1

BASIC THEORY

Unit Structure :

1.1 INTRODUCTION

As we already know, a differential equation - DE - is an equation relating the following three items :

- A function of one or more variables (the function being real valued or vector values).
- The independent variables of the function.
- A finite number of derivatives of the function.

The highest order of derivatives of the function appearing in the equation is the **order** of the differential equation.

Usually, the function in a differential equation is an unknown function; it is an are observable quantity of a real process and therefore we are interested in knowing the function. We use results and techniques of mathematical analysis along with our geometric intuition and tease the function out of the differential equation. We then speak of having solved the differential equation.

Depending on the nature of the function (in which a differential equation is set) we classify the differential equations in the following two types :

I) A function $X : I \to \mathbb{R}^n$ of a single real variable (say) t ranging in an open interval I gives rise to the succession of derivatives :

$$X = \left(\frac{d}{dt}\right)^0 X, \frac{dX}{dt}, \frac{d^2X}{dt^2}, \ldots \frac{d^kX}{dt^k} \ldots\ldots$$

Now a differential equation in the function $X(t)$ is therefore an equation of the type :

$$F\left(t, X \frac{dX}{dt}, \frac{d^2X}{dt^2}, ... \frac{d^kX}{dt^k}\right) = 0 \text{ .... (1)}$$

Such an equation is said to be an ordinary differential equation. Thus an **ordinary differential equation** is a differential equation in which the constituent function $X : t \mapsto X(t)$ is a function of a single real variable.

We often use the acronym ODE in place of the full term : ordinary differential equation.

II) On the other hand, there are differential equations in a function $u : x \mapsto u \;\; x$ of a real multivariable $x = \;\; x_1\, x_2 ..... x_n$ which ranges in an open subset Ω of $\mathbb{R}^n$. Such a function $u = u \;\; x$ gives rise to mixed partial derivatives :

$$\frac{\partial u}{\partial x_i}, 1 \leq i \leq n,$$

$$\frac{\overset{2}{\partial u}}{\partial x_i \partial x_j}\, j\, 1 \leq i, j \leq n$$

$$\vdots$$

$$D^\alpha u := \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} .... \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} u$$

for various multi-indices $\alpha = \;\; \alpha_1\, \alpha_2 .... \alpha_n$ with $\alpha_i \in \mathbb{Z}^+\, i = \;\; 0, 1, 2, ...$, the mixed partial derivative $D^\alpha u$ having the order $|\alpha| := \alpha_1 + \alpha_2 + ... + \alpha_n$.

Now a differential equation in such a function $u = u \;\; x$ of a multi - variable $x = \;\; x_1 ..... x_n$ ranging in an open subset Ω of $\mathbb{R}^n$ is an equation of the type :

$$F\left(x, u, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i\, \partial x_j} .... D^\alpha u : |\alpha| \leq m\right) = 0 \text{ .... (2)}$$

its order being m. Equation (2) is said to be a **partial differential equation** in $u(x)$ because it involves the mixed partial derivatives of u. We use the acronym PDE for this type of differential equations.

There is more about the setting of a differential equation : In a mathematical problem, a differential equation is accompanied by auxiliary data. A solution of a differential equation is required to satisfy this auxiliary data. To be more specific we are given a subset of the domain of a prospective solution and some of its derivatives of the solution at the points of this subset.

In case of the ODE, the auxiliary data is said to consist of **initial conditions**. An **initial value problem** consists of finding the solution of the ODE which satisfies the accompanying initial conditions. Often, the pair consisting of (a) an ordinary differential equation and (b) the initial conditions is referred as the initial value problem -IVP-.

Often the initial conditions are given at the end points of an interval which are then called the boundary conditions. Also, the resulting initial value problem is called a **boundary value problem**.

In the case of a partial differential equation, the accompanying auxiliary data is called the **Cauchy data** for PDE. The **Cauchy problem** for a given PDE consists of finding the solution of the PDE which satisfies the requirements of the given Cauchy data.

We will explain more about these terms initial conditions, Cauchy data etc. - at later stages.

Partial differential equations, being more intricate mathematical objects are studied by using the concepts and results of the ordinary differential equations. Therefore, a basic course on differential equations begins with a treatment of ordinary differential equations. In our treatment of the subject also, we will develop enough theory of ODE and then apply it to the partial differential equations.

Therefore, back to the theory of ODE.

1.2 ORDINARY DIFFERENTIAL EQUATIONS

To begin with, we reorganize the form (1) of the ODE in the following manner. Unraveling it, we separate the top order derivative and express it as a function of the remain variable quantities, namely, $t, X(t), \frac{dX}{dt}, \ldots.. \frac{d^{k-1}X}{dt^{k-1}}$, that is we form the equation

$$\frac{d^k X}{dt^k} = f\left(t, X, \frac{dX}{dt}, \frac{d^{k-1}x}{dt^{k-1}}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (3)$$

We regard the equation (3) as the standard form of an ODE (of course, the ODE has order = k.) Note that the function $X : t \mapsto X(t)$ is a vector valued function of the real variable t and as such it is a curve in $\mathbb{R}^n$. Each $X(t)$ has n components : $X(t) = X_1(t), X_2(t), \ldots., X_n(t)$ and therefore all the derivative of it has *n* components :

$$\frac{d^\ell X(t)}{dt^\ell} = \left(\left(\frac{d}{dt}\right)^\ell X_1(t), \left(\frac{d}{dt}\right)^\ell X_2(t), \ldots.. \left(\frac{d}{dt}\right)^\ell X_n(t)\right) \quad \text{for} \quad 1 \le \ell \le k.$$

Consequently, the function *f* appearing on the right hand side of (3) has *n* components : $f = f_1, f_2, \ldots, f_n$ each f_i being a real valued function.

Consequently the DE (3) is actually the following system of ODE in the functions :

$$t \mapsto X_1(t), t \mapsto X_2(t), ..., t \mapsto X_n(t)$$

$$\left.\begin{array}{l} \dfrac{d^k X_1}{dt^k} = f_1\left(t, X, \dfrac{dX}{dt}, ..., \dfrac{d^{k-1}X}{dt^{k-1}}\right) \\ \dfrac{d^k X_2}{dt^k} = f_2\left(t, X, \dfrac{dX}{dt}, ..., \dfrac{d^{k-1}X}{dt^{k-1}}\right) \\ \quad\vdots \\ \dfrac{d^k X_n}{dt^k} = f_n\left(t, X, \dfrac{dX}{dt}, ..., \dfrac{d^{k-1}X}{dt^{k-1}}\right) \end{array}\right\} \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

At this stage, we become more specific about the features of the ODE (3) (or equivalently about the system (4).)

Let I be an open internal and let Ω denote an open subset of $\mathbb{R}^n$. We consider the open sets

$$I \times \Omega \times \mathbb{R}^n \times ...... \times \mathbb{R}^n$$

(there being $k-1$ copies of $\mathbb{R}^n$ in the above Cartesian product). This set is being designated to accommodate the variable quantities :

$$t, X(t), \frac{dx}{dt}(t), \frac{d^{k-1}x}{dt^{k-1}}.$$

Clearly, the function *f* appearing on the right hand side of (3) must have this set as its domain of definition.

Choosing $t_0 \in I, x_0 \in \Omega$ and $w_1, w_2 ... w_{k-1}$ in $\mathbb{R}^n$, we form the initial condition $(t_0, x_0, w_1, w_2 ... w_{k-1})$. Now, the initial value problem for the ODE (3) is the following pair :

$$\left\{\begin{array}{c} \dfrac{d^k X}{dt^k} = f\left(t, X, \dfrac{dX}{dt}, ... \dfrac{d^{k-1}X}{dt^{k-1}}\right) \\ (t_0, x_0, w_1, ... w_{k-1}) \end{array}\right\} \ldots\ldots\ldots\ldots\ldots (5)$$

By a solution of the initial value problem (5) we mean an (at least) k times continuously differentiable function (= curve in Ω)

$X : J \to \Omega$, J being an open interval with $t_0 \in J < I$, which satisfies the following two items :

- The differential equation : $\frac{d^k X(t)}{dt^k} = f\left(t\,\frac{dX(t)}{dt}\ldots\frac{d^{k-1}X(t)}{dt^{k-1}}\right)$ for all $t \in J$
- The initial conditions :
  $x(t_0) = x_0, \frac{dX}{dt}(t_0) = w_1, \ldots., \frac{d^{k-1}X(t_0)}{dt^{k-1}} = w_{k-1}$

Remarks :

(I) Though the independent variable t of the function *X(t)* in the ODE (1) is stipulated to range in the interval I, we expect the solution $t \mapsto X(t)$ of the initial value problems (4) to be defined only on a sub-internal J of I (with $t_0 \in J$). Indeed, we come across concrete cases of the IVP in which a solution exists only on a sub-inter J of I and therefore, we grant this concession : a solution need be defined only on a sub-interval J of I.

II) Often, the initial conditions are expressed more explicitly in terms of the equations : $x(t_0) = x_0, \frac{dX}{dt}(t_0) = w_1, \ldots., \frac{d^{k-1}X(t_0)}{dt^{k-1}} = w_{k-1}$.

An important special case of the ODE (3) is $\frac{d^k X}{dt^k} = f\left(X, \frac{dX}{dt}, \frac{d^2X}{dt^2}, \frac{d^{k-1}X}{dt^{k-1}}\right)$ ………………… (6)

in which the function f is independent of the variable $t : f : \Omega \times \mathbb{R}^n \times \ldots. \times \mathbb{R}^n \to \mathbb{R}^n$. In this case we say that the ODE (6) is autonomous.

Returning to the initial value problem (5), there are two questions :

- Does the initial value problem (5) admit a solution at all?
- If it does is the solution unique?

Clearly, because *f* is the main ingredient of the ODE, the answer to both of there questions naturally depends on the properties of *f*, especially on its behaviour around the initial condition $t_0, x_0, w_1, \ldots\ldots. w_{k-1}$ (i.e. how if varies continuously, differentially etc. around $t_0, x_0, w_1, \ldots., w_{k-1}$ Following two examples illustrate that answers to both the questions are (in general) in the negative :

Let $f : \mathbb{R} \to \mathbb{R}$ be the function :

$$f(x) = -1 \quad x < 0$$
$$= +1 \quad x \geq 0$$

For this function we consider the (first order autonomous case of) the initial value problem : $\frac{dX}{dt} = f(X),\ x(o) = 0$.

We contend that this initial value problem has no solution.

For, if there was a solution $X : J \to \mathbb{R}$ with $0 \in J$, then $\frac{dX}{dt}(0) = f(0) = 1 > 0$ implies that the solution $t \mapsto X(t)$ is strictly monotonic increasing in a neighborhood $-\delta, \delta$ of O.

On the other hand, $\frac{dX}{dt} = -1\,(< 0) = f(t)$ for all $t \in -\delta, 0$ implies that $t \mapsto X(t) = -t$ is strictly monotonic decreasing in $-\delta, 0$. Thus, if the solution of the above IVP exists then it is strictly monotonic increasing as well as strictly monotonic decreasing. This prevents a solution!

- Now let $g : \mathbb{R} \to \mathbb{R}$ be the function

$$g\ x = 0 \qquad x \le 0$$
$$= x^{\frac{1}{3}} \qquad x \ge 0$$

For this *g*, we consider the autonomous initial value problem.

$$\frac{dX}{dt} = g\ X\ ,\ \ X\ 0\ = 0$$

Clearly, one solution of it is the function

$$X_1(t) = \begin{cases} \left(\frac{2}{3}\right)^{3/2} t^{3/2} & t \ge 0 \\ 0 & t \le 0 \end{cases}$$

Another solution of the same IVP is the function $X_0(t) \equiv 0$. Thus, the IVP has at least two distinct solutions. (In fact it has an infinitude of solutions. For each c > 0, the functions : $t \mapsto c\,X_1(t)$ is a solutions)

Of course, an IVP should admit a unique solution! In the following sections we will concentrate our attention on first order ODE and for such ODE we will introduce a condition - *f* being locally Lipschitz - which will guarantees both - existence and uniqueness of the solution.

Above, we have been discussing ODE of arbitrary order $n \in \mathbb{N} := 1, 2, 3, ...$ and the IVP associated with them. But there is a simplifying aspect of the ODE! Higher order ODE can be studied entirely in terms of first order ODE. (This point will be explained in detail in the last part of chapter 2). Therefore, for the time-being we will focus our interest on first order ODE only.

1.3 FIRST ORDER ODE

We begin some more generalities related to first order ODE. As in the preceding part, I denoted an open interval and Ω, an open subset of $\mathbb{R}^n$.

We consider a function :

$f: I \times \Omega \to \mathbb{R}^n$ ……………….. (7)

giving rise to the first order ODE :

$\dfrac{dX}{dt} = f(t, X)$ ……………………(8)

Note that for each $t \in I$ held fixed the map :

$f\ t, -\ : \Omega \to \mathbb{R}^n; x \mapsto f\ t, x$ is a vector field on Ω. Interpreting "t" as the time variable we call the function (7) a time dependant vector field on Ω. And, often, we call a solution $X : J \to \Omega$ of the ODE (8), an **integral curve** of the vector field *f*.

An initial condition for the ODE (8) consists of a pair t_0, x_0 with $t_0 \in I$, $x_0 \in \Omega$ and the associated initial value problem is :

$\dfrac{dX}{dt} = f(t, X), \quad x\ t_0\ = x_0$ ……………….. (9).

Finally, recall that a solution of (9) is an (at least) once continuously differentiable curve

$X : J \to \Omega$

(J being an open interval with $t_0 \in J\ CI$) satisfying :

$\dfrac{dX(t)}{dt} = f(t, X(t))$ for all $t \in I$ and the initial condition $X\ t_0\ = x_0$.

In the context of the IVP (9) we consider yet another equation the following integral equation in an unknown function $X : J \to \Omega$:

$X(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds \qquad t \in J$ ………………… (10).

Following result relates solutions of the IVP (9) and those of the integral equation (10) :

<u>Proposition 1</u> :

A continuously differentiable curve $X : J \to \Omega$

(J being an open subinterval of I with $t_0 \in J$) is a solution of the IVP (9) if and only if it satisfies the integral equation (10).

<u>Proof : (I)</u> - First suppose that the curve $X : J \to \Omega$ satisfies the integral equation (10). Putting $t = t_0$ in (10) we get :

$$\begin{aligned} X\ \ t_0\ &= x_0 + \int_{t_0}^{t_0} f(s, X(s))\, ds \\ &= x_0 + 0 \\ &= x_0 \end{aligned}$$

Thus X satisfies the initial condition.

Next, differentiating (10) we get

$$\begin{aligned} \frac{dX\ \ t}{dt} &= 0 + \frac{d}{dt}\int_{t_0}^{t} f(s, X(s))\, ds \\ &= f(t, X(t)) \end{aligned}$$

by fundamental theorem of integral calculus.

Above, we have verified that a solution $t \mapsto x\ \ t$ of the integral equation (1) is also a solution of the IVP (9). Conversely suppose, $t \mapsto x\ \ t$, $t \in J$ is a solution of the IVP (9). Integrating the identify.

$$\frac{dx}{dt}\ \ t\ = f(t\, x(t)\, t \in J$$

We get :

$$\begin{aligned} x\ \ t\ - x\ \ t_0\ &= \int_{t_0}^{t} \frac{d}{ds} X\ \ s\ \ ds \\ &= \int_{t_0}^{t} f(r, x\ \ s\ \ ds \end{aligned}$$

And therefore :

$$x\ \ t\ - x_0 = \int_{t_0}^{t} f(s, X\ \ s\ \ ds)$$

Thus we have :

$$x\ \ t\ = x_0 + \int_{t_0}^{t} f(s, X\ \ s\ \ ds$$

For all $t \in J$ proving that a solution of the IVP (9) is also a solution of the integral equation (10).

1.4 EXISTENCE AND UNIQUENESS OF SOLUTIONS (SCALAR CASE)

In this section, we consider the scalar case, (i.e. a single differential equation) of the initial value problem. Now Ω will be an open subset of $\mathbb{R}$ which, without loss of generality will be taken to be an open internal, we write J for Ω. Thus we have on function

$$f : I \times J \to \mathbb{R}$$

Along with $t_0 \in I$, $x_0 \in J$ giving rise to the scalar case of the initial value problem :

$$\frac{dX}{dt} = f(t, x) \qquad x(t_0) = x_0 \text{ .... } (11)$$

Following property of *f* ensures both, existence and uniqueness of the solution of (11).

<u>**Definition 1 :**</u> *f* is **locally lipschitz** on J, uniformly in $t \in I$ if the following two conditions are satisfied.

a) f is continuous on $I \times J$.
b) For each $t_0 \in I$, $x_0 \in J$ there exist finite numbers $\delta > 0, K > 0$ satisfying the following :

i) $t_0 - \delta, t_0 + \delta < I,\ x_0 - \delta, x_0 + \delta < J$ and

ii) $\left|f\ t, x - f(t, y)\right| \leq K\left|x - y\right|$ holds for all $t \in\ t_0 - \delta, t_0 + \delta$ and for all pairs x, y in $x_0 - \delta, x_0 + \delta$.

Remark : An **autonomous** ODE arises from a function $f : J \to \mathbb{R}$ which is independent of the time variable $t \in I$. For such a function, the condition (b) in the definition takes the following simpler from : for each $x_0 \in J$, there exist $\delta < 0, K > 0$ satisfying :

i) $x_0 - \delta, x_0 + \delta < J$ and

ii) $\left|f\ x - f(y)\right| \leq K\left|x - y\right|$ for all x, y in J.

Also note that this condition implies continuity of *f* at every $x_0 \in J$ and therefore there is no separate mention of condition (a) in the definition of local Lipschitz property of such a $f : J \to \mathbb{R}$.

Following proposition describes a broad class of functions with the locally Lipschitz property :

Proposition 2 : If $f : I \times J \to \mathbb{R}$ is continuously differentiable on its domain, then it has the locally Lipschitz property.

Proof : Let $t_0, x_0 \in I \times J$ be arbitrary. Using openness of $I \times J$, choose $\delta > 0$ such that $t_0 - \delta, t_0 + \delta \times x_0 - \delta, x_0 + \delta < I \times J$.

Now, the function $\frac{\partial f}{\partial x}$ is continuous on $I \times J$ and therefore it is bounded on the compact subset $t_0 - \delta, t_0 + \delta \times x_0 - \delta, x_0 + \delta$. We consider any constant *k > 0* with the property.

$$\left|\frac{\partial f}{\partial x}(t,x)\right| \leq K \text{ for all } t \in t_0 - \delta, t_0 + \delta \text{ and for all } x \in x_0 - \delta, x_0 + \delta .$$

Finally let $t \in t_0 - \delta, t_0 + \delta$ x, y in $x_0 - \delta, x_0 + \delta$ be arbitrary. By the mean value theorem of differential calculus, we have :

$$f\ t, y\ - f(t,x) = (y - x)\frac{\partial f}{\partial x}\ t, \mathbb{Z}$$

for some z between x and y. Therefore,

$$\left|f\ t, y\ - f(t,x)\right| \leq \left|y - x\right|\left|\frac{\partial f}{\partial x}(t,z)\right| \leq K\left|y - x\right|$$

(Since $t, z \in t_0 - \delta, t_0 + \delta \times x_0 - \delta, x_0 + \delta$ and therefore $\left|\frac{\partial f}{\partial x}(t, \mathbb{Z}\right| \leq K$. This proves the locally Lipschitz property of *f*.

Theorem 1 (Emil Picard) : If $f : I \times J \to \mathbb{R}$ is locally Lipschitz then for any $t_0 \in I$, $x_0 \in J$, the initial value problem.

$$\frac{df}{dx} = f(t,x), \qquad x(t_0) = x_0$$

has a solution $x :\ t_0 - \delta, t_0 + \delta \to J$ for some $\delta > 0$.

Proof : Choose $\delta > 0$ such that $t_0 - \delta, t_0 + \delta < I,\ x_0 - \delta, x_0 + \delta < J$ and there exists $K > 0$ for which $\left|f\ t, x\ - f(t,y)\right| \leq K\left|x - y\right|$ holds for all $t_0 - \delta, t_0 + \delta \times x_0 - \delta, x_0 + \delta$. We choose M > 0 such that $\left|f\ t, x\ \right| \leq M$ for all $t \in t_0 - \delta, t_0 + \delta$, $x \in x_0 - \delta, x_0 + \delta$.

Using the constants $\delta > 0$, K > 0, M > 0, chosen above, we choose one more constant η satisfying $0 < \eta < \min\left\{\frac{1}{K}, \frac{\delta}{M}\right\}$.

We define a sequence of functions :
$x_k : t_0 - \eta\, t_0 + \eta \to \mathbb{R}, k \in \mathbb{Z}+$ recursively as follows :

$$x_0\,(t) \equiv x_0$$

$$x_1(t) = x_0 + \int_{t_0}^{t} f(s\, x_0)\, ds$$

$$x_2\,(t) = x_0 + \int_{t_0}^{t} f(s\, x_1\,(s))\, ds$$

$$\vdots$$

$$x_{k+1(t)} = x_0 + \int_{t_0}^{t} f(s\, x_k\,(s))\, ds$$

$$\vdots$$

The sequence $x_k : k \in \mathbb{Z}^+$ of functions has the following two properties :

a) $x_k(t) \in x_0 - \delta, x_0 + \delta$ for each $t \in t_0 - \delta, t_0 + \delta$

b) $\left|x_{k+1}(t) - x_k(t)\right| \le \dfrac{MK^k \left|t - t_0\right|^{k+1}}{(k+1)!}$

Both these properties are derived using principle of mathematical induction and the locally Lipschitz property of *f*. Using property (b) we deduce that the sequence $x_k : k \in \mathbb{Z}^+$ is uniformly canchy on $t_0 - \eta, t_0 + \eta$. For if $t \in t_0 - \eta, t_0 + \eta\,, k \in \mathbb{Z}^t, p \in \mathbb{N}$, then $x_{k+p}(t) - x_k(t) = \sum_{j=k}^{k+p-1} \left[x_{j+1}(t) - x_j(t)\right]$ and therefore

$$\left|x_{k+p}(t) - x_k(t)\right| \le \sum_{j=k}^{k+p-1} \left[x_{j+1}(t) - x_j(t)\right]$$

$$\le \frac{M}{K} \sum_{j=k}^{k+p-1} \frac{K^{j+1}\left|t - t_0\right|^{j+1}}{(j+1)!}$$

$$\le \frac{M}{K} \sum_{j>k} \frac{K^j \eta^j}{j!}$$

$$\frac{M}{K} \sum_{j>k} \frac{K\eta^{\;j}}{j!}$$

$\rightarrow 0$ as $k \rightarrow \infty$.

This last observation is true because $\sum_{j>k} \dfrac{K\eta^{\;j}}{j!}$ is convergent, converging $t_0\, e^{k\eta}$. Note that in the inequalities :

$\left|x_{k+p}(t)-x_k(t)\right| \le \dfrac{M}{K}\sum_{j\ge k}\dfrac{K\eta^{\ j}}{j!}$ the right hand sides are independent of t and therefore $\left|x_{k+p}(t)-x_k(t)\right| \to 0$ uniformly on $t_0-\eta, t_0+\eta$ as $k\to\infty$, p being arbitrary. This completes the proof of our claim that the sequence $x_k : k\in\mathbb{Z}^t$ of functions is uniformly Cauchy on $t_0-\eta, t_0+\eta$.

