in a nontrivial manner" is added because some equations that appear to satisfy the above definition are really identities. That is, they are always true, no matter what the unknown function might be. An example of such an equation is :

$$\sin^2\left(\frac{dy}{dt}\right)\cos^2\left(\frac{dy}{dt}\right) = 1.$$

This equation is satisfied by every differential function of one variable. Another example is :

$$\left(\frac{dy}{dt} - y\right)^2 = \left(\frac{dy}{dt}\right)^2 - 2y\left(\frac{dy}{dt}\right) + y^2.$$

Classification of Differential Equations

Differential equations are classified in several different ways : **ordinary** or **partial; linear** or **nonlinear**. There are even special subclassifications: **homogeneous** or **nonhomogeneous; autonomous** or **nonautonomous; first-order, second-order,….,nth order.** Most of these names for the various types have been inherited from other areas of mathematics, so there is some ambiguity in the meanings. But the context of any discussion will make clear what a given name means in that context. There are reasons for these classifications, the primary one being to enable discussions about differential equations to focus on the subject matter in a clear and unambiguous manner. Our attention will be on ordinary differential equations. Some will be linear, some nonlinear. Some will be first-order, some second-order, and some of higher order than second. What is the order of a differential equation? As a rule, only those differential equations are considered which are algebraic in the differential coefficients.

Definition. The **order** of a differential equation is the order of the highest derivative that appears (nontrivially) in the equation.

At this early stage in our studies, we need only be able to distinguish ordinary from partial differential equations. This is easy: a differential equation is an **ordinary differential equation** if the only derivatives of the unknown function(s) are ordinary derivatives, and a differential equation is a **partial differential equation** if the only derivatives of the unknown function (s) are partial derivatives.

Example. Here are some ordinary differential equations :

$\frac{dy}{dt} = 1+y^2$ (first-order) [nonlinear]

$\frac{d^2y}{dt^2} + y = 3\cos t$ (second-order) [linear, nonhomogeneous]

$\frac{d^3y}{dt^3} + 3\frac{d^2y}{dt^2} - 5y = 0$ (third-order) [linear, homogeneous]

Example. Here are some partial differential equations :

$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y}$ (first-order in x and y)

$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$ (first-order in t; second-order in x)

$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ (second-order in x and y)

$\frac{\partial^2 u}{\partial x \partial y} = 3$ (second-order)

Linearity. We now introduce the important concept of **linearity** applied to such equations. This concept will help us to classify these equations still further.

Definition. An **ordinary differential** equation of order n, in the dependent variable y and the independent variable t, is said to be a linear equation which can be expressed in the form

$$a_0(t)\frac{d^n y}{dt^n} + a_1(t)\frac{d^{n-1}y}{dt^{n-1}} + \ldots + a_{n-1}(t)\frac{dy}{dt} + a_n(t)y = Q(t) \qquad (*)$$

where $a_0(t)$ is not identically zero on [a.b].

The right handed member Q(t) of (*) is called the **nonhomogeneous** term. If Q(t) is identically zero, equation (*) reduces to

$$a_0(t)\frac{d^n y}{dt^n} + a_1(t)\frac{d^{n-1}y}{dt^{n-1}} + \ldots a_n(t)y = 0 \qquad (**)$$

and is then called **homogeneous.** Thus a linear homogeneous differential equation of order n does not contain a term involving the independent variable alone.

Examples. The ordinary differential equations

(i) $\frac{d^2y}{dt^2} + 7\frac{dy}{dt} + 6y = 0$,

(ii) $t^7 \frac{d^3y}{dt^2} + 6t\frac{dy}{dt} + y = \sin t$,

are both linear.

Definition. An ordinary differential equation which is not linear is called a **nonlinear** ordinary differential equation.

Example. The following ordinary homogeneous **differential equations** are all nonlinear.

(i) $\frac{d^2y}{dt^2} + 7\frac{dy}{dt} + 6y^3 = 0$,

(ii) $\frac{d^2y}{dt^2} + 4\left(\frac{dy}{dt}\right)^3 + 6y = 0,$

(iii) $\frac{d^2y}{dt^2} + 6y\frac{dy}{dt} + 6y = 0$.

Definition. To say that y = g(t) is a solution of differential equation

$$F\left(t, y, \frac{dy}{dt}, ..., \frac{d^n y}{dt^n}\right) = 0$$

on an interval I means that

$$F(t, g(t), g'(t), \ldots, g^{(n)}(t)) = 0 ,$$

for every choice of t in the interval I. In other words, a solution, when substituted into the differential equation, makes the equation identically true for t in I.

Initial-value problem. An initial-value problem associated with a first order differential equation is of the form

$$\frac{dy}{dt} = f(t,y), \quad t \in I$$

$$y(t_0) = y_0 ,$$

for some point $t_0 \in I$.

An initial-value problem associated with a second order differential equation has the form

$$\frac{d^2y}{dt^2} = f\left(t, y, \frac{dy}{dt}\right), \quad t \in I$$

with initial conditions

$$y(t_0) = y_0 ,$$

$$y'(t_0) = \xi_0 ,$$

for some point $t_0 \in I$.

Integral Equations

An integral equation is an equation in which the unknown function, say u(t), appears under an integral sign. A general example of an integral equation in u(t) is

$$u(t) = f(t) + \int K(t,s)u(s)ds$$

where K(t,s) is a function of two variables called the kernel or nucleus of the integral equation. According to Bocher [1914], the name integral equations was suggested in 1888 by du Bois-Reymond, although the first appearance of integral equations is accredited to Abel for his thesis work on the Tautochrone, which was published in 1823 and 1826. There is also the opinion that such first appearance was in Laplace's work in 1782 as it shall make sense when we speak of the inverse Laplace transform. For example, the Laplace transform of the given (known) function f(t), $0 < t < \infty$, is

$$L\{f(t)\} = F(s) = \int_0^\infty e^{-st}f(t)dt, \quad s > a$$

provided that the integral converges for s > a. So, if we are now given F(s), say $F(s) = \frac{1}{s^2}$, s > 0 , and we are to find the original function (now as unknown) f(t), or the inverse Laplace transform of F(s), i.e., $f(t) = L^{-1}\{F(s)\}$,

$$\frac{1}{s^2} = \int_0^\infty e^{-st} f(t)dt ,$$

then we are against solving above integral equation in (the unknown) f(t). So it does make sense that integral equations started with Laplace, since he was, in the final analysis, after recovering the original function f(t) from knowing F(s). In our above example, f(t) = t.

In the same vein, Fourier in 1820 solved for the inverse f(t) of the following Fourier transform F(λ) of f(t), $-\infty < t < \infty$,

$$\mathcal{F}\{f\} = F(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda t} f(t)\, dt$$

as

$$f(t) = \mathcal{F}^{-1}\{F\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\lambda t} F(\lambda) d\lambda .$$

Hence in finding the (unknown) f(t), he solved an integral equation in f(t). With such an explicit solution f(t), it is not surprising that some historians consider this Fourier (inverse transform) result as the first very clear and reachable solution of an integral equation.

Some problems have their mathematical representation appear directly, and in a very natural way, in terms of integral equations. Other problems, whose direct representation is in terms of differential equations and their auxiliary conditions, may also be reduced to integral equations. Problems of a "hereditary" nature fall under the first category, since the state of the system u(t) at any time t depends by definition on all the previous states u(t−τ) at the previous times t−τ, which means that we must sum over them, hence involve them under the integral sign in an integral equation. We may then say that such problems, among others, have integral equations as their natural mathematical representation. The rest of the examples are problems that are formulated in terms of ordinary or partial differential equations with initial and/or boundary conditions that are reduced to an integral equation or equations. The advantage here is that the auxiliary conditions are automatically satisfied, since they are incorporated in the process of formulating the resulting integral equation. The other advantage of the integral equation form is in the case when both differential equations as well as integral equations forms do not have exact, closed-form solutions in terms of elementary known functions.

CLASSIFICATION OF INTEGRAL EQUATIONS

The most of the integral equations fall under two main categories : Volterra and Fredholm integral equations. A Volterra integral equation for the first kind is of the form

$$-f(x) = \int_a^x K(x,\xi)u(\xi)d\xi ,$$

and a Volterra integral equation of the second kind is of the type

$$u(x) = f(x) = \int_a^x K(x,\xi)u(\xi)d\xi .$$

A Fredholm integral equation of the first and second kind are, respectively,

$$-f(x) = \int_a^b K(x,\xi)u(\xi)d\xi ,$$

$$u(x) = f(x) + \int_a^b K(x,\xi)u(\xi)d\xi .$$

Initial value problems reduced to Volterra Integral Equations

Now, we shall illustrate in detail **how an initial value** problem associated with a linear-differential equation and auxiliary conditions **reduces to a Volterra** integral equation.

Example 1. Consider the initial-value problem

$$\frac{dy}{dt} = f(t, y) , \quad t \varepsilon I \qquad (1)$$

$$y(t_0) = y_0 \qquad (2)$$

Integrating (1) w.r.t. t from t_0 to t, we write

$$y(t) - y(t_0) = \int_0^t f(s, y(s))ds$$

$$\Rightarrow \quad y(t) = y_0 + \int_{t_0}^t f(s, y(s))ds \qquad (3)$$

which is a Volterra integral equation of the second kind. Conversely, the differentiation of (3) gives

$$\frac{dy}{dt} = f(t, y(t)) \text{ for all } t \varepsilon I \qquad (4)$$

Further, from (3), we write, on putting $t = t_0$ both sides,

$$y(t_0) = y_0 . \qquad (5)$$

That is, y(t) given by (4) also satisfies the initial condition in (5).

Example 2. Consider the initial value problem associated with the second-order differential equation

$$\frac{d^2y}{dt^2} = \lambda\, y(t) + g(t) , \qquad (1)$$

$$y(0) = 1,$$

$$y'(0) = 0 . \qquad (2)$$

We integrate equation (1) w.r.t. 't' over the **interval [0, t]**. We obtain

$$\frac{dy}{dt} - y'(0) = \lambda \int_0^t y(\xi)d\xi + \int_0^t g(\xi)d\xi$$

or $$\frac{dy}{dt} = \lambda \int_0^t u(\xi)d\xi + \int_0^t g(\xi)d\xi , \qquad (3)$$

using one initial condition given in (2). Integrating again, we find

$$y(t) - y(0) = \lambda \int_0^t\int_0^\xi y(s)ds + d\xi + \int_0^t\int_0^\xi g(s)ds\, d\xi$$

or $$y(t) = 1 + \int_0^t (t-s)\, g(s)ds + \lambda \int_0^t (t-s)y(s)ds$$

$$\Rightarrow \qquad y(t) = f(t) + \lambda\int_0^t K(t,s)\, y(s)ds, \qquad (4)$$

where $f(t) = 1 + \int_0^t (t-s)g(s)ds$ is the **non-homogeneous** term and $K(t, s) = t-s$ is the kernel of Volterra integral equation (4). Integral equation (4) automatically takes care of two auxiliary conditions in (2).

Now, we shall consider the initial value problem associated with the general second-order differential equation.

Example 3. $$\frac{d^2y}{dt^2} + A(t)\frac{dy}{dt} + B(t)y(t) = g(t), \qquad (1)$$

$$y(a) = c_1,$$
$$y'(a) = c_2\,. \qquad (2)$$

We write

$$\frac{d^2y}{dt^2} = -A(t)\frac{dy}{dt} - B(t)y(t) + g(t)\,.$$

We now integrate over the interval (a, t) to obtain

$$\frac{dy}{dt} - c_2 = -\int_a^t A(\xi)y'(\xi)d\xi - \int_a^t B(\xi)y(\xi)d\xi + \int_a^t g(\xi)d\xi$$

$$= \left[-A(\xi)y(\xi)\right]_a^t + \int_a^t A'(\xi)y(\xi)d\xi - \int_a^t B(\xi)y(\xi)d\xi + \int_a^t g(\xi)d\xi)$$

$$= \int_a^t \left[A'(\xi) - B(\xi)\right]y(\xi)d\xi + \int_a^t g(\xi)d\xi - A(t)y(t) + c_1A(a)\,. \qquad (3)$$

Integrating (3) again, we obtain

$$y(t) - c_1 - c_2(t-a) = \int_a^t (t-s)[A'(s) - B(s)]y(s)ds - \int_a^t A(s)y(s)ds$$
$$+ \int_a^t (t-s)g(s)ds + c_1A(a)[t-a]\,.$$

This implies

$$y(t) = \int_a^t [(t-s)\{A'(s) - B(s)\} - A(s)]y(s)ds + f(t) \qquad (4)$$

where the non-homogeneous term f(t) is

$$f(t) = \int_a^t (t-s)\, g(s)ds + (t-a)\,[c_1(A(a)) + c_2] + c_1\,. \qquad (5)$$

Equation (4) is a Volterra integral equation of the second kind of the type

$$y(t) = f(t) + \int_a^t K(t,s)y(s)ds\ , \qquad (6)$$

in which the kernel K(t, s) is given by

$$K(t,s) = (t-s)\,[A'(s)-B(s)] - A(s)\,. \qquad (7)$$

Integral equation (5) is equivalent to the given initial value problem and it takes care of auxiliary conditions in (2).

Exercise

Obtain the Volterra integral equation corresponding to each of the following initial value problems

(a) $y'' + \lambda y = 0$; $y(0) = 1, y'(0) = 0$

(b) $y'' + \lambda y = 0$; $y(0) = 0, y'(0) = 1$

(c) $y'' + y = \sin t$; $y(0) = 1, y'(0) = 1$

(d) $y'' - y + t = 0$; $y(0) = 1, y'(0) = 0$

(e) $y'' + \lambda y = f(t)$; $y(0) = 1, y'(0) = 0$

(f) $y'' + ty = 1$, $y(0) = y'(0) = 0$

(g) $y'' - 2ty' - 3y = 0$, $y(0) = 1, y'(0) = 0$.

2

EXISTENCE THEOREM

Let I denote an open interval on the real line $-\infty < t < \infty$, that is, the set of all real t satisfying $a < t < b$ for some real constants a and b. The set of all complex-valued functions having k continuous derivatives on I is denoted by $C^k(I)$. If f is a member of this set, one writes $f \varepsilon C^k(I)$, or $f \varepsilon C^k$ on I. The symbol ε is to be read "is a member of" or "belongs to." It is convenient to extend the definition of C^k to intervals I which may not be open. The real intervals $a < t < b$, $a \le t \le b$, $a \le t < b$, and $a < t \le b$ will be denoted by (a, b), [a,b], [a, b), and (a, b], respectively. If $f \varepsilon C^k$ on (a,b), and in addition the right-hand kth derivative of f exists at a and is continuous from the right at a, then f is said to be of class C^k on [a, b). Similarly, if the kth derivative is continuous from the left at b, then $f \varepsilon C^k$ on (a, b]. If both these conditions hold, one says $f \varepsilon C^k$ on [a, b].

A nonempty set S of points of the real (t, y) plane will be called **connected** if any two points of S can be joined by a **continuous curve** which lies entirely in S.

A non-empty set S of points of the ty-plane is called **open** if each point of S is an interior point of S.

An open and connected set in the ty-plane is called a **domain.**

A point P is called a **boundary point** of a domain D if every circle around P contains both points in D and points not in D.

A domain plus its boundary points will be called a **closed domain**.

If D is a domain in the real (t, y) plane, the set of all complex-valued functions f defined on D such that all kth-order partial derivatives $\partial^k f/\partial t^p \partial y^q$ (p+q = k) exist and are continuous on D is denoted by $C^k(D)$, and one writes $f \varepsilon C^k(D)$, or $f \varepsilon C^k$ on D.

The sets $C^0(I)$ and $C^0(D)$, the continuous functions on I and D, will be denoted by C(I) and C(D), respectively.

Let D be a domain in the (t, y) plane and suppose f is a real-valued function such that $f \varepsilon C(D)$. Then the central problem may be phrased as follows:

Problem. To find a differentiable function ϕ defined on a real t interval I such that

(i) $(t, \varphi(t) \varepsilon D \qquad (t \varepsilon I)$

(ii) $\frac{d}{dt}[\varphi(t)] = f(t, \varphi(t)) \qquad (t \varepsilon I)$

This problem is called an ordinary differential equation of the first order, and is denoted by

$$\frac{dy}{dt} = f(t,y) . \qquad \text{(E)}$$

If such an interval I and function φ exist, then φ is called a solution of the differential equation (E) on I. Clearly if φ is a solution of (E) on I, then $\varphi \varepsilon C^1$ on I, on account of (ii).

In geometrical language, (E) prescribes a slope f(t, y) at each point of D. A solution φ on I is a function whose graph [the set of all points (t, φ (t)), t ε I] has the slope f(t, φ(t)) for each t ε I.

Suppose (t_0, y_0) is a given point in D. Then an initial-value problem associated with differential equation (E) and this point (t_0, y_0) is formulated in the following way :

Initial-value Problem. To find an interval I containing t_0 and a solution $\varphi = \varphi(t)$ of differential equation

$$\frac{dy}{dt} = f(t, y) \qquad \text{for } (t, y) \in D$$

on I satisfying the initial condition

$$\varphi(t_0) = y_0 .$$

As seen earlier, this initial-value problem is completely equivalent to the finding of a continuous function $\varphi = \varphi(t)$ on I satisfying the integral equation

$$\varphi(t) = y_0 + \int_{t_0}^{t} f(s, \varphi(s))ds, \quad t \in I .$$

Remark. Given a continuous function f(t, y) on a domain D, the first question to be answered is "Whether there exists a solution of the differential equation

$$\frac{dy}{dt} = f(t, y) \quad \text{for all } t \in I\text{"}$$

The answer is YES, if interval I is properly prescribed.

Example. Consider the differential equation

$$\frac{dy}{dt} = y^2 , \qquad y(1) = -1 \quad \text{with } t_0 = 1 \in I .$$

A solution of this problem is

$$\varphi(t) = -t^{-1} .$$

However, this solution does not exist at t = 0, although

$$f(t,y) = y^2$$

is continuous at t = 0.

This example shows that any general existence theorem will necessarily have to be of LOCAL nature around t_0. Existence in the large can only be asserted under additional conditions on f.

Existence in the large. On what t-ranges does a solution of initial-value problem

$$\frac{dy}{dt} = f(t,y),$$

$$y(t_0) = y_0, \quad t \in I$$

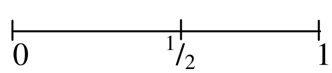
exist?

Let E be a subset of (t,y) space where

$$E = \{(t,y) |\ 0 \le t \le 1, |y| \le 1\} .$$

Consider I V P

$$\frac{dy}{dt} = f(t, y), \ y(0) = 1 .$$



A solution of this IVP may exist for $0 \le t \le {}^1/_2$ and increase from 0 to 1 as t goes from 0 to ${}^1/_2$, then one cannot expect to have an extension of solutions, y(t), for $t > {}^1/_2$.

Uniqueness of solutions

Let y be a scalar and consider the IVP

$$\frac{dy}{dt} = |y|^{1/2} , \ y(0) = 0 .$$

This IVP has more than one solution.

First solution : $y(t) = 0$ for all t.

Second solution: It has one parameter family of solutions defined by

$$y(t) = \begin{cases} 0 & \text{for } t \le c \\ \left(\dfrac{t-c}{2}\right)^2 & \text{for } t \ge c \end{cases}$$

where c is an arbitrary constant with $c \ge 0$. Thus, solution of this IVP is not unique.

Definition ($\in$-approximate solution)

Let f be a real-valued continuous function on a domain D in the (t,y) plane. An $\in$-approximate solution of an ODE of the first order

$$\frac{dy}{dt} = f(t,y)$$

on a t-interval I is a function $\varphi \,\varepsilon\, C(I)$ such that

i) $(t, \varphi(t)) \,\varepsilon\, D$ for all $t \,\varepsilon\, I$

ii) $\varphi \,\varepsilon\, C^1(I)$, except possibly for a finite set of points S on I where $\varphi'(t)$ may have simple discontinuities (i.e., at such points of S, the right and left limits of $\phi'(t)$ exist but are not equal),

iii) $|\varphi'(t) - f(t, \varphi(t))| < \in$ for $t \,\varepsilon\, I - S$.

Remark (1) When $\in = 0$, then it will be understood that the set S is empty, i.e., $S = \phi$. So (ii) holds for all $t \,\varepsilon\, I$.

(2) The statement (ii) implies that φ have a piecewise continuous derivative on I, and we shall denote it by

$$\varphi \,\varepsilon\, C_p^1(I).$$

(3) Consider the rectangle

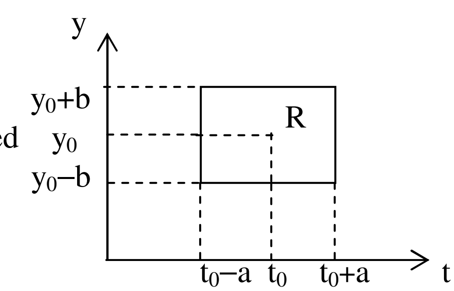
$$R = \{(t, y) : |t-t_0| \le a, |y-y_0| \le b, \ a > 0, b > 0\} \quad (1)$$

about the point (t_0, y_0).

Let $f \,\varepsilon\, C$ on the rectangle R. Since the rectangle R is a closed set, so the continuous function f on R is bounded. Let

$$M = \max |f(t,y)| \text{ on } R \qquad (2)$$

$$\text{Let } \alpha = \min\left(a, \frac{b}{M}\right) \qquad (3)$$



(Cauchy-Euler construction of an approximate solution).

Theorem 2.1. Let $f \,\varepsilon\, C$ on the rectangle R. Given $\in > 0$, there exists an **$\in$-approximate solution** φ of ODE of first order

$$\frac{dy}{dt} = f(t,y) \qquad (1)$$

on the interval $I = \{t : |t - t_0| \le \alpha\}$ such that $\varphi(t_0) = y_0$, α being some constant.

Proof. Let $\in$ 0 be given. We shall construct an $\in$–approximate solution for the interval $[t_0, t_0+ \alpha]$. A similar construction will define it for $[t_0 -\alpha, t_0]$.

This approximate solution will consist of a polygonal path starting at (t_0, y_0), i.e., a finite number of straight-line segments joined end to end.

Since f is continuous on the closed rectangle

$$R = \{(t,y) : |t-t_0| \le a,\ |y-y_0| \le b,\ a > 0,\ b > 0\} \quad (2)$$

So f is bounded and uniform continuous on R. Let

$$M = \max_R |f(t,y)| \quad (3)$$

and

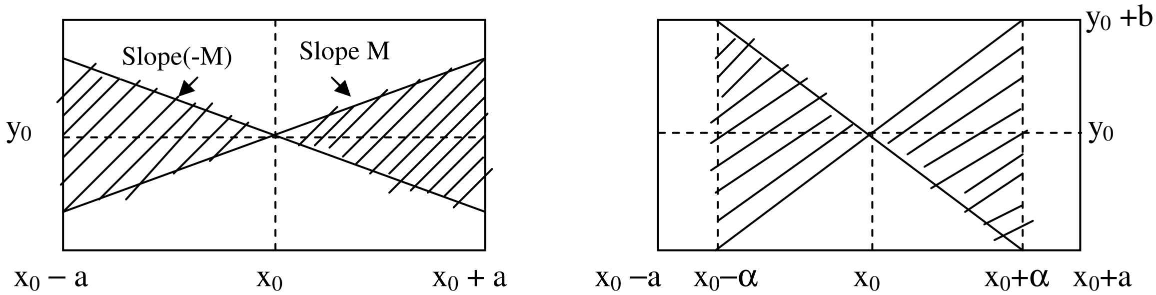
$$\alpha = \min(a, \frac{b}{M}) . \quad (4)$$

Then

(i) $\alpha = a$ if $M \le \frac{b}{a}$ (fig 2.1a)

(ii) $\alpha = \frac{b}{M}$ if $M \ge \frac{b}{a}$ (fig. 2.1b)

In the first case, we get a solution valid in the whole interval $|t - t_0| \le a$, whereas in the second of the interval case, we get a solution valid only on I, a sub-interval f$|t-t_0| \le a$.



(a) (b)

Fig. 2.1

We consider the second case when $M \ge \frac{b}{a}$. Since f is uniformly continuous on R, therefore, for given $\in > 0$, there exists a real number $\delta = \delta_\in = \delta(\in) > 0$ such that

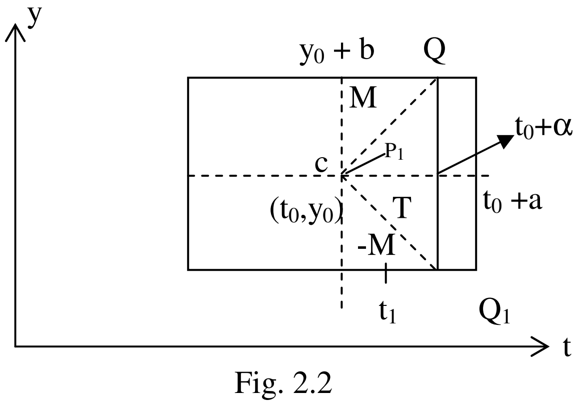


Fig. 2.2

$$|f(t,y) - f(\bar{t}, \bar{y})| \le \in \quad (5)$$

provided

$$|t - \bar{t}| \le \delta_\in ,\ |y - \bar{y}| \le \delta_\in \quad (6)$$

for $(t, y) \varepsilon R$ and $(\bar{t}, \bar{y}) \varepsilon R$.

Now divide the interval $[t_0, t_0 + \alpha]$ into n parts such that

$$t_0 < t_1 < \ldots < t_n = t_0 + \alpha$$

and

$$\max.|t_k - t_{k-1}| \le \min\left(\delta_\in, \frac{\delta_\in}{M}\right) . \qquad (7)$$

Starting from the point C (t_0, y_0), we construct a straight-line segment with slope $f(t_0, y_0)$ proceeding to the right of t_0 until it intersects the line $t = t_1$ at some point $P_1(t_1, y_1)$. Here, slope of line CP_1 is $f(t_0, y_0)$. This line segment, CP_1 , must lie inside the triangular region T bounded by the lines issuing from C with slope M and –M, and the line $t = t_0+\alpha$, as shown in the figure (2.2) above because, we have, in this case,

$$\alpha = \frac{b}{M} . \qquad (8)$$

Now, at the point $P_1(t_1, y_1)$, we construct to the right of t_1 a straight-line segment with slope $f(t_1, y_1)$ upto the intersection with line $t = t_2$, say at the point $P_2(t_2, y_2)$.

Continuing in this fashion, in a finite number of steps, the resultant path $\varphi(t)$ will meet the line $t = t_0 + \alpha$ at the point $P_n(t_n, y_n)$. Further, this polygon path (fig. 2.3) will lie entirely within the region T. This φ is the required $\in$–approximate solution.

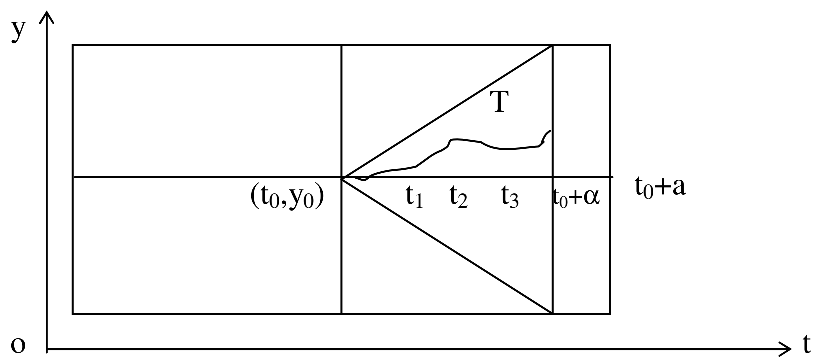


Fig. 2.3

Analytically, the solution function $\varphi(t)$ has the equation

$$\varphi(t) = \varphi(t_{k-1}) + f(t_{k+1}, \varphi(t_{k-1}))\,(t-t_{k-1}) \qquad (9)$$

for $t \in [t_{k-1}, t_k]$ and $k = 1,2,\ldots,n$, and $\phi(t_0) = y_0$. From the construction of the function ϕ, it is clear that $\varphi \,\varepsilon\, C'_p$ on $[t_0, t_0 + \alpha]$, and that

$$|\varphi(t) - \varphi(\bar{t})| \le M|t - \bar{t}| \qquad (10)$$

where t, $\bar{t}$ are in $[t_0, t_0 + \alpha]$.

If t is such that $t_{k-1} \le t \le t_k$, then equations (7) and (10) together imply that

$$\begin{aligned}|\varphi(t) - \varphi(t_{k-1})| &\le M|t-t_{k-1}| \\ &\le M|t_k-t_{k-1}| \\ &\le M.\frac{\delta_\in}{M} = \delta_\in .\end{aligned}$$

From equations (4), (5), (7) and (9), we obtain

$$\begin{aligned}|\varphi'(t) - f(t, \varphi(t))| &= |f(t_{k-1}, \varphi(t_{k-1})) - f(t, \varphi(t))| \\ &\le \in , \end{aligned} \qquad (12)$$

where $t_{k-1} \le t \le t_k$.

This shows that φ is an $\in$-approximate solution, as desired. This completes the proof.

Remark. After finding an "$\in$-approximate solution" of an IVP, one may prove that there exists a sequence of these approximate solutions which tend to a solution. To achieve this aim, the notion of an equicontinuous set of functions is required.

Definition (Equicontinuous set family of functions)

Statement. A set of functions F = {f} defined on a real interval I is said to be equicontinuous on I if, for given any $\in > 0$, there exists a real number $\delta = \delta_\in = \delta(\in) > 0$, independent of f ε F, such that

$$|f(t) - f(\bar{t})| < \in$$

whenever $|t - \bar{t}| < \delta_\in$ for $t, \bar{t} \in I$.

Note. In this definition, the choice of $\delta_\in$ does not depend on the member f of family F but is admissible for all f in the family F.

Theorem 2.2. (**Due to Ascoli**).

Statement. On a bounded interval I, let F = {f} be an infinite, uniformly bounded, equicontinuous set of functions. Prove that F contains a sequence which is uniformly convergent on I.

Proof. Let $\{r_k\}$, k = 1,2,…, be all the rational numbers present in the bounded interval I enumerated/listed in some order. The set of numbers $\{f(r_1) : f \,\varepsilon\, F\}$ is bounded, hence there exists a sequence of distinct functions $\{f_{n_1}\}, f_{n_1} \,\varepsilon\, F$, such that the sequence $\{f_{n_1}(r_1)\}$ is convergent.

Similarly, the set of numbers $\{f_{n_1}(r_2)\}$ has a convergent subsequence $\{f_{n_2}(r_2)\}$.

Continuity in this way, an infinite set of functions $f_{n_k} \,\varepsilon\, F$, n, k = 1,2,…, is obtained which have the property that $\{f_{n_k}\}$ converges at $r_1, r_2, .., r_k$.

Define

$$g_n = f_{nn}\,. \tag{1}$$

Now, it will be shown that $\{g_n\}$ is the required sequence which is uniformly convergent on I. Clearly, $\{g_n\}$ converges at each point r_k of the rationals on I. Thus, given any $\in > 0$, and each rational number $r_k \,\varepsilon\, I$, there exists an integer $N_\in(r_k)$ such that

$$|g_n(r_k) - g_m(r_k)| < \in \quad \text{for an, } m > N_\in(r_k)\,. \tag{2}$$

Since the set F is equicontinuous, there exists a real number $\delta = \delta_\in = \delta(\in) > 0$, which is independent of f ε F, such that

$$|f(t) - f(\bar{t})| < \in\,, \tag{3}$$

for

$$|t - \bar{t}| < \delta_\in \text{ and } t, \bar{t} \,\varepsilon\, I\,.$$

We divide the interval I into a finite number of subintervals $I_1, I_2, \dots, I_p$ such that the length of the largest subinterval is less than $\delta_\in$, i.e.,

$$\max.\{l(I_k) : k = 1,2,\dots,p\} < \delta_\in\,. \tag{4}$$

For each such subinterval I_k, choose a rational number $\bar{r}_k \,\varepsilon\, I_k$. If t ε I, then $t \,\varepsilon\, I_k$ for some suitable k. Hence, by (2) and (3), it follows that

$$|g_n(t) - g_m(t)| \le |g_n(t) - g_n(\bar{r}_k)| + |g_n(\bar{r}_k) - g_m(\bar{r}_k) + |g_m(\bar{r}_k) - g_m(t)|$$

$$< \in + \in + \in = 3\in\,, \tag{5}$$

provided that

$$m,n > \max\{N_\in(\bar{r}_1), N_\in(\bar{r}_2), \dots, N_\in(\bar{r}_p)\}\,.$$

This proves the uniform convergence of the sequence $\{g_n\}$ on I, where $g_n \in F$ for each $n \in N$. This completes the proof.

Remark. The existence of a solution to the initial-value problem, without any further restriction on the function f(t, y) is guaranteed by the following Cauchy-Peano theorem.

Theorem 2.3. (**known as Cauchy-Peano Existence theorem**).

Statement. If f ε C on the rectangle R, then there exists a solution $\varphi \in C^1$ of the differential equation

$$\frac{dy}{dt} = f(t,y)$$

on the interval $|t-t_0| \le \alpha$ for which $\varphi(t_0) = y_0$, where

$$R = \{(t,y) : |t-t_0| \le a, |y-y_0| \le b, \ a > 0, b > 0\}, \ \alpha = \min\left(a, \frac{b}{M}\right),$$

$$M = \max |f(t,y)| \text{ on } R .$$

Proof. Let $\{\in_n\}$, n = 1,2,…, be a monotonically decreasing sequence of positive real numbers which tends to 0 as $n \to \infty$. By theorem 2.1, for each such $\in_n$, there exists an $\in_n$–approximate solution, say φ_n, of ODE

$$\frac{dy}{dt} = f(t, y) \tag{1}$$

on the interval

$$|t-t_0| \le \alpha \ \text{ with } \varphi_n(t_0) = y_0 . \tag{2}$$

It is being given that

$$\alpha = \min\left(a, \frac{b}{M}\right) \tag{3}$$

$$M = \max|f(t,y)| \quad \text{for } (t,y) \in R \tag{4}$$

$$R = \{(t,y): |t-t_0| \le a, \ |y-y_0| \le b, \ a > 0 , b> 0 \}. \tag{5}$$

Further, from theorem 2.1, it follows that

$$|\varphi_n(t) - \varphi_n(\bar{t})| \le M |t - \bar{t}| \tag{6}$$

for t, $\bar{t}$ in $[t_0, t_0+\alpha]$.

Applying (6) to $\bar{t} = t_0$ and since, we know that

$$|t-t_0| \le \alpha \le \frac{b}{M} , \tag{7}$$

it follows that

$$|\varphi_n(t) - y_0| < b \quad \text{for all t in } \ |t-t_0| < \alpha . \tag{8}$$

This implies that the sequence $\{\varphi_n\}$ is uniformly bounded by $|y_0| + b$.

Further, (6) implies that the sequence $\{\varphi_n\}$ is an equicontinuous set. Hence, by the theorem 2.2, there exists a subsequence $\{\phi_{n_k}\}$, k = 1,2,.., of $\{\varphi_n\}$, converging uniformly on the interval $[t_0 - \alpha, t_0 + \alpha]$ to a limit function ϕ, which must be continuous since each ϕ_n is continuous.

Now, we shall show that this limit function ϕ is a solution of (1) which meets the required specifications. For this, we write the relation defining ϕ_n as an $\in_n$–approximate solution in an integral form, as follows :

$$\phi_n(t) = y_0 + \int_{t0}^{t} [f(s, \phi_n(s)) + \Delta_n(s)] \, ds \tag{9}$$

where

$$\Delta_n(s) = \phi'_n(s) - f(s, \phi_n(s)) \tag{10}$$

at those points where ϕ'_n exists, and $\Delta_n(s) = 0$, otherwise.

Because ϕ_n is an $\in_n$-approximate solution, so

$$|\Delta_n(s)| \le \in_n . \qquad (11)$$

Since f is uniformly continuous on R, and $\phi_{n_k} \to \phi$ uniformly on $[t_0 - \alpha, t_0 + \alpha]$ as $k \to \infty$, it follows that

$$f(t, \phi_{n_k}(t)) \to f(t, \phi(t))$$

uniformly on $[t_0 - \alpha, t_0 + \alpha]$, as $k \to \infty$.

Replacing n by n_k in (9), one obtains, in letting $k \to \infty$,

$$\phi(t) = y_0 + \int_{t0}^{t} f(s, \phi(s))ds . \qquad (12)$$

From (12), we get

$$\phi(t_0) = y_0 \qquad (13)$$

and upon differentiation, as f is continuous,

$$\frac{d\phi}{dt} = f(t, \phi(t)) . \qquad (14)$$

It is clear from (13) and (14) that ϕ is a solution of ODE (1) through the point (t_0, y_0) on the interval $|t - t_0| \le \alpha$ of class C^1. This completes the proof of the theorem.

Remarks. (1) If uniqueness of solution is assured, the choice of a subsequence in theorem 2.1 is unnecessary.

(2) It can happen that the choice of a subsequence is unnecessary even though uniqueness is not satisfied. Consider the example

$$\frac{dy}{dt} = y^{1/3} . \qquad (1)$$

There are an infinite number of solutions starting/issuing at the point C(0,0) which exist on I = [0,1].

For any c, $0 \le c \le 1$, the function ϕ_c defined by

$$\phi_c(t) = \begin{cases} 0 & 0 \le t \le c \\ \left[\dfrac{2(t-c)}{3}\right]^{3/2} & c < t \le 1 \end{cases} \qquad (2)$$

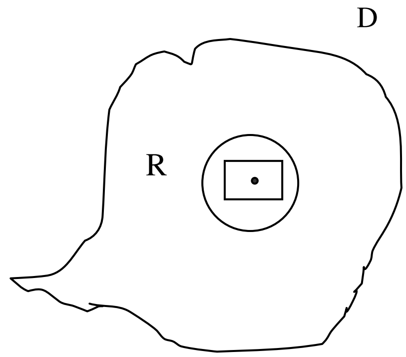
is a solution of (1) on I. If the construction of theorem 2.1 is applied to equation (1), one finds that the only polygonal path starting at the path C(0,0) is ϕ_1. This shows that this method cannot, in general, give all solutions of (1).

Theorem 2.4. Let $f \in C$ on a domain D in the (t, y) plane, and suppose (t_0, y_0) is any point in D. Then there exists a solution ϕ of

$$\frac{dy}{dt} = f(t,y) \quad \text{for } (t, y) \in D \ , \quad y(t_0) = y_0 \qquad (1)$$

on some t-interval containing t_0 in its interior.

Proof. Since domain D is open, there exists an r > 0 such that all points whose distance from $C(t_0, y_0)$ is less than r, are contained in D.



Let R be any closed rectangle containing $C(t_0, y_0)$, and contained in this open circle of radius r. Then theorem 2.2 applied to (1) on R gives the required result.

3

UNIQUENESS OF SOLUTIONS

We consider the following examples.

Example 1: Consider the initial value problem

$$\frac{dy}{dt} = y^{2/3} \text{ in } [0, 1],$$
$$y(0) = 0.$$

Hence $f(t, y) = y^{2/3}$ is a continuous function.
There are two solutions of it, namely,

$$y_1(t) \equiv 0,$$
$$y_2(t) = \frac{t^3}{27} \text{ in } [0, 1] .$$

Example 2: Consider the initial value problem

$$\frac{dy}{dt} = y^{1/2} \text{ in } [0, 1]$$
$$y(0) = 0.$$

In this problem $f(t,y) = y^{1/2}$ is continuous. This problem also have two solutions, namely,

$$y_1(t) \equiv 0,$$
$$y_2(t) = \frac{t^2}{4} \text{ in } [0, 1]$$

Example 3: Consider the initial value problem

$$\frac{dy}{dt} = y^{1/3} \text{ in } [0, 1],$$
$$y(0) = 0.$$

Here $f(t, y) = y^{1/3}$ is a continuous function. Further

$$y_1(t) \equiv 0,$$
$$y_2(t) = \left(\frac{2t}{3}\right)^{3/2},$$

are two solutions of the above initial value problem.
The above examples show that something more than the continuity of $f(t,y)$ in the differential equation

$$\frac{dy}{dt} = f(t, y) \text{ in } D$$

is required in order to guarantee that a solution passing through a given point $(t_0, y_0) \in D$ be unique.
A simple condition which permits one to imply uniqueness of solution is the Lipschitz condition, defined below.

Definition: Suppose f(t, y) is defined in a domain D in the (t, y) plane. If there exists a constant K > 0 such that

$$| f(t, y_1) - f(t, y_2) | \le K | y_1 - y_2 |$$

for every pair of points (t, y_1) and (t, y_2) in D, then f (t,y) is said to satisfy a Lipschitz condition w.r.t. y in D. The constant K is called the Lipschitz constant.

Notations: (1) The fact that f(t,y) satisfy Lipschitz condition is expressed as

$$f \in \text{Lip in D}.$$

(2) If, in addition $f \in C$ in D, we write as

$$f \in (C, \text{Lip}) \text{ in D}.$$

Remark. If $f \in$ Lip in D, then f is uniformly continuous in y for each fixed t, although nothing is implied concerning the continuity of f w.r.t. "t".

Definition(Convex set): A set $D \subseteq R^2$ is said to be convex set if D contains the line segment joining any two points in D.

Theorem (3.1): Let f(t, y) be such that $\frac{\partial f}{\partial y}$ exists and is bounded for all $(t, y) \in D$, where D is a domain or closed domain such that the line segment joining any two points of D lies entirely within D, Then f satisfies a Lipschitz condition, (with respect to y) in D, where the Lipschitz constant is given by

$$K = \underset{(t,y)\in D}{\text{lub}} \left| \frac{\partial f(t, y)}{\partial y} \right| .$$

Proof: Since $\frac{\partial f(t,y)}{\partial y}$ exists and is bounded for all $(t, y) \in D$, there exists a constant K (K > 0) such that

$$\underset{(t,y)\in D}{\text{lub}} \left| \frac{\partial f(t, y)}{\partial y} \right| = K . \tag{1}$$

Moreover, by the mean value theorem of differential calculus, for any pair of points (t, y_1), (t, y_2) in D there exists ξ , $y_1 < \xi_1 < y_2$, such that

$$f(t, y_1) - f(t, y_1) = (y_1 - y_2) \frac{\partial f(t, \xi_1)}{\partial y}, \tag{2}$$

for $(t, \xi) \in D$. Thus

$$\begin{aligned} | f(t, y_1) - f(t, y_2)| &= |y_1 - y_2| \left| \frac{\partial f(t, \xi_{1|})}{\partial y} \right| \\ &\le |y_1 - y_2| \left(\underset{(t,y)\in D}{\text{lub}} \left| \frac{\partial f(t, y)}{\partial y} \right| \right) \\ &= K | y_1 - y_2 |, \end{aligned}$$

This implies

$$| f(t, y_1) - f(t, y_2)| \le K | y_1 - y_2 | , \tag{3}$$

for all (t, y_1), (t, y_2) in D. This shows that f(t, y) satisfies a Lipschitz condition in D and K is the Lipschitz constant.

Remark 1 : The sufficient condition of the above theorem (3.1) is not necessary for f (t, y) to satisfy a Lipschitz condition in D. That is, there exists function f(t, y) such that f satisfy a Lipschitz condition in D but such that the hypothesis of theorem (3.1) is not satisfied.

(a) Consider the function f defined by

$$f(t, y) = t \mid y \mid , \tag{1}$$

where D is the rectangle defined by

$$D = \{(t,y) \mid t \mid \leq a, \mid y \mid \leq b\} \tag{2}$$

we note that

$$\begin{aligned} \mid f(t, y_1) - f(t, y_2) \mid &= |t| y_1| - t|y_2|| \\ &\leq \mid t \mid \mid y_1 - y_2 \mid \\ &\leq a|y_1 - y_2 \mid \end{aligned} \tag{3}$$

for all (t, y_1) and (t, y_2) in D. $(\Theta \mid |y_1| - |y_2| \mid \leq \mid y_1 - y_2 \mid)$

Thus f(t,y) satisfies a Lipschitz condition in D. However, the partial derivative $\frac{\partial f}{\partial y}$ does not exists at any point $(t, 0) \in D$ for which $t \neq 0$.

(b) Consider the function f(t, y) given by

$$f(t, y) = t^2 \mid y \mid \tag{1}$$

in the rectangle

$$R = \{\mid (t, y) \mid |t| \leq 1, |y| \leq 1\}. \tag{2}$$

We find

$$\begin{aligned} \mid f(t, y_1) - f(t, y_2) \mid &\leq \mid t^2 |y_1| - t^2 |y_2| \mid \\ &\leq \mid t^2 \mid \mid y_1 - y_2 \mid \\ &\leq \mid y_1 - y_2 \mid \end{aligned} \tag{3}$$

in R. So f(t, y) satisfies a Lipschitz condition, with Lipschitz constant 1. However, the partial derivative $\frac{\partial f}{\partial y}$ does not exist at any point (t, 0) in R for $t \neq 0$, as

$$\begin{aligned} &\lim_{h \to 0} \frac{f(t, 0+h) - f(t,0)}{h} \\ &= \lim_{h \to 0} \frac{t^2 \mid h \mid - t^2 .0}{h} \\ &= \lim_{h \to 0} \frac{t^2 \mid h \mid}{h} \end{aligned} \tag{4}$$

does not exist.

Remark 2 : The two initial value problems, discussed in the beginning of this lesson, did not have a unique solution, and this can now be attributed to the failure of the Lipschitz condition at $t = 0$.

Definition: The series $\sum_{n=1}^{\infty} u_n(t)$ is said to converge uniformly to a function u(t) on the interval $a \leq t \leq b$ if its sequence of partial sums, $\{S_n(t)\}$, converges uniformly to u(t) on the interval $a \leq t \leq b$.

Weierstrass M-test : Suppose $\{u_n(t)\}$ is a sequence of real valued functions defined on the interval $a \leq t \leq b$, and suppose

$$\mid u_n(t) \mid \leq M_n \quad \text{for all } n = 1, 2, 3, \ldots..$$

and for all $t \in [a, b]$. Then the series $\sum_{n=1}^{\infty} u_n(t)$ converges uniformly on the interval [a, b] if the series $\sum_{n=1}^{\infty} M_n$ of positive real numbers converges.

Note : The existence proof given in Cauchy-Peano existence theorem (2.1) is unsatisfactory in the respect that there is no constructive method given for obtaining a solution to the initial value problem. In particular, if f(t, y) satisfies a Lipschitz condition in addition to its continuity, a relatively simple yet a very useful method exists, known as the method of successive approximations, which deduces the existence and uniqueness of a solution of the given initial value problem. This method is also called Picard iteration method and which is given in the following theorem.

Theorem 3.2 : (Picard – Lindelof Theorem)

Statement: Let D be a domain of the ty-plane. Let (t_0, y_0) be an interior point of D. Let rectangle $R = \{(t, y) : | t - t_0 | \le a, | y - y_0 | \le b, a > 0, b > 0\}$ lie within D. Let f(t, y) be a real valued function which is continuous in D, satisfies a Lipschiz condition (w.r.t. y) in D and $M = \max | f(t, y) |$ in R. Then, there exists a unique solution $\phi = \phi(t)$ of the initial value problem

$$\frac{dy}{dt} = f(t, y) \text{ in D}, \; y(t_0) = y_0$$

on the closed interval $| t - t_0 | \le \alpha$, where $\alpha = \min\{a, \frac{b}{M}\}$.

Proof : The given initial value problem is equivalent to the following Volterra integral equation

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds, \qquad (1)$$

for $| t - t_0 | \le a$. Thus, a solution of the given I V P on $| t - t_0 | \le \alpha$ must satisfy (1) and conversely. Now, we define a sequence $\{\phi_k\}$ of successive approximations (Picard iterants) of the problem by the recurrence formulas

$$\phi_0(t) = y_0,$$

$$\phi_{k+1}(t) = y_0 + \int_{t_0}^{t} f(s, \phi_k(s))ds \quad , \qquad (2)$$

on the interval $| t - t_0 | \le \alpha$. Here, k = 0, 1, 2, …..

We shall be considering the interval $[t_0, t_0 + \alpha]$ only. Similar arguments hold for the interval $[t_0 - \alpha, t_0]$. Firstly, it will be shown that

(i) every $\phi_k(t)$ exists on $[t_0, t_0 + \alpha]$,

(ii) $\phi_k \in C^1$, and

(iii) $| \phi_k(t) - y_0 | \le M(t - t_0)$, (3)

for $t \in [t_0, t_0 + \alpha]$ and for all k. We shall prove it by the mathematical induction.

Obviously ϕ_0, being the constant function, satisfies these conditions. Now, we assume that φ_k does the same and we shall prove the requirements for ϕ_{k+1} .

By assumption, $f(t, \phi_k(t))$ is defined and continuous on the interval $[t_0, t_0 + \alpha]$. Hence, by formula (2),

$\phi_{k+1}(t)$ exists on $[t_0, t_0 + \alpha]$,

$\phi_{k+1}(t) \in C_s^1$ there,

and

$$| \phi_{k+1}(t) - y_0 | = | \int_{t_0}^{t} f(s, \phi_k(s))\, ds |$$

$$\leq \int_{t_0}^{t} | f(s, \phi_k(s)) | ds$$

$$= M(t - t_0) . \quad (4)$$

Therefore these properties are shared by the function $\phi_k(t)$ by induction, for all k.

