Ordinary Differential Equations
Modern Perspective

Ordinary Differential Equations
Modern Perspective

Mohan C Joshi
IIT Bombay

Preface

This book has culminated into its title as an outgrowth of series of undergraduate and graduate level courses on ordinary differential equations (ODE), taught by the author for the past several years. It is our aim to present a unified and comprehensive treatment to ODE based on widely used abstract formulations in the underlying spaces. The emphasis is more on qualitative analysis of solvability of ODE than the actual methodology of computing the closed form solution. In our presentation we presume only the basic knowledge of linear algebra and advanced calculus of the undergraduate level. Hence the text is suitable for graduate students of mathematics and other branches of science and engineering requiring a broad perspective of ODE. The contents have been class tested for several years in a regular course on ODE in the Department of Mathematics at IIT Bombay.

With this approach in mind, we have divided the text into eight chapters. Chapter 1 introduces ODE with several examples drawn from applied areas. These examples are continuously tracked at appropriate places in later chapters for illustrating the theory.

Chapter 2 is devoted to the development of mathematical structure required to deal with linear and nonlinear operators in function spaces. These operators will arise while discussing various facets of ODE.

Chapter 3 gives basic existence and uniqueness theory needed for initial value problems, the starting point for any discussion of ODE. The subsequent analysis rests heavily on the concept of transition matrix. It is, therefore, dealt with in detail in Chapter 4. While in Chapter 5, we use the properties of transition matrix in the study of stability theory of both linear and nonlinear systems.

Chapter 6 lays emphasis on the clasical theory of series solution for ODE. We focus on Legendre, Hermite and Bessel equations. Boundary value problems are investigated in Chapter 7 through the methods of Green's function and eigenfunction expansion. We give rather an elongated description on Green's function in view of its inherent importance.

The closing Chapter 8 is a fitting finale, giving interesting glimpse of vast applicability potential of ODE to control theory - a fertile inter-disciplinary area involving of science and engineering.

Acknowledgements

This book could not have been written without direct and indirect assistance from various institutional bodies within IIT Bombay. Firstly, we would like to acknowledge the financial support received from the Curriculum Development Program of the Institute. Secondly, we thank the Department of Mathematics and Industrial Mathematics Group for providing the necessary secretarial and computational help throughout the preparation of the manuscript while it went through the number of iterations. Several colleagues and students helped us from time to time in many ways giving correct references, proof-reading the manuscript etc. In this connection, we would like to acknowlege the help rendred by Amiya K Pani, M K Ramaprasad and Soumya Mohanty. Our special thanks to Asmit Pany for type setting the manuscript in LATEX.
Finally, this author would like to thank his wife for patiently supporting him in implementing the project.

Mohan C. Joshi
Powai, Mumbai

Contents

Chapter 1

Introduction To Ordinary Differential Equations

In this chapter we begin by introducing the concept of solvability of initial value problems (IVP) corresponding to first order differential equations (ODE) in $\Re^n$. We also give, along with appropriate examples of practical importance, all related notions like orbit or trajectory, direction field, isocline and solution curve of IVP.

The reader is also familiarized with difference equations and numerical solutions of ODE through the process of discretization. An interesting concept of neural solution of ODE is also touched upon. Towards the end, we recall standard methods of solving some classes of first order ODE.

1.1 First Order Ordinary Differential Equations

Let $\Omega \subset \Re^n$ be open and let $f : \Omega \to \Re^n$ be a mapping (not necessarily linear) given by

$$f(x_1, x_2, x_n) = (f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots x_n), \ldots, f_n(x_1, x_2, \ldots, x_n))$$

We shall refer to f as vector field on Ω. A vector field may also depend on an additional parameter t, that is, $f = f(t, \overline{x})$.

The main thrust of our study is the following differential equation

$$\frac{d\bar{x}}{dt} = f(t, \overline{x}) \tag{1.1.1}$$

where $f : I \times \Omega \to \Re^n, I$ a subinterval of $\Re$.

The value $\overline{x}(t) \in \Omega \subset \Re^n$ is sometimes referred as the state of the system described by the ordinary differential equation (ODE) - Eq.(1.1.1). The set Ω is then called the state space or phase space.

Definition 1.1.1 *The Initial Value Problem (IVP) corresponding to Eq. (1.1.1) is given by*

$$\frac{d\overline{x}}{dt} = f(t, \overline{x}(t)) \tag{1.1.1(a)}$$

$$\overline{x}(t_0) = \overline{x}_0 \tag{1.1.1(b)}$$

Definition 1.1.2 *A function $\overline{x}(t)$ is said to be a solution of the IVP - Eq. (1.1.1), if there exists an interval $J \subset I$, containing t_0 such that $\overline{x}(t)$ is differentiable on J with $\overline{x}(t) \in \Omega$ for all $t \in J$ and $\bar{x}(t)$ satisfies Eq. (1.1.1).*

As $\overline{x}(t)$ needs to satisfy Eq. (1.1.1) only on J, this solution is sometimes referred as a local solution. If $t \to f(t, .)$ is defined on the entire line $\Re$ and $\bar{x}(t)$ satisfies Eq. (1.1.1) on $\Re$, then $\bar{x}(t)$ is said to be a global solution of the IVP. In Chapter 3, we shall discuss the local and global solvability of IVP in detail.

At times, it is desirable to indicate the dependence of the solution $\overline{x}(t)$ on its initial value $\overline{x}_0$. Then we use the notation $\overline{x}(t, t_0, \overline{x}_0)$ for $\overline{x}(t)$.

Definition 1.1.3 *The orbit or the trajectory of the ODE- Eq.(1.1.1), is the set $\{\overline{x}(t, t_0, \bar{x}_0) : t \in J\}$ in the state space Ω. Whereas the solution curve is the set $\{(t, \overline{x}(t, t_0, \bar{x}_0)) : t \in J\} \subset I \times \Omega$.*

Definition 1.1.4 *The direction field of ODE - Eq. (1.1.1), is the vector field $(1, f(t, \overline{x}))$.*

It is clear from the above definitions that the orbit of the ODE - Eq. (1.1.1) is tangent to the vector field whereas the solution curve is tangent to the direction field at any point.

The following proposition gives the equivalence of solvability of IVP - Eq. (1.1.1) with the solvability of the corresponding integral equation.

Proposition 1.1.1 *Assume that $f : I \times \Omega \to \Re^n$ is continuous. $\bar{x}(t)$ is a solution the IVP - Eq. (1.1.1) iff $\bar{x}(t)$ is the solution of the integral equation*

$$\bar{x}(t) = \bar{x}_0 + \int_{t_0}^{t} f(s, \bar{x}(s))ds, \quad t \in J \tag{1.1.2}$$

Proof : If $\bar{x}(t)$ is a solution of Eq.(1.1.1), by direct integration on the interval $[t_0, t]$ for a fixed t, we get

$$\int_{t_0}^{t} \left(\frac{d\bar{x}}{ds}\right) ds = \int_{t_0}^{t} f(s, \bar{x}(s))ds$$

This gives

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

and hence we have Eq. (1.1.2).

Conversely, if Eq.(1.1.2) holds, $\bar{x}(t)$ given by Eq. (1.1.2) is differentiable and differentiating this equation we get

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

Also we have $\bar{x}(t_0) = \bar{x}_0$. ∎

Definition 1.1.5 *The set of points in Ω where $f(t, \bar{x}) = 0$, is called the set of equilibrium points of ODE - Eq. (1.1.1). It is clear that at these points the orbit or trajectory is a singleton.*

Example 1.1.1 *(Population Dynamics Model)*
We investigate the variation of the population $N(t)$ of a species in a fixed time span. An important index of such investigation is growth rate per unit time, denoted by

$$R(t) = \left[\frac{1}{N(t)}\right]\frac{dN}{dt} \tag{1.1.3}$$

If we only assume that the population of species changes due to birth and death, then growth rate is constant and is given by,

$$R(t) = R(t_0) = b - d$$

where b and d are birth and death rates, respectively.

If the initial population $N(t_0) = N_0$, we get the following initial value problem

$$\frac{dN}{dt} = R(t_0)N \tag{1.1.3(a)}$$

$$N(t_0) = N_0 \tag{1.1.3(b)}$$

The solution of the above IVP is given by

$$N(t) = N_0 \exp(R(t - t_0)), \quad t \in \Re$$

Although the above model may predict the population $N(t)$ in the initial stages, we realise that no population can grow exponentially for ever. In fact, as population grows sufficiently large, it begins to interact with its environment and also other species and consequently growth rate $R(t)$ diminishes. If we assume the form of $R(N)$ as equal to $a - cN(a, c$ positive constant$)$, we get the following IVP

$$\frac{dN}{dt} = N(a - cN) \tag{1.1.4a}$$

$$N(t_0) = N_0 \tag{1.1.4b}$$

This is known as logistic equation with a as the growth rate without environmental influences and c representing the effect of increase population density.

The equilibrium points are 0 *and* $a/c, C = a/c$ *is called the carrying capacity of the environment.*

One can easily solve the above ODE by the method of separation of variables and get

$$N(t) = \frac{\frac{a}{c}}{\left[1 + \frac{\left(\frac{a}{c} - N_0\right)}{N_0} \exp(-a(t - t_0))\right]}$$

It follows from the above representation that $\lim_{t \to \infty} N(t) = a/c = C$ *(carrying capacity). So the above solution has the representation*

$$N(t) = \frac{C}{\left[1 + \frac{(C - N_0)}{N_0} \exp(-a(t - t_0))\right]} \tag{1.1.5}$$

The solution curve of the above population model is as under.

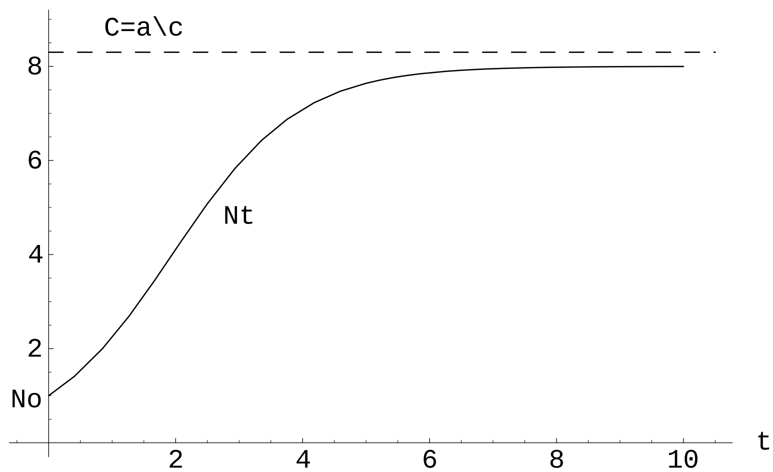

Figure 1.1.1: Growth of population $N(t)$

The parameters a *and* C *are usually not known but can be estimated by minimising the mean square deviation* $\phi(a, C)$ *between the known discrete population data* $N_d(t_m)$ *and theoretical data* $N(t_m)$:

$$\phi(a, C) = \sum_{m=1}^{M} \left[N(t_m) - N_d(t_m)\right]^2 \tag{1.1.6}$$

By using one of the algorithms of the unconstrained optimization techniques (refer Algorith 2.3.1), we can actually carry out the computation of a *and* C *from the given data.*

Example 1.1.2 *Consider the following ODE*

$$\frac{dx}{dt} = -tx \tag{1.1.7}$$

A general solution of the above ODE is given by

$$x(t) = ae^{-\frac{1}{2}t^2}, \ \ a \text{ is any constant}$$

The solution passing through (t_0, x_0) *in* (t, x) *plane is given by*

$$x(t) = x_0 e^{-\frac{1}{2}(t^2 - t_0^2)} \tag{1.1.8}$$

The following graphics demonstrates the solution curve and the direction field of the above ODE.

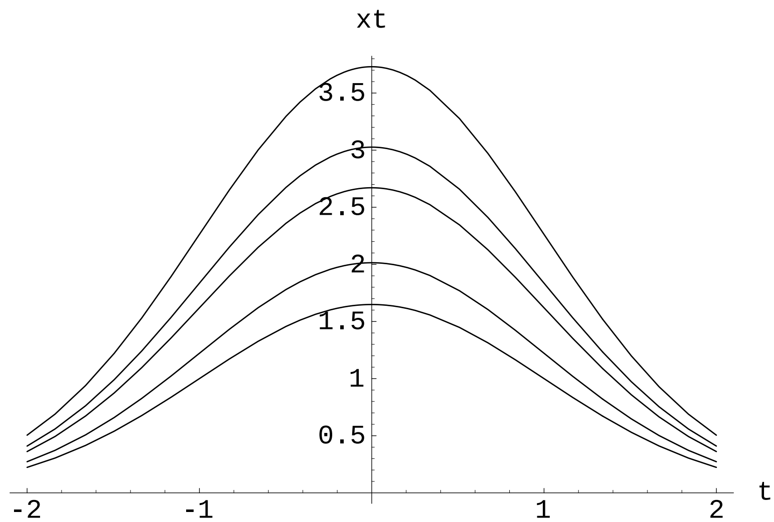

Figure 1.1.2: Solution curves $x(t)$

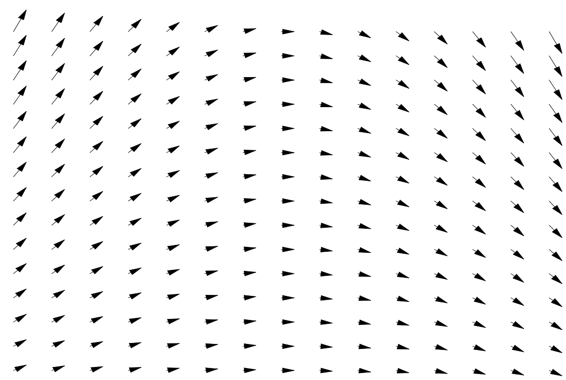

Figure 1.1.3: Direction field of the ODE

If IVP corresponding to Eq. (1.1.1) has a unique solution $\overline{x}(t, t_0, x_0)$ passing through $(t_0, \overline{x}_0)$, no two solution curves of the IVP can intersect. Thus the uniqueness theorem, to be dealt with in the Chapter 3, will be of utmost importance to us.

As we have seen through Example 1.1.2, the direction field for a given differential equation can be an effective way to obtain qualitative information about the behaviour of the system and hence it would be more appropriate to expand on this concept.

To sketch the direction field for a given ODE in 1-dimension, a device called isocline can be used.

Definition 1.1.6 *Isocline is a curve in $t-x$ plane through the direction field along which $p = f(t, x)$ is constant. The family of isoclines is then the family of curves $f(t, x) = p$ in $t - x$ plane.*

Example 1.1.3 *We wish to find the direction field of the following ODE*

$$\frac{1+\dfrac{dx}{dt}}{1-\dfrac{dx}{dt}} = 2\sqrt{t+x}$$

Solving for $\dfrac{dx}{dt}$, we get

$$\frac{dx}{dt}(1+2\sqrt{t+x}) = 2\sqrt{t+x} - 1$$

which gives

$$\frac{dx}{dt} = \frac{2\sqrt{t+x}-1}{2\sqrt{t+x}+1} = p$$

So iscolines are given by

$$\frac{2\sqrt{t+x}-1}{2\sqrt{t+x}+1} = p$$

That is

$$t + x = \frac{1}{2}\left[\frac{1+p}{1-p}\right]^2, -1 \leq p < 1 \tag{1.1.9}$$

Thus isoclines are all lines of slope -1 with $-1 \leq p < 1$. So we get the following graphics. The solution of the ODE passing through (t_0, x_0) is given by

$$(x-t) + \sqrt{t+x} = (x_0 - t_0) + \sqrt{t_0 + x_0} \tag{1.1.10}$$

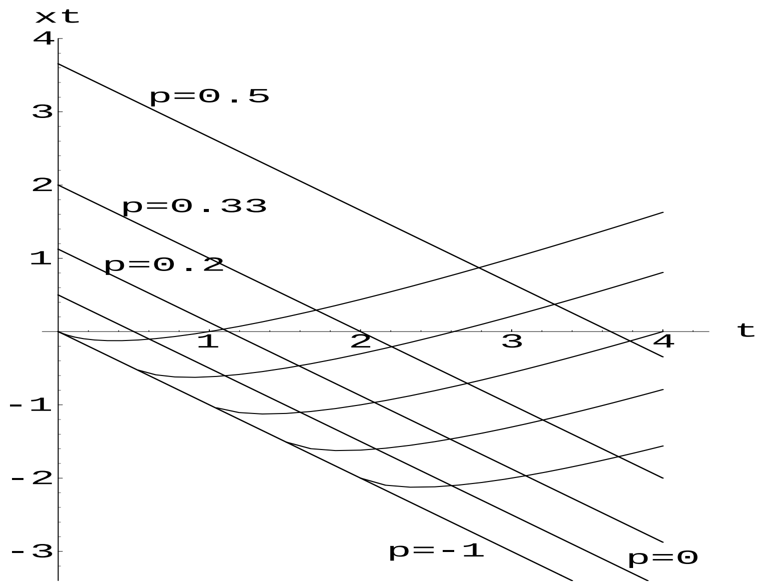

Figure 1.1.4: Isoclines with solution curves

1.2 Classification of ODE

The ODE

$$\frac{d\overline{x}}{dt} = f(t, \overline{x}) \tag{1.2.1}$$

is called linear if f has the form

$$f(t, \overline{x}) = A(t)\overline{x} + \overline{b}(t) \tag{1.2.2}$$

where $A(t) \in \Re^{n \times n}$ and $\overline{b}(t) \in \Re^n$ for all t. A linear ODE is called homogeneous if $\overline{b}(t) \equiv 0$ and if $A(t) = A$ (constant matrix), then we call it a linear ODE with constant coefficients. Linear ODE will be discussed in detail in Chapter 4.

The ODE - Eq. (1.2.1) is called autonomous if f does not depend on t. That is ODE has the form

$$\frac{d\overline{x}}{dt} = f(\overline{x}) \tag{1.2.3}$$

For autonomous ODE we have the following proposition (refer Mattheij and Molenaar [9]).

Proposition 1.2.1 *If $\overline{x}(t)$ is a solution of Eq. (1.2.3) on the interval $I = (a, b)$, then for any $s \in \Re$, $\overline{x}(t+s)$ is a solution of Eq. (1.2.3) on the interval $(a-s, b-s)$. Hence the trajectory $x(t)$ of Eq. (1.2.3) satisfies the following property*

$$\overline{x}(t, t_0, \overline{x}_0) = \overline{x}(t - t_0, 0, \overline{x}_0) \text{ for all } t_0 \in I$$

This implies that the solution is completely determined by the initial state $\bar{x}_0$ at $t = 0$.

A non-autonomous ODE - (1.2.1) is called periodic if

$$f(t+T, \overline{x}) = f(t, \overline{x}) \text{ for some } T > 0 \tag{1.2.4}$$

The smallest of such T is called its period. From Eq. (1.2.4), it follows that the form of the solution is not affected if we shift t_0 to $t_0 \pm nT$, $n \in \aleph$ (set of natural numbers). That is

$$\overline{x}\left(t \pm nT, t_0 \pm nT, \overline{x}_0\right) = \overline{x}(t, t_0, \overline{x}_0)$$

It is not necessary that a solution of periodic ODE is periodic as we see in the following example.

Example 1.2.1

$$\frac{d\overline{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \overline{x} + \begin{bmatrix} 0 \\ 2\cos t \end{bmatrix}$$

$f(t, \overline{x}) = A\overline{x} + \overline{b}(t)$, $\overline{b}(t) = (0, 2\cos t)$. *$f$ is periodic with period 2π but its solution (to be discussed in Chapter 4) is given by*

$$\overline{x}(t) = (t - \sin t, \sin t + t\cos t)$$

However, we have the following theorem regarding the periodic solution of the periodic ODE.

Proposition 1.2.2 *If a solution of a periodic ODE is periodic. Then it has the same period as the vector field.*

Proof : Assume that $f(t, \overline{x})$ has a period T and the solution $\overline{x}(t)$ has period S with $S \neq T$. Because $\bar{x}$ is periodic, its derivative $\dot{\overline{x}}(t) = f(t, \overline{x})$ will also be periodic with period S as well T. That is

$$f(t+S, \overline{x}) = f(t, \overline{x})$$

Also, T periodicity of f implies that

$$f(t + S - nT, \overline{x}) = f(t, \overline{x})$$

We choose n such that $0 < S - nT < T$. It implies that f is periodic with a period smaller than T, a contradiction. Hence $T = S$ and the solution $\bar{x}(t)$ has the same period as the vector field. ∎

Example 1.2.2 *(Predator-Prey Model)*
Predator-prey model represents the interaction between two species in an environment. For example, we shall focus on sharks and small fish in sea.
If the food resource of sharks (fish) is non-existent, sharks population exponentially decays and is increased with the existence of fish. So, the growth rate of sharks $\left(\frac{1}{S}\frac{dS}{dt}\right)$ *is modelled as* $-k$ *without the fish population and* $-k + \lambda F$ *with fish population* F*. Thus*

$$\frac{dS}{dt} = S(-k + \lambda F)$$

For the growth rate of fish, we note that it will decrease with the existence of sharks and will flourish on small plankton (floating organism in sea) without sharks. Thus

$$\frac{dF}{dt} = F(a - cS)$$

Thus, we have, what is called the famous Lotka - Volterra model for the predator-prey system

$$\frac{dF}{dt} = F(a - cS) \qquad (1.2.4(a))$$

$$\frac{dS}{dt} = S(-k + \lambda F) \qquad (1.2.4(b))$$

This reduces to the following autonomous system of ODE in $\Re^2$

$$\frac{d\bar{x}}{dt} = f(\bar{x})$$

where

$$\overline{x} = (F, S), f(\bar{x}) = (F(a - cS), S(-k + \lambda F))$$

The equilibrium points are given by

$$\hat{F} = k/\lambda, \hat{S} = a/c; \hat{F} = 0, \hat{S} = 0.$$

We shall have the discussion of the phase-plane analysis in the Chapter 5. However, if we linearize the system given by Eq.(1.2.4) around the equilibrium point

$\hat{F} = k/\lambda, \hat{S} = a/c$, we get the following linear system (refer Section 2.3 for linearization)

$$
\begin{aligned}
\frac{d\overline{x}}{dt} &= A\overline{x}(t) \\
\overline{x} &= (F-\hat{F}, S-\hat{S}) \\
A &= \begin{bmatrix} 0 & -\frac{kc}{\lambda} \\ \frac{a\lambda}{c} & 0 \end{bmatrix}
\end{aligned} \tag{1.2.5}
$$

The solution of the IVP corresponding to Eq. (1.2.5) is given by

$$S = \hat{S} + S_0 \cos wt + \frac{a\lambda}{cw} F_0 \sin wt \tag{1.2.6a}$$

$$F = \hat{F} + F_0 \sin wt - \frac{cw}{a} S_0 \cos wt \tag{1.2.6b}$$

(Refer Section 4.3 for solution analysis of this linear system).

Thus, the solution is periodic with period $\frac{2\pi}{w} = 2\pi(ak)^{-1/2}$. The following graphics depict the fish and shark population in $t - F/S$ plane.

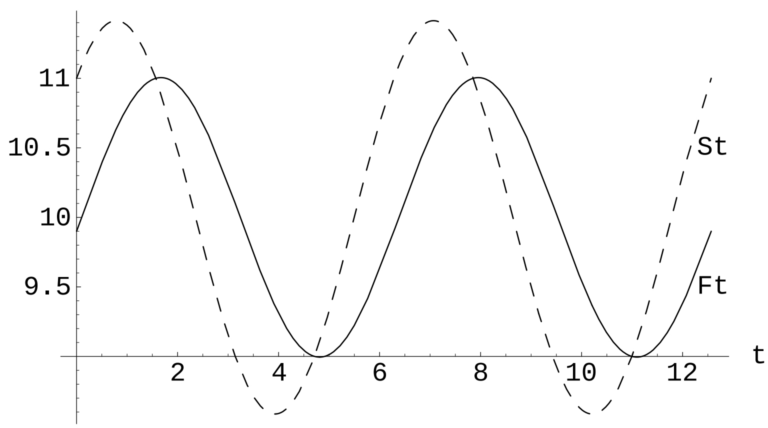

Figure 1.2.1: Fish and shark population with time

The average population of the general predator-prey system given by Eq. (1.2.4) is obtained as follows.

We have

$$\frac{1}{S}\frac{dS}{dt} = -k + \lambda F$$

This gives

$$ln\left[\frac{S(t)}{S(t_0)}\right] = -k(t-t_0) + \lambda \int_{t_0}^{t} F(\tau) d\tau \tag{1.2.7}$$

In view of the information obtained from the linearized model, we can assume that both S and F are periodic of period $T = t_f - f_0$. That is, $S(t_f) = S(t_0)$ and $F(t_f) = F(t_0)$ and hence Eq. (1.2.7) gives

$$0 = -kT + \lambda \int_{t_0}^{t_0+T} F(\tau)d\tau$$

This implies that

$$\frac{1}{T}\left[\int_{t_0}^{t_0+T} F(\tau)d\tau\right] = k/\lambda \qquad (1.2.8(a))$$

and similarly

$$\frac{1}{T}\left[\int_{t_0}^{t_0+T} S(\tau)d\tau\right] = a/c \qquad (1.2.8(b))$$

So in predator-prey system, no matter what the solution trajectory is, the average population remains around the equilibrium point.
It is also interesting to analyse the men's influence on the above ecosystem if we fish both predator and prey. Then, we have

$$\begin{aligned} \frac{dF}{dt} &= F(a - cS) - \sigma_1 F \\ \frac{dS}{dt} &= S(-k + \lambda F) - \sigma_2 S \end{aligned}$$

This is equivalent to

$$\begin{aligned} \frac{dF}{dt} &= F(a' - cS) \\ \frac{dS}{dt} &= S(-k' + \lambda F) \\ a' = a - \sigma_1, \qquad k' &= k + \sigma_2 \end{aligned}$$

Hence it follows that the average population of predator is $\frac{a'}{c} = \frac{a - \sigma_1}{c}$ *(decreases) and that of prey is* $\frac{k'}{c} = \frac{k + \sigma_2}{c}$ *(increases).*

1.3 Higher Order ODE

Assume that $x(t)$ is a scalar valued $n-$times continuously differentiable function on an interval $I \subset \Re$. Let us denote the derivative $\frac{d^k x}{dt^k}$ by $x^{(k)}(t)$.

We shall be interested in the following higher order ordinary differential equation

$$x^{(n)}(t) = g(t, x(t), x^{(1)}(t), \ldots, x^{(n-1)}(t)) \qquad (1.3.1)$$

where $g : I \times \Re \times \Re \times \ldots \times \Re \to \Re$ is a given mapping. We define a vector $\overline{x}(t)$ with components $x_i(t), i = 1, \ldots n$ by

$$x_i(t) = x^{(i-1)}(t), \; x_1(t) = x(t); \;\; 2 \leq i \leq n \qquad (1.3.2)$$

and the vector field $f(t, \bar{x})$ by

$$f(t, \bar{x}) = (x_2, x_3, \ldots, g(t, x_1, \ldots, x_n))$$

Then the higher order ODE - Eq. (1.3.1) is equivalent to the following first order ODE

$$\frac{d\overline{x}}{dt} = f(t, \overline{x})$$

A linear higher order ODE has the form

$$x^{(n)}(t) + a_{n-1}(t)x^{(n-1)}(t) + \ldots + a_0(t)x(t) = b(t) \tag{1.3.3}$$

where $a_i(t)(0 \leq i \leq n-1)$ and $b(t)$ are given functions. Then in view of Eq. (1.3.2) - Eq. (1.3.3) we have

$$\begin{aligned} \frac{dx_1}{dt} &= x_2(t) \\ \frac{dx_2}{dt} &= x_3(t) \\ &\vdots \\ \frac{dx_n}{dt} &= b(t) - a_0(t)x_1(t) - a_1(t)x_2(t) - \cdots - a_{n-1}(t)x_n(t) \end{aligned}$$

This is equivalent to the following first order system

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t) + \overline{b}(t) \tag{1.3.4}$$

where

$$\begin{aligned} A(t) &= \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-1} \end{bmatrix} \\ \overline{b}(t) &= (0, 0, \ldots, b(t)) \end{aligned}$$

The matrix $A(t)$ is called the companion matrix.

Example 1.3.1 *(Mechanical Oscillations)*
A particle of mass m is attached by a spring to a fixed point as given in the following diagram.