Using this last mentioned property of the sequence $x_k : k\in\mathbb{Z}^+$ we define a function.

$x:\ t_0-\eta, t_0+\eta \ \to\ x_0-\delta, x_0+\delta$

by putting $x\ \ t\ = \lim\limits_{k\to\infty} x_k(t), t\in\ t_0-\eta, t_0+\eta$. The function x, thus defined, is the uniform limit of the sequence $x_k : k\in\mathbb{N}$. Therefore we have :

$$\begin{aligned}
x\ \ t\ &= \lim_{k\to\infty} x_k(t)\\
&= \lim_{k\to\infty}\left[x_0+\int_{t_0}^{t} f(s, x_{k-1}(s)\ ds\right]\\
&= x_0+\lim_{k\to\infty}\int_{t_0}^{t} f(s, x_{k-1}(s)\ ds\\
&= x_0+\int_{t_0}^{t}\ \lim_{k\to\infty} f(s, x_{k-1}(s)\ \ ds\\
&= x_0+\int_{t_0}^{t} f(s, \lim_{k\to\infty} x_{k-1}(s)\ ds\\
&= x_0+\int_{t_0}^{t} f(s, x(s)\ ds
\end{aligned}$$

All the above equations being valid because of the uniform convergence of of $x_k\ t_0\ x$ on $t_0-\eta, t_0+\eta$.

Finally, the identity :

$x\ \ t\ = x_0+\int_{t_0}^{t} f(s, x(s)\ ds, t\in\ t_0-\eta, t_0+\eta$ derived above has the following two consequences.

1) Differentiation of the identity implies :

$$\frac{dx(t)}{dt} = \frac{d}{dt}\left[x_0 + \int_{t_0}^{t} f(s, x(s))ds\right]$$

$$= 0 + \frac{d}{ds}\int_{t_0}^{t} f(s, x(s))ds$$

$$= f(t, x(t))$$

by fundamental theorem of integral calculus.

2) $$x(t_0) = x_0 + \int_{t_0}^{t_0} f(s, x(s))ds$$

$$= x_0 + 0$$

$$= x_0$$

We have now verified that the function $x:[t_0-\eta, t_0+\eta] \to [x_0-\delta, x_0+\delta] < J$ is a solution of the given initial value problem.

<u>Remark :</u> The above theorem there proves that the functions x_k constructed above are approximate solutions of the initial value problem (11).

The sequence $\{x_k : k \in \mathbb{Z}^t\}$ is called Picard's scheme of approximate solutions of the initial value problem.

In the next chapter, we will generalize this result (the scalar case) so as to become applicable to a system of first order ODE. We will also prove that any two solutions of the initial value problem (11) agree on the overlap of their domains.

1.5 ILLUSTRATIVE EXAMPLES

<u>Example 1 :</u> Obtain Picard's scheme of approximate solutions of the initial value problem : $\frac{dx}{dt} = x, x(2) = 3$ and thereby obtain the solution of it.

<u>Solution :</u> This DE is an autonomous ODE with $f(t, x) = x, t_0 = 2$ and $x_0 = 3$. Therefore the approximate solutions are as follows :

$$x_0(t) = 3$$

$$x_1(t) = 3 + \int_2^t 3ds = 3 + 3\,(t-2)$$

$$x_2(t) = 3 + \int_2^t \; 3 + 3\,(s-2) \; ds = 3 + 3\,(t-2) + \frac{3(t-2)^2}{2!}$$

$$x_k(t) = 3 + \int_2^t \left[3 + \frac{3}{1!}(s-2) + \frac{3}{2!}(s-2)^2 + .. + \frac{3}{k!}(s-2)^k \right] ds$$

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

Therefore, the solution of the IVP is

$$x\;\;t\;\; = \lim_{k\to\infty} x_k\;\;t$$

$$= \lim_{k\to\infty} \left[3 + \frac{3}{1!}(t-2) + \frac{3}{2!}(t-2)^2 + ... + \frac{3}{k!}(t-2)^k \right]$$

$$= 3 \sum_{k\geq 0} \frac{t-z^{\;\;k}}{k!} = 3e^{(t-2)}$$

Example 2 : Obtain approximate solution (to within t^7) of the initial value problem :

$$\frac{dx}{dt} = xt + t^2, x(0) = 2$$

Solution : Here, $f\;\;t, x\;\; = xt + t^2, t_0 = 0, x_0 = 2$.

Therefore, $x_0(t) \equiv 2$

$$x_1(t) = 2 + \int_0^t (2s + s^2) ds$$

$$= 2 + \frac{t^2}{2} + \frac{t^3}{3}$$

$$x_2(t) = 2 + \int_0^t (2s + s^3 + \frac{s^4}{4} + s^2)\, ds$$

Using this observation, we get :

$$x_2(t) = 0 + \int_0^t \max\ s, x_1(s)\ ds$$

$$= \begin{cases} \int_0^t s ds & t \le 2 \\ \int_0^2 s ds + \int_2^t \frac{s^2}{2} ds, & t \ge 2 \end{cases}$$

$$= \begin{cases} \frac{t^2}{2} & t \le 2 \\ 2 + \int_0^{t-2} (r+2)^2 ds & t \le 2 \end{cases}$$

$$= \begin{cases} \frac{t^2}{2} & t \le 2 \\ +2 + 2(t-2)^2 + 2\frac{(t-2)^2}{2} + \frac{(t-2)^3}{3.2} & \text{if} \quad t \ge 2 \end{cases}$$

Again, note that $\max\ t, x_2(t)\ = \begin{cases} t & \text{if } t \le 2 \\ x_2(t) & \text{if } t \ge 2 \end{cases}$ and consequently, we get

$$x_3(t) = \begin{cases} \frac{t^2}{2} & t \le 2 \\ 2 + \frac{2(t-2)}{1!} + \frac{2(t-2)^2}{2!} + \frac{2(t-2)^3}{3!} + \frac{2(t-2)^4}{4!} & t \ge 2 \end{cases}$$

Using principle of mathematical induction, we get

$$x_k(t) = \frac{t^2}{2} \qquad t \le 2$$

$$= 2 + 2 + \frac{2(t-2)}{1!} + \frac{2(t-2)^2}{2!} + \ldots + \frac{2(t-2)^k}{k!} + \frac{2(t-2)^{k+1}}{(k+1)!} \quad t \ge 2$$

Noting that $\frac{t-2^{\ k}}{k!} \to 0$ as $k \to \infty$ for every $t \in \mathbb{R}$ we get :

$$x\ t\ = \lim_{k \to \infty} x_k(t) = \begin{cases} \frac{t^2}{2} & \text{if } t \le 2 \\ 2e^{t-2} & \text{if } t \ge 2 \end{cases}$$

1.6 EXERCISES

Obtain solutions to within t^5 of the following initial value problems :

1) $\frac{dx}{dt} = t + x^2$ $\quad x(0) = 0 \ for\ 1t1 < \frac{1}{2}$

2) $\frac{dx}{dt} = x^2$ $\quad x(0) = 1$

3) $\frac{dx}{dt} = 1 + x^2$ $\quad x(0) = 0$

4) $\frac{dx}{dt} = \frac{3x}{t}$ $\quad x(1) = 1$

5) $\frac{dx}{dt} = 2tx + x - x^3,$ $\quad x(0) = 0$

6) $\frac{dx}{dt} + xt = \frac{t}{x},$ $\quad x(0) = 2$

7) $\frac{dx}{dt} = \frac{x}{2} + x^3,$ $\quad x(0) = 1$

❖❖❖❖

2

SYSTEMS OF FIRST ORDER ODE

Unit Structure :

2.1 INTRODUCTION

Basic concepts related to differential equations such as systems of first order ordinary differential equations, the initial value problem associated with such a system, a local solutions of the initial value problem etc. were introduced in the first unit. At the end of the unit, we proved a result regarding the solution of a single first order ODE.

We will extend the results of a single ODE to a system of first order ODE and prove both, existence and uniqueness of solutions of an initial value problem associated with a system of first order ODE. We will then derive some simple results giving information about the nature of solution of such an initial value problem.

We will conclude the chapter by explaining how a system of higher order ODE can be reduced to a system of first order ODE. We can then invoke the existence / uniqueness theorems and apply them to the first order ODE and get some information about the solutions of the higher order ODE.

2.2 EXISTENCE AND UNIQUENESS OF SOLUTIONS

We will use the same notations which were introduced in Unit 1.

Let Ω be an open subset of $\mathbb{R}^n$, I an open interval and let $f : I \times \Omega \to \mathbb{R}^n$ be a time dependent vector field having components:

$$f_1, f_2, ..., f_n : I \times \Omega \to \mathbb{R}$$

The vector field gives rise to the first order ODE :

$$\frac{dx}{dt} = f\ \ t, X$$

… (1)

which when written in terms of its components becomes the following system of first order ODE :

$$\left.\begin{aligned} \frac{dx_1}{dt} &= f_1(t, X_1, \ldots\ldots, X_n) \\ \frac{dx_2}{dt} &= f_2(t, X_1, \ldots\ldots, X_n) \\ &\vdots \\ \frac{dx_n}{dt} &= f_n(t, X_1, \ldots\ldots, X_n) \end{aligned}\right\}$$

... (2)

Between the expressions (1) and (2) the compact form (1) is more convenient and therefore we will use it throughout this chapter, bearing in mind that it is the same as the system (2).

Recall, given $t_0 \in I, x_0 \in \Omega$ we have the initial value problem :

$$\frac{dx}{dt} = f(t, X), X(t_0) = x_0$$

… (3)

a solution of which is a continuously differentiable curve $X : J \to \Omega$ satisfying $\frac{dX(t)}{dt} = f(t, X(t)$ for all $t \in I, X(t_0) = x_0$. (J being an open interval with $t_0 \in J \subseteq I$).

Now, towards the existence / uniqueness of solution of (3) we introduce the locally Lipschitz property of *f*:

Definition :

The vector field f has the **locally Lipschitz property** if it satisfies the following two conditions :

a) *f* is continuous on its domain and

b) For $t_0 \in I, x_0 \in \Omega$, there exist two constants $\delta > 0, K > 0$ such that

i) $t_0 - \delta, t_0 + \delta\ \subseteq I, B\ \ x_0, \delta\ < \Omega$ and

ii) $\|f\ \ t\,x\ - f(t\,y)\| \leq K\,\|x - y\|$ for all $t \in\ t_0 - \delta, t_0 + \delta$ and for all x, y in $B\ \ x_0, \delta$.

Remark :

We will also consider vector fields $f : \Omega \to \mathbb{R}^n$ as a special case of $f : I \times \Omega \to \mathbb{R}^n$ in which $f\ \ t, x$ is independent of $t : f(t, x) = f(x)$. Recall such vector field give rise to the **autonomous** ODE : $\frac{dx}{dt} = f(x)$. Now, the definition of locally Lipschitz property for such $f : \Omega \rightarrow \mathbb{R}^n$ takes the following simpler form : For any $x_0 \in \Omega$, there exist $\delta > 0, K > 0$ such that $B\ \ x_0, \delta\ \ < \Omega$ and $\|f\ \ x\ \ - f(y)\| \leq K\|x - y\|$ for all $x, y,\ B\ \ x_0, \delta$.

Note that the above condition implies continuity of f at every $x_0 \in \Omega$ and as such there is no separate mention of continuity on f.

Now we have the following result :

Proposition 1: If $f : I \times \Omega \rightarrow \mathbb{R}^n$ is continuously differentiable on its domain, then it has the locally Lipschitz property.

The proof of this proposition is on lines similar to that of **Proposition 2** of **Unit 1**.

Theorem 1: (Local Existence of solutions) : If $f : I \times \Omega \rightarrow \mathbb{R}^n$ has the locally Lipschitz property then the initial value problem :

$\frac{dx}{dt} = f(t, x), x(t_0) = x_0$ has a solution $X :\ \ t_0 - \delta, t_0 + \delta\ \ \rightarrow \Omega$.

Proof : We give here a sketchy proof. (To fill up all the details that are left here, consult the proof of **Theorem 1** in **unit 1**)

For $t_0 \in I, x_0 \in \Omega$ choose b > 0 such that

$t_0 - b, t_0 + b\ \ \subseteq I, \overline{B}\ \ x_0, b\ \ \subseteq \Omega$ and choose $K > 0$ such that $\|f\ \ t\ x\ \ - f(t\ y)\| \leq K\|x - y\|$ holding for all $t \in\ \ t_0 - b, t_0 + b$ and for all x, y in $\overline{B}\ \ x_0, b$.

Using the fact that continuous functions are bounded on compact subsets of their domains, we choose a constant M > 0 such that $\|f\ \ t\ x\ \ \| \leq M$ holds for all $t \in\ \ t_0 - b, t_0 + b$ and for all $x \in \overline{B}\ \ x_0, b$.

We choose one more constant δ which satisfying $0 < \delta < \min\left\{\frac{1}{k}, \frac{b}{M}\right\}$.

Now, we define the sequence of maps $x_k :\ \ t_0 - b, t_0 + b\ \ \rightarrow \mathbb{R}^n$ recursively by putting :

$x_0\ (t) \equiv x_0$

$$x_1(t) = x_0 + \int_{t_0}^{t} f(s\ x_0)\, ds$$

$$x_2\ (t) = x_0 + \int_{t_0}^{t} f(s\ X_1\ (s))\, ds$$

$\vdots$

$$x_{k+1}(t) = x_0 + \int_{t_0}^{t} f(s\ X_k\ (s))\, ds \qquad \text{for each } t \in t_0 - b, t_0 + b\ .$$

$$\vdots$$

About the sequence $x_k : k \in \mathbb{Z}^+$ we have the following :

a) $x_k(t) \in \overline{B}\ x_0, b$ for each $t \in t_0 - b, t_0 + b$.

b) $\|x_{k+1}(t) - x_k(t)\| \leq \dfrac{MK^k}{(k+1)!}$ $t \in t_0 - b, t_0 + b\ , k \geq 0$

c) $x_k : k \in \mathbb{Z}^+$ is uniformly Cauchy on $t_0 - b, t_0 + b$ (verification of these properties is left for the reader). We consider the uniform limit of the sequence $x :\ t_0 - \delta, t_0 + \delta \rightarrow \overline{B}\ x_0, b \subseteq \Omega$ which is given by

$$x\ t = \lim_{k \to \infty} x_k(t) \qquad t \in t_0 - b, t_0 + b$$

Thus, $x(t) = \lim\limits_{k \to \infty} \left[x_0 + \int_{t_0}^{t} f(s, x_{k-1}(s)\, ds \right]$

$$= x_0 + \lim_{k \to \infty} \int_{t_0}^{t} f(s, x_k(s)\, ds \quad = x_0 + \int_{t_0}^{t} \lim_{k \to \infty} f(s, x_k(s) \quad ds$$

$$= x_0 + \int_{t_0}^{t} f(s, \lim_{k \to \infty} x_k(s)\, ds = x_0 + \int_{t_0}^{t} f(s, x(s)\, ds$$

Thus the function $t \mapsto x(t)$ satisfying the integral equation :

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s)\, ds \text{ for all } t \in t_0 - b, t_0 + b\ .$$

Finally the validity of this integral equation has the following two implications :

1) $X(t_0) = x_0 + \int_{t_0}^{t_0} f(s, x(s)\, ds\ = x_0 + 0 = x_0$

2) $\dfrac{dx}{dt} = 0 + \dfrac{d}{dt} \int_{t_0}^{t} f(s, x(s))\, ds\ = f(t, X(t))$

This now proves that the curve $X :\ t_0 - \delta, t_0 + \delta \rightarrow \Omega,$ thus obtained is a solution of the initial value problem.

2.3 UNIQUENESS OF A SOLUTION

We prove an inequality which will lead us to the uniqueness of the solutions :

Proposition 2 : (Gronwall's Inequality) : Let $f:[a,b]\to[0,\infty)$, $g:[a,b]\to[0,\infty)$ be continuous functions and $A\geq 0$, a constant satisfying

$$f(t)\leq A+\int_a^t f(s)\,g(s)\,ds \text{ for all } t\in[a,b]\text{. Then}$$

$$f(t)\leq A.e^{\int_a^t g(s)\,ds} \quad \text{for all } t\in[a,b].$$

Proof : First we assume A > 0 and put $h(t)=A+\int_a^t f(s)\,g(s)\,ds$ for all $t\in[a,b]$. Then $h(t)>0$ for all $t\in[a,b]$ and

$$\begin{aligned} h'(t) &= f(t)\cdot g(t)\\ &\leq h(t)\cdot g(t) \end{aligned}$$

that is, $\dfrac{h'(t)}{h(t)}\leq g(t)$ for all $t\in[a,b]$. Integrating this inequality over $[a,t]$ for $t\in[a,b]$ we get $\log\left(\dfrac{h(t)}{h(a)}\right)\leq\int_a^t g(s)\,ds$.

Nothing that $h(a)=A$, we get the desired inequality in this case. Now suppose A = 0. Then for each $n\in\mathbb{N}$ we have :

$$f(t)\leq\frac{1}{n}+\int_a^t f(s)\,g(s)\,ds\text{, for all } t\in[a,b)$$

Applying the above argument to $A=\dfrac{1}{n}$, we get

$$f(t)\leq\frac{1}{n}e^{\int_a^t g(s)\,ds}$$

for every $t\in[a,b]$ and for every $n\in\mathbb{N}$. Holding t fixed and taking limit of the last inequality as $n\to\infty$ we get

$$\begin{aligned} f(t) &\leq \frac{1}{n}0.e^{\int_a^t g(s)\,ds}\\ &= A\,e^{\int_a^t g(s)\,ds} \end{aligned}$$

Now we prove the following *essential uniqueness* result of the solution :

Proposition 3 : Let $x : J \to \Omega$, $y : J \to \Omega$ be two solutions of the initial value problem :

$$\frac{dx}{dt} = f(t, x) \quad x(t_0) = x_0 .$$

then $x(t) = y(t)$ for all $t\, J \cap J$.

Proof : Recall that both *x* and *y*, being solutions of the initial value problem, satisfy the integral equations on their domain intervals :

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds$$

$$y(t) = x_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

Therefore $x(t) - y(t) = \int_{t_0}^{t} \left[f\ \ s, x(s)\ \ - f\ \ s, y(s)\ \ \right] ds$ which implies :

$$\|x(t) - y(t)\| \le 0 + \int_{t_0}^{t} \|f\ \ s, x(s)\ \ - f\ \ s, y(s)\ \ \|$$

$$\le 0 + \int_{t_0}^{t} K \|x(s) - y(s)\| \text{ for all } t \ge t_0 .$$

Applying Gronwalls result with $A = 0$, we get

$$0 \le \|x(t) - y(t)\| \le 0 \text{ for all } t \le t_0 .$$

This gives the desired equality $x(t) = y(t)$ for all $t \in J \cap J$.

Towards the uniqueness of the solution of the initial value problem (3), we consider all the solutions of the initial value problem (3). Let the totality of them be denoted by : $x_\lambda : J_\lambda \to \Omega : \lambda \in \Lambda$ the solutions x_λ being thus indexed by a suitable indexing set Λ .

Above we have verified that any two solutions, say x_{λ_1} and x_{λ_2} are equal on the overlap $J_{\lambda_1} \cap J_{\lambda_2}$ of their domains. Therefore, we patch together all the solutions to get a maximal solution is the solution defined on the largest open interval. It is obtained as follows.

Let $J = U\{J_\lambda : \lambda \in \Lambda\}$ clearly J is an open sub internal of I with $x_0 \in J$ and all the solutions x_λ patch up to get a solution $x : J \to \Omega$ of the initial value problem (3).

Because, we consider all the solutions of (3) we get that J is the largest open internal on which the solution of (3) is defined. We summarize all this discussion in the following theorem.

Theorem 2 : (Uniqueness of the solution) :
The initial value problem (3) has a unique (maximal) solution defined on the largest open sub-interval *J*.

Clearly the solution is unique because it is defined on the largest and hence unique internal *J*.

From now onwards we will consider this unique solution defined on the maximal interval.

2.4 THE AUTONOMOUS ODE

We note a few simple properties of the autonomous ODE :

$$\frac{dx}{dt} = f(x)$$

determined by a locally Lipschitz and hence continuous vector field $f : \Omega \to \mathbb{R}^n$.

Now if $x : J \to \Omega$ is a solution of this autonomous ODE then $\frac{dx(t)}{dt} = f(x(t)) \quad t \in I$ and continuity of f and differentiability of $x(t)$ and hence its continuity together implies that $\frac{dx}{dt}(t)$ is continuous and thus the curve $x : J \to \Omega$ is continuously differential on I. This argument also gives the following result. If $f : \Omega \to \mathbb{R}^n$ is k times continuously differentiable then the solution $x : J \to \Omega$ is k + 1 times continuously differentiable on its domain interval.

Note that a solution of an autonomous ODE may not be defined for all $t \in \mathbb{R}$: We consider the initial value problem.

$$\frac{dx}{dt} = x^2, x(0) = 1.$$

Clearly its solution is $x(t) = \frac{1}{1-t}$ which is defined on $(-\infty, 1)$ only and not on the whole of $\mathbb{R}$.

It is a result that if the vector field f is compactly supported, then the solution of the initial value problem (3) is defined for all $t \in \mathbb{R}$. (We do not prove this result here).

2.5 SOLVED EXAMPLES

[Note : Recall, if $g:[a,b]\to\mathbb{R}^n$ is an integrable (vector valued) functions with components $(g_1,g_2,...,g_n)$ then $\int_a^b g(x)\,dx =$ $\left(\int_a^b g_1(x)dx, \int_a^b g_2(x)dx,....,\int_a^b g_n(x)dx,\right)$. Equivalently written in the columnal form, we have :

$$\int_a^b g(x)\,dx = \begin{bmatrix} \int_a^b g_1(x)\,dx \\ \int_a^b g_2(x)\,dx \\ \vdots \\ \int_a^b g_n(x)\,dx \end{bmatrix}$$

We will use these notations in this article.

Example 1 : Obtain approximate solutions (upto t^3) of the following initial value problem :

$$\frac{dx}{dt} = 2x+3y \quad x(0)=1$$

$$\frac{dy}{dt} = t+y \qquad y(0)=2$$

Solution :

We have : $x_0(t)=1,\;\; y_0(t)=2$

$$x_1(t) = 1+\int_0^t \left[2.x_0(s)+3y_0(s)\right]ds = 1+\int_0^t 8ds = 8t+1$$

$$y_1(t) = 2+\int_0^t \left[s+y_0(s)\right]ds$$

$$= 2+\int_s^t (s+2)\,ds = 2+2t+\frac{t^2}{2}$$

Thus $\begin{bmatrix} x_1(t) \\ y_1(t) \end{bmatrix} = \begin{bmatrix} 8t+1 \\ \frac{t^2}{2}+2t+2 \end{bmatrix}$.