Secondly it shall be proved that the sequence $\{\phi_k\}$of functions converges uniformly to a continuous function $\phi = \phi(t)$ on I. For this, we define

$$\Delta_k(t) = | \phi_{k+1}(t) - \phi_k(t) | , \quad (5)$$

for $t \in [t_0, t_0 + \alpha]$. From equations (2) and (5), we write

$$\Delta_k(t) = | \int_{t_0}^{t} \{f(s, \phi_k(s)) - f(s, \phi_{k-1}(s)\}ds |$$

$$\leq \int_{t_0}^{t} | f(s, \phi_k(s)) - f(s, \phi_{k-1}(s) | ds$$

$$\leq K \int_{t_0}^{t} | \phi_k(s) - \phi_{k-1}(s) | ds$$

$$= K \int_{t_0}^{t} \Delta_{k-1}(s) ds , \quad (6)$$

where K is a Lipschitz constant and we have used the fact that $f \in$ Lip on R.

Equation (3) gives for k = 1,

$$\Delta_0(t) = | \phi_1(t) - \phi_0(t) |$$
$$= | \phi_1(t) - y_0 |$$
$$\leq M(t - t_0) ,$$

and an easy induction (left as an exercise to a reader) on (6) implies that

$$\Delta_k(t) \leq \frac{M}{K} . \frac{(\alpha K)^{k+1}}{(k+1)!} , \qquad (\Theta\, t - t_0 \leq \alpha) \quad (7)$$

for all k and $t \in [t_0, t_0 + \alpha]$.

This shows that the terms of the series $\sum_{k=0}^{\infty} \Delta_k(t)$ are majorized by those of the power series for $\left(\frac{M}{K}\right)[e^{\alpha K} - 1]$, and therefore, by the Weierstrass M – test for uniform convergence, the series $\sum_{k=0} \Delta_k(t)$ is uniformly convergent on $[t_0, t_0 + \alpha]$.

Thus, the series

$$\phi_0(t) + \sum_{k=0}^{\infty} [\phi_{k+1}(t) - \phi_k(t)] , \quad (8)$$

is absolutely and uniform convergent on the interval $[t_0, t_0 + \alpha]$. Consequently, the sequence of its partial sums

$$S_{n+1}(t) = \phi_0(t) + \sum_{k=0}^{n-1} [\phi_{k+1}(t) - \phi_k(t)]$$
$$= \phi_n(t) , \quad (9)$$

is absolutely and uniformly convergent on the interval $[t_0, t_0 + \alpha]$, to a limit function, say $\phi(t)$,which is continuous on $[t_0, t_0 + \alpha]$.

Thirdly, it will be shown that this limit function $\phi(t)$ is a solution of desired problem in equation (1). As ϕ is continuous, so $f(s, \phi(s))$ exists for $s \in [t_0, t_0 + \alpha]$ and

$$\left| \int_{t_0}^{t} [f(s, \phi(s)) - f(s, \phi_k(s))]\, ds \right|$$

$$\leq \int_{t_0}^{t} | f(s, \phi(s)) - f(s, \phi_k(s)) |\, ds$$

$$\leq K \int_{t_0}^{t} | \phi(s) - \phi_k(s) |\, ds \tag{10}$$

as f satisfies the Lipschitz condition on R.
Since $\phi_k \to \phi$ uniformly on $[t_0, t_0 + \alpha]$, so

$$| \phi(s) - \phi_k(s) | \to 0 \text{ as } k \to \infty \tag{11}$$

uniformly on the interval $[t_0, t_0 + \alpha]$. Combining (10) and (11), it follows that

$$f(s, \phi_k(s)) \to f(s, \phi(s)) \text{ on } [t_0, t_0 + \alpha] \tag{12}$$

uniformly as $k \to \infty$. Consequently equations (2), (9) and (12) imply that, on taking $k \to \infty$, we get at once

$$\phi(t) = y_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds . \tag{13}$$

This proves that $\phi(t)$ is a solution of integral equations (1), and, therefore, a solution of the given initial – value problem on the interval $[t_0, t_0 + \alpha]$.

Finally, we shall prove that solutions of (1) is unique. If possible, suppose that $\Psi = \Psi(t)$ is another solution of integral equations (1). Then

$$\Psi(t) = y_0 + \int_{t_0}^{t} f(s, \Psi(s))\, ds, \tag{14}$$

on I. Let

$$N = \text{Max } | \phi(t) - \Psi(t) | \text{ on I.} \tag{15}$$

From equations (13) to (14), we obtain

$$| \phi(t) - \Psi(t) | = \left| \int_{t_0}^{t} \{f(s, \phi(t)) - f(s , \Psi(s))\}ds \right|$$

$$\leq \int_{t_0}^{t} | f(s, \phi(t)) - f(s , \Psi(s) |\, ds$$

$$\leq K \int_{t_0}^{t} | \phi(s) - \Psi(s) |\, ds . \tag{16}$$

Using (15) in (16), we get

$$| \phi(t) - \Psi(t)| \leq KN(t - t_0), \text{ for } t \in [t_0, t_0 + \alpha]. \tag{17}$$

By the repeated use of (16) and (17), we obtain

$$| \phi(t) - \Psi(t)| \leq \frac{K^n N(t - t_0)^n}{n!}, \; t \in [t_0, t_0 + \alpha] \tag{18}$$

and for all n = 1, 2, 3, ……. Since, the series $\sum \frac{K^n (t-t_0)^n}{n!}$ of positive terms converges for $t \in [t_0, t_0 + \alpha]$, therefore,

$$\frac{K^n (t-t_0)^n}{n!} \to 0 \text{ for all } t \in [t_o, t_o + \alpha] \tag{19}$$

as $n \to \infty$. From equations (18) and (19), we find

$$|\phi(t) - \Psi(t)| \le 0 \text{ for all } t_0 \le t \le t_0 + \alpha$$

This implies

$$\phi(t) = \Psi(t) \text{ for all } t_0 \le t \le t_0 + \alpha \tag{20}$$

thus, solution $\phi = \phi(t)$ of the given initial – value problem is unique the interval $[t_0 - t_0 + \alpha]$. We can carry through similar agreements on the interval $[t_0 - \alpha, t_0]$. This completes the proof of the theorem.

Remark: A significant and useful feature of this proof is that the sequence $\{\phi_n(t)\}$ converges uniformly to $\phi(t)$, which is the unique solution of the given initial – value problem. The proof is also constructive in that it provides a method of obtaining approximate solutions to any required degree of accuracy. However, from this point of view, it is rarely of practical value as the iterates usually converge too slowly to be useful. Further, the above theorem is only a local existence theorem in that, whatever the original interval of definition, $|t - t_0| \le a$, in general the solution is only guaranteed to exist in a smaller interval, $|t - t_0| \le \alpha$, where $\alpha \le a$.

Example 1. Consider the initial value problem $\frac{dy}{dt} = y$, $y(0) = 1$.

Solution. Integrating over the interval [0, t], we obtain

$$y(t) = 1 + \int_0^t y(s)ds , \tag{1}$$

which is a volterra integral of the second kind. Let

$$\phi_0(t) = 1 . \tag{2}$$

Then , by Picard's method

$$\phi_1(t) = 1 + \int_0^t \phi_0(s)\, ds$$

$$= 1 + \int_0^t 1\, ds$$

$$= 1 + t , \tag{3}$$

$$\phi_2(t) = 1 + \int_0^t \phi_1(s)ds$$

$$= 1 + \int_0^t (1 + s)\, ds$$

$$= 1 + t + \frac{t^2}{2!} , \tag{4}$$

$$\phi_3(t) = 1 + \int_0^t \phi_2(s)ds$$

$$= 1 + \int_0^t (1 + s + \frac{s^2}{2!})ds$$

$$= 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} . \qquad (5)$$

Continuing like this, we shall obtain

$$\phi_n(t) = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \ldots + \frac{t^n}{n!} . \qquad (6)$$

Taking the limit as $n \to \infty$, we get

$$\lim_{n\to\infty} \phi_n(t) = e^t . \qquad (7)$$

This implies that

$$\phi(t) = e^t \qquad (8)$$

is the unique solution of the given initial value problem, by the Picard's method of successive approximations.

Example 2: Solve the initial – value problem

$$\frac{dy}{dt} = t^2 y, \; y(0) = 1$$

By Picard method.

Solution. The corresponding integral equation is

$$y(t) = 1 + \int_0^t s^2 y(s)ds. \qquad (1)$$

Picard's iterates are

$$\phi_0(t) = 1 , \qquad (2)$$

$$\phi_1(t) = 1 + \int_0^t s^2 . 1ds$$

$$= 1 + \frac{t^3}{3} , \qquad (3)$$

$$\phi_2(t) = 1 + \int_0^t s^2\left(1+\frac{s^3}{3}\right) ds$$

$$= 1 + \frac{t^3}{3} + \frac{1}{2!}\left(\frac{t^3}{3}\right)^2, \qquad (4)$$

$$\phi_3(t) = 1 + \int_0^t s^2\left(1+\frac{s^3}{3} + \frac{1}{2!}\left(\frac{s^3}{3}\right)^2\right) ds$$

$$= 1 + \frac{t^3}{3} + \frac{1}{2!}\left(\frac{t^3}{3}\right)^2 + \frac{1}{3!}\left(\frac{t^3}{3}\right)^3, \qquad (5)$$

$$\vdots$$

$$\phi_n(t) = 1 + \int_0^t s^2 \phi_{n-1}(s)\, ds$$

$$= 1 + \frac{t^3}{3} + \frac{1}{2!}\left(\frac{t^3}{3}\right)^2 + \ldots + \frac{1}{n!}\left(\frac{t^3}{3}\right)^n . \quad (6)$$

The exact solution can easily be seen to be

$$\phi(t) = e^{(t^3/3)}, \quad (7)$$

to which the above approximate solutions converge.

Example 3: Solve the initial-value problem

$$\frac{dy}{dt} = t(y - t^2 + 2),\ y(0) = 1 ,$$

by Picard's method.

Solution. The integral equation, equivalent to the above initial value problem is

$$y(t) = 1 + \int_0^t s(y(s) - s^2 + 2)\, ds. \quad (1)$$

The approximate solutions are

$$\phi_0(t) = 1 ,$$

$$\phi_1(t) = 1 + \int_0^t s(3 - s^2)\, ds$$

$$= 1 + \frac{3t^2}{2} - \frac{1}{4}t^4 \quad (2)$$

$$\phi_2(t) = 1 + \int_0^t s\left(3 + \frac{s^2}{2} - \frac{s^4}{4}\right) ds ,$$

$$= 1 + \frac{3t^2}{2} + \frac{t^4}{8} - \frac{t^6}{24} \quad (3)$$

$$\phi_3(t) = 1 + \int_0^t s\left(3 + \frac{s^2}{2} + \frac{s^4}{8} - \frac{s^6}{24}\right) ds$$

$$= 1 + \frac{3t^2}{2} + \frac{t^4}{8} + \frac{t^6}{48} - \frac{t^8}{192} , \quad (4)$$

and so on.

The exact solution can be easily found to be

$$\phi(t) = t^2 + e^{t^2/2}$$

$$= t^2 + 1 + \frac{t^2}{2} + \frac{t^4}{8} + \frac{t^6}{48} + \frac{t^8}{384} + \ldots \text{to } \infty . \quad (5)$$

Example 4: Solve the differential equation

$$\frac{dy}{dt} = t\, y,\ y(0) = 1 ,$$

by the method of successive approximations.

Solution : We write f(t, y) = t y and the integral equation corresponding to the initial value problem is

$$y(t) = 1 + \int_0^t t\, y(t)\, dt. \quad (1)$$

The successive approximations are, given by,

$$\phi_0(t) = 1,$$

$$\phi_n(t) = 1 + \int_0^t t\,\phi_{n-1}(t)\,dt, \qquad (2)$$

for n = 1, 2, 3, …..

Thus

$$\phi_0(t) = 1\ , \qquad (3)$$

$$\phi_1(t) = 1 + \int_0^t s\,ds$$

$$= 1 + \frac{t^2}{2}, \qquad (4)$$

$$\phi_2(t) = 1 + \int_0^t s\left(1+\frac{s^2}{2}\right)ds$$

$$= 1 + \int_0^t \left(s+\frac{s^2}{2}\right)ds$$

$$= 1 + \frac{t^2}{2} + \frac{t^4}{2.4}. \qquad (5)$$

We shall establish by induction that

$$\phi_n(t) = 1 + \left(\frac{t^2}{2}\right) + \frac{1}{2!}\left(\frac{t^2}{2}\right)^2 + \ldots\ldots.. + \frac{1}{n!}\left(\frac{t^2}{2}\right)^2, \qquad (6)$$

for all n.

for n = 0, 1, 2; we have already checked the relation (6). Suppose that.

$$\phi_{n-1}(t) = 1 + \left(\frac{t^2}{2}\right) + \frac{1}{2!}\left(\frac{t^2}{2}\right)^2 + \ldots\ldots.. + \frac{1}{n!}\left(\frac{t^2}{2}\right)^{n-1} \qquad (7)$$

Then

$$\phi_n(t) = 1 + \int_0^t s\left\{1+\left(\frac{s^2}{2}\right) + \frac{1}{2!}\left(\frac{s^2}{2}\right)^2 + \ldots\ldots.. + \frac{1}{n-1!}\left(\frac{t^2}{2}\right)^{n-1}\right\}ds$$

$$= 1 + \int_0^t \left\{s+\frac{s^3}{2} + \frac{1}{2!}\ \frac{s^5}{2^2} + \ldots\ldots.. + \frac{1}{n-1!}\ \frac{s^{2n-1}}{2^{n-1}}\right\}ds$$

$$= 1 + \left\{\frac{t^2}{2}+\frac{t^4}{2.4} + \frac{1}{2!}\ \frac{t^6}{2^2.6} + \ldots\ldots.. + \frac{1}{n-1!}\ \frac{t^{2n}}{2^{n-1}(2n)}\right\}.$$

This implies

$$\phi_n(t) = 1 + \left(\frac{t^2}{2}\right) + \frac{1}{2!}\left(\frac{t^2}{2}\right)^2 + \frac{1}{3!}\left(\frac{t^2}{2}\right)^3 + \ldots\ldots.. + \frac{1}{n!}\left(\frac{t^2}{2}\right)^n. \qquad (8)$$

Therefore, by the principle of mathematical induction, the equality (6) is true for all n = 1, 2, 3,…… Moreover, we observe that $\phi_n(t)$ is the partial sum of the first (n + 1) terms of the infinite series expansion of the function

$$\phi(t) = e^{\frac{t^2}{2}}. \qquad (9)$$

Further this series converges for all real t. This means that

$$\phi_n(t) \to \phi(t) ,$$

for all real t. Hence, the function ϕ is the required solution.

Example 5: Solve the initial value problem

$$\frac{dy}{dt} = -y, \ y(0) = 1$$

by the method of successive approximations.

Solution : The integral equation equivalent to the given initial value problem is

$$y(t) = 1 + \int_0^t -y(s)\, ds = 1 - \int_0^t y(s)\, ds. \qquad (1)$$

The successive approximations given by Picard's method are

$$\phi_0(t) = 1,$$

$$\phi_{n+1}(t) = 1 - \int_0^t \phi_n(t)\, dt. \text{ for } n = 1, 2, \ldots\ldots\ldots \qquad (2)$$

We find

$$\phi_1(t) = 1 - \int_0^t ds$$

$$= 1 - t \ , \qquad (3)$$

$$\phi_2(t) = 1 - \int_0^t (1 - s)ds$$

$$= 1 - t + \frac{t^2}{2!} \ , \qquad (4)$$

$$\phi_3(t) = 1 - \int_0^t \left(1 - s + \frac{s^2}{2!}\right) ds$$

$$= 1 - t + \frac{t^2}{2!} - \frac{t^3}{3!}. \qquad (5)$$

By the induction, it may be verified that (left as an exercise to the readers)

$$\phi_n(t) = 1 - \frac{t}{1!} + \frac{t^2}{2!} - \frac{t^3}{3!} + \ldots + (-1)^n \frac{t^n}{n!} \quad . \qquad (6)$$

We observe that $\phi_n(t)$ is the partial sum of the first (n + 1) terms of the infinite series expansion of the function

$$\phi(t) = e^{-t} \ . \qquad (7)$$

Further this series converges for all real t. This means that

$$\phi_n(t) \to \phi(t) = e^{-t} ,$$

for all t. Hence the function ϕ, given in (7) , is the solution of the given problem.

Example 6: Solve the initial value problem

$$\frac{dy}{dt} = 2ty, \ y(0) = 1$$

by the method of successive approximations.

Solutions : The given initial – value problem is equivalent to the integral equation

$$y(t) = 1 + \int_0^t 2s \,.\, y(s)\, ds\,. \qquad (1)$$

The successive approximations are

$$\phi_0(t) = 1,$$

$$\varphi_{n+1}(t) = 1 + \int_0^t 2s\, \phi_n(s)\, ds. \qquad (2)$$

we find

$$\phi_1(t) = 1 + 2\int_0^t s \,.\, 1 ds \qquad (3)$$

$$= 1 + t^2,$$

$$\phi_2(t) = 1 + 2\int_0^t s(1 + s^2)\, ds$$

$$= 1 + t^2 + \frac{t^4}{2!}, \qquad (4)$$

$$\phi_3(t) = 1 + \int_0^t 2s\left(1 + s^2 + \frac{s^4}{2!}\right) ds$$

$$= 1 + t^2 + \frac{t^4}{2!} + \frac{t^6}{3!}. \qquad (5)$$

From the induction, we shall find (left as an exercise)

$$\phi_n(t) = 1 + t^2 + \frac{t^4}{2!} + \frac{t^6}{3!} + \ldots\ldots\ldots + \frac{t^{2n}}{n!}. \qquad (6)$$

we visualize that $\phi_n(t)$ is the partial sum of the first (n + 1) terms of the infinite series expansion of the function

$$\phi(t) = e^{t^2}. \qquad (7)$$

Further this series converges uniformly for all real t. This means that

$$\phi_n(t) \to \phi(t) = e^{t^2}$$

for all t. Hence the function $\phi(t)$ is the required solution of the given problem.

Example 7: Solve the initial value problem

$$\frac{dy}{dt} = y,\; y(1) = 1$$

by the method of successive approximations.

Solutions : The given problem is equivalent to the integral equation

$$y(t) = 1 + \int_1^t y(s)\, ds. \qquad (1)$$

The successive approximations are given by

$$\phi_0(t) = 1,$$

$$\phi_{n+1}(t) = 1 + \int_1^t \phi_n(s)\, ds\,. \qquad (2)$$

we find

$$\phi_0(t) = 1\,, \qquad (3)$$

$$\phi_1(t) = 1 + \int_1^t ds$$

$$= t \ , \qquad (4)$$

$$\phi_2(t) = 1 + \int_1^t s\, ds$$

$$= 1 + \int_1^t [(s-1)+1]\, ds. \qquad (5)$$

Here, it is convenient to have integrand occurring in the successive approximations in powers of (s – 1) rather than in powers of s (Θ t_0 = 1 and not zero). Therefore, (5) gives

$$\phi_2(t) = 1 + \left\{ s + \frac{(s-1)^2}{2} \right\} \quad = 1 + (t-1) + \frac{(t-1)^2}{2!}. \qquad (6)$$

$$\phi_3(t) = 1 + \int_1^t \left\{ 1 + (s-1) + \frac{(s-1)^2}{2} \right\} ds$$

$$= 1 + (t-1) + \frac{(t-1)^2}{2!} \ \frac{(t-1)^3}{3!} \ . \qquad (7)$$

By induction, we shall obtain (exercise)

$$\phi_n(t) = 1 + (t-1) + \frac{(t-1)^2}{2!} + \frac{(t-1)^3}{3!} + ... + \frac{(t-1)^n}{n!}. \qquad (8)$$

We note that $\phi_n(t)$ is the partial sum of the first (n + 1) terms of the infinite series expansion of the function

$$\phi(t) = e^{t-1}. \qquad (9)$$

Moreover this series converges for all real t. Therefore,

$$\phi_n(t) \to \phi(t) = e^{t-1}$$

for all t. Hence the function $\varphi(t)$, given in (8), is the required solution of the given problem.

4

THE n-th ORDER DIFFERENTIAL EQUATION

Let n be a positive integer. Let $f^1, f^2, ..., f^n$ be n real valued continuous functions defined on some domain D of the real $(t, y^1, y^2, \ldots, y^n)$ space, which is (n + 1) dimensional. As before, $t \in I$ and $y^i = y^i(t)$ for $t \in I$.

A system of n ordinary differential equations of the first order is of the type (in normal form)

$$\frac{dy^1}{dt} = f^1(t, y^1, y^2, y^3, \ldots, y^n)$$

$$\frac{dy^2}{dt} = f^2(t, y^1, y^2, \ldots, y^n)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{dy^n}{dt} = f^n(t, y^1, y^2, \ldots\ldots, y^n)$$

or written as

$$\frac{dy^i}{dt} = f^i(t, y^1, y^2, \ldots\ldots, y^n),\ 1 \le i \le n.$$

Initial – value problem: Let $(t_0, y_0^1, y_0^2, \ldots\ldots, y_0^n) \in D$. The initial – value problem consists of finding n differentiable functions $\phi^1(t), \phi^2(t), \ldots\ldots \phi^n(t)$ defined on a real t intend I such that

(i) $(t, \phi^1(t), \phi^2(t), \ldots\ldots, \phi^n(t)) \in D$ for all $t \in I$

(ii) $\dfrac{d\phi^i(t)}{dt} = f^i(t, \phi^1(t), \phi^2(t), \ldots\ldots, \phi^n(t))$ for all $t \in I$

(iii) $\phi^i(t_0) = y_0^i$ for $1 \le i \le n$.

Remark: The results so for obtained (for the case n = 1) can be carried over successfully the system of differential equations (5). Let R^n be the n – dimensional real Euclidean space with its elements $y = (y^1, y^2, \ldots\ldots, y^n)$, $y^i \in R$ for $1 \le i \le n$. In developing the theory of system of differential equations (5), we shall need to use a convenient measure of the magnitude (or norm) of $y = (y^1, y^2, \ldots\ldots, y^n)$. It is denoted by $| y |$ and defined as

$$| y | = | y^1 | + | y^2 | + \ldots\ldots\ldots + | y^n |$$

$$= \sum_{i=1}^{n} | y^i |.$$

We prefer to use this norm. The distance between two points y_1 and y_2 of R^n is defined to be $| y_2 - y_1 |$. Here, $| y |$ is a non – negative real number.

Note : Other definitions for the magnitude for a vector $y \in R^n$ are

$$| y | = \sqrt{\sum_{i=1}^{n} | y^i |^2},$$

$$| y | = \max. (| y^i |)$$

All are, of course, equivalent.

Result (Normed linear Space) : The linear space R^n is a normal linear space in which the norm function $| \, | : R^n \to R$, satisfies the following properties

(i) $| y | = 0$ if and only if $y = 0$;

(ii) $| y | \geq 0$;

(iii) $| y_1 + y_2 | \leq | y_1 | + | y_2 |$

(iv) $| \alpha y | = | \alpha | | y |$

for all y_1, y_2, y in R^n and $\alpha \in R$ is a scalar.

Definition (Metric Space): Let M be a non – empty set. A functions ρ: M x M $\to$ R is said to be a metric on M if it satisfies the following properties

(i) $\rho (x, y) = \rho (y, x)$

(ii) $\rho (x, y) = 0 \Leftrightarrow x = y$

(iii) $\rho (x, y) \leq \rho (x, z) + \rho (z, y)$

Then, the pair (M, ρ) is called a metric space or we simply say that M is a metric space with metric ρ.

Note : ρ is also termed as distance function.

Result : The normed linear space R^n is a metric space with metric ρ induced by its norm and defined as

$$\rho (y_1, y_2) = | y_1 - y_2 |.$$

Definition : A sequence of vectors $\{y_k\}$ in R^n is said to be convergent if it is convergent w. r. t. this distance function.

Remark : Sequence $\{y_k\}$ is convergent iff each of the component sequence $\{ y_k^i \}$, $1 \leq i \leq n$, is convergent.

Definition : A Banach space is a complete normed linear space, complete as a derived / induced metric space.

Definition : Lipschitz conditions in R^n

Suppose a vector function f(t, y) is defined on a domain D of the (t, y) space with $y \in R^n$ and $t \in I$. If there exists a constant $K > 0$ such that

$$| f(t, y_1) - f(t, y_2) | \leq K | y_1 - y_2 |$$

holds for every pair (t, y_1) and (t, y_2) in D, then the vector functions f (t,y) is said to satisfy a Lipschitz condition w. r. t. the variable y in D, and one write, as before,

$$f \in \text{Lip in D}.$$

Constant K is called Lipschitz constant.

Note :- The admissible values of K depend on the norm in the f – space as well as the norm in the y – space.

Definitions : $\in$ - approximate solution in R^n

Suppose $f \in C$ on a domain D in the (t, y) space. An $\in$-approximate solution of the vector differential equation

$$\frac{dy}{dt} = f(t, y); \; y \in R^n, \; t \in (a, b) = I$$

on an interval I is a vector functions $\phi(t) \in C$ on I such that

(i) $(t, \phi(t)) \in D$ for $t \in I$
(ii) $\phi \in C^1$ on I except possibly for a finite set of point S on I.
(iii) $| \phi'(t) - f(t, \phi(t) | \leq \in$ for $t \in I - S$.

Definition: A solution of vector differential equation on an interval I is a vector function $\phi = (\phi^1, \phi^2, \ldots, \phi^n)$ defined on I satisfying
(i) $(t, \phi(t)) \in D$ for $t \in I$
(ii) $\frac{d\phi}{dt} = f(t, \phi(t))$ for $t \in I$.

Definition: (Equi-continuous family)
Let $E \subset R^n$ be a subset. A family $F = \{f\}$ of functions f(y) defined on E is said to be equi-continuous if , for each $\in > 0$, there exists $\delta = \delta_\in > 0$ such that
$| f(y_1) - f(y_2) | \leq \in$
whenever $y_1, y_2 \in E$, $f \in F$ and $| y_1 - y_2 | \leq \delta$.

Observations : (1) The choice of $\delta_\in$ does not depend on f (y) but is valid for all admissible functions f (y) in the family F.
(2) The most frequently encountered equicontinuous families F will occur when all $f \in F$ satisfy the Lipschitz condition on the (t, y) space w.r.t. y. Here, there exists a Lipschitz constant K for all $f \in F$ and we may choose $\delta_\in = \frac{\in}{K}$, for given $\in > 0$.

Remarks: (1) In terms of the definitions introduced above, all the previous theorems discussed in chapters 1-3, are valid for the vector differential equation

$$\frac{dy}{dt} = f(t, y),\ y \in R^n$$

if, in their statements and proofs, y is replaced by the word "vector y", f is replaced by the word "vector function f" and the magnitude is understood in the sense of norm, defined above for vectors.
(2) The Ascoli theorem (2.2) is valid for vectors also. Therefore, it will be assumed from now onwards that these theorems stand proved for the more general vector differential equation

$$\frac{dy}{dt} = f(t, y),\ t \in I \text{ and } y \in R^n.$$

Now, we shall write the system of differential equations introduced earlier in vector notation. To achieve this, we introduce the vector y with components $y^i (i = 1, 2, \ldots, n)$, so that

$$y(t) = \begin{bmatrix} y^1(t) \\ y^2(t) \\ - \\ - \\ y^n(t) \end{bmatrix}.$$

The derivative of a vector valued function y(t) is $\frac{dy}{dt}$, and is defined to be the vector(or column-matrix/vector) whose components are $\frac{dy^i(t)}{dt}$, (i = 1, 2,….,n), so that

$$\frac{dy}{dt} = \begin{bmatrix} \frac{dy^1}{dt} \\ \frac{dy^2}{dt} \\ - \\ - \\ \frac{dy^n}{dt} \end{bmatrix}.$$

Similarly, we define the vector valued function f(t, y), a specified function of the vector y and real variable t, to be the vector whose components are $f^i(t, y)$, $1 \le i \le n$, so that

$$f(t, y) = \begin{bmatrix} f^1(t, y) \\ f^2(t, y) \\ - \\ - \\ f^n(t, y) \end{bmatrix}$$

It follows at once that first order system mentioned earlier can be written in the compact vector form

$$\frac{dy}{dt} = f(t, y).$$

Special case (Linear System) : When the first order system is linear, then functions $f^i(t, y^1, y^2, \ldots, y^n)$ are of the particular form

$$f^i(t, y^1, y^2, \ldots, y^n) = \sum_{k=1}^{n} a_{ik}(t)\, y^k,\ 1 \le i \le n$$

in which $a_{ik}(t)$ are continuous functions on I . In this case, this system is, in fact,

$$\frac{dy^1}{dt} = a_{11}(t)\, y^1 + a_{12}(t)\, y^2 + a_{13}(t)\, y^3 + \ldots\ldots a_{1n}(t)\, y^n .$$

$$\frac{dy^2}{dt} = a_{21}(t)\, y^1 + a_{22}(t)\, y^2 + a_{23}(t)\, y^3 + \ldots\ldots a_{2n}(t)\, y^n .$$

$$\vdots$$

$$\frac{dy^n}{dt} = a_{n1}(t)\, y^1 + a_{n2}(t)\, y^2 + a_{n3}(t)\, y^3 + \ldots\ldots a_{nn}(t)\, y^n .$$

This linear system can be put in the compact vector form

$$\frac{dy}{dt} = A(t)\, y$$

where the matrix A(t) is of type $n \times n$ and

$$A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) \ldots\ldots\ldots a_{1n}(t) \\ a_{21}(t) & a_{22}(t) \ldots\ldots\ldots a_{2n}(t) \\ - - - - & - - - - - - - - - - \\ a_{n1}(t) & a_{n2}(t) \ldots\ldots\ldots a_{nn}(t) \end{bmatrix}$$

and

$$y = \begin{bmatrix} y^1 \\ y^2 \\ - \\ - \\ y^n \end{bmatrix} \text{ with } \frac{dy}{dt} = \begin{bmatrix} \frac{dy^1}{dt} \\ \frac{dy^2}{dt} \\ - \\ - \\ \frac{dy^n}{dt} \end{bmatrix}.$$

In this case, f (t,y) satisfies a Lipschitz condition on the (n + 1) – dimensional region D, where $D = \{(t, y) \mid a \le t \le b,\ y \in R^n,\ |y| < \infty\}$.

(Here, D is not a domain, since it is not open). In fact, for (t, y_1) and (t, y_2) in D,

$|f(t, y_1) - f(t, y_2)| \le K\,|y_1 - y_2|$

$$\text{with } K = \max.\left\{\sum_{i=1}^{n} |a_{ik}(t)| : \begin{matrix} t \in [a, b] \\ k = 1,2,\ldots, n \end{matrix}\right\}$$

The above mentioned linear system is, therefore, expressible as

$$\frac{dy^i}{dt} = \sum_{k=1}^{n} a_{ik}(t)\, y^k,\ 1 \le i \le n.$$

Result I: For the linear system of differential equations

$$\frac{dy^\alpha}{dt} = \sum_{k=1}^{n} a_{\alpha k}(t)\, y^k,\ 1 \le \alpha \le n.$$

where the functions $a_{\alpha k}(t) \in C$ on [a, b], there exists one and only one solution $\phi(t)$ of this system on [a, b] passing through any point $(t_0, y_0) \in D$.

Result II: Let the functions $a_{ij}(t)$, (i, j = 1,2,….n), be continuous on an open interval I, which may be unbounded. Then there exists on I one and only one solution φ of the vector differential equation satisfying the initial condition $\phi(t_0) = y_0$, $t_0 \in I$, $|y_0| < \infty$.

Note : For the proofs of these results, the reader is advised to consult the book by Coddington and Levinson.

Reduction of nth order ODE to a first – order vector differential equation

This can be achieved by introducing variables to represent the derivatives appearing in the given nth order ODE.

Now, we consider a differential equation of nth order of the type

$$z^{(n)} = f(t, z, z^{(1)}, z^{(2)}, \ldots, z^{(n)})$$

Where $z^{(j)} = \frac{d^j z}{dt^j}$ (j = 1, 2, ..,n), t is a real variable on interval I, z and f are scalars (and not vector), and the function f is defined on a domain D of the real (n + 1)-dimensional space.

Problem : To find a function $\phi = \phi(t)$ defined on a real t-interval I possessing n derivatives there such that

(i) $(t, \phi(t), \phi^{(1)}(t), \phi^{(2)}(t), \ldots \phi^{(n-1)}(t)) \in D$ for all $t \in I$

(ii) $\phi^{(n)}(t) = f(t, \phi(t), \phi^{(1)}(t), \phi^{(2)}(t), \ldots \phi^{(n-1)}(t))$ for all $t \in I$.

If such an interval I and a function $\phi(t)$ exist, then $\phi(t)$ is said to be a solution of the given nth order differential equation on the interval I. If ϕ is a solution, clearly $\phi \in C^n$ on I. Note that ϕ is not a vector here.

Initial value problem :
Let $(t_0, z_0^1, z_0^2, \ldots, z_0^n) \in D$. Then the problem of finding a solution ϕ of the given nth order ODE on an interval I containing t_0 such that

$$\phi(t_0) = z_0^1,\ \phi^{(1)}(t_0) = z_0^2, \ldots.., \phi^{(n-1)}(t_0) = z^n$$

is called an initial – value problem.

Remark : The theory of the solution of nth order ODE can be reduced to the theory of a system of n first – order differential equations. For this, we make the substitutions

$$z = y^1$$
$$z^{(1)} = y^2$$
$$z^{(2)} = y^3$$
$$M$$
$$z^{(n-1)} = y^n$$

Then, we get the following system of n first order differential equations in n unknowns $y^1, y^2, \ldots.,$ y^n.

$$\frac{dy^1}{dt} = y^2$$
$$\frac{dy^2}{dt} = y^3$$
$$\ldots\ldots\ldots\ldots$$
$$\frac{dy^{n-1}}{dt} = y^n$$
$$\frac{dy^n}{dt} = f(t, y^1, y^2, \ldots., y^n) .$$

This system, in turn, is equivalent to the following first-order vector differential equation

$$\frac{dy}{dt} = f(t, y) ,$$

where

$$y = \begin{bmatrix} y^1 \\ y^2 \\ - \\ - \\ y^n \end{bmatrix}, \qquad f = \begin{bmatrix} y^2 \\ y^3 \\ - \\ - \\ y^n \\ f(t, y^1, y^2, \ldots., y^n) \end{bmatrix}.$$

Theorem (4.1) : Consider the nth order ordinary differential equation

$$\frac{d^n z}{dt^n} = f\left(t, z, \frac{dz}{dt}, \ldots, \frac{d^{n-1}z}{dt^{n-1}}\right)$$

where the function f is continuous and satisfies Lipschitz condition in a domain D of real (n + 1)-dimensional space. Let $(t_0, z_0^1, z_0^2, \ldots, z_0^n)$ be a point of D. Prove that there exists a unique solution $\phi(t)$ of the given nth order differential equation such that

$$\phi(t_0) = z_0^{\,1},\ \phi^{(1)}(t_0) = z_0^{\,2},\ldots\ldots,\phi^{(n-1)}(t_0) = z_0^{\,n},$$

defined on some interval around t_0.

Proof : Consider the substitutions

$$y^1 = z,\ y^2 = \frac{dz}{dt},\ y^3 = \frac{dz}{dt},\ldots\ldots,y^n = \frac{d^{n-1}z}{dt^{n-1}} \tag{1}$$

We shall now prove that, the given nth order differential equation is equivalent to the following system of n first order ordinary differential equations.

$$\frac{dy^1}{dt} = y^2$$

$$\frac{dy^2}{dt} = y^3$$

$$\vdots$$

$$\frac{dy^{n-1}}{dt} = y^n$$

$$\frac{dy^n}{dt} = f(t, y^1, y^2,\ldots\ldots,y^n). \tag{2}$$

Let $\phi = \phi(t)$ be a solution of the given nth order initial value problem. We define

$$\phi^1 = \phi,\ \phi^2 = \frac{d\phi}{dt},\ldots,\phi^n = \frac{d^{n-1}\phi}{dt^{n-1}}\ . \tag{3}$$

Let

$$y = \begin{bmatrix} \phi^1 \\ \phi^2 \\ - \\ - \\ \phi^n \end{bmatrix}, \tag{4}$$

be a vector of functions. Then y is a solution of the first order system (2) which satisfies the initial conditions

$$\phi^1(t_0) = z_0^{\,1},\ \phi^2(t_0) = z_0^{\,2},\ldots\ldots,\ \phi^m(t_0) = z_0^{\,m}\ . \tag{5}$$

Conversely, now it is assumed that a vector function $y = \begin{bmatrix} \phi^1 \\ \phi^2 \\ \vdots \\ \phi^n \end{bmatrix}$ is a solution of first order system (2) which satisfies initial conditions in (5). Then

$$\frac{d\phi^1}{dt} = \phi^2$$

$$\frac{d\phi^2}{dt} = \phi^3 = \frac{d^2\phi^1}{dt^2}$$

$$\frac{d\phi^3}{dt} = \phi^4 = \frac{d^3\phi^1}{dt^3}$$

M

$$\frac{d\phi^{n-1}}{dt} = \phi^n = \frac{d^{n-1}\phi^1}{dt^{n-1}} , \qquad (6)$$

and

$$\frac{d\phi^n}{dt} = f(t, \phi^1, \phi^2, \phi^3, \ldots, \phi^n)$$

or

$$\frac{d^n\phi^1}{dt^n} = f\left(t, \phi^1, \frac{d\phi^1}{dt}, \frac{d\phi^1}{dt}, \ldots, \frac{d^{n-1}\phi^1}{dt^{n-1}}\right). \qquad (7)$$

This shows that $y = \phi^1$ is a solution of the given nth order ordinary equation and this solution, using (5), satisfies the initial conditions

$$\phi^1(t_0) = z_0^1, \ \frac{d\phi^1}{dt}(t_0) = z_0^2, \ldots, \frac{d^{n-1}\phi^1}{dt^{n-1}}(t_0) = z_0^n \quad . \qquad (8)$$

This establishes the desired equivalence. This completes the proof.

Remark 1 : It also shows that the first component ϕ^1 of the vector

$$y = \begin{bmatrix} \phi^1 \\ \phi^2 \\ - \\ - \\ \phi^n \end{bmatrix}$$

is a solution of the given nth order ordinary differential equation .

Now $(t_0, z_0^1, z_0^2, \ldots, z_0^n)$ is a point of D, there exists an (n + 1)-dimensional rectangle, say R, about this point such that the function f satisfies the stated hypothesis in the rectangle R.

Thus, the system (2) of the first order equations satisfies all the hypothesis(necessary) in the rectangle R. So, there exists a unique solution $(\phi^1, \phi^2, \ldots, \phi^n)$ of system (2) which satisfies the conditions in (8) and this solution is defined on some sufficiently small interval around t_0. Thus, if we set

$$\phi = \phi^1,$$

the above shown equivalence gives the desired conclusion.

Remark 2: It is thus clear that all statements proved about the system (2) of n first-order equations carry over directly to statements about the nth order ordinary differential equation. In particular, we state the following theorem about linear equation.

Theorem (4.2) : Consider the linear ordinary differential equation

$$a_0(t)\frac{d^n y}{dt^n} + a_1(t)\frac{d^{n-1}y}{dt^{n-1}} + \ldots + a_{n-1}(t)\frac{dy}{dt} + a_n(t)\, y = F(t)$$

where $a_0, a_1, \ldots, a_n$ and F are continuous functions on the interval $a \le t \le b$ and $a_0(t) \ne 0$ on $a \le t \le b$. Let t_0 be a point of the interval $a \le t \le b$ and let $z_0^1, z_0^2, \ldots, z_0^n$ be n real constants. Prove that there exists a unique solution $\phi = \phi(t)$ of the above ODE such that
$\phi(t_0) = z_0^1, \phi^{(1)}(t_0) = z_0^2, \ldots, \phi^{(n-1)}(t_0) = z_0^n$.

Remark: Let $y = \begin{bmatrix} y^1 \\ y^2 \\ - \\ - \\ y^n \end{bmatrix}$, $A(t) = \begin{bmatrix} 0 & 1 & 0 \dots\dots\dots 0 & 0 \\ 0 & 0 & 1 \dots\dots\dots 0 & 0 \\ - & - & - - - - - - & - \\ 0 & 0 & 0 \dots\dots\dots 0 & 1 \\ \frac{-a_n}{a_0} & \frac{-a_{n-1}}{a_0} & \frac{-a_{n-2}}{a_0} \dots \frac{-a_2}{a_0} & \frac{-a_1}{a_0} \end{bmatrix}$,

$$Q = \begin{bmatrix} 0 \\ 0 \\ - \\ 0 \\ \frac{F(t)}{a_0(t)} \end{bmatrix}.$$

Then, the nth order ODE is equivalent to the following first order linear non-homogeneous vector differential equation

$$\frac{dy}{dt} = A(t)\, y + Q(t).$$

Definition: This vector equation is sometimes called the companion vector equation of the nth order scalar differential equation.

Note: For more details, readers are advised to consult the book by Ross, S.L.

5

MAXIMAL INTERVAL OF EXISTENCE

Suppose that the initial- valued problem

$$\frac{dy}{dt} = f(t, y),\ y(t_0) = y_0\ , \tag{1}$$

possesses a unique solution, say ϕ_0, on interval $| t - t_0 | \le h$, but nothing is implied about ϕ_0 outside this interval. Let

$$t_1 = t_0 + h,\ y_1 = \phi_0(t_1)\ . \tag{2}$$

Now, this point (t_1, y_1) is obviously a point of the rectangle R and it is also a point of the domain D in which the hypothesis of existence and uniqueness theorems are satisfied.

Thus we can reapply the existence and uniqueness theorems to conclude that the initial-value problem

$$\frac{dy}{dt} = f(t, y),$$

$$y(t_1) = y_1, \tag{3}$$

possesses a unique solution, say ϕ_1, which is defined on some interval $[t_1, t_1 + h_1]$, where $h_1 > 0$ (Fig. 5.1).

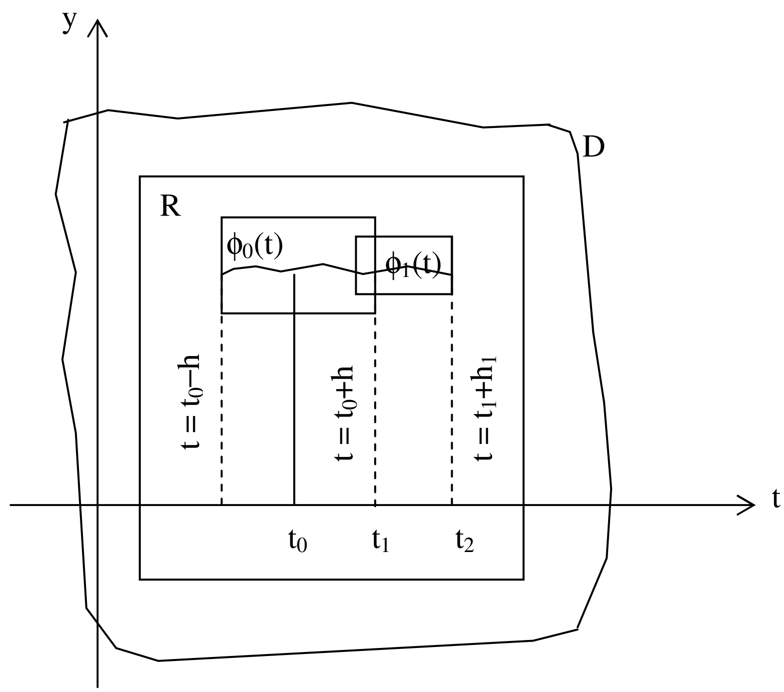


Fig. 5.1 : Continuation of solutions

Now let us define a function ϕ as follows:

$$\phi(t) = \begin{cases} \phi_0(t) & t_0 - h \le t \le t_0 + h = t_1 \\ \phi_1(t) & t_1 \le t \le t_1 + h_1 \end{cases} . \tag{4}$$

We now assert / claim that ϕ is a solution of the I V P (1) on the extended interval $[t_0 - h, t_1 + h_1]$. We note that the function $\phi(t)$ is continuous on this interval and is such that

$$\phi(t_0) = y_0. \tag{5}$$

For the interval $[t_0 - h, t_0 + h]$, I V P (1) imply

$$\phi_0(t) = y_0 + \int_{t_0}^{t} f(s, \phi_0(s))\, ds, \tag{6}$$

and hence

$$\phi(t) = y_0 + \int_{t_0}^{t} f(t, \phi(t))\, dt, \tag{7}$$

for all $t \in [t_0 - h, t_0 + h]$. On the interval $[t_0 + h, t_1 + h_1]$, we have

$$\phi_1(t) = y_1 + \int_{t_1}^{t} f(s, \phi_1(s))\, ds$$

or

$$\phi(t) = y_1 + \int_{t_1}^{t} f(s, \phi_1(s))\, ds , \tag{8}$$

using equation (3). From (7), we have

$$y_1 = \phi(t_1) = y_0 + \int_{t_0}^{t_1} f(s, \phi(s))\, ds . \tag{9}$$

From equations (8) and (9), we thus have

$$\phi(t) = y_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds, \tag{10}$$

on the interval $[t_0 + h, t_1 + h_1]$. Thus, combining the results (7) and (10), we see that ϕ satisfies the integral equation (7) on the extended interval $[t_0 - h, t_1 + h_1]$. Since ϕ is continuous on this interval, so is $f(s, \phi(s))$. Thus

$$\frac{d\phi(t)}{dt} = f(t, \phi(t)) , \tag{11}$$

on the extended interval $[t_0 - h, t_1 + h_1]$. Therefore, ϕ is a solution of the I V P (1) on this larger interval.

Definition: The solution ϕ, so defined, is called a **continuation of the solution** ϕ_0 to the interval $[t_0 - h, t_1 + h_1]$.

Remark : The process of continuation of a solution can be carried further. If we now apply the basic existence and uniqueness theorem again at the point $(t_1 + h_1, \phi(t_1 + h_1))$, we may thus obtains the continuation over the still longer interval $t_0 - h \le t \le t_2 + h_2$ where $t_2 = t_1 + h_1$ and $h_2 > 0$.

Repeating this process further, we may continue the solution over successively longer intervals $t_0 - h \le t \le t_n + h_n$, extending farther and farther to the right of $t_0 + h$. Also, in like manner, it may be continued over successively longer intervals extending farther and farther to the left of $t_0 - h$. Thus, repeating the process indefinitely on both the left and the right, we continue the solutions to successively longer intervals $[c_n, d_n]$, where $[t_0 - h, t_0 + h] = [c_0, d_0] \subset [c_1, d_1] \subset [c_2, d_2] \subset \ldots \subset [c_n, d_n] \subset \ldots$

Let $c = \lim c_n$, $d = \lim d_n$.
We thus obtain a largest open interval (c, d) over which the solution ϕ, satisfying the initial condition $\phi(t_0) = y_0$, may be defined.

Example : Consider the I V P

$$\frac{dy}{dt} = y^2,$$
$$y(-1) = 1 .$$

It has a solution $\phi(t) = \frac{-1}{t}$ through the point (-1, 1) and this solution eists on the interval [-1, 0] but cannot be continued further to the right. Because, in that case, ϕ does not stay within the region D, where $f(t, y) = y^2$ is bounded.

Maximal interval of existence:

Let f(t, y) be a continuous function on a (t, y)-set E. Let $\phi = \phi(t)$ be a solution of the differential equation

$$\frac{dy}{dt} = f(t, y) , \qquad (1)$$

on an interval I.
The interval I is called a right maximal interval of existence for ϕ if there does not exist an extension of $\phi(t)$ over an interval, say I_1, so that $\phi = \phi(t)$ remains a solution of (1), where I is a proper subset of I_1 and I, I_1 have different right end points.
A left maximal interval of existence for ϕ is defined similarly.

Definition: A maximal interval of existence of a solution of ODE (1) is an interval which is both a left and right maximal interval.

Kneser′s theorem (without proof): This theorem is about the case of non-unique solutions of initial value problems.

Statement: Let f(t, y) be continuous on the rectangle R,

$$R = \{(t, y) : t_0 \le t \le t_0 + a, | y - y_0 | \le b. a > 0, b > 0\}.$$

Let $| f(t, y) | \le M$ and $\alpha = \min(a, \frac{b}{M})$, and $t_0 \le c \le t_0 + \alpha$. Let S_c be the set of points y_c for which there is a solution y = y(t) of the initial value problem.

$$\frac{dy}{dt} = f(t, y), \; y(t_0) = y_0 \qquad \text{on } [t_0, c]$$

such that

$$y(c) = y_c.$$

Then the set S_c is a closed connected set.

Kamke′s Convergence theorem(Statement only):
Let f(t, y) be continuous for $t_0 \le t \le t_0 + a$ and all y. Let $t_0 < c \le t_0 + a$ and assume that all solutions y(t) of initial value problem

$$\frac{dy}{dt} = f(t, y) , \; y(t_0) = y_0$$

exist on $t_0 \leq t \leq c$. Then, the set S_c is a continuum.

Books recommended for reading (for Unit I : Chapters 1-5) are

1. E. A. Coddington and N. Levinson
Theory of ODE, McGraw Hill, NY(1955)
2. S. L. Ross
Differential Equations, John Wiley & Sons, Third Edition, 1984.
3. G. Birkhoff and G. C. Rota
ODE, 3^{rd} edition, John Wiley & Sons , NY, 1978.
4. E. L. Ince.
ODE, Dover Publications.
5. P. Hertman
ODE, John Wiley & Sons, NY, 1964
6. G. F. Simmons
Differential Equations with Applications and Historical Notes, McGraw Hill, 1991.