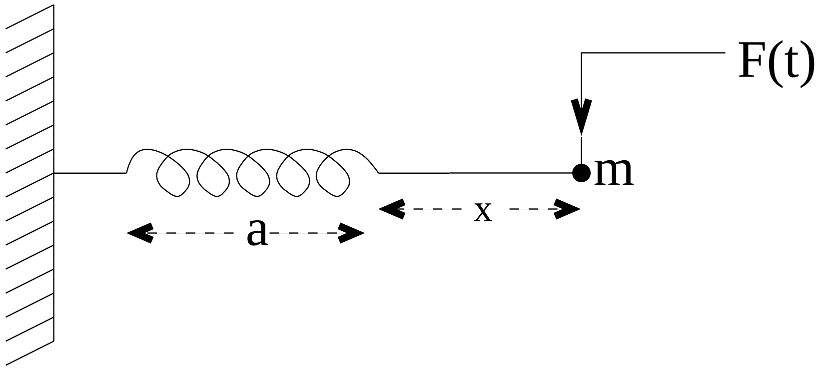

Figure 1.3.1: Mechanical oscillations

We assume that spring obeys Hook's law (tension is proportional to its extension) and resistance (damping) is proportional to the particle speed. The external force applied to the particle is $F(t)$. By equilibrium of forces we have

$$mF = m\frac{d^2x}{dt^2} + T + mk\frac{dx}{dt}$$

By Hook's Law, we get $T = \dfrac{\lambda x}{a}$ and hence

$$F = \frac{d^2x}{dt^2} + \frac{\lambda}{ma}x + k\frac{dx}{dt} \tag{1.3.5}$$

Equivalently, we get the following second order equation modelling the spring problem

$$\frac{d^2x}{dt^2} + k\frac{dx}{dt} + \omega^2 x = F(t), \quad \omega^2 = \frac{\lambda}{ma} \tag{1.3.6}$$

Case 1: No resistance and no external force (Harmonic Oscillator)

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \quad \omega^2 = \frac{\lambda}{ma} \tag{1.3.7}$$

Eq. (1.3.7) is equivalent to the following system of first order differential equations

$$\frac{d\overline{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -\dfrac{\lambda}{ma} & 0 \end{bmatrix} \overline{x} \tag{1.3.8}$$

It is easy to see the solution of the above system is given by (refer Section 4.4)

$$x(t) = x_0 \cos\omega(t - t_0) + \frac{x_0}{\omega}\sin\omega(t - t_0). \tag{1.3.9}$$

$$\dot{x}(t) = -x_0\omega\sin\omega(t - t_0) + x_0\cos\omega(t - t_0) \tag{1.3.10}$$

Equivalently, we have

$$x(t) = \alpha \cos(\omega t + \beta)$$

This is a simple harmonic motion with period $\dfrac{2\pi}{\omega}$ and amplitude α. ω is called its frequency.

We can easily draw plots of phase-space (normalized $\omega = 1$) and also solution curve of the ODE - Eq. (1.3.8).

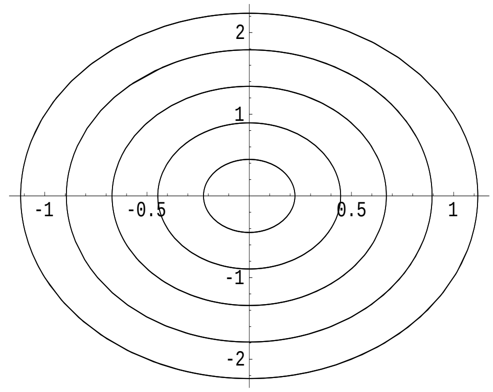

Figure 1.3.2: Phase space of the linear system

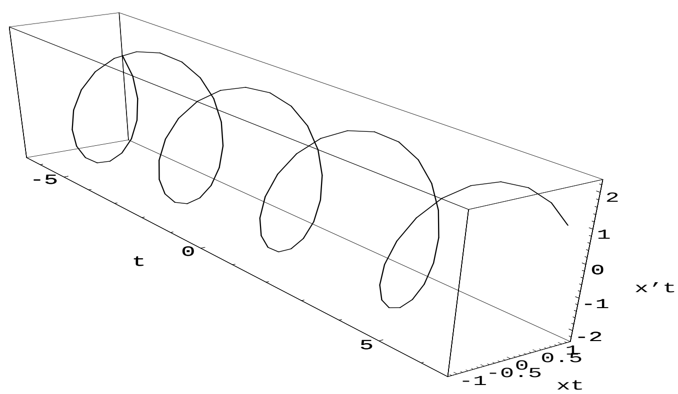

Figure 1.3.3: Solution curve of the linear system

Case 2: Solution with resistance

$$\frac{d\overline{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -k \end{bmatrix} \overline{x}(t) \tag{1.3.11}$$

(i) $k^2 - 4\omega^2 < 0$

The solution curve $x(t)$ is a damped oscillation given by (refer Section 4.4)

$$\begin{aligned} x(t) &= \exp(\frac{-kt}{2})(A\cos bt + B\sin bt) \\ b &= \frac{\sqrt{4\omega^2 - k^2}}{2} \end{aligned}$$

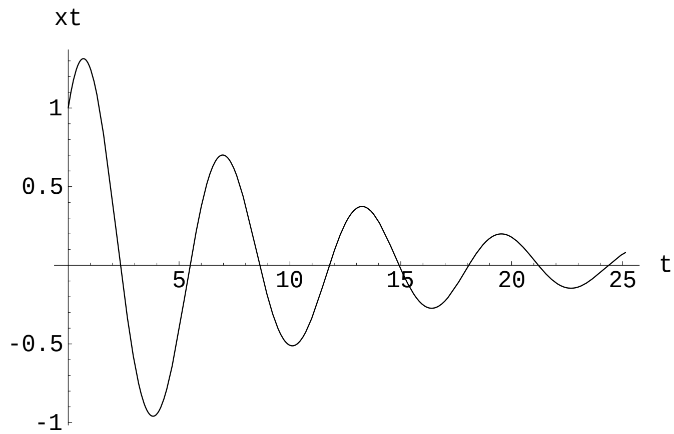

Figure 1.3.4: Solution curve for Case2(i)

The oscillation curves for the other cases are as under.
(ii) $(k^2 - 4\omega^2) = 0$

$$x(t) = A\exp(\frac{-kt}{2}) + Bt\exp(\frac{-kt}{2})$$

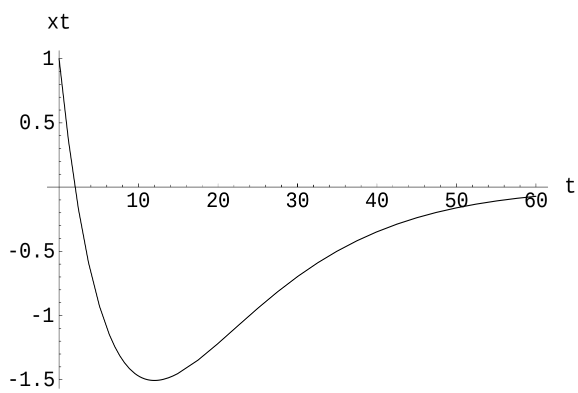

Figure 1.3.5: Solution surve for case2(ii)

(iii) $k^2 - 4\omega^2 > 0$

$$\begin{aligned} x(t) &= \exp(\frac{-kt}{2})\left[A_1 e^{ct} + A_2 e^{-ct}\right] \\ c &= \sqrt{k^2 - 4\omega^2} \end{aligned}$$

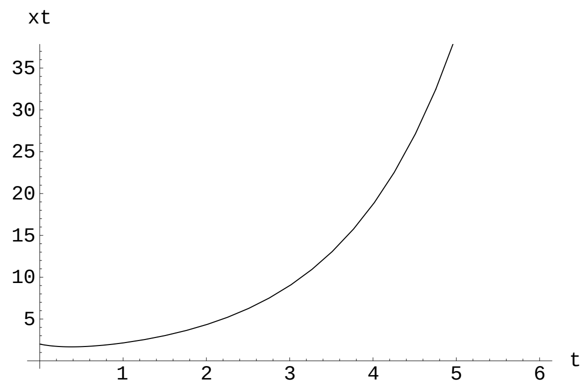

Figure 1.3.6: Solution curve for case2(iii)

Case 3: Effect of resistance with external force

$$\frac{d\overline{x}(t)}{dt} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -k \end{bmatrix} \overline{x} + \overline{F}(t), \overline{F}(t) = (0, F(t))$$

The solution $x(t)$ is given by

$$x = x_p + x_c$$

where x_c is the solution of the homogeneous equation with $F(t) = 0$. We have seen that $x_c \to 0$ as $t \to \infty$. A particular solution (response function) $x_p(t)$ of the above system is given by

$$x_p(t) = \frac{F_0}{D}\left[\cos\left(\beta t - \phi\right)\right]$$

corresponding to the external force (input function) $F(t) = F_0 \cos \beta t$.
The constant D is given by

$$D = \left[(\omega^2 - \beta^2)^2 + k^2\beta^2\right]^{1/2}$$

and

$$\sin\phi = \frac{k\beta}{D}, \cos\phi = \frac{\omega^2 - \beta^2}{D}$$

Thus the forced oscillations have the same time period as the applied force but with a phase change ϕ and modified amplitude

$$\frac{F_0}{D} = \frac{F_0}{[(\omega^2 - \beta^2)^2 + k^2\beta^2]^{1/2}}$$

Amplitude modification depends not only on the natural frequency and forcing frequency but also on the damping coefficient k.
As $k \to 0$ we have

$$x_p \to \frac{F_0}{(\omega^2 - \beta^2)} \to \infty \text{ as } \omega \to \beta.$$

For $k = 0$, the response is given by

$$x_p = \frac{F_0}{2\omega} t \sin wt$$

which implies that the system resonates with the same frequency but with rapidly increasing magnitude, as we see in the following graph.

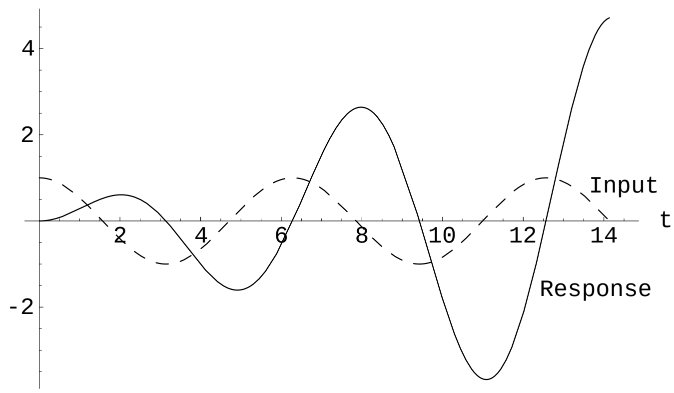

Figure 1.3.7: Input-Response curve

Example 1.3.2 *(Satellite Problem)*
A satellite can be thought of a mass orbiting around the earth under inverse square law field.

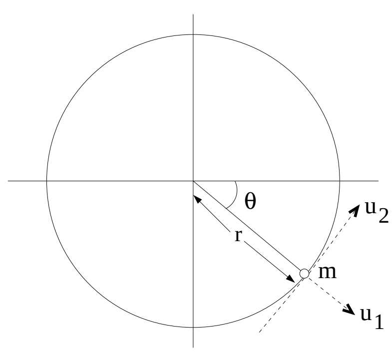

Figure 1.3.8: Satellite in orbit around the Earth

We assume that $m = 1$ and also the satellite has thrusting capacity with radial thrust u_1 and tangential thrust u_2.
Equating forces in normal and tangential direction on the orbiting satellite, we get

$$\begin{aligned}
\left[\frac{d^2r}{dt^2} - r(t)(\frac{d\theta}{dt^2})^2\right] &= -\frac{k}{r^2(t)} + u_1(t) \\
\left[r\frac{d^2\theta}{dt^2} + 2\frac{d\theta}{dt}\frac{dr}{dt}\right] &= u_2(t)
\end{aligned}$$

This gives a pair of second order differential equations

$$\frac{d^2r}{dt^2} = r(t)(\frac{d\theta}{dt})^2 - \frac{k}{r^2(t)} + u_1(t) \tag{1.3.12a}$$

$$\frac{d^2\theta}{dt^2} = -\frac{2}{r(t)}\frac{d\theta}{dt}\frac{dr}{dt} + \frac{u_2(t)}{r(t)} \tag{1.3.12b}$$

] If $u_1 = 0 = u_2$, then one can show that Eq. (1.3.12) has a solution given by $r(t) = \sigma,\ \theta(t) = \omega t$ ($\sigma,\ \omega$ are constant and $\sigma^3\omega^2 = k$). Make the following change of variables:

$$x_1 = r - \sigma,\ x_2 = \dot{r},\ x_3 = \sigma(\theta - \omega t),\ x_4 = \sigma(\dot{\theta} - \omega)$$

This gives

$$r = x_1 + \sigma,\ \dot{r} = x_2,\ \theta = \frac{x_3}{\sigma} + \omega t,\ \dot{\theta} = \frac{x_4}{\sigma} + \omega$$

So Eq. (1.3.12) reduces to

$$\begin{aligned}
\frac{dx_1}{dt} &= x_2 \\
\frac{dx_2}{dt} &= (x_1 + \sigma)(\frac{x_4}{\sigma} + \omega)^2 - \frac{k}{(x_1 + \sigma)^2} + u_1 \\
\frac{dx_3}{dt} &= x_4 \\
\frac{dx_4}{dt} &= -2\sigma(\frac{x_4}{\sigma} + \omega)\frac{x_2}{(x_1 + \sigma)} + \frac{u_2\sigma}{(x_1 + \sigma)}
\end{aligned} \tag{1.3.13}$$

Eq. (1.3.13) is a system of nonlinear ordinary diferential equations involving the forcing functions (controls) u_1 and u_2 and can be written in the compact form as

$$\frac{d\bar{x}}{dt} = f(\bar{x},\ \bar{u}),\ \overline{x}(t) \in \Re^4,\ \overline{u}(t) \in \Re^2 \tag{1.3.14}$$

Here f is a vector function with components f_1, f_2, f_3, f_4 given by

$$\begin{aligned}
f_1(x_1, x_2, x_3, x_4; u_1, u_2) &= x_2 \\
f_1(x_1, x_2, x_3, x_4; u_1, u_2) &= (x_1 + \sigma)(\frac{x_4}{\sigma} + \omega)^2 - \frac{k}{(x_1 + \sigma)^2} + u_1 \\
f_1(x_1, x_2, x_3, x_4; u_1, u_2) &= x_4 \\
f_1(x_1, x_2, x_3, x_4; u_1, u_2) &= -2\sigma(\frac{x_4}{\sigma} + \omega)\frac{x_2}{(x_1 + \sigma)} + \frac{u_2\sigma}{(x_1 + \sigma)}
\end{aligned}$$

We shall be interested in the solution, $\overline{x}(t) \in \Re^4$ of linearized equation corresponding to Eq. (1.3.14) in terms of the control vector $\overline{u}(t) \in \Re^2$.

1.4 Difference Equations

In some models (as we shall see subsequently), the state vector $\overline{x}(t)$ may not depend on t continuously. Rather, $\overline{x}(t)$ takes values at discrete set of points $\{t_1, t_2, \ldots, t_k, \cdots\}$

In such a situation we use difference quotients instead of differential quotients and that leads to difference equations. Suppose there exists a sequence of vector fields $f_i(\overline{x}) : \Omega \to \Re^n, \Omega \subseteq \Re^n$. Then the first order difference equation has the form

$$\overline{x}_{i+1} = f_i(\overline{x}_i), \bar{x}_i \in \Omega, i = 1, 2 \ldots \qquad (1.4.1(a))$$

If in addition

$$\overline{x}_0 = \overline{z}_0 \text{ (a given vector in } \Re^n) \qquad (1.4.1(b))$$

then Eq. (1.4.1(a)) - Eq. (1.4.1(b)) is called the IVP corresponding to a difference equation.

As for ODE, we can define orbit or trajectory and solution curve as discrete subsets of $\Re^n$ and $\Re^{n+1}$, respectively in an analogous way. A stationary point or equilibrium point of the difference equation Eq. (1.4.1) is a constant solution $\overline{x}$ such that

$$\overline{x} = f_i(\overline{x}), i \in \aleph$$

Eq. (1.4.1) is called linear if it has the form

$$\overline{x}_{i+1}(t) = A_i\overline{x}_i + \overline{b}_i, i \in \aleph$$

where $A_i \in \Re^{n \times n}$ and $\overline{b}_i \in \Re^n$.

For the scalar case $n = 1$, we have the linear difference equation

$$x_{i+1} = a_i x_i + b_i, i \in \aleph$$

Its solution is given by

$$x_i = (\Pi_{j=0}^{i-1}(a_j))x_0 + \sum_{j=0}^{i-1}(\Pi_{l=j+1}^{i} a_l)b_j$$

A k^{th} order difference equation in 1-dimension is given by

$$x_{i+1} = g_i(x_i, x_{i+1}, \ldots, x_{i+1-k}),\ i = k-1, k, \ldots$$

A linear k^{th} order difference equation is given by

$$x_{i+1} = \sum_{j=1}^{k} a_{ij} x_{i-j+1} + b_i.$$

This can be written as the first order system

$$\overline{x}_{i+1} = A_i \overline{x}_i + \overline{b}_i, \quad i \geq 0 \tag{1.4.2}$$

where

$$\bar{x}_i = \begin{bmatrix} x_i \\ \vdots \\ x_{i+k-1} \end{bmatrix}, \quad \bar{b}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ b_{i+k} \end{bmatrix}$$

$$A_i = \begin{bmatrix} 0 & 1 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \\ a_{i+k-1,k} & a_{i+k-1,k-1} & \cdots & a_{i+k-1,1} \end{bmatrix}$$

For the homogeneous case $\overline{b}_i = 0$, with $A_i = A$ for all i, we get

$$\overline{x}_{i+1} = A\overline{x}_i \tag{1.4.3}$$

The solution of this system is of the form

$$\overline{x}_i = r^i \overline{c}$$

where scalar r and vector $\overline{c}$ are to be determined.
Plugging this representation in Eq. (1.4.3) we get

$$(A - rI)\overline{c} = 0 \tag{1.4.4}$$

This is an eigenvalue problem. If A has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with linearly independent eigenvectors $\overline{c}_1, \ldots \overline{c}_n$, then the general solution of Eq. (1.4.3) is given by

$$\overline{x}_m = \sum_{i=1}^{n} d_i \lambda_i^m \overline{c}_i, \ \ m = 1, 2 \ldots \tag{1.4.5}$$

If the initial value is $\overline{x}_0$ then $\overline{d} = [d_1, \ldots, d_n]$ is given by $\overline{d} = C^{-1}(\overline{x}_0)$. Here $C = [\bar{c_1}, \bar{c_2}, \cdots, \bar{c_n}]$. So the solution to the IVP corresponding to Eq. (1.4.3) is completely given by Eq. (1.4.5).

Example 1.4.1 *(Discrete One Species Population Dynamics Model)*
We wish to measure population changes in a species with one year age distribution. Let $N_i(t)$ *denote the number of* i^{th} *year old species and let* b_i, d_i *the corresponding birth and death rates, respectively* $(0 \le i \le M)$. *Then we have*

$$\begin{aligned} N_0(t+\triangle t) &= b_0 N_0(t) + b_1 N_1(t) \\ &\quad + \ldots + b_M N_M(t),\ (\triangle t = 1) && (1.4.6(a)) \\ N_{i+1}(t+\triangle t) &= (1-d_i)N_i(t),\ 0 \le i \le M-1 && (1.4.6(b)) \end{aligned}$$

Let $\overline{N}(t) = (N_0(t), N_1(t), \ldots, N_M(t))$. *Then Eq. (1.4.6) becomes*

$$\overline{N}(t+\triangle t) = A\overline{N}(t) \tag{1.4.7}$$

where

$$A = \begin{bmatrix} b_0 & b_1 & \cdots & b_M \\ 1-d_0 & 0 & \cdots & 0 \\ 0 & 1-d_1 & & \cdots \\ \vdots & & \vdots & \cdots \\ 0 & 0 & \cdots & 1-d_{M-1} \end{bmatrix}$$

If we denote $\overline{N}_m = \overline{N}(m\triangle t)$, *then Eq. (1.4.7) can be viewed as a difference equation of the form*

$$\overline{N}_{m+1} = A\overline{N}_m \tag{1.4.8}$$

This is of the type given by Eq. (1.4.3) and hence the solution of the above equation is of the form

$$\overline{N}_m = \sum_{i=1}^{M} a_i \lambda_i^m \overline{\phi}_i$$

where $\{\lambda_i\}_{i=1}^M$ *and* $\{\overline{\phi}_i\}_{i=1}^M$ *are eigenvalues (distinct) and linearly independent eigenvectors of A, respectively.* $\overline{a} = [a_1, \ldots, a_m]$ *is a constant vector. If the initial population* $\overline{N}_0$ *is known, the constant vector* $\overline{a}$ *is given by* $\overline{a} = C\overline{N}_0$ *where the nonsingular matrix C is given by*

$$C = [\overline{\phi}_1, \overline{\phi}_2, \ldots, \overline{\phi}_M]$$

It is to be noticed that the population of each age group grows and decays depending upon the sign of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_M$.

1.5 Discretization

In this section we shall describe some well-known one-step numerical methods, which are used to solve ODE. However, we shall not be concerned with the stability and convergence aspects of these methods.
We also introduce to reader the concept of neural solution of ODE. This is also based on discretization but uses neural network approach (refer Haykin [6]).

It is interesting to observe that sometimes neural solution is a better way of approximating the solution of ODE.
Let us begin with scalar ordinary differential equation

$$\frac{dx}{dt} = f(t, x) \tag{1.5.1}$$

corresponding to the scalar field $f(t, x) : I \times \Re \to \Re$.
Let $t_0, t_1, \ldots, t_N$ be a set of time points (called grid-points) wherein we would like to approximate the solution values $x(t_i), 0 \leq i \leq N$.
Integrating the above equation on $[t_i, t_i + 1]$ we get

$$x(t_{i+1}) = x(t_i) + \int_{t_i}^{t_{i+1}} f(s, x(s))ds$$

If f is approximated by its values at t_i, we get the Euler's forward method

$$x(t_{i+1}) = x(t_i) + hf(t_i, x(t_i)),\ h = x_{i+1} - x_i \tag{1.5.2}$$

which is the so-called explicit, one-step method.
On the other hand, if f is approximated by its values at t_{i+1} we get Euler's backward method

$$x(t_{i+1}) = x(t_i) + hf(t_{i+1}, x(t_{i+1})) \tag{1.5.3}$$

which has to be solved implicitly for $x_{t_{i+1}}$.
If the integral $\int_{t_i}^{t_{i+1}} f(s, x(s))ds$ is approximated by the trapezoidal rule, we get

$$x(t_{i+1}) = x(t_i) + \frac{1}{2}h\left[f(t_i, x(t_i)) + f(t_{i+1}, x(t_{i+1}))\right] \tag{1.5.4}$$

which is again implicit.
Combining Eq. (1.5.2) and Eq. (1.5.4) to eliminate $x(t_{i+1})$, we get the Heun's method (refer Mattheij and Molenaar[9])

$$x(t_{i+1}) = x(t_i) + \frac{1}{2}h\left[f(t_i, x(t_i)) + f(t_{i+1}, x(t_i) + hf(t_i, x(t_i)))\right] \tag{1.5.5}$$

An important class of one-step numerical methods is Runge-Kutta method. Consider the integral equation

$$x(T) = x(t) + \int_{t}^{T} f(s, x(s))ds \tag{1.5.6}$$

Approximating the integral $\int_{t_i}^{t_{i+1}} f(s, x(s))ds$ by a general quadrature formula $\sum_{j=1}^{m} \beta_j f(t_{ij}, x(t_{ij}))$, we get

$$x(t_{i+1}) = x(t_i) + h\sum_{j=1}^{m} \beta_j f(t_{ij}, x(t_{ij})) \tag{1.5.7}$$

where $h = t_{i+1} - t_i$ and $t_{ij} = t_i + \rho_j h (0 \leq \rho_j \leq 1)$ are nodes on the interval $[t_i, t_{i+1}]$.
As $x(t_{ij})$ are unknowns, to find them we apply another quadrature formula for the integral $\int_{t_i}^{t_{ij}} f(s, x(s)) ds$ to get

$$x(t_{ij}) = x(t_i) + h \sum_{l=1}^{m} r_{jl} f(t_{il}, x(t_{il})) \tag{1.5.8}$$

Combining Eq. (1.5.7) and Eq. (1.5.8), we get the Runge-Kutta formula

$$\begin{aligned} x(t_{i+1}) &= x(t_i) + h \sum_{j=1}^{m} \beta_j k_j \qquad (1.5.9) \\ k_j &= f(t_i + \rho_j h, x(t_i) + h \sum_{l=1}^{m} r_{jl} k_l) \end{aligned}$$

Eq. (1.5.9) is explicit if $r_{jl} = 0, l \geq j$. If $\rho_1 = 0, \rho_2 = 1, r_{11} = 0 = r_{12}, r_{21} = 1, r_{22} = 0$ and $\beta_1 = 1/2 = \beta_2$, we get the Heun's formula given by Eq. (1.5.5). Also, it is easy to get the following classical Runge-Kutta method from Eq. (1.5.9)

$$x(t_{i+1}) = x(t_i) + h \left[\frac{1}{6} k_1 + \frac{1}{3} k_2 + \frac{1}{3} k_3 + \frac{1}{6} k_4 \right] \tag{1.5.10}$$

$$\begin{aligned} k_1 &= f(t_i, x(t_i)) \\ k_2 &= f(t_i + \frac{1}{2} h, x(t_i) + \frac{1}{2} h k_1) \\ k_3 &= f(t_i + \frac{1}{2} h, x(t_i) + \frac{1}{2} h k_2) \\ k_4 &= f(t_i + h, x(t_i) + h k_3) \end{aligned}$$

In case we have $f(t, \overline{x})$ as a vector field, we obtain a similar Runge - Kutta method, with scalars $x(t_i)$, $x(t_i + 1)$ being replaced by vectors $\overline{x}(t_i)$, $\overline{x}(t_{i+1})$, and scalars k_i being replaced by vectors $\overline{k}_i (1 \leq i \leq 4)$.

We now briefly discuss the concept of neural solution of the IVP associated with Eq. (1.5.1). A neural solution (refer Logaris et al [8]) $x_N(t)$ can be written as

$$x_N(t) = x_0 + t N(t, \overline{W}) \tag{1.5.11}$$

where $N(t, \overline{W})$ is the output of a feed forward neural network (refer Haykin [6]) with one input unit t and network weight vector $\overline{W}$. For a given input t, output of the network is $N = \sum_{i=1}^{k} v_i \sigma(z_i)$ where $z_i = w_i t - u_i$, w_i denotes the weight from the input unit t to the hidden unit i, v_i denotes the weight from

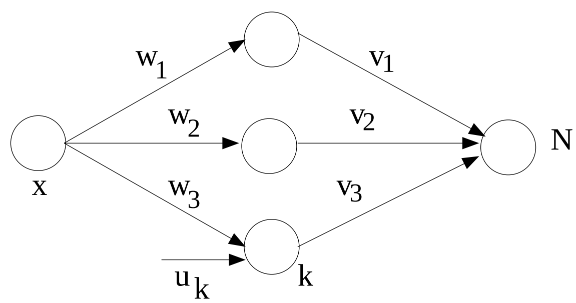

Figure 1.5.1: Neural network with one input

the hidden unit i to the output and u_i denotes the bias of the unit i, $\sigma(z)$ is the sigmoid transfer function

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

and k is the number of hidden neurons.

We first determine the network parameters w_i, v_i and u_i in such a manner that $x_N(t)$ satisfies Eq. (1.5.1) in some sense. For this we discretize the interval $[t_0, t_f]$ as $t_0 < t_i < \ldots < t_m = t_f$.
As $x_N(t)$ is assumed to be a trial solution, it may not be exactly equal to $f(t_j, x_N(t_j))$ at point t_j, $0 \leq j \leq m$. Hence we train the parameter vector $\overline{W}$ in such a way that it minimizes the error function $\phi(\overline{W})$ given as

$$\phi(\overline{W}) = \sum_{j=0}^{m} \left(\frac{dx_N}{dt}|_{t=t_j} - f(t_j, x_N(t_j)) \right)^2$$

One can use any unconstrained optimization algorithms (refer Joshi and Kannan [7]) to minimize ϕ w.r.t. the neural network parameters to obtain their optimal values u_i^*, v_i^*, w_i^* to get the neutral solution $x_N(t)$ given by Eq. (1.5.11)

This methodology will be clear from the following example, wherein we compute numerical solutions by Euler's and Runge-Kutta methods and also neural solution. Further, we compare these solutions with the exact solution.