Next $x_2(t) = 1+\int_0^t [2x_1(s)+3y_1(s)]ds = 1+\int_0^t \left[8+2\times 2s+\frac{3}{2}s^2\right]ds$

$$= \frac{t^3}{3}+11t^2+8t+1$$

$$y_2(t) = 2+\int_0^t [s+y_1(s)]ds = 2+\int_0^t \left[\frac{s^2}{2}+3s+2\right]ds$$

$$= 2+\frac{t^3}{3.2}+\frac{3t^2}{2}+2t = \frac{t^3}{6}+\frac{3}{2}+2t+2$$

Thus, $\begin{bmatrix} x_2(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} \frac{t^3}{3} + 11t^2 + 8t + 1 \\ \frac{t^3}{6} + \frac{3t^2}{2} + 2t + 2 \end{bmatrix}$ and so on.

Example 2 : Obtain approximate solution upto t^5 :

$$\frac{dx}{dt} = tx + t^2 y \qquad x(0) = 1$$

$$\frac{dy}{dt} = xy + t \qquad y(0) = 2$$

Solution : We have $x_0(t) = 1, y_0(t) = 2$

$$x_1(t) = 1 + \int_0^t \left[s + 2s^2 \right] = 1 + \frac{t^2}{2} + \frac{2t^3}{3}$$

$$y_1(t) = 2 + \int_0^t [2 + s] ds = 2 + 2t + \frac{t^2}{2}$$

Thus, $\begin{bmatrix} x_1(t) \\ y_1(t) \end{bmatrix} = \begin{bmatrix} \left(\frac{2t^3}{3} + \frac{t^2}{2} + 1 \right) \\ \left(\frac{t^2}{2} + 2t + 2 \right) \end{bmatrix}$

Next $x_2(t) = 1 + \int_0^t \left[x_1(s) + s^2 y_1(s) \right] ds = 1 + \int_0^t \left[s + \frac{5s^3}{2} + \frac{2s^4}{3} \right] ds$

$$= 1 + \frac{t^2}{2} + \frac{5t^4}{8} + \frac{2t^5}{15}$$

$$y_2(t) = 2 + \int_0^t [x_1(s) y_1(s) + s] ds$$

$$= 2 + \int_0^t \left[\frac{s^5}{3} + \frac{19s^4}{12} + \frac{7s^3}{3} + \frac{s^2}{2} + 2s + 2 \right] ds$$

$$= 2 + \frac{t^6}{18} + \frac{19t^5}{60} + \frac{7t^4}{12} + \frac{t^3}{6} + t^2 + 2t$$

and so on

Example 3 : Obtain approximate solutions (upto t^3) :

$$\frac{dx}{dt} = 2y + t \qquad x(1) = 1$$

$$\frac{dy}{dt} = 3z + t^2 \qquad y(1) = 2$$

$$\frac{dz}{dt} = xz \qquad z(1) = 3$$

Solution : We have : $\begin{bmatrix} x_0(t) \\ y_0(t) \\ z_0(t) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$$x_1(t) = 1 + \int_1^t (4+s)\,ds = 1 + 4t + \frac{t^2}{2} - 4 - \frac{1}{2} \;= \frac{t^2}{2} + 4t - \frac{9}{2}$$

$$y_1(t) = 2 + \int_1^t (3+s^2)\,ds = 2 + 9t + \frac{t^3}{3} - 9 - \frac{1}{3} = \frac{t^3}{3} + 9t - \frac{22}{3}$$

$$z_1(t) = 3 + \int_1^t 3ds = 3 + 3t - 3 = 3t$$

Thus,
$$\begin{bmatrix} x_1(t) \\ y_1(t) \\ z_1(t) \end{bmatrix} = \begin{bmatrix} \frac{t^2}{3} + 4t - \frac{9}{2} \\ \frac{t^3}{3} + 9t - \frac{22}{3} \\ 3t \end{bmatrix}$$

$$x_2(t) = 2 + \int_1^t \left(\frac{2s^3}{3} + 8s - \frac{44}{3} + s \right) ds$$

$$= 2 + \frac{t^4}{6} + \frac{9t^2}{2} - \frac{44t}{3} - \frac{1}{6} - \frac{9}{2} + \frac{44}{3}$$

$$= \frac{t^4}{6} + \frac{9t^2}{2} - \frac{44t}{3} - 8$$

$$y_2(t) = 2\int_1^t (9s + s^2)\,ds$$

$$= 2 + \frac{9t^2}{2} + \frac{t^3}{3} - \frac{9}{2} - \frac{1}{3}$$

$$= \frac{9t^2}{2} + \frac{t^3}{3} - \frac{17}{6}$$

$$z_2(t) = 3 + \int_1^t x_1(s)\, z_1(s)\, ds$$

$$= 3 + \int_1^t \left(\frac{s^2}{2} + 4s - \frac{9}{2} \right)(3s)\,ds$$

$$= 3 + \frac{3t^4}{2.4} + 4t^3 - \frac{27t^2}{4} - \frac{3}{2.8} - 4 + \frac{27}{4}$$

$$= \frac{3t^4}{8} + 4t^3 - \frac{27t^2}{4} - \frac{41}{16}$$

Thus, $\begin{bmatrix} x_2(t) \\ y_2(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} \frac{t^4}{6}+\frac{9t^2}{2}-\frac{44t}{3}-8 \\ \frac{t^3}{3}+\frac{9t^2}{2}-\frac{17}{6} \\ \frac{3t^4}{8}+4t^3-\frac{27t^2}{4}-\frac{41}{16} \end{bmatrix}$

2.6 HIGHER ORDER ODE

As mentioned earlier, we associate a first order ODE with a order k ODE and try to get information about the solutions of the higher order ODE in terms of those of the associated first order ODE. In particular, we are interested in a condition on the function $f = f(t, X, \frac{dx}{dt} ... \frac{d^{k-1}x}{dt^k})$ which will ensure existence and uniqueness of the initial value problem for the higher order ODE. In this article we explain the theory.

To begin with, we consider the ODE :

$$\frac{d^k x}{dt^k} = f(t, x, \frac{dx}{dt} ... \frac{d^{k-1}x}{dt^{k-1}})$$

… (4)

Together with the initial condition :

$$x(t_0) = x_0, \frac{dx}{dt}(t_0) = w_1 ... \frac{d^{k-1}x}{dt^{k-1}}(t_0) = w_{k-1}.$$

We introduce a new variable $y = (y_1, y_2, ..., y_k)$ where y_1 ranges in Ω and $y_2 \dots y_k$ ranging in $\mathbb{R}^n$. Next we define $F : I \times \Omega \times \mathbb{R}^n \times .... \times \mathbb{R}^n \to \mathbb{R}^{nk}$

by putting $F(t, y) = F(t, y_1, ..., y_n ....) = (y_2, y_3, ... y_K, f(t, y))$

Now, the given initial value problem for the order k ODE give rise to the following initial value problem in the first order ODE.

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

… (5)

the initial condition for it being $y(t_0) = (x_0, w_1, ..., w_{k-1})$.

Clearly the above first order ODE is actually the following system of first order ODE :

$$\left.\begin{aligned} \frac{dy_1}{dt} &= y_2 \\ \frac{dy_2}{dt} &= y_3 \\ &\vdots \\ \frac{dy_k}{dt^k} &= f(t, y) \\ &= f(t, y_1, y_2.....y_k) \end{aligned}\right\}$$

... (5*)

Note that taking $y_1 = x$ the system (5) reduces to the given order k ODE (4).

This shows that the solutions of the order k ODE (4) can be studied in terms of the solutions of the first order ODE(5). Note that if f (of the ODE (4)) is continuously differentiable on its domain, namely the set $I \times \Omega \times \mathbb{R}^n \times ..... \times \mathbb{R}^n$ then F is continuously differentiable on its own domain, consequently the existence and uniqueness result for the first order ODE applies to the initial value problem :

$$\frac{dy}{dt} = F(t, y)$$
$$y(t_0) = (x_0, w_1....w_{k-1})$$

The solution of which giving the solution of the above initial value problem :

$$\frac{d^k x}{dt} = f\left(t, x, \frac{dx}{dt} ... \frac{d^{k-1}x}{dt^{k-1}}\right)$$

$$x\ t_0\ = x_0, \frac{dx}{dt}\ t_0\ = w_1 ... \frac{d^{k-1}x}{dt^{k-1}}\ t_0\ = w_{k-1}.$$

We summarize this observation in the following :

Theorem 3 : If the function $f : I \times \Omega \times \mathbb{R}^n \times .... \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable on its domain of definition then the initial value problem.

$$\frac{d^k x}{dt^k} = f\left(t, x\, \frac{dx}{dt}, ..., \frac{d^{k-1}x}{dt^{k-1}}\right)$$

$$x\ t_0\ = x_0, \frac{dx}{dt}\ t_0\ = w_1, ..., \frac{d^{k-1}x}{dt^{k-1}}\ t_0\ = w_{k-1}$$

has a unique solution.

Illustrative Examples :

These examples explain how we obtain a system of first order ODE from a given higher order ODE.

1) The second order ODE : $\frac{d^2x}{dt^2} = k\,x, k$ being a given constant, gives rise to the system:

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ kx \end{bmatrix}$$

Moreover the initial condition $x\ t_0\ = x_0, \frac{dx}{dt}\ t_0\ = y_0$ gives the initial condition $\begin{bmatrix} x\ t_0 \\ y\ t_0 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ for the (reduced) first order system.

2) The initial value problem :

$$\frac{d^3x}{dt^3} + \frac{4d^2x}{dt^2} + 3t\frac{dx}{dt} + t^2x = t^5$$

$$x\ 0\ = 1, \frac{dx}{dt}\ 0\ = 2, \frac{d^2x}{dt^2}\ 0\ = 3$$

reduces to the following system of first order ODE along with the initial conditions.

$$\frac{dx}{dt} = y,\ x\ 0\ = 1$$

$$\frac{dy}{dt} = z,\ y\ 0\ = 2$$

$$\frac{dz}{dt} = -4z - 3ty - t^2x + t^5,\ z(0) = 3$$

3) The third order system of ODE : $\frac{d^3}{dt^3}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x.y \end{bmatrix}$

is equivalent to the following system of first order ODE in $z = \ z_1, z_2, z_3, z_4, z_5, z_6$

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{bmatrix} = \begin{bmatrix} z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_1 + z_2 \\ z_i.z_2 \end{bmatrix}$$

which is obtained by putting

$$z_1 = x, z_2 = y, z_3 = \frac{dx}{dt} \quad z_4 = \frac{dy}{dt} \quad z_5 = \frac{d^2x}{dt^2} \quad z_6 = \frac{d^2y}{dt^2}$$

2.7 EXERCISES

1) Prove : A continuously differentiable $f:\mathbb{R}^2 \to \mathbb{R}^2$ has the locally Lipschitz property.
2) Given a continuous, 2×2 matrix valued function $A:\mathbb{R}\to M_2\ \mathbb{R}$ let the time dependent vector field $f:\mathbb{R}\times\mathbb{R}^2\to\mathbb{R}^2$ be given by $f\ t,x\ = A(t)\,x$ for all $t\in\mathbb{R}, x\in\mathbb{R}^2$ prove that f is locally Lipschitz.
3) Give an example of a $f:\mathbb{R}^2\to\mathbb{R}^2$ which is continuous but not locally Lipschtiz.
4) Obtain approximate solutions of the following initial value problems.

a) $\dfrac{dx}{dt} = x^2y + x \quad x\ 1\ = 2$

$\dfrac{dy}{dt} = xy^2 + y \quad y\ 1\ = 3$

b) $\dfrac{dx}{dt} = 2x + y^2 \quad x\ 0\ = 2$

$\dfrac{dy}{dt} = 3y + 4x \quad y\ 0\ = 3$

c) $\dfrac{dx}{dt} = \dfrac{1}{y}, \quad x\ 0\ = 2$

$\dfrac{dy}{dt} = \dfrac{1}{x} \quad y\ 0\ = 3$

4) $\dfrac{dx}{dt} = \dfrac{3dy}{dt} + 4x, \quad x\ 0\ = -1$

$\dfrac{dy}{dt} = \dfrac{2dx}{dt} - y, \quad y\ 0\ = 2$

❖❖❖❖

3

LINEAR SYSTEMS OF ODE (I)

Unit Structure :

3.1 INTRODUCTION

We now consider following system of first order ordinary differential equations in the function

$x : t \mapsto x(t) = \quad x_1, (t), x_2(t), ..., x_n(t) \;$:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \\ &\vdots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n \end{aligned}\right\} \qquad \ldots (1)$$

where the coefficients a_{ij} appearing on the right hand sides of the system (1) are all constant real numbers. (The case in which $a_{ij} = a_{ij}(t), t \in \mathbb{R}$ will be discussed in the next chapter). Writing $X \;\; t$ in the columnal form :

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

and collecting the coefficients a_{ij} in a matrix A, that is $A = \left[a_{ij}\right] 1 \le i, j \le n$, we rewrite the system (1) in the matrix form :

$$\frac{dX}{dt} = A.X$$

For a given vector $w \in \mathbb{R}^n$ with $w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ we form the initial condition $X\ 0\ = w$. Thus, we now have the initial value problem :

$$\frac{dX}{dt} = A.X,\ X(0) = w \qquad \ldots (2)$$

Later on, we will consider an arbitrary $t_0 \in \mathbb{R}$ and the initial value problem :

$$\frac{dX}{dt} = A.X, \quad X(t_0) = w \qquad \ldots (2')$$

obtained by bringing t_0 in place of t = 0. The solution of this more general $IVP\ 2'$ is easily obtained from the solution of (2). Therefore, we treat the particular case (2) in detail first.

Here we are taking $f\ X\ = A.X$ in our model differential equation $\frac{dX}{dt} = f\ X$ (the first order, autonomous case) treating the matrix $A = \left[a_{ij}\right]$ as a linear transformation (= linear endomorphism) of $\mathbb{R}^n$, its action on a vector $x \in \mathbb{R}^n$ being given by

$$A.x = \begin{bmatrix} a_{11}, a_{12}, \ldots.\, a_{1n} \\ a_{21}, a_{22}, \ldots.\, a_{2n} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{n1}, a_{n2}, \ldots.\, a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots.\, a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots.\, a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots.\, a_{nn}x_n \end{bmatrix} \in \mathbb{R}^n.$$

$$\text{where } X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n.$$

Thus in our treatment, the symbol A is made to play a double rule : (i) A as the $n \times n$ matrix and (ii) A as a linear transformation (= linear transformation) of $\mathbb{R}^n$.

Note that when n = *1*, the system (1) reduces to the single differential equation $\frac{dx}{dt} = a\,x$.

The solution of the initial value problem $\frac{dx}{dt} = a\,x, \quad X(0) = w \;\; w \in \mathbb{R}$ in this one-dimensional case (i.e. n = 1) is the familiar function : $t \mapsto we^{ta}, t \in \mathbb{R}$. Recall :

$$e^{ta} = 1 + \frac{ta}{1!} + \frac{t^2a^2}{2!} + \frac{t^3a^3}{3!} + \ldots\ldots..(*)$$

The comparison between the on - dimensional initial value problem and its n-dimensional case suggests that we expect the solution of the IVP(2) to be a curve of the form

$X : \mathbb{R} \to \mathbb{R}^n, t \mapsto X(t) = e^{tA}.w\ldots\ldots(**)$ where e^{tA} is an $n \times n$ matrix which has the power series expansion :

$$I + \frac{tA}{1!} + \frac{t^2A^2}{2!} + \frac{t^3A^3}{3!} + \ldots\ldots..$$

suggested by (*) above. To carry forward the analogy, we call e^A, the **exponential** of A. We will define the new quantity e^A first. (Replacing A by tA for $t \in \mathbb{R}$, will then yield e^{tA}). Once this is accomplished, we will verify that the curve $X : t \mapsto X(t) = e^{tA}w$ is indeed the solution of the initial value problem (2).

3.2 THE EXPONENTIAL OF A LINEAR ENDOMORPHISM

Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear endomorphism having its matrix $\left[a_{ij}\right]$, we choose a finite constant C satisfying the inequality :

$$\|A(x)\| \leq C\|x\|$$

for all $x \in \mathbb{R}^n$ (e.g. $C = n^{3/2} \max \; \left|a_{ij}\right| : 1 \leq i, j \leq n$ will do the job). Note that the above inequality implies :

$\left\|A^k \;\; x \;\right\| \leq C^k \|x\|$ and $\left\|t^k A^k \;\; x \;\right\| \leq |t|^k \; C^k \|x\|$ for every $x \in \mathbb{R}^n$, every $t \in \mathbb{R}$ and every $k \in \mathbb{Z}^+$. In particular for the vector field $f : \mathbb{R}^n \to \mathbb{R}^n$ given by $f \;\; x \; = A.x, x \in \mathbb{R}^n$, we have :

$$\begin{aligned} \|f \;\; x \; - f(y)\| &= \|A \;\; x \; - A(y)\| \\ &= \|A \;\; x - y \;\| \\ &\leq C\|x - y\| \end{aligned}$$

for any *x, y*, in $\mathbb{R}^n$. In other words the vector field *f*(*x*)=*Ax* has the Lipschitz property. Consequently the initial value problem (2) has a unique solution.

Next, we define a map $B:\mathbb{R}^n \to \mathbb{R}^n$ as follows : Let $x \in \mathbb{R}^n$ be arbitrary. Then we have :

$$\|x\| + \frac{\|A(x)\|}{1!} + \frac{\|A^2(x)\|}{2!} + \frac{\|A^3(x)\|}{3!} + ...$$
$$\leq \|x\| + \frac{C\|x\|}{1!} + \frac{C^2\|x\|}{2!} + ...$$
$$= e^C \|x\| < \infty$$

This shows that the infinite sum :

$$x + \frac{A(x)}{1!} + \frac{A^2(x)}{2!} + \frac{A^3(x)}{3!} + ...$$

converges absolutely. We put :

$$B(x) = x + \frac{A(x)}{1!} + \frac{A^2(x)}{2!} + ...$$

Thus the map B; expressed in terms of *A*, is given by

$$B = I + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + ...$$

Note that each power A^k is a linear transformation of $\mathbb{R}^n$ and this implies the linearlity of B. in fact, for any a, b in $\mathbb{R}$, any *x, y* in $\mathbb{R}^n$, we have

$$B(ax+by) = (ax+by) + \frac{A}{1!}(ax+by) + \frac{A^2}{2!}(ax+by) + ...$$
$$= ax + by + \frac{a}{1!}A(x) + \frac{b}{2!}A(y) + \frac{a}{2!}A^2(x) + \frac{b}{2!}A^2(y) + ...$$
$$= a\left[x + \frac{A(x)}{1!} + \frac{A^2(x)}{2!} + ...\right] + b\left[y + \frac{A(y)}{1!} + \frac{A^2(y)}{2!} + ...\right]$$
$$= a\,B(x) + b\,B(y)$$

We adapt the notation e^A for B. occasionally we use the notation exp(A) for e^A. Thus $\exp(A)\,(x) = e^A(x) = x + \frac{A(x)}{1!} + \frac{A^2(x)}{2!} + ...$ for every $x \in \mathbb{R}^n$. Thus, each linear endomorphism A of $\mathbb{R}^n$ gives rise to the linear endomorphism e^A. Moreover, if t is any real number then *tA* is also a linear endomorphism and it gives rise to the exponential e^{tA} given by :

$$e^{tA} = I + \frac{tA}{1!} + \frac{t^2 A^2}{2!} + \frac{t^3 A^3}{3!} + ...$$

it being a linear endomorphism of $\mathbb{R}^n$ where

$$e^{tA} \ \ x \ = x + \frac{tA(x)}{1!} + \frac{t^2 A^2(x)}{2!} + \frac{t^3 A^3(x)}{3!} + ...$$

for every $x \in \mathbb{R}^n$.

Having defined the exponential e^A, we obtain the solution of the IVP(2) in terms of the exponentiation. Consider the map

$$X : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

given by $x \ \ t \ = e^{tA}(w)$

$$= w + \frac{tA(w)}{1!} + \frac{t^2 A^2(w)}{2!} + ... \qquad t \in \mathbb{R}$$

Because the infinite series defining each $X(t)$ converges absolutely we get that $t \mapsto X(t)$ is differentiable and the derivative $\frac{dX}{dt}(t)$ is obtained by termwise differentiation of the infinite series defining $X(t)$. Thus we have

$$\frac{dX(t)}{dt} = 0 + \frac{A(w)}{1!} + \frac{tA^2(w)}{1!} + \frac{t^2 A^3(w)}{2!} + ...$$

$$= A\left(w + \frac{tA(w)}{1!} + \frac{t^2 A^2(w)}{2!} + \frac{t^3 A^3(w)}{3!} + ...\right)$$

$$= A \ \ x(t)$$

Thus, we have : $\frac{dX\ (t)}{dt} = A \ \ x(t)$ for every $t \in \mathbb{R}$. Moreover we have : $X(0) = w + 0 + 0 ..... = w$.

This completes the proof that the map $X : \mathbb{R} \rightarrow \mathbb{R}^n$ given by $x \ \ t \ = e^{ta}(w)$ is a solution of the IVP(2). We summarize this observation in the following.

Theorem 1 : The map $X : \mathbb{R} \rightarrow \mathbb{R}^n$ given by $x \ \ t \ = e^{ta}(w)$ is the solution of the initial value problem $\frac{dX}{dt} = A.X, X(0) = w$. The proof of the following is self evident :

Corollary : The curve $X : \mathbb{R} \rightarrow \mathbb{R}^n$ given by $x \ \ t \ = e^{\ t - t_0 \ A}(w)$ for all $t \in \mathbb{R}$ is the solution of the initial value problem : $\frac{dx}{dt} = A.X, x(t_0) = w$.

3.3 PROPERTIES OF THE EXPONENTIAL

Following few properties help us sum the infinite series defining e^A and get the matrix of it.

I) If A is a diagonal matrix, say $A = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$

then we have $A^k = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}$

and therefore

$$e^A = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} + \frac{1}{1!}\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2 \end{bmatrix}$$

$$= \begin{bmatrix} \left(1+\lambda_1+\frac{\lambda_1^2}{2}+....\right) & & \\ & \left(1+\frac{\lambda_2}{1!}+\frac{\lambda_2^2}{2!}+....\right) & \\ & & \left(1+\frac{\lambda_n}{1!}+\frac{\lambda_n^2}{2!}+\frac{\lambda_n^3}{3!}+....\right) \end{bmatrix}$$

$$= \begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix}$$

II) If $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ then $e^A = e^a \begin{bmatrix} \cos b, -\sin b \\ \sin b \quad \cos b \end{bmatrix}$.

Proof : Let $\lambda = a + ib$ so that $a = \mathrm{Re}(\lambda), b = I_m(\lambda)$ and

$$A = \begin{bmatrix} \text{Re } \lambda & -I_m(\lambda) \\ I_m(\lambda) & \text{Re } \lambda \end{bmatrix}$$

Moreover, $A^2 = \begin{bmatrix} \text{Re } \lambda & -I_m(\lambda) \\ I_m(\lambda) & \text{Re } \lambda \end{bmatrix} \begin{bmatrix} \text{Re } \lambda & -I_m(\lambda) \\ I_m(\lambda) & \text{Re } \lambda \end{bmatrix}$

$$= \begin{bmatrix} \text{Re } \lambda \ \text{Re } \lambda - I_m(\lambda) - I_m(\lambda), -2\,\text{Re } \lambda \ I_m(\lambda) \\ 2\,\text{Re } \lambda \ I_m(\lambda), \text{Re } \lambda \ \text{Re } \lambda - I_m(\lambda) - I_m(\lambda) \end{bmatrix}$$

$$= \begin{bmatrix} \text{Re } \lambda^2 & -I_m(\lambda^2) \\ I_m(\lambda^2) & \text{Re } \lambda^2 \end{bmatrix}$$

In general $A^k = \begin{bmatrix} \text{Re } \lambda^k & -I_m(\lambda^k) \\ I_m(\lambda^k) & \text{Re } \lambda^k \end{bmatrix}$ holding for every $k \in \mathbb{Z}^+$.