6

DEPENDENCE OF SOLUTIONS ON INITIAL CONDITIONS AND PARAMETERS

Dependence of Solutions on Initial Conditions

Consider the first order I V P

$$\frac{dy}{dt} = y,$$
$$y(t_0) = y_0. \qquad (1)$$

It has the solution (exercise, to obtain it)

$$\phi(t) = y_0\ e^{t-t_0}\ , \qquad (2)$$

which passes through the point (t_0, y_0). The functions ϕ in (2) can be considered as function, not only $t \in I$, but of the coordinates of point (t_0, y_0), through which the solution curve passes. The solution function ϕ in (2), without any confusion /ambiguity can be written as

$$\phi(t, t_0, y_0) = y_0\ e^{t-t_0}\ . \qquad (3)$$

Now, we shall investigate the behavior of the solutions as functions of the initial conditions for the general problem.

Let $f(t, y)$ be continuous and satisfy a Lipschitz condition w. r. t. ′y′ in a domain D. Let (t_0, y_0) be a fixed point of D. Now, by Picard's existence and uniqueness theorem (3.2), the initial value problem

$$\frac{dy}{dt} = f(t, y),$$
$$y(t_0) = y_0\ , \qquad (1)$$

has a unique solution ϕ defined as some sufficiently small interval $[t_0 - h_0, t_0 + h_0]$ around t_0 . Now suppose that the initial y-value is changed from y_0 to Y_0. Our first concern is whether or not the new initial – value problem

$$\frac{dy}{dt} = f(t, y)\ ,$$
$$y(t_0) = Y_0\ , \qquad (2)$$

also has a unique solution on some sufficiently small interval $| t - t_0 | \le h_1$. If Y_0 is such that $| Y_0 - y_0 |$ is sufficiently small, then we can be certain that the problem (2) does possess a unique solution on some such interval $| t - t_0 | \le h_1$. In fact, let the rectangle R: $| t - t_0 | \le a$, $| y - y_0 | \le b$, lie in D and let Y_0 be such that $| Y_0 - y_0 | \le b/2$. Then, by Picards theorem, this problem has a unique solution ψ which is defined and contained in R for $| t - t_0 | \le h_1$, where $h_1 = \min(a, b/2M)$ and $M = \max | f(t, y) |$ for $(t, y) \in R$. Thus we may assume that there exists $\delta > 0$ and $h > 0$ such that for each Y_0 satisfying $| Y_0 - y_0 | \leq \delta$, problem (2) possesses a unique solution $\phi(x, Y_0)$ on $|t - t_0| \leq h$ (see Figure 6.1 below).

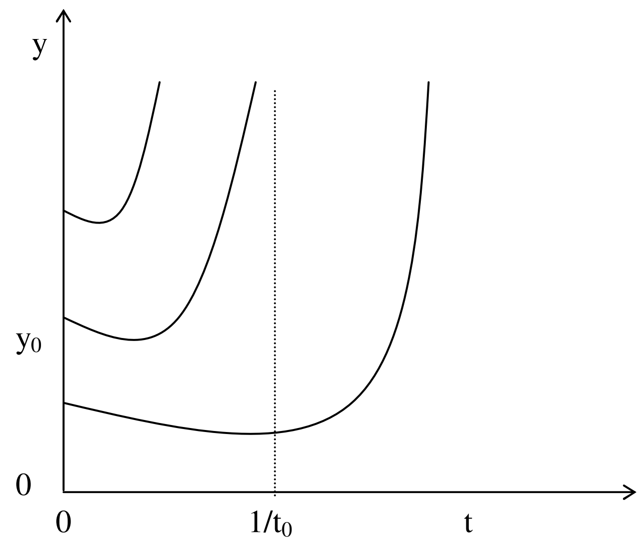


Fig.6.1 Plots of solution for different values of y_0

To illustrate this, consider the I VP

$$\frac{dy}{dt} = t\, y^3,$$
$$y(0) = y_0\ ,$$

Its solution is

$$\phi(t) = y_0(1 - y_0^2\, t^2)^{-1/2}\ ,$$

and is only defined for the interval $| t | < | y_0 |^{-1}$. Here, y_0 can be regarded as an arbitrary constant and, as y_0 varies, the solutions fill the entire t – y plane. The general solution is shown in the following figure (6.2). Nevertheless, for each particular value y_0, the corresponding unique solution is defined only over an interval whose size depends on y_0.

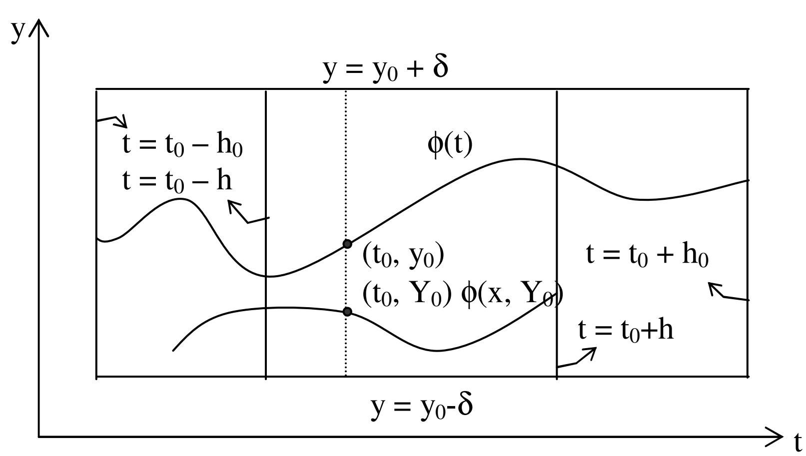


Fig. 6.2

We are now in a position to state the basic theorem concerning the dependence of solutions on initial conditions.

Theorem 6.1. Let f be continuous and satisfy a Lipschitz condition with respect to y, with Lipschitz constant k, in a domain D of the ty plane; and let (t_0, y_0) be a fixed point of D. Assume there exists $\delta > 0$ and $h > 0$ such that for each Y_0 satisfying $| Y_0 - y_0 | \leq \delta$ the I VP

$$\frac{dy}{dt} = f(t, y),$$
$$y(t_0) = Y_0,$$

possesses a unique solution $\phi(t, Y_0)$ defined and contained in D on $| t - t_0 | \leq h$. Let ϕ denote the unique solution of I VP when $Y_0 = y_0$, and $\tilde{\phi}$ denotes the unique solution of I VP when $Y_0 = \tilde{y}_0$, where $| \tilde{y}_0 - y_0 | = \delta_1 \leq \delta$. Prove that

$$| \tilde{\phi}(t) - \phi(t) | \leq \delta_1 e^{kh} \text{ on } | t - t_0 | \leq h.$$

Proof. From Picards theorem, we know that

$$\phi = \lim_{n\to\infty} \phi_n, \tag{1}$$

where

$$\phi_n(t) = y_0 + \int_{t_0}^{t} f[t, \phi_{n-1}(t)]dt \quad (n = 1, 2, 3,\ldots\ldots),$$

and $\phi_0(t) = y_0;\ | t - t_0 | \leq h$.
In like manner,

$$\tilde{\phi} = \lim_{n\to\infty} \tilde{\phi}_n, \tag{3}$$

where

$$\tilde{\phi}_n(t) = \tilde{y}_0 + \int_{t_0}^{t} f[t, \tilde{\phi}_{n-1}(t)]dt \quad (n = 1, 2, 3,\ldots\ldots), \tag{4}$$

and $\tilde{\phi}_0(t) = \tilde{y}_0;\ | t - t_0 | \leq h$.
We shall show by induction that

$$| \tilde{\phi}_n(t) - \phi_n(t) | \leq \delta_1 \sum_{j=0}^{n} \frac{K^j (t-t_0)^j}{j} \tag{5}$$

on $[t_0, t_0 + h]$, where K is the Lipschitz constant. We thus assume that on $[t_0, t_0 + h]$,

$$| \tilde{\phi}_{n-1}(t) - \phi_{n-1}(t) | \leq \delta_1 \sum_{j=0}^{n-1} \frac{K^j (t-t_0)^j}{j} \tag{6}$$

Then

$$| \tilde{\phi}_n(t) - \phi_n(t) = | \tilde{y}_0 + \int_{t_0}^{t} f[t, \tilde{\phi}_{n-1}(t)]dt - y_0 - \int_{t_0}^{t} f[t, \phi_{n-1}(t)]dt |$$
$$\leq | \tilde{y}_0 - y_0| + \int_{t_0}^{t} | f[t, \tilde{\phi}_{n-1}(t)] - f[t, \phi_{n-1}(t)] | dt.$$

Applying the Lipschitz condition, we have

$$| f[t, \tilde{\phi}_{n-1}(t)] - f[t, \phi_{n-1}(t)] | \leq K | \tilde{\phi}_{n-1}(t) - \phi_{n-1}(t);$$

and so, since

$$| \tilde{y}_0 - y_0 | = \delta_1,$$

therefore,

$$| \tilde{\phi}_n(t) - \phi_n(t) | \leq \delta_1 + k \int_{t_0}^{t} | \tilde{\phi}_{n-1}(t) - \phi_{n-1}(t) | dt. \tag{7}$$

Using the assumption (6), we have

$$| \tilde{\phi}_n(t) - \phi_n(t) | \le \delta_1 + k \int_{t_0}^{t} \delta_1 \sum_{j=0}^{n-1} \frac{K^j (t-t_0)^j}{j!} dt$$

$$= \delta_1 + k\delta_1 \sum_{j=0}^{n-1} \frac{k^j}{j!} \int_{t_0}^{t} (t - t_0)^j dt = \delta_1 [1 + \sum_{j=0}^{n-1} \frac{k^{j+1}(t-t_0)^{j+1}}{(j+1)!}]. \qquad (8)$$

Since

$$\delta_1 [1 + \sum_{j=0}^{n-1} \frac{K^{j+1}(t-t_0)^{j+1}}{(j+1)!}] = \delta_1 \sum_{j=0}^{n} \frac{K^j (t-t)^j}{j!}, \qquad (9)$$

we have

$$| \tilde{\phi}_n(t) - \phi_n(t) | \le \delta_1 \sum_{j=0}^{n} \frac{K^j (t-t_0)^j}{j!}, \qquad (10)$$

which is (6) with (n –1) replaced by n.
Also, on $[t_0, t_0 + h]$, we have

$$| \tilde{\phi}_1(t) - \phi_1(t) | = | \tilde{y}_0 + \int_{t_0}^{t} f[t, \tilde{y}_0]dt - y_0 - \int_{t_0}^{t} f[t, y_0]\, dt |$$

$$\le | \tilde{y}_0 - y_0 | + \int_{t_0}^{t} | f[t, \tilde{y}_0] - f[t, y_0] | dt$$

$$\le \delta_1 + \int_{t_0}^{t} K | \tilde{y}_0 - y_0 | dt$$

$$= \delta_1 + K \delta_1 (t - t_0). \qquad (11)$$

Thus (10.23) holds for n = 1. Hence the induction is complete and (5) holds on $[t_0, t_0 + h]$. Using similar arguments on $[t_0 - h, t_0]$, we have

$$| \tilde{\phi}_n(t) - \phi_n(t) | \le \delta_1 \sum_{j=0}^{n} \frac{K^j (t-t_0)^j}{j!}$$

$$\le \delta_1 \sum_{j=0}^{n} \frac{(Kh)^j}{j!}$$

for all t on $| t - t_0 | \le h$, n = 1, 2, 3, …. Letting $n \to \infty$, we have

$$| \tilde{\phi}(t) - \phi(t) | \le \delta_1 \sum_{j=0}^{\infty} \frac{(Kh)^j}{j!} \qquad (12)$$

But $\sum_{j=0}^{\infty} \frac{(Kh)^j}{j!} = e^{Kh}$; and so we have obtained the desired inequality

$$| \tilde{\phi}_1(t) - \phi_1(t) | \le \delta_1 = e^{Kh} \text{ on } | t - t_0 | \le h. \qquad (13)$$

This completes the proof of the theorem.

Cor : The solution $\phi(t, Y_0)$ of I V P is a continuous functions of the initial value Y_0 at $Y_0 = y_0$.

Proof :- It follows immediately from the results of the above theorem.

Remark : Thus under the conditions stated, if the initial values if the two solutions ϕ and $\tilde{\phi}$ differ by sufficiently small amount, then their values will differ by an arbitrarily small amount at every point of $| t - t_0 | \le h$. Geometrically, this means that if the corresponding integral curves are sufficiently close to each other initially, then they will be arbitrarily close to each other for all t such that $| t - t_0 | \leq h$.

Dependence on Parameters:

In many practical problems the dynamical system which the differential equation

$$\frac{dy}{dt} = f(t, y) \qquad (1)$$

describes contains external parameters, as well as the dependent variable y.

Assume that the right side member of the above ODE contains a parameter vector μ and μ-space has m real dimensions. Let D_μ be the domain of μ - space for which

$$| \mu - \mu_0 | < c \qquad (2)$$

where μ_0 is fixed and $c > 0$. The differential equation

$$\frac{dy}{dt} = f(t, y, \mu) \qquad (3)$$

will now be considered. Here, $f(t, y, \mu)$ is required to be a continuous function of t, y and μ for $(t, y) \in D$ and $\mu \in D_\mu$. The initial condition for the above differential equation is again the same, i.e.,

$$y(t_0) = y_0. \qquad (4)$$

Now, for each fixed value of μ, say $\mu = \mu_0$, there exists a unique solution which, however, depends on μ_0 as well. In order to incorporate the dependence of the solutions on initial conditions t_0 and y_0, as well as on μ, we adopt the device of allowing t_0 and y_0 to depend on μ so that

$$t_0 = t_0(\mu) \text{ and } y_0 = y_0(\mu) \qquad (5)$$

are continuous functions of μ. Before to proceed further, we extend the notion of the Lipschitz condition as follows:

$$| f(t, y_1; \mu) - f(t, y_2; \mu) | \leq L| y_1 - y_2 | \qquad (6)$$

uniformly for all (t, y_1), $(t, y_2) \in D$ and $\mu \in D_\mu$, L being a Lipschitz constant. The constant L is independent of t, y_1, y_2 and μ.

Theorem 6.2 : Let $f(t, y; \mu)$ be continuous for

$$| t - t_0 | \leq a, \quad | y - y_0 | \leq b, \quad | \mu - \mu_0 | \leq c$$

and satisfy a Lipschitz condition with constant L in this region. Let $y_0(\mu)$ and $t_0(\mu)$ be continuous functions of μ for $| \mu - \mu_0 | \leq c$ such that $y_0(\mu_0) = y_0$ and $t_0(\mu_0) = t_0$. Then there exists a $\delta > 0$ and $\in > 0$ such that the initial – value problem

$$\frac{dy}{dt} = f(t, y; \mu),$$

$$y(t_0(\mu)) = y_0(\mu) \qquad (1)$$

has a unique solution for $| t - t_0 | \leq \delta$ and $| \mu - \mu_0 | \leq \in$ which is a continuous function of t and μ.

Proof :- The initial – value problem (1) is equivalent to the following integral equation,

$$y(t; \mu) = y_0(\mu) + \int_{t_0(\mu)}^{t} f(y(s; \mu), s; \mu)\, ds. \qquad (2)$$

The proof is analogous to that of previous theorem (6.1). Thus, we define the iterates

$$\phi_0(t;\mu) = y_0(\mu),\ldots,\quad \phi_{n+1}(t;\mu) = y_0(\mu) + \int_{t_0(\mu)}^{t} f(\phi_n(s;\mu), s;\mu)\,ds. \quad (3)$$

n = 1,2,….

The first step is to choose regions over which t and μ can vary such that each iterate remains with in the domain of definition of $f(t, y; \mu)$. Let

$$\delta = \frac{1}{2}\min\{a, b/M\}, \quad (4)$$

where M is such that

$$|f(t, y;\mu)| \le M \text{ for } |t - t_0| \le a,\ |y - y_0| \le b \text{ and } |\mu - \mu_0| \le c. \quad (5)$$

Then choose $\in$ $(0 < \in\ \le c)$ so that, for

$$|\mu - \mu_0| \le \in.$$

We must have

$$|t_0(\mu) - t_0| < \frac{1}{2}\delta \text{ and } |y_0(\mu) - y| < \frac{1}{4}b. \quad (6)$$

It follows that, if $|t - t_0| < \delta$ then

$$|t - t_0(\mu)| \le |t - t_0| + |t_0(\mu) - t_0|$$
$$\le \frac{3}{2}\delta. \quad (7)$$

Next, it may be shown easily that

$$|\phi_{n+1}(t;\mu) - \phi_0(\mu)| \le M\,|t - t_0(\mu)|$$
$$\le \frac{3}{4}b,$$

for $|t - t_0(\mu)| \le \frac{3}{2}\delta$.

But then, for $|t - t_0| \le \delta$, it follows that

$$|\phi_{n+1}(t;\mu) - y_0| \le |\phi_{n+1}(t;\mu) - y_0(\mu)| + |y_0(\mu) - y_0| \le b.$$

Thus each iterate is well defined for $|t - t_0| \le \delta$ and $|\mu - \mu_0| \le \in$. Next, it can be shown that

$$|\phi_{n+1}(t;\mu) - \phi_n(t;\mu)| \le M\,L^n\,\frac{|t - t_0(\mu)|^{n+1}}{(N+1)!}, \quad (8)$$

and, hence, the sequence $\phi_{n+1}(t;\mu)$ is uniformly convergent to a limit function $\phi(t;\mu)$. The key new result here is that the continuity if $f(t, \mu; \mu)$ in t , y and μ, and the continuity if $y_0(\mu)$ and $t_0(\mu)$, ensures that each iterate is a continuous function of t and μ for $|t - t_0| \le \delta$ and $|\mu - \mu_0| \le \in$. Hence, since the convergence is uniform, $\phi(t;\mu)$ is also a continuous function of t and μ. The rest of the proof, that $x(t, \mu)$ solves the initial value problem, now follows immediately.

This completes the proof of the theorem.

7

DIFFERENTIAL INEQUATIONS

In the following r, u, v, U, V are scalars while y, z, f, g are n – dimensional vectors. Now, we state and prove Gronwall's inequality, which is one of the simplest and most useful results involving an integral inequation.

Theorem7.1 (GRONWALL'S INEQUALITY).

Statement: Let u(t) and v(t) be non – negative, continuous functions defined on closed interval [a, b]. Let $c \geq 0$ and

$$v(t) \leq c + \int_a^t v(s)\, u(s)\, ds, \qquad \text{for } a \leq t \leq b$$

then

$$v(t) \leq c \exp\left[\int_a^t u(s)ds\right] \qquad \text{for } a \leq t \leq b.$$

and, in particular, if c = 0, then $v(t) \equiv 0$.

Proof :- Case I : When c > 0.

Let $$V(t) = c + \int_a^t v(s)\, u(s)\, ds . \tag{1}$$

Then $$v(a) = c \qquad (2),$$

and by hypothesis

$$v(t) \leq V(t) , \tag{3}$$

and

$$V(t) \geq c > 0 \quad \text{on } [a, b], \tag{4}$$

as u and v are non – negative functions. Also, from (1), we have, on [a, b],

$$V'(t) = v(t)\, u(t)$$

$$\leq V(t)\, u(t),$$

using (3). This implies, using (4),

$$\frac{V'(t)}{V(t)} \leq u(t). \qquad (5)$$

Integrating (5) over [a, t], we get

$$V(t) \leq c \exp[\int_a^t u(s)\, ds]$$

or $$V(t) \leq V(t) \leq c\left[\exp\int_a^t u(s)ds\right]$$

or $$v(t) \leq c \exp\left[\int_a^t u(s)ds\right]. \qquad (6)$$

This proves the result.

Case II : When c = 0. Letting $c \to 0^+$ in (6), we get the desired result.
Restatement : Another form of Gronwall′s inequality is given below.

Statement: Let r(t) be continuous for $|t - t_0| \leq \delta$ and satisfy the inequalities

$$0 \leq r(t) \leq \in + \delta \left|\int_{t_0}^t r(s)\, ds\right|$$

for some non-negative constants $\in$ and δ. Then

$$0 \leq r(t) \leq \in \exp\{\delta|t - t_0|\}.$$

Proof: On taking $c = \in$, $t_0 = a$ and $u(t) = \delta$ in theorem 7.1, the result follows immediately.

Cor. 1. Let f(t, y) satisfy a Lipschitz condition with constant L for $y \in D$ and $|t - t_0| \leq \delta$. Let y(t) and z(t) be solutions of problem

$$\frac{dy}{dt} = f(t, y)$$

for $|t - t_0| \leq \delta$ such that

$$y(t_0) = y_0, \qquad z(t_0) = z_0,$$

where $y_0, z_0 \in D$. Then

$$|y(t) - z(t)| \leq |y_0 - z_0| \exp\{L|t - t_0|\}.$$

Proof :- In the integral equation formulation, it follows that

$$z(t) = z_0 + \int_{t_0}^t f(s, z(s))\, ds, \qquad y(t) = y_0 + \int_{t_0}^t f(s, y(s))\, ds. \qquad (1)$$

Subtracting, we see that

$$z(t) - y(t) = z_0 - y_0 + \int_{t_0}^t [f(s, z(s)) - f(s, y(s))]\, ds. \qquad (2)$$

Taking the norm of both sides and applying the Lipschitz condition, it follows that

$$0 \le | z(t) - y(t) |$$
$$\le | z_0 - y_0 | + \left| \int_{t_0}^{t} L | z(s) - y(s) | ds \right| . \qquad (3)$$

From Gronwall's inequality (7.1), with $r(t) = | z(t) - y(t) |$, $\in = | x_0 - y_0 |$ and $\delta = L$, the result follows immediately.

Cor. 2. Let f(t,y) satisfy a Lipschitz condition for $y \in D$ and $| t - t_0 | \le \delta$. Then the initial value problem has a unique solution, that is, there is at most one continuous function y(t) which satisfies

$$\frac{dy}{dt} = f(t, y),$$
$$y(t_0) = y_0.$$

Proof :- Putting $z_0 = y_0$ in Cor. 1, we see that $z(t) = y(t)$ for all $| t - t_0 | \le \delta$, thus establishing the uniqueness of the initial value problem, whenever f(t, y) satisfies a Lipschitz condition. Hence the result.

Note : Cor.1 also shows that the solutions of the initial value problem are continuous in the initial data, since it follows that if $y_0 \to z_0$, then $y(t) \to z(t)$ uniformly for all $| t - t_0 | \le \delta$. Thus, proveded the Lipschitz condition holds, the initial value problem is well–set.

The comparison Theorems:

Since most differential equations can not be solved in terms of elementary functions, it is important to be able to compare the unknown solutions of one differential equation with the known solutions of another. The following theorems give such comparisons.

Theorem 7.2. Let f(t, y) satisfy a Lipschitz condition for $t \ge a$. If the function $u = u(t)$ satisfies the differential inequality

$$\frac{dy}{dt} \le f(t, y) \quad \text{for } t \ge a \qquad (1)$$

and $v = v(t)$ is a solution of differential equation

$$\frac{dy}{dt} = f(t, y) \qquad (2)$$

satisfying the initial conditions

$$u(a) = v(a) = c_0 \qquad (3)$$

then $u(t) \le v(t)$ for $t \ge a$. (4)

Proof :- If possible, suppose that

$$u(t_1) > v(t_1) \qquad (5)$$

for some t_1 in the given interval. Let t_0 be the largest t in the interval $[a, t_1]$ such that

$$u(t) \le v(t).$$

Then

$$u(t_0) = v(t_0) . \qquad (6)$$

Let $\sigma(t) = u(t) - v(t).$ (7)

Then $\sigma(t_0) = 0, \quad \sigma(t_1) > 0$ (8A)

and $\sigma(t) \ge 0$ for $[t_0, t_1]$. (8B)

Also for $t_0 \le t \le t_1$,

$$\sigma'(t) = u'(t) - v'(t)$$
$$\le f(t, u(t)) - f(t, v(t)), \quad \text{using (1) \& (2)}$$
$$\le K| u(t) - v(t) |,$$
$$= K\,\sigma(t), \tag{8C}$$

where K is the Lipschitz constant for the function f.

Multiplying both sides of (8C) by e^{-Kt}, we write

$$0 \ge e^{-Kt}.\{\sigma'(t) - K\,\sigma(t)\}$$
$$= \frac{d}{dt}\{\sigma(t).\ e^{-Kt}\}\ .$$

This implies

$$\frac{d}{dt}\{\sigma(t).\ e^{-Kt}\} \le 0 \qquad \text{in } [t_0, t_1] \tag{9}$$

So, $\sigma(t).\ e^{-Kt}$ is a decreasing function for $[t_0, t_1]$.

Therefore $\quad \sigma(t).\ e^{Kt} \le \sigma(t_0).\ \ e^{-Kt_o} \qquad$ for all t in $[t_0, t_1]$

$\Rightarrow \sigma(t) \le \sigma(t_0)\ e^{K(t_o - t)}$

$\Rightarrow \sigma(t) \le 0 \qquad$ for all t in $[t_0, t_1]$, using (8A)

$\Rightarrow \sigma(t)$ vanishes identically zero in $[t_0, t_1]$.

This contradicts the assumption in (5) that $\sigma(t_1) > 0$. Hence, we conclude that

$$u(t) \le v(t)$$

for all t in the given interval.
This completes the proof.

Theorem 7.3. (Comparison Theorem).
Let u = u(t) and v = v(t) be solutions of differential equations

$$\frac{dy}{dt} = U(t, y),\ \frac{dz}{dt} = V(t, z) \tag{1}$$

respectively, where

$$U(t, y) \le V(t, y) \tag{2}$$

in the strip $a \le t \le b$ and U or V satisfies a Lipschitz conditions, and

$$u(a) = v(a). \tag{3}$$

Then $\quad u(t) \le v(t) \qquad$ for all $t \in [a, b]$. (4)

Proof :- Let V satisfy a Lipschitz conditions. Since

$$\frac{dy}{dt} = U(t, y) \le V(t, y),$$

the functions u(t) and v(t) satisfy the conditions of theorem 7.2 with V in place of f. Therefore, the inequality (4) follows immediately.

If U satisfies a Lipschitz condition, the functions

$$f(t) = -u(t), \quad g(t) = -v(t) \tag{5}$$

satisfies the differential equations

$$\frac{du}{dt} = -U(t, -u)\ ,$$

and

$$\frac{dv}{dt} = -V(t, -v)$$
$$\leq - U(t, -v), \qquad \text{using (2)} . \qquad (6)$$

Theorem 7.3 now yields the inequality

$$g(t) \leq \ f(t) \qquad \text{for } t \geq a$$
$$\Rightarrow -g(t) \geq -f(t) \qquad \text{for } t \geq a$$
$$\Rightarrow v(t) \geq u(t) \qquad \text{for } t \geq a$$
$$\Rightarrow u(t) \leq u(t) . \qquad \text{for } t \geq a$$

This completes the proof.

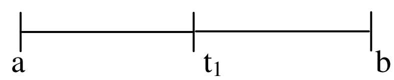
Remark : The inequality $u(t) \leq v(t)$ in this comparison theorem 7.3 can often be replaced by a strict inequality.

Corollary 1: In theorem 7.3, for any $t_1 > a$, either

$$u(t_1) < v(t_1) \quad \text{or} \quad u(t) \equiv v(t) \quad \text{for } a \leq t \leq t_1.$$

Proof :- By theorem 7.3, $u(t) \leq v(t)$ for all $t > a$. (1)

Let $t_1 > a$ be aby value of t.



If $u(t_1)$ is not less than $v(t_1)$, then $u(t_1) = v(t_1)$. (2)

Then, either u and v are identically equal for $a \leq t \leq t_1$, or else

$$u(t_0) < v(t_0) \qquad (3)$$

for some t_0 in the interval (a, t_1).

Let $\sigma_1(t) = v(t) - u(t)$ (4)

for $t \in [a, t_1]$. Then

$\sigma_1(t_0) > 0,$ (5)

and by theorem 7.3,

$$u(t) \leq v(t) \qquad \text{for } t \in [a, t_1]$$
$$\Rightarrow \sigma(t) \geq 0 . \qquad \text{for } t \in [a, t_1] \qquad (6)$$

Further, for $t \in [a, t_1]$,

$$\sigma_1'(t) = v'(t) - u'(t)$$
$$= V(t, v(t)) - U((t, u(t))$$
$$\geq V(t, v(t)) - V(t, u(t)) \quad (\Theta\, U \leq V \text{ given})$$
$$\geq - K\{v(t) - u(t)\}$$
$$\Rightarrow \sigma_1'(t) \geq - K\, \sigma_1$$
$$\Rightarrow (\sigma_1' + K\, \sigma_1) \geq 0 . \qquad (7)$$

Hence

$$\{e^{Kt}.\sigma_1(t)\}' = e^{Kt}\{\sigma_1'(t) + K\, \sigma_1(t)\} \geq 0,$$

using (7) for $t \in [a, t_1]$. This shows that the function $\phi(t) = e^{kt} + \sigma_1(t)$ is an increasing function on the interval $[a, t_1]$. So

$$\phi(t) \geq \phi(t_0) \qquad \text{for } t \in [t_0, t_1,]$$
$$\Rightarrow e^{kt}\, \sigma_1(t) \geq e^{Kt_0}\, \sigma_1(t_0)$$
$$\Rightarrow \quad \sigma_1(t) \geq \sigma_1(t_0)\ e^{-K(t-t_0)} > 0$$

$\Rightarrow \quad \sigma_1(t) > 0$ in $[t_0, t_1]$

$\Rightarrow v(t) - u(t) > 0$ in $[t_0, t_1]$

$\Rightarrow \quad v(t) > u(t)$ in $[t_0, t_1] \leq [a, b]$

which is a contradictions. Hence, u and v are identical for $a \leq t \leq t_1$. This completes the proof.

Cor 2. In theorem 7.3, assume that U, as well as V, satisfies a Lipschitz conditions and, instead of $u(a) = v(a)$, that $u(a) < v(a)$.
Then

$$u(t) < v(t) \qquad \text{for } t > a.$$

Proof :- The proof will be by contradiction.
If we had $u(t) \geq v(t)$ for some $t > a$, there would be a first $t = t_1 > a$, where

$$u(t) \geq v(t). \tag{1}$$

We define two functions

$$y = \phi(t) = u(-t),$$
$$z = \psi(t) = v(-t) . \tag{2}$$

Then ϕ and ψ satisfy the differential equations

$$\frac{dy}{dt} = -U(-t, y), \ \frac{dz}{dt} = -V(-t, z) \tag{3}$$

as well as the respective initial conditions, respectively,

$$\phi(-t_1) = \psi(-t_1). \tag{4}$$

Since

$$-U(-t, y) \geq -V(-t, y) . \tag{5}$$

We can apply theorem 7.3 in the internal $[-t_1, -a]$, knowing that the function $-U(-t, y)$ satisfies a Lipschitz condition. So, by theorem 7.3, we conclude that

$$\phi(t) \geq \psi(t) \text{ in } [-t_1, -a]$$
$$\Rightarrow \phi(-a) \geq \psi(-a)$$
$$\Rightarrow u(a) \geq v(a), \tag{6}$$

a contradiction. Hence, by contradiction, assertion of the corollary 2 holds. That is,

$$u(t) < v(t) \qquad \text{for } t > a. \tag{7}$$

Hence, the result.

8

MAXIMAL & MINIMAL SOLUTIONS, LYAPUNOV FUNCTIONS

Let U(t, u) be a continuous function on a plane (t, u) set E. by a maximal solution $u = u_M(t)$ of the initial value problem

$$\frac{du}{dt} = U(t, u),$$

$$u(t_0) = u_0 , \qquad (1)$$

is meant a solution of above I V P on a maximal interval of existence such that if u(t) is any solution of (1), then

$$u(t) \leq u_m(t) \qquad (2)$$

holds on the common interval of existence of u, u_M.

Similarly, a solution $u = u_m(t)$ of I V P (1), defined on a maximal interval of existence, is called a minimal solution of it if

$$u(t) \geq u_M(t) , \qquad (3)$$

for every other solution u(t) of I V P (1), and inequality (3) holds on the common interval of existence of u, u_m.

Note: (1) Every solution must remain between a maximal solution and a minimal solution.

(2) Sometimes the maximal and minimal solutions are the same over an interval. Then the solution is unique over any interval where this occurs.

Notation. Let f(t, y) be continuous on an open (t, y)- set Ω. Let u(t, y) be a real valued function defined in a vicinity of a point $(t_1, y_1) \in \Omega$. Let y(t) be a solution of the system

$$\frac{dy}{dt} = f(t, y) \qquad (1)$$

satisfying initial condition

$$y(t_1) = y_1. \qquad (2)$$

If u(t, y(t)) is differentiable at $t = t_1$, this derivative is **called the trajectory derivative of u(t, y)** at the point (t_1, y_1) along the orbit y = y(t) and is **denoted by** $\dot{u}(t_1, y_1)$.

Remark. When u(t, y) has continuous partial derivatives, its trajectory derivative exists and can be calculated without finding solutions of (1). In fact,

$$\dot{u}(t, y) = \frac{\partial u}{\partial t} + (\text{grad } u).\ f(t, y) \qquad (3)$$

where the dot on the right side of (3) signifies scalar multiplication and

$$\text{grad } u = \left(\frac{\partial u}{\partial y^1}, \frac{\partial u}{\partial y^2}, \ldots, \frac{\partial u}{\partial y^n} \right) \qquad (4)$$

is the gradient of u w.r.t. $y = (y^1, y^2, \ldots, y^n)$.
Now, we consider the autonomous system

$$\frac{dy}{dt} = f(y) \tag{5}$$

(i.e., when f does not depend on t explicitly). We assume that f(y) is defined on an open set containing $y = 0$.

Definition. Let V(y) be a scalar function of y, defined and continuous, with continuous partial derivatives at all points in a domain D containing the origin $y = 0$ and such that $V(0) = 0$.

(i) The function V(y) is **called positive definite** in D if $V(y) > 0$ for all other points y in D.
(ii) The function V(y) is **called positive semi definite** in D if $V(y) \geq 0$ for all other points in D.
(iii) The function V(y) is **called negative definite** in D if $V(y) < 0$ for all other points in D.
(iv) The function V(y) is **called negative semi definite** in D if $V(y) \leq 0$ for all other points in D.

Example : The function

$$V(y) = (y^1)^2 + (y^2)^2 + \ldots + (y^n)^2$$

is positive semi-definite.

Note: Let y(t) be a solution of the autonomous system (5) and consider the function $V(t) = V(y(t))$. Then the derivative of V along the orbit $y = y(t)$ is

$$\dot{V}(t) = \dot{V}(y(t)) = (\text{grad } V).\, f(y) \tag{6}$$

Definition : A real valued function V(y) defined on a neighborhood of $y = 0$ is called a **Lyapunov function** if
(i) V(y) has continuous partial derivatives;
(ii) $V(y) \geq 0$ according as $|y| \geq 0$;
(iii) the trajectory / orbit derivative of V satisfies inequality $\dot{V}(y) \leq 0$.

Result :- Let f(y) be continuous on an open set containing the point $y = 0$, $f(0) = 0$, and let there exist a Lyapunov function V(y). Then the solution $y = 0$ of the autonomous system

$$\frac{dy}{dt} = f(y)$$

is stable, in the sense of Lyapunov.

Remark 1. Roughly speaking, Lyapunov stability of the critical point $y = 0$ means that if a solution y(t) starts near $y = 0$, it remains near $y = 0$ in the future $(t \geq 0)$.

Remark 2. The proof of this result and other related results shall be discussed in detail in chapter 14 .

Non autonomous systems

For nonautonomous systems, the definition of Lypunov function is suitably modified.
Let f(t, y) be continuous for $t \geq T$, $|y| \leq b$ and satisfy

$$f(t, 0) = 0 \qquad \text{for } t \geq T. \tag{7}$$

Definition : A function V(t, y) defined for $t > T$, $|y| \leq b$ is **called a Lyapunov function** if

(i) V(t, y) has continuous partial derivatives;
(ii) V(t, 0) = 0 for $t \geq T$;
(iii) there exists a continuous functions W (y) on $| y | \leq b$ such that $W(y) \geq 0$ according as $| y | \geq 0$;
(iv) $V(t, y) \geq W(y)$ for $t \geq T$;
(v) the trajectory derivative of V satisfies

$$\dot{V}(t, y) \leq 0.$$

Remark : The scalar function V(y) can be regarded as a measure of the ″energy″ of the system

$$\frac{dy}{dt} = f(t, y)$$

and it seeks to demonstrate that either this ″energy″ decreases as $t \to \infty$, indicating stability, or increases as , indicating instability.

To illustrate the use of Liapunov functions, consider the following examples.

Example 1 : For n = 2, consider the system of equations

$$\frac{dy^1}{dt} = -2y^1(y^2)^2 - (y_1)^3, \quad \frac{dy^2}{dt} = -y^2 + (y^1)^2 - y^2.$$

We try the function

$V(y^1, y^2) = a (y^1)^2 + (y^2)^2$, with constant a to be determined. We find

$$\dot{V}(y^1, y^2) = -2 (y^2)^2 - 2a(y^1)^4 - 2(y^1)^2(y^2)^2.(2a - 1).$$

Choose a = 1, so that

$$\dot{V} = -2(y^2)^2 - 2(y^1)^2\{(y^1)^2 + (y^2)^2\}.$$

Now V is positive definite, while $\dot{V}$ is negative definite. Consequently, the zero solution ($y^1 = y^2 \equiv 0$) is uniformly and asymptotically stable. We note that any choice of $a \geq \frac{1}{2}$ would be just as useful here.

Example 2. Consider the two – dimensional plane autonomous system.

$$\frac{dy^1}{dt} = y^2, \quad \frac{dy^2}{dt} = -w^2 y^1 - \alpha (y^1)^2 y^2.$$

Putting $u = y^1$, this system is equivalent to the second order O D E

$$\frac{d^2u}{dt^2} + \alpha u^2 \frac{du}{dt} + w^2u = 0$$

This equation can be recognised as the equation for a simple harmonic oscillator of frequency w, with a nonlinear damping term. Here, we try

$$V(y^1, y^2) = \frac{1}{2}\{w^2(y^1)^2 + (y^2)^2\},$$

which may be interpreted as the energy of the undamped oscillator. We find

$$\dot{V}(y^1, y^2) = -\alpha (y^1)^2 (y^2)^2.$$

For $\alpha > 0$, $\dot{V}$ is negative semi definite, showing that, as expected, the energy is damped.

Also note here that, for $\alpha < 0 < \dot{V}$, is positive semi-definite. This example shows that, in seeking Liapunov functions, it is sometimes useful to identify V with the energy of the physical system which the equations describe.

The readers are advised to refer the following books for reading (lessons 7-9 of Unit II) :

1. Birkhoff, G. and Rota, G.C.
   Ordinary differential Equations, John Wiley and sons, Third edition (1978)
2. Ross, S.L.
   Differential Equations, John Wiley& sons, (1984)
3. Hartman, P.
   Ordinary differential Equations, John Wiley (1964).
4. Deo, S.G. and Raghavendra, V.
   Ordinary differential Equations and Stability Theory, Tata McGraw Hill, New Delhi, 1980.

9

LINEAR SYSTEMS AND VARIATION OF CONSTANTS

Definition (Norm of a matrix). Let $A = (a_{ij})$ be a $n \times n$ matrix of complex numbers. The norm of the matrix A, denoted by $|A|$, is defined as

$$|A| = \sum_i \sum_j |a_{ij}| \qquad (1)$$

Note : If $x \in C^n$, then $A x \in C^n$. For $x = (x_1, x_2, \ldots\ldots\ldots, x_n) \in C^n$, we have already defined

$$|x| = |x_1| + |x_2| + \ldots\ldots\ldots + |x_n| = \sum_j |x_i|. \qquad (2)$$

We note that the definition of $|x|$, given in (2), coincides with the definition, given in (1), when x is regarded as a row matrix.

Result. The norm of a matrix satisfies the following properties:
(i) $|A + B| \le |A| + |B|$
(ii) $|A B| \le |A| |B|$
(iii) $|A x| \le |A| |x|$
for x being a n–dimensional vector.

Notes :- (1) A unit matrix of order n is denoted by E_n. We find $|E_n| = n$.
(2) $|A| = |A^T| = |\overline{A}| = |A^*|$, where $A^* = (\overline{A})^T = \overline{(A^T)}$.

Definition:- The determinant of the matrix A is denoted by **det A** and trace of A is denoted by **tr A**, where

$$\text{tr } A = a_{11} + a_{22} + \ldots + a_{nn}.$$

Remark : Let $a_{ij}(t)$ be complex-valued functions of a real variable t on an interval I, for $1 \le i, j \le n$. Let

$$A(t) = [a_{ij}(t)]$$

denote a matrix function. If the elements of a matrix possess a property such as continuity, differentiability, or integrability, for brevity it is said that the matrix function A(t) has this property. In particular, If A(t) is differentiable, then

$$A'(t) = \frac{dA(t)}{dt} = [a'_{ij}(t)].$$

If A(t) is integrable on an interval I = [a, b], then

$$\int_a^b A(t)\,dt = \left[\int_a^b a_{ij}(t)dt\right].$$

Notes:- (1) $\overline{A}'(t)$ means $\overline{A'(t)}$ or $[\,\overline{A(t)}\,]'$. Here, dash indicates differentiation and not transpose. For transpose, we write $(\ldots)^T$.

(2) In this chapter, the quantities u, v, p are scalars; c, y, z, f and g are n-dimensional column vectors; and A, B, C, X, Y, Z are matrices. The scalars, components of the vectors, and elements of the matrices will be supposed to be complex-valued.

Definition (Characteristic polynomial) :- Let A be a square matrix of order n. Then, det(λ E – A) is a polynomial in λ of degree n, and is **called the characteristic polynomial of A**. Its roots are **called the characteristic roots or eigenvalues** of the matrix A.

Note:- If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A, then

$$\det(\lambda E - A) = \prod_{i=1}^{n} (\lambda - \lambda_i).$$

This gives $\det A = \prod_{i=1}^{n} \lambda_i$, where the product is taken over all roots.

Definition (Similar matrices) :- Let A and B be two matrices of the n × n type. Then A and B are said to **be similar** if there exists a non-singular matrix P such that

$$B = P A P^{-1}.$$

Result:- If A and B are similar matrices, then they have the same characteristic polynomial, because

$$\begin{aligned}\det(\lambda E - B) &= \det[P(\lambda E - A)P^{-1}] \\ &= (\det P)\,[\det(\lambda E - A)]\,\det(P^{-1}) \\ &= \det(\lambda E - A).\end{aligned}$$

In particular, the coefficients of the powers of λ in polynomial, det (λ E – A), are invariant under similarity transformation. Two of the most important invariant are

(i) det A ≡ determinant of A

(ii) tr A ≡ trace of A.

We now state the following fundamental result concerning the canonical form of a matrix.

Statement:- Every n × n complex matrix A is similar to a matrix of the form

$$J = \begin{bmatrix} J_0 & 0 & 0 \ldots\ldots\ldots 0 \\ 0 & J_1 & 0 \ldots\ldots\ldots 0 \\ - - - & - - - & - - - \\ 0 & 0 & 0 \ldots\ldots\ldots J_s \end{bmatrix}$$

where $J_0 = \text{dia}\,[\lambda_1, \lambda_2, \ldots, \lambda_q]$
is a diagonal matrix and

$$J_i = \begin{bmatrix} \lambda_{q+i} & 1 & 0 & 0 \ldots\ldots\ldots 0 & 0 \\ 0 & \lambda_{q+i} & 1 & 0 \ldots\ldots\ldots 0 & 0 \\ - - - & - - - & - - - & - - - & - - - \\ 0 & 0 & 0 & 0 \ldots\ldots\ldots \lambda_{q+i} & 1 \\ 0 & 0 & 0 & 0 \ldots\ldots\ldots 0 & \lambda_{q+i} \end{bmatrix}, \quad (i = 1, 2, \ldots, s)$$

The λ_j (j = 1, 2, …, q + s) are the characteristic roots of A, which need not all be distinct. If λ_j is a simple root, then it occurs in J_0, and therefore, if all the roots are distinct, A is similar to the diagonal matrix

$$J = \begin{bmatrix} \lambda_1 & 0 & 0..........0 \\ 0 & \lambda_2 & 0..........0 \\ - - - & - - - & - - - \\ 0 & 0 & 0..........\lambda_n \end{bmatrix}$$

Remark 1:- From the above result, it follows that

$$\det A = \prod_i \lambda_i \,,\ tr(A) = \sum_i \lambda_i$$

where the product and sum are taken over all roots, each root counted a number of times equal to its multiplicity.

Remark 2:- The matrices J_i are of the form

$$J_i = \lambda_{q+i} E_{ri} + Z_i$$

where J_i has r_i rows and r_i columns, and

$$Z_i = \begin{bmatrix} 0 & 1 & 0 & 0.........0 & 0 \\ 0 & 0 & 1 & 0.........0 & 0 \\ - - - & - - - & - - - & - - - & - - - \\ 0 & 0 & 0 & 0.........0 & 1 \\ 0 & 0 & 0 & 0.........0 & 0 \end{bmatrix}.$$

Remark 3:- An equally valid form is

$$J_i = \lambda_{q+i} E_{ri} + \gamma Z_i$$

where γ is any constant not zero.

Remark 4:- The matrix Z_i is **nilpotent** as $Z_i^{ri} = 0$.

Definition 1:- (Sequence of matrices)

Let $\{A_m\}$ be a sequence of matrices. It is said to be convergent if, given $\in > 0$, there exists a positive integer N such that

$$| A_p - A_q | < \in \text{ for } p, q > N$$

i.e., norm of $(A_p - A_q)$ is less than $\in$ whenever $p, q > N$.

Definition 2 (Limit Matrix) :- A sequence $\{A_m\}$ of matrices is said to have a limit matrix A if, given $\in > 0$, there exists a positive integer N (depending upon $\in$ only) such that

$$| A_p - A | < \in \text{ whenever } p > N .$$

Result:- Clearly, sequence $\{A_m\}$ of matrices is convergent iff each of the component sequences is convergent. This implies that the sequence $\{A_m\}$ of matrices is convergent iff there exists a limit matrix to which it tends.

Definition 3 (Infinite Series)

The infinite series

$$A_1 + A_2 + ...+ A_m + ... = \sum_{m=1}^{\infty} A_m$$

of matrices is said to be convergent if the sequence of its partial sums is convergent.

Result : The sum of this series of matrices is defined to be the limit matrix of the sequence of partial sums.

Definition 4(Exponential of a matrix A)

The infinite series

$$E + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots\text{..to } \infty = E + \sum_{m=1}^{\infty} \frac{A^m}{m!}$$

is convergent for all A, since for any positive integers p, q,

$$|S_p - S_q| = \left| \sum_{m=p+1}^{q} \frac{A^m}{m!} \right|$$

$$\leq \sum_{m=p+1}^{q} \frac{|A|^m}{m!}$$

and the latter represents the Cauchy difference for the series $e^{|A|}$, of real numbers, which is convergent for all finite $|A|$.

The sum of the above convergent series is denoted by

e^A = exponential of matrix A.

Remark. The exponential series is of great importance for the study of linear differential equations.

Note(1):- For matrices, it is not, in general, true that

$$e^{A+B} = e^A e^B.$$

But this relation is valid if matrices A and B commute.

Note(2):- It will be shown later on that

$$\det(e^A) = e^{tr(A)} \neq 0.$$

Hence, e^A is a non-singular matrix for all A.

Note(3):- Since matrices A and '-A' commute, so

$$e^{-A} = (e^A)^{-1}.$$

Result:- We know that every square matrix A satisfies its own characteristic equation

$$\det(\lambda E - A) = 0.$$

This result is sometimes very useful for the actual calculation of e^A. For example, let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Its characteristic equation is $\lambda^2 = 0$, which is satisfied by A, i.e.,

$$A^2 = 0.$$

$$\Rightarrow A^m = 0 \text{ for } m \geq 2.$$

Hence, $e^A = E + A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

Definition:- [Logarithm of a matrix]. Let B be a non-singular matrix. A logarithm of B is a matrix, say A, such that

$$e^A = B.$$

Remark (1) :- A is not unique, because, $e^A = e^{A+2\pi ikE}$ for $k = 0, \pm1, \pm2, \ldots$..

Remark (2) :- Indeed, if B is in the canonical form J, then A can be taken as

$$A = \begin{pmatrix} A_0 & 0 & 0.........0 \\ 0 & A_1 & 0.........0 \\ ---- & ---- & ---- \\ 0 & 0 & 0..........A_s \end{pmatrix}.$$

provided that $e^{A_i} = J_i$ for i = 0, 1, 2, …., s.

It can also be verified that a suitable matrix A_0 is given by

$$A_0 = \begin{bmatrix} \log\lambda_1 & 0.........0 \\ 0 & \log\lambda_2........0 \\ ---- & ---- \\ 0 & 0...........\log\lambda_q \end{bmatrix}.$$

Remark (2) :- For any matrix M, we have

$$(P\,M\,P^{-1})^k = P\,M^k\,P^{-1} \qquad \text{for } k = 1, 2,…..$$

Consequently, we get

$$P(e^M)\,P^{-1} = e^{PMP^{-1}}.$$

Results (1). Let $\Phi = \Phi(t)$ be any $n \times n$ matrix of functions defined on a real t-interval I (the functions may be real or complex valued). Let

$$\Phi(t) = (a_{ij}(t))_{n\times n}. \qquad (1)$$

Let A_{ij} = cofactor of a_{ji} in $\Phi(t)$.