Example 1.5.1 *Solve the IVP*

$$\begin{aligned} \frac{dx}{dt} &= x - x^2, t \in [0,1] \quad (1.5.12) \\ x(0) &= \frac{1}{2} \end{aligned}$$

The exact solution of this ODE (Bernoulli's differential equation) is given by

$$x(t) = \frac{1}{1+e^{-t}} \quad (1.5.13)$$

(Refer Section 6 of this chapter).

For neural solution we take three data points $t_j = 0, 0.5, 1$ and three hidden neurons with a single hidden layer network. The neural solution is of the form

$$x_N(t) = \frac{1}{2} + t\sum_{i=1}^{3} \frac{v_i}{\left[1 + e^{(-w_i t + u_i)}\right]} \tag{1.5.14}$$

The weight vector $\overline{W} = (w_1, w_2, w_3, v_1, v_2, v_3, u_1, u_2, u_3)$ is computed in such a way that we minimize $\phi(\overline{W})$ given by

$$\begin{aligned}
\phi(\overline{W}) &= \sum_{j=1}^{3}\left[\frac{dx_N(t)}{dt}|_{t=t_j} - f(t_j, x_N(t_j))\right]^2 \\
&= \sum_{j=1}^{3}\Bigg[\left(\frac{1}{2} + \frac{v_1}{1+e^{u_1 - w_1 t_j}} + \frac{v_2}{1+e^{u_2 - w_2 t_j}} + \frac{v_3}{1+e^{u_3 - w_3 t_j}}\right) \\
&\quad - \left(\frac{v_1}{1+e^{u_1 - w_1 t_j}} + \frac{v_2}{1+e^{u_2 - w_2 t_j}} + \frac{v_3}{1+e^{u_3 - w_3 t_j}}\right) t_j \\
&\quad + \left(\frac{e^{u_1 - w_1 t_j} v_1 w_1}{(1+e^{u_1 - w_1 t_j})^2} + \frac{e^{u_2 - w_2 t_j} v_2 w_2}{(1+e^{u_2 - w_2 t_j})^2} + \frac{e^{u_3 - w_3 t_j} v_3 w_3}{(1+e^{u_3 - w_3 t_j})^2}\right) t_j \\
&\quad + \left(\frac{1}{2} + \left(\frac{v_1}{1+e^{u_1 - w_1 t_j}} + \frac{v_2}{1+e^{u_2 - w_2 t_j}} + \frac{v_3}{1+e^{u_3 - v_3 t_j}}\right) t_j\right)^2\Bigg]^2
\end{aligned}$$

We have used the steepest descent algorithm (refer to Algorithm 2.3.1) to compute the optimal weight vector $\overline{W}$. The neural solution is given by Eq. (1.5.14). We note that the neural solution is a continuous one. We can compare the values of this solution at a discrete set of points with the ones obtained by Euler and Runge-Kutta method (refer Conte and deBoor [3]). The following tables give the distinction between the three types of approximate solutions.

Table 1.5.1: Euler solution and actual solution

Euler's method	Actual solution	Absolute difference
0.525,	0.52498	0.00002
0.54994	0.54983	0.00010
0.57469	0.57444	0.00025
0.59913	0.59869	0.00044
0.62315	0.62246	0.00069
0.64663	0.64566	0.00097
0.66948	0.66819	0.00129
0.69161	0.68997	0.00163
0.71294	0.71095	0.00199
0.73340	0.73106	0.00234

Table 1.5.2: Runge-Kutta solution and actual solution

Runge Kutta method	Actual solution	Absolute difference
0.52498	0.52498	1.30339^{-9}
0.54983	0.54983	2.64162^{-9}
0.57444	0.57444	4.05919^{-9}
0.59869	0.59869	5.59648^{-9}
0.62246	0.62246	7.28708^{-9}
0.64566	0.64566	9.15537^{-9}
0.66819	0.66819	1.12147^{-8}
0.68997	0.68997	1.34665^{-8}
0.71095	0.71095	1.58997^{-8}
0.73106	0.73106	1.84916^{-8}

Table 1.5.3: Neural solution and actual solution

Neural Network sol	Actual sol	Absolute Difference
0.52511	0.52498	0.00013
0.54973	0.54983	0.00011
0.57383	0.57444	0.00062
0.59741	0.59869	0.00127
0.62048	0.62246	0.00198
0.64302	0.64566	0.00264
0.66503	0.66819	0.00316
0.68650	0.68997	0.00348
0.70743	0.71095	0.00352
0.72783	0.73106	0.003231

1.6 Techniques for Solving First Order Equations

The general form of a first order linear differential equation is

$$a(t)\frac{dx}{dt} + b(t)x + c(t) = 0 \tag{1.6.1}$$

where $a(t)$, $b(t)$ and $c(t)$ are continuous functions in a given interval with $a(t) \neq 0$. Dividing by $a(t)$, we get the equivalent equation in normal form

$$\frac{dx}{dt} + P(t)x = Q(t) \tag{1.6.2}$$

This equation is solved by multiplying Eq. (1.6.2) by $e^{\int P(t)dt}$ and integrating, to give us the solution

$$x(t) = e^{-\int P(t)dt} \left[\int e^{\int P(t)dt} Q(t)dt + c \right] \tag{1.6.3}$$

Exact Equations

The differential equation

$$M(t,x)dt + N(t,x)dx = 0 \tag{1.6.4}$$

or

$$M(t,x) + N(t,x)\frac{dx}{dt} = 0 = M(t,x)\frac{dt}{dx} + N(t,x)$$

is called exact if there is a differentiable function $u(t,x)$ such that

$$\frac{\partial u}{\partial t} = M, \quad \frac{\partial u}{\partial x} = N \tag{1.6.5}$$

If Eq. (1.6.5) holds, we have

$$\begin{aligned} M(t,x)dt + N(t,x)dx &= \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial x}dx \\ &= d(u(t,x)) \end{aligned}$$

Hence, intergrating the above wquation we get the solution

$$u(t,x) = C$$

of the exact equation given by Eq. (1.6.4).

The following theorem gives a test for the exactness of a differential equation.

Theorem 1.6.1 *If M and N are continuously differentiable function of $(t,x) \in D \subseteq \Re^2$ (with no hole in D), then the differential equation*

$$M(t,x)dt + N(t,x)dx = 0 \ \textit{ is exact iff } \ \frac{\partial M}{\partial x} = \frac{\partial N}{\partial t} \ \textit{ in } \ D \tag{1.6.6}$$

Proof : If the differential equation is exact, we have $\frac{\partial M}{\partial x} = \frac{\partial}{\partial x}\left[\frac{\partial u}{\partial t}\right]$ and $\frac{\partial N}{\partial t} = \frac{\partial}{\partial t}\left[\frac{\partial u}{\partial x}\right]$. Since M and N are continuously differentiable in D, it follows that $\frac{\partial^2 u}{\partial x \partial t} = \frac{\partial^2 u}{\partial t \partial x}$ and hence $\frac{\partial M}{\partial x} = \frac{\partial N}{\partial t}$ in D.

Conversly, let Eq. (1.6.6) holds.

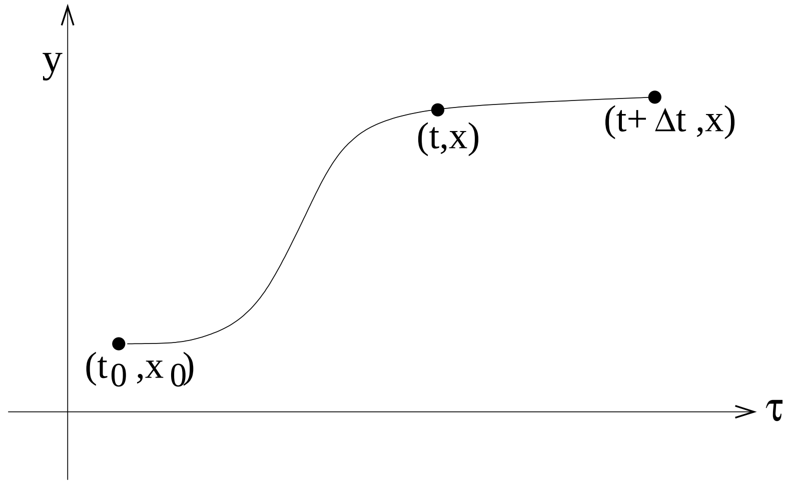

Figure 1.6.1: A curve in $\tau - y$ plane

We explicitily define the function $u(t,x)$ as the solution of the initial value problem

$$\begin{aligned} \frac{\partial u}{\partial t} &= M(t,x), \quad \frac{\partial u}{\partial x} = N(t,x) \\ u(t_0,x_0) &= u_0 \end{aligned}$$

Then, $u(t,x)$ is given by

$$u(t,x) = u_0 + \int_{(t_0,x_0)}^{(t,x)} [M(\tau,y)d\tau + N(\tau,y)dy]$$

where the integral on the RHS is a line integral along a path joining (t_0,x_0) and (t,x). It can be shown that this line integral is independent of path if Eq. (1.6.6) holds.
This gives

$$\begin{aligned} u(t+\triangle t,x) - u(t,x) &= \int_{(t,x)}^{(t+\triangle t,x)} [M(\tau,y)d\tau + N(\tau,y)dy] \\ &= \int_t^{t+\triangle t} M(\tau,y)d\tau \ \ (\text{as } \ y = x \text{ and } dy = 0) \\ &= M(\tau_1,x)\triangle t, \quad t < \tau_1 < t + \triangle t \end{aligned}$$

(by meanvalue theorem for integrals).
Hence

$$\frac{\partial u}{\partial t} = \lim_{\triangle t \to 0} \left[\frac{u(t+\triangle t,x) - u(t,x)}{\triangle t} \right] = \lim_{\triangle t \to 0} [M(\tau_1,x)] = M(t,x)$$

Similarly, we have

$$\frac{\partial u}{\partial x} = N(t,x)$$

∎

It may happen that the equation $M(t,x)dt + N(t,x)dx = 0$ is not exact but after multiplying both sides of this equation by a function $\mu(t,x)$, it becomes exact. Such a function is called an integrating factor. For example, $xdy - ydx = (x^2+y^2)dx$ is not exact. But dividing by (x^2+y^2), we get

$$\frac{xdy - ydx}{x^2+y^2} = dx$$

which is equivalent to

$$\frac{\frac{x}{x^2}dy - \frac{y}{x^2}dx}{1 + \frac{y^2}{x^2}} = dx$$

This is now an exact differential equation with solution

$$tan^{-1}(\frac{y}{x}) = x + c$$

Change of Variables

There are a few common types of equations, wherein substitutions suggest themselves, as we see below

[A]

$$\frac{dx}{dt} = f(at+bx), \quad a,b \in \Re \tag{1.6.7}$$

We introduce the change of variable $X = at + bx$, which gives

$$x = \frac{1}{b}(X - at), \quad \frac{dx}{dt} = \frac{1}{b}(\frac{dX}{dt} - a)$$

and hence Eq. (1.6.7) becomes

$$\frac{1}{b}(\frac{dX}{dt} - a) = f(X)$$
$$\text{or} \quad \frac{dX}{dt} = a + bf(X)$$

which can be easily solved.

[B]

$$\frac{dx}{dt} = f(\frac{x}{t}) \tag{1.6.8}$$

in which RHS depends only on the ratio $\frac{x}{t}$.

Introduce the change of variable $u = \frac{x}{t}$. Thus

$$x = ut, \qquad \frac{dx}{dt} = u + t\frac{du}{dt}$$

Then Eq. (1.6.8) becomes

$$u + t\frac{du}{dt} = f(u)$$

or

$$\frac{du}{dt} = \frac{f(u) - u}{t}$$

which can be easily solved.

[C] Consider the equation

$$\frac{dx}{dt} = f\left(\frac{at + bx + p}{ct + dx + q}\right), \quad a, b, c, d, p, q, \in \Re \tag{1.6.9}$$

in which RHS depends only on the ratio of the linear expression.

We substitute $T = t - h,\ X = x - k$.

Then $\frac{dx}{dt} = \frac{dX}{dt}$ where we choose h and k such that

$$ah + bk + p = 0 = ch + dk + q \tag{1.6.10}$$

Then $\frac{dX}{dT} = f\left(\frac{at + bX}{ct + dX}\right)$ which is of type **[B]**.

Eq. (1.6.10) can always be solved for h, k except when $ad - bc = 0$. In that case we have $cx + dy = m(ax + by)$ and hence Eq. (1.6.9) becomes

$$\frac{dx}{dt} = f\left[\frac{at + bx + p}{m(at + bx) + q}\right]$$

which is of the type **[A]**.

[D] The equation

$$\frac{dx}{dt} + P(t)x = Q(t)x^m \tag{1.6.11}$$

in which the exponent is not necessarily an integer, is called Bernoulli's equation. Assume that $m \neq 1$. Then introduce the change of variable

$X = x^{1-m}$. Then $x = X^{(\frac{1}{1-m})}$, $\dfrac{dx}{dt} = \dfrac{1}{1-m} X^{(\frac{m}{1-m})} \dfrac{dX}{dt}$. Hence Eq. (1.6.11) becomes

$$\frac{1}{1-m} X^{(\frac{m}{1-m})} \frac{dX}{dt} + P(t) X^{(\frac{1}{1-m})} = Q(t) X^{(\frac{m}{1-m})}$$

Equivalently

$$\frac{dX}{dt} + (1-m)P(t)X = (1-m)Q(t)$$

which is a linear differential equation in X.

Example 1.6.1 *Solve*

$$(t + x^2)dx + (x - t^2)dt = 0$$
$$M(t,x) = x - t^2, \quad N(t,x) = t + x^2$$
$$\frac{\partial M}{\partial x} = 1 = \frac{\partial N}{\partial t}$$

and hence this equation is exact.
So we write the solution $u(t,x)$ as

$$u(t,x) = u_0 + \int_{(t_0,x_0)}^{(t,x)} \left[(\tau - y^2)d\tau + (\tau + y^2)dy \right]$$

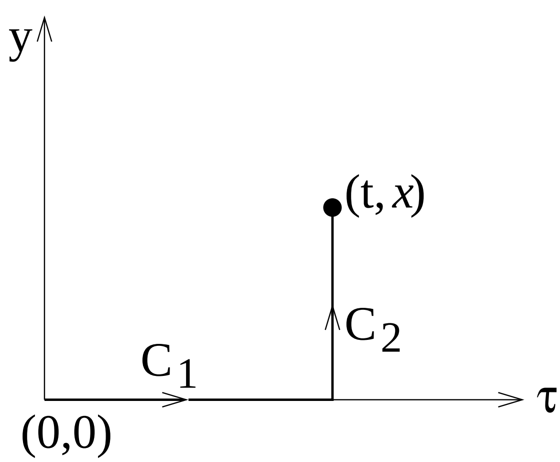

Figure 1.6.2: Path from $(0,0)$ to (t,x)

In the integrand on RHS, we go along the path C_1, C_2. Hence the integral become

$$\int_0^t (-\tau^2)d\tau + \int_0^x (t + y^2)dy = -\frac{t^3}{3} + tx + \frac{x^3}{3}$$

Hence, the solution is

$$\frac{-t^2}{3} + tx + \frac{x^3}{3} = C$$

Example 1.6.2 *Solve* $\dfrac{dx}{dt} = \dfrac{t+x-1}{t+4x+2}$

Solve $h+k-1 = 0 = h+4k+2$. *This gives* $h = 2, k = -1$. *Hence* $t = T+2, x = X-1$ *and* $\dfrac{dX}{dT} = \dfrac{T+X}{T+4X}$.

We now make substitution $X = TU$, *to get*

$$T\frac{dU}{dT} = \frac{1-4U^2}{1+4U}$$

which is equivalent to

$$\frac{1+4U}{1-4U^2}dU = \frac{dT}{T}$$

This yields

$$\begin{aligned} (1+2U)(1-2U)^3T^4 &= C \\ \text{or} \quad (T+2X)(T-2X)^3 &= C \\ \text{or} \quad (t+2x)(t-2x-4)^3 &= C \end{aligned}$$

Example 1.6.3 *Solve* $\dfrac{dx}{dt} + tx = t^3x^4$

This is a Bernoulli's equation. We make the change of variable $X = x^{-3}$ *and get an equivalent linear differential equation*

$$\frac{dx}{dt} - 3tX = -3t^3$$

The general solution of this equation is given by

$$\begin{aligned} X(t) &= e^{3\int tdt}\left[\int (3t^3)e^{-\int 3tdt} + C\right] \\ &= e^{\frac{3}{2}t^2}\left[\int (-3t^3)e^{-\frac{3}{2}t^2}dt + c\right] \end{aligned}$$

To compute $I = \int (-3t^3) e^{-\frac{3}{2}t^2} dt + c$, *put* $t^2 = u,\ 2tdt = du$. *This gives*

$$\begin{aligned} I &= \int e^{-\frac{3}{2}u}\left[-\frac{3}{2}udu\right] \\ &= -\frac{3}{2}\left[-\frac{2}{3}e^{-\frac{3}{2}u}u + \frac{2}{3}\int e^{-\frac{3}{2}u}du\right] \\ &= -\frac{3}{2}\left[-\frac{2}{3}e^{-\frac{3}{2}u}u - \frac{2}{3}\frac{2}{3}e^{-\frac{3}{2}u}\right] \\ &= e^{-\frac{3}{2}u}u + \frac{2}{3}e^{-\frac{3}{2}u} \\ &= e^{-\frac{3}{2}t^2}t^2 + \frac{2}{3}e^{-\frac{3}{2}t^2} \end{aligned}$$

and hence

$$\begin{aligned} X(t) &= e^{-\frac{3t^2}{2}}\left[e^{-\frac{3}{2}t^2}t^2 + \frac{2}{3}e^{-\frac{3}{2}t^2+c}\right] \\ &= \left[t^2 + \frac{2}{3} + ce^{\frac{3}{2}t^2}\right] \end{aligned}$$

Equivalently

$$x = \frac{1}{X^3} = \frac{1}{\left[t^2 + \frac{2}{3} + ce^{\frac{3}{2}t^2}\right]^3}$$

Example 1.6.4 *Solve the IVP*

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

We make the change of variable $X = \dfrac{1}{x}$ *and get an equivalent linear equation*

$$\frac{dX}{dt} + X = 1,\ \ X(0) = 2$$

The unique solution $X(t)$ *is given by*

$$\begin{aligned} X(t) &= e^{-t}\left[\int e^t dt + t\right] \\ &= e^{-t} + 1 \end{aligned}$$

This gives

$$x(t) = \frac{1}{1 + e^{-t}}$$

For more on techniques for solving first order differential equation refer Golomb and Shanks[4].

For more on real life problems giving rise to mathematical models generated by ordinary differential equations, refer to Braun et al[1], Burghe and Borrie[2] and Haberman[5].

1.7 Exercises

1. Sketch the direction field for the following differential equations.

   (a) $\left(\frac{dx}{dt}\right)^2 = x^2 + t^2$

   (b) $t\frac{dx}{dt} - x = \sqrt{\left(\frac{dx}{dt}\right)^2 - 1}$

   (c) $x^2 - (1 - t^2)\left(\frac{dx}{dt}\right)^2 = 0$

   Sketch the solution $x(t)$, wherever possible.

2. Consider the differential equation

$$\begin{aligned}\frac{dx}{dt} &= f(t, x(t)) \\ f(t, x) &= \begin{cases} \frac{2tx}{t^2 + x^2}, & (t, x) \neq (0, 0) \\ 0, & (t, x) = (0, 0) \end{cases}\end{aligned}$$

   (a) Sketch the direction field of the above equation.

   (b) Solve the above equation.

   (c) Find the number of solutions passing through $(0, 0)$ and $(x_0, y_0) \neq (0, 0)$.

3. Consider the differential equation

$$\frac{dx}{dt} = (x^2 - 1)t^p, \quad p \geq 0, \ t \geq 0$$

   (a) Determine the equilibrium points.

   (b) Sketch the vector field.

   (c) Find solution passing through $(0, -2)$, $(0, 0)$ and $(0, 2)$.

4. Consider the linear system in $\Re^2$ given by

$$\begin{aligned}\frac{dx_1}{dt} &= (x_2 + 1)\cos at \\ \frac{dx_2}{dt} &= x_1 + x_2\end{aligned}$$

   (a) Find the stationary points.

   (b) Sketch the vector field in $\Re^2$.

5. Solve following initial value problems.

(a) $t^2\dfrac{dx}{dt} + x(t+1) = t, \quad x(-1) = -2.$

(b) $(1+t^2)\left(\tan^{-1} t\right)\dfrac{dx}{dt} + x = t, \;\; x(-1) = -1$

(c) $2(1-t^2)\dfrac{dx}{dt} - 2tx = a(t^3 - t), \;\; x(0) = \dfrac{a}{3}$

6. Test if the following equations are exact. If so, solve them.

   (a) $\sin t \sin^2 x dt - (\cos t \cos 2x \tan x + \cos t \tan x)\, dx = 0.$

   (b) $\left[x log(t^2+x^2) + \dfrac{2t^2x}{t^2+x^2} + x\right] dx +$
   $\left[t + log(t^2+x^2) + \dfrac{2t^2x}{t^2+x^2} + t\right] dx = 0$

   (c) $\left[\dfrac{(t+x)^2}{1+(t+x)^2}\right] dt - \left[\dfrac{1}{1+(t+x)^2} - 1\right] dx = 0$

7. Using the line integrals, solve the following exact equations which pass through the given point.

   (a) $[2t + exp(t)\sin x]\, dt + exp(t)\cos x dx = 0, \;\; x(0) = \dfrac{\pi}{2}$

   (b) $(\cos t \cos x)\, dx - (\sin t \sin x)\, dx = 0, \;\; x(\dfrac{\pi}{2}) = \dfrac{\pi}{3}$

   (c) $\left[\dfrac{(t+x)}{t^2+x^2}\right] dt - \left[\dfrac{t-x}{t^2+x^2}\right] dx = 0, \;\; x(1) = 1$

8. Solve the following Bernoulli equations.

   (a) $\dfrac{dx}{dt} + x = tx^{3/2}$

   (b) $2tx^3\dfrac{dx}{dt} + t^4 - x^4 = 0$

   (c) $3\dfrac{dx}{dt} + t\sec x = x^4(1 - \sin t)^2$

9. Find all curves with the property that the angle of inclination of its tangent line at any point is three times the inclination of the radius vector to that point from the orgin.

10. Find the fimily of curves orthogonal to the family of cardiods $r = \lambda(1 + \cos\theta)$ in polar cordinates (r, θ).

11. In a model of epidemics, a single infected individual is introduced in to a community containing n individuals susceptible to the desease. Let $x(t)$ denote the number of uninfected individuals in the population at time t. If we assume that the infection spreads to all those susceptible, then $x(t)$ decreases from $x(0) = n$ to $x(0) = 0$. Give a mathematical model for the above problem and solve it.

References

[1] Martin Braun, Courtney S Coleman and Donald Drew, *Differential Equations Model, Ellis Horwood Ltd., 1981.*

[2] D. N. Burghe and M S Borrie, *Modelling with Differential Equation, Springer Verlag, 1983.*

[3] S. D. Conte and C deBoor, *Elementary Numerical Analysis- An Algorithmic Approach, McGraw Hill Book Co., 1981.*

[4] Michael Golomb and Merrill Shanks, *Elements of Ordinary Differneial Equations, McGraw Hill Book Company, 1965.*

[5] Richard Haberman, *Mathematical Models: Mechanical Vibrations, Population Dynamics and Traffic Flow, Prentice Hall, 1977.*

[6] S. Haykin, *Neural Networks, Macmillan College Publishing Company, Boston, 1994.*

[7] Mohan C Joshi and Kannan S Moudgalya, *Optimization - Theory and Practice, Narosa Publishing House Pvt. Ltd, 2004.*

[8] I. E. Logaris, A. Likas and D. I. Fotiadis, *Artificial Neural Networks for Solving Ordinary and Partial Differential Equations, IEEE Trans. Neural Networks, 9(5), 987-1000, 1998.*

[9] R.M.M Mattheij and J Molenaar, *Ordinary Differential Equations in Theory and Practice, John Wiley and Sons, 1996.*

Chapter 2

Mathematical Preliminaries

This chapter is meant for laying the foundation for abstract concepts needed for our subsequent discussion of ODE in function spaces.
We start with vector spaces and go on to define normed and inner product spaces and then dwell on notions like Schwartz inequality, Bessel's inequality, Parseval's equality, orthormal system and Fourier series.
Bounded linear and nonlinear operators in function spaces along with their much needed properties - Lipscntiz continuity, monotonicity, differentiability and compactness - are given a fair treatment.
We close by defining Dirac-delta function as a distribution function. This approach is essential for introducing Green's function in Chapter 7.

2.1 Finite and Infinite Dimensional Spaces

Definition 2.1.1 *Let X be a nonempty set which is closed with respect to binary operation $\overline{x}+\overline{y}$ $(\overline{x},\overline{y}\in X)$ and scalar multiplication $\alpha\overline{x}$ $(\alpha\in\Re,\overline{x}\in X)$. X is said to be a vector space if the following holds.*

1. *For any $\overline{x},\overline{y},\overline{z}\in X$*
   (a) $(\overline{x}+\overline{y})+\overline{z}=\overline{x}+(\overline{y}+\overline{z})$
   (b) $(\overline{x}+\overline{y})=(\overline{y}+\overline{x})$
   (c) $\exists$ *an element* $\overline{0}\in X$ *such that* $\overline{x}+\overline{0}=\overline{x}$
   (d) $\exists -\overline{x}\in X$ *such that* $\overline{x}+(-\overline{x})=\overline{0}$
2. *For any $\alpha,\beta\in\Re$ and $\overline{x},\overline{y}\in X$*
   (a) $\alpha(\overline{x}+\overline{y})=\alpha\overline{x}+\alpha\overline{y}$
   (b) $(\alpha+\beta)\overline{x}=\alpha\overline{x}+\beta\overline{x}$
   (c) $\alpha(\beta\overline{x})=(\alpha\beta)\overline{x}$
   (d) $1\overline{x}=\overline{x}$

Definition 2.1.2 *An inner product* $(\overline{x}, \overline{y})$ *in a vector space* X *is a function on* $X \times X$ *with values in* $\Re$ *such that the following holds.*

1. $(\overline{x}, \overline{x}) \geq 0$ *forall* $\overline{x} \in X$ *and equality holds iff* $\overline{x} = \bar{0}$
2. $(\overline{x}, \overline{y}) = (\overline{y}, \overline{x})$ *for all* $\overline{x}, \overline{y} \in X$
3. $(\alpha\overline{x} + \beta\overline{y}, \overline{z}) = \alpha(\overline{x}, \overline{z}) + \beta(\overline{y}, \overline{z})$; $\alpha, \beta \in \Re$, $\overline{x}, \overline{y}, \overline{z} \in X$

The vector space X *with an inner product defined is called inner product space.*

In an inner product space X, $\overline{x}$ *is said to be orthogonal to* $\overline{y}$ *if* $(\overline{x}, \overline{y}) = 0$. *This is denoted by* $\overline{x} \perp \overline{y}$.

Definition 2.1.3 *A vector space* X *is said to be a normed space if there exists a function* $\|\overline{x}\|$ *from* X *to* $\Re$ *such that the following properties hold.*

1. $\|\overline{x}\| \geq 0$ *for all* $\overline{x} \in X$ *and* $\|\overline{x}\| = 0$ *iff* $\overline{x} = 0$
2. $\|\alpha\overline{x}\| = |\alpha|\|\overline{x}\|$ *for all* $\overline{x} \in X$ *and* $\alpha \in \Re$
3. $\|\overline{x} + \overline{y}\| \leq \|\overline{x}\| + \|\overline{y}\|$ *for all* $\overline{x}, \overline{y} \in X$

In an inner product space, the induced norm is defined by

$$\|\overline{x}\|^2 = (\overline{x}, \overline{x}), \quad \overline{x} \in X$$

In a normed space X, *the induced metric* $d(\overline{x}, \overline{y})$ *(distance between two vectors) is defined as*

$$d(\overline{x}, \overline{y}) = \|\overline{x} - \overline{y}\| \;\; \forall \;\; \overline{x}, \overline{y} \in X$$

In view of Definition 2.1.3, this metric $d(\overline{x}, \overline{y})$ *satisfies the following properties.*

[P_1] $d(\overline{x}, \overline{y}) \geq 0$ *for all* $\overline{x}, \overline{y} \in X$ *and* $d(\overline{x}, \overline{y}) = 0$ *iff* $\overline{x} = \overline{y}$

[P_2] $d(\overline{x}, \overline{y}) = d(\overline{y}, \overline{x})$ *for all* $\overline{x}, \overline{y} \in X$

[P_3] $d(\overline{x}, \overline{z}) \leq d(\overline{x}, \overline{y}) + d(\overline{y}, \overline{z})$ *for all* $\overline{x}, \overline{y}, \overline{z} \in X$

It is possible to define a metric $d(x, y)$ *satisfying properties* $[P_1] - [P_3]$ *in any set without having a vector space structure. Such a set is called a metric space.*

Definition 2.1.4 *A metric space* X *is a non-empty set with a metrix* $d : X \times X \to \Re$ *defined on it, which satisfies the properties* $[P_1] - [P_3]$.