Consequently :

$$e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \ldots.$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{1!}\begin{bmatrix} \text{Re } \lambda & -I_m(\lambda) \\ I_m(\lambda) & \text{Re } \lambda \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} \text{Re } \lambda^2 & -I_m(\lambda^2) \\ I_m(\lambda^2) & \text{Re } \lambda^2 \end{bmatrix} + \ldots.$$

$$= \begin{bmatrix} [\text{Re } 1 + \frac{\text{Re } \lambda}{1!} + \frac{\text{Re } \lambda^2}{2!} + \ldots.] & -[\frac{I_m(\lambda)}{1!} + \frac{I_m \ \lambda^2}{2!} + \frac{I_m \ \lambda^3}{3!} \cdots] \\ [\frac{I_m(\lambda)}{1!} + \frac{I_m \ \lambda^2}{2!} + \ldots\ldots.] & [\text{Re } 1 + \frac{1}{1!}\text{Re } \lambda + \frac{1}{2!}\text{Re } \lambda^2 + \ldots] \end{bmatrix}$$

$$= \begin{bmatrix} [\text{Re } 1 + \frac{\text{Re } \lambda}{1!} + \frac{\text{Re } \lambda^2}{2!} + \ldots.] & -[0 + I_m \frac{(\lambda)}{1!} + I_m \frac{\lambda^2}{2!} + \ldots\ldots.] \\ [0 + I_m \frac{(\lambda)}{1!} + I_m \frac{\lambda^2}{2!} + \ldots\ldots.] & [\text{Re } 1 + \text{Re}\frac{\lambda}{1!} + \text{Re}\frac{\lambda^2}{2!} + \ldots\ldots.] \end{bmatrix}$$

$$= \begin{bmatrix} [\text{Re } 1 + \frac{\text{Re } \lambda}{1} + \frac{\text{Re } \lambda^2}{2!} + \ldots.] & -[I_m \ 1 + I_m \frac{\lambda}{1!} + I_m \frac{\lambda^2}{2!} + \ldots] \\ [I_m(1) + I_m \frac{(\lambda)}{1!} + I_m \frac{\lambda^2}{2!} + \ldots] & [\text{Re } 1 + \text{Re } 1 \ \frac{(\lambda)}{1!} + \text{Re } 1 \ \frac{\lambda^2}{2!} + \ldots] \end{bmatrix}$$

$$= \begin{bmatrix} \text{Re}(1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+...) & -I_m(1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+...) \\ I_m(1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+...) & \text{Re}\ (1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+...) \end{bmatrix}$$

$$= \begin{bmatrix} \text{Re}\ e^{\lambda} & -I_m\ e^{\lambda} \\ I_m\ e^{\lambda} & \text{Re}\ e^{\lambda} \end{bmatrix}$$

$$= \begin{bmatrix} e^a \cos b & -e^a \sin b \\ e^a \sin b & e^a \cos b \end{bmatrix}$$

$$= e^a \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}$$

III)If A and B are linear endomorphisms of $\mathbb{R}^n$ with the property $A \circ B = B \circ A$, then $e^{A+B} = e^A \circ e^B = e^B \circ e^A$.

Proof :

The classical binomial theorem applies to the powers $A+B^{\ k}; k \in \mathbb{Z}^+$:

$$A+B^{\ K} = \sum_{J=0}^{K} \binom{k}{j} A^j B^{k-j} .$$

Therefore $e^{A+B} = \sum_{m=0}^{\infty} \frac{1}{m!}\ A+B^{\ m} = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k=0}^{\infty} \frac{m!}{k!(m-k)!} \text{A}_0^k B^{m-k}$

$$= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{A^k}{k!} \frac{B^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty} \sum_{k+\ell=m}^{\infty} \frac{A^k}{k!} \frac{B^{\ell}}{e!}$$

$$= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{A^k}{k!} \frac{B^{\ell}}{e!} = \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} \right) \circ \left(\sum_{\ell=0}^{\infty} \frac{B^{\ell}}{e!} \right)$$

$$= e^A \circ e^B$$

It can be proved on similar lines that $e^{A+B} = e^B \circ e^A$

IV) Let $A = aI + B$ where a is a real number and B is a strictly upper triangular $n \times n$ matrix.

$$\begin{bmatrix} 0 & b_{12} & & b_{1n} \\ 0 & b_{23} & & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{n-1\,n-1} \\ & & & 0 \end{bmatrix}$$

So that $B^n = 0$ and $aI \circ B = B \circ aI$. Then we have

$$e^A = e^a \left(I + \frac{B}{1!} + \frac{B^2}{2!} + ... + \frac{B^{n-1}}{(n-1)!} \right).$$

V) Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be any linear transformation and let $B : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Then $e^{B \circ A \circ B^{-1}} = B \circ \ e^A \ \circ B^{-1}$.

Proof : For each $k \in \mathbb{Z}^+$ we have $(BAB^{-1})^k = B \circ A^k \circ B^{-1}$ and therefore :

$$\begin{aligned} e^{B \circ A \circ B^{-1}} &= I + \frac{1}{1!} B \circ A \circ B^{-1} + \frac{1}{2!} (B \circ AB^{-1})^2 + ... \\ &= B \circ I \circ B^{-1} + \frac{1}{1!} B \circ AB^{-1} + \frac{1}{2!} B \circ A^2 \circ B^{-1} + ... \\ &= B \circ \left(I + \frac{A}{1!} + \frac{A^2}{2!} + ... \right) \circ B^{-1} \\ &= B \circ e^A \circ B^{-1} \end{aligned}$$

VI) We recollect here a few elementary facts of linear algebra culminating in a formula relating two sets of coordinates on $\mathbb{R}^n$. These results will be used in a conjunction with (IV) above to solve systems of linear ODE.

Let a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ have all real and distinct eigen-values $\lambda_1, \lambda_2 .... \lambda_n$ with respective eigen-vectors $f_1, f_2 .... f_n : A(f_i) = \lambda_i f_i \quad 1 \le i \le n$. Now, λ_i are all distinct implies that the set $\{f_1, f_2 .... f_n\}$ is a vector basis of $\mathbb{R}^n$. Thus we have two vector bases of $\mathbb{R}^n$ now :

i) The standard vector basis $\{e_1, e_2 .... e_n\}$ with $e_i = (0, ...0, 1, 0....0)$ and

ii) The $\{f_1, f_2 .... f_n\}$ consisting of the eigen-vectors of A.

Let the linear transformation $B:\mathbb{R}^n \to \mathbb{R}^n$ be given by $B\ e_i\ = f_i\ 1 \le i \le n$. Clearly B is invertible. Putting $f_j = \sum_{i=1}^{n} b_{ij} e_i$ we get the matrix $\left[b_{ij}\right]$ it is the matrix of the linear map B (with respect to the standard basis $e_1, ... e_n$ of $\mathbb{R}^n$).

Now we note that the linear transformation $B^{-1} \circ A \circ B$ has the set $e_1, ... e_n$ as its eigen vectors with the respective eigen-values, $\lambda_i, ... \lambda_n$. Consequently the matrix of $B^{-1} \circ A \circ B$ with respect to the standard basis $e_1, ... e_n$ is diagonalized.

$$\left[B^{-1} A \circ B\right] = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Let $y_1, y_2 .... y_n$ be the coordinates on $\mathbb{R}^n$ determined by the vector basis $f_1, f_2 .... f_n$. As usual, $x_1, x_2 .... x_n$ are the Cartesian coordinates of $\mathbb{R}^n$ - they are the coordinates determined by the standard basis $e_1, ... e_n$ $\mathbb{R}^n$. Now we have :

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \left[B^{-1}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Where $\left[B^{-1}\right]$ is the matrix $\left[b_{ij}\right]^{-1}$

Examples :

In this section we use the theory developed in the preceding sections to solve linear systems of differential equations.

Example 1 :

a) Solve the following IVP.

$$\frac{dx}{dt} = 2x + y + z \qquad x(0) = 1$$
$$\frac{dy}{dt} = 2y + 2z \qquad y(0) = 2$$
$$\frac{dz}{dt} = 2z \qquad z(0) = 3$$

b) The same system of differential equations but take the initial conditions : $x\ 1\ = 1, y(1) = 2, z(1) = 3$.

Solution :

a) We rewrite the IVP in the form :

$$\frac{d}{dt}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}, \qquad \begin{bmatrix} x(0) \\ y(0) \\ z(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

We have : $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix} = 2I + \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

Next, note that $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and

$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}^k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ for all $k \geq 3$.

Therefore, we have

$$\exp\left(t\begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}\right) = e^{2t}\left(I + t\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} + \frac{t^2}{2}\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)$$
$$= e^{2t}\begin{bmatrix} 1 & t & t+t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{bmatrix}$$

According to the **Theorem 1** we have :

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = e^{2t} \begin{bmatrix} 1 & t & t+t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = e^{2t} \begin{bmatrix} 1+5t+3t^2 \\ 2+6t \\ 3 \end{bmatrix}$$

That is $x(t) = e^{2t}(1+5t+3t^2)$

$$y(t) = e^{2t}(2+6t)$$

$$z(t) = 3e^{2t}$$

To get the solution of (b) we apply the corollary to the **Theorem 1 w**hich suggests that the variable t in above is to be replaced by $t - 1$ this gives the solution of (b).

$$x(t) = e^{2(t-1)}\left[1+5(t-1)+3(t-1)^2\right]$$

$$= e^{2(t-1)}\left[3t^2 - t - 1\right]$$

$$y(t) = e^{2(t-1)}\left[2+6(t-1)\right]$$

$$= e^{2(t-1)}\left[6t^2 - 4\right]$$

$$z(t) = 3e^{2(t-1)}$$

<u>Example 2 :</u> Solve

$$\frac{dx}{dt} = 5x - 2y, \quad x(0) = 2$$

$$\frac{dy}{dt} = 2x + 5y, \quad y(0) = 3$$

Solution : We have :

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \qquad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

According to property (II) of **Section 3.3**, we have

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \exp\left(t\begin{bmatrix} 5 & -2 \\ 2 & 5 \end{bmatrix}\right)\begin{bmatrix} 2 \\ 3 \end{bmatrix} = e^{5t}\begin{bmatrix} \cos 2t, & -\sin 2t \\ \sin 2t, & \cos 2t \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Therefore $x(t) = e^{5t}(2\cos 2t - 3\sin 2t)$

$$y(t) = e^{5t}(2\sin 2t + 3\cos 2t)$$

Example 3 : Solve

$$\frac{dx}{dt} = x + 3z, \quad x(0) = 1$$
$$\frac{dy}{dt} = 4y, \quad y(0) = 2$$
$$\frac{dz}{dt} = -3z + z, \quad z(0) = 4$$

Solution : Because the middle equation is independent of *x, z,* we solve it (taking into account the initial condition on it). This gives $y\ t\ = 2e^{4t}$.

Next, we deal with the coupled pair of the remaining equation and the initial conditions on them :

$$\frac{d}{dt}\begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} x \\ z \end{bmatrix} \quad \begin{bmatrix} x(0) \\ z(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

This gives the solution :

$$\begin{bmatrix} x\ t \\ y\ t \end{bmatrix} = \mathrm{e}^{t}\begin{bmatrix} \cos(-3t) - \sin(-3t) \\ \sin(-3t) \quad \cos(-3t) \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix}$$
$$= e^{t}\begin{bmatrix} \cos 3t, & \sin 3t \\ -\sin 3t, & \cos 3t \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix}$$
$$= e^{t}\begin{bmatrix} \cos 3t + 4 & \sin 3t \\ -\sin 3t + 4 & \cos 3t \end{bmatrix}$$

That is $x\ t\ = e^{t}\ \cos 3t + 4\sin 3t$

$z\ t\ = e^{t}\ 4\cos 3t - \sin 3t$.

Putting together all of them, we get

$x\ t\ = e^{t}\ \cos 3t + 4\sin 3t$

$y\ t\ = 2e^{4t}$

$z\ t\ = e^{t}\ 4\cos 3t - \sin 3t$

3.4 EXERCISES

1) Compute the exponential of each of the matrixes.
(i) (ii) (iii)

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 4 \end{bmatrix} \qquad \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{bmatrix}$$

2) Obtain the matrix for $e^{tb}, t \in \mathbb{R}$ for the B given below.

(i)

$$B = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & -4 & 2 \end{bmatrix}$$

(ii)

$$B = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

3) Solve the following initial value problems.

i)

$$\frac{dx}{dt} = 3x + 4y \quad x\ 1 = 2$$

$$\frac{dy}{dt} = -4x + 3y \quad y\ 1 = 3$$

$$\frac{dz}{dt} = yz + 3 \quad z\ 1 = 4$$

ii)

$$\frac{dx}{dt} = 4x + 3y + 2z \quad x\ 1 = 2$$

$$\frac{dy}{dt} = 4y + z \quad y\ 1 = 5$$

$$\frac{dz}{dt} = 4z \quad z\ 1 = 6$$

iii)

$$\frac{dx}{dt} = 3x + 2y \quad x\ 1 = 4$$

$$\frac{dy}{dt} = -2x + 3y \quad y\ 1 = 3$$

4) Solve the initial value problem :

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 & -2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad \begin{matrix} x_1(0) = 1 \\ x_2(0) = 1 \\ x_3(0) = 1 \\ x_4(0) = 1 \end{matrix}$$

❖❖❖❖

4

LINEAR DIFFERENTIAL EQUATIONS

Unit structure :

4.1 OBJECTIVES

The learner will be able to:

- solve homogeneous and non-homogeneous second order differential equations.
- check linear dependence and linear independence of set of functions.
- learn definition of Wronskian and apply Wronskian to check linear dependence of set of functions.
- prove Abel's Identity and solve homogeneous and non-homogeneous n^{th} order differential equations.

4.2 INTRODUCTION

A second order linear differential equation has the general form

$$P(x)\frac{d^2y}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y = G(x) ----(4.2.1)$$

where P, Q, R and G are continuous functions.

There are two types of second order linear differential equation:

(a) Homogeneous linear equations: If in (4.2.1), $G(x) = 0$ then the differential equation is called Homogeneous linear differential equation.

(b) Non-homogeneous linear equations: If in (4.2.1), $G(x) \neq 0$ then the differential equation is called Non-homogeneous linear differential equation.

4.3 THE SECOND ORDER HOMOGENEOUS EQUATION

Consider the homogeneous equation of the form

$$P(x)\frac{d^2y}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y = 0 - - - - (4.3)$$

We say that $y^*(x)$ is a solution of (4.3) if it satisfies (4.3) that is

$$P(x)\frac{d^2y^*(x)}{dx^2} + Q(x)\frac{dy^*(x)}{dx} + R(x)y^*(x) = 0$$

Theorem 4.3: *If* $y_1(x)$ *and* $y_2(x)$ *are two solutions of the linear homogeneous equation* (4.3) *and if* c_1 *and* c_2 *are any constants, then*

$$y^*(x) = c_1y_1(x) + c_2y_2(x)$$

is also a solution of (4.3).

Proof. Since $y_1(x)$ and $y_2(x)$ are two solutions of (4.3), we have

$$P(x)\frac{d^2y_1(x)}{dx^2} + Q(x)\frac{dy_1(x)}{dx} + R(x)y_1(x) = 0$$

and

$$P(x)\frac{d^2y_2(x)}{dx^2} + Q(x)\frac{dy_2(x)}{dx} + R(x)y_2(x) = 0$$

Consider

$$P(x)\frac{d^2y^*(x)}{dx^2} + Q(x)\frac{dy^*(x)}{dx} + R(x)y^*(x)$$

$$= P(x)\frac{d^2(c_1y_1(x) + c_2y_2(x))}{dx^2} + Q(x)\frac{d(c_1y_1(x) + c_2y_2(x))}{dx} + R(x)(c_1y_1(x) + c_2y_2(x))$$

$$= P(x)\left[c_1\frac{d^2y_1(x)}{dx^2} + c_1\frac{d^2y_2(x)}{dx^2}\right] + Q(x)\left[c_1\frac{dy_1(x)}{dx} + c_1\frac{dy_2(x)}{dx}\right] + R(x)[c_1y_1(x) + c_2y_2(x)]$$

$$= c_1\left[P(x)\frac{d^2y_1(x)}{dx^2} + Q(x)\frac{dy_1(x)}{dx} + R(x)y_1(x)\right] + c_2\left[P(x)\frac{d^2y_1(x)}{dx^2} + Q(x)\frac{dy_1(x)}{dx} + R(x)y_1(x)\right]$$

$$= 0$$

Thus $c_1y_1(x) + c_2y_2(x)$ is also a solution of (4.3).

Remark 4.3: *If* $y_1(x)$ *and* $y_2(x)$ *are two linearly independent solutions of* (4.3) *and* $P(x)$ *is never* 0 *then the general solution is given by* $c_1y_1(x) + c_2y_2(x)$ *where* c_1 *and* c_2 *are arbitary constants.*

4.3.1 HOMOGENEOUS EQUATIONS WITH CONSTANT COEFFICIENTS

The equation of the form

$$A\frac{d^2y}{dx^2} + B\frac{dy}{dx} + Cy = 0 - - - - (4.3.1)$$

where A, B and C are constants and $A \neq 0$ is called homogeneous equations with constant coefficients.

Let $y = e^{kx}$ then substituting y in (4.3.1) we get

$$e^{kx}(ak^2 + bk + c) = 0$$

As $e^{kx} \neq 0$ we get that $y = e^{kx}$ is the solution of (4.3.1) if k is the root of

$$ak^2 + bk + c = 0 - - - - (4.3.2)$$

The equation (4.3.2) is called the auxiliary equation (or characteristic equation) of (4.3.1). The equation (4.3.2) is an algebraic equation in k whose solution can be found using the formula.

$$k_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad k_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Depending on the sign of $\sqrt{b^2 - 4ac}$ we have the following cases:

Case (i): $b^2 - 4ac > 0$

In this case the equation (4.3.2) has two real and distinct roots k_1 and k_2 and the general solution of (4.3.1) is given by $y = c_1e^{k_1x} + c_1e^{k_2x}$.

Case (ii): $b^2 - 4ac = 0$

In this case the equation (4.3.2) has only one distinct real root say k and the general solution of (4.3.1) is given by $y = c_1e^{kx} + c_1xe^{kx}$.

Case (iii): $b^2 - 4ac < 0$

In this case the equation (4.3.2) has two complex roots say $k_1 = a + ib$ and $k_2 = a - ib$ and the general solution of (4.3.1) is given by $y = e^{ax}(c_1 \cos bx + c_2 \sin bx)$.

Example 4.3.1.1. Solve the equation

$$y'' + y' - 6y = 0$$

Solution. Here the auxiliary equation is given as $k^2 + k - 6 = 0$

Thus, we have two distinct real roots $k_1 = 2, k_2 = -3$. Hence the general solution of the given equation is

$$y = c_1e^{2x} + c_2e^{-3x}$$

Example 4.3.1.2. Solve the equation

$$4y'' + 12y' + 9y = 0$$

Solution. Here the auxiliary equation is given as $4k^2 + 12k + 9 = 0$

Thus, we have $(2k + 3)^2 = 0$ which gives only one real root $k = -\frac{3}{2}$. Hence the general solution of the given equation is

$$y = c_1e^{-\frac{3}{2}x} + c_2xe^{-\frac{3}{2}x}$$

Example 4.3.1.3. Solve the equation

$$y'' - 6y' + 13y = 0$$

Solution. Here the auxiliary equation is given as $k^2 - 6k + 13 = 0$ which has no real roots. The complex roots of the auxiliary equation are given by

$$k_1 = 3 + 2i, \quad k_2 = 3 - 2i$$

and the general solution is given by

$$y = e^{3x}(c_1 \cos 2x + c_2 \sin 2x).$$

4.3.2 INITIAL VALUE PROBLEM FOR SECOND ORDER EQUATIONS

An initial value problem for second order equation (4.2.1) is to find a solution of the equation which also satisfies the initial condition $y(x_0) = y_0$ and $y'(x_0) = y_1$ where y_0 and y_1 are given constants.

For example, to find the solution of the initial value problem $y'' + y' - 6y = 0$ with $y(0) = 1$ and $y'(0) = 0$, we first find the root of the auxiliary equation $k^2 + k - 6 = 0$. From example (4.3.1.1) we know that the general solution is given by

$$y = c_1 e^{2x} + c_2 e^{-3x}$$

To satisfy the initial condition, we need

$$y(0) = c_1 + c_2 = 0$$

and

$$y'(0) = 2c_1 - 3c_2 = 0$$

which gives $c_1 = \frac{3}{5}$ and $c_2 = \frac{2}{5}$. Hence the solution of the initial value problem is given by

$$y = \frac{3}{5} e^{2x} + \frac{2}{5} e^{-3x}$$

Theorem 4.3.2[Uniqueness Theorem]: *If $P(x)$ and $Q(x)$ are continuous functions on an open interval I containing x_0, then the equation*

$$y'' + P(x)y' + Q(x)y = R(x)$$

with initial condition $y(x_0) = y_0$ and $y'(x_0) = y_1$ has a unique solution.

Proof. Let y and y^* be any two solutions of the equation $y'' + P(x)y' + Q(x)y = R(x)$ with initial conditions $y(x_0) = y_0$ *and* $y'(x_0) = y_1$. Let $m = y - y^*$. Then

$$m'' + P(x)m' + Q(x)m = 0$$

and $m(x_0) = m'(x_0) = 0$

Consider, $E(x) = m^2(x) + \left(m'(x)\right)^2$. Then $E(x) \geq 0$ and $E(x_0) = 0$. Differentiating $E(x)$ we get

$$\begin{aligned} E'(x) &= 2m(x)m'(x) + 2m'(x)m''(x) \\ &= 2m'(x)[m(x) + m''(x)] \\ &= 2m'(x)[m(x) - P(x)m'(x) - Q(x)m(x)] \\ &= -2P(x)\left(\left(m'(x)\right)^2\right) + 2m(x)m'(x)[1 - Q(x)] \end{aligned}$$

By Cauchy-Schwartz inequality,

$$m(x)m'(x)[1 - Q(x)] \leq (1 + |Q(x)|)\left(m^2(x) + \left(m'(x)\right)^2\right)$$

Thus,

$$\begin{aligned} E'(x) &\leq -2P(x)\left(\left(m'(x)\right)^2\right) + 2(1 + |Q(x)|)\left(m^2(x) + \left(m'(x)\right)^2\right) \\ &= (1 + |Q(x)|)m^2(x) + (1 + |Q(x)| + 2|P(x)|)\left(m'(x)\right)^2 \end{aligned}$$

Let $K \geq 1 + \max_{x\in I}\{|Q(x)| + 2|P(x)|\}$. Then

$$E'(x) \leq KE(x)$$

Claim: $E(x) = 0$ for all $x \in I$

Suppose, there exist some $x_1 \in I$ such that $E(x_1) > 0$ with $x_1 > x_0$. Then

$$\frac{d}{dx}\left(e^{-Kx}E(x)\right) = e^{-Kx}\left(E'(x) - KE(x)\right) \leq 0$$

Thus, $e^{-Kx}E(x)$ is a decreasing function and hence $e^{-Kx_1}E(x_1) \leq e^{-Kx_0}E(x_0) = 0$. This implies that $E(x_1) \leq 0$ which is a contradiction.
Thus, we get $E(x) = 0$ for all $x \in I$ which implies that $m(x) = 0$.