Let $\tilde{\Phi} = (A_{ij})$. (2)

If $\Phi'(t)$ exists and $\Phi(t)$ is non-singular at t, then $\Phi^{-1}(t)$ is differentiable at t and

$$\Phi^{-1} = \frac{\tilde{\Phi}}{\det\Phi}, \qquad (3)$$

and

$$\Phi\,\Phi^{-1} = \Phi^{-1}\,\Phi = E. \qquad (4)$$

Result (2) :- We find

$$(\Phi^{-1})' = -\Phi^{-1}\,\Phi'\,\Phi^{-1}, \qquad \det\Phi \neq 0. \qquad (5)$$

Solution:- We have

$$(\Phi\,\Phi^{-1})' = 0$$
$$\Rightarrow \quad \Phi'\,\Phi^{-1} + \Phi(\Phi^{-1})' = 0$$
$$\Rightarrow \quad \Phi(\Phi^{-1})' = -\,\Phi'\,\Phi^{-1}$$
$$\Rightarrow \quad (\Phi^{-1})' = -\,\Phi^{-1}\,\Phi'\,\Phi^{-1}.$$

Linear Systems

Now we shall be discussing some basic facts and results about linear systems of differential equations in the homogeneous case,

$$\frac{dy}{dt} = A(t)\,y, \qquad \text{(LH)}$$

and in the inhomogeneous / non-homogeneous case,

$$\frac{dy}{dt} = A(t)\,y + b(t). \qquad \text{(NH)}$$

Throughout the study, A(t) is a continuous $n \times n$ matrix and b(t) a continuous vector on a t-interval I = [a, b]. The linear homogeneous system (LH) is also called a linear homogeneous system of the nth order.

Linear Homogeneous Systems

We know that for given any y_0, and $t_0 \in$ I, there exists a unique solution $\varphi = \varphi(t)$ of linear homogeneous differential equation (LH) on I such that $\varphi(t_0) = y_0$.

Remarks :- (i) The z**ero** vector function on I is always a solution of system (LH). This solution is called the **trivial solution** of (LH).

(ii) If a solution of (LH) is zero for some $t_0 \in$ I, then, by uniqueness theorem, it must be zero throughout I .

Theorem 9.1. (Principle of Superposition) :- The set of all solutions of linear homogeneous system of nth order

$$\frac{dy}{dt} = A(t)\, y, \quad t \in I \qquad \text{(LH)}$$

form an n-dimensional vector space over the complex field.

Proof:- Let φ_1 and φ_2 be solutions of linear homogeneous system (LH) and c_1, c_2 be two complex numbers. Then

$$\varphi_1'(t) = A(t)\, \varphi_1(t),$$
$$\varphi_2'(t) = A(t)\, \varphi_2(t) . \qquad (1)$$

Now
$$\frac{d}{dt}\{c_1\, \varphi_1 + c_2\, \varphi_2\} = c_1 \frac{d\varphi_1}{dt} + c_2 \frac{d\varphi_2}{dt}$$
$$= c_1\, A\, \varphi_1 + c_2\, A\, \varphi_2$$
$$= A(c_1\, \varphi_1 + c_2\, \varphi_2) \qquad (2)$$

This shows that $c_1\, \varphi_1 + c_2\, \varphi_2$ is again a solution of (LH). Hence, the set of solutions forms a linear/ vector space over the field of complex numbers.

To show that this solution space is n-dimensional, it is required to establish a set of n linearly independent solutions $\varphi_1, \varphi_2,\ldots, \varphi_n$ such that every other solutions of (LH) is a linear combination (with complex coefficients) of solutions $\varphi_1 , \varphi_2,.., \varphi_n$.

We know that y-space is n-dimensional. Let ξ_i, i = 1, 2, …, n, be linearly independent points in this space. Then, by the existence theorem, for $t_0 \in$ I, there exists n solutions φ_i, $1 \le i \le n$, of linear homogeneous system (LH) such that

$$\varphi_i(t_0) = \xi_i, \qquad 1 \le i \le n . \qquad (3)$$

Now, we shall show that these solutions satisfy the required conditions.

If the solutions φ_i are linearly dependent, then there exists n complex numbers c_i, not all zero, such that

$$c_1\, \varphi_1(t) + c_2\, \varphi_2(t) + \ldots..+ c_n\, \varphi_n(t) = 0 \qquad \text{for all } t \in I. \qquad (4)$$

In particular, for $t = t_0$,

$$c_1\, \varphi_1(t_0) + c_2\, \varphi_2(t_0) + \ldots..+ c_n\, \varphi_n(t_0) = 0$$
$$\Rightarrow \quad c_1\, \xi_1 + c_2\, \xi_2 + \ldots..+ c_n\, \xi_n = 0 \ , \qquad (5)$$

and this contradicts the assumptions that the ξ_i are linearly independent. This contradiction shows that the solutions $\varphi_i(t)$ are linearly independent.

Let $\varphi = \varphi(t)$, $t \in$ I, be any solution of linear homogeneous system (LH) such that $\varphi(t_0) = \xi$. Then, there exists unique constants $k_1, k_2, \ldots, k_n$ such that

$$\xi = k_1\, \xi_1 + k_2\, \xi_2 + \ldots.+ k_n\, \xi_n \qquad (6)$$

as ξ belongs to n-dimensional y-space, for which ξ_i form a basis. Now, the function

$$k_1\varphi_1(t) + k_2\varphi_2(t) + \ldots. + k_n\varphi_n(t) , \qquad t \in I \qquad (7)$$

is a solution of (LH) on the t-interval I and this solution assumes the value ξ at $t = t_0$. Hence, by uniqueness theorem, this solution must be ϕ. That is,

$$\varphi = k_1\varphi_1 + k_2\varphi_2 + \ldots + k_n\varphi_n, \text{ on } I. \quad (8)$$

This shows that every solution φ is a unique linear combination of the n solutions φ_i.
Consequently, the solution space is n – dimensional. This completes the proof.

Definition(Fundamental Set):- Let

$$\frac{dy}{dt} = A(t)\,y, \quad t \in I \quad \text{(LH)}$$

be a linear homogeneous system of the nth order, A(t) being an $n \times n$ matrix. If $\varphi_1, \varphi_2, \ldots, \varphi_n$ form a set of n linearly independent solutions of the system (LH), they are said to form a **basis** or a **fundamental set** of solutions of the linear homogeneous system (LH).

Definition: (Fundamental Matrix)

Let $\Phi(t)$ be a matrix whose n columns are n linearly independent solutions of the linear homogeneous system (LH) on interval I. That is,

$$\Phi(t) = \begin{bmatrix} \varphi_1^{\,1}(t) & \varphi_2^{\,1}(t) \ldots\ldots\ldots \varphi_n^{\,1}(t) \\ \varphi_1^{\,2}(t) & \varphi_2^{\,2}(t) \ldots\ldots\ldots \varphi_n^{\,2}(t) \\ ------ & ---------- \\ \varphi_1^{\,n}(t) & \varphi_2^{\,n}(t) \ldots\ldots\ldots \varphi_n^{\,n}(t) \end{bmatrix}, \quad (1)$$

where

$$\phi_i(t) = (\phi_i^{\,1}(t), \phi_i^{\,2}(t), \ldots\ldots, \phi_I^{\,n}(t)) \quad (2)$$

for $1 \le i \le n$.
The matrix $\Phi(t)$ is called a fundamental matrix for the linear homogeneous system (LH). It is evident that $\Phi(t)$ satisfies the matrix – differential equation

$$\Phi'(t) = A(t)\,\Phi(t) \qquad \text{for } t \in I$$

which is associated with the given system (LH). We means that $\Phi(t)$ is a solution of the following associated matrix differential equation

$$X' = A(t)\,X, \quad t \in I. \quad \text{(M)}$$

The matrix $\Phi(t)$ is called a solution of matrix equation (M) on the interval I.

Remark: (1) It is now evident that a complete knowledge of the set of solutions of (LH) can be obtained if one knows a fundamental matrix for (LH), which is, of course, a particular solution of associated matrix – differential equation (M).
Remark : (2) The determinant of matrix $\Phi(t)$ is called the **Wronskian** of the system (LH) w.r.t. $\{\phi_1, \phi_2, \ldots, \phi_n\}$ and is denoted by $W(\phi_1, \phi_2, \ldots, \phi_n)$. It is a function of t.

Theorem 9.2 (Liouville's formula):-
Let A(t) be an $n \times n$ matrix with continuous elements on an interval I = [a, b], and suppose $\Phi(t)$ is a matrix of functions on I satisfying the matrix differential equation.

$$\Phi'(t) = A(t)\,\Phi(t), \quad t \in I.$$

Prove that $\det\{\Phi(t)\}$ satisfies the first order scalar differential equation

$$(\det \Phi)' = \{tr(A)\}\,(\det \Phi), \text{ on } I$$

and for $t_0, t \in I$,

$$\det \Phi(t) = \{\det \Phi(t_0)\} \exp\left[\int_{t_0}^{t} \{trA(s)\}\,ds\right].$$

Proof:- Let

$$\Phi(t) = (\phi_{ij}(t))_{n\times n}, \quad A(t) = (a_{ij}(t))_{n\times n}. \qquad (1)$$

Then the given matrix differential equation gives the following scalar differential equations

$$\phi_{ij}'(t) = \sum_{k=1}^{n} a_{ik}(t)\, \phi_{kj}(t), \qquad \text{for } i, j = 1, 2, \ldots, n. \qquad (2)$$

We know that the derivative of det{Φ(t)} is a sum of n determinants and given by

$$(\det \Phi(t))' = \begin{vmatrix} \phi_{11}'(t) & \phi_{12}'(t) \ldots\ldots\ldots \phi_{1n}'(t) \\ \phi_{21}(t) & \phi_{22}(t) \ldots\ldots\ldots \phi_{2n}(t) \\ ----- & ----- \\ \phi_{n1}(t) & \phi_{n2}(t) \ldots\ldots\ldots \phi_{nn}(t) \end{vmatrix}$$

$$+ \begin{vmatrix} \phi_{11}t) & \phi_{12}(t) \ldots\ldots\ldots \phi_{1n}(t) \\ \phi_{21}'(t) & \phi_{22}'(t) \ldots\ldots\ldots \phi_{2n}'(t) \\ ----- & ----- \\ \phi_{n1}(t) & \phi_{n2}(t) \ldots\ldots\ldots \phi_{nn}(t) \end{vmatrix} + \ldots\ldots + \begin{vmatrix} \phi_{11}(t) \ldots\ldots\ldots \phi_{1n}(t) \\ \phi_{21}(t) \ldots\ldots\ldots \phi_{2n}(t) \\ ----- \\ \phi_{(n-1)1}(t) \ldots\ldots\ldots \phi_{(n-1)1}(t) \\ \phi_{n1}'(t) \ldots\ldots\ldots \phi_{nn}'(t) \end{vmatrix}. \qquad (3)$$

Now, we shall consider each determinant on right hand side of equation (3), turn by turn. Using the values of ϕ_{ij}' from equation (2) in the first determinant on the right of (3), one gets

$$\Delta_1 = \begin{vmatrix} \sum_k a_{1k}\, \phi_{k1} & \sum_k a_{1k}\, \phi_{k2} \ldots\ldots \\ \phi_{21} & \phi_{22} \ldots\ldots\ldots \\ ----- & ----- \\ \phi_{n1} & \phi_{n2} \ldots\ldots\ldots \end{vmatrix}. \qquad (4)$$

This determinant remain unchanged if one subtracts from the first row a_{12} times the second row plus a_{13} times the third row up to a_{1n} times the nth row. This gives

$$\Delta_1 = a_{11}\, (\det \Phi). \qquad (5)$$

Carrying out a similar procedure with the second determinant, one gets

$$\Delta_2 = a_{22}(\det \Phi). \qquad (6)$$

After n steps, one gets

$$\Delta_n = a_{nn}\, (\det \Phi). \qquad (7)$$

Thus, from equations (3) to (7), one finally gets

$$(\det \Phi)' = (a_{11} + a_{22} + \ldots\ldots + a_{nn})\, (\det \Phi)$$
$$= (\text{tr } \Phi)\, (\det \Phi)\,. \qquad (8)$$

This proves the first part of the theorem.

Let $\quad u = \det \Phi,\ \alpha(t) = \text{tr } \Phi. \qquad (9)$

Then, equation (8) is of the form

$$\frac{du}{dt} - \alpha(t)\, u(t) = 0,$$

or
$$\frac{du}{u} = \alpha(t)\, dt$$

or
$$\Rightarrow \qquad u(t) = C \exp\left[\int_{t_0}^{t} \alpha(t)dt\right], \qquad (10)$$

where C is a constant of integration.

On putting $t = t_0$, both sides, we get

$$C = u(t_0) . \qquad (11)$$

Hence, we get

$$u(t) = u(t_0) \exp\left[\int_{t_0}^{t} \alpha(s)ds\right] .$$

Consequently,

$$\det[\Phi(t)] = \det[\Phi(t_0) \exp\left\{\int_{t_0}^{t} [tr(\Phi(s))]\right\} ds . \qquad (12)$$

This proves the theorem finally.

Theorem 9.3 :- A necessary and sufficient condition that a solution matrix Φ of differential equation

$$X' = A(t) X, \quad t \in I \qquad (M)$$

be a fundamental matrix is that

$$\det \{\Phi(t)\} \neq 0 \quad \text{for } t \in I.$$

Proof:- We know that, from Liouville theorem (9.2),

$$\frac{d}{dt}[\det\{\Phi(t)\}] = \{tr\ A(t)\} \{\det \Phi(t)\}, \qquad (1)$$

and

$$\det\{\Phi(t)\} = \det\{\Phi(t_0)\} \exp\left[\int_{t_0}^{t} tr(A(s)ds\right], \qquad (2)$$

for $t_0, t \in I$.

From (2), it follows that if

$$\det \Phi(t_0) \neq 0 \text{ for some } t_0 \in I, \qquad (3)$$

then

$$\det\{\Phi(t)\} \neq 0 \text{ for all } t \in I. \qquad (4)$$

Let $\Phi(t)$ be a fundamental matrix with column vectors $\varphi_1, \varphi_2, \ldots, \varphi_n$. Then vectors $\varphi_1, \varphi_2, \ldots, \varphi_n$ form a set of n linearly independent solutions of linear homogeneous differential equation

$$\frac{dy}{dt} = A(t)\ y, \quad t \in I \qquad (LH)$$

Let $\varphi = \varphi(t)$ be any non-trivial solution of (LH). Then there exists unique constants $c_1, c_2, \ldots, c_n$, not all zero, such that

$$\varphi(t) = c_1\varphi_1(t) + c_2\ \varphi_2(t) + \ldots + c_n\ \varphi_n(t) \text{ for all } t \in I. \qquad (5)$$

Equation (5) can be expressed as

$$\varphi(t) = \Phi\ (t)\ C \qquad (6)$$

where C is the column matrix / vector with components $c_1, c_2, \ldots, c_n$.

The relation (6) is a system of n linear non-homogeneous algebraic equations in the n unknowns $c_1, c_2, \ldots, c_n$, at any $t_0 \in I$, and has a unique solution for any choice of $\varphi(t_0)$. Consequently,

$$\det\{\Phi(t_0)\} \neq 0. \qquad (7)$$

Hence, by (2), it follows that

$$\det\{\Phi(t)\} \neq 0 \quad \text{for any } t \in I. \qquad (8)$$

Conversely, let us assume that $\Phi(t)$ be a solution matrix of matrix differential equation (M) such that $\det\{\Phi(t)\} \neq 0$ for $t \in I$. Then, the column vectors, say ϕ_i, of the matrix Φ are linearly independent at every $t \in I$. So, by definition, Φ is a fundamental matrix for (LH).
This completes the proof.

Theorem 9.4:- If Φ is a fundamental matrix of the linear homogeneous system

$$\frac{dy}{dt} = A(t)\, y \qquad \text{(LH)}$$

and C a complex constant non-singular matrix, then $\Phi\, C$ is again a fundamental matrix of (LH). Moreover, every fundamental matrix of (LH) is of this type for some non-singular constant matrix C.

Proof:- As $\Phi(t)$ is a fundamental matrix of the system (LH), so

$$\Phi'(t) = A(t)\, \Phi(t)\,, \qquad (1)$$

and $$\det\{\Phi(t)\} \neq 0 \qquad \text{for } t \in I. \qquad (2)$$

This implies

$$\Phi'(t)\, C = A(t)\, \Phi(t)\, C$$

$$\Rightarrow (\Phi\, C)' = A(t)\, \{\Phi\, C\}, \qquad (3)$$

C being a constant matrix.
This shows that $\Phi\, C$ is a solution of matrix differential equation

$$X' = A(t)\, X, \qquad t \in I. \qquad \text{(M)}$$

Since $$\det\{\Phi\, C\} = (\det \Phi)\, (\det C) \neq 0\,, \qquad (4)$$

because neither $\det \Phi = 0$ nor $\det C = 0$, being non-singular. It follows that $\Phi\, C$ is also a fundamental matrix.

Conversely, Let ψ be any other fundamental matrix of the system (LH). Then

$$\psi'(t) = A(t)\, \psi(t) \qquad (5)$$

Let $$\Phi^{-1}\, \psi = \chi. \qquad (6)$$

Then X is non-singular and

$$\psi = \Phi\, \chi \qquad (7)$$

and $$\psi' = \Phi'\, \chi + \Phi\, \chi' \qquad (8)$$

This implies

$$\begin{aligned} A\, \psi &= \psi' \\ &= \Phi'\, \chi + \Phi\, \chi' \\ &= A(\Phi\chi) + \Phi\, \chi' \\ &= A\, \psi + \Phi\, \chi'\,. \end{aligned}$$

This gives

$$\Phi\, \chi' = 0$$

$$\Rightarrow \quad \chi' = 0 \qquad (\Theta\ \Phi \text{ is non-singular})$$

$$\Rightarrow \quad \chi = \text{constant} = C, \text{ a non–singular matrix.}$$

$$\Rightarrow \psi = \Phi\, C, \quad \text{using (7)}$$

This completes the proof.

Note:- (1) If Φ is a fundamental matrix of (LH) and C is a constant non-singular matrix, then $C\, \Phi$ is not, in general, a fundamental matrix.

Note :- (2) Two different homogeneous systems can not have the same fundamental matrix, for in (LH),

$$A(t) = \Phi'(t)\, \Phi^{-1}(t)$$

Hence $\Phi(t)$ determines $A(t)$ uniquely. Although, the converse is not true.

Non homogeneous Linear Systems:

If a fundamental matrix Φ for the corresponding homogeneous linear system (LH) is known, then there is a simple method for calculating a solution of non homogeneous linear system (NH).

Theorem 9.5 :- If $\Phi(t)$ is a fundamental matrix for the homogeneous linear system

$$\frac{dy}{dt} = A(t)\, y, \quad t \in I \qquad \text{(LH)}$$

where A is a $n \times n$ matrix, then the function $\phi(t)$ defined by

$$\varphi(t) = \Phi(t)\left[\int_{t_0}^{t} \Phi^{-1}(s)b(s)ds\right], \qquad t \in I$$

is a solution of non-homogeneous linear system

$$\frac{dy}{dt} \quad A(t)\, y + b(t) \quad t \in I \qquad \text{(NH)}$$

satisfying

$$\varphi(t_0) = 0, \qquad t_0 \in I$$

Proof:- For any constant vector c, the function Φ c is a solution of homogeneous system (LH). The method here consists of considering c as a function of t on I such that

$$\varphi(t) = \Phi(t)\, c(t) \qquad (1)$$

is a solution of the non homogeneous system (NH). Then

$$\begin{aligned} \varphi'(t) &= \Phi'(t)\, c(t) + \Phi(t)\, c'(t) \\ &= \{A(t)\, \Phi(t)\}\, c(t) + \Phi(t)\, c'(t) \\ &= A(t)\, \varphi(t) + \Phi(t)\, c'(t)\,. \end{aligned} \qquad (2)$$

Comparing (2) with given (NH), it follows that

$$\Phi(t)\, c'(t) = b(t)$$

$$\Rightarrow \quad c'(t) = \Phi^{-1}(t)\, b(t)$$

$$\Rightarrow \quad c(t) = \int_{t_0}^{t} \Phi^{-1}(s)\, b(s)ds\,, \quad \text{for } t_0 \in I\,. \qquad (3)$$

Also $\qquad c(t_0) = 0. \qquad (4)$

For equations (1) and (3), it follows that

$$\varphi(t) = \Phi(t)\left[\int_{t_0}^{t} \Phi^{-1}(s)b(s)ds\right] \qquad (5)$$

is a solution of (NH) with $\varphi(t_0) = 0$. This completes the proof.

Remark (1) The formula (5) is called the **variation of constants** formula for (NH).

Remark (2) Under the assumptions of the above theorem, the solution $\varphi = \varphi(t)$ of non homogeneous linear system (NH) satisfying the initial condition

$$\varphi(t_0) = y_0, \qquad t_0 \in I, \ \ |\, y_0 \,| < \infty$$

is given by

$$\varphi(t) = \varphi_h(t) + \Phi(t)\left[\int_{t_0}^{t} \Phi^{-1}(s)b(s)ds\right], \qquad t \in I$$

where $\varphi_h(t)$ is that solution of homogeneous linear system (LH) on I satisfying the condition.

$$\varphi_h(t_0) = y_0.$$

10

REDUCTION OF THE ORDER OF A HOMOGENEOUS SYSTEM, LINEAR HOMOGENEOUS SYSTEMS WITH CONSTANT COEFFICIENTS, ADJOINT SYSTEMS

Consider the linear homogeneous differential equation

$$\frac{dy}{dt} = A(t)\, y, \quad t \in I \qquad \text{(LH)}$$

of order n, where A(t) is a matrix of type $n \times n$.
If m linearly independent solutions of (LH) are known, $0 < m < n$, the determination of all solutions of (LH) is reduced essentially to the problem of determining the solutions of a linear homogeneous system of (n-m) differential equations.

Suppose $\varphi_1, \varphi_2, \ldots\ldots, \varphi_m$ are m linearly independent vectors with are known solutions of (LH) on an interval I. Let φ_j have components ϕ_j^i $(i = 1, 2, \ldots, n)$.

Then the matrix (ϕ_j^i) is of the type $n \times m$ and the rank of this matrix, at each $t \in I$, is m, because of the linear independence of its columns. This means that for each $t \in I$, there is an $m \times m$ determinant in this matrix (ϕ_j^i) which does not vanish there.

Peck any $t_0 \in I$. W.l.o.g. we assume that the determinant of the sub-matrix Φ_m whose elements are ϕ_j^i $(i,j = 1, 2, \ldots, m)$ is not zero at t_0.

By the continuity of det Φ_m in its elements ϕ_j^i, and the continuity of the functions ϕ_j^i near t_0 (in the nbd of t_0), it follows that

$$\det \Phi(t) \neq 0 \qquad (1)$$

for t in some sub-interval $\tilde{I}$ containing t_0 and $\tilde{I} \subset I$.

The reduction process will be outlined for interval $\tilde{I}$. (The idea behind this process is a modification of the method of variation of constants/parameters).

Outlines of the reduction procedure

Let U be the matrix with the vectors $\varphi_1, \varphi_2, \ldots.., \varphi_m$ as its first m columns and the vectors $e_{m+1}, \ldots, e_n$ for its last (n - m) columns, where e_j is the column vector with all elements 0 except for the jth which is 1. That is,

$$U = \begin{bmatrix} \varphi_1^1 \ldots\ldots \phi_m^1 & 0 \ldots\ldots 0 \\ \varphi_1^2 \ldots\ldots \phi_m^2 & 0 \ldots\ldots 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ \varphi_1^n \ldots\ldots \varphi_m^n & 0 \quad 1 \end{bmatrix}_{n\times n}. \qquad (2)$$

In view of (1), the matrix U is non-singular on $\tilde{I}$. Now, we make the substitution

$$y = U\,w \qquad (3)$$

is made in (LH). We note that, corresponding to $w = e_j$ (j = 1, 2, …,m), the transformation (3) yields $y = \varphi_j$.

Thus, the substitution (3) may be expected to yield a system in w which will have e_j (j = 1, 2,…,m) as solutions. The use of (3) in (LH) gives

$$U'w + U\,w' = A\,U\,w \qquad (4)$$

Writing this out gives

$$\sum_{j=1}^{m} (\varphi_j^i)'\, w^j + \sum_{j=1}^{m} \varphi_j^i\, (w^j)' = \sum_{j=1}^{m} \sum_{k=1}^{n} a_{ik}\, \phi_j^k\, w^j + + \sum_{k=m+1}^{n} a_{ik}\, w^k \qquad (5a)$$

for i = 1, 2, ….,m ,

and

$$\sum_{j=1}^{m} (\phi_j^i)'\, w^j + (w^i)' + \sum_{j=1}^{m} \phi_j^i (w^j)' = \sum_{j=1}^{m} \sum_{k=1}^{n} a_{ik}\, \phi_j^k\, w^j + + \sum_{k=m+1}^{n} a_{ik}\, w^k \qquad (5b)$$

for i = m+1, m+2,…..,n.

As the vectors ϕ_j with components ϕ_j^i are solutions of system (LH), so

$$(\varphi_j^i)' = \sum_{k=1}^{n} a_{ik}\, \phi_j^k \qquad (i = 1, 2,….,n;\ j = 1, 2,…m) \qquad (6)$$

Using (6), relations (5a, b) results in

$$\sum_{j=1}^{m} \varphi_j^i (w^j)' = \sum_{k=m+1}^{n} a_{ik}\, w^k \qquad (i = 1, 2,…,m) \qquad (7a)$$

and

$$(w^i)' + \sum_{j=1}^{m} \phi_j^i (w^j)' = \sum_{k=m+1}^{n} a_{ik}\, w^k \qquad (i = m + 1, 2,…,n) \qquad (7b)$$

Since

$$\det \Phi_m \neq 0 \text{ on } \tilde{I}, \qquad (8)$$

the set of equations in (7a) may be solved for $(w^j)'$, $1 \le j \le m$, in terms of known quantities ϕ_j^i, a_{ik} and w^k (k = m + 1,…,n). These values of $(w^j)'$, so obtained, may then be put into the set of formulas of (7b). This process gives a set of first order differential order equations satisfied by w^i (i = m + 1,…,n) of the type

$$(w^i)' = \sum_{k=m+1}^{n} b_{ik}\, w^k \qquad (i = m + 1, ….,n)\,, \qquad (9)$$

which is a linear system of order n−m, on $\tilde{I}$.

Suppose $\tilde{\psi}_{m+1}$,….., $\tilde{\psi}_n$ is a fundamental set on $\tilde{I}$ for the system (9). Let ψ_j^i be the components of $\tilde{\psi}_j$ (for i, j = m + 1,…,n). Let $\tilde{\psi}_{n\text{-}m}$ denote the matrix with elements ψ_j^i, i.e.,

$$\tilde{\psi}_{n\text{-}m} = (\psi_j^i) \qquad (10)$$

is a matrix of order n - m. Clearly,

$$\det \tilde{\psi}_{n\text{-}m}(t) \neq 0 \text{ on } \tilde{I}. \qquad (11)$$

For each j = m + 1, …,n; let ψ_j^i (i = 1, 2,…..,m) be solved for by integration from the relations (in 7a)

$$\sum_{j=1}^{m} \varphi_j^i (\psi_j^p)' = \sum_{k=m+1}^{n} a_{ik}\, \psi_k^p \qquad (12)$$

for i = 1, 2,…..,m and p = m + 1,….,n.
Let ψ_p (p = m + 1,….,n) denote the vectors having components ψ_p^i (i = 1, 2,…,n). Let

$$\psi_p = e_p \qquad (p = 1, 2, \ldots, n) \tag{13}$$

Now, ψ_p (p = 1, 2,…,n) satisfy system (9) and the first set of equations of (7a), they must also satisfy the second set of equations in (7b). Therefore, ψ_p (p = 1, 2,….,n) are solutions of (7a, b). If Ψ is the matrix with columns ψ_p, p = 1, 2,….,n, and if

$$\Phi = U\,\Psi\,, \tag{14}$$

then Φ is a matrix solution of system (LH) on $\tilde{I}$. As U is non-singular and

$$\det \Psi = \det \tilde{\Psi}_{n-m} \text{ on } \tilde{I}\,, \tag{15}$$

it follows that Φ is non-singular on $\tilde{I}$. Hence, Φ is a fundamental solution of system (LH) on $\tilde{I}$. This completes the reduction procedure and it is summarized in the following theorem:

Theorem 10.1 :- :Let $\varphi_1, \varphi_2, \ldots, \varphi_m$ (m < n) be m known linearly independent solutions of system (LH) with φ_j (j = 1, 2, …,m) having components ϕ_j^i (i = 1, 2, …,n). Assume the determinant of the matrix with elements ϕ_j^i is not zero on some sub - interval $\tilde{I}$ of I. Then the construction of a set of n linearly independent solutions of (LH) on $\tilde{I}$ can be reduced to the solution of a linear system (9) of order n -–m, plus quadratures (integrations) (12), using the substitution (3).

Linear homogeneous systems with constant coefficients

Consider the linear homogeneous system

$$\frac{dy}{dt} = A\,y \tag{LHC}$$

in which A is an $n \times n$ constant matrix.
Let $y_1 \neq 0$ be a constant vector and λ be a complex number. By substituting

$$y = y_1 e^{\lambda t} \tag{1}$$

into equation (1), we at once get

$$Ay_1 = \lambda\, y_1. \tag{2}$$

Equation (2) shows that λ is an eigenvalue of the matrix A with corresponding non-zero eigen vector y_1. Thus, to each eigenvalue λ of A, there corresponds at least one solution of system (LHC) of the (1). If the matrix A has n linearly independent eigen vectors $y_1, y_2, \ldots, y_n$ belonging to the respective eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$; then

$$\phi = (y_1\ e^{\lambda_1 t}, y_2\ e^{\lambda_2 t}, \ldots, y_n\ e^{\lambda_n t}) \tag{3}$$

is a fundamental matrix for the system (LHC).

Theorem 10.2 :- A fundamental matrix Φ for the homogeneous linear system

$$\frac{dy}{dt} = A\,y \tag{LHC}$$

where A is an $n \times n$ constant matrix, is given by

$$\Phi(t) = e^{tA}. \qquad (|\,t\,| < \infty)$$

Proof:- We have

$$e^{(t+\Delta t)A} = e^{tA+\Delta tA} = e^{tA}.e^{\Delta tA} \tag{1}$$

because t A and ΔtA commutes. Further

$$\frac{e^{(t+\Delta t)A} - e^{tA}}{\Delta t} = e^{tA}\left(\frac{e^{\Delta tA} - E}{\Delta t}\right) \qquad (2)$$

Taking limit as $\Delta t \to 0$, we immediately obtain.

$$\frac{d}{dt}(e^{At}) = A\, e^{At}. \qquad (3)$$

This shows that

$$\Phi(t) = e^{At} \qquad (4)$$

is a solution of the given linear homogeneous system with constant coefficients. Since

$$\Phi(0) = E, \qquad (5)$$

it follows that

$$\det\{\Phi(t)\} = \exp\{t(\text{tr } A)\} \neq 0. \qquad (6)$$

This shows that $\Phi(t)$ is a fundamental matrix for the given system.
Hence, the proof is complete.

Theorem 10.3:- The solution φ of the linear homogeneous system with constant coefficients

$$\frac{dy}{dt} = A\, y \qquad \text{(LHC)}$$

satisfying the initial condition

$$\varphi(t_0) = y_0 \qquad (|\, t_0\,| < \infty,\ |\, y_0\,| < \infty)$$

is given by

$$\varphi(t) = \{e^{(t-t0)A}\}y_0$$

Proof:- The successive approximations for a solution of the initial value problem are

$$\phi_0(t) = y_0\,, \qquad (1)$$

$$\phi_n(t) = y_0 + \int_{t_0}^{t} A\,\phi_{n-1}(s)\, ds, \quad \text{for } n \geq 1. \qquad (2)$$

An induction shows that (left as an exercise for readers)

$$\phi_n(t) = \left\{E + A(t-t_0) + \frac{A^2(t-t_0)^2}{2!} + ... + \frac{A^n(t-t_0)^n}{n!}\right\} y_0\ . \qquad (3)$$

The sequence $\{\varphi_n(t)\}$ converges uniformly on any bounded t-interval to the function

$$\phi(t) = \{\, e^{(t-t_0)A}\,\}y_0, \qquad (4)$$

which is then a solution of the given initial value problem. This completes the proof.

Note:- For the inhomogeneous initial value problem

$$\frac{dy}{dt} = A\, y + b(t),$$

$$y(t_0) = y_0\,,$$

the solution is

$$\varphi(t) = \{e^{(t-t0)A}\}y_0 + \int_{t_0}^{t} \{e^{(t-s)A}\}\, b(s)\, ds.$$

Form of the fundamental matrix

Let J be the canonical form of the given matrix A. Then, there exists a non – singular constant matrix P such that

$$\text{or}\quad \left.\begin{matrix} J = P^{-1}A\,P \\ AP = PJ \end{matrix}\right\}. \tag{1}$$

Then $\quad e^{tA} = e^{t(P\,J\,P^{-1})} = e^{P(t\,J)\,P^{-1}}$

$$= P[e^{tJ}]P^{-1}, \tag{2}$$

and J has the form

$$J = \begin{bmatrix} J_0 & 0 & 0\ldots\ldots\ldots 0 \\ 0 & J_1 & 0\ldots\ldots\ldots 0 \\ ------ & ------ & ------ \\ 0 & 0 & 0\ldots\ldots\ldots J_s \end{bmatrix}, \tag{3}$$

where J_0 is a diagonal matrix with diagonal elements $-\lambda_1, \lambda_2, \ldots.., \lambda_q$, and

$$J_i = \begin{bmatrix} \lambda_{q+i} & 1 & 0\ldots\ldots\ldots 0 & 0 \\ 0 & \lambda_{q+i} & 1\ldots\ldots\ldots 0 & 0 \\ ------ & ------ & ------ & ------ \\ 0 & 0 & 0\ldots\ldots\ldots \lambda_{q+i} & 1 \\ 0 & 0 & 0\ldots\ldots\ldots 0 & \lambda_{q+i} \end{bmatrix}, \text{ for } i = 1, 2, \ldots, s. \tag{4}$$

is an $r_i \times r_i$ matrix $(n = q + r_1 + r_2 + \ldots.. + r_s)$. It follows that

$$e^{tJ} = \begin{bmatrix} e^{tJ_0} & 0\ldots\ldots\ldots 0 \\ 0 & e^{tJ_1}\ldots\ldots\ldots 0 \\ ------ & ------ \\ 0 & 0\ldots\ldots\ldots e^{tJ_s} \end{bmatrix}. \tag{5}$$

It is easy to see that

$$e^{tJ_0} = \begin{bmatrix} e^{t\lambda_1} & 0\ldots\ldots\ldots 0 \\ 0 & e^{t\lambda_2}\ldots\ldots\ldots 0 \\ ------ & ------ \\ 0 & 0\ldots\ldots\ldots e^{t\lambda_q} \end{bmatrix}, \tag{6}$$

and

$$e^{tJ_i} = e^{t\lambda_{q+i}} \begin{bmatrix} 1 & t & \frac{t^2}{2!}\ldots\ldots\ldots\frac{t^{r_i-1}}{r_{i-1}!} \\ 0 & 1 & t\ldots\ldots\ldots\frac{t^{r_i-2}}{r_{i-2}!} \\ ------ & ------ & ------ \\ 0 & 0\ldots\ldots\ldots 0 & 1 \end{bmatrix}. \tag{7}$$

Thus, if the canonical form (3) of the matrix A is known, then a fundamental matrix e^{tA} of system (LHC) is given explicitly by (2), where e^{tJ} is to be calculated from equations (5) to (7).

Adjoint Systems

Let Φ be a fundamental matrix for the system (LH), then Φ is non-singular and

$$\Phi\Phi^{-1} = E$$

$$\Rightarrow \Phi'\,\Phi^{-1} + \Phi\,(\Phi^{-1}) = 0$$

$$\Rightarrow \quad \Phi(\Phi^{-1})' = -\Phi'\,\Phi^{-1}$$

$$\Rightarrow \quad (\Phi^{-1})' = -\Phi^{-1}\,\Phi'\,\Phi^{-1}\,. \tag{1}$$

Since Φ is a fundamental matrix for the system (LH), so

$$\Phi' = A\,\Phi\,. \tag{2}$$

From (1) and (2), it follows that

$$(\Phi^{-1})' = \Phi^{-1}\,(A\,\Phi)\,\Phi^{-1}$$

$$= -\Phi^{-1}\,A\,. \tag{3}$$

Taking conjugate transpose, we get

$$[(\Phi^*)^{-1}]' = -A^*\,(\Phi^*)^{-1}\,. \tag{4}$$

This shows that $(\Phi^{-1})^* = (\Phi^*)^{-1}$ is a fundamental matrix for the linear homogeneous system

$$\frac{dy}{dt} = -A^*(t)\,y, \qquad t \in I \tag{AS}$$

Definition:- The system (AS) is called the adjoint to system (LH), and the matrix equation

$$X' = -\,A^*(t)\,X, \qquad t \in I \tag{AM}$$

is called the adjoint to matrix equation

$$X' = A(t)\,X,\ t \in I\,. \tag{M}$$

Remark:- The relationship is symmetric, for (LH) and (M) are the adjoints to (AS) and (AM), respectively.

Theorem (10.4) :- If Φ is a fundamental matrix for linear homogeneous system

$$\frac{dy}{dt} = A(t)\,y, \quad t \in I, \tag{LH}$$

then ψ is a fundamental matrix for its adjoint system

$$\frac{dy}{dt} = -\,A^*(t)\,y, \qquad t \in I, \tag{AS}$$

if and only if $\psi^*\,\Phi = C$,

where C is a constant non-singular matrix.

Proof:- Conditions is necessary.

Since Φ is a fundamental matrix for (LH), so, by definition, $(\Phi^*)^{-1}$ is a fundamental matrix for the linear homogeneous system (AS). Also ψ is an another fundamental matrix for the same linear homogeneous system (AS). So, ψ is of the type

$$\psi = (\Phi^*)^{-1}\,D\,, \tag{1}$$

where D is some constant non-singular matrix. From equation (1), we write

$$\Phi^*\,\psi = D$$

$$\Rightarrow \psi^*\,\Phi = D^*$$

$$\Rightarrow \psi^*\,\Phi = C\,, \tag{2}$$

where $C = D^*$ is some constant non-singular matrix. This shows that the condition is necessary.

Condition is sufficient.

Now, suppose that Φ is a fundamental matrix for (LH) and satisfies the condition. Then, the given condition gives

$$\psi^* = C\,\Phi^{-1}$$

or

$$\psi = (\Phi^*)^{-1}\,C^*. \tag{3}$$

Since $(\Phi^*)^{-1}$ is a fundamental matrix for the adjoint system (AS), hence, $(\Phi^*)^{-1}\,C^*$ and consequently ψ is a fundamental matrix of the adjoint system (AS). This completes the proof.

Remark:- If $A^* = -A$ (1)

Then $(\Phi^*)^{-1}$ is also a fundamental matrix for (LH). It follows that
$\Phi = (\Phi^*)^{-1} C$, for some constant non-singular matrix C

$$\Rightarrow \Phi^* \Phi = C . \quad (2)$$

Equation (2) implies that, in particular, the Euclidean length of any solution vector Φ of linear homogeneous system (LH) is constant.

11

FLOQUET THEORY

Periodicity of solutions of differential systems is an interesting aspect of qualitative study. We shall study certain characterizations for the existence of such solutions. Consider a linear system

$$\frac{dy}{dt} = A(t)y, \qquad -\infty < t < \infty \qquad (1)$$

where A(t) is a $n \times n$ matrix of complex valued continuous functions of real variable t.

Definition. A solution y(t) is periodic with period w, $w \neq 0$, when

$$y(t + w) = y(t) \qquad \text{for all } t.$$

Definition. If

$$A(t + w) = A(t) \qquad \text{for all } t, w \neq 0,$$

then the coefficient matrix is periodic with a period w and the linear system (1) is termed as linear system with periodic coefficients or simply as a periodic system.

Note : Now an interesting question is, when does the system (1) admit periodic solutions and, if it admits a periodic solution, what can be said about the matrix A?

Theorem(11.1): The necessary and sufficient conditions for the system

$$\frac{dy}{dt} = Ay, \qquad -\infty < t < \infty$$

where A is an $n \times n$ constant matrix, to admit a non – zero periodic solution of period w is that the matrix $(E - e^{Aw})$ is singular.

Proof:- We know that the general non – zero solution of the given system is

$$y(t) = e^{At}c, \qquad (1)$$

where c is an arbitrary non – zero constant vector. By definition, y(t) is periodic if and only if

$$y(t) = y(t + w)$$

or $$e^{At}c = e^{A(t+w)}c$$

or $$(E - e^{Aw})c = 0 . \qquad (2)$$

Since c is a non – zero vector, it follows that the given system has a non – zero periodic solution of period w if and only if the matrix $E - e^{Aw}$ is singular. This completes the proof.

Theorem (11.2) (Representation theorem): If $\Phi(t)$ is a fundamental matrix for the periodic system

$$\frac{dy}{dt} = A(t)y, \qquad -\infty < t < \infty$$

with period w and

$$\Psi(t) = \Phi(t + w), \qquad -\infty < t < \infty$$

then Ψ is also a fundamental matrix for the same system. Moreover, corresponding to every such Φ, there exists a periodic non – singular matrix P with period w, and a constant matrix R sung that

$$\Phi(t) = P(t)e^{eR} .$$

Proof:- As $\Phi(t)$ is a fundamental matrix for the given system, so

$$\Phi'(t) = A(t)\,\Phi(t), \qquad -\infty < t < \infty \tag{1}$$

Now, using (1), we obtain

$$\begin{aligned}\Psi'(t) &= \Phi'(t+w)\\ &= A(t+w)\,\Phi(t+w)\\ &= A(t)\,\Psi(t), \qquad -\infty < t < \infty\end{aligned} \tag{2}$$

because A(t) is periodic with period w as given. Equation (2) shows that $\Psi(t)$ is also a solution matrix of the given system. Since

$$\det\{\Psi(t)\} = \det\{\Phi(t+w)\} \neq 0\,, \tag{3}$$

it follows that $\Psi(t)$ is also a fundamental matrix for the given system.

Now $\Phi(t + w)$ and $\Phi(t)$ are two fundamental matrices of the same given linear system, so there exists a constant non- singular matrix C such that

$$\Phi(t+w) = \Phi(t)C\,. \tag{4}$$

As C is a non–singular matrix, there exists a constant matrix R such that one can write

$$C = e^{wR}\,. \tag{5}$$

(here, wR, is called a logarithm of C)

From equations (4) and (5), one obtains

$$\Phi(t+w) = \Phi(t)\,e^{wR}. \tag{6}$$

We define a matrix P(t) by the relation

$$P(t) = \Phi(t)\,e^{-tR}\,. \tag{7}$$

Then P(t) is a non – singular matrix as both $\Phi(t)$ and e^{-tR} are non – singular matrices. Moreover,

$$\begin{aligned}P(t+w) &= \Phi(t+w)\,e^{-(t+w)R},\\ &= \Phi(t)\,e^{wR}\,.\,e^{-(t+w)R},\\ &= \Phi(t)\,e^{-tR},\\ &= P(t).\end{aligned} \tag{8}$$

This shows that the matrix P(t) is periodic with period w. Further, from equation (7), we write

$$\Phi(t) = P(t)\,e^{tR}\,. \tag{9}$$

This completes the proof of the theorem.

Remarks:- (1) The representation of a fundamental matrix $\Phi(t)$, as given by (9), is of great interest. In this representations, $\Phi(t)$ has been expressed as a product of a periodic matrix P(t) with the same period w and matrix e^{tR}, where R is a constant matrix.

(2) Neither the matrix R nor its eigen values are uniquely determined by the given periodic system. On the other hand, the eigen values of e^R are uniquely determined by the given system.

(3) The eigenvalues of $C = e^{wR}$ are called the eigenvalues of the given system.

(4) The eigenvalues of R are called characteristic exponents.

Significance of theorem: Suppose a fundamental matrix Φ of given periodic is known over an interval of length w, say $0 \leq t \leq w$. Then, $\Phi(t)$ is at once determined over the entire domain $(-\infty, \infty)$ by relation (9). This process of extension is as follows.

(i) A constant non – singular matrix C is given by

$$C = \Phi^{-1}(0)\,\Phi(w)\,.$$

(ii) A constant matrix R is given by

$$R = \frac{1}{w}\,(\log C)\,.$$

(iii) A periodic matrix P(t) is now determined by the relation

$$P(t) = \Phi(t)\, e^{-tR} \text{ over the interval } (0, w) .$$

(iv) Since, P(t) is periodic with period w, so P(t) is determined at once over the entice interval $(-\infty, \infty)$.

(v) Consequently, $\Phi(t)$ is determined over $(-\infty, \infty)$ through the relation (9).

Theorem 11.3. Find the explicit form that a set of n linearly independent solution vectors of the periodic system $\frac{dy}{dt}$ A(t)y, $-\infty < t < \infty$,

assumes.

Solution:- Let w be the period of the given system. Let $\Phi(t)$ be a fundamental matrix of the given periodic solutions. Then, by the representation theorem (11.2),

$$\Phi(t) = P(t)\, e^{tR} , \qquad (1)$$

where P(t) is a periodic non-singular matrix with period w and R is a constant matrix. Suppose R is similar to a matrix J of the form

$$J = \begin{bmatrix} J_0 & 0 & 0 \ldots\ldots\ldots\ldots 0 \\ 0 & J_1 & 0 \ldots\ldots\ldots\ldots 0 \\ ------ & ------ & ------ \\ 0 & 0 & 0 \ldots\ldots\ldots\ldots J_s \end{bmatrix} \qquad (2)$$

where J_0 is a diagonal matrix with diagonal elements $\rho_1, \rho_2, \ldots\ldots, \rho_q$, and

$$J_i = \begin{bmatrix} \rho_{q+i} & 1 & 0 & 0 \ldots\ldots\ldots 0 & 0 \\ 0 & \rho_{q+i} & 1 & 0 \ldots\ldots\ldots 0 & 0 \\ ---- & ---- & ---- & ---- & ---- \\ 0 & 0 & 0 & 0 \ldots\ldots\ldots \rho_{q+i} & 1 \\ 0 & 0 & 0 & 0 \ldots\ldots\ldots 0 & \rho_{q+i} \end{bmatrix} \qquad (3)$$

for i = 1, 2, ……, s. Here, ρ_j, (j = 1, 2, …….., q + s), are the eigenvalues of the matrix R, which need not all be distinct. Since R is similar to J, there exists a constant non-singular matrix T such that

$$T^{-1} R T = J . \qquad (4)$$

Put $\qquad \Phi_1 = \Phi\, T . \qquad (5)$

Then

$$\begin{aligned} \Phi_1(t) &= \Phi\, T \\ &= \{P(t)\, e^{tR}\}T \\ &= \{P(t)T\}\{T^{-1}\, e^{tR}\, T\} \\ &= P_1(t)\, \{e^{t(T^{-1} R T)}\} \\ &= P_1(t)\, e^{tj} , \end{aligned} \qquad (6)$$

where $\qquad P_1 = P\, T, \qquad (7)$

is periodic with period w, i.e ,

$$\rho_1(t + w) = \rho_1(t) \qquad \text{for all t.} \qquad (8)$$

From equation (2), the matrix e^{tJ} will have the form

$$e^{tJ} = \begin{bmatrix} e^{tJ_0} & 0 \ldots\ldots\ldots 0 \\ 0 & e^{tJ_1} \ldots\ldots 0 \\ - - - - & - - - - - - \\ 0 & 0 \ldots\ldots e^{tJ_s} \end{bmatrix}, \tag{9}$$

where

$$e^{tJ_0} = \begin{bmatrix} e^{t\rho_1} & 0 \ldots\ldots\ldots 0 \\ 0 & e^{t\rho_2} \ldots\ldots 0 \\ - - - - & - - - - - - \\ 0 & 0 \ldots\ldots e^{t\rho_q} \end{bmatrix}, \tag{10}$$

$$e^{tJ_i} = e^{t\rho_{q+i}} \begin{bmatrix} 1 & t \ldots\ldots\ldots \dfrac{t^{ri-1}}{(ri-1)!} \\ 0 & 1 \ldots\ldots\ldots \dfrac{t^{ri-2}}{(ri-2)!} \\ - - - - & - - - - - - - - \\ 0 & 0 \ldots\ldots\ldots\ldots 1 \end{bmatrix}, \qquad i = 1,2,\ldots,s \quad q + \sum_i r_i = n \tag{11}$$

Clearly $\lambda_i = e^{w\rho_i}$. (12)

While the ρ_i are not uniquely determined, but real parts are. From (6), it follows that the columns $\varphi_1, \varphi_2, \ldots\ldots \varphi_n$ of Φ_1, which form a set of n linearly independent solutions of the given periodic system, are of the form

$$\varphi_1(t) = e^{t\rho_1} p_1(t)$$

$$\varphi_2(t) = e^{t\rho_2} p_2(t)$$

$$- - - - - - - - -$$

$$\varphi_q(t) = e^{t\rho_q} p_q(t)$$

$$\varphi_{q+1}(t) = e^{t\rho_{q+1}} p_{q+1}(t)$$

$$\varphi_{q+2}(t) = e^{t\rho_{q+1}} \{ t p_{q+1}(t) + p_{q+2}(t) \}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\varphi_{q+r_1}(t) = e^{t\rho_{q+1}} \left\{ \frac{t^{r_1-1}}{(r_1-1)!} p_{q+1}(t) + \ldots\ldots\ldots\ldots + t q_{q+r_1-1} + p_{q+r_1}(t) \right\}$$

$$\varphi_{n-r_s+1}(t) = e^{t\rho_{q+s}} p_{n-r_s+1}(t)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\varphi_n(t) = e^{t\rho_{q+s}} \left\{ \frac{t^{r_s-1}}{(r_s-1)!} p_{n-r_s+1}(t) + \ldots\ldots\ldots\ldots + p_n(t) \right\}, \tag{13}$$

where $p_1, p_2, \ldots\ldots\ldots p_n$ are the periodic column vectors of P_1.
Hence the result.