In a metric space, without the vector space structure, we can easily define the concepts of convergence, Cauchy convergence, completeness etc., as we see below.

Definition 2.1.5 *A sequence* $\{x_k\}$ *in a metrix space* (X, d) *is said to be convergent to* $x \in X$ *if* $d(x_k, x) \to 0$ *as* $k \to \infty$. $\{x_k\}$ *is said to be Cauchy if* $d(x_k, x_\ell) \to$ *as* $k, l \to \infty$.

Definition 2.1.6 *A metrix space (X, d) is said to be complete if everyCauchy sequence in (X, d) converges.*

A complete normed space is called a Banach space, whereas a complete inner product space is called a Hilbert space.

In a normed space X, it is possible to define infinite sum $\sum_{i=1}^{\infty} \overline{x}_i$. We say that $\overline{s} = \sum_{i=1}^{\infty} \overline{x}_i$ iff the sequence of partial sums $\overline{s}_n = \sum_{i=1}^{n} \overline{x}_i \in X$, converges to $\overline{s} \in X$.

Definition 2.1.7 *A set S of vectors $\{\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n\}$ in a vector space X is said to be linearly dependent (l.d.) if $\exists$ scalars $\alpha_i \in \Re$ (not all zero) such that*

$$\alpha_1\overline{x}_1 + \alpha_2\overline{x}_2 + \ldots + \alpha_n\overline{x}_n = \overline{0}.$$

Otherwise, this set is called linearly independent (ℓ.i.).

If the set S consists of infinite elements of the spaces X. Then S is said to be ℓ.i. if every finite subset of S is ℓ.i.

Definition 2.1.8 *A set of vectors $S = \{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n\}$ in a vector space X is said to be a basis for X if*

1. *S is ℓ.i. and*
2. *$L[S] = \{\alpha_1\overline{x}_1 + \alpha_2\overline{x}_2 + \cdots + \alpha_n\overline{x}_n; \alpha_i \in \Re\}$, called the linear span of S, equals X.*

The number of elements in a basis is unique (refer Bapat[1]).

Definition 2.1.9 *A vector space X is said to be finite dimensional if it has a basis consisting of finite number of elements and this number is called its dimension and is denoted by* $\dim X$.

A vector space X which is not finite dimensional, is called infinite dimensional space.

Example 2.1.1 *In the space $\Re^n$ of n-tuples, the Euclidean norm and inner product are defined as*

$$\|\overline{x}\|^2 = x_1^2 + x_2^2 + \ldots + x_n^2, \ \ x = (x_1, x_2, \ldots, x_n) \qquad (2.1.1(a))$$

$$(\overline{x}, \overline{y}) = x_1y_1 + x_2y_2 + \ldots x_ny_n, \ \ x = (x_1, x_2, \ldots, x_n), \ y = (y_1, y_2, \ldots, y_n) \qquad (2.1.1(b))$$

In terms of the matrix notation we can represent $(\overline{x}, \overline{y})$ as

$$(\overline{x}, \overline{y}) = \overline{y}^\top \overline{x} = \overline{x}^\top \overline{y}$$

if we treat $\overline{x}, \overline{y}$ as column vectors and $\overline{x}^\top, \overline{y}^\top$ as row vectors.

One can show that $\Re^n$ is complete with respect to the norm induced by the inner product defined by Eq. (2.1.1(b)) and hence it is a Hilbert space.

Remark 2.1.1 *It is possible to define other norms in the space $\Re^n$, as we see below*

$$\|\overline{x}\|_1 = |x_1| + |x_2| + \ldots + |x_n|$$

$$\|\overline{x}\|_\infty = \max_{1\leq i\leq n}\{|x_i|\}$$

Definition 2.1.10 *Two different norms $\|\overline{x}\|_a$ and $\|\overline{x}\|_b$ in a normed space X are said to be equivalent if $\exists$ positive constants α, β such that*

$$\alpha\|\overline{x}\|_a \leq \|\overline{x}\|_b \leq \beta\|\overline{x}\|_a$$

One can show that all norms in $\Re^n$ are equivalent (refer Limaye[8]) and hence $\Re^n$ is also complete w.r.t. $\|\overline{x}\|_1$ and $\|\overline{x}\|_\infty$, defined earlier.

Example 2.1.2 *In the space $C[0,1]$ of all real valued continuous functions, one can define norm and inner product as under*

$$\|f\|_2^2 = \int_0^1 f^2(t)dt \tag{2.1.2(a)}$$

$$(f,g) = \int_0^1 f(t)g(t)dt \tag{2.1.2(b)}$$

One can show that the space $C[0,1]$ w.r.t. the above norm is not complete and hence it is not a Banach space. To see this consider the sequence $\{f_n(t)\}$ of continuous functions on $[0,1]$ defined as

$$f_n(t) = t^n, t \in [0,1]$$

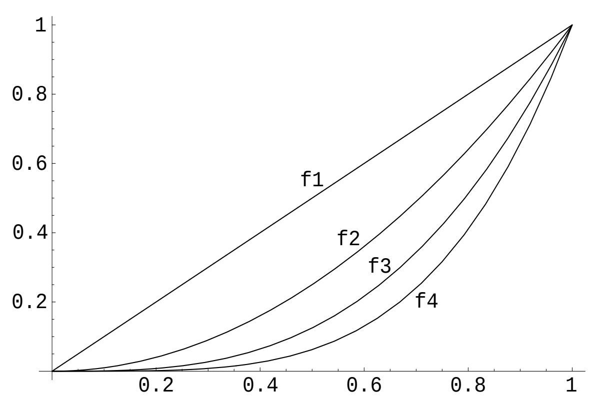

Figure 2.1.1: A sequence which is Cauchy but not complete

$$\|f_n - f_m\|_2^2 = \int_0^1 (t^n - t^m)^2 dt \to 0 \text{ as } n, m \to \infty$$

and hence $\{f_n\}$ *is Cauchy but* $f_n(t)$ *converges to* $f(t)$ *defined as*

$$f(t) = \begin{cases} 1, & t = 1 \\ 0, & t \neq 1 \end{cases}$$

which is not a continuous function and hence the space is $C[0,1]$ *is not complete w.r.t. the above norm.*

However, if $\|f\|_\infty$ *is defined as* $\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$, *then* $C[0,1]$ *is complete with respect to* $\|f\|_\infty$ *norm and hence it is a Banach space.*

Example 2.1.3 *Let* $L_2[a,b]$ *denote the space of all square integrable functions.* $L_2[a,b] = \{f : [a,b] \to \Re, \int_a^b f^2(t)dt < \infty\}$

Let the norm and inner product be defined as in Eq. (2.1.2). Then $L_2[a,b]$ *is a Hilbert space.*

Remark 2.1.2 $\Re^n$ *is finite dimensional, where as* $C[a,b)$ *and* $L_2[a,b]$ *are infinite dimensional spaces*

In an inner product space, we have the following identities.

1. *Pythagorus identity:*

$$\overline{x} \perp \overline{y} \text{ iff } \|\overline{x}+\overline{y}\|^2 = \|\overline{x}\|^2 + \|\overline{y}\|^2 \tag{2.1.3}$$

2. *Parallelogram law:*

$$\|\overline{x}+\overline{y}\|^2 + \|\overline{x}-\overline{y}\|^2 = 2[\|\overline{x}\|^2 + \|\overline{y}\|^2] \tag{2.1.4}$$

Parallelogram law is relevant in view of the following theorem.

Theorem 2.1.1 *A normed space* X *is an inner product space iff the norm in* X *satisfies the parallelogram law given by Eq. (2.1.4).*

Refer Goffman and Pedrick [5] for the proof of the above Theorem.

Example 2.1.4 $X = C[0,1]$ *with*

$$\|f\| = \sup_{t\in(0,1]} |f(t)|, f \in X \quad t \in [0,1]$$

Take $f = t, g = 1$. *One can check that parallegram law is not satisfied and hence the Banach space* $C[0,1]$ *with respect to the above norm can never be made a Hilbert space.*

Theorem 2.1.2 (Schwartz inequality)
Let X *be an inner product space. Then for all* $\overline{x}, \overline{y}, \in X$, *we have the following inequality*

$$|(\overline{x},\overline{y})| \leq \|\overline{x}\|\|\overline{y}\| \tag{2.1.5}$$

Equality holds iff $\overline{x}$ *and* $\bar{y}$ *are l.d.*

Proof : If any of the elements $\overline{x}, \overline{y}$ is zero, we are through. So, assume that $\overline{x}, \overline{y} \neq \overline{0}$. By normalizing $\overline{y}$ as $\overline{y}/\|\overline{y}\| = \bar{e}$, the above inequality reduces to

$$|(\overline{x}, \overline{e})| \leq \|\overline{x}\| \text{ for all } \overline{x} \in X \tag{2.1.6}$$

So, it suffices to show Eq. (2.1.6). We have

$$\begin{aligned} 0 &\leq (\overline{x} - (\overline{x}, \overline{e})\overline{e}, \overline{x} - (\overline{x}, \overline{e})\overline{e}) \\ &= (\overline{x}, \overline{x}) - (\overline{x}, \overline{e})^2 \end{aligned}$$

This gives Eq. (2.1.5). Also, if equality holds in Eq. (2.1.5), we get

$$(\overline{x} - (\overline{x}, \overline{y})\overline{y}, \overline{x} - (\overline{x}, \overline{y})\overline{y}) = 0$$

This implies that

$$\overline{x} - (\overline{x}, \overline{y})\overline{y} = 0$$

and hence $\overline{x}, \overline{y}$ are l.d.

On the other hand if $\bar{x}$, $\bar{y}$ are l.d., then Eq. (2.1.5) is true. This proves the theorem. ∎

Let M be a closed subspace of a Hilbert space X. Given an element $\overline{x} \in X$, we wish to obtain $\overline{z} \in M$ which is closest to M. We have the following theorem in this direction.

Theorem 2.1.3 *Suppose $\overline{x} \in X$ and M a closed subspace of X. Then $\exists$ a unique element $\overline{z} \in M$ such that*

$$\|\overline{x} - \overline{z}\| = \inf_{\overline{y} \in M} \|\overline{x} - \overline{y}\|$$

Proof : Let $d = \inf_{\overline{y} \in M} \|\overline{x} - \overline{y}\|$. Then $\exists$ a sequence $\{\bar{y}_n\} \in M$ such that

$$\|\overline{x} - \overline{y}_n\| \to d$$

By parallelogram law, we have

$$\begin{aligned} \|\overline{y}_n - \overline{y}_m\| &= \|(\bar{y}_n - \overline{x}) - (\overline{y}_m - \overline{x})\|^2 \\ &= [2\|\overline{y}_n - \overline{x}\|^2 + 2\|\overline{y}_m - \overline{x}\|^2 + \| - 2\overline{x} + \overline{y}_n + \overline{y}_m\|^2] \\ &\leq [2\|\overline{y}_n - \overline{x}\|^2 + 2\|\overline{y}_m - \overline{x}\|^2 - 4d^2] \\ &\to \ 2d^2 + 2d^2 - 4d^2 = 0 \end{aligned}$$

That is, $\{\overline{y}_n\}$ is Cauchy and since M is a closed subspace of X, it is complete and hence $\overline{y}_n \to \overline{z} \in M$. It is clear that

$$\|\overline{x} - \overline{z}\| = \lim_{n \to \infty} \|\overline{x} - \overline{y}_n\| = d$$

Uniqueness of $\overline{z}$ is easy to prove. ∎

This gives us the following theorem. In this theorem $M^{\perp}$ represent the spaces of all elements orthogonal to M.

Theorem 2.1.4 *Let M be a closed subspace of X. Then $\overline{z} \in M$ is the point of minimum distance from $\overline{x} \in X$ iff $\overline{x} - \overline{z} \in M^{\perp}$.*

This, in turn, implies that every $x \in X$ has a unique representation

$$\overline{x} = \overline{z} + \overline{w}, \quad \overline{z} \in M, \overline{w} \in M^{\perp}$$

This representation is written as

$$X = M \oplus M^{\perp}$$

X is called the direct sum of M and $M^{\perp}$.

Definition 2.1.11 *A set $S \subset X$ is called an orthonormal set if $\overline{e}_{\alpha}, \overline{e}_{\beta} \in S \Rightarrow (\overline{e}_{\alpha}, \overline{e}_{\beta}) = 0, \alpha \neq \beta$ and $\|\overline{e}_{\alpha}\| = 1 \; \forall \, \overline{e}_{\alpha} \in S$.*

We have the following property for orthonormal sets in X.

Theorem 2.1.5 *Let $\{\overline{e}_{\alpha}\}_{\alpha \in A}$ be a collection of orthonormal elements in X. Then, for all $\overline{x} \in X$, we have*

$$\sum_{\alpha \in A} |(\overline{x}, \overline{e}_{\alpha})|^2 \leq \|\overline{x}\|^2 \tag{2.1.7}$$

and further $\left(\bar{x} - \sum_{\alpha \in A} (\overline{x}, \overline{e}_{\alpha}) \overline{e}_{\alpha}\right) \perp \overline{e}_{\beta} \;\; \forall \;\; \beta \in A$.

The inequality Eq. (2.1.7) is referred as Bessel's inequality.

Definition 2.1.12 *Let S be an orthonormal set in a Hilbert space X. Then S is called a basis for X (or a complete orthonormal system) if no other orthonormal set contains S as a proper set.*

The following theorem presents the most important property of a complete orthonormal set.

Theorem 2.1.6 *Let X be a Hilbert space and let $S = \{\overline{e}_{\alpha}\}_{\alpha \in A}$ be an orthonormal set in X. Then the following are equivalent in X.*

(1) $S = \{\overline{e}_{\alpha}\}_{\alpha \in A}$ is complete.

(2) $\overline{0}$ is only vector which is orthogonal to every $\overline{e}_{\alpha} \in S$. That is, $\overline{x} \perp \overline{e}_{\alpha}$ for every $\alpha \Rightarrow \overline{x} = \overline{0}$.

(3) Every vector $\overline{x} \in X$ has Fourier series expansion $\overline{x} = \sum\limits_{\alpha \in A} (\overline{x}, \overline{e}_{\alpha}) \overline{e}_{\alpha}$.

(4) Every vector $\overline{x} \in X$ satisfies the Parseval's equality

$$\|\overline{x}\|^2 = \sum_{\alpha \in A} \|(\overline{x}, \overline{e}_{\alpha})|^2 \tag{2.1.8}$$

Proof : (1) $\Rightarrow$ (2)
Suppose (2) is not true, then $\exists\ \overline{x} \neq 0$ such that $\overline{x} \perp \overline{e}_\alpha \ \ \forall \alpha \in A$. Defining $\overline{e} = \overline{x}/\|\overline{x}\|$, we get an orthonormal set $S \cup \{\overline{e}\}$ which properly contains S. This contradicts the completeness of S. (2) $\Rightarrow$ (3)
$\overline{x} - \sum\limits_{\alpha \in A} (\overline{x}, \overline{e}_\alpha)\overline{e}_\alpha$ is orthogonal to $\overline{e}_\beta$ for every β. But by (2), it follows that it must be the zero vector $\overline{0}$ and hence

$$\overline{x} = \sum_{\alpha \in A} (\overline{x}, \overline{e}_\alpha)\overline{e}_\alpha$$

(3) $\Rightarrow$ (4)
We have

$$\begin{aligned}
\|\overline{x}\|^2 = (\overline{x}, \overline{x}) &= (\sum_{\alpha \in A} (\overline{x}, \overline{e}_\alpha)\overline{e}_\alpha, \sum_{\beta \in A} (\overline{x}, \overline{e}_\beta)\overline{e}_\beta) \\
&= \sum_{\alpha \in A} |(\overline{x}, \overline{e}_\alpha)|^2
\end{aligned}$$

(4) $\Rightarrow$ (1)
If $S = \{\overline{e}_\alpha\}_{\alpha \in A}$ is not complete, it is properly contained in an orthonormal set $\hat{S}$ and let $\overline{e} \in \hat{S} - S$. then Parseval's equation Eq. (2.1.8) gives

$$\|\overline{e}\|^2 = \sum_{\alpha \in A} |(\overline{e}, \overline{e}_\alpha)|^2 = 0$$

contradicting the fact that $\overline{e}$ is a unit vector. ∎

Example 2.1.5 *In $X = L_2[-\pi, \pi]$, the collection of functions*

$$S = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos t}{\sqrt{\pi}}, \frac{\cos 2t}{\sqrt{\pi}}, \cdots \frac{\sin t}{\sqrt{\pi}}, \frac{\sin 2t}{\sqrt{\pi}}, \cdots \right\}$$

is a complete orthonormal set and every $f \in L_2[-\pi, \pi]$ has a Fourier series expansion

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos kt + b_k \sin kt] \qquad (2.1.9)$$

where

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos kt dt \quad \textit{and} \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin kt dt, \ \ k = 0, 1, 2, \ldots$$

The fact that the above set S forms an orthonormal basis in $L_2[-\pi, \pi]$ will be proved in Chapter 7. We also note that for any $f \in L_2[-\pi, \pi]$ we have the Parseval's relation

$$\int_{\pi}^{\pi} f^2(t) dt = \frac{(f, 1)^2}{2\pi} + \sum_{n=1}^{\infty} \frac{1}{\pi} [(f, \cos nt)^2 + (f, \sin nt)^2]$$

The Fourier series representation given by Eq. (2.1.9) will be used while discussing the solution of boundary value problems.

For a piecewise continuous function $f(t)$ defined on $[-\pi, \pi]$,

$$f(t) = \begin{cases} 1, & t \geq 0 \\ -1, & t \leq 0 \end{cases}$$

we have the following Fourier series representation along with its graphics (Figure 2.1.2)

$$f(t) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin(2n+1)t}{2n+1}$$

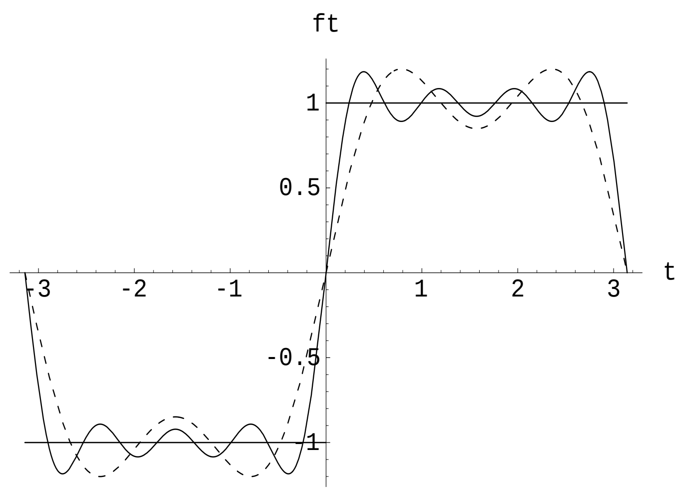

Figure 2.1.2: Sketch of first few terms of the Fourier series

We now give the following Gram-Schmidt procedure which helps in producing an orthonormal collection $\{\overline{e}_1, \overline{e}_1, \ldots, \overline{e}_n\}$ out of the $\ell.i.$ collection $\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n\}$. We first obtain orthogonal vectors $\{\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n\}$, inductively

$$\begin{aligned} \bar{y}_1 &= \bar{x}_1 \\ &\vdots \\ \bar{y}_i &= \bar{x}_i - \alpha_{i,1}\, \bar{y}_1 - \alpha_{i,2}\, \bar{y}_2, \ldots, \alpha_{i,i-1}\, \bar{y}_{i-1} \end{aligned}$$

where

$$\alpha_{i,j} = \frac{(\bar{x}_i, \bar{y}_j)}{(\bar{y}_j, \bar{y}_j)}, \quad i \leq j \leq i-1$$

This is continued inductively for $i+1, i+2, , \ldots, n$. It is clear that

$$L[\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n] = L[\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n]$$

for each n. Normalizing $\bar{y}_i$, we get the orthonormal collection $\{\bar{e}, \bar{e}_2, \ldots, \bar{e}_n\}$, $\bar{e}_i = \bar{y}_i/\|\bar{y}_i\|$.

Example 2.1.6 *In the space $L_2[-1,1]$, the set of all polynomials $\{P_0(t), P_1(t), \ldots\}$, called Legendre polynomials, which are obtained by Gram-Schmidt orthonormalisation procedure, is an orthonormal basis for $L_2[-1,1]$. Computation for p_0, p_1, p_2, p_3, p_4 is given below. The higher degree Legendre polynomials are inductively obtained.*

$$
\begin{aligned}
p_0(t) &= 1, \;\; p_1(t) = t \\
p_2(x) &= t^2 - \alpha_0 p_0(t) - \alpha_1 p_1(t), \alpha_0 = \frac{(t^2,1)}{(1,1)} = \frac{1}{3}, \\
&\quad \alpha_1 = \frac{(t^3,t)}{(t,t)} = 0
\end{aligned}
$$

and hence

$$
\begin{aligned}
p_2(t) &= t^2 - \frac{1}{3} \\
p_3(t) &= t^3 - \alpha_0 p_0(t) - \alpha_1 p_1(t) - \alpha_2 p_2(t), \alpha_0 = \frac{(t^3,1)}{(1,1)} = 0, \\
&\quad \alpha_1 = \frac{(t^3,t)}{(t,t)} = \frac{3}{5}, \alpha_2 = 0
\end{aligned}
$$

and hence

$$
\begin{aligned}
p_3(t) &= t^3 - \frac{3}{5}t \\
p_4(t) &= t^4 - \alpha_0 p_0(t) - \alpha_1 p_1(t) - \alpha_2 p_2(t) - \alpha_3 p_3(t), \\
&\quad \alpha_0 = \frac{(t^4,1)}{(1,1)} = \frac{1}{5}, \;\; \alpha_1 = 0, \\
&\quad \alpha_2 = \frac{(t^4, p_2(t))}{(p_2(t), p_2(t))} = \frac{6}{7}, \;\; \alpha_3 = 0
\end{aligned}
$$

and hence

$$
p_4(t) = 0 = t^4 - \frac{6}{7}t^2 - \frac{1}{5}
$$

Normalising these polynomials we get the Legendre polynomials $P_i(t)$, $1 \le i \le 4$, given by $P_0(t) = 1$, $P_1(t) = t$, $P_2(t) = \frac{1}{2}(3t^2 - 1)$, $P_3(t) = \frac{1}{2}(5t^3 - 3t)$, $P_4(t) = \frac{1}{8}(35t^4 - 30t^2 + 3)$.

For $f \in L_2[-1,1]$, we have the Legendre series representation

$$f(t) = \sum_{i=0}^{\infty} a_i P_i(t), a_i = (f, P_i) = \int_{-1}^{1} f(t) P_i(t) dt, \quad i \geq 0$$

For $f(t) = \sin \pi t$ we have the following graphical representation of first few terms of the Legendre series.

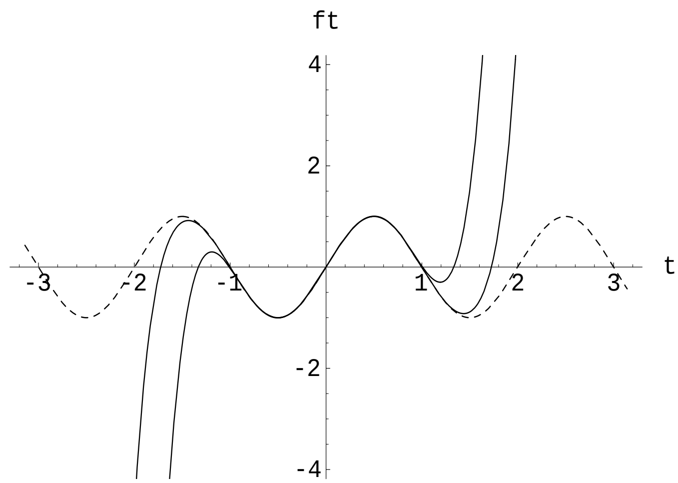

Figure 2.1.3: Sketch of First few of terms of Legendre series for $\sin \pi t$

Legendre polynomials will be used while discussing series solution of Legendre differential equation, to be discussed in Chapter 6.

2.2 Linear and Nonlinear Operators

In this section, the spaces under consideration will be normed spaces, unless otherwise stated.

Definition 2.2.1 *$T : X \to Y$ is said to be linear if*

$$T(\alpha_1 \overline{x}_1 + \alpha_2 \overline{x}_2) = \alpha_1 T\overline{x}_1 + \alpha_2 T\overline{x}_2, \forall \alpha_1, \alpha_2 \in \Re \ \forall \, \overline{x}_1, \overline{x}_2 \in X$$

Otherwise, T is called nonlinear.

Definition 2.2.2 *$T : X \to Y$ is said to be continuous at $\overline{x} \in X$ if*

$$\overline{x}_n \to \overline{x} \text{ in } X \Rightarrow T\overline{x}_n \to T\overline{x} \text{ in } Y$$

T is said to be continuous on X if T is continuous at every $\overline{x} \in X$.

One can show that a linear operator T is continuous on X iff $\exists$ a constant $k > 0$ such that

$$\|T\overline{x}\| \leq k\|\overline{x}\| \ \ \forall \ \ \overline{x} \in X \tag{2.2.1}$$

A linear T satisfying Eq. (2.2.1) is said to be a bounded operator. $\mathcal{B}(X, Y)$ will denote the space of all bounded linear operators from X to Y. One can show that $\mathcal{B}(X, Y)$ is also normed space with $\|T\|$ defined by

$$\|T\| = \sup\{\frac{\|T\overline{x}\|}{\|\overline{x}\|}, \overline{x} \neq \bar{0}\} \tag{2.2.2}$$

$\mathcal{B}(X)$ will denote the space of bounded linear operator from X into itself.

Example 2.2.1 *Let $X = \Re^n$ and $T : \Re^n \to \Re^n$ be defined by a matrix $(\alpha_{ij})_{n\times n}$*

$$[T\overline{x}]_i = \sum_{j=1}^{n} \alpha_{ij}x_j, \overline{x} = (x_1, \ldots, x_n) \in \Re^n$$

T is obviously linear. Also

$$\|T\overline{x}\|_1 \leq \left(\max_i \sum_{j=1}^{n} |\alpha_{ij}| \right) \|\overline{x}\|_1$$

Thus T is a bounded linear operator on $\Re^n$, equipped with 1-norm. But on $\Re^n$, all norms are equivalent and hence T is also a bounded linear operator on $\Re^n$, equipped with the Euclidean norm $\|\overline{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$.

Example 2.2.2 *For a discussion on boundary value problems, we shall be encountering a linear operator of the form*

$$[Tf](t) = \int_0^1 k(t,s)f(s)ds, f \in L_2[0,1]$$

Assume that $k : [0,1] \times [0,1] \to \Re$ is square integrable : $\int_0^1 \int_0^1 k^2(t,s)dtds < \infty$. We have

$$\begin{aligned} |Tf(t)|^2 &\leq \int_0^1 |k(t,s)f(s)|^2 ds \\ &\leq \int_0^1 |k(t,s)|^2 ds \int_0^1 |f^2(s)|ds \end{aligned}$$

(by the application of Schwartz inequality to $f(s)$ and $k(.,s) \in L_2[0,1]$). This inequality implies that

$$\int_0^1 |Tf(t)|^2 dt \leq (\int_0^1 |k(t,s)|^2 dtds)(\int_0^1 |f^2(s)|ds)$$

Denote by $k^2 = \int_0^1 \int_0^1 k(t,s)|^2 dtds$. *Then we get*

$$\|Tf\| \leq k\|f\| \;\; \forall\, f \in L_2[0,1]$$

That is, $T \in \mathcal{B}(L_2[0,1])$.

The normed space $\mathcal{B}(X, \Re)$ of all continuous linear operators from X to $\Re$ will play a significant role throughout the solvability analysis of ordinary differential equations. This space is called the conjugate space or dual space of X and is denoted by X^*. Elements of this space are called continuous linear functionals or simply functionals.

As claimed before, X^* is a normed space with norm of a functional $f \in X^*$ defined as

$$\|f\| = \sup_{\|\overline{x}\| \leq 1} |f(\overline{x})|$$

We also have the second conjugate $X^{**} = (X^*)^*$ defined in a similar way. One can show that $X \subseteq X^{**}$. However, if $X = X^{**}$, then X is said to be reflexive.

Definition 2.2.3 *A one to one and onto linear mapping from* X *to* Y *is called isomorphism. The spaces* X *and* Y *are then called isomorphic.*

If X and Y are finite dimensional spaces of the same dimension, then one can define an isomorphism between X and Y by defining a linear operator which maps basis to basis. With this notion, it follows that every finite dimensional space of dimension n is isomorphic to $\Re^n$.