4.4 LINEAR DEPENDENCE AND INDEPENDENCE OF SOLUTIONS

A set of solutions $\{f_1, f_2, \cdots, f_n\}$ of a differential equation is linearly independent on an interval I if and only if the only values of the scalars $c_1, c_2, \cdots, c_n$ such that

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0, \quad \forall\, x \in I$$

are $c_1 = c_2 = \cdots = c_n = 0$.
A set which is not linearly independent is called a linearly dependent set.

4.4.1 WRONSKIAN, A FORMULA FOR THE WRONSKIAN

Let $f_1, f_2, \cdots, f_n$ be functions in $C^{n-1}(I)$ where I is an interval. Then Wronskian of these functions is defined by

$$W[f_1, f_2, \cdots, f_n](x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

For example, if $f_1(x) = \sin x$ and $f_2(x) = \cos x$ on $(-\infty, \infty)$ then

$$W[f_1, f_2](x) = \begin{vmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{vmatrix} = -1$$

Theorem 4.4.1: *Let $f_1, f_2, \cdots, f_n$ be functions in $C^{n-1}(I)$ where I is an interval. If $W[f_1, f_2, \cdots, f_n]$ is non-zero at some point $x_0 \in I$ then $\{f_1, f_2, \cdots, f_n\}$ is linearly independent on I.*
Proof. Consider $c_1, c_2, \cdots, c_n$ such that

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0$$

for all $x \in I$.
Differentiating $n - 1$ times, we get the system of equation

$$\begin{gathered} c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0, \\ c_1 f_1'(x) + c_2 f_2'(x) + \cdots + c_n f_n'(x) = 0, \\ \vdots \\ c_1 f_1^{n-1}(x) + c_2 f_2^{n-1}(x) + \cdots + c_n f_n^{n-1}(x) = 0 \end{gathered}$$

where the unknowns are $c_1, c_2, \cdots, c_n$. The determinant of the coefficients of this system is the Wronskian of the functions $f_1, f_2, \cdots, f_n$. We know that if the determinant of the coefficients of a homogeneous system is non-zero, then the system has a unique solution.
Thus if $W[f_1, f_2, \cdots, f_n](x_0) \neq 0$ then the homogeneous system has a unique solution which is $c_1 = c_2 = \cdots = c_n = 0$.

4.4.2 ABEL'S IDENTITY

Let $y_1(x)$ and $y_2(x)$ be two solutions of $y'' + P(x)y' + Q(x)y = 0$. Then the Wronskian is given by $y_1(x)y_2'(x) - y_1'(x)y_2(x)$. Then

$$W'(x) = y_1(x)y_2''(x) - y_1''(x)y_2(x)$$

Since y_1 and y_2 are solutions of the given equation, we have

$$y_1''(x) = -P(x)y_1'(x) - Q(x)y_1(x)$$

and

$$y_2''(x) = -P(x)y_2'(x) - Q(x)y_2(x)$$

Substituting this in $W'(x)$ we get

$$W'(x) = -P(x)W(x)$$

That is the Wronskian satisfies the first order linear equation

$$W'(x) + P(x)W(x) = 0$$

which on solving gives

$$W(x) = W(x_0)e^{\int_{x_0}^{x} P(t)dt}$$

This formula is known as the Abel's Identity.

4.5 THE SECOND ORDER NONHOMOGENEOUS EQUATIONS

Theorem 4.5.1: *If $y_p(x)$ is any solution of $P(x)y'' + Q(x)y' + R(x)y = S(x)$. Then $y^*(x) = y_c(x) + y_p(x)$ is the general solution of $P(x)y'' + Q(x)y' + R(x)y = S(x)$ where $y_c(x)$ is the general solution of $P(x)y'' + Q(x)y' + R(x)y = 0$.*

Proof. Consider

$$\begin{aligned} P(x)y^{*''} + Q(x)y^{*'} + R(x)y^* &= P(x)\left(y_c''(x) + y_p''(x)\right) + Q(x)\left(y_c'(x) + y_p'(x)\right) \\ &\quad + R(x)\left(y_c(x) + y_p(x)\right) \\ &= \left(P(x)y_c''(x) + Q(x)y_c'(x) + R(x)y_c(x)\right) \\ &\quad + \left(P(x)y_p''(x) + Q(x)y_p'(x) + R(x)y_p(x)\right) \\ &= 0 + S(x) \\ &= S(x) \end{aligned}$$

Thus, $y^*(x)$ is the solution of $P(x)y'' + Q(x)y' + R(x)y = 0$. Here $y_c(x)$ is called the complementary function.

Remark: Thus, to solve a non-homogeneous differential equation, we just need to find the particular solution $y_p(x)$.

The following table gives a list of particular solutions for some familiar $S(x)$.

$S(x)$	$y_p(x)$
ke^{ax}	Ce^{ax}
$k_0 + k_1x + \cdots + k_nx^n$	$C_0 + C_1x + \cdots + C_nx^n$
$k\cos ax$	$C\cos ax + D\sin ax$
$k\sin ax$	$C\cos ax + D\sin ax$

Example 4.5.1.1: Solve the equation

$$y'' + 3y' + 2y = 4x^2 + 1$$

Solution. We first find $y_c(x)$. Consider the auxiliary equation

$$k^2 + 3k + 2 = 0$$
$$(k+2)(k+1) = 0$$

Hence the complementary function is $y_c = C_1e^{-2x} + C_2e^{-x}$.
Now, from the above table the particular solution is $y_p(x) = Ax^2 + Bx + C$ where the constants A, B and C are to be determined.
Consider

$$y_p'' + 3y_p' + 2y_p = 2Ax^2 + (6A + 2B)x + (2A + 3B + 2C) = 4x^2 + 1$$

Then comparing the coefficients of powers of x we get

$$2A = 4$$
$$(6A + 2B) = 0$$
$$(2A + 3B + 2C) = 1$$

which on solving gives $A = 2, B = -6, C = \frac{15}{2}$. Hence the general solution of the given equation is

$$y = C_1e^{-2x} + C_2e^{-x} + \left(2x^2 - 6x + \frac{15}{2}\right)$$

4.6 THE HOMOGENEOUS EQUATIONS OF ORDER n

A linear differential equation of order n is an equation of the form

$$P_n(x)y^{(n)} + P_{n-1}(x)y^{(n-1)} + \cdots + P_1(x)y' + P_0(x)y = 0 - - - (4.6)$$

is called a Homogeneous Equation of order n.
Remarks:

1. If f is a solution to (A) then Cf is also a solution to (4.6).
2. If f and g are the solutions to (A) then $C_1f + C_2f$ is also a solution to (4.6).
3. If $f_1, f_2, \cdots, f_k$ are the solutions to (4.6) then $C_1f_1 + C_2f_2 + \cdots + C_kf_k$ is also a solution to (4.6).

Example 4.6.1: Find the general solution of $y''' + 4y'' - 7y' - 10y = 0$.
Solution. Here the auxiliary equation is given by

$$k^3 + 4k^2 - 7k - 10 = 0$$

whose roots are given by $k_1 = -1, k_2 = 2, k_3 = -5$. Thus, the general solution is given by

$$y = C_1e^{-x} + C_2e^{2x} + C_3e^{-5x}$$

where C_1, C_2, C_3 are arbitrary constants.

4.7 INITIAL VALUE PROBLEM FOR n^{th} ORDER EQUATIONS

An initial value problem for n^{th} order equation (4.6) is to find a solution of the equation which also satisfies the initial condition $y(x_0) = y_0$, $y'(x_0) = y_1, \cdots, y^{(n-1)}(x_0) = y_{n-1}$ where $y_0, y_1, \cdots, y_{n-1}$ are given constants.

Example 4.7.1: Find the solution of the initial value problem $y''' + 4y'' - 7y' - 10y = 0$ with $y(0) = -3, y'(0) = 12$ and $y''(0) = -36$.
Solution. Here, the general solution is given by

$$y = C_1e^{-x} + C_2e^{2x} + C_3e^{-5x}$$

Now, we use the initial conditions to determine the constants C_1, C_2, C_3. Thus we get

$$C_1 + C_2 + C_3 = -3$$
$$-C_1 + 2C_2 - 5C_3 = 12$$
$$C_1 + 4C_2 + 25C_3 = -36$$

which on solving gives $C_1 = -\frac{5}{2}, C_2 = 1, C_3 = -\frac{3}{2}$. Hence the solution to the initial value problem is

$$y = -\frac{5}{2}e^{-x} + e^{2x} - \frac{3}{2}e^{-5x}$$

4.8 THE NONHOMOGENEOUS EQUATIONS OF ORDER n

General form of the n^{th} order non-homogeneous equation is given by
$P_n(x)y^{(n)} + P_{n-1}(x)y^{(n-1)} + \cdots + P_1(x)y' + P_0(x)y = f(x) - - -$ (4.8)
where $f(x)$ is a given function.
Similar to second order non-homogeneous, the general solutions of (4.8) is also given as

$$y = y_c + y_p$$

where y_c is the general solution of corresponding homogeneous equation and y_p is the particular solution of (4.8).
Example 4.8.1: Solve $y''' + 4y'' - 7y' - 10y = 100x^2 - 64e^{3x}$
Solution. Here the general solution of $y''' + 4y'' - 7y' - 10y = 0$ is given by

$$y_c = C_1e^{-x} + C_2e^{2x} + C_3e^{-5x}$$

Now for $f(x) = 100x^2 - 64e^{3x}$ we consider a trial function similar to the structure of $f(x)$ as

$$y_p = A_0 + A_1x + A_2x^2 + Be^{3x}$$

Substituting this to the given non-homogeneous equation, we get

$$(-10A_0 - 7A_1 + 8A_2) + (-10A_1 - 14A_2)x - 10A_2x^2 + 32Be^{3x}$$
$$= 100x^2 - 64e^{3x}$$

On comparing the coefficients, we get

$$(-10A_0 - 7A_1 + 8A_2) = 0$$
$$(-10A_1 - 14A_2) = 0$$
$$-10A_2 = 100$$
$$32B = -64$$

which on solving gives

$$A_0 = -\frac{89}{5}, A_1 = 14, A_2 = -10, B = -2$$

Hence the general solution is

$$y = C_1e^{-x} + C_2e^{2x} + C_3e^{-5x} + \left(-\frac{89}{5} + 14x - 10x^2 - 2e^{3x}\right)$$

4.9 EXERCISE

1. Solve $y''' + 4y'' - y' - 4y = 0$
2. Solve $y''' + 4y'' - 3y' - 18y = 0$
3. Solve $y''' + 6y'' + 12y' + 8y = 0$
4. Solve $y''' - 5y'' - y' + 5y = 10x - 63e^{-2x} + 29\sin 2x$
5. Solve $y^{(4)} + 8y'' + 16y = 64x\sin 2x$
6. Solve the initial value problem: $y^{(4)} + 2y''' - 2y' - y = 24xe^{-x} + 24e^{x} - 8\sin x$, with $y(0) = -2, y'(0) = 0, y''(0) = 6, y'''(0) = 10$

4.10 SUMMARY

In this unit, solving homogeneous and non-homogeneous second order differential equations are discussed. Linear dependence and linear independence of set of functions is discussed. Definition of Wronskian and applications of Wronskian to check linear dependence of set of functions is discussed. Abel's Identity is discussed and methods to solve homogeneous and non-homogeneous n^{th} order differential equations are also discussed in this unit.

References :

1. Differential Equations with Applications and Historical Notes by G.F. Simmons.
2. Advanced Differential Equations by M.D. Raisinghania.
3. ODE and PDE by M.D. Raisinghania.
4. Schaum's outline of Differential Equations by Richard Bronson.

*****

5

LINEAR SYSTEMS OF ODE (II)

Unit Structure :

5.1 INTRODUCTION

We consider a generalization of the type of systems studied in the preceding chapter. The new systems to be studied will be inhomogeneous, linear, first order systems with time dependant coefficients.

Throughout the chapter, I denotes an open interval. We consider a family of continuous functions :

$$a_{ij} : I \to \mathbb{R}, \quad 1 \le i, j \le n$$
$$u_i : I \to \mathbb{R}, \quad 1 \le i \le n$$

This family gives rise to the following system of non-homogeneous ODE :

$$\left.\begin{aligned} \frac{dX_1}{dt} &= a_{11}(t)X_1 + ... + a_{1n}(t)X_n + u_1(t) \\ \frac{dX_2}{dt} &= a_{21}(t)X_1 + ... + a_{2n}(t)X_n + u_2(t) \\ &\vdots \\ \frac{dX_n}{dt} &= a_{n1}(t)X_1 + ... + a_{nn}(t)X_n + u_n(t) \end{aligned}\right\} \quad \dots (1)$$

We also consider the same system but without the $u_i(t)$:

$$\left.\begin{aligned}\frac{dX_1}{dt} &= a_{11}(t)X_1 + ... + a_{1n}(t)X_n \\ \frac{dX_2}{dt} &= a_{21}(t)X_1 + ... + a_{2n}(t)X_n \\ &\vdots \\ \frac{dX_n}{dt} &= a_{n1}(t)X_1 + ... + a_{nn}(t)X_n\end{aligned}\right\} \quad ...(2)$$

We call (2) the homogeneous part of the system (1).

Our method of obtaining solutions of (1) consists of obtaining (i) a particular solution of the inhomogeneous system (1), then (ii) obtain the space of all solutions of the homogeneous system (2) and then combine (i) and (ii) to get all the solutions of the system (1).

We use the following abridged notations to which the operations of linear algebra will be applicable.

For each $t \in I, A(t)$ is the $n \times n$ matrix $\left[a_{ij}(t)\right], u(t)$ denotes the column $\begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix}$ and as usual $X(t)$ is the column $\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$.

In terms of these notations the systems (1) and (2) take the following compact forms :

$$\frac{dx}{dt} = A(t)\, X + u(t) \quad \ldots (1)$$

$$\frac{dx}{dt} = A(t)\, X \text{, the homogenous part of the above} \quad \ldots (2)$$

5.2 THE INITIAL VALUE PROBLEM

Given $t_0 \in I, x_0 \in \mathbb{R}^n$, we consider the IVP :

$$\frac{dx}{dt} = A(t)\, X + u(t) \ldots X(t_0) = x_0 \quad \ldots (3)$$

Note that the vector field $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ given by $f(t,x) = A(t)x + y(t)$ is locally Lipschitz :

Justification : Let $t_0 \in I, x_0 \in \mathbb{R}^n$ be arbitrary. Choose $\delta > 0$ such that $[t_0 - \delta, t_0 + \delta] \subseteq I$. Now the map $A : I \to \mathrm{M}_n(\mathbb{R})$ being continuous on its domain I, is bounded one the compact interval $t \in [t_0 - \delta, t_0 + \delta]$ and for any x, y in $B(x_0, \delta)$ we get.

$$\begin{aligned} f(t,x) - f(t,y) &= A(t)\,x + u(t) - A(t)\,y - u(t) \\ &= A(t)\,(x - y) \end{aligned} \quad \text{and therefore}$$

$$\begin{aligned} \|f(t,x) - f(t,y)\| &= \|A(t)(x - y)\| \\ &\leq \|A(t)\|.\|x - y\| \\ &\leq K\,\|x - y\| \end{aligned}$$

For all $t \in [t_0 - \delta, t_0 + \delta]$ and for all x, y in $B(x_0, \delta)$.

Therefore, the basic existence and uniqueness results are applicable. The IVP (3) has a unique solution defined on the largest open sub-interval (α, β) of I. We prove that $(\alpha, \beta) = \mathrm{I}$.

Proposition 1 :

The solution of the IVP (3) is defined on the whole of I.

Proof :

(Sketchy, by contradiction method). Assume the contradictory : $(\alpha, \beta) \underset{\neq}{C} I$, say, $\beta <$ right hand end point of I, so that $[t_0, \beta] \subseteq I$.

Now, being the solution of the IVP (3) the curve $X : (\alpha, \beta) \to \mathbb{R}^n$ satisfies the integral equation :

$$X(t) = x_0 + \int_{t_0}^{t} A(s)\,X(s)\,ds + \int_{t_0}^{t} u(s)\,ds$$

Using continuity of the maps $A : I \to M_n(\mathbb{R}), u : I \to \mathbb{R}^n$ we get a finite constant M such that $\|A(s)\| \leq M, \|u(s)\| \leq M$ for all $s \in [t_0, \beta]$ and therefore, we have :

$$\|X(t)\| \leq \|x_0\| + \int_{t_0}^{t} \|A(s)\,X(s)\|\,ds + \int_{t_0}^{t} \|u(s)\|\,ds$$

$$\leq \|x_0\| + \int_{t_0}^{t} M \|X(s)\| \, ds + M\beta.$$

By Gronwall's lemma, we get :

$\|x(t)\| \leq \left(\|x_0\| + M\beta\right) e^{M\beta}$ for all $t \in [t_0, \beta)$. Thus, the set $\{X(t) : t \in [t_0, \beta)\}$ is a bounded subset of $\mathbb{R}^n$ and therefore, the limit $\lim_{t \to \beta} X(t)$ exists. We call it $y \in \mathbb{R}^n$.

Having arrived at the point y in $\mathbb{R}^n$, we consider the initial value problem :

$$\frac{dx}{dt} = A(t)\, X + u(t) \qquad X(\beta) = y.$$

Let $X : (\beta - \eta, \beta + \eta) \to \mathbb{R}^n$ $(\eta > 0)$ be a solution of this IVP. Clearly the two solutions :

$X : (\alpha, \beta) \to \mathbb{R}^n, X : (\beta - \eta, \beta + \eta) \to \mathbb{R}^n$ agree on the overlap and therefore, they patch up to give a solution : $X : (\alpha, \beta + \eta) \to \mathbb{R}^n$ which contradicts the assumed maximality of the interval (α, β). Therefore, we must have : $\beta =$ right hand end point of I. Similar reasoning leades us to $\alpha =$ left hand end point of I and therefore $I = (\alpha, \beta)$.

Thus, every solution of (2) whatever be the initial condition is defined on the whole of $\mathbb{R}$.

5.3 THE SOLUTION OF THE HOMOGENEOUS EQUATION

We consider the set of all the solutions of the homogeneous equation (2). Let the set be denoted by V.

Proposition 2 : The set V has the structure of a n dimensional vector space.
Proof : Let a, b in $\mathbb{R}$, X, Y, in V be arbitrary. We prove that $aX + bY$ also is in V :

$$\frac{d}{dt}(aX + bY) = a\frac{dX}{dt} + b\frac{dY}{dt}$$
$$= a\,(A(t)\,X) + b\,A(t)\,Y$$
$$= A(t)\,(aX + bY) \qquad \text{because } A(t) \text{ is linear.}$$

Thus, $\frac{d}{dt}\left(aX+bY\right)=A(t)\left(aX+bY\right)$ i.e. $aX+bY\in V$.

This shows that V is a real vector space. Actually V is isomorphic with $\mathbb{R}^n$, the brief explanation of which is as follows.

Choose $t_0\in I$ arbitrarily and hold it fix. For each $x\in\mathbb{R}^n$ we consider the unique solution of the initial value problem :

$$\frac{dx}{dt}=A(t)\,X,\,X(t_0)=x.$$

We denote the unique solution of it by X_x where we have attached the suffix *x* to the solution X_x to indicate the dependence of the solution on the initial condition.

Now, we have an association rule $x\mapsto X_x$ associating the unique $X_x\in X$ with each $x\in\mathbb{R}^n$. In other words, we have the map :

$$\mathbb{R}^n\to V;\,x\mapsto X_x \qquad \ldots (4)$$

(Which associates each $x\in\mathbb{R}^n$, the element X_x of V). It is easy to show that this map is an isomorphism. First the linearity of the map : Let a, b in $\mathbb{R}$, *x, y* in $\mathbb{R}^n$ be arbitrary. We consider the two curves :

$$aX_x+bX_y:I\to\mathbb{R}^n \text{ and } X_{ax+by}:I\to\mathbb{R}^n.$$

It is clear that both are solutions of the IVP with the same initial condition *ax* + *by* and therefore by the uniqueness of the solution, we get the desired equality.

Clearly $X_x=0$ implies $x=0$. The implies that the linear map (4) is injective. Finally, let X be any element of V. Let $X\left(t_0\right)=x_0$. Then $X=X_{x_0}$ showing that the map (4) is a surjective map.

We have explained now that the map (3) is linear, it is injective and surjective as well. Therefore (4) is a linear isomorphism between V and $\mathbb{R}^n$, i.e. V is indeed n-dimensional real vector space.

We consider a vector basis $\left(X_1\,X_2\,.....\,X_n\right)$ of the solution space V. We call it a **fundamental system of solutions** of the homogeneous ODE (2).

Clearly for each $t \in I$, the vectors $X_1(t), X_2(t), ..., X_n(t)$, are linearly independent vectors of $\mathbb{R}^n$. Putting them along the columns of a $n \times n$ maxtrix, we denote the resulting $n \times n$ matrix by $W(t)$ thus :

$W(t) = \begin{bmatrix} \vdots & \vdots & & \vdots \\ X_1(t) & X_2(t) & & X_n(t) \\ \vdots & \vdots & & \vdots \end{bmatrix}$ or if the vector $X_j(t)$ has the coordinates : $X_j(t) = x_{1j}(t), x_{2j}(t), ..., x_{nj}(t)$ then $W(t) = \left[x_{ij}(t)\right]; 1 \leq i, j \leq n$.

We call the resulting map $W : I \to M_n \; \mathbb{R}$ a **fundamental matrix of solutions** of the homogeneous part (2). Note that for each $t \in I, W(t)$ is an invertible matrix. Here is a simple example :

We consider the 2 dimensional case in which the 2×2 matrix $A(t)$ is the constant matrix $A(t) = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$ for all $t \in \mathbb{R}$. It gives rise to the system of homogeneous ODE :

$$\left.\begin{aligned} \frac{dx}{dt} &= 2x - 3y \\ \frac{dx}{dt} &= 3x + 2y \end{aligned}\right\} \qquad \ldots (*)$$

Putting $X_1(t) = e^{2t}\begin{bmatrix} \cos 3t \\ \sin 3t \end{bmatrix}, X_2(t) = e^{2t}\begin{bmatrix} -\sin 3t \\ \cos 3t \end{bmatrix}, t \in \mathbb{R}$, we get the fundamental system X_1, X_2 of solution space of (*) and the resulting fundamental matrix $W : \mathbb{R} \to M_2(\mathbb{R})$ given by $W(t) = e^{2t}\begin{bmatrix} \cos 3t, -\sin 3t \\ \sin 3t, \; \cos 2t \end{bmatrix}$ for all $t \in \mathbb{R}$.

5.4 THE INHOMOGENEOUS EQUATION

We now consider the inhomogeneous ODE (1) and its solution space. To begin with, we have the following result relating the solutions of the two equations (1) and (2).