12

HIGHER ORDER LINEAR EQUATIONS

Let $a_1(t)$, $a_2(t)$,……,$a_n(t)$, b(t) be continuous, real – or complex valued functions defined on a real t – interval I = [a, b]. Now, we shall be considering the linear homogeneous differential equation

$$u^{(n)} + a_1(t)u^{(n-1)} + \dots + a_{n-1}(t)u^{(1)} + a_n(t)u = 0, \quad (1)$$

and the corresponding inhomogeneous equation

$$u^{(n)} + a_1(t)u^{(n-1)} + \dots + a_{n-1}(t)u^{(1)} + a_n(t)u = b(t). \quad (2)$$

The treatment of these nth order linear differential equations reduces to the systems

$$\frac{dy}{dt} = A(t)\, y, \quad (3)$$

and

$$\frac{dy}{dt} = A(t)\, y + f(t), \quad (4)$$

where

$$y = (u^{(0)}, u^{(1)}, \dots, u^{(n-1)})^T, \qquad u^{(0)} = u, \quad (5)$$

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \dots\dots 0 \\ 0 & 0 & 1 & 0 \dots\dots 0 \\ - - - & - - - & - - - & - - - \\ 0 & 0 & 0 & 0 \dots\dots 1 \\ -a_n & -a_{n-1} & \dots\dots & -a_1 \end{bmatrix} \quad (6)$$

$$f(t) = (0, 0, \dots, 0, b(t))^T. \quad (7)$$

We shall summarize the essential facts for this important case, in detail.

Initial value problem

The I V P consisting of differential equation (1) together with initial conditions

$$u(t_0) = u_0,\ u'(t_0) = u_1, \dots, u^{(n-1)}(t_0) = u_{n-1}, \text{ for } t_0 \in I,$$

where $u_0, u_1, \dots, u_{n-1}$ are arbitrary numbers, has a unique solution u = u(t) on the interval I = [a, b]. In particular, if

$$u_0 = 0,\ u_1 = 0, \dots, u_{n-1} = 0,$$

then

$$u(t) \equiv 0 \text{ on } [a, b].$$

Wronskian

Let $\phi_1(t)$, $\phi_2(t)$, ….,$\phi_n(t)$ be n solutions of the linear differential equation (1). Then, the matrix

$$\Phi = \begin{bmatrix} \phi_1 & \phi_2 \dots\dots \phi_n \\ \phi_1' & \phi_2' \dots\dots \phi_n' \\ - - - & - - - \\ \phi_1^{(n-1)} & \phi_2^{(n-1)} \dots \phi_n^{(n-1)} \end{bmatrix} \quad (8)$$

is known as a **solution matrix for equation** (1).

The determinant of this matrix Φ is called the **Wronskian** of equation (1) w.r.t. solutions $\varphi_1, \varphi_2, \ldots, \varphi_n$. It is denoted by

$$W(\varphi_1, \varphi_2, \ldots, \varphi_n).$$

It is a function of t on the interval I for fixed functions $\varphi_1, \varphi_2, \ldots, \varphi_n$. Its value at t is denoted by

$$W(t) = W(\varphi_1, \varphi_2, \ldots, \varphi_n)(t).$$

As shown/discussed earlier, we have

$$W(t) = W(t_0) \exp\left[\int_{t_0}^{t} tr(A(s))ds\right], \qquad t \in I$$

and $\quad tr\{A(s)\} = -a_1(s).$

So, $$W(t) = W(t_0) \exp\left[-\int_{t_0}^{t} (a_1(s))ds\right], \qquad \text{for } t \in I. \tag{9}$$

Remark. We denote

$$L_n \equiv \frac{d^n}{dt^n} + a_1(t)\frac{d^{n-1}}{dt^{n-1}} + \ldots + a_{n-1}(t)\frac{d}{dt} + a_n(t)$$

Then equation (1) is expressible as

$$L_n u = 0 . \tag{L}$$

Theorem (12.1): A necessary and sufficient condition that n solutions $\varphi_1, \varphi_2, \ldots, \varphi_n$ of differential equation

$$L_n\, u = 0 \tag{L}$$

on an interval I be linearly dependent there is that

$$w(\varphi_1, \varphi_2, \ldots, \varphi_n)(t) = 0 \qquad \text{for all } t \in I.$$

Further, show that every solution of differential equation (L) is a suitable linear combination of any n linearly independent solutions of (L).

Proof:- Let $\varphi_1, \varphi_2, \ldots, \varphi_n$ be linearly dependent on I. There, there exists constants $c_1, c_2, \ldots, c_n$; not all zero, such that

$$c_1\varphi_1(t) + \ldots.. + c_n\,\varphi_n(t) = 0 \qquad \text{for all } t \in I \tag{1}$$

Consequently,

$$c_1\varphi_1^{(k)}(t) + c_2\varphi_2^{(k)}(t) \ldots.. + c_n\,\varphi_n^{(k)}(t) = 0 \qquad \text{for all } t \in I \text{ and } k = 1, 2, \ldots, n-1 \tag{2}$$

This is a homogeneous system of linear equations which has a non zero solution. So, we must have

$$W(\varphi_1, \varphi_2, \ldots, \varphi_n)(t) = 0 \qquad \text{for all } t \in I, \tag{3}$$

as constants c_i are all not zero.

This proves that the condition is necessary.

Now, assume that the condition is satisfied. Then, the homogeneous matrix equation

$$\Phi\, C = 0, \tag{4}$$

has a non zero solution since

$$W = \det \Phi(t) = 0 \qquad \text{for all } t \in I. \tag{5}$$

Let $$C = (k_1, k_2, \ldots\ldots, k_n)^T, \tag{6}$$

be a non-zero solution of system (4).

We define a function f(t) as

$$f(t) = k_1\,\varphi_1(t) + k_2\,\varphi_2(t) + \ldots\ldots + k_n\,\varphi_n(t)\,, \qquad t \in I. \tag{7}$$

Then f(t) is a solution of equation (L) satisfying the initial conditions

$$f^{(k)}(t_0) = 0, \quad \text{for } k = 0, 1, 2, \ldots, n-1, \quad t_0 \in I. \tag{8}$$

By uniqueness theorem, we must have

$$f(t) \equiv 0, \quad \text{for all } t \in I.$$

This implies

$$k_1 \varphi_1(t) + k_2 \varphi_2(t) + \ldots + k_n \varphi_n(t) = 0 \quad \text{for } t \in I. \tag{9}$$

Hence, solutions $\varphi_1, \varphi_2, \ldots, \varphi_n$ are linearly dependent on I.

We know that every solution of vector system

$$\frac{dy}{dt} = A(t)\, y \quad \text{on } I, \tag{10}$$

where

$$A = \begin{pmatrix} 0 & 1 & 0 \ldots\ldots\ldots 0 \\ 0 & 0 & 1 \ldots\ldots\ldots 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \ldots\ldots\ldots 1 \\ -a_n & -a_{n-1} & \ldots\ldots\ldots -a_1 \end{pmatrix} \tag{11}$$

is a linear combination of n linearly independent vector solutions, and equations (L) and (10) are equivalent.

So, every solution of $L_n u = 0$ is a linear combination of any n linearly independent solution of $L_n u = 0$. This completes the proof.

Definition: A set of n linearly independent solutions of differential equation (L) is called a **basis/a fundamental set**.

Cor1: A necessary and sufficient condition that n solutions $\varphi_1, \varphi_2, \ldots, \varphi_n$ of

$$L_n u = 0 \quad \text{on } I$$

be linearly independent is that

$$W(t) \neq 0 \quad \text{for } t \in I.$$

Cor. 2 : If $\varphi_1, \varphi_2, \ldots, \varphi_n$ are n solutions of

$$L_n u = 0 \quad \text{on } I,$$

then $W(\varphi_1, \varphi_2, \ldots, \varphi_n)(t)$ is either identically zero on I or nowhere zero.

Theorem (12.2) : Suppose $\varphi_1, \varphi_2, \ldots, \varphi_n$ are n functions which possess continuous nth order derivatives on a real t-interval I, and

$$W(\varphi_1, \varphi_2, \ldots, \varphi_n)(t) \neq 0 \quad \text{on } I.$$

Then there exists a unique homogeneous differential equation of order n for which these functions form a fundamental set, namely,

$$(-1)^n \frac{W(u, \varphi_1, \varphi_2, \ldots, \varphi_n)}{W(\varphi_1, \varphi_2, \ldots, \varphi_n)} = 0.$$

Proof:- Consider the equation

$$W(u, \varphi_1, \varphi_2, \ldots, \varphi_n) = 0.$$

In the determinant form, it is

$$\begin{vmatrix} u & \varphi_1 & \varphi_2 \ldots\ldots\varphi_n \\ \frac{du}{dt} & \varphi_1' & \varphi_2' \ldots\ldots\varphi_n' \\ M & & \\ \frac{d^n u}{dt^n} & \varphi_1^{(n)} & \varphi_2^{(n)} \ldots\ldots\varphi_n^{(n)} \end{vmatrix} = 0, \qquad (1)$$

on interval I. On expanding by the first column, we see that equation (1) is a linear differential equation in u and coefficient of $\frac{d^n u}{dt^n}$ is $(-1)^n W(\varphi_1, \varphi_2,\ldots,\varphi_n)$, which is not zero by hypothesis. Hence, equation (1) is a linear homogeneous differential equation of order n. From (1), we see that

$$u = \varphi_i(t), \qquad t \in I, \quad i = 1, 2,\ldots,n \qquad (2)$$

are n solutions of equation (1) as two columns of (1) then become identical in the determinant on left side. In view of the hypothesis that

$$W(\varphi_1, \varphi_2,\ldots,\varphi_n) \neq 0 \quad \text{on I}, \qquad (3)$$

It follows that solutions $\varphi_1, \varphi_2,\ldots,\varphi_n$ form a fundamental set for differential equation (1) on the interval I.

The uniqueness of the equation follows from the fact that the corresponding vectors $\hat{\varphi}_i$ with components $\varphi_i^{(0)}, \varphi_i^{(1)},\ldots,\varphi_i^{(n-1)}$ determine the coefficient matrix uniquely of the associated system

$$\frac{dy}{dt} = A(t)\, y, \qquad (4)$$

with

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & 0\ldots\ldots\ldots.0 \\ 0 & 0 & 1 & 0\ldots\ldots\ldots.0 \\ - & - & - & - \\ 0 & 0 & 0 & 0\ldots\ldots\ldots..1 \\ -a_n & -a_{n-1} & -a_{n-2} & \ldots\ldots\ldots\ldots.-a_1 \end{bmatrix}, \qquad (5)$$

when the given equation is expressed as

$$\frac{d^n u}{dt^n} + a_1(t)\frac{d^{n-1}u}{dt^{n-1}} + a_2(t)\frac{d^{n-2}u}{dt^{n-2}} + \ldots\ldots + a_n(t)\, u = 0, \qquad t \in I. \quad (6)$$

We know that there is a one-to-one correspondence between linear equations of order n and linear system of type (4) and (5). This completes the proof.

Reduction of order

A direct procedure is suggested by the following process, which is the variation of constants adapted to

$$\frac{d^n u}{dt^n} + a_1(t)\frac{d^{n-1}u}{dt^{n-1}} + \ldots\ldots + a_n(t)\, u = 0. \qquad (1)$$

Let φ_1 be a known solution of differential equation (1). Then, the substitution

$$y = v(t)\,\varphi_1(t), \qquad (2)$$

gives a linear differential equation of the nth order in v which has v = 1 as a solution since φ_1 is a solution of equation (1). Thus, the coefficient of v in the new transformed equation must be zero. Let

$$w = v' = \frac{dv}{dt}. \qquad (3)$$

Then, the above obtained equation is a differential equation in w and it is now of order (n-1).

If φ_2 is another solution of equation (1), which is independent of φ_1, then $\left(\frac{\varphi_2}{\varphi_1}\right)'$ is a solution of the (n-1)st order equation in w, which can, by a repetition of the above, be reduced to an equation of order (n-2), etc., and so on.

Adjoint Equations

Associated with the form operator

$$L_n = a_0(t)\frac{d^n}{dt^n} + a_1(t)\frac{d^{n-1}}{dt^{n-1}} + \ldots\ldots + a_n(t), \qquad t \in I \qquad (1)$$

there is another linear operator of order n, denoted by L_n^+ and **called the adjoint of L_n**, defined by

$$L_n^+ = (-1)^n \frac{d^n}{dt^n}[\ \bar{a}_0(t)\ .\ \ ] + (-1)^{n-1}\frac{d^{n-1}}{dt^{n-1}}[\ \bar{a}_1(t)\ .\ \ ] + \ldots..+\ \bar{a}_n(t), \qquad (2)$$

For $t \in I$. If g(t) is any function defined on I which is such that $\{\ \bar{a}_k\ g(t)\}$ has n-k derivatives on I (for k = 0, 1, 2,….,n), then

$$L_n^+ g = (-1)^n \frac{d^n}{dt^n}[\ \bar{a}_0(t)\ g(t)\ ] + (-1)^{n-1}\frac{d^{n-1}}{dt^{n-1}}[\ \bar{a}_1(t)\ g(t)\ ] + \ldots..+\ \bar{a}_n(t)\ g(t)\ . \qquad (3)$$

Definition: The differential equation

$$\bar{L}_n\ v = 0, \qquad t \in I, \qquad (4)$$

is called the adjoint equation to

$$L_n\ u = 0 \text{ on } I.$$

If $a_k(t) \in C^{n-k}$ on I and $\varphi(t)$ is a solution of equation (4) with n derivatives on I, then using the product rule of differentiation, we get

$$L_n^+\ \varphi = (-1)^n\ \bar{a}_0\ \varphi^{(n)} + \ldots\ldots\ldots = 0\ , \qquad (5)$$

and by dividing by $(-1)^n\ \bar{a}_0$, one sees that $\phi(t)$ is a solution of a differential equation of order n of the type considered earlier.

Special Case: When $a_0(t) = 1$ for all $t \in I$.
The nth order differential equation is

$$L_n\ u = u^{(n)} + a_1(t)\ u^{(n-1)} + \ldots..+ a_n\ u = 0, \qquad t \in I \qquad (6)$$

The system associated with (6) is

$$\frac{dy}{dt} = A(t)\ y, \quad t \in I \qquad (7)$$

with its adjoint system as

$$\frac{dy}{dt} = -A^*(t)\ y, \quad t \in I \qquad (8)$$

where $$A^* = -\begin{bmatrix} 0 & 0 \ldots\ldots\ldots 0 & \bar{a}_n \\ -1 & 0 \ldots\ldots\ldots 0 & \bar{a}_{n-1} \\ 0 & -1 \ldots\ldots\ldots 0 & \bar{a}_{n-2} \\ \vdots & \vdots \quad \vdots & \\ 0 & 0 \ldots\ldots\ldots 0 & \bar{a}_2 \\ 0 & 0 \ldots\ldots -1 & \bar{a}_1 \end{bmatrix}. \qquad (9)$$

In terms of components, equations (8) and (9) give

$$(y^1)' = \bar{a}_n y^n,\ (y^k)' = -y^{k-1} + \bar{a}_{n-k+1} y^n, \qquad (10)$$

where $y = (y^1, y^2, \ldots, y^n)$, $(k = 2, 3, \ldots, n)$.

Thus, if $\varphi^1, \varphi^2, \ldots, \varphi^n$ is a solution of system (10) for which $(\phi^k(t))^{(k)}$ and $[\bar{a}_{n-k+1}\, \phi^n(t)]^{(k-1)}$ exists, one obtains, by differentiating kth relation in equation (10), (k-1) times and solving for $(\varphi^n)^{(n)}$,

$$(\varphi^n)^{(n)} - (\bar{a}_1 \varphi^n)^{(n-1)} + \ldots\ldots + (-1)^n (\bar{a}_n \varphi^n) = 0. \qquad (11)$$

Therefore, φ^n satisfies the nth order differential equation

$$L_n^+ y = 0, \qquad (12)$$

which is just the adjoint equation to differential equation (6).

Remark : The importance of L_n^+ is due to an interesting relation connecting L_n and L_n^+, which is indispensible for the study of boundary value problems.

Theorem (12.3):(Lagrange's identity) :

In the nth order differential operator

$$L_n \equiv a_0(t) \frac{d^n}{dt^n} + a_1(t) \frac{d^{n-1}}{dt^{n-1}} + \ldots\ldots + a_n(t), \quad t \in I,$$

suppose $a_k(t) \in C^{n-k}$ on I (k = 0, 1, 2, …..,n). If u, v are two complex functions on I possessing n derivatives there, then

$$\bar{v} L_n u - u \overline{L_n^+ v} = \frac{d}{dt}\{P(u, v)\},$$

where P(u, v) is a form in $(u, u', \ldots\ldots, u^{(n-1)})$ and $(v, v', \ldots\ldots, v^{(n-1)})$ given by

$$P(u, v) = \sum_{m=1}^{n} \sum_{j=1}^{m} (-1)^{j-1}.\, u^{(m-j)} (a_{n-m}.\, \bar{v})^{j-1}.$$

Proof:- Consider the expression

$$U^{(m-1)} V - U^{(m-2)} V' + \ldots\ldots + (-1)^{m-2} U' V^{(m-2)} + (-1)^{m-1}.U.V^{(m-2)}$$

for m = 0, 1, 2,……,n.

Then
$$\frac{d}{dt}[U^{(m-1)} V - U^{(m-2)} V' + \ldots\ldots + (-1)^{m-2} U' V^{(m-2)} + (-1)^{m-1}.U.V^{(m-1)}]$$
$$= [U^{(m)} V + U^{(m)} V']$$
$$- \quad [U^{(m-1)} V' + U^{(m-2)} V'']$$
$$+ \ldots\ldots + (-1)^{m-2}[U'' V^{(m-2)} + U' V^{(m-1)}] + (-1)^{m-1}[U' V^{(m-1)} + U V^{(m)}]$$
$$= U^{(m)} V + (-1)^{(m-1)} U V^{(m)}, \quad \text{for } m = 0, 1, 2, \ldots, n. \qquad (1)$$

This implies

$$V U^{(m)} = (-1)^m V^{(m)} U + \frac{d}{dt}[U^{(m-1)} V - U^{(m-2)} V' + \ldots\ldots + (-1)^{m-2} U' V^{(m-2)} + (-1)^{m-1}.U.V^{(m-1)}]$$

for m = 0, 1, 2,….,n. (2)

Applying (2) with $U = u$, $v = a_0 \bar{v}$, $m = n$;

$U = u$, $v = a_1 \bar{v}$, $m = n - 1$;

$U = u,\ v = a_{n-1}\bar{v},\ m = 1;$

successively, we obtain

$$(a_0\bar{v})u^{(n)} = (-1)^n \cdot u \cdot (\bar{v} a_0)^{(n)} + \frac{d}{dt}\left[u^{(n-1)}(\bar{v} a_0) - u^{(n-2)}(\bar{v} a_0)' + \ldots + (-1)^{n-2} u'(\bar{v} a_0)^{(n-2)} + (-1)^{n-1} \cdot u \cdot (\bar{v} a_0)^{(n-1)}\right]$$

$$(a_1\bar{v})u^{(n-1)} = (-1)^{n-1} \cdot u \cdot (\bar{v} a_1)^{(n-1)} + \frac{d}{dt}\left[u^{(n-2)}(\bar{v} a_1) - u^{(n-3)}(\bar{v} a_1)' + \ldots + (-1)^{n-2} \cdot u \cdot (\bar{v} a_1)^{(n-2)}\right]$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$(a_{n-1}\bar{v})u' = -u \cdot (\bar{v} a_{n-1})' + \frac{d}{dt}[u \cdot (v\, a_{n-1})]$$

$$(a_n\bar{v})u = u(\bar{v} a_n).$$

Adding all these expressions vertically, one obtains

$$\bar{v} L_n u - u\,\overline{L_n^+ v} = \frac{d}{dt}\left[\sum_{j=1}^{n}(-1)^{j-1}u^{(n-j)}(a_0\bar{v})^{(j-1)}\right] + \frac{d}{dt}\left[\sum_{j=1}^{n-1}(-1)^{j-1}u^{(n-1-j)}(a_1\bar{v})^{(j-1)}\right] + \ldots + \frac{d}{dt}\left[\sum_{j=1}^{1}(-1)^{j-1}u^{(1-j)}(a_n\bar{v})^{(j-1)}\right].$$

This implies

$$\bar{v} L_n u - u\,\overline{L_n^+ v} = \frac{d}{dt}\left[\sum_{m=1}^{n}\left\{\sum_{j=1}^{m}(-1)^{j-1} \cdot u^{(m-j)} \cdot (a_{n-m} \cdot \bar{v})^{(j-1)}\right.\right] = \frac{d}{dt}\{P(u, v)\} \tag{3}$$

where

$$P(u, v) = \sum_{m=1}^{n}\left[\sum_{j=1}^{m}(-1)^{j-1} \cdot u^{(m-j)} \cdot (a_{n-m} \cdot \bar{v})^{(j-1)}\right] \tag{4}$$

This completes the proof.

Definition: $P(u, v)$ is called the **BILINEAR Concomitant associated with operator** L_n and (3) is called the **Lagranges identity**.

Corollary (Grean's Formula) :
If the a_k in L_n, and u, v are the same as in theorem (12.3) then for any $t_1, t_2 \in I$,

$$\int_{t_1}^{t_2}(\bar{v} L_n u - u\,\overline{L_n^+ v})\,dt = P(u, v)\,|_{t=t_2} - P(u, v)\,|_{t=t_1}.$$

Proof:- Integrating the Lagrange's identity from t_1 to t_2, we obtain the above result, known as **Green's formula**.

Applications of Lagrange's identity :

If ψ is a known non – trivial solution of

$$L_n^+ v = 0 \text{ on } I,$$

then the problem of finding a non – trivial solutions of the differential equation

$$L_n u = 0 \text{ on } I,$$

is reduced, by Lagrenge's identity, to finding a function φ on I satisfying an ordinary differential equation of order n – 1, namely.

$$P(u, \psi) = \text{constant},$$

i.e.,

$$\sum_{m=1}^{n}\left[\sum_{j=1}^{n} (-1)^{j-1} u^{(m-j)} (a_{n-m}.\overline{\psi})^{(j-1)} = c, \; c = \text{constant}.\right.$$

The Non-homogeneous linear differential equation of order n :

On a real t – interval I, suppose $a_1(t), \ldots, a_n(t)$ and b(t) are continuous functions, and consider the equation

$$u^{(n)} + a_1(t)u^{(n-1)} + \ldots + a_n(t)u = b(t) . \qquad (1)$$

The system associated with this equation is given by

$$\frac{dy}{dt} = A(t)y + f(t), \qquad t \in I, \qquad (2)$$

where A(t) is a matrix given by

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \ldots\ldots 0 \\ 0 & 0 & 1 & 0 \ldots\ldots 0 \\ 0 & 0 & 0 & 1 \ldots\ldots 0 \\ \hline -a_n & -a_{n-1} \ldots\ldots & & -a_1 \end{bmatrix}, \; f(t) = \begin{bmatrix} 0 \\ 0 \\ M \\ 0 \\ b(t) \end{bmatrix}. \qquad (3)$$

Thus, the system (2) associated with inhomogeneous equation (1), is linear and non-homogeneous. The existence and uniqueness of solutions of system (2) can be interpreted, as usual, as existence and uniqueness results for the nth order non-homogenous differential equation (1).

The Linear equation of order n with constant coefficient :

Now we consider the special case, when the coefficients

$$a_1, a_2, \ldots, a_n ,$$

are all constants. Then the interval I may be assumed to be the entire real t – axis, i.e., $I = (-\infty, \infty)$. In this case, the nth order differential equation

$$u^{(n)} + a_1 u^{(n-1)} + \ldots + a_n u = 0, \qquad \text{(LH)}$$

has its associate system as

$$\frac{dy}{dt} = Ay , \qquad (1)$$

where A is the constant matrix given by

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}. \tag{2}$$

The characteristic polynomial for the matrix A is

$$f(\lambda) = \det(\lambda E_n - A), \tag{3}$$

which is of degree n in λ.

Theorem (12.4) : The characteristic polynomial for the matrix, given above, is

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots\ldots + a_n.$$

Proof:- we shall prove it by induction.

For n = 1, A = ($-a_1$) and so

$$\begin{aligned} f(\lambda) &= \det(\lambda E_1 - A) \\ &= |\lambda + a_1| \\ &= \lambda + a_1. \end{aligned} \tag{1}$$

So the result is true for n = 1.

Assume that the result is true for n – 1. Then

$$f(\lambda) = \det(\lambda E_n - A) = \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ - & - & - & - & - & - \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & a_2 & \lambda + a_1 \end{vmatrix}. \tag{2}$$

Expanding (2) by the first column, we notice that the coefficient of λ is a determinant of order n – 1 which is equal to $\det(\lambda E_{n-1} - A_1)$, where

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ - & - & - & - & - \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & a_1 \end{bmatrix}. \tag{3}$$

Therefore, by assumption of the result for n – 1, we have

$$\det(\lambda E_{n-1} - A_1) = \lambda^{n-1} + a_1\lambda^{n-2} + \ldots\ldots + a_{n-1}. \tag{4}$$

The only other non – zero element in the first column is a_n whose cofactor in the determinant (2) is 1.

Hence, expansion of (2) becomes

$$\begin{aligned} f(\lambda) &= \det(\lambda E_n - A) \\ &= \lambda(\lambda^{n-1} + a_1\lambda^{n-2} + \ldots\ldots + a_{n-1}) + a_n \cdot 1 \\ &= \lambda^n + a_1\lambda^{n-1} + \ldots\ldots + a_{n-1} + a_n, \end{aligned}$$

which proves the result by induction.

Remark : This theorem (12.4) shows that $f(\lambda)$ can be obtained from $L_n u$ by formally changing $u^{(k)}$ to λ^k .

Theorem (12.5) :- Let $\lambda_1, \lambda_2, \ldots\ldots \lambda_s$, be the distinct roots of the characteristic equation

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \ldots\ldots + a_n = 0,$$

and suppose λ_i has multiplicity m_i (i = 1, 2, ……..,s). Then a fundamental set for the nth order homogeneous linear equation

$$L_n u = u^{(n)} + a_1 u^{(n-1)} + \ldots\ldots + a_n u = 0$$

is given by the n functions

$$t^k\, e^{t\lambda_i}, \qquad (k = 0, 1, 2, \ldots\ldots m_{i-1};\ i = 1, 2, \ldots\ldots, s).$$

Proof:- We shall prove the theorem directly. From the theory of polynomial equations, we know that if λ_i is a root of $f(\lambda) = 0$ with multiplicity m_i, then λ_i is also a root of the (m – 1) equations

$$f'(\lambda) = 0,\ f''(\lambda) = 0,\ \ldots\ldots,\ f^{(m-1)}(\lambda) = 0 . \qquad (1)$$

It is clear that

$$L_n(e^{t\lambda}) = f(\lambda)e^{t\lambda}, \qquad (2)$$

And in general

$$L_n(t^k e^{t\lambda}) = L_n\left[\frac{\partial^k}{\partial\lambda^k} e^{t\lambda}\right]$$

$$= \frac{\partial^k}{\partial\lambda^k}\left[L_n(e^{t\lambda})\right]$$

$$= \frac{\partial^k}{\partial\lambda^k}\left[f(\lambda)e^{t\lambda}\right],$$

$$= \left[f^{(k)}(\lambda) + k f^{(k-1)}(\lambda)t + \frac{k(k-1)}{2!} f^{(k-2)}(\lambda)t^2 + \ldots + f(\lambda)t^k\right]e^{t\lambda}, \qquad (3)$$

using Leibnitz rule. From equations (1) and (3), it follows that, for any fixed i,

$$L_n(t^k e^{t\lambda_i}) = 0, \qquad (4)$$

for k = 0, 1, 2, ……, m_{i-1}. This proves that functions $t^k\, e^{t\lambda_i}$ are solutions of nth order homogeneous equation.

Now, it will be proved that these functions are linearly independent. If possible, suppose that they are not linearly independent. There exists constants c_{ik}, not all zero, such that

$$\sum_{i=1}^{s}\sum_{k=0}^{m_i-1} c_{ik} t^k\, e^{t\lambda_i} = 0$$

or

$$\sum_{i=1}^{\sigma} P_i(t)\, e^{t\lambda_i} = 0\,, \qquad (5)$$

where the $P_i(t)$ are polynomials and $\sigma \le s$ is chooses so that

$$P_\sigma(t) \neq 0 \text{ while } P_{\sigma+i}(t) \equiv 0 \text{ for } i \ge 1. \qquad (6)$$

Divide the above expression by $e^{t\lambda_i}$ and differentiate enough times so that the polynomial $P_1(t)$ becomes zero. Note that the degrees and the non - identically vanishing nature of the polynomials multiplying $e^{(\lambda_i - \lambda_1)t}$, i > 1, do not change under this operation. Thus there are results

$$\sum_{i=2}^{\sigma} Q_i(t)\, e^{t\lambda_i} = 0, \qquad (7)$$

where $Q_i(t)$ has the same degree as $P_i(t)$ for $i \ge 2$.

Repeating the procedure results finally in a polynomial F(t) of a degree equal to that of $P_\sigma(t)$ such that

$$F(t) = 0 \text{ for all } t. \qquad (8)$$

This is impossible, since a polynomial (of finite degree) can vanish only at isolated points.

Thus, the solutions (3) are linearly independent. This completes the proof.

Books recommended for reading for chapters 9 - 12 are

(1)	S.L. Ross	Differential Equations
(2)	E.A. Coddington and N.Levinson	Theory of Ordinary Differential Equations
(3)	P. Hertman	Differential Equations.

13

NONLINEAR DIFFERENTIAL EQUATIONS, PLANE AUTONOMOUS SYSTEMS

There have been two major trends in the historical development of differential equations.

The firs and the oldest is to find explicit solutions, either in closed form-which is rarely possible – or in terms of power series.

In the second, one abandons all hopes of solving equations in any traditional sense, and instead concentrates on a search for qualitative information about the general behaviour of solution.

The qualitative theory of nonlinear equations is totally different. It was founded by **Poincare around 1880,** in connection with his work in **celestial mechanics**. Very little of a general nature is known about nonlinear equations.

Why should one be interested in nonlinear differential equation? The basic reason is that many physical systems and the equations that describe them are simply nonlinear from the onset. The usual linearization are approximating devices that are partly confessions or defeat of the practical view that half a loaf is better than none.

Since any higher order differential equation can be transformed into a system of first order equations, we will restrict ourselves to such systems.
Consider the first order system.

$$\frac{dy}{dt} = f(t, y)\ , \qquad \ldots(1)$$

where, to avoid unnecessary complications, we shall suppose that $f(t, y)$ is defined and continuous for all y and all $t \geq t_0$, and satisfies a Lipschitz condition in y in any bounded domain. Then, for the initial-value problem.

$$y(t_0) = y_0\ , \qquad \ldots(2)$$

the uniqueness and existence theorems show that there is a unique solution

$$y = y\ (t\ ;\ y_0\ ,\ t_0) \qquad (t_0 \leq t < T), \qquad \ldots(3)$$

where $y(t\ ;\ y_0,\ t_0)$ is defined for all $t \geq t_0$. This will not be a significant restriction as in the applications to follow y(t) will either be a constant or a periodic function of t.

Stability is connected with the question as to whether solutions which are in some sense close to y(t) at some instant will remain close for all subsequent times. Clearly, stability is a desirable property is dynamical processes, modelled here by the system of equations (1), are often subject to small, unpredictable disturbances. Unstable solutions are thus extremely difficult to realize either experimentally or numerically, as an arbitrarily small disturbance will eventually cause large deviation from the unstable solution.

Autonomous systems and the phase plane

Definition. A system of differential equations

$$\frac{dy}{dt} = f(y)\ , \qquad \ldots(1)$$

in which the independent variable t does not occur explicitly, is called **autonomous**. Thus, the characteristic property of autonomous systems is that the function f do not depend on the independent variable. When the variable t is thought of as representing time, autonomous system are thus steady stationary.

The main point of this topic deals with the geometry of solutions of differential equations on a plane, i.e., n = 2. That is, the vector y in system (1) is two-dimensional.

Definition. A point $y = y_0$ is called **a singular or stationary point** of the autonomous system (1) if

$$f(y_0) = 0, \qquad \ldots(2)$$

and **a regular point** if

$$f(y_0) \neq 0. \qquad \ldots(3)$$

Remark. The second order system (1), in general, correspond with second order non linear differential equation of the form

$$\frac{d^2x}{dt^2} = f\left(x, \frac{dx}{dt}\right). \qquad \ldots(1)$$

If we imagine a simple dynamical system (with one degree of freedom) consisting of a particle of unit mass (m = 1) moving on the x-axis, and if $f\left(x, \frac{dx}{dt}\right)$ is the force acting on it, then equation (1) is the equation of motion of the particle.

If we introduce the variable

$$y = \frac{dx}{dt}\ , \qquad \ldots(2)$$

then second order equation (1) is replaced by the equivalent first order autonomous two-dimensional system

$$\begin{cases} \frac{dx}{dt} = y \\ \frac{dy}{dt} = f(x, y) \end{cases} \qquad \ldots(3)$$

We shall see that a good deal can be learned about the solutions of (1) by studying the solutions of system (3).

Definitions : The values of x (position/distance) and $\frac{dx}{dt}$ (velocity), which at each instant characterize the state of the system, are called **its phases.** The plane of the variables x and $\frac{dx}{dt}$ is called the **phase plane.**

Note : When t is regarded as a parameter, then in general a solution of system (3) is a pair of functions

$$x = x(t), \qquad y = y(t) \qquad \ldots(4)$$

defining a curve in the xy-plane, which is simply the **phase plane** mentioned above. We shall be interested in the total picture/idea formed by these curves in the phase plane.

Remark. More generally, we study autonomous two dimensional systems of the form

$$\frac{dx}{dt} = F(x, y), \frac{dy}{dt} = G(x, y)$$

where F and G are continuous and have continuous first partial derivatives through the xy-plane, which is called the phase plane of the system.

Given a solution x = x(t), y = y(t) of the above system, we can plot the points (x(t), y(t)) in the phase plane, obtaining a graph or curve having x(t) and y(t) as parametric functions.

To add a dynamic element to the geometry, think of the point (x(t), y(t)) as moving along this curve as t increases, endowing the curve with a sense of direction or orientation.

Definition. The oriented locus of points (x(t), y(t)) in the phase plane, formed from a solution of the given two-dimensional plane autonomous system, is called **a path |orbit| trajectory** of the system.

Thus, a path is a directed curve in the phase plane. In figures, we will use arrows to indicate the direction in which the path is traced out as t increases.

Remark : Returning now to general considerations, we will state some facts about trajectories. Consider the plane autonomous system

$$\frac{dy}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y) \qquad \ldots(1)$$

where F and G are continuous with continuous first partial derivatives for all points (x, y) in the phase plane.

Fact 1. There is a trajectory through each point (x_0, y_0) in the phase plane.

This is true because the initial-value problem

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$
$$x(t_0) = x_0,$$
$$y_0(t_0) = y_0 \qquad (2)$$

has a solution. This solution yields a trajectory through the point (x_0, y_0) in the phase plane.

Fact 2. If x = φ (t), y = ψ(t) is a solution of the plane system (1) and c is a constant, then

$$x = \varphi(t + c),$$
$$y = \psi(t + c), \qquad (3)$$

is also a solution of the system (1). When solutions are related in this way, we call their trajectories **translations of each other**. These two trajectories have the same graph but (φ(t), ψ(t)) and (φ(t + c), ψ(t + c)) arrive at a given geometric point at different times. Infact, if the first arrives at P at time t_0, the second gets there at time $t_0 - c$.

Thus, the only way two trajectories of a plane system (1) can pass through the same point is if each is a translation of the other. That is, **two trajectories through the same point are translations of each other**.

Fact 3. A trajectory of a plane autonomous system (1) may not cross itself. The only exception to this occurs at those points of the phase plane where both F and G vanish. In the neighbourhood of such points (singular points), solutions exhibit particularly interesting behaviour.

Fact 4. Trajectories may, however, be closed curves. Closed paths represent periodic solutions of the system.

Critical point of a plane autonomous system.

Definition. A point (x_0, y_0) of the phase plane is called a critical point of the plane autonomous system (1) if

$$F(x_0, y_0) = 0, \text{ and } G(x_0, y_0) = 0.$$

A critical point is also sometimes called an equilibrium point/stationary point.

Note. At such a point, the **unique solution is the constant solution**.

$$x = x_0 \text{ , and } y = y_0 \text{ .}$$

A constant solution does not define a path, so, no path goes through a critical point.

Isolated Critical Point

Definition. A critical point (x_0, y_0) is called **isolated** if there is a disk of positive radius about (x_0, y_0) containing no other critical point of the system.

Remark 1. The means that there are no other critical points of the system arbitrarily close an isolated critical point to (x_0, y_0). We will only deal with isolated critical points, and so will take the phrase “critical point” to mean “isolated critical point”. We will also assume that F and G and their first partial derivatives are continuous throughout the plane, unless explicit exception is made.

Remark 2. Suppose now that (x_0, y_0) is a critical point of the given system, and consider the initial value problem consisting of the system and the conditions $x(t_0) = x_0$ $y(t_0) = y_0$.

Since $\quad x'(t_0) = F(x_0, y_0) = 0$,

and

$$y'(t_0) = G(x_0, y_0) = 0,$$

the trajectory through (x_0, y_0) can never leave this point. We conclude that a **trajectory through a critical point consists of just the single critical point**.

Remark 3 : Further, since different trajectories of an autonomous system cannot cross each other, no other trajectory can pass through a critical point. Thus, **a trajectory beginning at a noncritical point can never reach a critical point**. It may, however, approach arbitrarily close to a critical point, and it may do so in variety of ways. We will pursue this idea in the next section, where we will see that understanding the behaviour of trajectories near a critical point yields useful information about properties of solutions of the system.

We will now consider how do trajectories/ paths behave as $t\to\infty$ or as $t\to-\infty$. This is the question of asymptotic behaviour. To develop these ideas we will assume that $P_0(x_0, y_0)$ is an isolated critical point of the given plane autonomous system (1).

Definition (Trajectory approaching a critical point). Let x = x(t), y = y(t) be a solution which parametrically represents the path C, and let $P_0(x_0, y_0)$ be a critical point of the given plane autonomous system.

$$\frac{dx}{dt} = F(x, y),$$

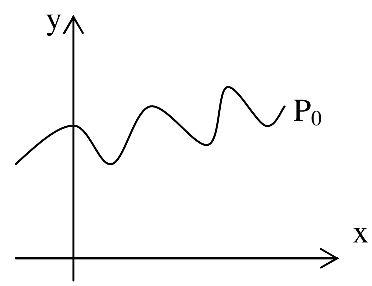
$$\frac{dy}{dt} = G(x, y) \ .$$

The path C approaches the critical point $P_0(x_0, y_0)$ if and only if

$$\lim_{t\to\infty} x(t) = x_0 \text{ and } \lim_{t\to\infty} y(t) = y_0 \ ,$$

or

$$\lim_{t\to-\infty} x(t) = x_0 \text{ and } \lim_{t\to-\infty} y(t) = y_0 \ .$$



Explanation. This definition is illustrated in the adjoining figure. Think of the trajectory as the path of a particle. For the trajectory to approach the critical point $P_0(x_0, y_0)$, the following must be true.

Given a circle C about P_0, we must be able to find a time, say, t_C such that the particle is within the disk enclosed by C at all times later than t_C, or at all times before t_C. Notice that, since the trajectory cannot cross a critical point, the particle never actually reaches P_0 − it simply comes arbitrarily close in the limit. Further, this limit is irrespective of the solution that is actually used to represent the trajectory. If $x = x_1(t)$, $y = y_1(t)$ is another solution defining the same trajectory, then one must have the same limits.

Definition (Trajectory entering a critical point). A trajectory {(x(t), y(t))} enters the critical point $P_0(x_0, y_0)$ of the plane autonomous system

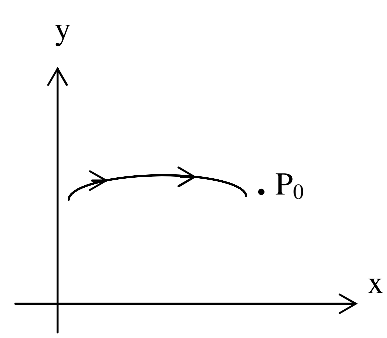
$$\frac{dx}{dt} = F(x, y),$$

$$\frac{dy}{dt} = G(x, y) \ ,$$

if and only if the trajectory approaches P_0, and also the quotient

$$\frac{y(t) - y_0}{x(t) - x_0} \ ,$$

has a finite limit as $t\to\infty$ or as $t\to-\infty$.



Explanation. The ratio in this definition is the slope of the straight line from the particle at the point P(x(t), y(t)) at time t to the critical point $P_0(x_0, y_0)$. For the trajectory to enter P_0, it must not only approach P_0, but must do so along a definite direction, given by this limit. That is, as the trajectory approaches P_0, it is more nearly moving along the line through P_0 having this slope. We note that in this case, the trajectory approaches P_0 along a specific direction.

Remark. Without loss of generality, **we will take the critical point (x_0 , y_0) to be the origin 0(0, 0) in further discussions.**

If necessary, we make use of the linear transformation

$$\xi = x - x_0,$$
$$\eta = y - y_0,$$

which transforms the point (x_0, y_0) of the xy-plane into the origin (0,0) in the $\xi\eta$-plane.

Theorem (3.1). For any continuously differential function V = V(x, y), each integral curve of the plane. autonomous system

$$\frac{dx}{dt} = \frac{\partial V(x,y)}{\partial y},$$
$$\frac{dy}{dt} = -\frac{\partial V(x,y)}{\partial x},$$

lies one some level curve

$$V(x, y) = \text{constant}.$$

Proof. Along any solution curve, we have

$$\begin{aligned}\frac{dV}{\alpha t} &= \frac{\partial V}{\partial x}.\frac{dx}{dt} + \frac{\partial V}{\partial y}.\frac{dy}{dt}\\ &= \frac{\partial v}{\partial x}.\frac{\partial v}{\partial t} - \frac{\partial v}{\partial y}.\frac{\partial v}{\partial x}\\ &= 0, \qquad \ldots(1)\end{aligned}$$

using the given plane autonomous system. Consequently

$$V(x(t), y(t)) = \text{constt.} \qquad \ldots(2)$$

The associated steady flow is divergence free, because

$$\begin{aligned}\text{div}\left(\frac{dx}{dt}\hat{i} + \frac{dy}{dt}\hat{j}\right) &= \text{div}\left(\frac{dv}{dy}\hat{i} - \frac{dv}{dx}\hat{j}\right)\\ &= \frac{\partial^2 V}{\partial x \partial y} - \frac{\partial^2 V}{\partial y \partial x}\\ &= 0. \qquad \ldots(3)\end{aligned}$$

In fluid mechanics, such a steady flow (1) is called **incompressible,** and the function V = V(x, y) is called its **stream function.** Also

$$\begin{aligned}&\left(\frac{dx}{dt}\hat{i} + \frac{dy}{dt}\hat{j}\right).\ \text{grad V}\\ &= \left(\frac{\partial V}{\partial y}\hat{i} + \frac{\partial V}{\partial x}\hat{j}\right)\left(\frac{\partial V}{\partial x}\hat{i} - \frac{\partial V}{\partial y}\hat{j}\right)\\ &= 0, \qquad \ldots(4)\end{aligned}$$

which shows that the level curves of V as the **orthogonal trajectories to the gradient lines of V (or path lines).**

The main advantage of the representation

$$\frac{dx}{dt} = \frac{\partial V}{\partial y},$$

$$\frac{\partial y}{\partial t} = -\frac{\partial V}{\partial x}\ , \qquad \ldots(5)$$

over the differential equation

$$\frac{dy}{dx} = -\frac{\partial V/\partial x}{\partial V/\partial y}\ , \qquad \ldots(6)$$

is the following. Whereas the solution curves of (6) terminate whenever $\frac{\partial V}{\partial y}$ vanishes, those of (5) terminate only where the function V has a critical point in the sense that grad V = 0. This happens exactly where the autonomous system (5) has critical points.

Illustration of theorem (13.1). If we set

$$V = -(x^2 + y^2)/2\ , \qquad (1)$$

we get the system

$$\frac{dx}{dt} = -y,\ ,$$

$$\frac{dy}{dt} = x\ , \qquad (2)$$

having circular streamlines.

If $\mu(x, y)$ is non-vanishing, then the system.

$$\frac{dx}{dt} = -\mu y,$$

$$\frac{dy}{dt} = \mu x\ , \qquad (3)$$

also has circles for solution curves. Thus, we can construct a wide variety of autonomous systems having the same solution curves in this way.

Example. Consider the autonomous system

$$\frac{dx}{dt} = -y, \quad \frac{dx}{dt} = x, \qquad (1)$$

whose solutions are

$$x(t) = r\cos(t+c),$$

$$y(t) = r\sin(t+c)\ , \qquad (2)$$

where r and c are arbitrary constants. The graphs of these solutions are concentric circles, with centre at the origin, whose equations are

$$x^2 + y^2 = r^2\ . \qquad (3)$$

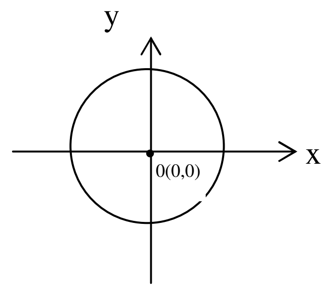
The corresponding first order equation, after eliminating t, is

$$\frac{dy}{dt} = \frac{x}{y}\ , \qquad (4)$$

with solutions

$$y(x) = \pm\sqrt{r^2 - x^2}\,, \qquad (5)$$

which are defined only for $|x| < |r|$. Whereas the function $-x/y$ is undefined where $y = 0$, the functions $F(x, y) = -y$ and $G(x, y) = x$ of the given system are defined throughout the plane. This gives an obvious advantage of the system (1) over the differential equation (4).
The circles (3) form a regular curve family in the **"punctured"** xy-plane, the critical point O(0,0) being deleted.



Example. For the autonomous system

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = -x\,, \qquad \ldots(1)$$

i) find the real critical points of the system ,
ii) obtain the differential equation that gives the slope of the tangent to the paths of the system,
iii) solve the above equation obtained in (ii) to obtain one parameter family of path.

Solution. We have

$$\frac{d^2x}{dt^2} = \frac{dy}{dt} = -x$$

$$\Rightarrow \qquad \frac{d^2x}{dt^2} + x = 0\,. \qquad \ldots(2)$$

Similarly

$$\frac{d^2y}{dt^2} + y = 0\,. \qquad \ldots(3)$$

The general solutions of equations (2) and (3) are, respectively,

$$x(t) = A\cos t + B\sin t\,, \qquad \ldots(4)$$
$$y(t) = C\cos t + D\sin t\,, \qquad \ldots(5)$$

where A, B, C, D are constants. From equations (1), (4) and (5), we get

$$\left.\begin{array}{l} C\cos t + D\sin t = -A\sin t + B\cos t \\ -A\cos t - B\sin t = -C\sin t + D\cos t \end{array}\right] \qquad \ldots(6)$$

This gives

$$C = B,\ D = -A\,. \qquad \ldots(7)$$

Hence, solutions of the system (1) are

$$\left.\begin{array}{l} x(t) = A\cos t + B\sin t \\ y(t) = B\cos t - A\sin t \end{array}\right], \qquad \ldots(8)$$

where A and B are arbitrary constants.

If these solutions are required to satisfy the supplementary initial conditions, say,

$$x(0) = 0,\ y(0) = 1\,, \qquad \ldots(9)$$

then, we shall get

$$A = 0,$$
$$B = 1.$$

Thus, solutions are

$$x(t) = \sin t,$$
$$y(t) = \cos t . \qquad \ldots(10)$$

Further, if the solutions for x and y satisfy the supplementary conditions, say,

$$x(0) = -1,\ y(0) = 0 ,$$

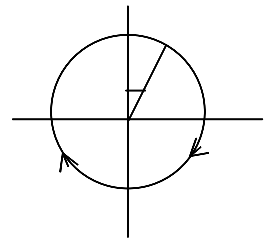
then they assume the form

$$x(t) = \sin (t - \pi/2),$$
$$y(t) = \cos(t - \pi/2) . \qquad \ldots(11)$$

We see that solutions given by (10) is different from the solution (11), but both of them define the same path is the xy-plane as the path is translatory invariant.

Eliminating the parameter t, we get

$$x^2 + y^2 = 1 . \qquad \ldots(12)$$



The path is a circle with centre (0, 0) and radius 1. As t increases from zero onwards, the path is traced in the clockwise direction.