Example 2.2.3 *Let* $X = \Re^n$. *Then* $X^* = \Re^n$ *(up to an isomorphism). To see this we proceed as follows.*

Let $S = \{\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_n\}$ *be an orthonormal basis for* $\Re^n$. *Define a set* $S^* = \{\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_n\} \subset X^*$ *as follows:*

$$\overline{f}_j(\overline{e}_i) = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

It is clear that $\{\overline{f}_j\}_{j=1}^n$ *are* $\ell.i.$ *Also, if* $\overline{f} \in X^*$, *one can write*

$$\overline{f} = \overline{f}(\bar{e}_1)\bar{f}_1 + \overline{f}(\bar{e}_2)\overline{f}_2 + \ldots + \overline{f}(\bar{e}_n)\overline{f}_n$$

Thus S^* *is a basis for* X^* *and* S^* *consists of* n $\ell.i.$ *vectors. So* X^* *is equal* $\Re^n$ *(up to an isomorphism).*

Example 2.2.4 *We denote by* $L_p[0,1]$ *(lp) the space of* p^{th} *power absolutely integrable functions (*p^{th} *power absolutely summable sequences),* $1 < p < \infty$. *One can show that* $L_p^*[0,1](lp^*) = L_q[0,1](lq), \dfrac{1}{p} + \dfrac{1}{q} = 1$. *(Refer Royden [10]). From this result it follows that* $L_p(lp)$ *spaces are reflexive spaces. In particular, we have* $L_2^*[0,1] = L_2[0,1](l_2^* = l_2)$.

Definition 2.2.4 *Let $T \in \mathcal{B}(X,Y)$. We define the conjugate operator $T^* : Y^* \to X^*$, of T as follows:*

$$[T^* f][\overline{x}] = f(T\overline{x}), f \in X^*, \overline{x} \in X$$

One can show that T^* is also bounded and $\|T^*\| = \|T\|$.
However, if X is a Hilbert space, then corresponding to $T \in \mathcal{B}(X,Y)$ we define $T^* \in \mathcal{B}(Y,X)$ as :

$$(\overline{x}, T^*\overline{y})_X = (T\overline{x}, \overline{y})_Y \quad \forall\, \overline{x} \in X,\ \overline{y} \in Y. \tag{2.2.3}$$

We say that $T \in \mathcal{B}(X)$ is self-adjoint if $T^* = T$.

Definition 2.2.5 *Let X be Hilbert space and $T \in \mathcal{B}(X)$ be a self adjoint operator. T is said to be positive semi-definite if*

$$(T\bar{x}, \bar{x}) \geq 0 \;\; \text{for all} \;\; \bar{x} \in X$$

T is called positive definite if the above inequality holds and is strict for $\bar{x} \neq \bar{0}$.
For $T \in \mathcal{B}(X,Y)$, the subspaces $N(T)$ and $R(T)$ are defined as follows

$$N(T) = \{\overline{x} \in X : T\overline{x} = 0\}, \;\; R(T) = \{\overline{y} \in Y : \overline{y} = T\overline{x},\ \overline{x} \in X\}$$

We have the following relationship between $N(T^*)$ and $R(T)$ and vice-versa.

Theorem 2.2.1 *Let X and Y be Hilbert spaces and let $T \in \mathcal{B}(X,Y)$. Then $N(T^*) = R(T)^{\perp}$ and $R(T^*)^{\perp} = N(T)$.*

Proof : It is sufficient to prove the first one. Assume that $\overline{x} \in N(T^*)$, then $T^*\bar{x} = 0$ and hence $(T^*\overline{x}, \overline{y}) = 0$ for all $\overline{y} \in X$. This implies that $(\overline{x}, T\overline{y}) = 0$ for all $\overline{y} \in X$. That is, $\overline{x} \in R(T)^{\perp}$ and hence $N(T^*) \subset R(T)^{\perp}$. Reversing the above argument, one can show that $R(T)^{\perp} \subset N(T^*)$ and hence the first part of the result. ∎

This theorem can be extended to get the following result.

Theorem 2.2.2 *Let X and Y be Hilbert spaces. If $T \in \mathcal{B}(X,Y)$, then $\overline{R(T)} = N(T^*)^{\perp}$ and $\overline{R(T^*)} = N(T)^{\perp}$.*

Remark 2.2.1 *If $R(T)$ is closed, we get*

$$R(T) = N(T^*)^{\perp}, R(T^*) = N(T)^{\perp}$$

For operators $T \in \mathcal{B}(X)$ with closed range in a Hilbert space, we get the orthogonal projector P from X to $M = R(T)$, if T^*T is invertible. This is done as follows by making use of Theorem 2.1.4.

Let $M = R(T)$. By Theorem 2.1.4 we can write $z \in X$ as follows:

$$\overline{z} = \overline{u} + \overline{v}, \;\; \overline{u} \in R(T) = M, \quad \overline{v} \in R(T)^{\perp} = N(T^*)$$

$\overline{v} \in N(T^*)$ implies that $T^*(\overline{z} - T\overline{x}) = 0$ and hence $T^*\overline{z} = T^*T\overline{x}$. This gives

$$\overline{x} = (T^*T)^{-1}T^*\overline{z} \tag{2.2.4}$$

The operator $P : X \to R(T) = M$ is called the orthogonal projector and is given by

$$P\overline{z} = \overline{u} = T\overline{x} = [T(T^*T)^{-1}T^*]\overline{z}$$

This element

$$\overline{u} = T(T^*T)^{-1}T^*\overline{z} \in R(T) = M \tag{2.2.5}$$

minimizes the functional $f(\overline{x}) = \|\overline{z} - T\overline{x}\|$ in $R(T) = M$.

Example 2.2.5 *Let $X = \Re^n$ and A be the matrix (α_{ij}). Then $A^* = (\beta_{ij})$ where $\beta_{ij} = \alpha_{ji}$. That is A^*, reduces to the transpose $A^\top$ of the matrix A. We have*

$$\begin{aligned}(A\overline{x}, \overline{y}) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}x_j y_i \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} \alpha_{ji}x_i y_j \\ &= \sum_{i=1}^{n}(\sum_{j=1}^{n} \alpha_{ji}y_j)x_i\end{aligned}$$

As $\beta_{ij} = \alpha_{ji}$, we get

$$\begin{aligned}(A\overline{x}, \overline{y}) &= \sum_{i=1}^{n}(\sum_{j=1}^{n} \beta_{ij}y_j)x_i \\ &= (\overline{x}, B\overline{y}), B = (\beta_{ij})\end{aligned}$$

This implies that $A^* = B = (\beta_{ij}) = (\alpha_{ji}) = A^\top$

Example 2.2.6 *If $X = \Re^n$ and $A = (a_{ij})$ is a $m \times n$ matrix with $(A^\top A)_{m\times m}$ invertible, then Eq. (2.2.5) gives the element $\overline{u} \in R(A)$ nearest to $R(A)$ from $\bar{z} \in \Re^n$. This formulation can be used in linear data fitting model leading to the least square approximation, as we see below.*

Data fitting model is used in measuring relationship between two physical quantities. Let the first quantity be represented by variables $x_1, x_2, \ldots, x_n$, giving the vector $\overline{x} \in \Re^n$, the second quantity by $\overline{y}$ and the functional relationship be given by

$$\overline{y} = f(\overline{x}), \quad \overline{x} = (x_1, \ldots, x_n)$$

We assume that the observed data f_k approximate the value $f(\overline{x}^{(k)})$. The problem of data fitting is to recover the function $f(\overline{x})$ from the given data $f_k, k = 1, 2, \ldots, m$. We try to fit the data by a linear function $F(\overline{x})$ given by

$$F(\bar{x}) = a_0 + a_1x_1 + \cdots + a_nx_n \tag{2.2.6}$$

We shall compute parameters a_k *so that the deviations*

$$d_k = f_k - F(\overline{x}^{(k)}; a_0, a_1, \ldots, a_n),\ k = 1, \ldots, m$$

are as small as possible.

Hence we minimize the error d_k *w.r.t. the 2-norm*

$$\|\bar{d}\|^2 = \sum_{k=1}^{m} d_k^2, \quad \bar{d} = (d_1, d_2, \ldots, d_m)$$

From the linear model given by Eq. (2.2.6), we have

$$\begin{aligned} \|\bar{d}\|^2 &= \sum_{k=1}^{m} (f_k - F(\bar{x}^{(k)}))^2 \\ &= \sum_{k=1}^{m} (f_k - \sum_{i=0}^{n} a_i x_i^{(k)})^2 \end{aligned}$$

Let $\bar{a} = (a_0, a_1, \ldots, a_n), \bar{z} = (f_1, f_2, \ldots, f_n)$

$$A = \begin{bmatrix} 1 & x^{(1)} & x^{(2)} & \cdots & x^{(n)} \\ 1 & x_2^{(1)} & x_2^{(2)} & \cdots & x_2^{(n)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_m^{(1)} & x_m^{(2)} & \cdots & x_m^{(n)} \end{bmatrix}.$$

Then the least square problem (data fitting problem) reduces to finding $\bar{u}^* = A\overline{a}^*$ *such that*

$$\|\overline{z} - \overline{u}^*\| = \inf_{\overline{a} \in \Re^n} \|\overline{z} - A\overline{a}\|$$

By using Eq. (2.2.5), the vector $\overline{u}^*$ *is given by*

$$\bar{u}^* = A(A^\top A)^{-1} A^\top \bar{z}$$

if $(A^\top A)$ *is invertible.*

Remark 2.2.2 *From the point of view of applications to differential equations, it is not possible to define the linear operator* $T = \dfrac{d}{dt}$ *on the whole space* $L_2[0,1]$. *In such a case we first define the domain* $D(T)$ *as follows:*

$$D(T) = \{f \in L_2[0,1] : \dot{f} = \frac{df}{dt} \text{ exists a.e. with } \dot{f} \in L_2[0,1]\}$$

$D(T)$ *is called the space* $AC[0,1]$ *of all absolutely continuous functions on* $[0,1]$. *It can be shown that* $D(T)$ *is dense in* $L_2[0,1]$. *However, it is not a bounded operator as it maps the sequence* $f_n(t) = \left\{\dfrac{\sin n\pi t}{n}\right\}$ *which converges to zero in* $L_2[0,1]$ *to a sequence* $Tf_n(t) = \pi \cos n\pi t$ *which is not convergent.*

Definition 2.2.6 *Let X be a Hilbert space and $T : D(T) \to X$ be such that $D(T)$ is dense in X. Then we define $D(T^*)$ to consist of $\overline{y} \in X$ for which the linear functional*

$$f(\overline{x}) = (T\overline{x}, \overline{y}), \quad \overline{x} \in D(T) \subset X$$

is continuous in $D(T)$. By Riesz representation theorem (refer Bose and Joshi [2]), there exists a unique $\overline{y}^* \in X$ such that

$$(T\overline{x}, \overline{y}) = (\overline{x}, \overline{y}^*) \;\; \forall\, \overline{x} \in D(T)$$

The adjoint T^* of T is defined on $D(T^*)$ by $T^*\overline{y} = \overline{y}^*$. In other words the adjoint T^* of T for densely defined operators is given by

$$(T\overline{x}, \overline{y}) = (\overline{x}, T^*\overline{y}), \;\; \forall\, \overline{x} \in D(T),\; \overline{y} \in D(T^*)$$

T is said to be self-adjoint iff $D(T^*) = D(T)$ and $T^* = T$.

Example 2.2.7 *Let $X = L_2[0,1]$ and $Tf(t) = -\ddot{f}(t) = -\dfrac{d^2 f}{dt^2}$, where $D(T)$ is defined as under*

$$D(T) = \{f \in L_2[0,1]\} : \dot{f}, \ddot{f} \in L_2[0,1] \text{ with } f(0) = f(1) = 0\}$$

One can show that $D(T)$ is dense in $L_2[0,1]$ and

$$\begin{aligned}
(Tf, g) &= -\int_0^1 \ddot{f}(t)g(t)dt = -\dot{f}(t)g(t)\Big|_0^1 + \int_0^1 \dot{f}(t)\dot{g}(t)dt \\
&= -\dot{f}(1)g(1) - \dot{f}(1)g(0) + f(t)\dot{g}(t)\Big|_0^1 - \int_0^1 f(t)\ddot{g}(t)dt \\
&= -\dot{f}(1)g(1) + \dot{f}(1)g(0) + f(1)\dot{g}(1) - f(0)\dot{g}(0) - \int_0^1 f(t)\ddot{g}(t)dt,\; f \in D(T) \\
&= -\dot{f}(1)g(1) + \dot{f}(1)g(0) - \int_0^1 f(t)\ddot{g}(t)dt \\
&= -\int_0^1 f(t)\ddot{g}(t)dt
\end{aligned}$$

if $g(0) = g(1) = 0$.

Hence

$$(Tf, g) = (f, Tg)$$

where $T^*g = -\ddot{g}$ and $D(T^*) = \{g \in X : \dot{g}, \ddot{g} \in L_2, g(0) = g(1) = 0\}$.

Thus $T^* = T$ and $D(T) = D(T^*)$. So, T is a self-adjoint operator and is densely defined. Also one can show that T is closed. That is if $x_n \to x$ in $X = L_2[0,1]$ and $Tx_n \to y$ in X. Then $x \in D(T)$ and $y = Tx$.

Definition 2.2.7 *$T \in \mathcal{B}(X,Y)$ is said to be compact if it maps bounded sets $B \subset X$ onto relatively compact set $T(B)$ ($\overline{T(B)}$ is compact). Equivalently, T is compact iff for every bounded sequence $\{\bar{x}_n\} \subset X, \{T\bar{x}_n\}$ has a convergent subsequence in Y.*

Definition 2.2.8 *$T \in \mathcal{B}(X,Y)$ is called finite rank operator if the range of T is finite dimensional.*

Theorem 2.2.3 *Every finite rank operator $T \in \mathcal{B}(X,Y)$ is compact.*

Proof : As $R(T)$ is finite dimensional, we have $R(T) = L[\overline{y}_1, \overline{y}_2, \ldots, \overline{y}_N]$. Consider now a bounded sequence $\{\overline{x}_n\} \subset X$, then $T\overline{x}_n = \sum_{j=1}^{N} \alpha_j^{(n)} \overline{y}_j$ for some scalars $\alpha_j^{(n)}$.

Since T is bounded $\{\alpha_j^{(n)}\}$ is bounded for each j and hence $\exists$ subsequence $\{\alpha_j^{(n_k)}\}$ such that $\alpha_j^{(n_k)} \to \alpha_j, 1 \leq j \leq N$. It now follows that $T\overline{x}_{nk} = \sum_{j=1}^{N} \alpha_j^{(n_k)} \overline{y}_j \to \overline{y} = \alpha_1 \overline{y}_1 + \alpha_2 \overline{y}_2 + \ldots + \alpha_N \overline{y}_N$. This proves the compactness of T. ∎

The following theorem gives the compactness property of the adjoint of a compact operator and also the limit of compact operators. X and Y are assumed to be Hilbert spaces.

Theorem 2.2.4 *(1) The composition of a compact operator with a bounded operator is compact.*

(2) Limit of a sequence of compact operator is compact.

(3) $T \in \mathcal{B}(X,Y)$ is compact iff $T^* \in \mathcal{B}(Y,X)$ is compact.

Refer Bose and Joshi [2] for the proof of this theorem.

Example 2.2.8 *Consider the operator $T: L_2[0,1] \to L2[0,1]$, as defined in Example 2.2.2:*

$$[Tf](t) = \int_0^1 k(t,s) f(s) ds, f \in L_2[0,1]$$

with

$$\int_0^1 \int_\infty^1 [k(t,s)]^2 dt ds < \infty.$$

Let $\{\phi_i(t)\}$ be a complete orthonormal set in $L_2[0,1]$, then the set $\{\phi_i(t)\phi_j(s)\}_{i,j=1}^{\infty}$ is a complete orthonormal set in $L_2([0,1] \times [0,1])$. Hence

$$k(t,s) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \phi_i(t) \phi_j(s)$$

where $a_{ij} = \int_0^1 \int_0^1 k(t,s)\phi_i(t)\phi_j(s) dt ds$. By Parseval's equality, we have

$$\int_0^1 \int_0^1 [k(t,s)]^2 dt ds = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (a_{ij})^2$$

Further, we have

$$\int_0^1 \int_0^1 \left[k(t,s) - \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\phi_i(t)\phi_j(s) \right]^2 dtds$$

$$= \int_0^1 \int_0^1 \left[\sum_{i,j=n+1}^{\infty} [a_{ij}\phi_i(t)\phi_j(s)] \right]^2 dtds$$

$$= \sum_{i,j=n+1}^{\infty} |a_{ij}|^2 \quad \to 0$$

since $\sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{\infty} |a_{ij}|^2 < \infty$.

Now define $K_n(t,s) = \sum\limits_{i,j}^{n} a_{ij}\phi_i(t)\phi_j(s)$.

Then $T_n : X \to X$ defined by

$$(T_n f)(t) = \int_0^1 K_n(t,s) f(s) ds$$

is a finite rank operator and hence it is compact. Also $T_n \to T$ in $\mathcal{B}(X)$ and hence by the previous theorem T is compact.

We have the following theorem, called Fredholm Alternative, giving the solvability of the operator equation involving compact operators.

Consider three related equations in a Hilbert space X (over the field of complex numbers)

$$T\overline{x} - \lambda \overline{x} = 0 \qquad (H)$$
$$T^*\overline{z} - \overline{\lambda}\overline{z} = 0 \qquad (H^*)$$
$$T\overline{y} - \lambda \overline{y} = f \qquad (NH)$$

Theorem 2.2.5 *1. Either both (H) and (H^*) have only trivial solutions or both have nontrivial solutions.*

2. The necessary and sufficient condition for (NH) to have a nontrivial solution is that f be orthogonal to every solution of (H^*).
In particular, if λ is not an eigenvalue of T then $\overline{\lambda}$ is not an eigenvalue of T^* and (NH) has one and only one solution for each $f \in X$.

We now state the following theorem which gives orthonormal basis for a Hilbert space X arising out of eigenvectors of $T \in \mathcal{B}(X)$, refer Bose and Joshi[2] for the proof.

Theorem 2.2.6 *Let T be a self-adjoint compact operator on a Hilbert space. Then the eigenvectors of T form a complete orthonormal basis for X.*

Corollary 2.2.1 *Let T be a compact and self-adjoint operator in a Hilbert space X. Then T is positive semidefinite (positive definite) iff all eigenvalues of T are nonnegative (positive).*

Example 2.2.9 *Let $X = \Re^n$ and $A : \Re^n \to \Re^n$ be given by a symmetric matrix $(a_{ij})_{n\times n}$. Then $A \in \mathcal{B}(\Re^n)$ is compact self-adjoint operator and hence the eigenvectors of A form a basis.*

Let

$$A\overline{\phi}_i = \lambda_i \overline{\phi}_i, 1 \le i \le n$$

where $\{\overline{\phi}_i\}_{i=1}^n$ are orthonormal.

Let $P = [\overline{\phi}_i, \overline{\phi}_2, \ldots, \overline{\phi}_n]$. Then P is an orthogonal matrix ($P^\top P = I$) and

$$AP = [A\overline{\phi}_i, \ldots, A\overline{\phi}_n] = [\lambda_1 \overline{\phi}_1, \ldots \lambda_n \overline{\phi}_n] = PD$$

where $D = diag[\lambda_1, \ldots, \lambda_n]$.

This gives us $P^{-1}AP = D$. Thus A is similar to a diagonal matrix. This is called diagonalisation procedure and will be used in Chapter 4 while computing transition matrix of a given system of differential equations. However, it may be pointed out A can be diagonalised even if A is not self-adjoint provided it has distinct eigenvalues. Assume that the eigenvectors corresponding to distinct eigenvalues are linearly independent and hence defining P as before, we get $P^{-1}AP = D$. But P is no longer an orthogonal matrix, it is merely nonsingular.

We now examine the properties of a nonlinear operator $F : X \to Y$, X and Y are normed spaces.

Definition 2.2.9 *$F : X \to Y$ is said to be continuous if $\overline{x}_n \to \overline{x}$ in $X \Rightarrow F\overline{x}_n \to F\overline{x} \ \forall x \in X$. F is said to be bounded if it maps bounded sets in X into bounded sets in Y.*

Compact nonlinear operators are defined in the same way as compact linear operators.

Remark 2.2.3 *For nonlinear operators continuity is not equivalent to boundedness, refer Joshi and Bose [6].*

Example 2.2.10 *Consider the following ODE*

$$\frac{dx}{dt} = f(t, x(t)), t \in I = [t_0, t_f] \tag{2.2.7}$$

with $f : I \times \Re \to \Re$ continuous and bounded.

Then by Proposition 1.1.1, the solvability of Eq. (2.2.7) is equivalent to the solvability of the equation

$$x(t) = x(t_0) + \int_{t_0}^{t} f(s, x(s))ds, \quad t \in I = [t_0, t_f] \tag{2.2.8}$$

The RHS of Eq. (2.2.8) gives rise to a nonlinear operator F on $C[t_0, t_f]$ defined as follows

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

We have

$$\begin{aligned} |[Fx](t_1) - [Fx](t_2)| &\leq \int_{t_1}^{t_2} |f(s, x(s))|ds \\ &\leq m|t_1 - t_2| (m = \sup_{(s,x)\in I\times\Re} |f(s, x)|) \end{aligned}$$

For this inequality it follows that Fx is continuous and hence $Fx \in C[t_0, t_f]$ if $x \in C[t_0, t_f]$. Further

$$|Fx_k(t) - Fx(t)| \leq \int_{t_0}^{t} |f(s, x_k(s)) - f(s, x(s)|ds$$

Let $x_k \to x$ in $C[t_0, t_f]$. This implies that $x_k(s) \to x(s)$ in $\Re$ for all $s \in [t_0, t_1]$. Since $(s, x) \to f(s, x)$ is continuous, it follows by bounded convergence theorem, that

$$|Fx_k(t) - Fx(t)| \leq \int_{t_0}^{t} |f(s, x_k(s)) - f(s, x, s)|ds \to 0$$

for all $t \in [t_0, t_f]$. One can also show that this convergence is uniform and hence

$$\|Fx_k - Fx\| = \sup_{t\in[t_0,t_f]} |Fx_k(t) - Fx(t)| \to 0$$

Hence F is a continuous nonlinear operator from $X = C[t_0, t_f]$ in to itself. Similarly, one can show that F is bounded on X.

The above result also holds true if $f(t, \overline{x})$ is defined on $I \times \Omega \to \Omega$ where Ω is an open subset of $\Re^n$. One then denotes by $C[t_0, t_f]$ the space of all $\Re^n-$ valued continuous functions and the norm of the elements $\overline{x} \in C[t_0, t_f]$ is defined as

$$\|\overline{x}\| = \sup_{t\in[t_0,t_f]} \|\overline{x}(t)\|_{\Re^n}$$

2.3 Lipschitz Continuity and Differentiability

In this section too the spaces under consideration will be normed space.

Definition 2.3.1 *$F : X \to Y$ is said to be Lipschitz continuous if $\exists$ constant $k > 0$ such that*

$$\|F\overline{x} - F\overline{y}\| \leq k\|\overline{x} - \overline{y}\| \;\; \forall\, \overline{x}, \overline{y} \in X \tag{2.3.1}$$

Definition 2.3.2 *$F : X \to Y$ is said to be differentiable at $\overline{x} \in X$ if $\exists\ A \in \mathcal{B}(X, Y)$ such that*

$$F(\overline{x} + \overline{h}) - F(\overline{x}) = A\overline{h} + w(\overline{x}, \overline{h}) \tag{2.3.2}$$

where $\lim_{\|\bar{h}\| \to 0} \dfrac{\|w(\overline{x}, \overline{h})\|}{\|\overline{h}\|} = 0$

A is called the derivative of F at $\overline{x}$ and is denoted by $F'(\overline{x})$. Its value at $\overline{h}$ will be denoted by $F'(\overline{x})\overline{h}$. One can show that F is differentiable at $\overline{x} \in X$ with derivative A iff

$$\lim_{t \to 0} \frac{F(\overline{x} + t\overline{h}) - F(\overline{x})}{t} = A\overline{h}\ \ \forall\ \ \overline{h} \in X$$

and $\bar{x} \to F'(\bar{x})$ is continuous.

Equivalently, writing $\phi(t) = F(\overline{x} + t\overline{h})$ for $\overline{x}, \overline{h} \in X$, we see that F has derivative $F'(\bar{x})$ iff

$$\frac{d}{dt}[\phi(t)]_{t=0} = F'(x)\overline{h}, \text{ for all } \bar{h} \in X$$

We have the following variant of the meanvalue theorem for $F : X \to Y$. Refer Joshi and Bose[6] for the proof of this theorem.

Theorem 2.3.1 *Let $F : X \to Y$ be differentiable at every $\bar{x}$ in X. Then for points $\bar{x}, \bar{x} + \bar{h} \in X$ and $\bar{e} \in Y^*$, there exists a constant τ, $0 < \tau < 1$, such that*

$$\big(\bar{e}, F(\bar{x} + \bar{h}) - F(x)\big) = \big(\bar{e}, F'(\bar{x} + \tau\bar{h})\bar{h}\big)$$

This immediately gives the following Corollary.

Corollary 2.3.1 *If $\|F'(\bar{x})\| \le K\ \forall\ \bar{x} \in X$ then F is Lipschitz continuous.*

Example 2.3.1 *Let $F : \Re^n \to \Re^n$ be given by $F(x_1, x_2, \dots, x_n) = (f_1(x_1, x_2, \dots, x_n), \dots, f_m(x_1, x_2, \dots x_n))$. Let $A = (a_{ij})$ be the $m \times n$ (matrix, which we want to compute) such that $F'(x) = A$.*

We choose $\bar{h} \in \Re^n$ as j^{th} coordinate vector $\overline{e}_j = (0, \dots, 1, \dots, 0)$ Then

$$\lim_{t \to 0} \| \frac{1}{t} F(\overline{x} + t\overline{h}) - F(\overline{x}) - tA\overline{h}\| = 0$$

$$\Leftrightarrow \lim_{t \to 0} | \frac{1}{t}[f_i(\overline{x} + \overline{e}_j) - f_i(\overline{x}) - ta_{ij}]| = 0, 1 \le i \le m, 1 \le j \le n$$

This shows that $F(\overline{x})$ is differentiable at $\overline{x} \in \Re^n$ iff $f_i(\overline{x})$ has continuous partial derivatives at $\overline{x} \in \Re^n$ and $F'(x) = A = \left(\dfrac{\partial f_i}{\partial x_j}(\overline{x})\right)$, if partial derivative are *continuous. So we have*

$$F'(\overline{x}) = \begin{bmatrix} \partial_1 f_1(\bar{x}) & \cdots & \partial_n f_1(\bar{x}) \\ \partial_1 f_2(\bar{x}) & \cdots & \partial_n f_2(\bar{x}) \\ \vdots & & \vdots \\ \partial_1 f_m(\bar{x}) & \cdots & \partial_n f_m(\bar{x}) \end{bmatrix}$$

In case f is a functional, that is, f is a mapping from $\Re^n$ to $\Re$, then $f'(\overline{x}) = \nabla f(\overline{x})$, called the gradient of f at $\overline{x}$.

Example 2.3.2 *Let X be a Hilbert space and let $f : X \to \Re$ be given by*

$$f(\overline{x}) = \frac{1}{2}(A\overline{x}, \overline{x}), \qquad \overline{x} \in X$$

We assume that A is a bounded self-adjoint linear operator. The derivative of f is computed as under.

$$\begin{aligned} f'(\overline{x})\overline{h} &= \lim_{t\to 0}\left[\frac{f(\overline{x}+t\overline{h}) - f(\overline{x})}{t}\right] \\ &= \lim_{t\to 0}\frac{1}{2}\left[\frac{(A(\overline{x}+t\overline{h}), (\overline{x}+t\overline{h})) - (A\overline{x}, \overline{x})}{t}\right] \\ &= \lim_{t\to 0}\frac{1}{2}\left[\frac{t\left[(A\overline{x}, \overline{h}) + (\overline{x}, A\overline{h})\right] + t^2(A\overline{h}, \overline{h})}{t}\right] \\ &= \lim_{t\to 0}\frac{1}{2}\left[(A\overline{x}, \overline{h}) + (A^*\overline{x}, \overline{h}) + t(A\overline{h}, \overline{h})\right] \\ &= \frac{1}{2}\left(\left(\frac{A+A^*}{2}\right)\overline{x}, \overline{h}\right) \\ &= (A\overline{x}, \overline{h}) \end{aligned}$$

as $A^* = A$.
This implies that $f'(\overline{x}) = A\overline{x} \;\forall\; \overline{x} \in X$. This derivative f' of a functional f is also denoted as ∇f.
From this representation it also follows that $f''(x) = A$.