Proposition 3 : Let $Y : I \to \mathbb{R}^n$ be a solution of the inhomogeneous system (1).

a) If $X : I \to \mathbb{R}^n$ is a solution of the homogeneous system (2) then $X + Y$ is a solution of the inhomogeneous system (2).

b) Then the solution of the (inhomogeneous) initial value problem :

$$\frac{dx}{dt} = A(t)\, X + u(t),\, X(t_0) = x_0$$

is given by $X(t) = Y(t) + W(t) \int_{t_0}^{t} W(s)^{-1}\, u(s)\, ds$ for all $t \in I$.

The proof of the theorem is a straight forward application of the fundamental theorem of integral calculus (applied to integration of vector valued functions).

Proof :(a) We have $X\ t_0\ = Y\ t_0\ + W\ t_0\ \int_{t_0}^{t} W\ s\ ^{-1}\, u(s)\, ds$

$$= x_0 + W\ t_0\ \ 0 \qquad \text{since} \int_{t_0}^{t} *\ ds = 0$$

$$= x_0$$

(b) First note that

$$\frac{d}{dt} W\ t\ = \frac{d}{dt} \begin{bmatrix} \vdots & \vdots & & \vdots \\ X_1(t) & X_2(t) & & X_n(t) \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$= \begin{bmatrix} \vdots & \vdots & & \vdots \\ \frac{dX_1(t)}{dt} & \frac{dX_2(t)}{dt} & & \frac{dX_n(t)}{dt} \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$= \begin{bmatrix} \vdots & \vdots & & \vdots \\ A(t)X_1(t) & A(t)X_2(t) & & A(t)X_n(t) \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$= A(t) \begin{bmatrix} \vdots & \vdots & & \vdots \\ X_1(t) & X_2(t) & & X_n(t) \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$= A(t)\, W(t)$$

Now, we have :

$$\frac{d}{dt} X(t) = \frac{d}{dt}\left[Y(t) + W(t) \int_{t_0}^{t} W(s)^{-1} u(s) ds \right]$$

$$= \frac{d}{dt} Y(t) + \frac{d}{dt} W(t) \int_{t_0}^{t} W(s)^{-1} u(s)\, ds + W(t) \frac{d}{dt} \int_{t_0}^{t} W(s)^{-1} u(s)$$

$$= A(t)\, Y(t) + A(t)\, W(t). \int_{t}^{t} W^{-1}(s)\, u(s)\, ds + W(t)\, W^{-1}(t)\, u(t)$$

$$= A(t) \left[Y(t) + W(t) \int_{t_0}^{+} W(s)^{-1} u(s) ds \right] + u(t)$$

$$= A(t)\, x(t) + u(t).$$

Thus $\frac{d\,X(t)}{dt} = a(t)\, x(t) + u(t)$ all $t \in I$ and $X(t_0) = x_0$.

Illustrative Example :

(We do not solve the example completely we indicate only a few steps leaving further details for the reader to settle.)

Solve :

$$\frac{dx}{dt} = 2x - 3y + 2 \quad x(0) = 1$$

$$\frac{dy}{dt} = 3x + 2y + t \quad y(0) = 2$$

Solution (Incomplete) :

We have : $W(t) = e^{2t} \begin{bmatrix} \cos 3t & -\sin 3t \\ \sin 3t & \cos 3t \end{bmatrix}$

And $Y(t) = e^{2t} \begin{bmatrix} \cos 3t & -\sin 3t \\ \sin 3t & \cos 3t \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$$= e^{2t} \begin{bmatrix} \cos 3t - 2\sin 3t \\ \sin 3t + 2\cos 3t \end{bmatrix}.$$

There fore

$$X(t) = e^{2+} \begin{bmatrix} \cos 3t - \sin 3t \\ \sin 3t + \cos 3t \end{bmatrix} + e^{2t} \begin{bmatrix} \cos 3t & -\sin 3t \\ \sin 3t & \cos 3t \end{bmatrix} \int_0^t e^{-2s} \begin{bmatrix} \cos 3s & \sin 3s \\ -\sin 3s & \cos 3s \end{bmatrix} \begin{bmatrix} 2 \\ s \end{bmatrix} ds.$$

$$= e^{2t}\begin{bmatrix}\cos 3t - 2\sin 3t \\ \sin 3t + 2\cos 3t\end{bmatrix} + e^{2t}\begin{bmatrix}\cos 3t & -\sin 3t \\ \sin 3t & \cos 3t\end{bmatrix}\int_0^t e^{-2s}\begin{bmatrix}\cos 3s & \sin 3s \\ -\sin 3s & \cos 3s\end{bmatrix}\begin{bmatrix}2 \\ s\end{bmatrix} ds$$

$$= e^{2t}\begin{bmatrix}\cos 3t - 2\sin 3t \\ \sin 3t + 2\cos 3t\end{bmatrix} + e^{-2t}\begin{bmatrix}\cos 3t & -\sin 3t \\ \sin 3t & \cos 3t\end{bmatrix}\begin{bmatrix}* \\ **\end{bmatrix}$$

where in the last column, $* = \int_0^t e^{-2s}(2\cos 3s + s\sin 3s)\, ds$

and $** = \int_0^{t} e^{-2s}(s\cos 3s - 2\sin 3s) ds$

5.5 HIGHER ORDER ODE

As usual, I denotes an open interval, for a natural number n, we consider a single ODE.

$$\frac{d^n x}{dt^n} + a_1(t)\frac{d^{n-1}}{dt^{n-1}} + a_2(t)\frac{d^{n-2}}{dt^{n-2}} + \ldots + a_n(t)\, x = b(t) \qquad \ldots (5)$$

in an unknown function $x : I \to \mathbb{R}$ the coefficients $a_1, a_2, \ldots, a_n$, *b* being smooth functions on *I*. Equation (5) is **linear** because the left hand side of it is a linear combination of $x, \frac{dx}{dt}, \ldots \frac{d^n x}{dt^n}$. Again, if b $\equiv 0$ then we say that the equation (5) is homogeneous.

Recall the initial valve problem for (5) is the following. Given to $\in I, x_0, x_1 \ldots x_n$ all constant real numbers, find a *n* times continuously differentiable function $x : J \to \mathbb{R}$, *J* being an open interval with $t_0 \in J \subseteq I$ such that the following two requirements are satisfied :

(i) $\frac{d^n x(t)}{dt^n} + a_1(t)\frac{d^{n-1}x(t)}{dt^{n-1}} + \ldots. + a_n(t)\, x(t) = b(t)$ for all $t \in J$

and (ii) $x(t_0) = x_0 , \frac{dx}{dt}(t_0) = x_1 \ldots \frac{d^{n-1}x(t_0)}{dt^{n-1}} = x_{n-1.}$

We will reduce the ODE (5) to a linear system of first order ODE and get information of solutions of the former in terms of those of the reduced system. Towards this aim, we consider the following object.

$$Y = \begin{bmatrix} x \\ \frac{dx}{dt} \\ \vdots \\ \frac{d^{n-1}x}{dt^{n-1}} \end{bmatrix}, \text{ (ii) } A(t) = \begin{bmatrix} \ddots & +1 & & & \\ & \ddots & +1 & & \\ & & \ddots & & \\ & & & & +1 \\ -a_1(t) & \dots & \dots & & -a_n(t) \end{bmatrix}$$

$$\text{(iii) } u(t) = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ b(t) \end{bmatrix}$$

In the matrix A(t) in (ii) their being zeros at all the vacant places, including the main diagonal, and the + 1 entries being just above the main diagonal and parallel to it. Now, we consider the system

$$\frac{dx}{dt} = A(t)\,x + u(t) \qquad \dots (5')$$

Along with its homogeneous part : $\frac{dx}{dt} = A(t)\,x$ $\qquad \dots (6)$

Note that the given (order n) ODE (5) is equivalent to the first order system (5') while the homogeneous part of (5) is equivalent to (6). We recall the results of the preceding sections obtained for the linear systems, now applicable to (5') which we transcribe them so as to become applicable to the equation (5).

Thus we consider a fundamental system $Y_1, \dots, Y_n$ of the solution space of (6). This system yields functions $x_1,\ x_2, \dots . x_n \ : I \to \mathbb{R}$ such that

$$Y_1 = \begin{bmatrix} x_1 \\ \frac{dx_1}{dt} \\ \vdots \\ \vdots \\ \frac{d^{n-1}x_1}{dt^{n-1}} \end{bmatrix}, Y_2 = \begin{bmatrix} x_2 \\ \frac{dx_2}{dt} \\ \vdots \\ \vdots \\ \frac{d^{n-1}x_2}{dt^{n-1}} \end{bmatrix} \dots\dots\dots Y_n = \begin{bmatrix} x_n \\ \frac{dx_n}{dt} \\ \vdots \\ \vdots \\ \frac{d^{n-1}x_n}{dt^{n-1}} \end{bmatrix}$$

Now, we have the following important facts :

(1) $Y_1, \dots, Y_n$ are linearly independent solutions of (6) implies $x_1 \dots x_n$ are solutions of the homogeneous part of (5). Moreover any solution x of the

homogeneous part of (5) is expressible as a linear combination of the functions $x_1\, x_2 \dots x_n$.

$$x = c_1 x_1 + c_2 x_2 + \quad + c_n\, x_n.$$

(2) The solutions $x_1 \dots x_n$ (of the homogeneous part of (5) are linearly independent over I.

$$d_1 x_1 + d_2 x_2 + \dots + d_n\, x_n \equiv 0$$

implies $d_1 = d_2 = \dots = d_n = 0$.

This proves the following important :

Proposition 4 : The solution space of the homogeneous part of (5) is a n-dimensional real vector space.

Now, given any set $f_1 \,\dots f_n : I \to \mathbb{R}$ of *n* times continuously differentiable functions, we associate with it the function :

$$W = W\,(f_1 \dots f_n)\; : I \to \mathbb{R}$$

given by : $W(t) = \det\left[\dfrac{d^j f_i}{dt^j}(t)\right]$ for all $t \in I$

The function *W* is called the Wronskian of the family $\{f_1 \dots\; f_n\}$.

Note that when the functions $x_1 \dots. x_n \;\; I \to \mathbb{R}$ form a vector berries of the solution space the matrix :

$$W = W(x_1 \dots x_n) : I \to \mathbb{R}$$

given by W(t) = det $\left[\dfrac{d^j}{dt^j} x_i\,(t)\right]$ for all t $\in I$ is the fundamental matrix of the homogenous part (6) of $(5')$

5.6 A SOLUTION OF THE NON-HOMOGENEOUS EQUATION

Suppose, a fundamental system $\{x_1 \dots\dots x_n\}$ of solutions of the homogeneous equation (6) is found. We discuss a method – attributed to Lagrange-which yields a solution of the non-homogeneous ODE (5).

Recall, W(t) stands for the fundamental matrix with its (ij)th entry $\dfrac{d^j}{dt^j} x_i(t)$.

For each i, $1 \le i \le n$, we consider the $n \times n$ matrix denoted by $W_j(t)$ obtained from W(t) by replacing its *j*th column by the column

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ b_{(t)} \end{bmatrix}$$

We adopt the notations : $D(t)$ for $det(W(\text{t}))$ and $D_i(\text{t})$ for $det(W_i(\text{t}))$.

Now, to obtain the desired solution we consider a function $x : I \rightarrow \mathbb{R}$ which is in the form

$$x(t) = v_1(t)x_1(t) + v_2(t)x_2(t)\, t\, . + v_n(t)\, x_n(t) = \sum_{j=1}^{n} v_j(t)x_j(t),$$

where $v_j : I \rightarrow \mathbb{R}\ (1 \leq j \leq n)$ are unknown functions which are required to satisfy a number of identities.

These identities will determine the functions v_j which in turn specify the x which will be the desired solution. Now differentiating $x(t)$. We get

$$\frac{dx}{dt} = \sum_{j=1}^{n} v_j \frac{dx_j}{dt} + \sum_{j=1}^{n} \frac{dv_j}{dt} x_j$$

The first requirement on the v_j is :

$$\sum_{j=1}^{n} \frac{dv_j}{dt} x_j = 0 \qquad \dots \text{(i)}$$

So that we are left with

$$\frac{dx}{dt} = \sum_{i=1}^{n} v_j \frac{dx_j}{dt} \qquad \dots (*)$$

Now differentiating (*) above, we get

$$\frac{d^2x}{dt^2} = \sum_{j=1}^{n} v_j \frac{d^2x_j}{dt^2} + \sum_{j=1}^{n} \frac{dv_j}{dt} \frac{dx_j}{dt}$$

The second requirement on the v_j is

$$\sum_{j=1}^{n} \frac{dv_j}{dt} \frac{dx_j}{dt} = 0 \qquad \dots \text{(ii)}$$

Leaving us with $\dfrac{d^2x}{dt^2} = \sum_{j=1}^{n} v_j \dfrac{d^2x_j}{dt^2}$ … (**)

Continuing this procedure we get analogous identities, the requirement on v_j at the last stage being :

$$\sum_{j=1}^{n} \frac{dv_j}{dt} \frac{d^{n-1}x_j}{dt^{n-1}} \equiv b.$$

Thus, we have the following two strings of identities :

$$\left.\begin{aligned} x &= \sum_{j=1}^{n} v_j\, x_j \\ \frac{dx}{dt} &= \sum_{j=1}^{n} v_j \frac{dx_j}{dt} \\ &\vdots \\ \frac{d^n x}{dt^n} &= \sum_{j=1}^{n} v_j \frac{d^n x_j}{dt^n} \\ &\quad + \sum_{j=1}^{n} \frac{dv_j}{dt} \frac{d^n x_j}{dt^n} \end{aligned}\right\} \ldots (I) \qquad \left.\begin{aligned} \sum_{j=1}^{n} \frac{dv_j}{dt} x_j &= 0 \\ \sum_{j=1}^{n} \frac{dv_j}{dt} \frac{dx_j}{dt} &= 0 \\ &\vdots \\ \sum_{j=1}^{n} \frac{dv_j}{dt} \frac{d^n x_j}{dt^n} &= b(t) \end{aligned}\right\} \ldots (II)$$

Multiplying the equations in (I) by a_n, a_{n-1},…., a_1, 1 and adding we get

$$\frac{d^n x}{dt^n} + a_1(t)\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_{n-1}(t)x + a_n(t) = b(t)$$

thus showing that the function $x(t)$ is a solution of the inhomogeneous equation.
On the other hand, using the simultaneous equations in (ii), we get :

$$\frac{dv_j}{dt} = \frac{D_j(t)}{D(t)}$$

and therefore $v_j(t) = \int_{t_0}^{t} \frac{D_j(s)}{D(s)} ds \quad , 1 \le j \le n$

This leads us to the desired solution

$$x(t) = \sum_{j=1}^{n} v_j(t)\, x_j(t), \quad t \in I.$$

EXERCISES :

(1) Prove that each solution of the inhomogeneous equation (5) is defined on I.
(2) Prove : If a solution $x : I \to \mathbb{R}$ of the homogeneous equation (6) vanishes at same $t_0 \in I$, then $x \equiv 0$

(3) Solve : $\frac{dx}{dt} = 5x + 3 \quad x(0) = 1$

$\frac{dy}{dt} = 3x + 2y + t \quad y(0) = 2$

(4) Solve : $\frac{dx}{dt} = -x + 3y + 4t \quad x(0) = 1$

$$\frac{dy}{dt} = 3x + y + 4 \quad y(0) = 1$$

(5) Same D. E as above but $x(1) = 1$, $y(1) = 1$.

6

METHOD OF POWER SERIES

Unit Structure

6.1 INTRODUCTION

In this chapter, we study a type of second order ODE (scalar case) which gives solutions in the form of absolutely convergent power series. These ODE contain in their form, functions (e.g. the coefficient functions) which are *analytic* in sense that they admit absolutely convergent power series expansions. Naturally the method of solving such DE makes use of techniques and properties of absolutely convergent power series. Therefore we call this method the **method of power series**.

The reader will realize that this method applies not only to second order linear ODE, it actually is applicable to a wider class of ODE of any order.

Recall, at the elementary level we could solve simple DE in terms of elementary function such as the polynomials, the logarithm function, the exponential functions, the trigonometric functions and so on. But soon we find that things start going the opposite way : Differential equations generate new functions as their solutions. Such functions are called **special functions**. Most of these functions are in the form of power series and as such are obtained by the methods of power series. There is a more powerful method which is called **Frobaneous method**. We discuss briefly this method also. Using this method we introduce two special functions : (i) the Legendre polynomials and (ii) the Bessel functions. We derive some of their properties.

We begin our treatment of special functions by recalling basic facts of power series.

6.2 POWER SERIES (A QUICK REVIEW)

A **power series** is an infinite sum of the type:

$$a_0 + a_1(t - t_0) + a_2(t - t_0)^2 + \ldots\ldots \ldots\ldots + a_k(t - t_0)^k + \ldots\ldots\ldots$$

$$= \sum_{k \in \mathbb{Z}^+} a_k (t - t_0)^k \qquad \ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

where $a_0, a_1, a_2, \ldots\ldots\ldots, a_k, \ldots\ldots\ldots$ are real constants. It is absolutely convergent if there exists a $r > 0$ such that

$$\sum_{k \in \mathbb{Z}^+} |a_k| r^k < \infty$$

The lub of all $r > 0$ satisfying the above inequality is the **radius of convergence** of the power series (1), we denote it by R. If (1) is absolutely convergent with R as its radius of convergence, it follows that for each

$t \in (t_0 - R, t_0 + R)$ the infinite sum

$$\sum_{k \in \mathbb{Z}^+} a_k (t - t_0)^k$$

converges, giving rise the function :

$$f : (t_0 - R, t_0 + R) \longrightarrow \mathbb{R} \ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

where

$$f(t) = \sum_{k \in \mathbb{Z}^+} a_k (t - t_0)^k$$

for each $t \in (t_0 - R, t_0 + R)$.

The function (2) is the **sum function** of the (absolutely convergent) power series (1). It is a basic result that the sum function (2) is infinitely differentiable on its domain and the k^{th} derivative ($k \in \mathbb{Z}^+$) is obtained by differentiating the infinite series *termwise*. In particular we have :

$$\left(\frac{d}{dt}\right)^k f(t_0) = k!\; a_k$$

Consequently the power series (1) becomes

$$f(t) = \sum \frac{(t - t_0)^k}{k!} \left(\frac{d}{dt}\right)^k f(t_0) \ldots\ldots\ldots \text{(3)}$$

An important implication of (3) is the following result :

If the functions

$$f : (t_0 - R, t_0 + R) \longrightarrow \mathbb{R}$$

$$g : (t_0 - R, t_0 + R) \longrightarrow \mathbb{R}$$

admit the power series expansions :

$$f(t) = \sum_{k \in \mathbb{Z}^+} a_k (t - t_0)^k$$

$$g(t) = \sum_{k \in \mathbb{Z}^+} b_k (t - t_0)^k s$$

then we have the following basic fact :

$$f(t) \equiv g(t)$$

if and only if $a_k = b_k$ for all $k \in \mathbb{Z}^+$ that is, if and only if

$$\left(\frac{d}{dt}\right)^k f(t_0) = \left(\frac{d}{dt}\right)^k g(t_0)$$

holds for all $k \in \mathbb{Z}^+$

We make use of this basic result in what is to follow in this chapter.

6.3 METHOD OF POWER SERIES.

We consider a second order ODE of the type :

$$\frac{d^2x}{dt^2} + P(t)\frac{dx}{dt} + Q(t)x = 0 \text{ ……… (4)}$$

where the coefficient functions P(t) and Q(t) admit absolutely convergent power series expansions on an interval (-R, R) :

$$\left.\begin{aligned} P(t) &= a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \cdots + a_k t^k + \cdots \\ Q(t) &= b_0 + b_1 t + b_2 t^2 + b_3 t^3 + \cdots + b_k t^k + \cdots \end{aligned}\right\} \text{……….. (5)}$$

Recall, according to the theory of linear second order homogeneous ODE discussed in Unit4, the solution space of the ODE (4) is a two dimensional vector space.

Now, because the functions P(t), Q(t) admit absolutely convergent power series expansions - we call such functions **analytic** – we expect a solution of (4) also to be analytic :

$$x(t) = \sum c_n t^n \text{ ……… (6)}$$

We prove below the following two results :

(a) Indeed, a solution $t \longmapsto x(t)$ of (4) has a power series expansion (6) and obtain the constants c_k, $k \geq 2$ in terms of the constants a_k, b_k, k, l in $\mathbb{Z}^+$ (The constants c_0, c_1, will play the role of the arbitrary constants in the solution of the second order ODE (4).)

(b) The infinite series (6) is absolutely convergent in the interval (-R, R).

We proceed to prove these two claims.