The differential equations which gives the tangent to the path, say C, is obtained by eliminating t between the given equations. We find

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt},$$
$$= -\frac{x}{y} , \qquad \ldots(13)$$

provided $(x, y) \neq (0, 0)$. The one-parameter family of solutions of equation (13) is given by

$$x^2 + y^2 = \alpha^2,\ \alpha = \text{constt.} , \qquad \ldots(14)$$

which are the parameter family of paths of the given system in the phase plane. The only critical point of the given autonomous system (1) is the origin (0, 0) where F = G = 0 .

14

CLASSIFICATION OF CRITICAL POINTS AND THEIR STABILITY

We will now distinguish four kinds of critical points, according to the way trajectories behave in their vicinity. Let 0(0, 0) be an isolated critical point of a plane autonomous system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y) \qquad \ldots(1)$$

where F and G have continuous first partial derivatives for all (x, y) in the phase plane.

Type I : Center or Rotation Point

Definition. A critical points 0(0, 0) is said to be a **center** of the system (1) if there exists a neighbourhood of 0(0, 0) which contains a countably infinite number of closed paths $C_1, C_2, C_3, \ldots, C_n, \ldots$, each of which contain 0(0, 0) in its interior and diameter of C_n tends to zero as $n \to \infty$. However, 0(0, 0) is not approached by any path either as $t \to \infty$ or $t \to \infty$.

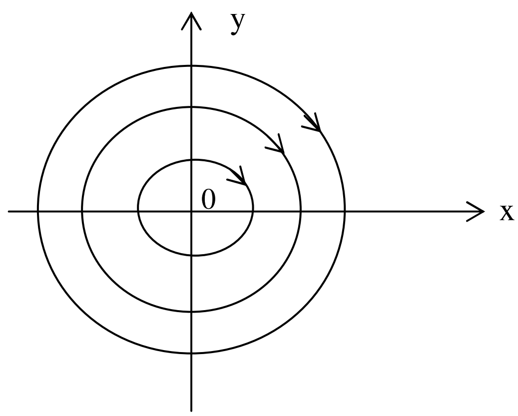


Fig. (14.1)

Explanation. In the above figure (14.1), the origin 0(0,0) is the critical point which is a center of the given system. It is surrounded by an infinite family of closed paths, members of which are arbitrarily near to 0(0, 0), but 0(0, 0) is not approached by any path as either $t \to \infty$ or as $t \to -\infty$.

Type II : Saddle Point

Definition. An isolated critical point 0(0, 0) of the system (1) is called a **saddle point** of the system (1) if there exists a neighbourhood of 0(0, 0) such that

(i) there exists two paths which approach and enter 0(0, 0) from opposite directions as $t \to \infty$ and there exists two other paths which approach and enter 0(0, 0) from different opposite directions as $t \to -\infty$,

y

(ii) in each of the four domains between any two of the four paths in (i) there are infinitely many paths which are arbitrarily close to 0(0, 0) but do not approach 0(0, 0) either as $t \to +\infty$ or as $t \to -\infty$.

Explanations. The figure (14.2) shows the critical point 0(0, 0) as **saddle point** which is such that

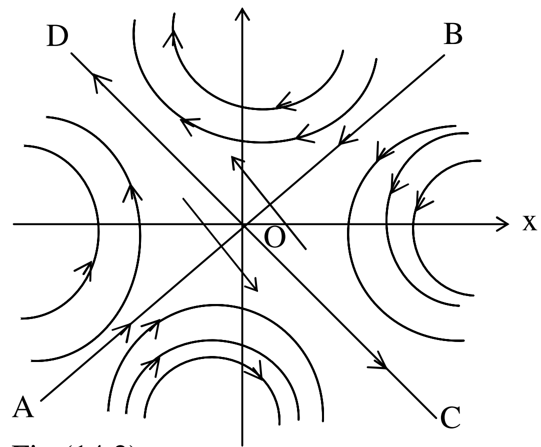


Fig (14.2)

(i) it is approached and entered by two half-line paths AO and BO as $t \to \infty$, these paths form the geometric curve AB, this solution tending to origin at $t \to +\infty$,

(ii) it is approached and entered by two half-line paths C0 and D0 as $t \to -\infty$, forming the geometric curve CD,

(iii) between the four half-line paths described in (i) and (ii), there are four domains R_1, R_2, R_3 and R_4 where each of these domains contains an infinite family of semi-hyperbolic paths which do not tend to 0(0, 0) as either $t \to +\infty$ or $t \to -\infty$, but which become asymptotic to one or another of the four half-line paths as $t \to +\infty$ or $t \to -\infty$.

Type 3. Spiral point (focal point)

Definition. The isolated critical point 0(0, 0) of the plane autonomous system (1) is called a **spiral point (or focal point)** if there exists a neighbourhood of 0(0, 0) such that every path C in this nbd has the following properties.

(i) C is defined for all $t > t_0$ (or for all $t < t_0$) for some number t_0 ;

(ii) C approaches 0(0, 0) as $t \to +\infty$ (or as $t \to -\infty$); and

(iii) C approaches 0(0, 0) in a spiral-like manner, winding around 0(0, 0) an infinite number of times as $t \to +\infty$ (or as $t \to -\infty$).

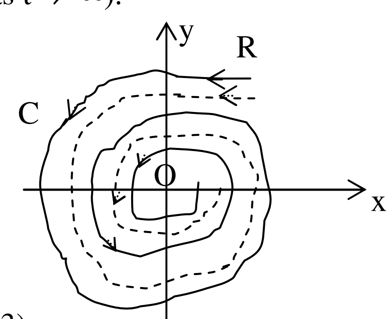


Fig. (14.3)

Explanation. The above figure (14.3) shows the critical point 0(0, 0) as a spiral/focal point which is approached in a spiral-like manner by an infinite family of paths as $t \to +\infty$ or as $t \to -\infty$. We observe that while the paths approach 0(0, 0) they do not enter it.

A point, say R, tracing such a path C approaches 0(0, 0) as $t \to +\infty$ (or as $t \to -\infty$), but the line OR does not tend to a definite direction, since the path C constantly winds about 0(0, 0).

Type (4) Node Point

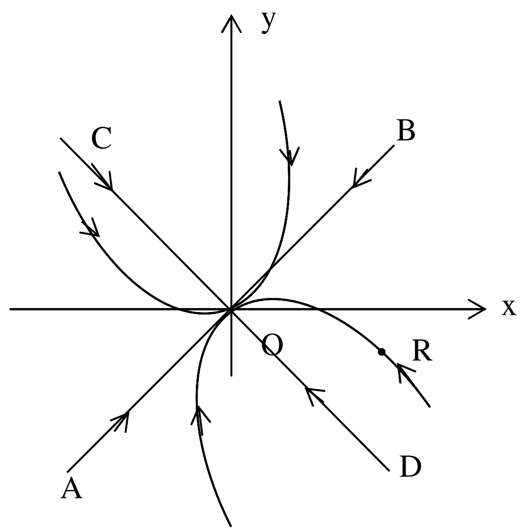
Definition. The isolated critical point 0(0, 0) of the plane autonomous system (1) is called a node if there exists a neighbourhood of 0(0, 0) such that every path C in this nbd has the following properties.

(i) C is defined for all $t > t_0$ (or for all $t < t_0$) for some number t_0 ;

(ii) C approaches 0(0, 0) as $t \to +\infty$ (or as $t \to -\infty$]; and

(ii) C enters 0(0, 0) as $t \to +\infty$ (or as $t \to -\infty$).

Explanation. The figure below shows the critical point 0(0, 0) as a **node point** which is not only approached, but also entered by an infinite family of paths as $t \to +\infty$ or as $t \to -\infty$.



(Fig. 14.4)

Here, a representative point R tracing such a path not only approaches 0 but does so in such a way that the line OR tends to a **definite direction** as $t \to +\infty$ or as $t \to -\infty$.

For the node shown here, there are Four "**rectilinear paths**" A0, B0, C0, and D0. All other paths are like **"semiparabolas".** As each of these semi parabolic like paths approaches 0 (0,0) its slope approaches that of line AB.

STABILITY OF CRITICAL POINTS

Let 0(0, 0) be a critical point of the plane autonomous system

$$\frac{dx}{dt} = F(x, y),$$

$$\frac{dy}{dt} = G(x, y) . \qquad \ldots(1)$$

We will now define what it means for a critical point 0(0, 0) of the system (1) to be stable. Intuitively, 0 is stable if trajectories that are close to 0 (0,0) at some time remain "close" at all later times.

Definition. Stable critical point

Let C be a path system (1) defined parametrically by its solution

$$x = f(t),\ y = g(t). \qquad \ldots(2)$$

Let $$D(t) = \sqrt{[f(t)]^2 + [g(t)]^2}, \qquad \ldots(3)$$

denote the distance OR between the critical point 0(0, 0) and the point R(f(t), g(t)) on C. Then the critical point 0(0, 0) is said to be **stable** if for each $\in > 0$, there exists a number $\delta > 0$ such that the following is true :

Every path C for which

$$D(t_0) < \delta \qquad \text{for some value } t_0 \qquad \ldots(4)$$

is defined for all $t \geq t_0$ and is that

$$D(t) < \in \qquad \text{for } t \leq + < \infty. \qquad \ldots(5)$$

Explanation. We now analyze all aspects of definition of the **stable point** in detail with reference to the figure (14.5) given below, where 0(0, 0) is the isolated critical point.

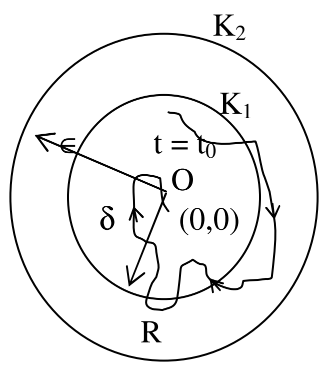


Fig. (14.5)

According to this definition, the point 0 (0,0) is stable if, corresponding to every positive real number $\in$, we can ensure the existence of another positive real number δ(depending on $\in$) which does "something" for us.

To explain "something", we must understand what the inequalities (4) and (5) means.

Accordingly to inequality (4), it means that the distance between the critical point 0 (0,0) and the point R_0 (corresponding to $t = t_0$) on the path C must be less than δ. This implies that the point R_0 lies within the circle K_1 of radius δ about 0 (0,0).

Likewise, the inequality (5) means that the distance between 0(0,0) and any point R on the path C, for $t \geq t_0$, is less than $\in$. Here, obviously $\delta \leq \in$. It implies that for all $t > t_0$, the points, like R, on the path C lie within the circle K_2 of radius $\in$ about 0(0, 0).

This explains the meaning of "something". When the critical point 0(0, 0) is stable, then every path C which is inside the circle K_1 at $t = t_0$, will remain inside the circle K_2 of radius $\in$, for all $t > t_0$.

Roughly speaking, if every path C stays as close to the critical point 0(0, 0) as we want it t_0 (i.e., within distance $\in$) after it once gets close enough (i.e., within distance δ), then the point 0(0, 0) is stable.

Definition. A critical point is called **unstable** if it is not stable

Illustrations. (i) For stable critical points, we point out the centre, the spiral (focal) point and node point in previous figures are all stable. Of these three, the focal point and node are asymptotically stable.

(ii) If the directions of the paths in figures for focal and node points has been reversed, then the spiral point and the node of these respective figures would have been unstable. The saddle point of the above figure is unstable.

Remark. Trajectories coming close to a stable point need not actually approach this point. But if they do approach this point, then we call the point asymptotically stable.

Definition. Asymptotically stable critical point.

Let 0(0, 0) be an isolated critical point of the plane autonomous system

$$\frac{dx}{dt} = F(x, y),$$

$$\frac{dy}{dt} = G(x, y) \quad . \qquad \ldots(1)$$

Let C be a path of system (1); and let

$$x = f(t),\ y = g(t) \qquad \ldots(2)$$

be a solution of system (1) representing C parametrically. Let

$$D(t) = \sqrt{[f(t)]^2 + [g(t)]^2}\ , \qquad \ldots(3)$$

denote the distance between the critical point 0(0, 0) and the point R(f(t), g(t)) on the curve C. The critical point 0(0, 0) **is called asymptotically stable** if

(i) it is stable and

(ii) there exists a number $\delta_0 > 0$ such that if

$$D(t_0) < \delta_0\ , \qquad \ldots(4)$$

for some value of t_0, then

$$\lim_{t\to+\infty} f(t) = 0,$$

$$\lim_{t\to+\infty} g(t) = 0\ . \qquad \ldots(5)$$

To analyse this definition, we note that condition (i) requires that the critical point 0(0, 0) must be stable. That is, every path C will stay as close to (0, 0) as we desire after it once gets sufficiently close.

But asymptotic stability is a stronger condition than mere stability. For, in addition to stability, the condition (ii) requires that every path that gets sufficiently close to (0, 0) ultimately approaches (0, 0) as $t\to+\infty$.

Linear plane autonomous system

Eventually, we want to use these concepts to study the behaviour of solutions of nonlinear systems of differential equations. However, for linear systems, it is possible to state a definite criterion, which we will use later when we linearize problems.

Consider the linear system

$$\frac{dx}{dt} = ax + by,$$
$$\frac{dy}{dt} = cx + dy, \qquad \ldots(1)$$

where a, b, c, d are real numbers, and

$$ad - bc \neq 0 . \qquad \ldots(2)$$

It is obvious that 0(0, 0) is the only critical point of the system (1), hence 0 (0,0) is an isolated critical point of it. We now seek solutions of the system (1) of the form

$$x(t) = c_1 e^{\lambda t},$$
$$y(t) = c_2 e^{\lambda t} , \qquad \ldots(3)$$

where c_1, c_2 are arbitrary constants and λ is a parameter. The substitution of (3) into system (1) at once leads to the quadratic equation in λ, namely,

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0, \qquad \ldots(4)$$

which is called the **characteristic equation of the linear plane autonomous system (1)**. By virtue of condition (2), $\lambda = 0$ is not a root of quadratic equation (4).

Let λ_1 and λ_2 be the roots of the characteristic equation (4). We now state and prove the following theorem which shall determine the nature of the critical point 0(0, 0).

Theorem (14.1). The critical point 0(0, 0) of the linear system (1) is

(i) a node point if λ_1 and λ_2 are real, unequal, and of the same sign ;

(ii) a saddle point if λ_1 and λ_2 are real, unequal, and of the opposite sign ;

(iii) a node point if λ_1 and λ_2 are real and equal;

(iv) a spiral point if λ_1 and λ_2 are conjugate complex with real part not zero;

(v) a center if λ_1 and λ_2 are pure imaginary.

The stability of the critical point 0(0, 0) of the linear system (1) is determined by the following theorem.

Theorem (14.2). The critical point 0(0, 0) of the linear system (1) is

(i) asymptotically stable if λ_1 and λ_2 are real and negative or conjugate complex with negative real parts ;

(ii) a stable, but not asymptotically stable if λ_1 and λ_2 are pure imaginary ;

(iii) unstable, if either of λ_1, λ_2 is real and positive or if λ_1 and λ_2 are conjugate complex with positive real parts.

The results of the above theorems (14.1) and (14.2) are summarized in the following table.

Table

Nature of roots λ_1 and λ_2 of characteristic equation $\lambda^2 - (a+d)\lambda + (ad - bc) = 0$	Nature of critical point 0(0, 0) of linear system $\frac{dx}{dt} = ax + by$, $\frac{dy}{dt} = cx + dy$	Stability of critical point 0(0, 0)
Real, unequal, and of same sign	Node	Asymptotically stable if roots are negative; unstable if roots are positive
Real, unequal, and of opposite	Saddle point	Unstable

sign		
Real and equal	Node	Asymptotically sable if roots are negative; unstable if roots are positive
Conjugate complex but not pure imaginary	Spiral point	Asymptotically stable if real part of roots is negative; unstable if real part of roots is positive
Pure imaginary	Center	Stable, but not asymptotically stable

Example. Determine the nature of the critical point 0(0, 0) of the linear system

$$\frac{dx}{dt} = 2x - 7y,$$

$$\frac{dy}{dt} = 3x - 8y.$$

Also determine whether or not the point is stable.

Solution. The given plane autonomous system is linear and 0(0, 0) is the only critical point of it. We seek solution of the system of the type

$$\left.\begin{aligned} x(t) &= c_1 e^{\lambda t} \\ y(t) &= c_2 e^{\lambda t} \end{aligned}\right] , \qquad \ldots(1)$$

where c_1 and c_2 are arbitrary constants and λ is a parameter. Substituting (1) in the given system, we get at once

$$\lambda c_1 = 2c_1 - 7c_2 \text{ and } \lambda c_2 = 3c_1 - 8c_2 .$$

This implies

$$(\lambda-2)\, c_1 + 7\, c_2 = 0$$

$$3c_1 + (-8-\lambda)\, c_2 = 0 \qquad \ldots(2)$$

To have a non trial solution of linear homogeneous system (2), we must have

$$\begin{vmatrix} \lambda-2 & 7 \\ 3 & -8-\lambda \end{vmatrix} = 0 .$$

This implies

$$\lambda^2 + 6\lambda + 5 = 0 , \qquad \ldots(3)$$

which is the characteristic equation of the given problem. Its roots are

$$\lambda_1 = -5,$$

$$\lambda_2 = -1 , \qquad \ldots(4)$$

which are real, unequal, and of the same sign (both negative). Therefore, the critical point 0(0, 0) of the given plane autonomous system is a NODE point. Since, λ_1 and λ_2 are real and negative, therefore, by Table above, the critical point 0(0, 0) is **asymptotically stable**. Hence the result.

Example. Determine the nature of the critical point (0, 0) of the linear system

$$\frac{dx}{dt} = 2x + 4y,$$

$$\frac{dy}{dt} = -2x + 6y$$

and determine whether or not the point is stable.

Solution. It is obvious that 0(0, 0) is the only critical point of the given plane autonomous system. We seek a solution of the given system of the type

$$x(t) = c_1\, e^{\lambda t},$$

$$y(t) = c_2 e^{\lambda t}, \qquad \ldots(1)$$

where c_1 and c_2 are arbitrary contents and λ is a parameter. Substitution of (1) in the given system gives

$$\lambda c_1 = 2c_1 + 4c_2, \lambda c_2 = -2c_1 + 6c_2$$

or

$$\left.\begin{aligned}(\lambda - 2)c_1 + (-4)\, c_2 = 0\\ 2\, c_1 + (\lambda - 6)\, c_2 = 0\end{aligned}\right]. \qquad \ldots(2)$$

To have a non-trivial solution of homogeneous system (2), we must have

$$\begin{vmatrix} \lambda - 2 & -4 \\ 2 & \lambda - 6 \end{vmatrix} = 0$$

or $$\lambda^2 - 8\lambda + 20 = 0, \qquad \ldots(3)$$

which is the characteristic equation whose roots are

$$\lambda_1 = 4 + 2i,$$
$$\lambda_2 = 4 - 2i. \qquad \ldots(4)$$

These eigenvalues are conjugate pair but not pure imaginary. So the critical point 0(0, 0) is a SPIRAL POINT. Since the real part of conjugate roots is positive, therefore, the critical point 0(0, 0) is **unstable** by Table given above.

Exercise I. Show that the origin is an unstable node of the linear system

$$\frac{dx}{dt} = 3x + y,$$
$$\frac{dy}{dt} = x + 3y.$$

Exercise II. Show that the origin is an unstable saddle point of the linear system

$$\frac{dx}{dt} = -x + 3y,$$
$$\frac{dy}{dt} = 2x - 2y.$$

Exercise III. Show that the origin is a stable centre but not asymptotically stable center of the linear system

$$\frac{dx}{dt} = 3x + y,$$
$$\frac{dy}{dt} = -13x - 3y.$$

Exercise IV. Show that the origin is an asymptotically stable node of the linear system

$$\frac{dx}{dt} = -5x + y,$$
$$\frac{dy}{dt} = x - 5y.$$

15

CRITICAL POINTS OF ALMOST LINEAR SYSTEMS, DEPENDENCE ON A PARAMETER AND LIAPUNOV'S DIRECT METHOD FOR NONLINEAR SYSTEMS

Let 0(0, 0) be an isolated critical point of the nonlinear real plane autonomous system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y)\,. \qquad \ldots(1)$$

The main result of this section enables us to draw conclusions about the behaviour of solutions of the system (1) when this system is not too different from a linear system. To decide what "not too different" means, consider systems of the special form

$$\frac{dx}{dt} = ax + by + P_1(x, y)\,,$$
$$\frac{dy}{dt} = cx + dy + Q_1(x, y)\,, \qquad \ldots(2)$$

where

(i) a, b, c, d are real constants and

$$ad - bc \neq 0\,, \qquad \ldots(3)$$

(ii) functions $P_1(x, y)$ and $Q_1(x, y)$ have continuous first order partial derivative for all (x, y), and are such that

$$\lim_{(x,y)\to(0,0)} \frac{P_1(x,y)}{\sqrt{x^2+y^2}} = 0,$$
$$\lim_{(x,y)\to(0,0)} \frac{Q_1(x,y)}{\sqrt{x^2+y^2}} = 0\,. \qquad \ldots(4)$$

Definition. The functions $P_1(x, y)$ and $Q_1(x, y)$ are called **perturbations,** and the system is referred to as the perturbed system corresponding to the linear system

$$\frac{dx}{dt} = ax + by,$$
$$\frac{dy}{dt} = cx + dy\,. \qquad \ldots(5)$$

Note. (1) The assumption (4) is

$$P_1 = 0(r),$$
$$Q_1 = 0(r)\,,$$

as $r \to 0+$, where $r = \sqrt{x^2 + y^2}$. …(4a)

This guarantees that the perturbations tend to zero faster than the linear terms in (2). Also, it is easily seen that this condition and condition (3) imply that the origin 0 (0,0) is an isolated critical point for the system (2). Such a critical point is also termed as **"simple critical point".**

Note. (2) One would suspect that the behaviour of the paths of the nonlinear system (2), near 0(0, 0), would be **similar** to that of the paths of the related/corresponding linear system (5), obtained from nonlinear system (2) by neglecting the nonlinear terms.

In other words, it would mean that the nature of the critical point 0(0, 0) of the nonlinear system (2) should be similar to that of the linear system (5), under the conditions mentioned in (3) and (4). In such a situation, the system (2) is called **"almost linear".**

The following theorem (without proof) enables us to draw conclusions about an almost linear system by examining the associated linear system.

Theorem (15.1). Let 0(0, 0) be an isolated critical point of the nonlinear system

$$\frac{dx}{dt} = ax + by + P_1(x, y),$$

$$\frac{dy}{dt} = cx + dy + Q_1(x, y), \qquad \ldots(1)$$

where a, b, c, d are real constants, and

$$ad - bc \neq 0, \qquad \ldots(2)$$

and $P_1(x, y)$ and $Q_1(x, y)$ have continuous first order partial derivatives for all (x, y) and are such that

$$\lim_{(x,y)\to(0,0)} \frac{P_1(x,y)}{\sqrt{x^2+y^2}} = 0,$$

$$\lim_{(x,y)\to(0,0)} \frac{Q_1(x,y)}{\sqrt{x^2+y^2}} = 0. \qquad \ldots(3)$$

Consider the corresponding linear system

$$\frac{dx}{dt} = ax + by,$$

$$\frac{dy}{dt} = cx + dy, \qquad \ldots(4)$$

obtained from (1), by neglecting the nonlinear terms $P_1(x, y)$ and $Q_1(x, y)$. 0(0, 0) is also an isolated critical point of the associated linear system (4). Let λ_1 and λ_2 be the roots of the characteristic equation

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0, \qquad \ldots(5)$$

of the associated linear system (4). Then the following conclusions hold.

(1) The critical point 0 (0, 0) of nonlinear system (1) is of the same type as that of the associated linear system (4) in the following cases.

(i) If λ_1 and λ_2 are real, unequal, and of the same sign, then not only is O(0, 0) a node of linear system (4), but also 0(0, 0) is a node of nonlinear system (1).

(ii) If λ_1 and λ_2 are real and unequal, and of opposite sign, then 0(0, 0) is a saddle point of both systems (1) and (4).

(iii) If λ_1 and λ_2 are real and equal and the system (4) is not such that

$$a = d \neq 0, \; b = c = 0,$$

then 0(0, 0) is a node point of both systems (1) and (4).

(iv) If λ_1 and λ_2 are conjugate complex with real part not zero**,** then 0(0, 0) is a spiral point of both systems (1), and (4).

(2) The critical point 0(0, 0) of the nonlinear system (1) is not necessarily of the same type as that of the linear system (4) in the following cases :

(v) If λ_1 and λ_2 are real and equal and the system (4) is such that

$$a = d \neq 0, \; b = c = 0,$$

then although 0(0, 0) is node of linear system (4), but the point 0(0, 0) may be either a node or a spiral point of nonlinear system (1).

(vi) If λ_1 and λ_2 are pure imaginary, then although 0(0, 0) is a center of linear system (4), but the point 0(0, 0) may be either a center or a spiral point of nonlinear system (1) .

Remark. Although the critical point 0(0, 0) of the non-linear system (1) is of the **same type** as that of the corresponding linear system (4) in cases (i) to (iv) of the above conclusion, the actual **appearance of the paths** is somewhat different

Example 1. Consider the nonlinear plane autonomous system

$$\frac{dx}{dt} = 4x + 2y - 4xy ,$$

$$\frac{dy}{dt} = x + 6y - 8x^2y . \qquad (1)$$

It is obvious that the origin 0(0, 0) is an isolated critical point of it. Here,

$$P_1(x, y) = -4xy$$

and

$$Q_1(x, y) = -8 x^2y \qquad (2)$$

are the perturbation functions. However,

$$\lim_{(x,y)\to(0,0)} \frac{-4xy}{\sqrt{x^2 + y^2}} = 0 ,$$

and

$$\lim_{(x,y)\to(0,0)} \frac{-8x^2y}{\sqrt{x^2 + y^2}} = 0 . \qquad (3)$$

Therefore, the given nonlinear system is almost linear. The associated linear system is

$$\frac{dx}{dt} = 4x - 2y ,$$

$$\frac{dy}{dt} = x + 6y , \qquad (4)$$

and the corresponding characteristic equation (left as an exercise) is

$$\lambda^2 - 10\lambda + 26 = 0 , \qquad (5)$$

with roots

$$\lambda_1, \lambda_2 = 5 \pm i. \qquad (6)$$

From the above theorem (15.1), it follows that the origin 0 (0,0) is a spiral point of the given nonlinear system.

Example 2. Consider the nonlinear autonomous system

$$\frac{dx}{dt} = x - 3y + 4xy ,$$

$$\frac{dy}{dt} = x + 7y - xy^4. \qquad (1)$$

It is routine to verify that the origins 0(0, 0) is the only critical point of it and the required conditions for the above system to be almost linear hold (left as an exercise). The associated linear system is

$$\frac{dx}{dt} = x - 3y,$$

$$\frac{dy}{dt} = x + 7y . \qquad (2)$$

It can be checked that the characteristic equation associated with this linear system is (left as an exercise)

$$\lambda^2 - 8\lambda + 10 = 0\,, \tag{3}$$

and its roots are

$$\lambda_1, \lambda_2 = 4 \pm \sqrt{6}\,. \tag{4}$$

Hence, by the theorem (15.1) stated above, the origin is 0(0,0) a node point of the given nonlinear system.

The following theorem (without proof), based on the work of the Russian engineer and Mathematician, Alexander M. Liapunov, enables us to draw conclusions about stability of the origin for almost linear systems satisfying the hypothesis of the proceeding theorem (15.1).

Theorem (15.2). Under the conditions of the preceding theorem (15.1), the following conclusions hold.

1. If λ_1 and λ_2 are either real and both negative, or complex with negative real parts, then the **origin 0(0, 0) is asymptotically stable for both systems**–linear and almost linear.

2. If λ_1 and λ_2 are both positive, or both complex with positive real parts, then the origin **0(0, 0) is an unstable critical point of both systems**–linear and almost linear.

Example 1. Consider the nonlinear plane autonomous system

$$\frac{dx}{dt} = -x + y + x^3y\,,$$
$$\frac{dy}{dt} = -2x - 3y - x^2y^2\,. \tag{1}$$

It can be verified that this system is almost–linear (left as an exercise). The associated linear system is

$$\frac{dx}{dt} = -x + y,$$
$$\frac{dy}{dt} = -2x - 3y\,. \tag{2}$$

It can be checked that the characteristic equations associated with this linear system is (left as an exercise)

$$\lambda^2 + 4\lambda + 5 = 0, \tag{3}$$

with roots

$$\lambda_1, \lambda_2 = -2 \pm i. \tag{4}$$

Since the roots are complex with negative real parts, we conclude that the critical point 0(0, 0) is an asymptotically stable spiral point of both linear and almost linear systems.

Example 2. Consider the nonlinear system

$$\frac{dx}{dt} = x + 4y - x^2\,,$$
$$\frac{dy}{dt} = 6x - y + 2xy. \tag{1}$$

It is obvious that the origin 0(0, 0) is a critical point of the given system. Further it can be checked that this system is almost linear (left as an exercise). The associated linear system is

$$\frac{dx}{dt} = x + 4y,$$
$$\frac{dy}{dt} = 6x - y. \tag{2}$$

It can be verified that the associated characteristic equation is (left as an exercise)

$$\lambda^2 - 25 = 0\,,$$

with roots

$$\lambda_1, \lambda_2 = \pm 5. \tag{3}$$

Since these eigenvalues are real, unequal and of opposite sign, so the critical point 0(0, 0) of the given nonlinear system is a saddle point. Further, it is also concluded that this critical point is unstable.

Example 3. Find all the real critical points of the nonlinear system

$$\frac{dx}{dt} = 8x - y^2,$$
$$\frac{dy}{dt} = -6y + 6x^2, \tag{1}$$

and determine the type and stability of each of these critical points.

Solution. The critical points of the given system are given by the following system of algebraic equations

$$8x - y^2 = 0,$$
$$-6y + 6x^2 = 0 .$$

Solving these equations, one finds (left as an exercise) that there are two real critical points, namely, 0(0, 0) and $P_0(2, 4)$.

Critical point 0(0, 0). It may be checked that the given system is almost linear (left as an exercise). The corresponding linear system is

$$\frac{dx}{dt} = 8x,$$
$$\frac{dy}{dt} = -6y. \tag{3}$$

It can be verified that (exercise) the characteristic equation of this linear system is

$$\lambda^2 - 2\lambda - 48 = 0, \tag{4}$$

with roots $\lambda_1, \lambda_2 = 8, -6$. Since the roots are real, unequal, and of opposite sign, therefore, it is concluded that the critical point 0(0, 0) of the given nonlinear system is a saddle point. Further, it is also concluded that this saddle point is unstable.

Critical point P_0(2, 4). We make the transformation

$$\alpha = x - 2,$$
$$\beta = y - 4 . \tag{5}$$

which transforms the critical point (x = 2, y = 4) into the origin ($\alpha = 0, \beta = 0$) in the $\alpha\beta$-plane. The given nonlinear system now becomes

$$\frac{d\alpha}{dt} = 8\alpha - 8\beta - \beta^2 ,$$
$$\frac{d\beta}{dt} = 24\alpha - 6\beta + 6\alpha^2, \tag{6}$$

which is almost linear (left as an exercise). Its associated linear system is

$$\frac{d\alpha}{dt} = 8\alpha - 8\beta,$$
$$\frac{d\beta}{dt} = 24\alpha - 6\beta . \tag{7}$$

The corresponding characteristic equation is (exercise)

$$\lambda^2 - 2\lambda + 144 = 0, \tag{8}$$

with roots $\lambda_1, \lambda_2 = 1 \pm \sqrt{143}\, i$, which are conjugate complex with real part not zero. Thus, the critical point ($\alpha = 0$, $\beta = 0$) of the nonlinear system is a spiral point (why ?) and is unstable (why ?). Consequently, the critical point P_0 (2, 4) of the given nonlinear system is an unstable spiral point. Hence, the result.

Dependence on a parameter

Consider a differential equation of the form

$$\frac{d^2x}{dt^2} = f(x, \lambda), \qquad \ldots(1)$$

where f is analytic for all values of x and λ. The equation (1) is equivalent to the following non-linear plane autonomous system

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = f(x, \lambda). \qquad \ldots(2)$$

For each fixed value of the parameter λ, the critical points of the system (2) are the points of the type (x_c , 0), where the abscissas x_c are the roots of the equation

$$f(x, \lambda) = 0, \qquad \ldots(3)$$

considered as an equation in the unknown x.

In general, as λ varies continuously through a given range of values, the corresponding x_c vary and hence so do the corresponding **critical points, paths, and solutions of the system (2).**

A value of the parameter λ at which two or more critical points coalesce into less than their previous number is called a **bifurcation value / critical value of the parameter λ**. At such a value, the nature of the corresponding paths changes abruptly.

Note. In determining both the critical values of the parameter λ and the critical points of the system (2), it is often very useful to investigate the graph of the relation

$$f(x, \lambda) = 0, \qquad \ldots(4)$$

in the $x\lambda$-plane.

Theorem (15.3) (without proof)

Let $$m\frac{d^2x}{dt^2} = F(x), \qquad \ldots(1)$$

be the differential equations of a conservative dynamical system, F being analytic for all values of x. Consider the equivalent autonomous system

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = \frac{F(x)}{m}. \qquad \ldots(2)$$

Let (x_c, 0) be a critical point of this system. Let V(x) be the potential energy function of the dynamical system (1) and defined by

$$V(x) = -\int_0^x F(x)\, dx. \qquad \ldots(3)$$

Then the following conclusions hold.

(1) If V has a relative minimum at $x = x_c$, then the critical point (x_c, 0) is a center and is stable.

(2) If V has a relative maximum at $x = x_c$, then the critical point (x_c, 0) is a saddle point and is unstable.

(3) If V has a horizontal inflection point at $x = x_c$, then the critical point (x_c, 0) is a "degenerate" type called a cusp and is unstable.

Example. Examine the critical points of the nonlinear plane autonomous system

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = x^2 - 4x + \lambda, \qquad (1)$$

where λ is a parameter. Also find the critical values of the parameter λ.

Solution. The critical points of this system are the points $P_1(x_1, 0)$ and $P_2(x_2, 0)$ where x_1 and x_2 are the roots of the quadratic equation

$$x^2 - 4x + \lambda = 0, \qquad (2)$$

in the unknown x. We find

$$x_1, = 2\sqrt{4-\lambda},$$
$$x_2 = 2-\sqrt{4-\lambda_0}. \qquad (3)$$

Thus, the critical points of the given system are the points

$$P_1\,(2+\sqrt{4-\lambda}, 0) \text{ and } P_2\,(2-\sqrt{4-\lambda}, 0). \qquad (4)$$

For $\lambda < 4$, the roots x_1, x_2 are real and distinct.

For $\lambda = 4$, the roots x_1, x_2 and real are equal, each equal to 2.

For $\lambda > 4$, the roots are complex.

Thus, for $\lambda < 4$, the critical points P_1 and P_2 are real and distinct. As $\lambda \to 4-$, these two critical points approach each other ; and at $\lambda = 4$, they coalesce into the one single critical point (2, 0). For $\lambda > 4$, there are no real critical points. Therefore, the value $\lambda = 4$ is the **critical value** of the parameter λ.

Now we will consider the three cases separately.

Case I. When $\lambda < 4$. For each fixed value, say λ_0, of λ such that $\lambda_0 < 4$, the critical points P_1 and P_2 are the real distinct points with coordinates

$$P_1\,(2+\sqrt{4-\lambda_0^2}, 0),\ P_2(2-\sqrt{4-\lambda_0}, 0)$$

The corresponding potential energy function $V(x, \lambda_0)$ is given by

$$V(x, \lambda_0) = -\int_0^x (x^2 - 4x + \lambda_0)\,dx = -\frac{x^3}{3} + 2x^2 - \lambda_0 x. \qquad \ldots(3)$$

Then $$V'(x, \lambda_0) = -x^2 + 4x - \lambda_0,$$
$$V''(x, \lambda_0) = -2x + 4. \qquad \ldots(4)$$

Hence

$$V''(2+\sqrt{4-\lambda_0}, \lambda_0) = 0,\ V''(2-\sqrt{4-\lambda_0}, \lambda_0) > 0 \qquad \ldots(5)$$

This shows that the potential energy $V(x, \lambda_0)$ has a relative minimum at $x = 2 - \sqrt{4-\lambda_0}$. Consequently, the critical point $(2-\sqrt{4-\lambda_0}, 0)$ is a centre and is stable.

Similarly, one finds that (left as an exercise) the critical point $(2+\sqrt{4-\lambda_0}, 0)$ is a saddle point and is unstable.

Case. 2. When $\lambda = 4$. There is only one critical point $P_1(2, 0)$ of the given system. From equation (4), we find

$$V'(2, 4) = 0,$$
$$V''(2, 4) = 0. \qquad \ldots(6)$$

However,

$$V'''(x, \lambda_0) = -2, \text{ for all } x \text{ and } \lambda_0. \qquad \ldots(7)$$

Therefore,

$$V'''(2, 4) \neq 0. \qquad \ldots(8)$$

Thus, the potential energy function V has a horizontal inflection point at x = 2. So, the critical point $P_1(2, 0)$ is a cusp and is unstable.

Case. 3. When $\lambda > 4$. For each fixed value λ_0 of λ such that

$$\lambda_0 > 4,$$

there are no real critical points of the given system. This completes the solution.

Liapunov's direct method to study stability

Earlier, we have studied the stability of linear and almost linear systems. There are occasions when almost linear technique fails to give useful information, or when it is inconvenient or difficult to solve. An alternative approach is to consider the fully nonlinear system, using Liapunov's direct method. This method is useful for studying the stability of more general plane autonomous systems.

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y). \qquad \ldots(1)$$

It is assumed that this system has an isolated critical point at the origin 0(0, 0) and that functions F(x, y) and G(x, y) have continuous first partial derivatives for all (x, y).

Definition. Let E = E(x, y) have continuous first order partial derivatives at all points (x, y) in a domain D containing the origin 0(0, 0). The derivative of E w.r.t. the system (1) is denoted by $\dot{E}$ and defined by

$$\dot{E}(x, y) = \frac{\partial E(x, y)}{\partial x} F(x, y) + \frac{\partial E(x, y)}{\partial y} G(x, y) \qquad \ldots(2)$$

Remark. Let C be a path of the non-linear system (1). Let

$$x = f(t),\ y = g(t), \qquad \ldots(3)$$

be an arbitrary solution of system (1) defining C parametrically. Let E = E(x, y) have continuous first partial derivatives for all (x, y) in a domain, say D, containing the curve C.

Now, E is a composite function of t along C. Using the chain rule, the derivative of E w.r.t "t" along the curve C is given by

$$\frac{dE}{dt}(f(t), g(t)) = E_x(f(t), g(t))\ \frac{df}{dt} + E_y(f(t), g(t))\frac{dg}{dt}$$
$$= E_x(f(t), g(t))\ F(f(t), g(t)) + E_y(f(t), g(t))G(f(t), g(t))$$
$$= \dot{E}(f(t), g(t)) \qquad \ldots(4)$$

Thus, we see that the derivative of E (f(t), g(t)) w.r.t. "t" along the path C is equal to the derivative of E w.r.t. the system (1) evaluated at x = f(t), y = g(t).

Definition. Let the plane autonomous nonlinear system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y), \qquad \ldots(1)$$

have an isolated critical point at the origin 0(0, 0). Let F and G have continuous first order partial derivatives for all (x, y).

Let E = E(x, y) be a function which is **positive definite** for all (x, y) in a domain D containing the origin 0(0, 0) and such that the derivative $\dot{E}(x, y)$ of E w.r.t the system (1) is **negative semi-definite** for all $(x, y) \in D$.

Then the function E is called a **Liapunov function for the plane autonomous system (1) in D.**

We now state two theorems (without proof) on the stability of the critical point 0(0, 0) of the nonlinear plane autonomous systems with the help of Liapunov functions.

Theorem (15.4). Let the nonlinear system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$

have an isolated critical point at the origin 0(0, 0) and that F(x, y) and G(x, y) have continuous first order partial derivatives for all (x, y). If there exists a Liapunov function E = E(x, y) for this system in some domain D containing the origin 0(0, 0), then the critical point 0(0, 0) of this system is stable.

Theorem (15.5). Let the nonlinear system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$

have an isolated critical point at the origin 0(0, 0) and that F(x, y) and G(x, y) have continuous first order partial derivatives for all (x, y). If there exists a Liapunov function E = E(x, y) for this system in some domain D containing the origin 0(0, 0) such that the derivative of E w.r.t. the given system, i.e.,

$$\dot{E}(x, y) = \frac{\partial E}{\partial x}F(x, y) + \frac{\partial E}{\partial y}G(x, y)$$

is negative definite in D, then the critical point 0(0, 0) is asymptotically stable.

Example. Construct a Liapunov function of the form $Ax^2 + By^2$ (where A and B are constants) for the plane autonomous nonlinear system

$$\frac{dx}{dt} = -x + y^2,$$
$$\frac{dy}{dt} = -y + x^2,$$

and use it to determine whether the critical point 0(0, 0) of this system is asymptotically stable or at least stable.

Solution. Consider the function

$$E(x, y) = x^2 + y^2 \qquad \ldots(1)$$

Then E(0, 0) = 0 and

$$E(x, y) > 0 \text{ for all } (x, y) \neq (0, 0). \qquad \ldots(2)$$

Therefore, the function E(x, y) is positive definite in every domain D containing the origin 0(0, 0). We find

$$\frac{\partial E}{\partial x} = 2x, \frac{\partial E}{\partial y} = 2y. \qquad \ldots(3)$$

The derivative of E(x, y) w.r.t. the given system, denoted by $\dot{E}(x, y)$, is given by

$$\dot{E}(x, y) = \frac{\partial E}{\partial x}F(x, y) + \frac{\partial E}{\partial y}G(x, y)$$
$$= 2x(-x + y^2) + 2y(-y + x^2)$$
$$= -2(x^2 + y^2) + 2(x^2y + xy^2), \qquad \ldots(4)$$

since, as in the given system

$$F(x, y) = -x + y^2,$$
$$G(x, y) = -y + x^2. \qquad \ldots(5)$$

Clearly $\dot{E}(0, 0) = 0$. Now, we observe the following :

If $x < 1$ and $y \neq 0$, then $xy^2 < y^2$; if $y < 1$ and $x \neq 0$, then $x^2 y < x^2$. Thus, if $x < 1$, $y < 1$, and $(x, y) \neq (0, 0)$, then

$$x^2y + xy^2 < x^2 + y^2, \qquad \ldots(6)$$

and hence

$$-(x^2 + y^2) + (x^2y + xy^2) < 0. \qquad \ldots(7)$$

Thus, in every domain D containing (0, 0) and such that $x < 1$ and $y < 1$, $\dot{E}(x, y)$ is negative definite and hence negative semi-definite. Therefore, E defined by (1) is a Liapunov function for the given system of nonlinear differential equations. Further, applying theorems (15.4) and (15.5), we see that the critical point 0(0, 0) is asymptotically stable.

Remark. (1) Liapunov's direct method is indeed "direct" in the sense that it does nto require any previous knowledge about the solution of nonlinear system or the type/nature of its critical point 0(0, 0). Instead, if one can construct a Liapunov function for the given nonlinear system, then one can "directly" obtain information about the stability of the critical point 0(0, 0). However, there is no general method for constructing a Liapunov function, although methods for doing so are available for certain class of equations.

(2) In seeking Liapunov functions, it is sometimes useful to identify E with the energy of the physical system which the differential equations describe.

16

PERIODIC SOLUTIONS, BENDIXSON THEOREM, INDEX OF A CRITICAL POINT

We are interested in obtaining criteria for the existence, or otherwise, of periodic solutions of a planar autonomous system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y) \quad . \qquad \ldots(1)$$

Clearly, the orbits or paths or trajectories of periodic solutions are closed curves in the phase plane.

Theorem (16.1). Show that periodic solutions and closed paths of the planar autonomous system (1) are very closely related.

Proof. First of all, if

$$x = f_1(t),\ y = g_1(t) \qquad \ldots(1)$$

where f_1 and g_1 are not both constant functions, is a periodic solutions of system,

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y), \qquad \ldots(2)$$

then the path which this solution defines is a closed path.
Secondly, let C be a closed path of the system (2) defined by a solution

$$x = f(t),\ y = g(t), \qquad \ldots(3)$$

and suppose that

$$f(t_0) = x_0,\ g(t_0) = y_0\ . \qquad \ldots(4)$$

Since C is a closed curve, there exists a value,

$$t_1 = t_0 + T,\ \text{(say)},\ T > 0, \qquad \ldots(5)$$

such that

$$f(t_1) = x_0,$$
$$g(t_1) = y_0\ . \qquad \ldots(6)$$

Now, the pair

$$x = f(t + T),$$
$$y = g(t + T), \qquad \ldots(7)$$

is also a solution of the same system (2). Further, at $t = t_0$, this later solution assumes the value

$$f(t_0 + T) = f(t_1) = x_0,$$
$$g(t_0 + T) = g(t_1) = y_0\ . \qquad \ldots(8)$$

Therefore, by uniqueness theorem, the two solutions, given by (3) and (7), are identically for all t. That is,

$$f(t+T) = f(t),$$
$$g(t+T) = g(t), \text{ for all t.} \quad \ldots(9)$$

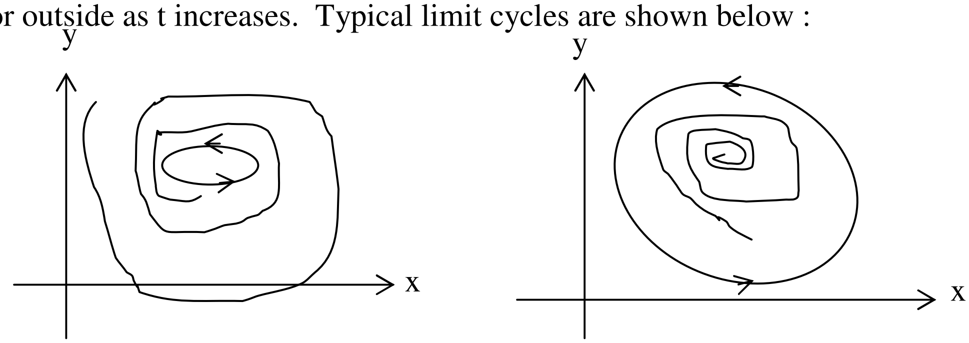
This shows that the solution (3), defining the closed curve C, is a periodic solution. This completes the proof.

Definition. (Limit Cycle)

For a system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$

a limit cycle is a closed trajectory that has non-closed trajectories spiraling towards it from either inside or outside as t increases. Typical limit cycles are shown below :



For a given planar autonomous system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$

we need a theorem giving sufficient conditions for the existence of a limit cycle of this system. The Poincare-Bendixson theorem is one of the few general theorems of this nature. Before stating this theorem, however, we shall state and prove a theorem on the non-existence of closed paths/trajectories of the given system, due to Bendixson.

Theorem 16.2 (Known as Bendixson non-existence theorem). Let D be a domain in the xy-plane. Let the plane autonomous system be

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$

where F and G have continuous first order derivatives in D. Suppose that $\frac{\partial F(x,y)}{\partial x} + \frac{\partial G(x,y)}{\partial y}$ has the same sign throughout D. Then the given system has no closed path in the domain D.

Proof. Suppose that C is the orbit of a periodic solution of the given plane autonomous system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y), \quad \ldots(1)$$

Let S be the region enclosed by C. Then, by Green's theorem in the plane, we have

$$\iint_S \left\{\frac{\partial F}{\partial x}+\frac{\partial F}{\partial y}\right\} dxdy = \int_C \{F(x,y)\, dy - G(x,y)\, dx\}, \qquad \ldots(2)$$

where the line integral is taken in the positive sense. But, on the orbit C, the system of given differential equations holds, i.e.,

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y), \qquad \ldots(3)$$

holds on C.

Let x = f(t), y = g(t) be the parametric equaths of the path C ; and let T denote the period of this solution. Then

$$\frac{df(t)}{dt} = F(f(t), g(t)),$$
$$\frac{dg(t)}{dt} = G(f(t), g(t)), \qquad \ldots(4)$$

along C. Hence

$$\int_C [F(x, y)dy - G(x, y)\, dx]$$
$$= \int_0^T \left[F[f(t), g(t)]\frac{dg(t)}{dt} - G[f(t), g(t)]\frac{df(t)}{dt}\right] dt$$
$$= \int_0^T \{F[f(t), g(t)]\, G[f(t), g(t)] - G[f(t), g(t)].\ F[f(t), g(t)]\}\, dt$$
$$= 0 .$$

Thus,

$$\iint_S \left\{\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right\} dxdy = 0. \qquad \ldots(5)$$

But the integrand in (5) has the same sign throughout the domain D. Hence, this double integral must be non-zero. This contradiction proves that D can contain no closed path of the given system.

Note : Theorem (16.2) may be restated in the following form :

"There are no periodic solutions in any domain where $\left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)$ is of one sign".

Example. Let F(x, y) = 3x + 4y + x^3 ,

$$G(x, y) = 5x - 2y + y^3 , \qquad (1)$$

and consider the system

$$\frac{dx}{dt} = F(x, y), \quad \frac{dy}{dt} = G(x, y). \qquad (2)$$

The functions F(x, y) and G(x, y) and their derivatives are continuous throughout the phase plane. Further,

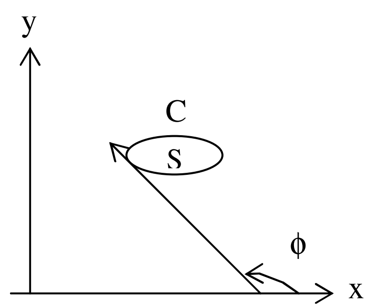
$$\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y} = 3 + 3x^2 - 2 + 3y^2$$
$$= 1 + 3\,(x^2 + y^2)$$
$$> 0\,, \qquad (3)$$

for all (x, y). We conclude that the given autonomous system has no closed trajectory and hence no periodic solution.