Remark 2.3.1 *In dealing with minimization problem involving several variables we shall encounter convex functionals as in Example 2.3.2. A function $f : X \to \Re$ is said to be convex if*

$$f(\alpha\overline{x} + (1-\alpha)\overline{y}) \le \alpha f(\overline{x}) + (1-\alpha)f(\overline{y}), 0 < \alpha < 1$$

for all $\overline{x}, \overline{y} \in X$. X is any normed space.

Remark 2.3.2 *One can show that f is strictly convex on a Hilbert space if $f^{''}(x)$ is positive definite (refer problem number 8 in sectoin 2.6). Hence the functional of Example 2.3.2 is strictly convex, if A is positive definite.*

Example 2.3.3 (Least Square Approximation - Data Fitting) *In Example 2.2.5 involving problems of least square approximations, we minimize $\|\overline{d}\|^2$ given by*

$$\|\overline{d}\|^2 = \|\overline{z} - A\overline{a}\|^2 = \sum_{k=1}^{m}(f_k - \sum_{i=0}^{n} a_i x_i^{(k)})^2$$

where $\overline{z}$ and A are as defined in the example. Let $\phi(\overline{a})$ be defined as follows

$$\begin{aligned} \phi(\overline{a}) &= (\overline{z} - A\overline{a}, z - A\overline{a}) \\ &= (\overline{z}, \overline{z}) - 2(A\overline{a}, \overline{z}) + (A\overline{a}, A\overline{a}) \\ &= (\overline{z}, \overline{z}) - 2(A\overline{a}, \overline{z}) + (A^\top A\overline{a}, \overline{a}), \quad \bar{a} \in \Re^n \end{aligned}$$

Proceeding as in Example 2.3.2, we get

$$\nabla\phi(\overline{a}) = -2A^\top\overline{z} + 2A^\top A\overline{a}$$

It is clear that $A^\top A$ is positive semi-definite. If $A^\top A$ is invertible then $A^\top A$ is positive definite and hence f is strictly convex and its critical point is the minimizer of ϕ. Also $\overline{a}^*$ is the critical point of ϕ iff $\nabla\phi(\overline{a}^*) = 0$

This implies that $A^\top\overline{z} = A^\top A\overline{a}^*$ and hence $\overline{a}^* = (A^\top A)^{-1}A^\top\overline{z}$.

This is the same result, which we obtained earlier.

Remark 2.3.3 *While solving differential equations, we shall also be estimating a set of parameters which will arise in the defining model. Estimating parameters requires the minimization of a functional involving these parameters.*

Such functionals f may not be convex and hence there is a need to generate algorithms which will computationally capture the minimizer (local / global). We give below one such algorithm - called the steepest descent algorithm - which captures the local minimizer of a function f defined on the entire space $\Re^n$.

Steepest Descent Algorithm 2.3.1

Step 1	:	Start with $\overline{x}^{(k)}(k = 0)$
Step 2	:	Set $\overline{d}^{(k)} = -\overline{g}^{(k)}, \overline{g}^{(k)} = \nabla f(\overline{x}^{(k)})$
Step 3	:	Minimize 1-dimensional function $\phi(\alpha) = f(\overline{d}^{(k)} + \alpha\overline{d}^{(k)})$ to get $\alpha^{(k)}$
Step 4	:	Set $\overline{x}^{(k+1)} = \overline{x}^{(k)} + \alpha^{(k)}\overline{d}^{(k)}$
Step 5	:	Use any stopping criterion and set $\overline{x}^* = \overline{x}^{(k+1)}$. Otherwise set $k = k + 1$ or GO TO STEP 2
Step 6	:	Compute $f(\overline{x}^*), \overline{x}^*$ is the point of optimum.

In Example 2.3.2, the iteration procedure for the functional

$$f(x) = \frac{1}{2}(A\overline{x}, \overline{x}), \quad \overline{x} \in \Re^n$$

is given by

$$\overline{x}^{(k+1)} = \overline{x}^{(k)} - \frac{\left(\overline{g}^{(k)}, \overline{g}^{(k)}\right)}{\left(A\overline{g}^{(k)}, \overline{g}^{(k)}\right)}\overline{g}^{(k)} \tag{2.3.3}$$

Example 2.3.4 *Let $F : X = C[0,1] \to C[0,1]$ be defined by*

$$[Fx](t) = \int_0^1 k(t,s)f(s,x(s))ds$$

where k and f satisfy the following properties

(i) *$k(t,s)$ is continuous on $[0,1] \times [0,1]$.*

(ii) *$x \to f(t,x)$ is differentiable mapping from $\Re$ to $\Re$ for all $t \in [0,1]$ such that*

$$|\frac{\partial f}{\partial x}(t,x)| \leq M \text{ for all } (t,x) \in [0,1] \times \Re$$

As in Example 2.2.11, we can prove that F is a continuous mapping of $C[0,1]$ to itself. We now compute the derivative of F.

$$\begin{aligned}
[F'(x)\overline{h}] &= \frac{d}{d\alpha}[F(x+\alpha h)]_{\alpha=0} \\
&= \lim_{\alpha \to 0}[\frac{F(x+\alpha h) - F(x)}{\alpha}] \\
&= \lim_{\alpha \to 0}\int_0^1 k(t,s)\frac{[f(s,(x+\alpha h)(s)) - f(s,\bar{x}(s))]}{\alpha} \\
&= \int_0^1 k(t,s)\frac{\partial}{\partial x}f(s,x(s))h(s)ds
\end{aligned}$$

by Lebesgue dominated convergence theorem. Thus F is differentiable at x with $F'(x)$ as a linear integral operator generated by the kernel $g(t,s) = k(t,s)\frac{\partial f}{\partial x}(s,x(s))$ and is given by

$$F'(x)h(t) = \int_0^1 g(t,s)h(s)ds.$$

Also, as $g(t,s)$ is a bounded on $[0,1] \times [0,1]$, it follows that $\|F'(x)\| \leq M$ and hence by Corollary 2.3.1 it follows that F is Lipschitz continuous on $X = L_2[0,1]$.

Let F be differentiable at $\overline{x}_0 \in X$, then it follows that we have the following expansion for $F(\overline{x})$ with $F'(\bar{x}_0) \in \mathcal{B}(X)$.

$$F(\overline{x}) = F(\overline{x}_0) + F'(\overline{x}_0)(\overline{x} - \overline{x}_0) + w(\overline{x}, \overline{h}) \tag{2.3.4}$$

where $\dfrac{\|w(\overline{x}, \overline{h}\|}{\|\overline{h}\|} \to 0$ as $\|\overline{h}\| \to 0$.

Definition 2.3.3 *If F is differentiable at $\overline{x}_0$ then the function*

$$\hat{F}(\overline{x}) = F(\overline{x}_0) + F'(\overline{x}_0)(\overline{x} - \overline{x}_0) \tag{2.3.5}$$

is called the linearization of F at $\overline{x}_0$.

Remark 2.3.4 *The linearization process can be used to find the zeros of the nonlinear function $F(\overline{x})$ in a neighbourhood of the operating point $\overline{x}_0 \in X$ if $[F^{'}(\bar{x}_0)]$ is invertible.*

By linearization, it suffices to compute the zeros $\hat{F}(\overline{x})$ and hence

$$0 = \hat{F}(\overline{x}) = F(\overline{x}_0) + F'(\overline{x}_0)(\overline{x} - \overline{x}_0).$$

This gives

$$\overline{x} = \overline{x}_0 + \left[F'(x_0)\right]^{-1} F(\overline{x}_0)$$

Hence we get the Newton's algorithm to find the zeros of $F(\overline{x})$, starting with an initial guess $\bar{x}_0$.

Algorithm 2.3.2 - Newton's Algorithm

Step 1 Set $k = 0$ and input initial guess $\overline{x}_k$ and the tolerance $\epsilon > 0$

Step 2 $\overline{x}^{(k+1)} = \overline{x}^{(k)} - [F'(\overline{x}^{(k)})]^{-1} F(\overline{x}^{(k)})$

Step 3 If $\|\overline{x}^{(k+1)} - \overline{x}^{(k)}\| < \in$, STOP and set $\overline{x}^* = \overline{x}^{(k)}$ as the computed zero of $F(\bar{x})$. Otherwise increment k by 1.

Step 4 GO To Step 2

Example 2.3.5 *We have shown in Example 1.3.2 that the motion of a satellite orbiting around the earth is given by an ordinary differential equation (ODE) in $\Re^4$:*

$$\frac{d\bar{x}}{dt} = f(\bar{x}, \bar{u}), \bar{x} \in \Re^4, \bar{u} \in \Re^2 \tag{2.3.6}$$

$\bar{x}(t) = (x_1(t), x_2(t), x_3(t), x_4(t))$ represents the state vector and $\bar{u}(t) = (u_1(t), u_2(t))$ is the control vector (thrusts on the satellite), f is given by

$$f(\bar{x}, \bar{u}) = (f_1(\bar{x}, \bar{u}), f_2(\bar{x}, \bar{u}), f_3(\bar{x}, \bar{u}), f_4(\bar{x}, \bar{u}))$$

where

$$f_1 = x_2, \quad f_2 = (x_1 + \sigma)(\frac{x_4}{\sigma} + w)^2 - \frac{k}{(x_1 + \sigma)^2} + u_1,$$
$$f_3 = x_4, \quad f_4 = -2\sigma(\frac{x_4}{\sigma} + w)\frac{x_2}{(x_1 + \sigma)} + \frac{u_2\sigma}{(x_1 + \sigma)}$$

Using linearization procedure for $f(\bar{x},\bar{u})$, described by Eq. (2.3.5), around $\bar{x}=0,\bar{u}=0$ we get

$$\hat{f}(\bar{x},\bar{u}) = f'_x(\bar{0},\bar{0})\bar{x} + f'_u(\bar{0},\bar{0})\bar{u}$$

$$A = f'_x(\bar{0},\bar{0}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} & \frac{\partial f_1}{\partial x_4} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4} \\ \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_3}{\partial x_4} \\ \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4} \end{bmatrix}_{(\bar{o},\bar{o})} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3w^2 & 0 & 0 & 2w \\ 0 & 0 & 0 & 1 \\ 0 & -2w & 0 & 0 \end{bmatrix} \tag{2.3.7}$$

$$B = f'_u(\overline{0},\overline{0}) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \\ \frac{\partial f_3}{\partial u_1} & \frac{\partial f_3}{\partial u_2} \\ \frac{\partial f_4}{\partial u_1} & \frac{\partial f_4}{\partial u_2} \end{bmatrix}_{(\bar{o},\bar{o})} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.3.8}$$

Here σ is normalized to 1. So the linearized ODE, corresponding to Eq. (2.3.6) is

$$\frac{d\bar{x}}{dt} = A\bar{x} + B\bar{u}$$

where A and B are given by Eq. (2.3.7) and Eq. (2.3.8), respectively.

2.4 Fixed Point Theorems and Monotone Operators

Definition 2.4.1 *Let X be a metric space and $F : X \to X$. An element $x \in X$ is said to be a fixed point of the mapping F if $Fx = x$.*

We begin with the following constructive fixed point theorem by Banach, known as Banach contraction mapping principle. This theorem will be one of the important instruments in obtaining existence theorems for ordinary differential equations.

Definition 2.4.2 *Let (X,d) be a metric space. $F : X \to X$ is said to be contraction if F is Lipschitz continuous with Lipschitz constant strictly less than*

1. That is, there exists a positive constant $\alpha < 1$ such that

$$d(Fx, Fy) \leq \alpha d(x, y) \quad \text{for all } x, y \in X \tag{2.4.1}$$

If X is a normed space, then (2.4.1) is equivalent to the following inequality

$$\|F\bar{x} - F\bar{y}\| \leq \alpha \|\bar{x} - \bar{y}\| \quad \text{for all } \bar{x},\ \bar{y} \in X \tag{2.4.2}$$

Theorem 2.4.1 *Let F be a contraction mapping of a complete metric space X into itself. Then F has a unique fixed point $x^* \in X$. Moreover if $x_0 \in X$ is any arbitrary point, and $\{x_n\}$ a sequence defined by*

$$x_{n+1} = Fx_n \ (n = 0, 1, 2, \ldots)$$

then $\lim\limits_{n\to\infty} x_n = x^*$ and $d(x_n, x^*) \leq \dfrac{\alpha^n}{1-\alpha} d(x_1, x_0)$.

Proof : We have

$$\begin{aligned} d(x_k, x_{k-1}) &= d(Fx_{k-1}, Fx_{k-2}) \\ &\leq \alpha d(x_{k-1}, x_{k-2}) \\ &\leq \alpha d(x_1, x_0) \end{aligned}$$

This gives

$$\begin{aligned} d(x_\ell, x_m) &\leq d(x_\ell, x_{\ell-1}) + \ldots + d(x_{m+1}, x_m) \\ &\leq [\alpha^m + \alpha^{m+1} + \ldots + \alpha^{l-1}] d(x_1, x_0) \\ &\leq \alpha^m [1 + \alpha + \alpha^2 + \ldots] d(x_1, x_0) \\ &= \left[\frac{\alpha^m}{1-\alpha}\right] d(x_1, x_0), l > m \end{aligned}$$

$\to 0$ as $m \to \infty$

This implies that $\{x_n\}$ is Cauchy in X and as X is complete, it converges to $x^* \in X$. As $x_n \to x^*$, $x_{n+1} = Fx_n$ and F is continuous, it follows that $x^* = Fx^*$.

The uniqueness of the fixed point x^* follows from the fact that for two fixed point x^* and y^*, we have

$$\begin{aligned} d(x^*, y^*) &= d(Fx^*, Fy^*) \\ &\leq \alpha d(x^*, y^*) \\ &< d(x^*, y^*) \end{aligned}$$

and hence $d(x^*, y^*) = 0$.

The error estimate $d(x_n, x^*)$ is given by

$$\begin{aligned} d(x_n^*, x^*) &\leq d(x_m, x^*) + d(x_m, x_n) \\ &\leq d(x_m, x^*) + \frac{\alpha^n}{1-\alpha} d(x_1, x_0), m > k \end{aligned}$$

Letting $m \to \infty$, we get

$$d(x_n, x^*) \leq \frac{\alpha^n}{1-\alpha} d(x_1, x_0)$$

∎

As a Corollary of this theorem, we obtain the solvability of the linear equation

$$\overline{x} - A\overline{x} = \overline{y} \tag{2.4.3}$$

in a Banach space X. Here, $A \in \mathcal{B}(X)$ with $\|A\| < 1$ and $\bar{y} \in X$.

Corollary 2.4.1 *Let $A \in \mathcal{B}(X)$ be such that $\|A\| < 1$. Then $(I - A)$ is bounded and invertible with*

$$\|(I-A)^{-1}\| \leq \frac{1}{1 - \|A\|} \tag{2.4.4}$$

Proof : We first show that

$$\overline{x} - A\overline{x} = \overline{y}, \ \ \overline{y} \in X$$

has a unique solution. Let $F\overline{x} = A\overline{x} + \overline{y}$. Then

$$\|F\overline{x}_1 - F\overline{x}_2\| = \|A\overline{x}_1 - A\overline{x}_2\| \leq \|A\| \|\overline{x}_1 - x_2\| \text{ for all } \overline{x}_1, \overline{x}_2 \in X$$

As $\alpha = \|A\| < 1$, it follows that F is a contraction and hence it has a unique fixed point $\overline{x}$ in the Banach space X (complete metric space). That is, Eq. (2.4.3) has a unique solution $\overline{x}$. Further, let

$$\overline{x} = (I - A)^{-1}\overline{y}$$

This implies that $\overline{x} - A\overline{x} = \overline{y}$, $\overline{y} \in X$ arbitrary and hence

$$\begin{aligned} \|\overline{x}\| &\leq \|A\overline{x}\| + \|\overline{y}\| \\ &\leq \|A\| \|\overline{x}\| + \|\overline{y}\| \end{aligned}$$

This gives

$$\|\bar{x}\| \leq \left[\frac{1}{1 - \|A\|}\right] \|\overline{y}\|$$

Hence

$$\|(I-A)^{-1}\overline{y}\| \leq \left[\frac{1}{1 - \|A\|}\right] \|\overline{y}\|$$

for all $\bar{y} \in X$.

This implies that

$$\|(I-A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

∎

Example 2.4.1 *Solve the linear system*

$$A\bar{x} = \bar{b} \tag{2.4.5}$$

in $\Re^n$, $A = (\alpha_{ij})_{n\times n}$. *Expanding Eq. (2.4.5) in terms of its component we get*

$$\begin{array}{l} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}xn = b_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}xn = b_2 \\ \quad\vdots \\ \alpha_{n_1}x_1 + \alpha_{n_2}x_2 + \ldots + \alpha_{n_n}x_n = b_n \end{array} \tag{2.4.6}$$

Assume that $\alpha_{ii} > 0$ *and the matrix* A *is diagonally dominant:*
$\sum_{\substack{i=1 \\ i\neq j}}^{n} |\alpha_{ij}| < \alpha_{ii}, 1 \le i \le n.$

Then Eq. (2.4.6) is equivalent to

$$\begin{aligned} x_1 &= \frac{b_1}{\alpha_{11}} - \frac{\alpha_{12}x_2}{\alpha_{11}} \ldots - \frac{\alpha_{1n}x_n}{\alpha_{11}} \\ x_2 &= \frac{b_2}{\alpha_{22}} - \frac{\alpha_{21}x_1}{\alpha_{22}} - \ldots - \frac{\alpha_{2n}x_n}{\alpha_{22}} \\ &\vdots \\ x_n &= \frac{b_n}{\alpha_{nn}} - \frac{\alpha_{n},x_1}{\alpha_{nn}} - \ldots - \frac{\alpha_{n,n-1}x_{n-1}}{\alpha_{nn}} \end{aligned} \tag{2.4.7}$$

Define

$$B = -\begin{pmatrix} 0 & \frac{\alpha_{12}}{\alpha_{11}} & \cdots & \frac{\alpha_{1n}}{\alpha_{11}} \\ \frac{\alpha_{21}}{\alpha_{22}} & 0 & \cdots & \frac{\alpha_{2n}}{\alpha_{22}} \\ & \vdots & & \\ \frac{\alpha_{n1}}{\alpha_{nn}} & \frac{\alpha_{n_2}}{\alpha_{nn}} & \cdots & 0 \end{pmatrix}, \bar{c} = \begin{pmatrix} \frac{b_1}{\alpha_{11}} \\ \frac{b_2}{\alpha_{22}} \\ \vdots \\ \frac{b_n}{\alpha_{nn}} \end{pmatrix}$$

Then the solvability of Eq. (2.4.7) is equivalent to the solvability of the matrix equation

$$\bar{x} - B\bar{x} = \bar{c}, \quad \|B\| < 1.$$

Hence by the above corollary, this has a unique solution $\bar{x}^*$ *and the iteration scheme*

$$\bar{x}^{(k+1)} = \bar{c} + B\bar{x}^{(k)}$$

converges to $\bar{x}^*$.
This gives us the well known Jacobi's iteration scheme (refer Conte and deBoor [4]).

$$
\begin{aligned}
x_1^{(k+1)} &= \frac{b_1}{\alpha_{11}} - \frac{\alpha_{12}}{\alpha_{11}} x_2^{(k)} \cdots - \frac{\alpha_{2n}}{\alpha_{11}} x_n^{(k)} \\
x_2^{(k+1)} &= \frac{b_n}{\alpha_{22}} - \frac{\alpha_{21}}{\alpha_{22}} x_1^{(k)} \cdots - \frac{\alpha_{2,n}}{\alpha_{22}} x_n^{(k)} \\
\vdots &= \vdots \qquad -\vdots \qquad\qquad -\vdots \\
x_n^{(k+1)} &= \frac{b_n}{\alpha_{nn}} - \frac{\alpha_{n1}}{\alpha_{nn}} x_1^{(k)} \cdots \frac{\alpha_{n,n-1}}{\alpha_{nn}} x_n^{(k)}
\end{aligned}
$$

For the diagonally dominant matrix

$$
A = \begin{pmatrix} 24 & 3 & 9 & 7 \\ 1 & 12 & -3 & -4 \\ 0 & 1 & 22 & -3 \\ 7 & -4 & -3 & 24 \end{pmatrix}, \quad \bar{b} = (1, 1, 2, -4)
$$

we get the following iterated solution $\bar{x}^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, x_4^{(n)})$ of $A\bar{x} = \bar{b}$.

Table 2.4.1: Iterated solution of the matrix equation

n	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$	$x_4^{(n)}$	$\|x - x^{(n)}\|$
0	1	0	-1	2	
1	-0.16667	0.41667	0.36364	-0.58333	5.53030
2	0.02336	-0.00631	-0.00758	-0.00316	1.56439
3	0.04622	0.07844	0.09077	-0.17548	0.37828
4	0.04901	0.04368	0.06341	-0.15573	0.08465
5	0.05785	0.04319	0.06769	-0.16575	0.02363
6	0.05923	0.04018	0.06634	-0.16788	0.00786
7	0.06073	0.03902	0.06619	-0.16895	0.00389
8	0.06125	0.03850	0.06610	-0.16960	0.00178
9	0.06154	0.03822	0.06603	-0.16986	0.00089

It may happen that $F : (X, d) \to (X, d)$ is not a contraction but some power of it is a contraction. We still get a unique fixed point, as we see in the following corollary.

Corollary 2.4.2 *Let (X, d) be a complete metric space. $F : (X, d) \to (X, d)$ be such that F^N is a contraction for some positive integer N. Then F has a unique fixed point.*

Proof : By the contraction mapping principle, F^N has a unique fixed point x^*. That is, $\exists\ x^* \in X$ such that $F^N x^* = x^*$. Applying F to both sides, we get $F^{N+1}(x^*) = F(x^*)$ and hence $F^N(F(x^*)) = F(x^*)$. This implies that $F(x^*)$ is also a fixed point of F^N and uniqueness implies that $Fx^* = x^*$. Also every fixed point of F is also a fixed point of F^N and hence fixed point of F is unique. ∎

Example 2.4.2 *Let $F : C[a,b] \to C[a,b]$ be defined by*

$$[Fx](t) = \int_a^t x(s)ds$$

We have

$$[F^k x](t) = \frac{1}{k-1!}\int_a^t (t-s)^{k-1} x(s)ds$$

This gives

$$\begin{aligned}
|(F^k x - F^k y)(t)| &\leq \frac{1}{k-1!}\int_a^t (t-s)^{k-1}|x(s)-y(s)|ds \\
&\leq \frac{(b-a)^k}{k!}\sup_{s\in[a,b]}|x(s)-y(s)| \\
&= \frac{(b-a)^k}{k!}\|x-y\|
\end{aligned}$$

This implies that

$$\|F^k x - F^k y\| \leq \frac{(b-a)^k}{k!}\|x-y\| \ \forall\ x,y \in X = C[a,b]$$

As $\frac{(b-a)^k}{k!} \to 0$ as $k \to \infty$, it follows that $\exists$ positive integer N such that $\frac{(b-a)^N}{N!} < 1$. Thus there exist a positive integer N such that $\alpha = \frac{(b-a)^N}{N!} < 1$ and $\|F^N x - F^N y\| \leq \alpha\|x-y\| \ \forall \ \ x,y \in X = C[a,b]$. That is F^N is a contraction. However, F is not a contraction if $(b-a) > 1$.

At times, we shall encounter mappings in a normed space, which are not contractions but have compactness property. For such operators, we have the following Schauder's theorem, giving a fixed point of F.

Theorem 2.4.2 *Let K be a nonempty closed convex subset of a normed space X. Let F be a continuous mapping of K into a compact subset of K. Then F has a fixed point in K.*

Another important notion which will be used in solvability analysis of differential equation is the concept of monotone operators in Hilbert spaces. We shall now define this concept and obtain some abstract results which subsequently will be used in Chapter 3 while dealing with existence and uniqueness theorems. In the following, the space X is assumed to be a Hilbert space.

Definition 2.4.3 *$F : X \to X$ is said to be monotone if*

$$(Fx_1 - Fx_2, x_1 - x_2) \geq 0 \quad \text{for all} \;\; x_1, x_2 \in X \tag{2.4.8}$$

F is called strictly monotone if the above inequality holds and is strict for $x_1 \neq x_2$. F is called strongly monotone if $\exists$ constant $c > 0$ such that

$$(Fx_1 - Fx_2, x_1 - x_2) \geq c\|x_1 - x_2\|^2 \text{ for all } x_1, x_2 \in X \tag{2.4.9}$$

If F is not defined on the whole space, then we verify Eq. (2.4.8) - Eq. (2.4.9) on $D(F)$, the domain of F.

Definition 2.4.4 *A multivalued mapping $F : X \to 2^X$ is called monotone if*

$$(\bar{y}_1 - \bar{y}_2, \bar{x}_1, \bar{x}_2) \geq 0 \;\; \forall \;\; \bar{x}_1, \bar{x}_2 \in D(F) \quad \text{and} \quad \bar{y}_1 \in F\bar{x}_1, \;\; \bar{y}_2 \in F\bar{x}_2$$

A monotone mapping F is called maximal monotone if it has no proper monotone extensions.

Example 2.4.3 *$A : L^2[0,1]$ be the differential operator $-\dfrac{d^2}{dt^2}$ with $D(A)$ defined as*

$$D(A) = \{x \in L^2[0,1] : \dot{x}(t), \ddot{x}(t) \in L^2[0,1]; \quad x(0) = x(1) = 0\}$$

A has dense domain and it is self adjoint

$$(A\bar{x}, \bar{y}) = \left(- \left(\frac{d^2x}{dt^2} \right), y(t) \right) = \left(x, - \left(\frac{d^2y}{dt^2} \right) \right); x, y \in D(A)$$

A is monotone as

$$\begin{aligned} (A\overline{x}, \overline{x}) &= -\int_a^1 \frac{d^2x}{dt^2} x(t)dt \\ &= \int_0^1 (\frac{dx}{dt})^2 dt, \quad \bar{x} \in D(A) \\ &\geq 0 \end{aligned}$$

Example 2.4.4 *Let $f(s,x)$ be a function defined on $[0,1] \times \Re$ to $\Re$ be such that $f(s,x)$ is measurable w.r.t. s for all $x \in \Re$ and continuous w.r.t. x for almost all $s \in [0,1]$. Let $F : X = L_2[0,1] \to L_2[0,1]$ be defined by*

$$[Fx](s) = f(s, x(s))$$

If we assume a growth condition of the type

$$|f(s,x)| \leq a(s) + b|x|, \quad a \in L_2[0,1], b > 0$$

on f, then one can show that F maps X into itself and is continuous and bounded (refer Joshi and Bose[6]). Further, assume that the mapping $x \to f(s,x)$ is monotone increasing for almost all $s \in [0,1]$. Then we have

$$(F\bar{x}_1 - F\bar{x}_2, \bar{x}_1 - \bar{x}_2) = \left(\int_0^1 (f(s, x_1(s)) - f(s, x_2(s)))(x_1(s) - x_2(s))ds \right) \tag{2.4.10}$$

As the integrand in the above equation is nonnegative for all $x_1, x_2 \in X$, it follows that F is monotone.

Example 2.4.5 *Let $F : X \to X$ be a contraction. Then $(I - F)$ is strongly monotone.*

$$
\begin{aligned}
((I-F)\bar{x}_1 - (I-F)\bar{x}_2, \bar{x}_1 - \bar{x}_2) &= (\bar{x}_1 - \bar{x}_2, \bar{x}_1 - \bar{x}_2) - (F\bar{x}_1 - F\bar{x}_2, \bar{x}_1 - \bar{x}_2) \\
&\geq \|\bar{x}_1 - \bar{x}_2\|^2 - \|F\bar{x}_1 - F\bar{x}_2\| \ \|x_1 - x_2\| \\
&\geq \|\bar{x}_1 - \bar{x}_2\|^2 - \alpha\|\bar{x}_1 - \bar{x}_2\|^2 \\
&= (1-\alpha)\|\bar{x}_1 - \bar{x}_2\|^2 \ \ \forall \ \bar{x}_1, \bar{x}_2 \in X
\end{aligned}
$$

We have the following solvability result due to Browder[3], concerning strongly monotone continuous operators on X.

Theorem 2.4.3 *Let $F : X \to X$ be a strongly monotone continuous operator. Then $F\bar{x} = \bar{y}$ has a unique solution for any $y \in X$. Moreover F^{-1} is Lipschitz continuous and monotone.*

In view of Example 2.4.5, we get the following Corollary.