Differentiating the power series (6) for the solution x(t), we get

$$\frac{dx}{dt}(t) = \sum_{k\geq 1} k c_k t^{k-1}$$

$$\frac{d^2x}{dt^2} = \sum_{k\geq 2} k(k-1) c_k t^{k-2}$$

for all t in (-R, R). Substituting these power series expansions along with those for P(t), Q(t) in the given DE we get :

$$\left\{\sum_{k\geq 2} k(k-1)c_k t^{k-1} + \left(\sum_{l\geq 0} a_l t^l\right)\cdot\left(\sum_{m\geq 1} m c_m t^{m-1}\right) + \left(\sum_{l\geq 0} b_l t^l\right)\cdot\left(\sum_{m\geq 1} c_m t^{m-1}\right)\right.$$

In above the coefficient of the power t^{n-2} for $n \geq 2$ is :

$n(n-1)c_n + a_0(n-1)c_{n-1} + a_1(n-2)c_{n-2} + \ldots\ldots\ldots + a_{n-2}c_1 + b_0c_{n-2} + b_1c_{n-3} + \ldots\ldots\ldots + b_{n-2}c_0$

Equating it with zero we get the following succession of equations :

$$n(n-1)c_n = -[a_0(n-1)c_{n-1} + a_1(n-2)c_{n-2} + \ldots.. + a_{n-2}c_1]$$
$$-[b_0c_{n-2} + b_1c_{n-3} + \ldots\ldots\ldots + b_{n-2}c_0] \ldots\ldots\ldots (*)$$

for $n \geq 2$. These equations show that the constants $(c_n : n \geq 2)$ can be obtained recursively in terms of *arbitrary constants* c_0, c_1 and the *given* constants $a_0, a_1, a_2, \ldots\ldots\ldots, b_0, b_1, b_2, \ldots\ldots\ldots$ Thus the solution (6) is *formally* obtained. It remains only to prove that the formal series (6) is absolutely convergent for $|t| < R$ and hence determines a function $x: (-R, R) \longrightarrow \mathbb{R}$ which then becomes the solution of (2). Towards the justification of this claim, we have :

$$n(n-1)|c_n| \leq |a_0|(n-1)|c_{n-1}| + |a_1|(n-2)|c_{n-2}| + \ldots\ldots\ldots + |a_{n-2}||c_1|$$
$$+ |b_0||c_{n-2}| + |b_1||c_{n-3}| + \ldots\ldots\ldots + |b_{n-2}||c_1|$$
$$\leq (n-1)\,[|a_0||c_{n-1}| + |a_1||c_{n-2}| + \ldots\ldots\ldots + |a_{n-2}||c_1| + \ldots]$$
$$+ [|b_0||c_{n-2}| + |b_1||c_{n-3}| + \ldots\ldots\ldots + |b_{n-2}||c_0|] \ldots\ldots\ldots (6)$$

Now, let a number r satisfying 0 < r < R be arbitrary chosen. Also, choose one more constant say S with 0 < r < S < R. By the absolute convergence of the two series in (5) in (-R, R) and by the choice S < r, we have

$$\sum_{n\in\mathbb{Z}^+} |a_n| S^n < \infty \text{ and } \sum_{n\in\mathbb{Z}^+} |b_n| S^n < \infty$$

We choose a D > 0 such that

$$\sum_{n\in\mathbb{Z}^+} |a_n| S^n \leq D \text{ and also } \sum_{n\in\mathbb{Z}^+} |b_n| S^n \leq D \text{ .... (7)}$$

Consequently, we have $|a_n| \leq \frac{D}{S^n}$, $|b_n| \leq \frac{D}{S^n}$ for all $n \in \mathbb{Z}^+$. Next, we consider an arbitrary $m \in \mathbb{N}$ (to be fixed later) and for this m, another constant M, (again, larger enough but finite) so that the following inequalities hold for $0 \leq k \leq m-1$:

$$|c_k| \leq \frac{M}{r^k} \text{ .... (8)}$$

Substituting the estimates (7), (8) in the inequality (6) we get :

$$m(m-1)|c_m| \leq (m-1)\left[\frac{DM}{r^{m-1}} + \frac{D}{S}\frac{DM}{r^{m-2}} + \cdots + \frac{D}{S^{m-1}} M\right]$$

$$+ \left[\frac{DM}{r^{m-1}} + \frac{D}{S}\frac{DM}{r^{m-2}} + \cdots + \frac{D}{S^{m-1}} M\right]$$

$$= (m-1)\frac{MrD}{r^m}\left[1 + \frac{r}{s} + \cdots + \left(\frac{r}{m}\right)^{m-1}\right]$$

$$+ \frac{M^2 rD}{r^m}\left[1 + \frac{r}{s} + \left(\frac{r}{s}\right)^2 + \cdots + \left(\frac{r}{m}\right)^{m-2}\right]$$

$$\leq (m-1)\frac{MrD}{r^m}\left[1 + \frac{r}{s} + \left(\frac{r}{s}\right)^2 + \cdots\right]$$

$$+ \frac{M}{r^m} r^2 D\left[1 + \frac{r}{s} + \left(\frac{r}{s}\right)^2 + \cdots\right]$$

$$= \frac{M}{r^m}\left[\frac{rD(m-1) + r^2 D}{m(m-1)\left(1 - \frac{r}{s}\right)}\right]$$

Therefore, we have

$$|c_m| \leq \frac{M}{r^m}\left[\frac{rD(m-1) + r^2 D}{m(m-1)\left(1 - \frac{r}{s}\right)}\right] \text{ .... (**)}$$

At this stage we fix m. It should be so large that the expansion in the above last inequality (**) is ≤ 1. With this choice of m, we have $|c_m| \leq \frac{M}{r^m}$. This inequality together with the inequalities (*) imply :

$$|c_k| \leq \frac{M}{r^k}$$

for $0 \leq k \leq m$. Now, application of principle of mathematical induction and the inequality (6) together imply that, the inequalities (8) are true for all $k \in \mathbb{Z}^+$. This ensures that the series (6) defining the function $t \longmapsto x(t)$ is absolutely convergent for all t with $|t| < r$. Again, this is true for all r with $0 < r < R$ and therefore the series in (6) is absolutely convergent for all $t \in (-R, R)$. This leads us to the following :

Theorem 1 :The function

$$x : (-R, R) \longrightarrow \mathbb{R}$$

given by

$$x(t) = \sum_{k \in \mathbb{Z}^+} c_k t^k$$

where c_0, c_1 are arbitrary constants and the c_k, $k > 2$ satisfying (*).

Remark :Of all the constants $(c_k : k \in \mathbb{Z}^+)$ in (6) the constants c_2, c_3, ………., c_k, ………. are expressed in terms of the constants a_k, b_k, $k \in \mathbb{Z}^+$,the last constants namely c_0, c_1 remaining unspecified. They are the two *arbitrary constants* of the second order ODE (4).

6.4 ILUSTRATIVE EXAMPLES

(I) The DE :$\frac{d^2x}{dt^2} + 10x = 0$. Here, $P(t) \equiv 0$ and $Q(t) \equiv 10$.

Let $x(t) = c_0 + c_1t + c_2t^2 + ……….$ be a solution of the equation. Then we get $\frac{d^2x}{dt^2} = 2.1c_2 + 3.2c_3t + 4.3c_4t^2 + ……….$ and therefore $2.1c_2 + 3.2c_3t + 4.3c_4t^2 + ………. + 10(c_0 + c_1t + c_2t^2 + ……) = 0$, that is :
$(10c_0 + 2.1c_2) + (10c_1 + 3.2c_3)t + (10c_2 + 4.3c_4)t^2 + …………… + (10c_k + (k+2)(k+1)c_{k+2})t^k + ………. = 0$

Therefore we get :

$10c_k + (k+2)(k+1)c_{k+2} = 0$ for $k = 2, 3, ……….$ This gives :

$c_{k+2} = \frac{-10c_k}{(k+2)(k+1)}$ for all $k \in \mathbb{Z}^+$.

This recurrence relation gives the following succession :

$c_2 = \frac{-10c_0}{2}$ $c_3 = \frac{-10c_1}{3.2}$ $c_4 = \frac{-10c_2}{4.3} = \frac{10^2c_0}{4!}$

$c_5 = \frac{-10c_3}{5.4} = \frac{10^2c_1}{5!}$, $c_6 = \frac{-10c_4}{6.5} = \frac{10^2c_0}{6!}$, $c_7 = \frac{-10c_5}{7.6} = \frac{-10^2c_1}{7!}$

…………

$c_{2k} = \frac{-10^k c_0}{2k!}$ ……… $c_{2k+1} = \frac{(-10)^k c_1}{(2k+1)!}$………

Therefore,

$x(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 + \ldots\ldots\ldots$

$= (c_0 + c_2 t^2 + c_4 t^4 + \ldots\ldots\ldots) + (c_1 t + c_3 t^3 + c_5 t^5 \ldots\ldots\ldots)$

$$= c_0\left[1 - \frac{10t^2}{2!} + \frac{10^2 t^4}{4!} - \frac{10^3 t^6}{6!} + \cdots\right]$$

$$+ c_1\left[t - \frac{10t^3}{3!} + \frac{10^2 t^5}{5!} - \ldots\right]$$

$$= c_0\left[1 - \frac{10t^2}{2!} + \frac{10^2 t^4}{4!} - \frac{10^3 t^6}{6!} + \cdots\right]$$

$$+ \frac{c_1}{\sqrt{10}}\left[\sqrt{10}t - \frac{(\sqrt{10})^3 t^3}{3!} + \frac{(\sqrt{10})^5 t^5}{5!} - \frac{(\sqrt{10})^7 t^7}{7!} + \cdots\right]$$

$$= c_0 \cos\sqrt{10}t + \frac{c_1}{\sqrt{10}} \sin\sqrt{10}t$$

$$= A \cos\sqrt{10}t + B \sin\sqrt{10}t$$

Where $A = c_0$, $B = \frac{c_1}{\sqrt{10}}$ are arbitrary constants.

(II) Solve $\frac{d^2x}{dt^2} + 2t\frac{dx}{dt} + 4x = 0$

Solution : Let $x(t) = c_0 + c_1 t + c_2 t^2 + \ldots\ldots\ldots$

Now, we have :

$\frac{d^2x}{dt^2} = 2.1c_2 + 3.2c_3 t + 4.3c_4 t^2 + \ldots\ldots + (k + 2)(k + 1)c_{k+2} t^k + \ldots\ldots$

and therefore,

$2t\frac{d^2x}{dt^2} = 2c_1 t + 2.2c_2 t^2 + \ldots\ldots\ldots + 2kc_k t^k + \ldots\ldots\ldots$

$4x(t) = 4c_0 + 4c_1 t + 4c_2 t^2 + \ldots\ldots\ldots + 4c_k t^k + \ldots\ldots\ldots$

Therefore,

$\frac{d^2x}{dt^2} + 2t\frac{d^2x}{dt^2} + 4x = (12c_2 + 4c_0) + (6c_3 + 6c_1)t + (4.3c_4 + 8c_2)t^2$

$+ \ldots\ldots\ldots + [(k + 2)(k + 1)c_{k+2} + 2(k + 2)c_k]t^4$

$+ \ldots\ldots\ldots = 0$

Equating the coefficients of the powers of t with zero, we get

$c_2 = \frac{-c_0}{2}$, $c_3 = -c_1$, $c_4 = \frac{-2c_2}{3} = \frac{c_0}{3}$, ………, $c_{k+2} = \frac{-2c_k}{(k+1)}$, ………

This gives the solution :

$x(t) = c_0 + c_1 t - \frac{c_0}{2} t^2 - c_1 t^3$ ……….

(III) We consider here a first order ODE, the solution of which is to be obtained following a similar procedure :

$$\frac{dX}{dt} = 5x.$$

Assuming the solution to be the power series :
$x(t) = c_0 + c_1 t + c_2 t^2 + \ldots\ldots\ldots + c_k t^k \ldots\ldots\ldots$
We get :
$\frac{dx}{dt} = c_1 + 2c_2 t + 3c_3 t^2 + \ldots\ldots\ldots + (k + 1)c_{k+1} t^k + \ldots\ldots\ldots$
Substituting these power series in the given differential equations, we get :
$c_1 + 2c_2 t + 3c_3 t^2 + \ldots\ldots\ldots + (k + 1)c_{k+1} t^k + \ldots\ldots\ldots$
$= 5c_0 + 5c_1 t + 5c_2 t^2 + \ldots\ldots\ldots + 5c_k t^k + \ldots\ldots\ldots$
Equating the coefficients, we get :
$c_1 = 5c_0$, $2c_2 = 5c_1$, $3c_3 = 5c_2$, ……….,$kc_k = 5c_{k-1}$, for all $k \geq 1$ and therefore $c_k = \frac{5^k c_0}{k!}$ for all $k \geq 1$.

This gives :
$x(t) = c_0 + \frac{c_0 5t}{1!} + \frac{c_0 (5t)^2}{2!} + \frac{c_0 (5t)^3}{3!} + \ldots\ldots\ldots$
$= c_0 e^{5t}$.

6.5 LEGENDRE EQUATIONS, LEGENDRE POLYNOMIALS

For an arbitrary real number α, we consider the differential equation

$$(1 - t^2)\frac{d^2x}{dt^2} + 2t\frac{dx}{dt} + \alpha(\alpha + 1)\, x = 0 \;\ldots\ldots\ldots\; (9)$$

for |t| < 1. We rewrite it in the form

$$\frac{d^2x}{dt^2} - \frac{2t}{(1 - t^2)}\frac{dx}{dt} + \frac{\alpha(\alpha + 1)}{(1 - t^2)} x = 0 \;\ldots\ldots\ldots\; (9')$$

The two equivalent forms (9) and (9') of the differential equation are called the **Legendre equation** involvingthe parameter α. It is a particular case of the ODE (4) in which $P(t) = \frac{-2t}{1 - t^2}$, $Q(t) = \frac{\alpha(\alpha+1)}{(1 - t^2)}$both admitting power series solution in (-1, 1) i.e. R = 1. According to Theorem 1 the Legendre equation has a solution given by an infinite power series converging absolutely in the interval (-1, 1). The resulting function (which depends on the parameter α) is called the **Legendre function**.

It can be proved that the coefficients c_n, $n \in \mathbb{Z}^+$ in the expansion

$$x(t) = \sum_{n \in \mathbb{Z}^+} c_n t^n$$

of the solution of (9) satisfy the recurrence relations :

$$c_n = \frac{[(n-1)(n-2)-\alpha(\alpha+1)]\, c_{n-2}}{n(n-1)}$$

for $n \geq 2$. In particular, if the parameter α takes an integral value say $\alpha = m$, then $(n - 1)(n - 2) - \alpha(\alpha + 1) = 0$ for $n = m + 1$ and consequently, $c_{m+1} = 0$. This further implies that one solution of the Legendre equation is a polynomial. Because the Legendre equation is a homogeneous linear differential equation, the polynomial solution of it is determined to within a multiplicative constant. A particular polynomial solution of it denoted by $P_m(t)$ is the polynomial :

$$P_m(t) = \frac{-1}{2^m.m!}\frac{d^m}{dt^m}(t^2-1)^m$$

$P_m(t)$ is the **Legendre polynomial of degree m**.

6.6 THE FROBANEOUS METHOD

We consider a homogeneous linear second order ODE of the type :

$$t^2\frac{d^2x}{dt^2} + t\,P(t)\frac{dx}{dt} + Q(t)x = 0 \quad \ldots\ldots\ldots (10)$$

Like the ODE (9) it is more general than (4) because of the coefficients t^2 of $\frac{d^2x}{dt^2}$ and t of $\frac{dx}{dt}$. Again, the functions P(t) and Q(t) have the power series expansions :

$$P(t) = a_0 + a_1t + a_2t^2 + \ldots\ldots\ldots$$

$$Q(t) = b_0 + b_1t + b_2t^2 + \ldots\ldots\ldots$$

both the power series being absolutely convergent in an interval (-R, R).

It turns out that the solution is in the form of power series with $t = 0$ as a singular point of the solution. The method of getting a solution of (10) is called the **Frobaneous method**. It is explained below.

We expect the solution of (10) to be a function of type :

$$x(t) = t^s(c_0 + c_1t + c_2t^2 + \ldots\ldots\ldots) \quad \ldots\ldots\ldots (*)$$

wheres is a real number and c_0 is non-zero. We have therefore to find s, c_0, c_1, ………

Assuming the series (*) to be absolutely convergent in (-R, R) we consider its derivatives :

$$\left.\begin{aligned} \frac{dX}{dt} &= st^{s-1}(c_0 + c_1t + c_2t^2 + \cdots) \\ &+ t^s\,(c_1 + 2c_2 + 3c_3t^2 + \cdots) \end{aligned}\right\} \ldots\ldots\ldots (**)$$

$$\left.\begin{aligned} \frac{d^2X}{dt^2} &= s(s-1)t^{s-2}(c_0 + c_1t + c_2t^2 + \cdots) \\ &+ 2st^{s-1}(c_1 + 2c_2 + 3c_3t^2 + \cdots) \\ &+ t^s\,(2.1c_2 + 3.2c_3t + 4.3c_4t^2 + \cdots) \end{aligned}\right\} \ldots\ldots\ldots\ (***)$$

Substituting these power series expansions for x(t), $\frac{dX}{dt}$, $\frac{d^2X}{dt^2}$ in equation (10) we get :

$$t^s \cdot \begin{pmatrix} c_0s(s-1) + c_1s(s+1)t + c_2(s+1)(s-1)t^2 + \cdots \\ + [a_0 + a_1t + a_2t^2 + \cdots]\cdot[sc_0 + (s+1)c_1t + (s+2)c_2t^2 + \cdots] \\ + [b_0 + b_1t + b_2t^2 + \cdots]\cdot[c_0 + c_1t + c_2t^2 + \cdots] \end{pmatrix}$$

$\equiv 0$

Because the factor $t^s \neq 0$, we get that the expansion within the bracket must be identically zero. Therefore the coefficient of each power of t in above must be zero. This gives the following succession of equations :

$$\left.\begin{aligned} c_0(s-1) + c_0a_0s + c_0b_0 = 0 \\ c_1s(s+1) + c_1a_0(s+1) + c_0a_1s + c_1b_0 + c_0b_0 = 0 \\ \vdots \\ c_n(s+n-1)(s+n) + c_n(s+1)a_0 + \cdots + c_nb_0 + \cdots \\ = 0 \end{aligned}\right\} \ldots\ldots\ldots\ (****)$$

We solve these equations to get the values of c_0, c_1, c_2, ………

To begin with, we consider the *first* equation in the set (****). Since $c_0 \neq 0$, we get

$$s(s-1) + a_0s + b_0 = 0$$

This equation is called the **indicial equation**. This equation, which is a quadratic equation in s, when solved it gives two values for s to be substituted in the solution (*).

We then consider arbitrary $c_0 \neq 0$ and using the succession of equations in (****) we obtain c_0, c_1, c_2, ………..

The procedure described in above is applied in the next section where we obtain a family of special functions called **Bessel functions**.

6.7 BESSEL FUNCTIONS

For any $p \in \mathbb{Z}^+$ we consider the ODE :

$$t^2\frac{d^2x}{dt^2} + t\frac{dx}{dt} + (t^2 - p^2)x = 0$$

theODE being called the **Bessel's equation**.

Clearly, the indicial equation of this D.E. is :

$$s(s-1) + s - p^2 = 0$$

It gives s ± p. Therefore, one solution of the Bessel's equation, denoted by $J_p(t)$ is of the form :

$$J_p(t) = t_p(c_0 + c_1 t + c_2 t^2 + \ldots\ldots\ldots)$$

where c_0 is an arbitrary constant. Taking $c_0 = \frac{1}{2^p \cdot p!}$ we get

$$J_p(t) = t^p\left(\frac{1}{2^p \cdot p!} + c_1 t + c_2 t^2 + c_3 t^3 + \cdots\right)$$

Substituting the power series expansion of $J_p(t)$, $\frac{d}{dt}\big(J_p(t)\big)$, $\frac{d^2}{dt^2}\big(J_p(t)\big)$, in the differential equation (11) we get

$s(s-1)c_0 + (s+1)sc_1 t + (s+2)(s+1)c_2 t^2 + \ldots\ldots\ldots$

$+ sc_0 + (s+1)c_1 t + (s+2)c_2 t^2 + \ldots\ldots\ldots$

$- p^2 c_0 - p^2 c_1 t - p^2 c_2 t^2 - p^2 c_3 t^3 - \ldots\ldots\ldots$

$= 0$

Equating coefficients of powers of t gives :

$[s(s-1) + s - p^2] \cdot c_0 = 0$

$[(s+1)s + s + 1 - p^2]c_1 = 0$

.

.

.

$[(s+n)(s+n-1) + (s+n) - p^2]c_n + c_{n+2} = 0 \quad n \geq 2$

The first of these equations in the indicial equation giving s = ±p, the second gives $c_1 = 0$ and the last equation gives $c_n = \frac{-c_{(n-2)}}{(p+n)^2 - p^2}$. Therefore the Bessel functions are given by

$$J_p(t) = \frac{t^p}{2^p p!}\left[1 - \frac{t^2}{2(2p+2)} + \frac{t^4}{2.4(2p+2)(2p+4)} \ldots\ldots\ldots\right]$$

When s = -p, we get the relation :

$c_n = \frac{-c_{n-2}}{(-p+n) - p^2}$ for $n \geq 2$.

EXERCISES :

Obtain solutions in the form of power series of the following D.E.

(i) $\frac{dX}{dt} = -2tx$

(ii) $t\frac{dX}{dt} - 3x = k$ (= constant)

(iii) $\frac{d^2X}{dt^2} + 6x = 0$

(iv) $\frac{d^2X}{dt^2} - 3\frac{dX}{dt} + 2x = 0$

(v) $(t-2)\frac{dX}{dt} = tx$

(vi) $\frac{d^2X}{dt^2} - 4\frac{dX}{dt} + (4t^2 - 2)x = 0$

*****

7

STURM – LIOUVILLE THEORY

Unit Structure :

7.1 Introduction :

Recall, the initial value problem for an ordinary differential equation:

$$\frac{d^k X}{dt^k} = f(t, X, \frac{dX}{dt}, \ldots \frac{d^{k-1}X}{dt^{k-1}})$$

is to obtain a solution $X : I \to \mathbb{R}$ of it when the values :

$X(t_o) = X_0, \frac{dX}{dt}(t_o) = v_1, \ldots, \frac{d^{k-1}}{dt^{k-1}} t_0 = v_{k-1}$ of the solution and its derivatives are prescribed at a single point to of its domain internal. A **Boundary Value Problem** (BVP), is another fundamental problem in the theory of ODE in which the solution of an ODE is required to satisfy a number of conditions at two points of its domain (the two points actually being the boundary points of the domain interval.)

In this chapter, we will study an important type of boundary value problems associated with a certain type of linear second order ODE; we call the BVP the "Sturm Liouille eigenvalue problem." The resulting theory is very vast and makes use of results from functional analysis. Therefore, we only outline the theory introducing the concepts and stating the results without proof. We will illustrate the scope of the theory by using it to solve the vibrating string problem.

2. The sturmain Boundary Value Problem :

In the following, I stands for the interval $[a,b]$. All the functions $t \mapsto X(t)$, $t \mapsto Y(t)$, $t \mapsto Z(t)$ etc appearing in the discussion are assumed to be defined on internals containing I.

We use the following notations :

C(*I*) is the vector space of all continuous functions $X: I \to R, C^1(I)$ is the subspace of *C*(*I*) which are continuously differentiable on *I* while $C^2(I)$ consist of those $X: I \to \mathbb{R}$ in *C*(*I*) which are twice continuously differentiable on *I*.

To introduce the type of boundary value problems we want to discuss, we consider the following data:

(I) The functions *p, q, r :* $I \to \mathbb{R}$ with the following properties :

(i) *p* is continuously differentiable on *I* i.e. $p \in C^1(I)$

(ii) *q, r* are continuous on I.

(iii) *p*(*t*) > 0 for all $t \in I$.

(II) Constants $\alpha_1, \alpha_2, \beta_1, \beta_2$ with ${\alpha_1}^2 + {\alpha_2}^2 > O$ and $\beta_1^2 + \beta_2^2 > O$ (Note that $\alpha_1^2 + \alpha_2^2 > O$ is equivalent to the property that of the pair $\alpha_1\ \alpha_2$, at least one is non-zero. The other inequality also has similar interpretation).

(III) Arbitrary constants η_1, η_2.

Using the functions in (I) we construct the linear, second order ordinary differential equation :

$$(pX')' + qX = r \qquad \dots (1)$$

i.e. $\frac{d}{dt}\left(p(t)\frac{dX}{dt}(t)\right) + q(t)X(t) = r(t), t \in I.$

in the unknown function $t \mapsto X(t)$

The equation : $(PX')' + qX = 0 \qquad \dots (2)$

is the homogeneous part of it.

We require the solution function $t \mapsto X(t)$ of (1) to satisfy the boundary conditions :

$$\left.\begin{aligned} \alpha_1 X(a) + \alpha_2 X'(a) &= \eta_1 \\ \beta_1 X(b) + \beta_2 X'(b) &= \eta_2 \end{aligned}\right\} \qquad \dots (3)$$

Taking together the DE(1) and the boundary conditions (3) we get the pair

$$(pX')' + qX = r$$

$$\alpha_1 X(a)+\alpha_2 X'(a)=\eta_1,\ \beta_1 X(b)+\beta_2 X'(b)=\eta_2 \qquad \ldots (4)$$

The pair (4) is said to constitute the **Sturmian boundary value problem**.