Theorem (16.3). The orbit C of a periodic solution of a plane autonomous system must enclose atleast one critical point.

Proof. The proof will be by contradiction. Suppose that the orbit C contains a region, say S, with no critical points of the plane autonomous systems

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y)\,. \qquad \ldots(1)$$



Then $\qquad F^2 + G^2 \neq 0$ in S. $\qquad \ldots(2)$

Let ϕ be the angle between the tangent vector to C and the x-axis, as shown in the figure above. Then clearly

$$\int_C d\phi = 2\pi\,. \qquad \ldots(3)$$

But $\qquad \tan\phi = \dfrac{dy}{dx}$

$$= \frac{dy/dt}{dx/dt}$$
$$= \frac{F(x, y)}{G(x, y)}\,. \qquad \ldots(4)$$

This gives (left as an exercise)

$$d\phi = \frac{FdG - GdF}{F^2 + G^2}\,. \qquad \ldots(5)$$

The Green's theorem in the plane gives

$$\int_C d\,\phi = \int_C \frac{FdG - GdF}{F^2 + G^2}$$
$$= \iint_S \left\{\frac{\partial}{\partial F}\left(\frac{F}{F^2 + G^2}\right) + \frac{\partial}{\partial G}\left(\frac{G}{F^2 + G^2}\right)\right\} dFdG\,. \qquad \ldots(6)$$

But the integrand in the integral over S is identically zero (left as an exercise). This leads to a contradiction. Hence the result holds. This completes the proof.

Note. (1) The second theorem and shows that the location of critical points may be an indicator of the presence of periodic solutions.

(2) The two theorems have (16.2) and (16.3) provided negative criteria. They help us to determine if an autonomous system does not have a closed trajectory in a region of the plane.

The next theorem (without proof), known as **Poincare-Bendixson theorem**, gives sufficient conditions for a closed trajectory to exist. It also relates the existence of periodic solutions to the location of critical points.

Theorem 16.4. (Poincare-Bendixson Theorem). Let the given plane autonomous system be

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$

where F and G have continuous first partial derivative in a domain D of the xy-plane. Let $R \subseteq D$ be a closed bounded region containing no critical points of this system, and such that there is an orbit C which lies in R for all $t \geq 0$. Then either C is a closed orbit, or C approaches a closed orbit spirally as $t \to \infty$.

Remark. In either case, there exists a closed path of the given system in the region R.

The Index of a critical point

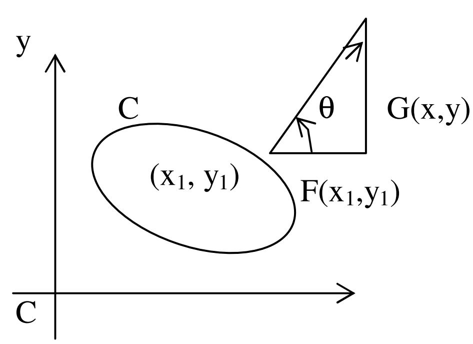
Simple closed curve. A closed curve having no double points is called a simple closed curve
Example (i) A circle is a simple closed curve.
Example (ii) The curve ∞ is not a simple closed curve.
Consider the system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y), \qquad \ldots(1)$$

where F and G have continuous first order partial derivatives for all (x, y). Assume that all of the critical points of this system are isolated. Now consider a simple closed curve, say C, which passes through no critical points of this system. Consider a point (x_1, y_1) on the curve C and the vector

$$\vec{a} = F(x_1, y_1)\,\hat{i} + G(x_1, y_1)\,\hat{j}\,, \qquad \ldots(2)$$

at the point (x_1, y_1). Let this vector make an angle θ with the positive x-direction.



Now, let (x_1, y_1) describe the curve C once in the positive (counterclockwise) direction and return to its original position. As the point (x_1, y_1) describes the curve C, the vector (2) changes continuously, and consequently the angle θ also varies continuously.

When the point (x_1, y_1) reaches its original position, the angle θ will have changed by an amount, Δθ say.

Definition. Let θ denote the angle from the positive x-direction to the vector $F(x_1, y_1)\hat{i} + G(x_1, y_1)\hat{j}$ defined by the system

$$\frac{dx}{dt} = F(x, y),$$
$$\frac{dy}{dt} = G(x, y),$$

at the point (x_1, y_1). Let Δθ denote the total change in θ as (x_1, y_1) describes the simple closed curve C once in the anticlockwise direction. The number

$$I = \frac{\Delta\theta}{2\pi}$$

is called the **index of the curve** C w.r.t the given autonomous system.

Remarks. (1) Clearly Δθ is either equal to zero or a positive or negative integral multiple of 2π. Hence, the index I. is either 0 or a positive or negative integer.

(2) If the vector $F(x_1, y_1)\ \hat{i} + G(x_1, y_1)\ \hat{j}$ merely oscillates but does not make a complete rotation as (x_1, y_1) describes C, then I = 0.

(3) If the net change Δθ in θ is a decrease, then I < 0.

Definition. By the **index of an isolated critical point** (x_0, y_0) of the given autonomous system, we mean the index of a simple enclose curve C which encloses (x_0, y_0) but no other critical point of the given system.

Remark. (1) The index of a node/centre/spiral point is I = 1 .

(2) The index of a saddle point is I = −1.

(3) Let (x_0, y_0) be an isolated critical point of a given system. Then all simple closed curves enclosing (x_0, y_0) but containing no other critical point of the given system, have the same index.

For more information/detail, the reader is advised to refer to the following books.

Books Suggested

(1)	S. L Ross	Differential Equations
(2)	R. Grim Shaw	Nonlinear Ordinary Differential Equations, CRC Press, 1993
(3)	P. Hertman	Differential Equations.

17

PRELIMINARIES ABOUT LINEAR SECOND ORDER DIFFERENTIAL EQUATIONS

One of the most frequently occurring types of differential equation in mathematics and physical sciences is the linear second order differential equation of the form

$$\frac{d^2u}{dt^2} + g(t)\frac{du}{dt} + f(t)\,u = h(t)\,, \qquad \ldots(1)$$

or of the form

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = h(t). \qquad \ldots(2)$$

Unless otherwise specified, it is assumed that the functions $f(t)$, g(t), h(t) and $p(t) \neq 0$, q(t) in these equations are continuous real/complex valued functions of some real variable t on some interval I, which can be bounded or unbounded.

Note 1. The form (2) is more general since (1) can be written as

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + p(t)\,f(t)\,u = p(t)\,h(t), \qquad \ldots(3)$$

provided p(t) is defined as

$$p(t) = \exp\left[\int_a^t g(s)\,ds\right], \qquad \ldots(4)$$

for some $a \in I$.

Note 2. If p(t) is continuously differentiable, then (2) can be written as

$$\frac{d^2u}{dt^2} + \left\{\frac{p'(t)}{p(t)}\right\}\frac{du}{dt} + \left\{\frac{q(t)}{p(t)}\right\}u = \left\{\frac{h(t)}{p(t)}\right\}, \qquad \ldots(5)$$

which is of form (1).

Remark. When the function p(t), in equation (2), is continuous but does not have a continuous derivative, then equation (2) can not be written in the form (1). In such a situation, equation (2) is to be interpreted as the first order linear system for the binary vector

$$y = \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}, \qquad \ldots(6)$$

where

$$y^1 = u,\ y^2 = p(t)\frac{du}{dt}, \qquad \ldots(7)$$

and equations are

$$\frac{dy^1}{dt} = \frac{1}{p(t)} y^2,$$

$$\frac{dy^2}{dt} = -q(t)y^1 + h(t). \qquad \ldots(8)$$

Equations in (8) can be put in the form

$$\frac{dy}{dt} = A(t)\, y + b(t)\,, \qquad \ldots(9)$$

where $$A(t) = \begin{bmatrix} 0 & 1/p(t) \\ -q(t) & 0 \end{bmatrix}, \quad b(t) = \begin{bmatrix} 0 \\ h(t) \end{bmatrix}. \qquad \ldots(10)$$

In other words, a solution u = u(t) of equation (2) is a continuously differentiable function such that $p(t)\frac{du}{dt}$ has a continuous derivative satisfying (2).

When $p(t) \neq 0$, q(t), h(t) are continuous, the standard existence and uniqueness theorems for linear systems are applicable to system(9), hence equation (2).

Particular Case. When $p(t) \equiv 1$ on I. Then equation (2) becomes

$$\frac{d^2u}{dt^2} + q(t)\, u = h(t) \qquad \ldots(11)$$

Reduction of equation (2) to an equation of the type (11)

When $p(t) \neq 0$ is real valued, then we change the independent variables through the relation

$$s = \int_a^t \frac{d\xi}{p(\xi)} + \text{const.}$$

i.e., $$ds = dt/p(t), \qquad \ldots(12)$$

for some $a \in I$.

The function s = s(t) has a derivative

$$\frac{ds}{dt} = \frac{1}{p(t)} \neq 0\,,$$

and so s = s(t) is strictly monotone. Hence s = s(t) has an inverse function

$$t = t(s)\,,$$

define on some s-interval.

In terms of the new independent variable s, the differential equation (2) becomes, as derived below :

$$\frac{du}{ds} = \frac{du}{dt} . \frac{dt}{ds}$$

$$= p(t)\frac{du}{dt}\,,$$

$$\frac{d^2u}{ds^2} = \frac{d}{dt}\left[p(t)\frac{du}{dt}\right] . \frac{dt}{ds}$$

$$= p(t)\frac{d}{dt}\left[p(t)\frac{du}{dt}\right]. \qquad \ldots(13)$$

So, equation (2) becomes, after multiplying by p(t) throughout,

$$\frac{d^2u}{ds^2} + p(t)q(t)\,u = p(t)h(t)\,, \qquad \ldots(14)$$

where t in p(t) q(t) and p(t) h(t) is replaced by the function t = t(s). Thus, equation (14) is of the type (11), under transformation given by (12).

Reduction of equation (1) to an equation of the type (11)

We change the dependent variable u into z by the relation

$$u(t) = z(t)\exp\left[-\frac{1}{2}\int_a^t g(\xi)d\xi\right], \qquad \ldots(15)$$

for some $a \in I$. Then

$$\frac{du}{dt} = z.\exp\left[-\frac{1}{2}\int_a^t g(\xi)d\xi\right].\left\{-\frac{1}{2}g(t)\right\} + \frac{dz}{dt}\exp\left[-\frac{1}{2}\int_a^t g(\xi)d\xi\right]$$

$$= \left\{-\frac{1}{2}z(t).g(t) + \frac{dz}{dt}\right\}\exp\left[-\frac{1}{2}\int_a^t g(\xi)d\xi\right], \qquad \ldots(16)$$

and

$$\frac{d^2u}{dt^2} = \left\{-\frac{1}{2}z(t)g(t) + \frac{dz}{dt}\right\}\exp\left\{-\frac{1}{2}\int_a^t g(\xi)d\xi\right\}\left\{-\frac{1}{2}g(t)\right\}$$

$$+\left\{-\frac{1}{2}\frac{dz}{dt}g(t) - \frac{1}{2}z(t)g'(t) + \frac{d^2z}{dt^2}\right\}\exp\left\{-\frac{1}{2}\int_a^t g(\xi)d\xi\right\}$$

$$= \left[\frac{1}{4}z(t)g^2(t) - g(t)\frac{dz}{dt} - \frac{1}{2}z(t)g^1(t) + \frac{d^2z}{dt^2}\right].\exp\left\{-\frac{1}{2}\int_a^t g(\xi)d\xi\right\}. \qquad \ldots(17)$$

On substituting, equation (1) is transformed to

$$\frac{d^2z}{dt^2} + \left[f(t) + \frac{g^2(t)}{4} - \frac{g^1(t)}{2}\right]z = h(t).\exp\left[\frac{1}{2}\int_a^t g(\xi)d\xi\right], \qquad \ldots(18)$$

which is of type (11). This completes the reduction process.

Note. Generally, the second order differential equations to be considered will be assumed to be of the form (2) or (11)

Example. 1. The simplest differential equations of the type (11) are

$$u'' = 0,$$
$$u'' + \sigma^2 u = 0,\ u'' - \sigma^2 u = 0,\ \sigma \neq 0. \qquad \ldots(1)$$

The general solutions of these equations are, respectively,

$$u = c_1 + c_2 t,\ u(t) = c_1\cos\sigma t + c_2\sin\sigma t, u(t) = c_1e^{\sigma t} + c_2e^{-\sigma t}\ , \qquad \ldots(2)$$

where c_1 and c_2 are arbitrary constants.

Example. 2. Consider the differential equation

$$u'' + bu' + au = 0, \qquad \ldots(1)$$

a and b being constants.

Then $u(t) = e^{\lambda t}$ is a solution of equation (1) iff λ satisfies

$$\lambda^2 + b\lambda + a = 0. \qquad \ldots(2)$$

Hence, the general solution of equation (1) is

$$u(t) = e^{-\frac{bt}{2}}(c_1 + c_2 t)$$

or $$u(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \qquad \ldots(3)$$

according as quadratic equation (2) has a double root $\lambda = -\frac{1}{2}b$ or distinct roots

$$\lambda_1, \lambda_2 = -\frac{b}{2} \pm \sqrt{\frac{1}{4}b^2 - a}\ .$$

When a and b are real and $\frac{1}{4}b^2 - a < 0$, then non real exponents in the solution can be avoided by writing

$$u(t) = e^{-bt/2}\left\{c_1 \cos\left(\sqrt{a - \frac{1}{4}b^2}\right).t + c_2 \sin\left(\sqrt{a - \frac{1}{4}b^2}\right)t\right\} \qquad \ldots(4)$$

Remark. The substitution

$$u(t) = z(t)\, e^{-bt/2}$$

reduces differential equation (1) to

$$\frac{d^2z}{dt^2} + \sigma^2 z = 0,$$

where

$$\sigma^2 = a - \frac{1}{4}b^2.$$

Definition. (Non-oscillatory and oscillatory functions)

A real valued function *f*(t) defined and continuous in an interval [a, b] is said to be non-oscillatory in [a, b] if *f*(t) has not more than one zero in [a, b].

If *f*(t) has atleast two zeros in [a, b], then *f*(t) is said to be oscillatory in [a, b].

Examples. (i) consider the function

$$f(t) = Ae^{-t} + Be^{-t}$$

for $t \geq 0$, A and B are constants. Then *f*(t) is non-oscillatory.

(ii) Let $f(t) = \sin t,\ t \geq 0$.

Then f(t) is oscillatory in $[0, 4\pi]$.

Definition. (Non-oscillatory and oscillatory differential equations)

A second order differential equation

$$\frac{d^2u}{dt^2} + p(t)\frac{du}{dt} + q(t)\ y = h(t),\ t \geq 0$$

is called “non-oscillatory” if every solution u = u(t) of it, is non-oscillatory. Otherwise, differential equations is called oscillatory.

Example. (1) $u'' + u = 0$ is oscillatory.
Its general solution is

$$u(t) = A \cos t + B \sin t , t \geq 0 .$$

W.l.o.g., we can assume that both A and B are non-zero constants, otherwise, u(t) is trivially oscillatory.
In that case, u(t) has a zero at

$$t = n \pi + \tan^{-1} (A/B), \text{ for } n = 0, 1, 2, 3, \ldots.$$

So, this equation is oscillatory.
Example. (2) consider the linear equation

$$u'' - u = 0, \text{ for } t = 0.$$

Its general solution is $u(t) = Ae^{t} + Be^{-t}$, A & B are constants. This solution is non-oscillatory. Hence, this equation is non-oscillatory.

Definition. Let $f(t)$ and g(t) be two real valued defined and continuous functions in some interval [a, b]. Then ***f*(t) is said to oscillate more rapidly than g(t)** if the number of zeros of $f(t)$ in [a, b] exceed the number of zeros of g(t) in [a, b] by more than one.

Example. Let $f(t) = \sin 2t$ in $[0, 4\pi]$
$g(t) = \sin t$ in $[0, 4\pi]$
Then zeros of $f(t)$ are only half as far apart as the zeros of g(t). So $f(t)$ oscillates more rapidly than g(t) in the interval $[0, 4\pi]$.

Remark. Qualitative properties of solutions of differential equations assume importance in the absence of closed form solutions.

In case the solutions are not expressible in terms of the usual "known functions", an analysis of the differential equation is necessary to find the facets of the solutions.

One such qualitative property, which has wide applications, is the oscillation of solutions.

Prüfer transformation/ Polar coordinate transformation

This transformation is applicable to linear homogeneous second order differential equations. This transformation yields an equivalent system of two first order differential equations. This transformation changes an equation from Liouville normal form to two successive ordinary differential equations. It is often used to obtain information about the zeros of solutions.

Procedure. Suppose we have the Sturm-Liouville equation

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\, u = 0, \qquad \ldots(1)$$

defined on the interval I with $p(t) > 0$, $p(t) \in C^1$, and q(t) continuous. If we think of this single second order equation (1) as two first order differential equations for the unknowns $\{u, u'\}$, then we can change the dependent variables from $\{u, u'\}$ to $\{\rho(t), \phi(t)\}$ by

$$p(t)\, u'(t) = \rho(t) \cos \phi(t) ,$$
$$u(t) = \rho(t) \sin \phi(t) . \qquad \ldots(2)$$

This substitution (2) is called **Prufer substitution**. Equation (2) gives

$$\rho(t) = [u^2 + (p\,u')^2]^{1/2} > 0\,,$$

$$\phi(t) = \tan^{-1}\left(\frac{u}{p\,u'}\right)\,. \qquad \ldots(3)$$

ρ is called the **amplitude** and ϕ the **phase variable**. Here u and u' cannot vanish simultaneously (for a non-trivial solution), so $\rho > 0$. The correspondences $(pu', u) = (\rho, \phi)$ defined by (2) and (3) are analytic with nonvanishing Jacobian.

We now derive an equivalent system of differential equations for $\rho(t)$ and $\phi(t)$.

From equation (3), we have

$$\cot\phi = pu'/u\,. \qquad \ldots(4)$$

Differentiating (4) w.r.t. 't', we get

$$-\operatorname{cosec}^2\phi\,\frac{d\phi}{dt} = \frac{(pu')'}{u} - \frac{pu'^2}{u^2}$$

$$= -q(t) - \frac{1}{p}\cot^2\phi\,.$$

This implies

$$\frac{d\phi}{dt} = \frac{1}{p(t)}\cos^2\phi + q(t)\sin^2\phi\,. \qquad \ldots(5)$$

Differentiating the relation

$$\rho^2 = u^2 + (pu')^2 \qquad \ldots(6)$$

w.r.t 't' and simplifying, we obtain

$$\frac{d\rho}{dt} = \left[\frac{1}{p(t)} - q(t)\right]\rho\sin\phi\cos\phi$$

$$= \frac{1}{2}\left[\frac{1}{p(t)} - q(t)\right]\rho\sin 2\phi\,. \qquad \ldots(7)$$

The system, consisting of first order differential equations (5) and (7), is equivalent to the second order differential equation (1) in the sense that every non-trivial solution of this system defines a unique solution of the differential equation (1) by the Prufer substitution, and conversely. This system is called the **Prufer system** associated with the self-adjoint differential equation (1).

The differential equation (5) of the Prufer system is a first-order differential equation in ϕ and t alone, and not containing the other dependent variable ρ. If a solution $\phi = \phi(t)$ of first order ordinary differential equation (5) is known, then a corresponding solution of first order ordinary differential equation (7) is obtained as

$$\rho(t) = \rho(a)\exp\left[\frac{1}{2}\int_a^t \left\{\frac{1}{p(s)} - q(s)\right\}\sin 2\varphi(s)ds\right] \qquad \ldots(8)$$

Note. 1. Each solution of the Prufer system (5) & (7) depends on two constants, namely,

(i) the initial amplitude $\rho_0 = \rho(t_0)$

(ii) the initial phase $\phi(t_0) = \phi_0$. Changing the constant ρ_0 just multiplies a solution u(t) of equation(1) by a constant factor.

Thus, the zero of any solution $u = u(t)$ of (1) can be located by studying only the differential equation (5).

Note. 2. An advantage of differential equation (5) is that any solution of (5) exists on the whole interval I where p, q are continuous. This is clear from the relations between solutions of (1) and (5).

Note. 3. For the study of zeros of u(t), the Prufer transformation is particularly useful since

$$u(t_0) = 0 \text{ if and only if } \phi(t_0) \equiv 0 \pmod{\pi}. \qquad \ldots(9)$$

This shows that the zero of any solution u(t) of equation (1) occur where the phase function $\phi(t)$ assumes the values 0, $\pm\pi$, $\pm 2\pi$, …, i.e., at all points where $\sin \phi(t) = 0$. At each of these points

$$\cos^2 \phi = 1 , \qquad \ldots(10)$$

and $\frac{d\phi}{dt}$ is positive when $p(t) > 0$, by equation (5).

Note. 4. When $\phi(t) \equiv 0 \pmod{\pi}$, then the relation (4) is not defined. But the final equations (5) and (7) can still be derived by differentiating the relation

$$\tan \phi = \frac{u}{pu'} . \qquad \ldots(10)$$

Illustration. Consider the linear second order homogeneous ordinary differential equation

$$t u'' - u' + t^3 u = 0 . \qquad \ldots(1)$$

It can be written in Liouville normal form as

$$\frac{d}{dt}\left[\frac{1}{t}\frac{du}{dt}\right] + t u = 0, \qquad \ldots(2)$$

giving

$$\left.\begin{aligned} p(t) &= 1/t, \\ q(t) &= t. \end{aligned}\right\} \qquad \ldots(3)$$

Therefore, first equation of Prufer system becomes

$$\begin{aligned} \frac{d\phi}{dt} &= t \cos^2 \phi + t \sin^2 \phi \\ &= t. \end{aligned} \qquad \ldots(4)$$

Solving (4), we obtain

$$\phi(t) = \frac{t^2}{2} + C, \qquad \ldots(5)$$

where C is an arbitrary constant.

Then, the second differential equation of the Prufer system yield (left as an exercise)

$$\rho(t) = \rho(a). \qquad \ldots(6)$$

Therefore, we conclude, from the Pruffer transformation, that

$$u(t) = \rho(a) \sin \left[\frac{t^2}{2} + c\right], \qquad \ldots(7)$$

or

$$u(t) = \frac{(a)\sin(t^2/2 + C)}{\sin(a^2/2 + C)} , \qquad \ldots(8)$$

as the solution of equation (1).

Definition. Consider two differential equations

$$\frac{d}{dt}\left(p_1(t)\frac{du}{dt}\right) + q_1(t)\,u = 0, \qquad \ldots(1)$$

and

$$\frac{d}{dt}\left(p_2(t)\frac{du}{dt}\right) + q_2(t)\,u = 0, \qquad \ldots(2)$$

where $p_i(t)$ and $q_i(t)$ are real-valued continuous functions defined on an interval I such that

$$p_1(t) \geq p_2(t) > 0, \qquad \ldots(3)$$

and

$$q_1(t) \leq q_2(t) \qquad \ldots(4)$$

Then the differential equation (2) is called a **Sturm majorant** for differential equation (1) on the interval I. And differential equation (1) is known as a **Sturm minorant** for equation (2).

Remark. If, in addition,

$$q_1(t) < q_2(t)\,, \qquad \ldots(5)$$

or

$$p_1(t) > p_2(t) > 0 \text{ and } q_2(t) \neq 0\,, \qquad \ldots(6)$$

holds at some point t of I then equation (2) is called a strict **Sturm majorant** for equation (1) on the interval I.

18

BASIC FACTS OF LINEAR SECOND ORDER DIFFERENTIAL EQUATIONS

We shall write the following homogeneous and inhomogeneous scalar differential equations

$$\frac{d}{dt}\left(p(t)\frac{du}{dt}\right) + q(t)\,u = 0 \qquad \ldots(1)$$

$$\frac{d}{dt}\left(p(t)\frac{dw}{dt}\right) + q(t)\,w = h(t) \qquad \ldots(2)$$

as the binary vector differential equations, respectively,

$$\frac{dx}{dt} = A(t)\,x, \qquad \ldots(3)$$

$$\frac{dy}{dt} = A(t)\,y + \begin{pmatrix} 0 \\ h(t) \end{pmatrix}, \qquad \ldots(4)$$

where

$$x = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} u(t) \\ p(t)\frac{du}{dt} \end{pmatrix} \qquad \ldots(5)$$

and

$$y = \begin{pmatrix} y^1 \\ y^2 \end{pmatrix} = \begin{pmatrix} w \\ p(t)\frac{dw}{dt} \end{pmatrix} \qquad \ldots(6)$$

are the vectors and A(t) is the 2×2 matrix

$$A(t) = \begin{pmatrix} 0 & 1/p(t) \\ -q(t) & 0 \end{pmatrix}. \qquad \ldots(7)$$

We find

$$\text{tr}\,\{A(t)\} = 0. \qquad \ldots(8)$$

Unless, otherwise stated, it is assumed that $0 \neq p(t)$, $q(t)$, $h(t)$, and other coefficient functions are continuous, complex-valued functions on a t-interval I (which may or may not be closed and/or bounded).

Result I. If to $\in$ I and c_1 and c_2 are arbitrary complex numbers, then the initial value problem

$$\frac{d}{dt}\left[p(t)\frac{dw}{dt}\right] + q(t)\,w = h(t) \qquad \ldots(1)$$

$$w(t_0) = c_1,$$

$$w'(t_0) = c_2, \qquad \ldots(2)$$

has a unique solution which exists on whole interval I.

Results II. If to $\in$ I, then the initial-value problem

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = 0,\ t \in I \qquad \ldots(1)$$

$$u(t_0) = 0,$$
$$u'(t_0) = 0, \qquad \ldots(2)$$

has trivial solution, namely,

$$u(t) \equiv 0\ , \qquad \ldots(3)$$

on I, as the unique solution.

Result III. If $u(t) \neq 0$ is a solution of differential equation

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} + q(t)\,u = 0,\ t \in I, \qquad \ldots(1)$$

then the zeros of u(t) cannot have a cluster point in I.

Proof. If possible, suppose that u(t) have a cluster point in I. Then, u(t) has an infinite number of zeros in some bounded interval [a, b], contained in I.

Let $$E = \{t : t \in [a, b] \text{ such that } u(t) = 0\}\ . \qquad \ldots(2)$$

Then E is a bounded and infinite subset of real number. Therefore, by the Bolzano-Weierstrass theorem, the set E has a limit point, say t_0. Moreover $t_0 \in [a, b] \subset I$. Thus, there exists a sequence of distinct points of E(hence in I) which converges to t_0 .

Let $t \to t_0$ through the sequence of zeros of u. Then the continuity of u implies that

$$\lim_{t \to t_0} u(t) = u(t_0), \qquad \ldots(3)$$

where $t \to t_0$ in the above sense. This gives

$$u'(t_0) = 0. \qquad \ldots(4)$$

Moreover, by definition,

$$u'(t_0) = \lim_{t \to t_0} \frac{u(t) - u(t_0)}{t - t_0}\ . \qquad \ldots(5)$$

When $t \to t_0$ through the zero of u, then R.H.S. of (5) becomes zero so,

$$u'(t_0) = 0\ . \qquad \ldots(6)$$

Thus u is a solution of (1) satisfying the initial conditions (4) and (6). Hence, by uniqueness theorem, u (t) $\equiv$ 0 for all $t \in [, b] \subset I$.

This is a contradiction. Hence, the result follows immediately.

Result IV. (Superposition principles)

(A) Let $u = u_1(t), u_2(t)$ be two solutions of homogeneous linear differential equation

$$\frac{d}{dt}\left(p(t)\frac{du}{dt}\right) + q(t)\,u = 0,\ t \in I, \qquad \ldots(1)$$

and c_1 and c_2 be constants. Then $c_1\,u_1(t) + c_2\,u_2(t)$ is a solution of (1).

(B) If $w_0(t)$ is a solution of non homogeneous linear differential equation

$$\frac{d}{dt}\left(p(t)\frac{dw}{dt}\right) + q(t)\,w = h(t),\ t \in I, \qquad \ldots(2)$$

then $w_1(t)$ is also a solution of (2) if and only if

$$u(t) = w_1(t) - w_0(t) \qquad \ldots(3)$$

is a solution of the corresponding homogeneous differential equation.

Results V. If $u_1(t)$, $u_2(t)$ are two solutions of homogeneous differential equation, then the corresponding vector solutions

$$x_1 = \begin{pmatrix} u_1(t) \\ p(t)\,u_1'(t) \end{pmatrix},$$

$$x_2 = \begin{pmatrix} u_2(t) \\ p(t)\,u_2'(t) \end{pmatrix},$$

of binary system

$$\frac{dx}{dt} = A(t)\,x\,, \qquad \ldots(4)$$

are linearly independent at every value of t if and only if $u_1(t)$ and $u_2(t)$ are linearly independent in the sense that if c_1 and c_2 are constants such that

$$c_1\, u_1(t) + c_2\, u_2(t) \equiv 0, \text{ for all } t,$$

then $c_1 = c_2 = 0$.

Result VI. If $u_1(t)$, $u_2(t)$ are solution of homogeneous differential equation, then there is a constant c(depending on $u_1(t)$ and $u_2(t)$) such that their Wronskian W(t) is given by

$$\begin{aligned} W(t) &= W(u_1(t), u_2(t)) \\ &= u_1(t)\, u_2'(t) - u_1'(t)\, u_2(t) \\ &= \frac{c}{p(t)}\,. \end{aligned} \qquad \ldots(1)$$

Proof. A solution matrix for linear binary homogeneous system

$$\frac{dx}{dt} = A(t)x, \qquad \ldots(2)$$

is

$$X(t) = \begin{pmatrix} u_1(t) & u_2(t) \\ p(t)u_1'(t) & p(t)u_2'(t) \end{pmatrix},$$

and

$$\det X(t) = p(t)\, W(t). \qquad \ldots(3)$$

Hence, the result follows, immediately.

Result VII. Lagrange Identity. Consider the pair of differential equations.

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} + q(t)\, u(t) = f(t), \qquad \ldots(1)$$

$$\frac{d}{dt}\left\{p(t)\frac{dv}{dt}\right\} + q(t)\, v(t) = g(t)\,, \qquad \ldots(2)$$

where $f = f(t)$ and $g = g(t)$ are continuous functions on interval I. Multiplying the second relation (2) by u(t), the first (1) by u(t), and the results subtracted, one finds

$$\frac{d}{dt}\left[p(t)\left\{u\frac{dv}{dt} - v\frac{du}{dt}\right\}\right] = gu - fv\,. \qquad \ldots(3)$$

The relation (3) is called the **Lagrange identity**. Its integrated from

$$[p(u\, v' - u'v)\,]_a^t = \int_a^t\ (gu - fv)\, ds\,, \qquad \ldots(4)$$

where $[a, t] \subset I$, is called **Green’s formula**.

Special Case. If $f = g = 0$, i.e., when u & v are solutions of homogeneous differential equations, then (3) implies that

$$p(t)\,\{u(t)\,v'(t) - v(t)\,u'(t)\} = \text{constt} = c, \text{ say.}$$

This gives

$$u(t)\,v'(t) - v(t)\,u'(t) = \frac{c}{p(t)}. \qquad \ldots(5)$$

This also proves, the result VI for homogeneous differential equations.

(a) In particular, it shows that u(t) and v(t) are linearly independent solutions of homogeneous differential equation iff $c \neq 0$, in equation (5).

(b) If $p(t) = \text{const} \neq 0$, (e. g., $p(t) \equiv 1$), then equation (5) shows that the Wronskian of any pair of solutions of homogeneous second order linear differential equations is a constant.

Result VIII. If one solution $u(t) \not\equiv 0$ of differential equation

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} + q(t)\,u = 0\,, \qquad \ldots(1)$$

is known, then the determination (atleast, locally) of other solution v(t) of (1) is obtained by considering a certain (reduction of order) differential equation of first order.
Dividing relation (5) of result VII by $u^2(t)$, we get

$$\frac{d}{dt}\left\{\frac{v(t)}{u(t)}\right\} = \frac{c}{p(t)u^2(t)}, \quad c = \text{const.}. \qquad \ldots(2)$$

Integrating

$$\frac{v(t)}{u(t)} = c_1 + c\int_a^t \frac{ds}{p(s)u^2(s)}\,,$$

this gives

$$v(t) = c_1\,u(t) + c\,u(t)\left[\int_a^t \frac{ds}{p(s)u^2(s)}\right], \qquad \ldots(3)$$

for $a \in I$ and for all $t \in I$.

Theorem. 18.1. Let $u(t) \not\equiv 0$ be a real valued solution of homogeneous linear second order differential equation

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} + q(t)\,u = 0\,,$$

on the interval [a, b], where $p(t) > 0$ and q(t) are real-valued and continuous. Let u(t) have exactly $n(\geq 1)$ zeros, say,

$$t_1 < t_2 < \ldots < t_n \qquad \text{on } [a, b].$$

Let $\phi(t)$ be a continuous function defined by

$$\phi(t) = \tan^{-1}\left(\frac{u}{pu'}\right),$$

and

$$0 \leq \phi(a) < \pi\,.$$

Prove that

$$\phi(t_k) = k\pi,\ k = 1, 2, \ldots, n$$

and

$$\phi(t) > k\pi \text{ for } t \in [t_k, b],\ k = 1, 2, \ldots, n$$

Proof. From the hypothesis we write

$$\tan \phi = \frac{u}{pu'} . \qquad \ldots(1)$$

Differentiating (1), we shall obtain (left as an exercise to the reader)

$$\frac{d\varphi}{dt} = \frac{1}{p(t)} \cos^2 \phi + q(t) \sin^2 \phi . \qquad \ldots(2)$$

We note that at a t-value, say t_0,

$$u(t_0) = 0$$

iff $\phi(t_0) = 0$

iff $\phi(t_0) \equiv 0 \pmod \pi$ ...(3)

and, then $\cos^2\phi(t_0) = 1$ . ...(4)

So equation (2) gives

$$\frac{d\varphi}{dt} = \frac{1}{p(t_0)} > 0, \qquad \ldots(5)$$

as p(t) > 0. This shows that $\phi(t)$ is an increasing function in the neighbourhoods of points where

$$\phi(t) = j\pi \qquad \ldots(6)$$

for some integer j. It follows that from equation (3) that

$$\phi(t_k) = k\pi, \qquad \ldots(7)$$

and $\phi(t) > \phi(t_k) = k\pi$ for $t \in [t_k, b]$ . ...(8)

This proves the theorem.

19

THEOREMS OF STURM AND ZEROS OF SOLUTIONS

This lesson deals in the main, with differential equations of the type

$$L(u) \equiv \frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = 0, \qquad \ldots(1)$$

in which p(t) and q(t) are, throughout the closed interval $a \le t \le b$, continuous real functions of the real variable t. p(t) does not vanish, and may therefore be assumed to be positive, and has a continuous first derivative throughout the interval.

The fundamental existence theorem has established the fact that differential equation (1) has one and only one continuous solution with a continuous derivative which satisfies the initial conditions

$$u(t_0) = c_1,\ u'(t_0) = c_2, \qquad \ldots(2)$$

where t_0 is any point of the closed interval [a, b]. But valuable as the existence theorem is from the theoretical point of view, it supplies little or no information as to the nature of the solution whose existence it guarantees.

It is important from the point of view of physical applications, and not without theoretical interest, to determine the number of zeros which the solution has in the interval [a, b]. This problem was first attacked by Sturm in 1836. The theory based upon his work is now regarded as classical. The "theorems of comparison", which form the core of the present lesson, are fundamental, and serve as the basis of a considerable body of further investigation. In one of the previous lessons, we have proved that no continuous solution of equation (1) have an infinite number of zeros in [a, b] without being identically zero.

The phrase "comparison theorems" for differential equations is used in the sense stated below :

"If a solution of a differential equation has a certain known property then the solutions of a second differential equation have the same or some related property under certain conditions".

These theorems have many interesting implications in the theory of oscillations.

Theorem (19.1) Let $u_1(t)$ and $u_2(t)$ be two linearly independent solutions of

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = 0\,,$$

in [a, b], with $p(t) > 0$. Prove that $u_1(t)$ and $u_2(t)$ do not admit common zeros.

Proof. If possible, suppose that solutions $u_1(t)$ and $u_2(t)$ admit a common zero at $t = t_0$, say, $t_0 \in [a, b]$. Then

$$u_1(t_0) = u_{12}(t_0) = 0\,. \qquad \ldots(1)$$

From Abel's lemma, we have

$$p(t)\,[u_1(t)\,u'_2(t) - u_1'(t)\,u_2'(t)] = \text{constt} = c\,, \qquad \ldots(2)$$

for all $t \in [a, b]$. In particular, for $t = t_0$, we obtain

$$p(t_0)\,[u_1(t_0)\,u_2'(t_0) - u_1'(t_0)\,u_2(t_0)] = c$$

or $$c = 0\,. \qquad \ldots(3)$$

Hence, from equations (2) and (3), we obtain

$$p(t)\,[u_1(t)\,u_2(t) - u_1(t)\,u_2(t)] = 0, \text{ for all } t \in [a, b]$$

$$\Rightarrow \quad u_1(t)\,u_2(t) - u_1(t)\,u_2(t) = 0 \text{ for all } t \in [a, b], \text{ as } p(t) > 0$$

$$\Rightarrow \quad W(u_1, u_2)(t) = 0 \text{ for all } t \in [a, b]\,. \qquad \ldots(4)$$

Thus, it follows that u_1 and u_2 are linearly dependent (why) which is a contradiction to the hypothesis. Hence u_1 and u_2 can't have common zero. This completes the proof.

Restatement. If non-trivial solutions $u_1(t)$ and $u_2(t)$ of differential equation

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = 0,\ t \in [a, b],$$

have a common zero on [a, b], then they are linearly dependent on [a, b].

Theorem. (19.2). Let u(t) be a non-trivial solution of the differential equation.

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = 0 \text{ in } [a, b].$$

Prove that the zeros of u(t) are isolated.

Proof. Let $t = t_0$ be a zero of u(t). Then

$$u(t_0) = 0\,. \qquad \ldots(1)$$

Since u(t) is a non-trivial solution of differential equation,

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = 0,\ t \in [a, b] \qquad \ldots(2)\backslash$$

it follows that

$$u'(t_0) \neq 0, \qquad \ldots(3)$$

otherwise, by uniqueness theorem, $u(t) \equiv 0$, which is not the case. Now there are two possible cases.

Case 1. Where $$u'(t_0) > 0\,. \qquad \ldots(4)$$

Since the derivative of u(t) is continuous and positive at $t = t_0$, it follows that the function u(t) is strictly increasing in some neighbourhood of $t = t_0$. This means that $t = t_0$ is the only zero of u(t) in that neighbourhood. This shows that the zero, $t = t_0$, of u(t) is isolated.

Case 2. When $$u'(t_0) < 0\,. \qquad \ldots(5)$$

The proof is similar (left as an exercise) to that of case I. This completes the proof.

Theorem. (19.3). Let $u_1(t)$ and $u_2(t)$ be non-trivial linearly dependent solutions of differential equation

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\,u = 0,$$

on [a, b], and $p(t) > 0$. Then the zeros of $u_1(t)$ and $u_2(t)$ are identical.

Proof. Since $u_1(t)$ and $u_2(t)$ are linearly dependent on [a, b], so there exists constants c_1 and c_2, not both zero, such that

$$c_1\,u_1(t) + c_2\,u_2(t) = 0, \qquad \ldots(1)$$

for all $t \in [a, b]$. Now, we shall show that neither c_1 nor c_2 is zero.
If $c_2 = 0$, then (1) gives

$$c_1 u_1(t) = 0 \text{ for all } t \in [a, b]$$

$$\Rightarrow \quad c_1 = 0, \qquad \ldots(2)$$

as $u_1(t) \not\equiv 0$ on [a, b]. This is a contradiction to the assumption that both of c_1 and c_2 are not zero.

So

$$c_2 \neq 0. \qquad \ldots(3)$$

Similarly (exercise),

$$c_1 \neq 0. \qquad \ldots(4)$$

Let $t = t_0$ be a zero of $u_1(t)$. Then

$$u_1(t_0) = 0. \qquad \ldots(5)$$

Equation (1) gives

$$c_1 u_1(t_0) + c_2 u_2(t_0) = 0$$

$$\Rightarrow \quad c_2 u_2(t_0) = 0$$

$$\Rightarrow \quad u_2(t_0) = 0 . \qquad \ldots(6)$$

This shows that t_0 is also a zero $u_2(t)$. Thus, every zero of $u_1(t)$ is also a zero of $u_2(t)$. Similarly, (exercise) every zero of $u_2(t)$ is also a zero of $u_1(t)$. Hence, $u_1(t)$ and $u_2(t)$ both have the same zeros. This completes the proof.

Illustration. Consider the differential equation

$$\frac{d^2u}{dt^2} + u = 0, \qquad \ldots(1)$$

with $p(t) \equiv 1$, $q(t) \equiv 1$ on every interval [a, b] of real line.

Let

$$u_1(t) = A \sin t, \qquad \ldots(2)$$

$$u_2(t) = B \sin t, \qquad \ldots(3)$$

where A and B are arbitrary constants. Then $u_1(t)$ and $u_2(t)$ are two non-trivial linearly dependent solutions of the given differential equation. These solutions have the following common zeros at

$$t = \pm n\pi, \; n = 0, 1,2,3,\ldots$$

and no other zero.

Theorem. (19.4). (known as Separation theorem). Let $u_1(t)$ and $u_2(t)$ be two real-valued non-trivial linearly independent solutions of differential equation

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q(t)\, u = 0 ,$$

on the interval [a, b], with $p(t) > 0$. Then the zeros of $u_1(t)$ separate and are separated by those of $u_2(t)$.

Proof. Let $t = t_1, t_2$ be two consecutive zeros of $u_1(t)$ on [a, b] so,

$$u_1(t_1) = u_1(t_2) = 0, \; a \le t_1 < t_2 < b. \qquad \ldots(1)$$

Since u_1 and u_2 are linearly independent on [a, b], so they do not admit common zeros (why). In particular,

$$u_2(t_1) \neq 0, \; u_2(t_2) \neq 0. \qquad \ldots(2)$$

We shall now show that u_2 has one zero in the open interval (t_1, t_2). If possible, suppose that it does not happen. Then, the quotient function $\left(\frac{u_1}{u_2}\right)$ satisfies (how) all the requirements of Rolle's theorem on the interval $[t_1, t_2]$, which are

i) $\frac{u_1}{u_2}$ is continuous in $[t_1, t_2]$

ii) $\frac{u_1}{u_2}$ has continuous derivative in (t_1, t_2)

iii) $\left(\frac{u_1}{u_2}\right)(t_1) = \left(\frac{u_1}{u_2}\right)(t_2) = 0$.

So, $\left(\frac{u_1}{u_2}\right)^{\copyright}$ has atleast one zero say, t = c, in (t_1, t_2). That is

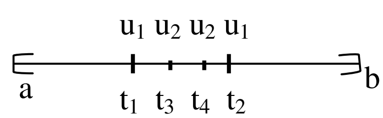
$$\left[\frac{d}{dt}\left(\frac{u_1}{u_2}\right)\right]_{t=0} = 0$$

$$\Rightarrow \qquad \frac{W(u_1, u_2)(c)}{g[u_2(c)]^2} = 0$$

$$\Rightarrow \qquad W(u_1, u_2)(c) = 0, \qquad \ldots(3)$$

which implies that solution u_1 and u_2 are linearly dependent on [a, b]. This contradicts the hypothesis that u_1 and u_2 are linearly independent on [a, b]. This contradiction proves that $u_2(t)$ has atleast one zero in the open interval (t_1, t_2). So, the zero of u_1 are separated by zero of u_2.

Now we shall show that u_2 has exactly one zero in the open interval (t_1, t_2). If possible suppose that $u_2(t)$ has two consecutive zeros t_3, t_4 $(t_3 < t_4)$ in the open interval (t_1, t_2).



On interchanging the role of solutions u_1 and u_2, the just above proved conclusion shows that there is atleast one zero, say $t = t_5$, of $u_1(t)$ in the open interval (t_3, t_4) with $t_1 < t_5 < t_2$.

This contradicts the assumption that t_1 and t_2 were two consective zero of $u_1(t)$. This contradiction proves that $u_2(t)$ has exactly one zero between two consective zeros of $u_1(t)$. Similarly, between two consective zeros of $u_2(t)$, there will be exactly one zero of $u_1(t)$. This shows that the zeros of u_1 separate the zero of $u_2(t)$. Hence the proof is complete.

Restatement. Prove that the zeros of two real linearly independent solutions of a linear differential equation of the second order separate one another.

Note. This theorem may be stated roughly as follows : "The zeros of all solutions of a given differential equation oscillate equally rapidly".

This statement implies that the number of zeros of any solution in an interval $[\alpha, \beta] \subset [a, b]$, cannot exceed the number of zeros of any independent solution in the same interval by more than one.

Illustration. Consider the differential equation

$$\frac{d^2u}{dt^2} + u = 0. \qquad \ldots(1)$$

Here $p(t) \equiv 1$, $q(t) \equiv 1$. Then

$$u_1(t) = \sin t, \qquad \ldots(2)$$
$$u_2(t) = \cos t, \qquad \ldots(3)$$

are two linearly independent solutions of the given differentiation equation (1). Therefore, between any two consective zeros of one of these two solutions, there is precisely one zero of the other solution. We know that zero of $u_1(t)$ are

$$t = n\,\pi \text{ for } n = 0, 1, 2, \ldots. \qquad \ldots(4)$$

and zeros of $u_2(t)$ are

$$t = (2\,m + 1)\,\frac{\pi}{2} \text{ for } m = 0, 1, 2, \ldots \qquad \ldots(5)$$

which are separated by each other.

Exercise. (1) Show that between any two consective real zeros of sin 3t + cos 3t, there is precisely one zero of sin 3t–cos 3t and conversely. State the result which is used to show this

Exercise. (2) Show that the zero of

$$u_1(t) = A \sin t + B \cos t$$
$$u_2(t) = C \sin t + D \cos t,$$

separate one another provided that

$$AD - BC \neq 0.$$

A, B, C, D being real constants.

Exercise. (3) Use the Sturm separation theorem to show that between any two consective zeros of sin 2 t cos 2t, there is precisely one zero of sin 2t – cos 2t.

Theorem (19.5). (known as Sturm's Fundamental Theorem). Consider the differential equations

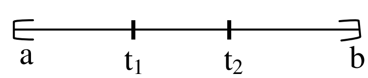
$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q_1(t)\,u = 0, \qquad \ldots(1)$$
$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + q_2(t)\,u = 0, \qquad \ldots(2)$$

on the interval [a, b] such that $p(t) > 0$ have a continuous derivative on [a, b], and $q_1(t) < q_2(t)$ be continuous functions on [a, b]. Let $u_1(t)$ and $u_2(t)$ be respective non-trivial solutions of equations (1) and (2). Prove that between any two consective zeros of $u_1(t)$ on [a, b], there lies at least one zero of $u_2(t)$.

Proof. Let $t_1 < t_2$ be two consective zeros of $u_1(t)$ on [a, b]. Then

$$u_1(t_1) = 0 = u_1(t_2). \qquad \ldots(3)$$

By hypothesis, we have



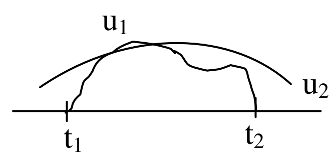
$$\frac{d}{dt}[p(t)\,u_1'(t)] + q_1(t)\,u_1(t) = 0, \quad \ldots(4)$$
$$\frac{d}{dt}[p(t)\,u_2'(t)] + q_2(t)\,u_2(t) = 0, \quad \ldots(5)$$

for all $t \in [a, b]$. Multiplying (4) by $u_2(t)$ and (5) by $u_1(t)$, and then subtracting, we obtain

$$\frac{d}{dt}[p(t)\,\{u_1'(t)\,u_2(t) - u_1(t)\,u_2'(t)\}] = [q_2(t) - q_1(t)]\,u_1(t)\,u_2(t), \qquad \ldots(6)$$

for all $t \in [a, b]$. Integrating w.r.t. 't' over the interval $[t_1, t_2]$, we obtain after using (3),

$$p(t_2)\,u_1'(t_2)\,u_2(t_2) - p(t_1)\,u_1'(t_1)\,u_2(t_1) = \int_{t_1}^{t_2} \{q_2(t) - q_1(t)\}\,u_1(t)\,u_2(t)\,dt. \ldots(7)$$



If possible, on the contrary assume that $u_2(t)$ does not have any zero in the open interval (t_1, t_2) .
W.l.o.g, we can assume that $u_1(t)$ and $u_2(t)$ are positive in the open interval (t_1, t_2), i.e.,

$$u_1(t) > 0,\ u_2(t) > 0 \text{ in } (t_1, t_2). \quad \ldots(8)$$

By hypothesis $\quad p(t_2) > 0,\ p(t_1) > 0. \quad \ldots(9)$

Also, by assumption

$$u_2(t_1) \geq 0,\ u_2(t_2) \geq 0. \quad \ldots(10)$$

As t_1 and t_2 are consective zeros of u_1, so

$$u_1'(t_1) > 0,\ u_1'(t_2) < 0. \quad \ldots(11)$$

From equations (7) and (9) to (11), it follows that

$$\int_{t_1}^{t_2} \{q_2(t) - q_1(t)\}\ u_1(t)\ u_2(t)\ dt \leq 0. \quad \ldots(12)$$

Also,

$$q_2(t) - q_1(t) > 0, \quad \ldots(13)$$

on $[t_1, t_2]$, and

$$u_1(t) > 0,\ u_2(t) > 0 \quad \ldots(14)$$

on (t_1, t_2) by assumption. From equations (13) and (14), it follows that

$$\int_{t_1}^{t_2} \{q_2(t) - q_1(t)\}\ u_2(t)\ u_2(t)\ dt > 0. \quad \ldots(15)$$

This contradicts (12). This contradiction proves that our assumption that $u_2(t)$ has no zero in (t_1, t_2) is wrong. So, $u_2(t)$ has atleast one zero between two consective zeros of $u_1(t)$.