Corollary 2.4.3 *Let $F : X \to X$ be a contraction on the Hilbert space X. Then $\bar{x} = F\bar{x} + \bar{y}$ has a unique solution $\bar{x}$ for every $\bar{y} \in X$ and this solution $\bar{x}$ varies continuously w.r.t. $\bar{y}$.*

Definition 2.4.5 *$F : X \to X$ is called coercive if $\lim\limits_{\|\bar{x}\| \to \infty} \dfrac{(F\bar{x}, \bar{x})}{\|\bar{x}\|} = \infty$.*

For functions f defined on $\Re$, this corresponds to the condition $\lim\limits_{x \to \infty} f(x) = \infty$ and $\lim\limits_{x \to -\infty} f(x) = -\infty$.

We have the following result due to Minty [9].

Theorem 2.4.4 *Let $F : X \to X$ be continuous, monotone and coercive. Then $R(F) = X$.*

While dealing with differential equations, we shall encounter operators, which are only densely defined. For example, $F = L + F_0$ where L is densely defined linear maximal monotone operator and F_0 is monotone operator defined on the whole space. Then the following theorem gives the surjectivity of $F = L + F_0$. Refer Joshi and Bose [6] for the proof.

Theorem 2.4.5 *Let F_1 be a maximal monotone operator with dense domain containing 0 and F_2 a single valued operator which is defined on the whole space and is continuous, bounded and coercive. Then $R(F) = X$ where $F = F_1 + F_2$.*

2.5 Dirac-Delta Function

During the course of our discussion on boundary value problems, we shall encounter Dirac-delta function $\delta(x)$, which is a symbolic function representing the charge density corresponding to a unit charge at the origin. To specify it, we take a continuously distributed charge on the real axis with charge density

$$\rho_\epsilon(t) = \frac{1}{\pi}\left[\frac{\epsilon}{t^2+\epsilon^2}\right]$$

$\epsilon > 0$ and $\rho_\epsilon(t)$ is small for small ϵ but for a peak height $\dfrac{1}{\epsilon\pi}$ at the origin.

The cumulative charge distribution $\gamma_\epsilon(t)$ is given by

$$\gamma_\epsilon(t) = \int_{-\infty}^{t} \rho_\epsilon(x)dx = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(t/\epsilon)$$

Hence

$$\lim_{t\to\infty} \gamma_\epsilon(t) = \int_{-\infty}^{\infty} \rho_\epsilon(x)dx = 1 \text{ for all } \epsilon$$

So, the total charge for the densily distribution $\rho_\epsilon(t)$ on the line is 1, whereas

$$\lim_{\epsilon\to 0} \rho_\epsilon(t) = \begin{cases} 0, & t \neq 0 \\ \infty, & t = 0 \end{cases}$$

This limiting charge is denoted by $\delta(t)$. That is

$$\delta(t) = \lim_{\epsilon\to 0} \rho_\epsilon(t)$$

So, vaguely the Dirac-delta function $\delta(t)$ can be defined as

$$\delta(t) = \begin{cases} 0, & t \neq 0 \\ \infty, & t = 0 \end{cases}$$

such that $\displaystyle\int_{-\infty}^{\infty} \delta(t)dt = 1$

Obviously, this is not a function.

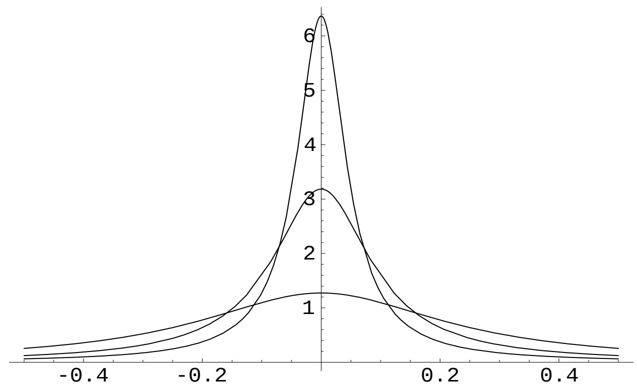

Figure 2.5.1: Charge density

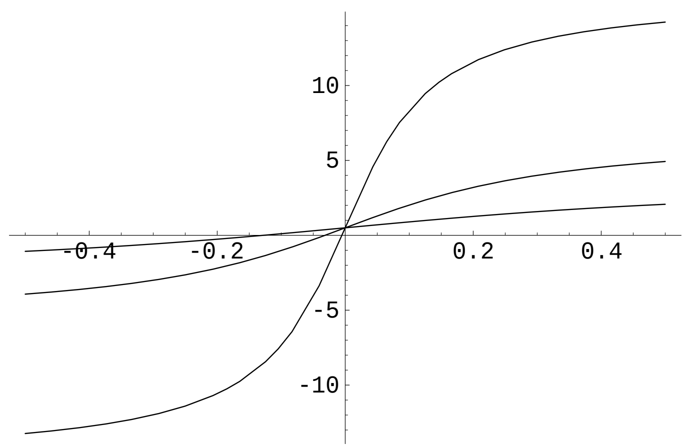

Figure 2.5.2: Cumulative charge density

The description of a point charge by considering it as a limit of the sequence $\rho_\epsilon(t)$ is not quite satisfactory, mathematically. For, one can take various types of sequences to get the limit $\delta(t)$, yet these sequences can behave differently at $t = 0$.

To be more mathematically rigorous, we define what we call the action of $\rho_\epsilon(t)$ on a class of test functions which consists of functions continuous at the origin and well behaved at infinity.

Definition 2.5.1 *The support of a function $f(t)$ is the closure of the set of all points where $f(t)$ is different from zero, that is,*

$$\sup f = \overline{\{t : f(t) \neq 0\}}$$

Definition 2.5.2 *The space $\mathcal{D}(\Re)$ of test functions consists of all real valued functions $\phi(t)$ which are infinitely times differentiable and which have compact support. This space is a vector space with respect to pointwise addition and scalar multiplication.*

Example 2.5.1

$$\phi(t) = \begin{cases} 0, & t \geq 1 \\ \exp\left(\dfrac{1}{t^2 - 1}\right), & |t| < 1 \end{cases} \tag{2.5.1}$$

$\phi(t)$ is infinite times differentiable and derivatives of all order vanish for $|t| \geq 1$.

Definition 2.5.3 *A sequence of function $\{\phi_n(t)\}$ in $\mathcal{D}(\Re)$ is said to converge to zero if*

(i) all $\phi_n(t)$ have the same compact support,

(ii) sequence $\{\phi_n(t)\}$ converges to zero uniformly and for any positive integer k, the derivatives $\{\phi_n^{(k)}(t)\}$ also converges uniformly to zero.

Example 2.5.2 *The sequence $\left\{\dfrac{\phi(t)}{n}\right\}_{n=1}^{\infty}$ with $\phi(t)$ defined by Eq. (2.5.1) converges to zero in $\mathcal{D}(\Re)$. Whereas, the sequence $\left\{\phi\left(\dfrac{t}{n}\right)\right\}$ does not converge to zero in $\mathcal{D}(\Re)$ as these functions do not have the same compact support.*

Definition 2.5.4 *A continuous linear function on the space $\mathcal{D}(\Re)$ is called a distribution. The space of all distributions is denoted by $\mathcal{D}'(\Re)$. The action of $f \in \mathcal{D}'(\Re)$ is denoted by (f, ϕ).*

Definition 2.5.5 *A sequence of functions $\{f_n\}_{n=1}^{\infty}$ in $\mathcal{D}'(\Re)$ is said to converge if for every $\phi \in \mathcal{D}(\Re)$ the sequence of scalars $\{(f_n, \phi)\}$ converges to the limit $\{(f, \phi)\}$. One can show that f is also a distribution.*

Example 2.5.3 *Define a functional δ on $\mathcal{D}(\Re)$ by $\langle \delta, \phi \rangle = \phi(0), \forall \phi \in \mathcal{D}(\Re)$*

It is easy to see that δ is linear and continuous and hence $\delta \in \mathcal{D}'(\Re)$. The distribution functional δ is defined as above is called the Dirac-delta function.

Example 2.5.4 *Dipole distribution $\tilde{\delta}$ is defined as $(\tilde{\delta}, \phi) = \phi'(0) \ \forall \ \phi \in \mathcal{D}(\Re)$*

Definition 2.5.6 *A locally integrable function f generates a distribution via the definition*

$$(f, \phi) = \int_{-\infty}^{\infty} f(t)\phi(t) \ \ \forall \ \phi \in \mathcal{D}(\Re)$$

This is called a regular distribution. A distribution which is not a regular distribution is called a singular distribution.

Theorem 2.5.1 *The Dirac-delta distribution is a singular distribution. For the proof of this fact refer Bose and Joshi [2]*

We can define translation, scale expansion, derivative of a distribution as follows.

Definition 2.5.7 *The translation $f(t-a)$ and scale expansion $f(at)$ of a distribution is given by*

$$
\begin{aligned}
(f(t-a), \phi(t)) &= (f(t), \phi(t+a)) \;\; \forall \;\; \phi \in \mathcal{D}(\Re) \\
(f(at), \phi(t)) &= \frac{1}{|a|}(f(t), \phi(t/a)), \;\; a \neq 0 \;\; \forall \;\; \phi \in \mathcal{D}(\Re)
\end{aligned}
$$

Example 2.5.5 *The translation $\delta(t-a)$ and scale expansion $\delta(-t)$ are given by*

$$(\delta(t-a), \phi(t)) = (\delta, \phi(t+a)) = \phi(a)$$

and

$$
\begin{aligned}
(\delta(-t), \phi(t)) &= (\delta, \phi(-1)) = \phi(0) \\
&= (\delta, \phi(t)) \quad \forall \, \phi \in \mathcal{D}(\Re)
\end{aligned}
$$

Definition 2.5.8 *For any distribution f, its derivative f' is defined as*

$$(f', \phi) = (f, -\phi') = -(f, \phi') \; \forall \; \phi \in \mathcal{D}(\Re)$$

By repeated application of the above definition, we get

$$(f^{(n)}, \phi) = (-1)^n (f, \phi^{(n)}) \; \forall \; \phi \in \mathcal{D}(\Re)$$

Here $f^{(n)} = \dfrac{d^n f}{dt^n}$.

Example 2.5.6 *Let $H(t)$ be the Heaviside function*

$$
H(t) = \begin{cases} 0, & t < 0 \\ 1/2, & t = 0 \\ 1, & t > 0 \end{cases}
$$

$H(t)$ generates a distribution given by

$$
\begin{aligned}
(H(t), \phi(t)) &= \int_{-\infty}^{\infty} H(t)\phi(t)dt \\
&= \int_{0}^{\infty} \phi(t)dt
\end{aligned}
$$

Its derivative $H'(t)$ is given by

$$
\begin{aligned}
(H'(t), \phi(t)) &= -(H, \phi') \\
&= -\int_{0}^{\infty} \phi'(t)dt = \phi(0) \\
&= (\delta, \phi(t))
\end{aligned}
$$

Hence we have $H' = \delta$.

Example 2.5.7 *Let $f(t)$ be differentiable except at points $a_1, \ldots, a_n$ where $f(t)$ jumps by the respective amount $\triangle f_1, \triangle f_2, \ldots, \triangle f_n$. We assume that f' is locally integrable wherever it exists. Then the distributional derivative f' is computed as follows*

$$\begin{aligned}(f', \phi) = -(f, \phi') &= -\int_{-\infty}^{\infty} f(t)\phi'(t)dt \\ &= -\int_{-\infty}^{\infty} f(t)\phi'(t)dt - \int_{a_1}^{a_2} f(t)\phi'(t)dt \ldots \\ &\quad -\int_{a_n}^{\infty} f(t)\phi'(t)dt\end{aligned}$$

Integrating by parts and noting that $f(a_k^+) - f(a_k^-) = \triangle f_k$, we get

$$(f', \phi) = \int_{-\infty}^{\infty} \frac{df}{dt}\phi(t)dt + \int_{a_1}^{a_2} \frac{df}{dt}\phi(t)dt + \ldots$$
$$+ \int_{a_n}^{\infty} \frac{df}{dt}\phi(t)dt + \sum_{k=1}^{n} \triangle f_k \phi(a_k)$$

Therefore

$$(f', \phi) = \int_{-\infty}^{\infty} [f']\phi(t)dt + \sum_{k=1}^{n} \triangle f_k \phi(a_k)$$

This gives $f' = [f'] + \sum_{k=1}^{n} \triangle f_k \delta(t - a_k)$.

Here $[f']$ indicates the usual derivative of f, wherever it exists.

Example 2.5.8 *If $g(t)$ is infinitely differentiable, then $g(t)f^{(n)}(t)$ can be found as follows*

$$\begin{aligned}(gf^{(n)}(t), \phi) &= (f^{(n)}, g\phi) \\ &= (-1)^n (f, \frac{d^n}{dt^n}(g\phi))\end{aligned}$$

Example 2.5.9 *Let L be a linear differential operator given by*

$$\begin{aligned}L &= a_0(t)\frac{d^n}{dt^n} + a_1(t)\frac{d^{n-1}}{dt^{n-1}} + \ldots + a_{n-1}(t)\frac{d}{dt} + a_n \\ &= \sum_{k=0}^{n} a_{n-k}(t)\frac{d^k}{dt^k}\end{aligned}$$

with coefficients $a_k(t)$ infinite times differentiable. Then for the distribution f, we have

$$\begin{aligned}(Lf, \phi) &= (\sum_{k=0}^{n} a_{n-k}(t)f^{(k)}, \phi) \\ &= (-1)^k \left(f, \frac{d^k}{dt^k}\left(a_{n-k}(t)\phi(t)\right)\right) \\ &= (f, L^*\phi)\end{aligned}$$

where $L^*\phi = \sum_{k=0}^{n}(-1)^k \frac{d^k}{dt^k}(a_{n-k}(t)\phi(t))$.

The n^{th} order differential operator L^* is known as the formal adjoint of L.

Example 2.5.10 *Consider the sequence* $\{\sin nt\}_{n=1}^{\infty}$. *As it is locally integrable, it generates a sequence of distributions. This sequence does not converge point wise except at* $t = 0$. *However, we have*

$$\begin{aligned}
|(\sin nt, \phi(t))| &= |\int_{-\infty}^{\infty} \sin nt\phi(t)dt| \\
&\leq \frac{1}{n}\int_{-\infty}^{\infty} |\cos nt\phi'(t)|dt \\
&\leq \frac{1}{n}\int_{-\infty}^{\infty} |\phi'(t)|dt \\
&\to 0 \text{ as } n \to \infty
\end{aligned}$$

So $\{\sin nt\}$ *converges in* $\mathcal{D}'(\Re)$ *to zero distribution.*

Example 2.5.11 *Let* $f_n(t) = \begin{cases} 0, & |t| \geq \frac{1}{n} \\ \frac{n}{2}, & |t| < \frac{1}{n} \end{cases}$

$$\begin{aligned}
(f_n, \phi) &= \int_{-\infty}^{\infty} f_n(t)\phi(t)dt = \int_{-\frac{1}{n}}^{\frac{1}{n}} \frac{n}{2}\phi(t)dt \\
&= \frac{n}{2}\int_{-\frac{1}{n}}^{\frac{1}{n}} \phi(t)dt
\end{aligned}$$

This implies that

$$\begin{aligned}
\lim_{n\to\infty}(f_n, \phi) &= \lim_{n\to\infty}\frac{n}{2}\int_{-1/n}^{1/n} \phi(t)dt \\
&= \phi(0) = (\delta, \phi)
\end{aligned}$$

Thus the sequence of regular distribution $\{f_n\}$ *converges in* $\mathcal{D}'(\Re)$ *to the delta distribution whereas the corresponding functions converge almost everywhere to zero.*

2.6 Exercises

1. Test the convergence of the following sequences of the functions in the space $L_2[0, 1]$. Compute the limit function, if it exists.

   (a) $x_n(t) = \sin n\pi t$

(b) $x_n(t) = t^n$

(c) $x_n(t) = t \sin n\pi t$

(d) $x_n(t) = \dfrac{\sin n\pi t}{n}$

2. Let $A \in \Re^{n\times n}$ and let $N(A)$ and $R(A)$ denote the null space and range space of A, respectively. Show that

$$\Re^n = N(A) \oplus R(A)$$

3. Let $M \subset L_2\,[-\pi, \pi]$ be the linear span of $\{1,\ t^2,\ \cos t\}$. Find an element in M closest to $\cos 2t \in L_2[-\pi,\ \pi]$.

4. Show that

$$[Lx](t) = \int_0^t x(s)ds$$

is a bounded linear operator in $L_2[0, 1]$. Also, find its adjoint L^* and show that L is a singular operator.

5. Let $L : L_2[0, 1] \to L_2[0, 1]$ be defined by

$$[Lx](t) = \int_0^1 (s + \sin t)^2 x(s)ds$$

From the first principle, show that L is compact.

6. Linearize the following nonlinear operators around the given points.

(a) $F(x_1, x_2) = \big(exp(x_1x_2) + \cos x_1 \sin x_2,\ x_2^2 x_1 exp(x_1, x_2)\big)$
$(x_1, x_2) = (0, 0)$

(b) $G(x_1, x_2) = (x_1 + \sin x_1 \cos x_2,\ \cos x_1 \sin x_2 + x_1x_2)$
$(x_1, x_2) = (0, 0)$

7. Determine if the function $F : \Re^3 \to \Re^3$ defined as

$$F(x_1, x_2, x_3) = \frac{1}{10}\left(\sin^2 x_1 + \cos x_2 \cos x_3, \sin x_2 \cos x_3, \cos x_1 \cos x_2 \cos x_3\right)$$

is a contraction in $\Re^3$ w.r.t 1-norm in $\Re^3$.

8. Let $f : X \to \Re$ (X, a Hilbert space) be differentiable. Show that f is convex iff $\bar{x} \to \bigtriangledown f(\bar{x})$ is monotone for all $\bar{x} \in X$.

9. Show that

$$\Psi(t) = \begin{cases} exp(-(t^{-2}))exp(-(t-a)^{-2}), & 0 \le t \le a \\ 0, & t \le 0, \text{ or } t \ge a \end{cases}$$

is a test function.

10. Show that

    (a) $g(t)\delta'(t) = g(0)\delta'(t) - g'(t)\delta(t), \quad g \in C^\infty$

    (b) $\delta'(b-a) = \int \delta'(b-a)\delta(b-a)dt$

    (c) $\frac{d}{dt}(sgnt) = 2\delta(t)$

References

[1] R B Bapat, *Linear Algebra and Linear Models, Hindustan Book Agency, 1993.*

[2] R. K. Bose and M. C. Joshi, *Methods of Mathematical Physics, Tata McGraw Hill Publishing Co. Ltd., 1984.*

[3] F. E. Browder, *The Solvability of Nonlinear Functional Equations, Duke-Math Journal 30, 557-566, 1963.*

[4] S. D. Conte and C. deBoor, *Elemenatry Numerical Analysis - An Algorithmic Approach, McGraw Hill Book Co., 1981.*

[5] Casper Goffman and George Pedrick, *First Course in Functional Analysis, Prentice Hall of India, 1974.*

[6] Mohan C Joshi and Ramendra K Bose, *Some Topics in Nonlinear Functional Analysis, Wiley Eastern Ltd, 1985*

[7] M. A. Krasnoselskii, *Topological Methods in the Theory of Nonlinear Integral Equation, Pergaman Press, 1964.*

[8] Balmohan V Limaye, *Functional Analysis, New Age International, 1996.*

[9] G. Minty, *Monotone (Nonlinear) Operator in a Hilbert Space, Duke Math Journal, 29, 341-346, 1962.*

[10] H. L. Royden, *Real Analysis, Macmillan Publications, 1988.*

Chapter 3

Initial Value Problems

This chapter primarily deals with the existence and uniqueness theory of initial value problems. The notions of - Picard iterates, maximal interval of existence, continuation of solution and continuity of solution with respect to initial condition - are the prime targets of our investigation.

3.1 Existence and Uniqueness Theorems

Let I be an interval in $\Re$ and Ω an open set containing $(t_0, \overline{x}_0) \in I \times \Omega$. Let $f : I \times \Omega \mapsto \Re^n$ be a function, not necessarily linear. As defined earlier (Definition 1.1.1), by Initial Value Problem (IVP), we mean the following differential equation with initial condition

$$\frac{d\overline{x}}{dt} = f(t, \overline{x}(t)) \tag{3.1.1a}$$

$$\overline{x}(t_0) = \overline{x}_0 \tag{3.1.1b}$$

Also, let us recall what we mean by the solution of the IVP - Eq. (3.1.1). $\overline{x}(t)$ is a solution of Eq. (3.1.1), if $\exists$ an interval $J \subset I$, containing 't_0' such that $\overline{x}(t)$ is differentiable on J with $\overline{x}(t) \in \Omega$, $\forall\, t \in J$ and $\overline{x}(t)$ satisfies Eq. (3.1.1).
The following theorem gives the existence of unique solution of Eq. (3.1.1), requiring Lipschitz continuity on $f(t, \overline{x})$ with respect to $\overline{x}$.

Theorem 3.1.1 *Let $E = I \times \Omega \subseteq \Re \times \Re^n$ be an open set containing $(t_0, \overline{x}_0)$ and let $f : E \mapsto \Re^n$ be continuous. Further assume that $\overline{x} \mapsto f(t, \overline{x})$ is Lipschitz continuous :*
$\exists$ a constant M such that

$$||f(t, \overline{x}) - f(t, \overline{y})|| \le M||\overline{x} - \overline{y}|| \ \ \forall\, t \in I \text{ and } \forall\, \overline{x},\ \overline{y} \in \Omega \tag{3.1.2}$$

Then the IVP given by Eq. (3.1.1) has a unique solution.

Proof : As shown in Proposition 1.1.1, solvability of Eq. (3.1.1) is equivalent to the solvability of the following integral equation

$$\overline{x}(t) = \overline{x}_0 + \int_{t_0}^{t} f(s, \overline{x}(s))ds \tag{3.1.3}$$

in some appropriate subinterval $J \subset I$.
Let B denote a closed ball with centre at $(t_0, \bar{x}_0)$ and contained in E. Let $m = \sup_{(t,\overline{x}) \in B} |f(t, \bar{x})|$. '$m$' is finite as f is continuous on a compact set $B \subseteq I \times \Re^n$.
Choose two positive numbers 'δ' and 'τ' such that $M\tau < 1$, $m\tau < \delta$ and

$$\{ \ (t, \bar{x}) \in I \times \Re^n : |t - t_0| \leq \tau, \ ||\overline{x} - \overline{x_0}|| \leq \delta \ \} \subseteq B.$$

Denote by $J = [t_0 - \tau, t_0 + \tau]$ and $\Omega_0 = \{ \ \overline{x} \in \Omega : ||\overline{x} - \overline{x}_0|| \leq \delta \ \}$. Define $X \subset C(J)$ to be set of all continuous functions on J with values in $\Omega_0 \subseteq \Omega$. Then X is a closed subset of $C(J)$ and hence a complete metric space with metric induced by the sup-norm of $C(J)$. We define a nonlinear operator F on X as follows

$$[F\overline{x}](t) = \overline{x}_0 + \int_{t_0}^{t} f(s, \overline{x}(s))ds, \ t \in J \tag{3.1.4}$$

It is clear that $F\overline{x}$ is a continuous function defined on J. Also,

$$\begin{aligned}
\|[F\overline{x}](t) - \overline{x}_0\|_{\Re^n} &= \| \int_{t_0}^{t} f(s, \overline{x}(s))ds \|_{\Re^n} \\
&\leq \int_{t_0}^{t} \|f(s, \overline{x}(s))\|_{\Re^n} ds \\
&\leq m \int_{t_0}^{t} ds \\
&\leq m\tau \\
&\leq \delta
\end{aligned}$$

This implies that $[F\overline{x}](t) \in \Omega_0 \ \forall \, t \in J$ and hence F maps X into itself.
Further, $\forall \bar{x}, \ \bar{y} \in X$, we have,

$$\begin{aligned}
||F\overline{x}(t) - F\overline{y}(t)||_{\Re^n} &\leq \int_{t_0}^{t} ||f(s, \overline{x}(s)) - f(s, \overline{y}(s))||_{\Re^n} ds \\
&\leq M \int_{t_0}^{t} ||\overline{x}(s) - \overline{y}(s)||_{\Re^n} ds \\
&\leq M \ ||\overline{x} - \overline{y}|| \, (t - t_0) \\
&\leq \tau M \ ||\overline{x} - \overline{y}||
\end{aligned}$$

and hence

$$||F\overline{x} - F\overline{y}|| \leq \tau M ||\overline{x} - \overline{y}|| \qquad \forall \, \overline{x}, \ \overline{y} \in X$$

As $\alpha = \tau M < 1$, it follows that F is a contraction on the complete metric space X. Invoking Theorem 2.4.1, it follows that F has a unique fixed point $\overline{x}^* \in X$. That is, $\exists\, \bar{x}^* \in X = C(J, \Omega_0)$ such that

$$\overline{x}^*(t) = F\overline{x}^*(t) = \overline{x}_0 + \int_{t_0}^{t} f(s, \overline{x}^*(s))ds$$

and $\overline{x}^*$ is unique.
This proves that $\overline{x}^*(t)$ is unique solution of Eq. (3.1.3) and hence that of Eq. (3.1.1). ∎

Corollary 3.1.1 *The iterations $\overline{x}^{(k)}(t)$, defined as*

$$\overline{x}^{(k)}(t) = \overline{x}_0 + \int_{t_0}^{t} f(s, \overline{x}^{(k-1)}(s))ds \tag{3.1.5}$$

with $\overline{x}^{(0)}(t) = \overline{x}_0$, converge to the unique solution $\overline{x}^*(t)$ of IVP given by Eq. (3.1.1).
The iterates defined by Eq. (3.1.5) are called Picard iterates.

Proof : This follows from Theorem 2.4.1, wherein the converging iterates are defined as

$$\overline{x}^{(k)} = F\overline{x}^{(k-1)},\ \overline{x}^{(0)} \in X, \ \ \text{(arbitrary)}$$

Here F is defined by (3.1.4). We note that $\overline{x}^{(0)}(t) = \overline{x}_0 \in X = C(J, \Omega_0)$. ∎

We now consider an n^{th} order differential equation

$$\frac{d^n x}{dt^n} + a_{n-1}(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1(t)\frac{dx}{dt} = f(t, x(t)) \tag{3.1.6a}$$

with initial values

$$x(0) = x_1^{(0)},\ x^{(1)}(0) = x_2^{(0)}, \cdots, x^{(n-1)}(0) = x_n^{(0)} \tag{3.1.6b}$$

Here $x^{(k)}(t)$ refer to the k^{th} derivative of $x(t)$.