In above, taking $r\equiv 0, \eta_1=0=\eta_2$ we get the **homogeneous Sturmian boundary value problem:**

$$(P X')'+q X=0$$

$$\alpha_1 X(a)+\alpha_2 X'(a)=0=\beta_1 X(b)+\beta_2 X'(b) \qquad \ldots (5)$$

In above for every $X\in C^2(I)$, we put

$$L(X)=(pX')+q X$$

$$=pX''+p'X'+qX.$$

Note that $X\in C^2(I)$ implies that $L(X)$ is continuous. Thus we get the map

$$L: C^2(I)\to C(I)$$

of the indicated vector spaces. Clearly L is linear.

Now for any X, Y in $C^2(I)$ we have

$$L(X)(t).Y(t)-X(t).L(Y)(t)$$

$$=\frac{d}{dt}\left[p(t)\left(\frac{d X(t)}{dt}.Y(t)-X(t).\frac{dY}{dt}(t)\right)\right]$$

for all $t\in I$. We refer to this equality as the *Lagrange Identity*. Integrating this identity over the interval $[a,b]$, we get

$$\int_a^b\left[L(X)(t).Y(t)-X(t).L(Y)(t)\right]dt$$

$$=p(b)\left[X'(b).Y(b)-X(b).Y'(b)\right]$$

$$-p(a)\left[X'(a)Y(a)-X(a)Y'(a)\right] \qquad \ldots (*)$$

Moreover, if both the functions X, Y satisfy the boundary conditions in (5) then it follows that the R.H. S. of (*) is zero and consequently we have :

$$\int_a^b L(x)(y).Y(t)\,dt=\int_a^b X(t)\,L(x)(t)\,dt \qquad \ldots (**)$$

We will explain more about this equality (**) at a later stage.

Here is a short list of properties of the spaces of solutions of the boundary value problem (4) and its homogeneous part (5).

(1) A finite linear combination :

$$C_1 X_1+C_2 X_2+\ldots+C_n X_n$$

of solutions $X_1, X_2, \ldots, X_n$ of the B.V.P. (5) is also a solution of the BVP (5).

(2) If X and Y are solutions of the in homogeneous BVP (5) then the difference $X - Y$ is a solution of the homogeneous boundary value problem (5).

(3) If $X : I \to R$ is a solution solution of the BVP (5) and $Y : I \to \mathbb{R}$ a solution of the non-homogeneous B.V.P. (4) then X + Y is a solution of the BVP (4).

(4) Finally, let $Y : I \to \mathbb{R}$ be a fixed solution of the inhomogeneous BVP (4). Then every other solution $\tilde{Y} : I \to \mathbb{R}$ of (4) can be expressed in the form $Y = X + Y$ for a unique solution X of the homogeneous BVP (5).

At this stage we describe a condition which ensures a unique solution of the boundary value problem (4). Towards this aim, recall that the solution space of the second order, linear homogeneous ODE (2) is a 2 dimensional vector space and we call a basis of this vector space a **fundamental system** of the ODE.

We choose a fundamental system X_1, X_2 of the ODE (2). Next, using it and the constants $\alpha_1, \alpha_2, \beta_1, \beta_2$, of the boundary conditions of (5) we form the quantity.

$$R_{11} = \alpha_1 X_1 \ a \ + \alpha_2 X_1' \ a$$

$$R_{12} = \beta_1 X_1 \ b \ + \beta_2 X_1' \ b$$

$$R_{21} = \alpha_1 X_2 \ a \ - \alpha_2 X_2' \ b$$

$$R_{22} = \beta_1 X_2 \ b \ + \beta_2 X_2' \ b$$

and we consider the determinant $\det \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ which we denote by W. Now we have the following result.

Theorem 1 : The boundary value problem (4) has a unique solution if and only if $W \neq 0$.

Proof :

Let X_1, X_2 be a fundamental system of solutions of the DE(2). Using X_1, X_2 and the variation of constants formula we choose a particular solution $Y^* : I \to \mathbb{R}$ of the differential equation (1). Now a general solution of (1) has the form:

$$Y \ t \ = Y^* \ t \ + C_1 Y_1 \ t \ + C_2 Y_2 \ t \ , \ t \in I \qquad \ldots (*)$$

C_1 , C_2 being some constants.

Now we consider (*) to be a solution of the BVP (6). Clearly, (*) is a solution of BVP (6) if and only if the constants C_1, C_2 satisfy the following simultaneous equations:

$$\left.\begin{aligned} \eta_1 &= R_1\, Y^* + C_1 R_{11} + c_2 R_{12} \\ \eta_2 &= R_2^* + C_1\, R_{21} + C_2\, R_{22} \end{aligned}\right\} \qquad \ldots(**)$$

where $R_1 Y^* = \alpha_1 Y^*(a) + \alpha_2 Y^{*\prime}(a)$, $R_2 Y^* = \beta_1 Y^*(b) + \beta_2 Y^{*\prime}(b)$

(here, of course $Y^{*\prime}$ being he derivative $\dfrac{dY^*}{dt}$).

Clearly the equations (**) are satisfied if and only if the matrix of the coefficients of C_1, C_2 in (**) is non-singular, that is, if and only if:

$$W = \det\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \neq O$$

Here is a simple illustrative case.
The boundary value problem :

$$\frac{d^2X}{dt^2} + X = 1, \quad 0 \le t \le \pi$$

$$X(0) + X'(0) = \eta \quad X'(\pi) = \nu$$

η, ν being some constants.
We claim that the BVP has a unique solution.
In fact, here, we have $\alpha_1 = \alpha_2 = 1$, $\beta_1 = 0$, $\beta_2 = 1$,

$X_1(t) \equiv \sin t, X_2(t) = \cos t$ is a fundamental system of solutions of the homogeneous part. $\dfrac{d^2X}{dt^2} + X = 0$. Therefore, we get :

$R_{11} = 1$, $R_{12} = 0 = R_{21}$ while $R_{22} = 1$. This gives $W = 1 \neq 0$. giving existence and uniqueness of the solution. In fact $X(t) = 1 + C_1 \cos t + C_2 \sin t$ is a general solution of the ODE in the BVP above Now
$\eta = X(0) + X'(0) = C_1 + C_2$. and $\nu = X'(\pi) = C_2$ give
$\nu = C_2$ and $C_1 = \eta - \nu$ Therefore the unique solution of the above B.V.P. is
$X(t) = 1 + (\eta - \nu)\, cost + \nu \sin t$.

Next, let the functions $p, q : I \to \mathbb{R}$, the constants $\alpha_1 \alpha_2, \beta_1, \beta_2$ and the differential operator $L : C^2(I) \to C(I)$, all be as in the preceding section. In addition let $r : I \to \mathbb{R}$ be a continuous function with $r(t) > 0$ holding for all $t \in I$.

For a real number λ we consider the linear homogeneous ODE.

$$L(X)+\lambda r\,X=0 \qquad \ldots (6)$$

along with the boundary conditions :

$$\alpha_1 X(a)+\alpha_2 X'(a)=0$$
$$\beta_1 X(b)+\beta_2 X'(b)=0$$

Thus we have the linear homogeneous boundary value problem :

$$\left.\begin{array}{l} p\,X''+p'X'+(q+\lambda r)X=0 \\ \alpha_1 X(a)+\alpha_2 X'(a)=0=\beta_1 X(b)+\beta_2 X'(b) \end{array}\right\} \ldots (7)$$

which involves the real parameter λ.

A Value λ of the parameter for which a non-zero solution $X=X_\lambda$ of (7) exists is called an $\boxed{eigen-value}$ of the boundary value problem (7).

Sturm – Liouville eigen-value problem consists of getting the get E consisting of all the eigen values λ, the corresponding eigen-functions X_λ and studying the function space $C(I)$ in terms of the eigen functions. $\{X_\lambda : \lambda \in E\}$. We will state the main theorem, proving only a part of it, and use it to solve the vibrating string problem in the next section.

Before going further, we consider the following three concepts:

(1) The inner product $<-,->: C(I)\times C(I) \rightarrow [0,\infty)$ given by $<\phi,\Psi> = \int_a^b \phi(t)\,\Psi(t)\,\sigma(t)\,dt$, ϕ,Ψ in $C(I)$

(2) The norm $\|\ \|: C(I) \rightarrow [o,\infty)$ given by $\|\phi\| = +\sqrt{<\phi,\phi>}$ for all $\phi \in C(I)$

(3) The uniform norm $\|\ \|_\infty : C(I) \rightarrow [o,\infty)$ given by $\|\phi\|_\infty = \text{lub}\{|\phi(t)| : t \in I\}$.

Note that if $C=\sqrt{\int_a^b r(t)\,dt.}$, then we have $\|\phi\| \le C\,\|\phi\|_\infty$ ……. (8)

holds for all $\phi \in C(I)$.

Now, for $\lambda \in \mathbb{R}$, we consider the following two subspaces of C(I) :

(1) $V(\lambda)$ is the solution space of the (second order, linear, homogeneous) ODE (6).

(2) $W(\lambda)$ is the subspace of $V(\lambda)$ consisting of all the solutions of the boundary value problem (7). If $W(\lambda) \neq (0)$ then we call $W(\lambda)$ the **eigen-space** of the B.V.P (7) with λ as its **eigen-value.**.

We prove the following three properties of the spaces $W(\lambda), \lambda \in \mathbb{R}$.

(I) Each $W(\lambda)$ is a **proper** subspace of $V(\lambda)$ (Thus if $W(\lambda) \neq (0)$ then it is a 1-dimensional subspace of $V(\lambda)$)

(II) If λ_1, λ_2 are in $\mathbb{R}$, with $\lambda_1 \neq \lambda_2$, and if $\phi \in W(\lambda_1), \Psi \in W(\lambda_2)$ then $< \phi, \Psi > = 0$ $(\phi \perp \Psi$ with respect to the inner product $< , >$.

(III) There is a countable subset $\{\lambda_k : k \in \mathbb{Z}^+\}$ of real numbers such that $W(\lambda_k) \neq (0)$ for each $k \in \mathbb{Z}$ and $W(\lambda) = 0$ if $\lambda \neq \lambda_k$ for any $\lambda \in \mathbb{R}$.

We prove property (I) : choose a fundamental system $\{X_1, X_2\}$ of solutions of the equation (6). Their linear dependence implies

$$\det \begin{bmatrix} X_1(t) & X_2(t) \\ X_1'(t) & X_2'(t) \end{bmatrix} \neq 0 \text{ for all } t \in I.$$

In particular we have

$$\det \begin{bmatrix} X_1(a) & X_2(a) \\ X_1'(a) & X_2'(a) \end{bmatrix} \neq 0$$

Now, if both X_1, X_2 were in W(λ), then they would satisfy the boundary conditions. In particular, $\alpha_1 X_1(a) + \alpha_2 X_2'(a) = 0$ and $\alpha_1 X_2(a) + \alpha_2 X_2'(a) = 0$

But $\alpha_1^2 + \alpha_2^2 > o$ implies $\det \begin{bmatrix} X_1(a) & X_2(a) \\ X_1'(a) & X_2'(a) \end{bmatrix} = 0$ which contradicts the above stipulation of linear independence. This proves that W(λ) is a proper subspace of $V(\lambda)$ (i.e. it is either the {0} subspace or it is one dimensional subspace of $V(\lambda)$.)

Property (II) follows by the property $\langle L(X), Y \rangle = \langle X, L(Y) \rangle$ for all *X*, *Y* in *C* (*I*).

Property (III) is a consequence of the fact that (*C* (*I*), $\| \; \|$) is a separable metric space.

Now we state without proof the main theorem.

Theorem 2 (Sturm – Lionville) :

(1) The boundary value problem (7) has a non-zero solution only for a countable (finite or denumerable) set of $\lambda \in \mathbb{R}$ i.e. the solution spaces $W(\lambda)$ are non-trivial only for a countable collection of real numbers λ.

Let $E = \{\lambda_1, \lambda_2, ..., \lambda_k, ...\}$ be the subset of $\mathbb{R}$ consisting of those λ such that $W(\lambda) \neq (0)$ if and only if $\lambda = \lambda_k$ for some $\lambda_k \in E$. (λ_k are the eigen values of the Sturm-Liouville problem and $W(\lambda_k)$ are the eigen spaces).

(2) The set E has no limit point.

(3) Each $W(\lambda_k)$ is a 1 dimensional subspace of $C^2(I)$. For each $\lambda_k \in E$ we choose $X_k \in W(\lambda_k)$ with $\|X_k\| = 1$. Now we have $W(X_k) = \mathbb{R}.X_k$.

(4) If $k \neq e$, then $X_k \perp X_e$ ie $< X_k, X_e > = 0$.

(5) If $X \in C^2(I)$ then $X(t) = \sum_k < X(t), X_k(t) > X_k(t)$

for all $t \in I$ where the convergence of to infinite series to $X(t)$ is uniform in $t \in I$.

(6) If $X \in C(I)$ is such that $< X_k, X > = 0$ holds for all X_k then $X \equiv 0$.

The proof of this theorem makes use of the properties of a compact operator on a separable Hilbert Space and as such it is to be studied from a suitable advanced text-book on ODE (which usually refer to text-books of functional analysis).

We state here two more results without proof :

Theorem 3 (Sturm-Liouville Separation Theorem.)

The zeros of two linearly independent solutions of $L(X) = 0$ separate each other.

Thus if X_1 *and* X_2 are two independent solutions of $L(X) = 0$, their between two consecutive zeros of X_1 is a zero of X_2 and between two consecutive zeros of X_2 is a zero of X_1.

Theorem 4 (The Comparison Theorem)

Consider two eigen-value problems of Sturm Liouville with the respective date p, q, r, α, β and $p^*, q^*, r^*, \alpha^*, \beta^*$ over the same interval $I = [a, b]$.

If $p \geq p^*$, $q^* \leq q$, $r^* \leq r$ in I and $0 \leq \alpha \leq \alpha^* \leq \pi$, $0 < \beta^* \leq \beta \leq \pi$ hold with strict inequality in at least one place, then the corresponding eigenvalues satisfy $\lambda_n > \lambda_n^*$ for all n.

In the next section, we study the dynamics of an elastic vibrating string in which we will use the results of the Sturm-Liouville Theorem.

7.3 VIBRATIONS OF AN ELASTIC STRING

The vibrating motion of a stretched elastic string is governed by a partial differential equation called the (one-dimensional) $\boxed{\textit{wave equation}}$. Wave equation is solved by a method called method of separation of variables. The resulting analysis makes use of the Sturm Liouville theory. We will therefore study the problem of the vibrating elastic string as an application of the Sturm-Liouville Theory.

We first explain the PDE, the wave equation of the vibrating string.

A string of natural length L is held horizontally along the X-axis of a vertical XOY-plane. Its ends, A, B remain tied to the points (0, 0) and (L, 0) respectively. The string is plucked slightly and then is set in motion in such a way that each point C of the string vibrates vertically. We study the vibrating motion of the string as the collective vertically oscillating motion of each point C of the string.

Therefore we consider an arbitrary point C of the string Let $l\ A,C\ = x$ (The real number x and the point C determine each other and therefore we may take of "the point x" instead of "the point C".)

Let at an instant $t \geq 0$, $Y(t, x)$ be the instantaneous y-coordinate of the point C (Since the oscillatory motion of C is only in the vertical direction, x-coordinate of C remains constant.) Therefore the (oscillatory) motion of the point C is described by the function $t \mapsto Y(t,x)$ and the motion of the whole string is given by the function $t, x \mapsto Y\ t, x \quad t \geq 0\ \ 0 \leq x \leq 1$

… (9)

Now, the basic equations of motion enable us to derive the equation

$$\frac{\partial^2 Y}{\partial t^2} = c^2 \frac{\partial^2 Y}{\partial x^2} \quad t \geq 0,\ 0 \leq x \leq L \qquad \ldots(10)$$

satisfied by the function $Y(t, x)$. Equation (10) is the wave equation which is satisfied by the vibrating string. In (10) c is a constant determined by the mass of the string its elastic properties and the gravitational constant.)

Suppose the string was plucked slighlty and released with initial velocity (= initial velocity of each point C). so that the string executes the vibration motion as described above. We consider two continuous functions :

$f:[0,L]\to\mathbb{R},\ \ g:[0,L]\to\mathbb{R}$ describing initial position and initial velocity of the string, that is :

$$Y(0,x)=f(x),\frac{\partial Y}{\partial t}(0,x)=g(x),0\le x\le L.$$

Now, we have the initial / boundary value problem for the function $(t,x)\mapsto Y(t,x)$;

$$\frac{\partial^2 Y}{\partial t^2}=c^2\frac{\partial^2 Y}{\partial x^2}\quad t\ge 0,\ 0\le x\le L \qquad \ldots(10)$$

$$Y(t,0)=0=Y(t,L)\qquad t\ge 0,$$

$$Y(0,x)=f(x),\frac{\partial Y}{\partial t}(0,x)=g(x),0\le x\le L.$$

We want to find out the function $(t,x)\mapsto Y(t)$

To begin with, we consider all the solutions of the wave equation which are of the type *X*(*x*). *T*(*t*):

$$Y(t,x)=X(x)T(t).$$

Now, $\frac{\partial^2 Y}{\partial t^2}(t,x)=X(x)\ddot{T}(t),\frac{\partial^2 Y}{\partial x^2}=\ddot{X}(x),T(t)$ (the dots indicating differentiation (twice) with respect to the appropriate variable)

Now the equation takes the form

$$X(x)\ddot{T}(t)=c^2\ddot{X}(x)T(t)$$

Assuming $X(x)\ne 0, T(t)\ne 0 \ for\, t>0, 0<x<L,$ we get

$$\frac{1}{c^2}\frac{\ddot{T}(t)}{T(t)}=\frac{\ddot{X}(x)}{X(x)} \qquad \ldots(11)$$

This shows that the common value in (11) is independent of *t*, *x* i.e. it must be a constant say *d*.

$$\frac{1}{c^2}\frac{\ddot{T}(t)}{T(t)}=\frac{\ddot{X}(x)}{X(x)}=d \qquad \ldots(12)$$

Now if $d\ge 0$ we would get $\ddot{T}(t)=c^2 d\,T(t)$ with $c^2 d\ge 0$. If $d=0$, we would get *T*(*t*) = *A*t + *B* for some constants $a\ne 0$, *B* (*A* is not zero because, otherwise $T(t)\equiv B$ which gives *Y*(*t*, *x*) = *B*. X(x) implying that the motion of the string is independent of t ie the string is station any!). On the other hand, if d > o, we would get

$$T\ t = Ae^{\sqrt{d}.ct} + Be^{-\sqrt{d}.ct} \quad t \geq 0$$

In either case (i.e d > 0, or d = 0) the factor T(t) becomes unbounded as *t* ranges in $0,\infty$. This renders *Y*(*t*, *x*) also unbounded (ie the string stretching limitless, another physical impossibility!). Therefore we are left with the possibility d < 0. We put $d = -\lambda^2$ for $\lambda \in \mathbb{R}$. Then the above ODE (12) take the form :

$$\ddot{T}\ t = -\lambda^2 c^2 T\ t\ ,\ \ddot{X}\ x = -\lambda^2 X\ x \qquad ..\ (13)$$

We consider the second equation :

$$\ddot{X}\ x = -\lambda^2 X\ x\ , \quad 0 \leq x \leq L.$$

Its general solution is :

$$X\ x = x\cos\lambda x + \beta\sin\lambda x \quad \alpha,\beta \text{ being constants. This gives:}$$

$Y\ t,x = T\ t\ \ \alpha\cos\lambda x + \beta\sin\lambda x$ with the condition $T\ t \neq 0$. Now $0 \equiv Y\ t,0 = x.T\ t$ implies $\alpha = 0$ and therefore $Y\ t,x = \beta.T\ t\ \sin\lambda x$ with $\beta \neq 0, T\ t \not\equiv 0.$

But we have *Y(t, L)* = 0 and therefore $\sin\lambda L = 0$ which implies that $\lambda L = k\pi$ *for* $k \in \mathbb{Z}$. Therefore the parameter λ can take the values $\lambda k = \dfrac{k\pi}{L}, k \in \mathbb{Z},$.

This shows that :

$$T\ t = T_k\ t = \alpha_k.\cos\left(\frac{\pi cL}{k}\right)t + \beta_k \sin\left(\frac{\pi cL}{k}\right)t$$

and $X\ x = X_k\ x = Sin\left(\dfrac{\pi L}{k}\right)x, k \in \mathbb{Z}$.

Thus, we get a sequence of solutions :

$$Y_k\ t\,x = \left[\alpha_k.\cos\left(\frac{\pi cL}{k}\right)t + \beta_k.\sin\left(\frac{\pi cL}{k}\right)t\right]\sin\left(\frac{\pi L}{K}\right)x$$

for $k \in \mathbb{Z}$, $X\alpha_{k,}\ \beta_k$ being arbitrang constants. Now the general solution *Y(t, x)* will be a linear combination of all of them.

$$Y\ t,x = \sum_{k\in\mathbb{Z}}\left[\alpha_k.\cos\left(\frac{\pi cL}{k}\right)t + \beta_k.\sin\left(\frac{\pi cL}{k}\right)t\right]\sin\left(\frac{\pi L}{k}\right)x$$

Differentiating the infinite series partially with respect to t gives :

$$\frac{\partial Y}{\partial t}\ t,x = \sum_{k\in\mathbb{Z}}\left(\frac{\pi CL}{k}\right)\left[-\alpha_k.\sin\left(\frac{\pi CL}{k}\right)t + \beta_k\cos\left(\frac{\pi CL}{k}\right)t\right]\sin\left(\frac{\pi L}{k}\right)x$$

In particular, we have

$$f(x) = Y(0,x) = \sum_{k\in\mathbb{Z}} \alpha_k \sin\left(\frac{\pi L}{k}\right)x$$

$$g(x) = \frac{\partial Y}{\partial t}(0,x) = \sum_{k\in\mathbb{Z}} \beta_k \sin\left(\frac{\pi L}{k}\right)x$$

These are nothing but the Fourier expansions of the given functions *f(x), g(x),* α_k, β_k being their Fourier coefficients which are calculated using the standard trigonometric identities. Now we have the solution of the vibrating string problem:

$$Y(t,x) = \sum_{k\in\mathbb{Z}} \left[\alpha_k . \cos\left(\frac{\pi CL}{k}\right)t + \beta_k . \sin\left(\frac{\pi CL}{k}\right)t\right] \sin\left(\frac{\pi L}{k}\right)x$$

7.4 Unit End Exercises :

Find the eigen values and the eigen functions of the following boundary value problems :

(1) $\ddot{X} + \lambda X = 0 \qquad X(0) = 0,\ \dot{X}(1) = 0$

(2) $\ddot{X} + \lambda X = 0 \qquad X(0) = 0, X(L) = 0 \text{ for } L > 0.$

(3) $\ddot{X} + \lambda X = 0, \qquad X(0) = 0 \quad \dot{X}(L) = 0 \text{ for } L > 0.$

(4) $(t\dot{X})^{\cdot} + \lambda t^{-1} = 0 \qquad X(1) = 0 \ \dot{X}(e) = 0$

(Hint : Try $t = e^r,\ r \in \mathbb{R}$)

(5) $\left(e^{2t}\dot{X}\right)^{\cdot} e^{2t}(\lambda + 1)X = 0, \quad X(0) = 0 = X(\pi)$

(Hint: Take $X = \overset{-t}{e}u$)

(6) $\left(t^{-1}\dot{X}\right)^{\cdot} + (\lambda + 1)t^{-3}X = 0,\ X(1) = 0, X(e) = 0.$