This completes the proof.

Remark. (1) In particular, if $u_1(t)$ and $u_2(t)$ are both zero at $t = t_1$ then the theorem (19.5) shows that $u_2(t)$ vanishes again before the consective zero of $u_1(t)$ appears. Thus, $u_2(t)$ oscillates more rapidly that $u_1(t)$.

Remark. (2) Theorem (19.5) is also termed as "Sturm's comparison theorem".

Example. State Sturm's fundamental comparison theorem Verify it in the case of real solutions of the differential equation equations

$$\frac{d^2u}{dt^2} + A^2\, u = 0,$$

and

$$\frac{d^2u}{dt^2} + B^2\, u = 0,$$

where A and B are constants such that $B > A > 0$.

Solution. Let

$$u_1(t) = \sin At, \quad \ldots(1)$$

$$u_2(t) = \sin Bt\ . \quad \ldots(2)$$

Then $u_1(t)$ and $u_2(t)$ are real solution of given equations, respectively. Consecutive zeros of $u_1(t)$ are

$$\frac{n\pi}{A},\ (n+1)\frac{\pi}{A} \text{ for } n = 0, \pm 1, \pm 2, \ldots \quad \ldots(3)$$

By Sturm's comparison theorem (19.5) with $p(t) \equiv 1$, $q_1(t) = A^2$, $q_2(t) = B^2$, $q_2>q_1$), the solution $u_2(t)$ has atleast one zero, say ξ_n, between the zeros $\frac{n\pi}{A}$ and $(n+1)\frac{\pi}{A}$ of $u_1(t)$. That is,

$$\frac{n\pi}{A} < \xi_\eta < (n+1)\frac{\pi}{A},\ n = 0, \pm 1, \pm 2, \ldots, \qquad \ldots(4)$$

we making take

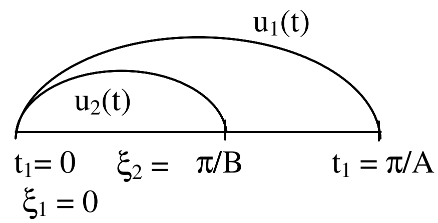
$$\xi_\eta = \frac{(n+1)\pi}{B}, \qquad \ldots(5)$$

as zero of $u_2(t)$.
In particular, for n = 0,

$$t_1 = 0,\ t_2 = \frac{\pi}{A} \qquad \ldots(6)$$

are two consective zeros of $u_1(t)$. The zero $t = \frac{\pi}{B}$ of $u_2(t)$ lies between t_1 and t_2, as

$$0 < \frac{\pi}{B} < \frac{\pi}{A}\ . \qquad \ldots(8)$$



This verifies the results of comparison

theorem. Consequently, the solutions of differential equation (2) oscillates more rapidly than those of differential equation (1).

Theorem (19.6) In the differential equation

$$\frac{d^2u}{dt^2} + q(t)\, u = 0, \qquad \ldots(1)$$

let q(t) be real-valued continuous, and satisfying

$$0 < m \le q(t) \le M. \qquad \ldots(2)$$

If $u = u(t) \not\equiv 0$ is a solution with a pair of consecutive zeros $t = t_1, t_2$ ($t_1 < t_2$), prove that

$$\frac{\pi}{\sqrt{M}} \le t_2 - t_1 \le \frac{\pi}{\sqrt{m}}. \qquad \ldots(3)$$

Solution. Consider the differential equations with constant coefficients

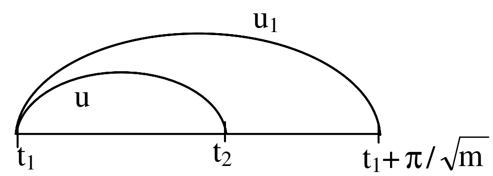
$$\frac{d^2u}{dt^2} + m\, u = 0, \qquad \ldots(4)$$

$$\frac{d^2u}{dt^2} + Mu = 0. \qquad \ldots(5)$$

we note that

$$u_1(t) = \sin \sqrt{m}\,(t - t_1) \qquad \ldots(6)$$

is a solution of differential equation (4) such that $u_1(t_1) = 0$ and $u_1\left(t_1 + \frac{\pi}{\sqrt{m}}\right) = 0$. That is, t_1 and $t_1 + \frac{\pi}{\sqrt{m}}$ are consecutive zeros of solution $u_1(t)$ of differential equations (4). It is also given that



$$q(t) \geq m > 0\ .$$

Hence, by Sturm's comparison theorem (19.5) the zeros t_1 and $t_1 + \frac{\pi}{\sqrt{m}}$ of $u_1(t)$ are separated by zero $t = t_2$ of $u(t)$. So

$$t_1 < t_2 \leq t_1 + \frac{\pi}{\sqrt{m}}$$

$$\Rightarrow \qquad t_2 - t_1 \leq \frac{\pi}{\sqrt{m}}\ . \qquad \ldots(7)$$

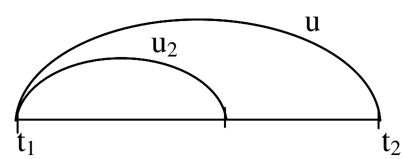
we again note that

$$u_2(t) = \sin\sqrt{M}\,(t - t_1) \qquad \ldots(8)$$

is a solution of differential equation (5) with consective zeros $t = t_1,\ t_1 + \frac{\pi}{\sqrt{M}}$. Also

$$M \geq q(t) > 0\ . \qquad \ldots(9)$$

Hence, by Sturm's comparison theorem (19.5), the consecutive zeros t_1, t_2 of $u(t)$ are separated by the zero $t = t_1 + \frac{\pi}{\sqrt{M}}$ of $u_2(t)$.



That is

$$t_1 < t_1 \ \frac{\pi}{\sqrt{M}} \leq t_2$$

$$\Rightarrow \qquad \frac{\pi}{\sqrt{M}} \leq t_2 - t_1\ . \qquad \ldots(10)$$

Hence from equation (9) and (10), the result follows.

Remark. Theorems (19.4) to (19.6) provide information about the interlacing of zeros.

Example. Given the differential equation

$$\frac{d^2u}{dt^2} + q(t)\,u = 0, \qquad \ldots(1)$$

where $q(t) > 0$ on $[a, b]$. Let q_m denote the minimum value of $q(t)$ on $[a, b]$. If

$$q_m > \frac{k^2\pi^2}{(b-a)^2}, \qquad \ldots(2)$$

show that every real solutions of the given equation has atleast k zeros on $a \leq t \leq b$.

Solution. Consider the differential equation

$$\frac{d^2u}{dt^2} + \frac{k^2\pi^2}{(b-a)^2}\,u = 0, \qquad \ldots(3)$$

which has a solution

$$u_1(t) = \sin\left(\frac{k\pi}{b-a}\right)(t-a)\,, \qquad \ldots(4)$$

and its zero are given by

$$t-a = \frac{n(b-a)}{k\pi}\pi = \frac{n(b-a)}{k}, \text{for } n = 0, \pm 1 \pm 2, \ldots$$

i.e.
$$(t-a) = 0, \pm\frac{b-a}{k}, \pm\frac{2(b-a)}{k}, \pm\frac{3(b-a)}{k}, \ldots. \qquad \ldots(5)$$

we note that the zeros

$$t = a, \frac{b-a}{k}+a, \frac{2(b-a)}{k}+a, \ldots, \frac{k(b-a)}{k}+a, \qquad \ldots(6)$$

$$(\equiv t_1, t_2, t_3, \ldots, t_{k+1} \text{ say})$$

of $u_1(t)$ are consective, lie in $a \le t \le b$ and $(k+1)$ in number.
Since q_m is the minimum value of $q(t)$ an $a \le t \le b$, so,

$$q(t) \ge q_m > \frac{k^2\pi^2}{(b-a)^2} \text{ on } a \le t \le b\,,$$

$$\Rightarrow \qquad q(t) > \left(\frac{k\pi}{b-a}\right)^2 \text{ on } a \le t \le b. \qquad \ldots(7)$$

Now, we apply the Sturm comparison theorem (19.5) to differential equations (1) and (3) with inequality (7). By this theorem, there is atleast one zero of every solution of equation (1) between any two consective zeros of $u_1(t)$ in $a \le t \le b$. Consequently, there are atleast k zero of every solution of equation (1) on [a, b].

This completes the solution.

Modifications due to Picone (without proof)

Picone (1909) considered the more general case which compares the rapidity of the oscillations of the solutions of the two differential equations

$$\frac{d}{dt}\left\{p_1(t)\frac{du}{dt}\right\} - q_1(t)\,u = 0, \qquad \ldots(1)$$

$$\frac{d}{dt}\left\{p_2(t)\frac{dv}{dt}\right\} + q_2(t)\,v = 0, \qquad \ldots(2)$$

wherein

$$p_1(t) \ge p_2(t) > 0, \qquad \ldots(3)$$

and

$$q_1(t) \le q_2(t), \qquad \ldots(4)$$

in [a, b]. He showed that between any two consective zeros of u, there is atleast one zero of v.

Article. (19.7) Find conditions that the solutions of differential equation

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} - q(t)\,u = 0 \qquad \ldots(1)$$

for t in [a, b] may be oscillatory or non-oscillatory.

Solution. The functions $p(t)$ and $q(t)$ in equation (1) are being supposed to be continuous and bounded in the closed interval [a, b]. Let M_p and M_q be the upper bounds of $p(t)$ and $q(t)$ in [a, b] respectively. Let m_p and m_q be their lower bounds, respectively. Then, throughout [a, b], we have

$$M_p \ge p(t) \ge m_p > 0, \qquad \ldots(2)$$

and

$$M_q \geq q(t) \geq m_q \ . \qquad \ldots(3)$$

Result. 1. Consider, the differential equation (with constant coefficients)

$$\frac{d}{dt}\left\{m_p \frac{du}{dt}\right\} - m_q\, u = 0, \qquad \ldots(4)$$

which may be written as

$$\frac{d^2u}{dt^2} - \left(\frac{m_q}{m_p}\right) u = 0. \qquad \ldots(5)$$

By Picone theorem, the solutions of equation (1) do not oscillate more rapidly in [a, b] than the solution of (4). We observe that equation (4), in its alternative form (5), is immediately integrable. Solutions of (5) are as follows

(i) If $m_q > 0$, there is the exponential solution $\exp\left[\left\{\sqrt{\frac{m_q}{m_p}}\right\}t\right]$, which has no zero in [a, b]. Similarly, if $m_q = 0$, the comparison solution may be taken as unity. Hence, if $m_q \geq 0$ the solutions of (4) are non-oscillatory. This leads to the conclusion that if $q(t) \geq 0$ throughout the interval [a, b], the solutions of the given differential equation (1) are non-oscillatory.

(ii) If $m_q < 0$, there is the oscillatory solution $\sin\left\{\left(\sqrt{-\frac{m_q}{m_p}}\right)t\right\}$. The interval between its consective zeros, or between consective zeros of any other solution of the comparison equation, is $\left\{\left(\sqrt{-\frac{m_p}{m_q}}\right)\pi\right\}$. If, therefore,

$$\pi\sqrt{\left(-\frac{m_p}{m_q}\right)} > b - a, \qquad \ldots(6)$$

no solution of the given equation can have more than one zero in the interval [a, b]. Consequently the solutions of the given differential equation (1) are non-oscillatory provided

$$-\frac{m_q}{m_p} < \frac{\pi^2}{(b-a)^2}\,. \qquad \ldots(7)$$

Result II. Now, we consider a second comparison equation (with constant coefficients)

$$\frac{d}{dt}\left\{M_p \frac{du}{dt}\right\} - M_q\, u = 0, \qquad \ldots(8)$$

or equivalently

$$\frac{d^2u}{dt^2} - \left(\frac{M_q}{M_p}\right) u = 0. \qquad \ldots(9)$$

Then the solutions of differential equation (1) oscillate atleast as rapidly as those of differential equation (9).

Let M_q be negative ; then the solutions of (9) are oscillatory, and the interval between consective zeros of any solution is

$$\left\{\sqrt{-M_p/M_q}\right\}\pi .$$

It follows that a sufficient condition that the solutions of the given differential equation (1) should have atleast m zeros in the interval [a, b] is that

$$m\,\pi\sqrt{(-M_p/M_q)} \le b - a$$

or

$$-\left(\frac{M_q}{M_p}\right) \ge \frac{m^2\pi^2}{(b-a)^2}. \qquad \ldots(10)$$

In particular, a sufficient condition that the differential equation (1) should possess a solution which oscillates in the interval [a, b] is that

$$-\left(\frac{M_q}{M_p}\right) \ge \frac{\pi^2}{(b-a)^2}.$$

Theorem. (19.8) Consider the differential equations

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} + q_1(t)\,u = 0, \qquad \ldots(1)$$

$$\frac{d}{dt}\left\{p(t)\frac{dv}{dt}\right\} + q_2(t)\,v = 0, \qquad \ldots(2)$$

where $p(t) > 0$, $q_2(t) > q_1(t)$ on $a \le t \le b$. Furthermore, either

(i) $$\frac{u'(a)}{u(a)} \ge \frac{v'(a)}{v(a)},\ u(a) \ne 0,\ v(a) \ne 0, \qquad \ldots(3)$$

or

(ii) $$u(a) = 0,\ v(b) = 0. \qquad \ldots(4)$$

Then v(t) has atleast as many zeros in [a, b] as u(t). In the case (3), if the zeros of u(t) are $t_1, t_2, \ldots, t_n$ with $a < t_1 < t_2 < \ldots < t_n \le b$ and the zeros of v(t) are $\xi_1, \xi_2, \ldots, \xi_m$ with $a < \xi_1 < \xi_2 < \ldots < \xi_m \le b$, then

$$\xi_k < t_k . \qquad \ldots(5)$$

Proof. By the fundamental comparison theorem, between any two zeros of u(t), there is atleast one zero of v(t). Thus, v(t) has atleast n−1 zeros in [a, b]. It is sufficient to show that v(t) has a zero lying in the interval $[a, t_1]$.

In case (ii), this is obvious as t = a is also a zero of v. In case (i), we assume, w. l. o. g., that $u(t) > 0$, $v(t) > 0$ in (a, t_1). We have

$$[p(t)\,\{v(t)\,u'(t) - v'(t)\,u(t)\}]_{t=a}^{t=t_1}$$

$$= p(t_1)\,u'(t)\,v(t_1) - p(a)\left\{\frac{u'(a)v(a) - u(a)v'(a)}{u(a)v(a)}\right\}.\,u(a)\,v(a)$$

$$= p(t_1)\,u'(t_1)\,v(t_1) - p(a)\left\{\frac{u'(a)}{u(a)} - \frac{v'(a)}{v(a)}\right\}u(a)\,v(a)$$

$$> 0 \qquad \ldots(6)$$

as $$\frac{u'(a)}{u(a)} \ge \frac{v'(a)}{v(a)},\ p(a) \ge 0,\ p(t_1) \ge 0,\ u'(t_1) < 0,\ v(t_1) > 0,\ u(a) > 0,\ v(a) > 0,$$

and

$$\int_a^{t_1} \{q_2(t) - q_1(t)\}\, u(t)\, v(t) \le 0. \qquad \ldots(7)$$

This is a contradiction, hence proof is complete.

Theorem. (19.9) Consider the differential equations

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} + q_1(t)\, u = 0, \qquad \ldots(1)$$

$$\frac{d}{dt}\left\{p(t)\frac{dv}{dt}\right\} + q_2(t)\, v = 0, \qquad \ldots(2)$$

where $p(t) > 0$, $q_2(t) > q_1(t)$ on $[a, b]$. Further, let either

(i) $$\frac{u'(a)}{u(a)} \ge \frac{v'(a)}{v(a)},\ u(a) \ne 0,\ v(a) \ne 0\ , \qquad \ldots(3)$$

or

(ii) $$u(a) = 0\ v(b) = 0. \qquad \ldots(4)$$

Suppose that $u(t)$ and $v(t)$ have the same number of zeros in $[a, b]$. Then, show that

$$\frac{u'(b)}{u(b)} > \frac{v'(b)}{v(b)},\ \text{if } u(b) \ne 0.$$

Proof. Since $u(b) \ne 0$ and $v(t)$ has as many zeros in $[a, b]$ as $u(t)$, it follows that

$$v(b) \ne 0, \qquad \ldots(5)$$

since, in this case, $t_n < b$ and $v(t)$ has atleast as many zeros as $u(t)$ in $[a, b]$ by virtue of comparison theorem.

Applying Green's identity t_0 the interval $[t_n, b]$, we have

$$\{p(t)\, [u'(t)\, v(t) - u(t)\, v'(t)]\}_{t_n}^{b} = \int_{t_n}^{b} [q_2(t) - q_1(t)]\, u(t)\, v(t)\ . \qquad \ldots(6)$$

W.l.o. g, we may assume that $u(t) > 0$, $v(t) > 0$ in $[t_n, b]$. Then

$$\int_{t_n}^{b} [q_2(t) - q_1(t)]\, u(t)\, v(t) > 0. \qquad \ldots(7)$$

So, using (6) and (7), we have

$$p(b)\, [u'(b)\, v(b) - u(b)\, v'(b)] > p(t_n)\, [u'(t_n)\, v(t_n)], \qquad (\because u(t_n) = 0)$$

$$> 0 \qquad (\because p(t_n) > 0,\ v(t_n) > 0\ u'(t_n) > 0)$$

$$\Rightarrow \quad u'(b)\, v(b) > u(b)\, v'(b)$$

$$\Rightarrow \quad \frac{u'(b)}{u(b)} > \frac{v'(b)}{v(b)}. \qquad [\because v(b) > 0,\ u(b) > 0]$$

This proves the result.

Theorem (19.10). Let the coefficient functions in differential equations

$$\frac{d}{dt}\left(p_1(t)\frac{du}{dt}\right) + q_1(t)\, u = 0, \qquad \ldots(1)$$

$$\frac{d}{dt}\left(p_2(t)\frac{du}{dt}\right) + q_2(t)\, u = 0, \qquad \ldots(2)$$

be continuous on the interval $I = [a, b]$ and let

$$p_1(t) \ge p_2(t) > 0 \text{ and } q_1(t) \le q_2(t). \qquad \ldots(3)$$

Let $u = u_1(t) \not\equiv 0$ be a solution of (1) and let $u_1(t)$ have exactly n (≥ 1) zeros $t = t_1 < t_2 < \ldots < t_n$ in $[a, b]$. Let $u = u_2(t) \not\equiv 0$ be a solution of (2) satisfying

$$\frac{p_1(a)u_1'(a)}{u_1(a)} \ge \frac{p_2(a)u_2'(a)}{u_2(a)}\ . \qquad \ldots(4)$$

Then $u_2(t)$ has atleast n zeros on $a < t \le t_n$. Furthermore, $u_2(t)$ has at least n zeros on $a < t < t_n$ if either the inequality (4) holds or equation (2) is a strict Sturm majorant for equation (1) on the interval $a \le t \le t_n$.

Proof. In view of (4), it is possible to define a pair of continuous function $\phi_1(t)$ and $\phi_2(t)$ on the interval I = [a, b] by

$$\phi_j(t) = \tan^{-1}\left(\frac{u_j(t)}{p_j(t)u_j^{©}(t)}\right) \qquad j = 1, 2, \qquad \text{...(5)}$$

with $$0 \le \phi_1(a) \le \phi_2(a) < \pi . \qquad \text{...(6)}$$

Then, the analogue of first order differential equation in Prufer transformation, is

$$\frac{d\varphi_j}{dt} = \frac{1}{p_j(t)} \cos^2 \phi_j(t) + q_j(t) \sin^2 \phi_j(t) \equiv f_j (t, \phi_j), \text{ say} . \qquad \text{...(7)}$$

Since the continuous functions f_j (t, ϕ_j) are smooth as functions of the variable ϕ_j, the solutions of two differential equations in (7) are uniquely determined by their initial conditions.
From inequalities in (3), it follows that

$$f_1(t, \phi) \le f_2(t, \phi) , \qquad \text{...(8)}$$

for $a \le t \le b$ and all ϕ. Also, the initial conditions satisfy the inequality in (6). So by comparison theorem for differential inequalities, it follows that

$$\phi_1(t) \le \phi_2(t) \text{ for } t\in [a, b] . \qquad \text{...(9)}$$

In particular, by hypothesis

$$\phi_1(t_n) = n\pi. \qquad \text{...(10)}$$

From (9) and (10), it follows that

$$n\,\pi \le \phi_2(t_n) \qquad \text{...(11)}$$

Hence it follows that $u_2(t)$ has atleast n zeros on the interval (a, t_n].

Proof of last part of the theorem.

Suppose first that the sign of inequality holds in (4). Then

$$\phi_1(a) < \phi_2(a) \qquad \text{...(12)}$$

Let $\phi_{20}(t)$ be the solution of second equation (for j = 2) in (7) satisfying the initial condition

$$\phi_{20}(a) = \phi_1(a), \qquad \text{...(13)}$$

so that

$$\phi_{20}(a) < \phi_2(a). \qquad \text{...(14)}$$

Since solutions of equation (7), for j = 2, are uniquely determined by initial conditions, so

$$\phi_{20}(t) < \phi_2(t) \text{ for } t\in \{a, b]. \qquad \text{...(15)}$$

From equation (9) and (15), we find

$$\phi_1(t) \le \phi_{20}(t) < \phi_2(t) \text{ for } t \in [a, b] , \qquad \text{...(16)}$$

and equations (10) and (16) imply

$$\phi_2(t_n) > n\pi . \qquad \text{...(17)}$$

Hence $u_2(t)$ has n zeros on the interval (a, t_n] .

Now, consider the case that equality holds in (4) but equation (2) is a strict Sturm majorant for equation (1) on the interval $a \le t \le t_n$. Then, by definition, either

$$q_1(t) < q_2(t) \qquad \text{...(18a)}$$

or

$$p_1(t) > p_2(t) > 0 \text{ and } q_2(t) \ne 0 \qquad \text{...(18b)}$$

holds at some point of [a, t_n].
Write equation (7) for j = 2, as

$$\phi_2' = \frac{1}{p_1}\cos^2\phi_2 + q_1\sin^2\phi_1 + \in(t) \qquad \ldots(19)$$

where

$$\in(t) = \left(\frac{1}{p_2} - \frac{1}{p_1}\right)\cos^2\phi_2 + (q_2 - q_1)\sin^2\phi_2\,. \qquad \ldots(20)$$

If the assertion/result is false, it follows from the case first considered that

$$\phi_1(t) = \phi_2(t) \text{ for } a \le t \le t_n. \qquad \ldots(21)$$

Hence,

$$\phi_1'(t) = \phi_2'(t)$$

and so

$$\in(t) = 0 \text{ for } a \le t \le t_n\,. \qquad \ldots(22)$$

Since

$$\sin\phi_2(t) = 0$$

only at the zeros of $u_2(t)$, it follows, from equation (20) and (22), that

$$q_2(t) = q_1(t) \text{ for } a \le t \le t_n. \qquad \ldots(23)$$

Consequently, from equation (20), (22) and (23), we have

$$\left(p_2^{-1} - p_1^{-1}\right)\cos^2\phi_2(t) = 0. \qquad \ldots(24)$$

Since $p_2^{-1}(t) - p_1^{-1}(t) > 0$ at some point t, it follows that

$$\cos\phi_2(t) = 0 \qquad \ldots(25)$$

$$\Rightarrow \quad u_2'(t) = 0. \qquad \ldots(26)$$

If (18a) does not hold at any t on $[a, t_n]$, it follows that (18b) holds at some point t and hence on some subinterval of $[a, t_n]$. But then

$$u_2'(t) = 0,$$

on this interval, thus

$$(p_2 u_2')' = 0,$$

on this interval. But this contradicts

$$q_2(t) \ne 0$$

on this interval. This contradiction proves the theorem.

Remark. The above discussed comparison theorems aim at comparing the distribution of the zero of the solution of the given differential equation with the distribution of the zeros of the solution of some "other" differential equation.

20

STURM-LIOUVILLE BOUNDARY VALUE PROBLEMS (SLBVP)

By a boundary problem in its general sense is meant the question as to whether a given differential equation possesses or does not possess solutions which satisfy certain boundary, or end-point, conditions.
SLBV problems represent a class of linear BVPs. These problems arise naturally, for instance, when separation of variables is applied to the wave equation, the potential equations, or the diffusion equation. The importance of these problems lies in the fact that they generate sets of orthogonal functions. The sets of orthogonal functions are useful in the expansion of a certain class of functions.

The Sturm-Liouville differential equation is

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + [q(t) + \lambda\, r(t)]\, u = 0\,, \qquad (1)$$

where p, p′, q, and r are continuous and

$$p(t) > 0\,,\ \ r(t) > 0\ \,, \qquad (2)$$

on the interval [a,b] and λ a real number. Equation (1) is equivalently written as

$$L\,[u(t)] = -\lambda\, u(t), \qquad (3)$$

where the Sturm-Liouville operator, L, is defined by

$$L = \frac{1}{r(t)}\left\{\frac{d}{dt}\left[p(t)\frac{d}{dt}\right] + q(t)\right\} \qquad (4)$$

The parametric λ is called an **eigenvalue** of the differential equation.

Given a specific set of boundary conditions, there may be specific values of λ for which equation (1) has a non-trivial solution called **eigenfunction**. For different types of boundary conditions, different types of behaviour are possible.

Note : The operator L, as defined above, is formally self-adjoint. The boundary conditions are of the form

$$\left.\begin{aligned} \alpha_1\, u(a) + \alpha_2 u'(a) = 0 \\ \beta_1\, u(b) + \beta_2\, u'(b) = 0 \end{aligned}\right\}, \qquad (5)$$

where α_1 and α_2 are constants (not both zero), and β_1 and β_2 are constants (not both zero).
The boundary conditions of the form

$$\left.\begin{aligned} u(a) = u(b) \\ u(a) = u'(b) \end{aligned}\right\}, \qquad (6)$$

are called **periodic**.

Example. Find non-trivial solutions of the SLBVP

$$\frac{d^2u}{dt^2} + \lambda u = 0,$$
$$u(0) = u(\pi) = 0. \tag{1}$$

Solution. λ is a real number. We consider separately the three cases : $\lambda = 0$, $\lambda < 0$ and $\lambda > 0$.

Case 1. $\lambda = 0$. In this case, a solution of given differential equation is

$$u(t) = c_1 t + c_2 \tag{2}$$

The boundary condition $u(0) = 0$ implies

$$c_2 = 0 . \tag{3}$$

The boundary condition $u(\pi) = 0$ implies

$$c_1\pi = 0$$

or $$c_1 = 0 . \tag{4}$$

This shows that, in this case, $\lambda = 0$ is not an eigenvalue as it corresponds to only trivial solution.

Case 2. When $\lambda < 0$. We write $\lambda = -\alpha^2$, $\alpha > 0$ (5)

In this case, solution of given equation is of the type

$$u(t) = c_1 e^{\alpha t} + c_2 e^{-\alpha t} \tag{6}$$

Now the boundary condition $u(0) = 0$ implies

$$c_2 = -c_1 , \tag{7}$$

and the boundary condition $u(\pi) = 0$ gives

$$c_1(e^{\alpha\pi} - e^{-\alpha\pi}) = 0$$

or $$c_1(e^{2\alpha\pi} -1) = 0 .$$

To have a non-trivial solution, c_1 should not be zero (otherwise, c_2 will also be zero), so that

$$e^{2\alpha\pi} = 1$$

or $$\alpha = 0$$

or $$\lambda = 0 , \tag{8}$$

which is not possible. Thus, no negative real number can be an eigen-value of the given SLBVP.

Case 3. When $\lambda > 0$. We write $\lambda = \alpha^2$, $\alpha > 0$. (9)

A solution of the given differential equation is

$$u(t) = c_1 \sin \alpha t + c_2\cos \alpha t$$

The boundary condition $u(0) = 0$ implies

$$c_2 = 0 , \tag{10}$$

and the condition $u(\pi) = 0$ gives

$$c_1\sin \alpha\pi = 0$$

or $\sin \alpha\pi = 0$ (otherwise, if $c_1 = 0$, no non-trivial solution exist)

or $$\alpha = \pm n$$

or $$\lambda = n^2 , \quad n = 1, 2, 3,\ldots \tag{11}$$

Thus, the eigenvalues of the given SLBVP are given by (11). The corresponding eigenfunction, upto a multiplicative constant, is

$$u(t) = \sin nt , \tag{12}$$

for n = 1, 2,3,…

This completes the solution.

Exercise. Find the eigenvalues and eighenfunctions of each of the following SLBVP's.

(i) $\dfrac{d^2u}{dt^2} + \lambda u = 0$, $u(0) = u(\pi/2) = 0$.

(ii) $\dfrac{d^2u}{dt^2} + \lambda u = 0$, $u(0) = 0, u(L) = 0$, $L > 0$.

(iii) $\dfrac{d}{dt}\left[t\dfrac{du}{dt}\right] + \dfrac{\lambda}{t}u = 0$, $u(1) = 0$, $u(e^{\pi}) = 0$

(iv) $\dfrac{d}{dt}\left[(t^2+1)\dfrac{du}{dt}\right] + \dfrac{\lambda}{t^2+1}u = 0$, $u(0) = 0 = u(1)$.

Theorem 20.1. Prove that the eigenvalues of a SLBVP are discrete.

Proof. Let $u_1(t; \lambda)$ and $u_2(t; \lambda)$ be two linearly independent solutions (for fixed λ) of a SLBV problem consisting of a differential equation.

$$\frac{d}{dt}\left\{p(t)\frac{du}{dt}\right\} + \{q(t) + \lambda r(t)]u = 0 , \tag{1}$$

and the boundary conditions

$$u(a) = u(b) = 0 . \tag{2}$$

Then any solution of differential equation (1) can be expressed as

$$u(t; \lambda) = Au_1(t; \lambda) + Bu_2(t; \lambda) , \tag{3}$$

the linear combination of u_1 and u_2 . The constants A and B are determined by requiring $u(t; \lambda)$, in (3), to satisfy the boundary conditions in (2). Application of these boundary conditions leads to

$$\begin{aligned} A\, u_1(a; \lambda) + B\, u_2(a; \lambda) &= 0 \\ A\, u_1(b; \lambda) + B\, u_2(b ; \lambda) &= 0 , \end{aligned} \tag{4}$$

or a matrix equation

$$\begin{pmatrix} u_1(a;\lambda) & u_2(a;\lambda) \\ u_1(b;\lambda) & u_2(b;\lambda) \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} . \tag{5}$$

Thus, equations (4) can be solved for A and B only if the determinant of the matrix of coefficients in (5) vanishes. Otherwise, the only solution is $A = B = 0$, which yields the trivial solution $u(t) = 0$. The condition for a non-trivial solution of (4) is therefore given by

$$\begin{vmatrix} u_1(a;\lambda) & u_2(a;\lambda) \\ u_1(b;\lambda) & u_2(b;\lambda) \end{vmatrix} = u_1(a,\lambda)\, u_2(b; \lambda) - u_1(b; \lambda)\, u_2(a; \lambda) = 0 . \tag{6}$$

If we consider $u_1(t; \lambda)$ and $u_2(t; \lambda)$ to be analytic functions of λ, then the determinant itself is an analytic function of λ . Therefore, the zeros of the determinant, by the theory of complex-valued functions, must be isolated (how).

Since the zeros of the determinant correspond to allow solutions of the SLBVP, we conclude that the eigenvalues of (1) and (2) are discrete.

This completes the proof.

Definition 1. Two functions u(t) and v(t) are said to be **orthogonal** w.r.t. a weight function w(t) on the interval $a \le t \le b$ if

$$\int_a^b w(t)\, u(t)v(t) = 0 .$$

Definition 2. Let $\{\phi_n(t)\}$ be a sequence of functions on $a \le t \le b$. Then these functions are called **mutually orthogonal** w.r.t. a weight function w(t) on $a \le t \le b$ if

$$\int_a^b w(t)\,\phi_n(t)\,\phi_m(t)\,dt = 0 \quad , \quad \text{for } n \neq m \,.$$

Illustration (1). Two functions $f(t) = \sin t$ and $g(t) = \sin 5t$ are orthogonal w.r.t the weight function $w(t) = 1$ on the interval $0 \le t \le \pi$, for

$$\int_0^\pi \sin t \;\sin 5t\, dt = \frac{1}{2}\int_0^\pi \{\cos 4t - \cos 6t\}dt$$

$$= \frac{1}{2}\left[\frac{\sin 4t}{4} - \frac{\sin 6t}{4}\right]_0^\pi$$

$$= 0 \;.$$

(2) Consider the sequence $\{\phi_n\}$ of functions, where

$$\phi(t) = \sin nt \,, \quad n = 1, 2, 3, \ldots$$

on the interval $0 \le t \le \pi$. The set $\{\phi_n\}$ is an orthogonal system w.r.t. the weight function $w(t) = 1$ on the interval $0 \le t \le \pi$, for $m \neq n$,

$$\int_0^\pi \sin mt \;\sin nt\, dx = \frac{1}{2}\left[\frac{\sin(m-n)t}{m-n} - \frac{\sin(m+n)t}{m+n}\right]_{t=0}^{t=\pi}$$

$$= 0 \,.$$

Theorem 20.2. (**Orthogonality of characteristic functions**) :

(1) Consider the Sturm-Liouville problem consisting of the differential equation

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + [q(t) + \lambda\, r(t)]u = 0 \quad , \tag{1}$$

where p, q, r are real functions such that p has a continuous derivative, q and r are continuous, and $p(t) > 0$ and $r(t) > 0$ for all $t \in [a,b]$; and λ is a parameter independent of t; and

(II) the conditions $A_1 u(a) + A_2 u'(a) = 0$

$$B_1\, u(b) + B_2\, u'(b) = 0 \,, \tag{2}$$

where A_1, A_2, B_1, B_2 are real constants such that A_1 and A_2 are not both zero and B_1 and B_2 are not both zero.

(III) Let λ_m and λ_n be any two distinct characteristic values of this problem. Let ϕ_m be a characteristic function corresponding to λ_m and let ϕ_n be a characteristic function corresponding to λ_n .

Prove that the characteristic functions ϕ_m and ϕ_n are orthogonal with respect to the weight function $r(t)$ on the interval $a \le t \le b$.

Proof. Since ϕ_m is a characteristic function corresponding to λ_m, the function ϕ_m satisfies the differential equation (1) with $\lambda = \lambda_m$. Thus

$$\frac{d}{dt}\left[p(t)\phi_m'(t)\right] + [q(t) + \lambda_m\, r(t)]\phi_m(t) = 0 \,. \tag{3}$$

Similarly

$$\frac{d}{dt}\left[p(t)\phi_n'(t)\right] + [q(t) + \lambda_n r(t)]\phi_n(t) = 0 \,. \tag{4}$$

Multiplying both sides of (3) by $\phi_n(t)$ and both sides of (4) by $\phi_m(t)$ and then subtracting the results, we obtain

$$\phi_n(t)\,\frac{d}{dt}\left[p(t)\phi'_m(t)\right] - \phi_m(t)\,\frac{d}{dt}\left[p(t)\phi'_n(t)\right]$$
$$= (\lambda_n - \lambda_m)\, r(t)\, \phi_n(t)\, \phi_m(t)\,. \tag{5}$$

This implies

$$\frac{d}{dt}\left[p(t)\{\phi'_m(t)\phi_n(t) - \phi_m(t)\phi'_n(t)\}\right] = (\lambda_n - \lambda_m)\, r(t)\, \phi_n(t)\phi_m(t)\,. \tag{6}$$

We now integrate both sides of this identity from a to b to obtain

$$(\lambda_n - \lambda_m)\int_a^b r(t)\,\phi_n(t)\,\phi_m(t)dt = p(b)[\phi_n(b)\phi'_m(b) - \phi_m(b)\phi'_n(b)]$$
$$- p(a)[\phi_n(a)\phi'_m(a) - \phi_m(a)\phi'_n(a)]\,. \tag{7}$$

Since ϕ_m and ϕ_n are characteristic functions of the problem, they satisfy the supplementary conditions of the problem. That is,

$$\left.\begin{aligned} A_1\phi_m(a) + A_2\,\phi'_m(a) = 0 \\ B_1\phi_m(b) + B_2\phi'_m(b) = 0 \end{aligned}\right\}, \tag{8}$$

and

$$\left.\begin{aligned} A_1\phi_n(a) + A_2\phi'_n(a) = 0 \\ B_1\phi_n(b) + B_2\phi'_n(b) = 0 \end{aligned}\right\}. \tag{9}$$

Now we discuss all the possible cases :

Case 1. If $A_2 = 0$, $B_2 = 0$. Then $A_1 \neq 0$ and $B_1 \neq 0$, since both of A_1 and A_2 are not zero and both of B_1 and B_2 are not zero. Conditions (8) and (9) yields

$$\phi_m(a) = 0\,,\ \ \phi_m(b) = 0\,,\ \ \phi_n(a) = 0\,,\ \ \phi_n(b) = 0\ \ . \tag{10}$$

From equations (7) and (10), we see that

$$(\lambda_n - \lambda_m)\int_a^b r(t)\,\phi_n(t)\phi_m(t)dt = 0\ \ . \tag{11}$$

Case II. If $A_2 = 0$ but $B_2 \neq 0$. Then $A_1 \neq 0$. Let

$$\frac{B_1}{B_2} = \alpha\,. \tag{12}$$

Then conditions (8) and (9) gives

$$\phi_m(a) = 0\,,\ \ \phi_n(a) = 0\ \ . \tag{13}$$

and

$$\alpha\,\phi_m(b) + \phi'_m(b) = 0\,,\ \ \alpha\phi_n(b) + \phi'(b) = 0\ \ . \tag{14}$$

Using results (13) and (14), the right hand side of (7) takes the form

$$p(b)[\phi_n(b)\,\{-\alpha\phi_m(b)\} - \phi_m(b)\,\{-\alpha\phi_n(b)\}]$$
$$= p(b)\,[-\alpha\,\phi_n(b)\phi_m(b) + \alpha\phi_m(b)\phi_n(b)]$$
$$= 0\ \ ,$$

Hence

$$(\lambda_n - \lambda_m)\int_a^b r(t)\,\phi_n(t)\,\phi_m(t)\,dt\ = 0\,. \tag{15}$$

Case III. If $A_2 \neq 0$, $B_2 = 0$. This case is similar to Case II.

Case IV. If $A_2 \neq 0$, $B_2 \neq 0$. Let

$$\alpha_1 = A_1/A_2\,,$$
$$\alpha_2 = B_1/B_2\,. \tag{16}$$

Then conditions (8) and (9) takes the form

$$\alpha_1\phi_m(a) + \phi'_m(a) = 0\ \ ,$$

$$\alpha_2\phi_m(b) + \phi'_m(b) = 0 \ ,$$
$$\alpha_1\phi_n(a) + \phi'_n(a) = 0 \ ,$$
$$\alpha_2\phi_n(b) + \phi'_n(b) = 0 \ . \qquad (17)$$

Using (17) the right hand side of (7) becomes

$$p(b)[\phi_n(b)\ \{-\alpha_2\phi_m(b)\} - \phi_m(b)\ \{-\alpha_2\phi_n(b)\}]$$
$$-p(a)[\phi_n(a)\{-\alpha_1\phi_m(a)\}-\phi_m(a)\{-\alpha_1\phi_n(a)\}]$$
$$= 0 \ .$$

Hence

$$(\lambda_n - \lambda_m) \int_a^b r(t)\,\phi_n(t)\phi_m(t)dt = 0 \ . \qquad (18)$$

Thus in all possible cases, we conclude that

$$(\lambda_n - \lambda_m) \int_a^b r(t)\,\phi_n(t)\ \phi_m(t)dt = 0 \ . \qquad (19)$$

Since λ_m and λ_n are distinct eigenvalues, therefore,

$$\int_a^b r(t)\,\phi_n(t)\ \phi_m(t)\ dt = 0 \ . \qquad (20)$$

This proves that eigenfunctions $\phi_m(t)$ and $\phi_n(t)$ are orthogonal w.r.t. the weight function r(t) on [a,b].

Illustration. Consider the already discussed SL-problem

$$\frac{d^2u}{dt^2} + \lambda u = 0 \qquad (1)$$
$$u(0) = 0 \ , \ u(\pi) = 0 \ . \qquad (2)$$

We know that corresponding to each eigenvalue $\lambda_n = n^2$, this is an eigenfunction c_n sin nt, c_n being an arbitrary non zero constant. Let

$$\phi_n(t) = \sin nt \ . \qquad (3)$$

Then $\{\phi_n\}$ is the sequence of eigenfunctions.

Then by the theorem (20.2), the infinite set $\{\phi_n\}$ is an orthogonal system with respect to the weight function r(t), where

$$r(t) = 1 \quad \text{for all } t\ \varepsilon\ [0,\pi] \ . \qquad (4)$$

That is

$$\int_0^\pi \sin t \sin mt\ dt = 0 \quad \text{for } n \neq m,\ \ n, m\ \varepsilon\ N \ . \qquad (5)$$

Theorem 20.3. Prove that the eigenvalues of a SLBV problem are real.

Proof. Let λ_n be an eigenvalue corresponding to the eigenfunction $\phi_n(t)$ of the given SL-BV problem. Then, by definition,

$$\frac{d}{dt}\left[p(t)\frac{d\phi_n}{dt}\right] + [q(t) + \lambda_n\ r(t)]\ \phi_n(t) = 0 \ , \qquad (1)$$

and

$$A_1\phi_n(a) + A_2\phi'_n(a) = 0 \ ,$$
$$B_1\phi_n(b) + B_2\phi'_n(b) = 0 \ . \qquad (2)$$

We know that p(t), q(t) and r(t) are real valued functions of t over the interval [a,b]. So, taking the complex conjugate of equations (1) and (2), we obtain

$$\frac{d}{dt}\left[p(t)\frac{d\phi_n}{dt}\right] + [q(t) + \bar{\lambda}_n\, r(t)]\;\bar{\phi}_n = 0 \;, \tag{3}$$

and

$$A_1\;\bar{\phi}_n(a) + A_2\;\bar{\phi}'_n(0) = 0 \;,$$
$$B_1\;\bar{\phi}_n(b) + B_2\;\bar{\phi}'_n(b) = 0 \;, \tag{4}$$

since A_1, B_1, A_2, B_2 are real constants.

This shows that $\bar{\phi}_n$ is also an eigenfunction, corresponding to an eigenvalue $\bar{\lambda}_n$ of the same SLBVP. So, from theorem (20.2), it follows that

$$(\lambda_n - \bar{\lambda}_n)\int_a^b r(t)\,\phi_n(t)\;\bar{\phi}_n(t)\,dt = 0 \;,$$

or

$$(\lambda_n - \bar{\lambda}_n)\int_a^b r(t)\,|\phi_n(t)|^2\,dt = 0 \;. \tag{5}$$

Since $r(t) > 0$, and $|\phi_n(t)| \neq 0$, being a non-trivial solution, so we must have,

$$\lambda_n - \bar{\lambda}_n = 0 \;,$$

or

$$\lambda_n = \bar{\lambda}_n \;. \tag{6}$$

This shows that the eigenvalues are real.
This completes the proof.

Theorem 20.4. Consider the SLBVP consisting of differential equation

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] + [q(t) + \lambda r(t)]\;u = 0 \;, \tag{1}$$

for $t \in [a,b]$ and boundary conditions

$$u(a) = 0 \;,\; u(b) = 0 \;. \tag{2}$$

Prove that there exists an infinite sequence of eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n \ldots$, of this problem with the properties

$$0 < \lambda_1 < \lambda_2 < \ldots. < \lambda_n < \ldots.$$

and

$$\lambda_n \to +\infty \;,$$

and the eigenfunction $\phi_n(t)$, corresponding to eigenvalue λ_n, has precisely n zeros on the interval [a, b].

Proof. By hypothesis of a SLBVP, functions $p(t)$, $p'(t)$, $q(t)$ and $r(t)$ are continuous on [a,b] and

$$p(t) > 0 \;,\; r(t) > 0 \;,\;\; t \in [a, b]. \tag{3}$$

λ is a parameter which is independent of t. Now, we extend the domain of definition of functions $p(t)$, $q(t)$, and $r(t)$ from [a,b] to $[a, \infty]$, by defining

$$p(t) = p(b),$$
$$q(t) = q(b) \;,$$
$$r(t) = r(b) \;, \tag{4}$$

for all $t > b$.

Now, on the extended interval $a \le t < \infty$, we have $p(t) > 0$ and $r(t) > 0$, and $p(t)$, $p'(t)$, $q(t)$, $r(t)$ are continuous. We transform the given SLBVP by the substitution

$$\xi = \int_a^t \frac{ds}{p(s)} \;, \tag{5}$$

which changes the independent variable t to a new independent variable ξ . Equation (5) gives

$$\frac{d\xi}{dt} = \frac{1}{p(t)} \ . \tag{6}$$

Now

$$\frac{du}{dt} = \frac{du}{d\xi} \cdot \frac{d\xi}{dt}$$

gives

$$p(t)\ \frac{du}{dt} = \frac{du}{d\xi} \ , \tag{7}$$

and

$$\frac{d}{dt}\left[p(t)\frac{du}{dt}\right] = \left(\frac{d^2u}{d\xi^2}\right)\frac{1}{p(t)} \ . \tag{8}$$

Accordingly, equation (1) becomes

$$\frac{d^2u}{d\xi^2} + p_1(\xi)\,[q_1(\xi) + \lambda r_1(\xi)]\ u(\xi) = 0 \ , \tag{9}$$

where

$$\begin{aligned} p_1(\xi) &= p(t) \ , \\ q_1(\xi) &= q(t) \ , \\ r_1(\xi) &= r(t) \ . \end{aligned} \tag{10}$$

The boundary conditions (2) now becomes

$$\begin{aligned} u(\xi = 0) &= 0 \ , \\ u(\xi = c) &= 0 \ . \end{aligned} \tag{11}$$

Here

$$c = \int_a^b \frac{ds}{p(s)} \ . \tag{12}$$

Now, ξ increases steadily from 0 to ∞ as t increases from a to ∞ . Further

$$p_1(\xi) > 0 \ , \quad r_1(\xi) > 0 \ , \tag{13}$$

using (5). Moreover $p_1(\xi)$, $q_1(\xi)$ and $r_1(\xi)$ are continuous on the interval $0 \le \xi < \infty$.

For λ fixed but arbitrary, let $u(\xi, \lambda) \not\equiv 0$ be a solution of differential equation (9) with the property that

$$u(0, \lambda) = 0 \ . \tag{14}$$

It is clear that for λ negative and sufficiently large numerically

$$u(\xi, \lambda) \neq 0 \quad \text{on } 0 < \xi \le c \ .$$

Indeed, λ may be chosen so that

$$p_1(\xi)\ [q_1(\xi) + \lambda r_1(\xi)] < 0 \ , \ \text{on } 0 < t \le c \ .$$

Furthermore, it is clear that for λ positive and sufficiently large, the solution $u(\xi, \lambda)$ will have a first zero following $\xi = 0$.

Now from Sturm comparison theorem, it follows that the zeros $\alpha_n(\lambda)$, of $u(\xi, \lambda)$, decreases steadily as λ increases and that $\alpha_n(\lambda) \to 0$ as $\lambda \to \infty$. Then, let λ be chosen so that the first positive zero $\alpha_1(\lambda)$ lies to the right of $\xi = c$.

As λ increases, α_1 decreases and for precisely one value of λ, say λ_1, we have $\alpha_1 = c$. The number λ_1 is the first characteristic number of the system consisting of equations (9) and (11). The corresponding solution $u_1(\xi, \lambda_1)$ is the first characteristic function.

As λ continues to increase, the zero α_1 zeros on to the interval $0 < \xi < c$ and $\alpha_2(\lambda)$, the second positive zero of $u(\xi, \lambda)$ moves towards $\xi = c$ and coincides with c for

$$\lambda = \lambda_2 > \lambda_1 . \qquad (15)$$

Accordingly, λ_2 is the second characteristic number and $u_2(\xi,\lambda_2)$ is the corresponding characteristic function.

Continuing this process, we get an infinite sequence of characteristic numbers

$$\lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots \text{ to } \infty \qquad (16)$$

and a corresponding infinite sequence of characteristic functions

$$u_1(\xi, \lambda_1) , u_2(\xi, \lambda_2) ,\ldots, u_n(\xi, \lambda_n) , \ldots$$

for the system (9) and (11).

Using the transformation (5), we get the desired result.

Note : The readers are advised the following books for reading (for chapters 17-20).

1. Ross S.L. Differential Equations.

2. Ince, E.L. Ordinary Differential Equations.

3. Birkhoff G. and Rota, G.C. Ordinary Differential Equations.

4. Hertman, P. Differential Equations.