Corollary 3.1.2 *Let $\Omega \subseteq \Re$ be open and $f : I \times \Omega \mapsto \Re$ be continuous with $x \mapsto f(\ .\ ,\ x\ )$ Lipschitz continuous with Lipschitz constant M. Further, assume that $a_i(t),\ 1 \le i \le n-1$ are continuous functions on I. Then n^{th} order initial value problem given by Eq. (3.1.6) has a unique solution.*

Proof : We shall reduce Eq. (3.1.6) in to an initial value problem in $\Re^n$. For this, set

$$x_1 = x,\ x_2 = \dot{x}_1 = \dot{x},\ x_3 = \dot{x}_2 = \ddot{x}_1 = \ddot{x},\ \cdots,\ \ x_n = \dot{x}_{n-1}$$

Proceeding as in section (1.3), the IVP given by Eq. (3.1.6) reduces to the following IVP on $I \times \Omega \times \Omega \cdots \Omega \subseteq I \times \Re^n$

$$\frac{d\overline{x}}{dt} = F(t, \overline{x}(t)) \tag{3.1.7a}$$

$$\overline{x}(0) = \overline{x}_0 \tag{3.1.7b}$$

Here $\bar{x} = (x_1, x_2, \cdots, x_n),\ \bar{x}_0 = \left(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}\right) \in \Re^n$ and
$F(t, \overline{x}) = (x_2, x_3, \ldots, x_n, f(t, x_1) - a_1(t)x_2 - \cdots - a_{n-1}(t)x_n)$. It is clear that F is continuous on $I \times \Omega^n$ and further $\overline{x} \mapsto F(., \overline{x})$ is Lipschitz continuous as we see below.

$$\begin{aligned}
||F(t,\overline{x}) - F(t,\overline{y})||^2_{\Re^n} &\leq |x_2 - y_2|^2 + |x_3 - y_3|^2 + \cdots + |x_n - y_n|^2 \\
&+ n\left[|f(t,x_1) - f(t,y_1)|^2 + |a_1(t)|^2|x_2 - y_2|^2 \right.\\
&\quad \left. + \cdots + |a_{n-1}|^2|x_n - y_n|^2\right] \\
&\leq nM^2|x_1 - y_1|^2 + (1 + n|a_1(t)|^2)|x_2 - y_2|^2 \\
&\quad + \cdots + \left(1 + n|a_{n-1}(t)|^2\right)|x_n - y_n|^2 \\
&\leq \alpha^2\left[|x_1 - y_1|^2 + |x_2 - y_2|^2 + \cdots + |x_n - y_n|^2\right].
\end{aligned}$$

where $\alpha^2 = max\left(nM^2, 1 + na_1^2,\ 1 + na_2^2,\ \cdots, 1 + na_{n-1}^2\right)$,
with $a_i = \sup_{t \in I} |a_i(t)|, \quad 1 \leq i \leq n.$
Hence we have,

$$||F(t,\overline{x}) - F(t,\overline{y})||^2_{\Re^n} \leq \alpha^2 ||\overline{x} - \overline{y}||^2_{\Re^n}$$

Thus the function F appearing in the IVP given by Eq. (3.1.7), satisfies all conditions of Theorem 3.1.1 and hence it has a unique solution. This in turn implies that the n^{th} order differential equation given by Eq. (3.1.6) has a unique solution. ∎

Example 3.1.1

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

Clearly, the function $f(x) = x^2$ is Lipschitz continuous on a closed, bounded subset of $\Re$. Hence by Theorem 3.1.1 this IVP has a unique solution. The Picard's iterates $x^{(k)}(t)$ are given by

$$\begin{aligned}
x^{(k)}(t) &= x_0 + \int_0^t f(s, x^{(k-1)}(s))ds \\
&= 1 + \int_0^t \left[x^{(k-1)}(s)\right]^2 ds
\end{aligned}$$

So, we get the following Picard iterates for the above problem.

$$\begin{aligned}
x^{(1)}(t) &= 1 + t \\
x^{(2)}(t) &= (1 + t + t^2) + \frac{t^3}{3} \\
&\vdots \\
x^{(n)}(t) &= (1 + t + t^2 + \cdots + t^k) + o(t^{k+1})
\end{aligned}$$

We can actually compute the solution of the above IVP by the method of separation of variables and obtain $x(t) = \frac{1}{1-t}$*. The largest interval on which this solution is defined in* $(-\infty, 1)$*. For small* $'t'$*, the function* $x(t)$ *has the series expansion given by*

$$x(t) = 1 + t^2 + t^3 + \cdots + t^n + \cdots .$$

(refer Chapter 6).
Thus the Picard iterates $x^{(k)}(t)$ *do converge to the actual solution* $x(t) = 1 + t^2 + t^3 + \cdots$ *in a small interval* $|t| < 1$ *around* $t = 0$*, as envisaged in the Corollary 3.1.1.*
The following graph gives first three Picard iterates.
The actual solution $x(t) = \frac{1}{1-t}$ *is plotted in Figure 3.1.2.*

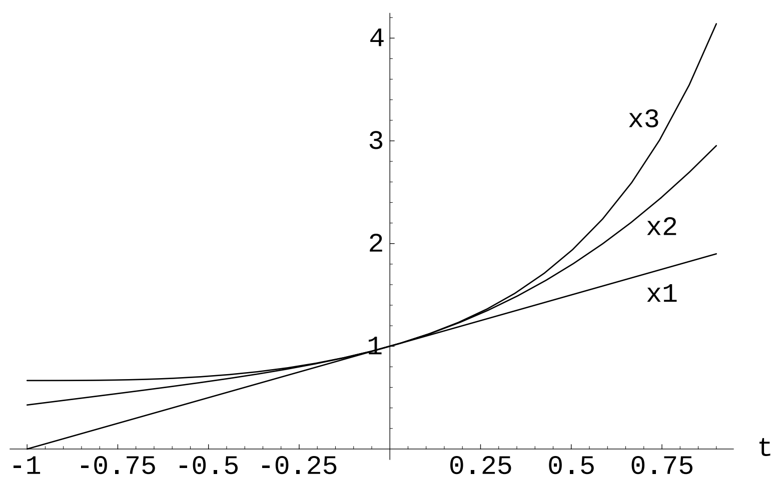

Figure 3.1.1: First three Picard iterates

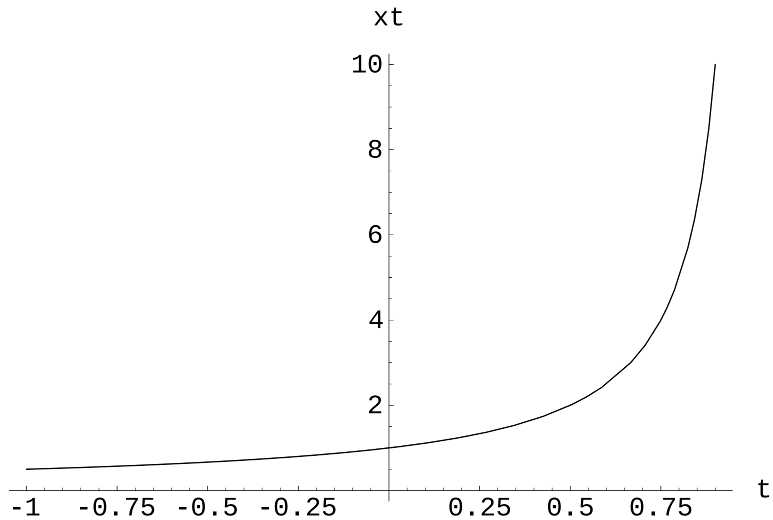

Figure 3.1.2: Actual solution $x(t) = \dfrac{1}{1-t}$

Example 3.1.2

$$\frac{dx}{dt} = f(x), \quad x(0) = 0$$

where $f(x)$ is given by

$$f(x) = \begin{cases} 2\sqrt{x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

As $\dfrac{f(x)}{x} = \dfrac{2}{\sqrt{x}}, \quad x > 0, \dfrac{f(x)}{x} \to \infty$ as $x \to 0_+$, and hence f is <u>*not*</u> Lipschitz continuous at '0'. Hence Theorem 3.1.1 is not applicable. However, by inspection one can easily compute the following solutions of above IVP :

$$x_1(t) = \begin{cases} t^2, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

$$x_2(t) = 0 \qquad \forall\, t \in \Re$$

That is, the above IVP has at least two distinct solutions or in other words, the solution is not unique.

Remark 3.1.1 *This example concludes that there may exist a solution of IVP of the type Eq. (3.1.1) without Lipschitz continuity assumptions on the nonlinear function $f(t, \bar{x})$. However it may not be unique. This point will discussed in detail in section 3.4.*

Example 3.1.3

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

$$f(t,x) = \begin{cases} \dfrac{2tx}{t^2+x^2} & (t,x) \neq (0,0) \\ 0 & (t,x) = (0,0) \end{cases}$$

The solution of the above differential equation passing through $(0,0)$ *is given by the following curve in* (t,x)*-plane*

$$(x-c)^2 - t^2 = c^2, c \quad \textit{is arbitrary constant} \tag{3.1.8}$$

The solution curve is plotted as under.

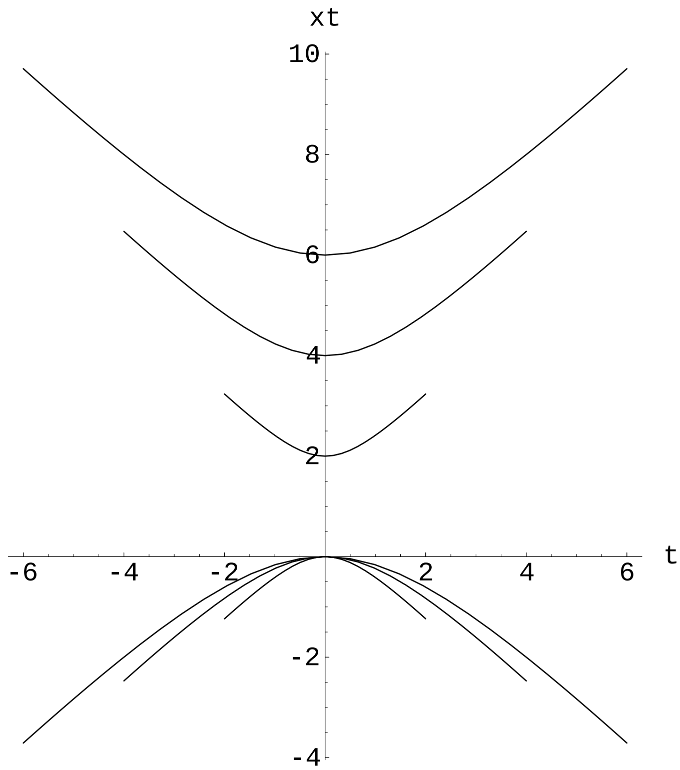

Figure 3.1.3: Solution curve for the above example

It may noted that the above IVP has a unique solution passing through the point away form $(0,0)$. It can be easily seen that $(t,x) \mapsto f(t,x)$ is not even continuous at $(0,0)$. Yet, there are infinitely many solutions passing through $(0,0)$, of the above differential equation.

Example 3.1.4

$$\frac{d^2x}{dt^2} = \cos x, \quad x(0) = 0, \quad \dot{x}(0) = 0$$

As $f(x) = \cos x$ *is Lipschitz continuous, it follows by Corollary 3.1.2 that the above second order initial value problem has a unique solution. However, a close form solution of the above problem is difficult to obtain. One can, therefore, resorts to Picard iterative procedure to obtain iterates* $x^{(k)}(t)$ *converging to the unique solution of the above second order IVP.*

$$x_1(t) = x(t),\ x_2(t) = \dot{x};\ \bar{x}(t) = (x_1(t), x_2(t))$$

$$\frac{d\bar{x}}{dt} = F(\bar{x}(t)), F(\overline{x}) = (x_2, \cos x_1), \quad \overline{x}(0) = \bar{0}$$

We have

$$\begin{aligned}
x_1^{(1)}(t) &= \int_0^t x_2^{(0)}(s)ds = 0 \\
x_2^{(1)}(t) &= \int_0^t \cos x_1^{(0)}(s)ds = \int_0^t ds = t \\
x_1^{(2)}(t) &= \int_0^t x_2^{(1)}(s)ds = \int_0^t sds = \frac{t^2}{2} \\
x_2^{(2)}(t) &= \int_0^t \cos x_1^{(1)}(s)ds = \int_0^t ds = t \\
x_1^{(3)}(t) &= \int_0^t x_2^{(2)}(s)ds = \int_0^t sds = \frac{t^2}{2} \\
x_2^{(3)}(t) &= \int_0^t \cos x_1^{(2)}(s)ds = \int_0^t \cos \frac{s^2}{2} ds
\end{aligned}$$

One can approximate $\cos\left(\frac{s^2}{2}\right)$ *by series expansion and obtain* $x_2^{(3)}(t)$ *and* $\left(x_1^{(4)}(t), x_2^{(4)}(t)\right)$*, up to some accuracy.*

$$\begin{aligned}
x_1^{(4)}(t) &= \frac{t^2}{2} - \frac{t^6}{240} + \frac{t^{10}}{34560} \\
x_2^{(4)}(t) &= t - \frac{t^5}{40} + \frac{t^9}{3456}
\end{aligned}$$

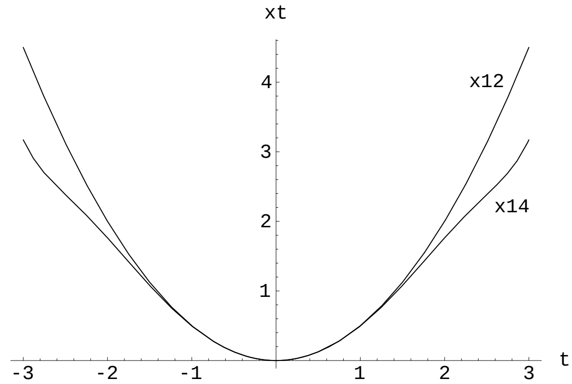

Figure 3.1.4: Graph of iterates $x_1^{(2)}(t), x_1^{(4)}(t)$

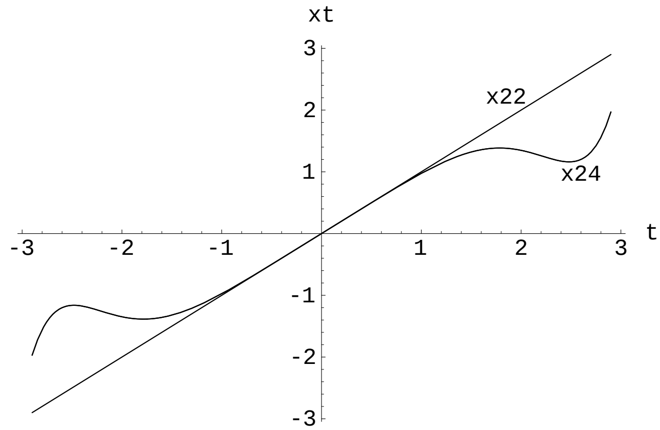

Figure 3.1.5: Graph of iterates $x_2^{(2)}(t), x_2^{(4)}(t)$

There do exist examples where the nonlinear function $f(t,\bar{x})$ under consideration is not Lipschitz continuous but satisfies other nice property like monotonicity with respect to the variable $\bar{x}$. We shall now state and prove the existence and uniqness theorem for IVP of the type given by Eq. (3.1.1) under the monotonicity assumption.
Using the notion of monotonicity of Chapter 2 (Definition 2.4.3), we shall say that $\bar{x} \mapsto f(t,\bar{x})$ is monotone in $\Re^n$ if

$$(f(t,\bar{x}_1) - f(t,\bar{x}_2), \overline{x}_1 - \bar{x}_2)_{\Re^n} \geq 0 \quad \forall \ \bar{x}_1, \bar{x}_2 \in \Re^n \tag{3.1.9}$$

It is said to be strong monotone if $\exists \, c > 0$ such that

$$(f(t,\bar{x}_1) - f(t,\bar{x}_2), \overline{x}_1 - \bar{x}_2)_{\Re^n} \geq c \quad \|\bar{x}_1 - \bar{x}_2\|^2_{\Re^n} \ \ \forall \ \bar{x}_1, \bar{x_2} \in \Re^n \tag{3.1.10}$$

In the IVP given by Eq. (3.1.1), we assume that $t_0 = 0$ and $\bar{x}_0 = 0$.

Definition 3.1.1 *A function $\bar{x}(t) \in L_2[0, t_f]$ is said to satisfy Eq. (3.1.1) in almost everywhere (a.e) sense if $\bar{x}(t)$ is differentiable for all $t \in [0, t_f] = I$ and satisfies Eq. (3.1.1) for all $t \in I$ except on a set of measure zero in I.*

Theorem 3.1.2 *Let $f(t, \overline{x}) : [0, t_f] \times \Re^n \mapsto \Re^n$ be continuous and $\overline{x} \mapsto f(t, \bar{x})$ be strongly monotone and further it satisfies a growth condition of the form*

$$\|f(t, \overline{x})\|_{\Re^n} \leq a(t) + b\|x\|_{\Re^n}, a \in L_2[0, t_f], \quad b > 0 \tag{3.1.11}$$

Then there exist unique $\bar{x}(t) \in L_2[0, t_f]$ satisfying Eq. (3.1.1) in almost everywhere sense.

Proof : We are interested in the solvability of the IVP of the form

$$\begin{aligned} \frac{d\overline{x}}{dt} + f(t, \overline{x}) &= 0 \\ \overline{x}(0) &= \overline{0} \end{aligned} \tag{3.1.12}$$

in the interval $I = [0, t_f]$, in almost every where sense. Set $X = L_2[0, t_f]$. Here $L_2[0, t_f]$ is defined to be the set of all square integrable functions with values in $\Re^n$. Define L as $L\bar{x} = \frac{d\bar{x}}{dt}$ with $D(L) = \{\overline{x} \in X : \frac{d\bar{x}}{dt} \in X \text{ and } \overline{x}(0) = 0\}$. Then $D(L)$ is dense in X and L is monotone as

$$\begin{aligned} (L\bar{x}, \bar{x}) = (L\overline{x}, \bar{x})_{L_2[0,t_f]} &= \int_0^{t_f} \left(\frac{d\bar{x}}{dt}, \overline{x}(t)\right)_{\Re^n} dt \\ &= ||\overline{x}(t_f)||^2_{\Re^n} - \int_0^{t_f} \left(\overline{x}(t), \frac{d\bar{x}}{dt}\right)_{\Re^n} dt \\ &= ||\bar{x}(t_f)||^2_{\Re^n} - (L\bar{x}, \bar{x}) \end{aligned}$$

This implies that

$$(L\bar{x}, \bar{x}) = \frac{\|\overline{x}(t_f)\|}{2} \geq 0 \;\; \forall \;\; \overline{x} \in L_2[0, t_f]$$

Further, one can show that L is maximal monotone as $R(L + \lambda I) = X$ for $\lambda > 0$ (refer Pazy[6]).

We now define the nonlinear operator N on X as in Example 2.4.4:

$$[N\overline{x}] = f(t, \bar{x}(t)) \tag{3.1.13}$$

N is a bounded, continuous operator on $X = L_2[0, t_f]$ (refer Joshi and Bose

[4]). Further, it is strongly monotone as we have

$$\begin{aligned}(N\bar{x}_1 - N\bar{x}_2, \bar{x}_1 - \bar{x}_2) &= \int_0^{t_f} (f(t, \bar{x}_1(t)) - f(t, \bar{x}_2(t)),\ (\bar{x}_1(t) - \bar{x}_2(t)))\, dt \\ &\geq c\int_0^{t_f} ||\overline{x}_1(t) - \overline{x}_2(t)||^2 dt \\ &= c||\overline{x}_1 - \overline{x}_2||^2\end{aligned}$$

for all $\bar{x_1}, \bar{x_2} \in X$.
Thus the solvability of equation Eq. (3.1.12) reduces to the solvability of the operator equation

$$L\overline{x} + N\overline{x} = 0 \tag{3.1.14}$$

in the space $X = L_2[0, t_f]$. $L : D(L) \to X$ is a linear maximal monotone operator with dense domain. N is a bounded, continuous strongly monotone operator defined on the whole space X. Hence by Theorem 2.4.5, it follows that Eq. (3.1.14) has a unique solution and hence the unique solvability of Eq. (3.1.12), in almost everywhere sense in X. ▮

Example 3.1.5

$$\begin{aligned}\frac{dx}{dt} + \alpha x + \cos x &= 0 \\ x(0) &= 0\end{aligned}$$

$f(x) = \alpha x + \cos x$ *is strongly monotone if* $\alpha > 1$.
Hence it is follows by Theorem 3.1.2 that the above initial value problem has a unique solution $x(t)$*, (a.e. for* t*) in a finite subinterval of* $[0, \infty)$*. However, Theorem 3.1.1 is not applicable.*

3.2 The Maximal Interval of Existence

Theorem 3.1.1 and Theorem 3.1.2 establish the fact that under suitable continuity and Lipschitz continuity amd monotoncity assumptions on $f(t, x)$, the IVP of the form Eq. (3.1.1) has a unique solution $x(t)$ on some interval surrounding t_0. In this section we shall show that the solution $x(t)$ is defined on the maximal interval of existence (α, β).
Assumption 1: Let $\Omega \subseteq \Re^n$ be an open set containing $\overline{x}_0$. Let $f : \Re \times \Omega \to \Re^n$ satisfies the assumptions of Theorem 3.1.1

Theorem 3.2.1 *Let* f *satisfies Assumption 1. If* $\overline{x}(t)$*,* $\overline{y}(t)$ *are solutions of the IVP given by Eq. (3.1.1) on the intervals* I_1*,* I_2*. Then* $t_0 \in I_1 \cap I_2$ *and if* I *is any open interval around '*t_0*' contained in* $I_1 \cap I_2$*, then* $\overline{x}(t) = \overline{y}(t) \quad \forall\, t \in I$.

Proof : It is clear that $t_0 \in I_1 \cap I_2$. If $t_0 \in I \subset I_1 \cap I_2$, then by Theorem 3.1.1, $\exists$ some open interval $I_a = (a - t_0, a + t_0) \subset I$ such that Eq. (3.1.1) has a unique solution. In view of the uniqueness of solutions on I_a, $\overline{x}(t) = \overline{y}(t)$ on $I_a \subset I \subset I_1 \cap I_2$.

Define $I^* = \cup I_a$. Then I^* is the largest open interval contained in I on which $\overline{x}(t) = \overline{y}(t)$. If I^* is properly contained in I, then one of the end points t^* of I^* is contained in $I \subset I_1 \cap I_2$. As both $\overline{x}(t)$ and $\overline{y}(t)$ are continuous on I, it follows that,

$$\lim_{t \to t^*} \overline{x}(t) = \lim_{t \to t^*} \overline{y}(t) = \bar{x}^*$$

Now examine the initial value problem

$$\frac{d\overline{x}}{dt} = f(t, \overline{x}(t)), \ \ \overline{x}(t^*) = \bar{x}^* \tag{3.2.1}$$

Eq. (3.2.1) has a unique solution on some interval $I_0 = (t^* - a^*, \ t^* + a^*) \subset I$. Again by uniqueness on $I_0 \subset I \subset I_1 \cap I_2$, it follows that, $\overline{x}(t) = \overline{y}(t)$ for $t \in I^* \cup I_0$ and I^* is proper subinterval of $I^* \cup I_0$ and hence a contradiction of the fact that I^* is the largest open interval contained in I on which $\overline{x}(t) = \overline{y}(t)$. Hence $I^* = I$ and thereby implying that $\overline{x}(t) = \overline{y}(t) \ \forall \ t \in I$. ∎

Theorem 3.2.2 *Let f satisfies Assumption 1. Then for each $\overline{x}_0 \in \Omega$, $\exists$ a maximal interval I^* on which the IVP given by Eq. (3.1.1) has a unique solution. Furthermore, the maximal interval J is open, that is $I^* = (\alpha, \ \beta)$.*

Proof : As in Theorem 3.2.1, we define $I^* = \cup I_a$ where I_a is an open interval surrounding t_0, on which Eq. (3.1.1) has a unique solution $\overline{x}_a$. Define $\overline{x} : I^* \mapsto \Omega$ as follows,

$$\overline{x}(t) = \overline{x}_a(t), \qquad t \in I_a$$

This is well defined. for, if $t \in I_1 \cap I_2$, where I_1 and I_2 are two intervals on which Eq. (3.1.1) has solutions $\overline{x}_1(t)$ and $\overline{x}_2(t)$. Then by Theorem 3.2.1 it follows that $\overline{x}_1(t) = \overline{x}_2(t)$. Also $\overline{x}(t)$ is a unique solution of Eq. (3.1.1) as $\overline{x}_a(t)$ is a unique solution of Eq. (3.1.1) for $t \in I_a$.

By definition, I^* is maximal. Also let the maximal interval of existence be not open say of the form $(\alpha, \ \beta]$ then we can extend the solution of Eq. (3.1.1) on $(\alpha, \ \beta + a]$ with $a > 0$ as in the proof of Theorem 3.2.1. ∎

If $(\alpha, \ \beta)$ is the maximal interval of existence for IVP given by Eq. (3.1.1), then $t_0 \in (\alpha, \ \beta)$. The intervals $[t_0, \ \beta)$ and $(\alpha, t_0]$ are called the right and the left maximal intervals of existence respectively. Without loss of generality we can concentrate on the right maximal interval of existence. We wish to find out if the right maximum interval of existence $[t_0, \ \beta)$ can be extended to $[t_0, \infty)$. This will then lead to global solvability of IVP given by Eq. (3.1.1). The following theorem is an important one in that direction.

Theorem 3.2.3 *Let Ω and f be as defined in Theorem 3.2.2 with $\overline{x}_0 \in \Omega$. Let $(\alpha,\ \beta)$ be the maximal interval of existence of the solution $\overline{x}(t)$ of the IVP given by Eq. (3.1.1). Assume that $\beta < \infty$. Then given any compact set $K \subset \Omega$, $\exists\, t \in (\alpha,\ \beta)$ such that $\overline{x}(t) \notin K$.*

Proof : Let K be compact set in $\Re^n$. Since f is continuous on $(\alpha,\ \beta) \times K$. It follows that

$$m = \sup_{(t,\overline{x}) \in (\alpha,\ \beta)\times K} \|f(t,\overline{x})\| < \infty$$

If possible, let the solution $\overline{x}(t)$ of Eq. (3.1.1) be such that $\overline{x}(t) \in K \ \ \forall\, t \in (\alpha,\ \beta)$ with $\beta < \infty$. For $\alpha < t_1 < t_2 < \beta$, we have

$$\begin{aligned} \|\overline{x}(t_1) - \overline{x}(t_2)\| &\le \int_{t_1}^{t_2} \|f(s,\overline{x}(s))\| ds \\ &\le m\|t_2 - t_1\| \end{aligned}$$

As $t_1,\ t_2 \to \beta$, it follows that $\|\overline{x}(t_1) - \overline{x}(t_2)\| \to 0$ which by Cauchy's criterion of convergence in $\Re^n$ implies that $\lim_{t\to\beta-} \overline{x}(t)$ exists. Let $\overline{x}_1$ be this limit. Since $\overline{x}(t) \in K$ and K is compact, it follows that $\overline{x}_1 \in K \subset \Omega$. Now define $\overline{y}(t)$ on $(\alpha,\ \beta]$ by

$$\overline{y}(t) = \begin{cases} \overline{x}(t), & t \in (\alpha,\ \beta) \\ \overline{x}_1, & t = \beta \end{cases}$$

$\overline{y}(t)$ is differentiable on $(\alpha,\ \beta]$ with $\dot{\overline{y}}(\beta) = f(\beta, \overline{y}(\beta))$, as $\overline{y}(t)$ satisfies the integral equation

$$\overline{y}(t) = \overline{x}_0 + \int_{t_0}^{t} f(s,\overline{y}(s)) ds$$

Thus $\overline{y}$ is a continuation of the solution $\overline{x}(t)$ on $(\alpha,\ \beta]$.
We now examine the IVP

$$\left.\begin{aligned} \frac{d\overline{x}}{dt} &= f(t,\overline{x}(t)) \\ \bar{x}(\beta) &= \bar{x}_1 \end{aligned}\right\} \tag{3.2.2}$$

By Theorem 3.1.1 the equation (3.2.2) has a unique solution $\overline{z}(t)$ on some open interval $(\beta - a, \beta + a)$ and hence $\overline{z}(t) = \overline{y}(t)$ on $(\beta - a,\ \beta)$ and at the end point β we have $\overline{z}(\beta) = \overline{x}_1 = \overline{y}(\beta)$.
Now we define a function $\overline{v}(t)$ as

$$\overline{v}(t) = \begin{cases} \overline{y}(t), & t \in (\alpha,\ \beta] \\ \overline{z}(t), & t \in [\beta, \beta + a) \end{cases}$$

Then $\overline{v}(t)$ is a solution of the IVP given by Eq. (3.1.1) on $(\alpha,\ \beta + a)$, contradicting the assumption that $(\alpha,\ \beta)$ is the maximal interval of existence for given IVP. Hence, if $\beta < \infty$, $\exists$ a $t \in (\alpha,\ \beta)$ such that $\overline{x}(t) \notin K$. ∎

As a Corollary of the above result, we immediately get the following theorem, called continuation theorem. The process of continuation is already described in the proof of Theorem 3.2.3.

Theorem 3.2.4 *Let Ω, f be as defined in Theorem 3.2.3 with $\overline{x}_0 \in \Omega$. Let $[t_0,\ \beta)$ be the right maximal interval of existence of solution $\overline{x}(t)$ of the IVP given by Eq. (3.1.1).*
Assume that $\exists$ a compact set $K \subset \Omega$ such that

$$\left\{\overline{y} \in \Re^n \ / \ \overline{y} = \overline{x}(t) \quad \text{for } t \in [\,t_0,\ \beta)\,\right\} \subset K.$$

Then $\beta = \infty$. That is,IVP given by Eq. (3.1.1) has a solution $\overline{x}(t)$ on $[t_0, \infty)$. That is, $\overline{x}(t)$ is a global solution.

Corollary 3.2.1 *Suppose $\Omega = \Re^n$ and satisfies the following condition:*
$\exists\ \rho > 0$ such that

$$||f(t, \overline{x})||_{\Re^n} \leq |a(t)|\ ||\overline{x}|| \ \ \forall\ ||\overline{x}||_{\Re^n} > \rho$$

with $a(t)$ as a continuous function.Then the initial value problem given by Eq. (3.1.1) has a global solution.

Proof : We have

$$\frac{d\overline{x}}{dt} = f(t, \overline{x}(t))$$

Taking scalar product with $\overline{x}(t)$ we get

$$\begin{aligned}\left(\overline{x}(t), \frac{d\overline{x}}{dt}\right) = \left(\overline{x}(t), \frac{d\overline{x}}{dt}\right)_{\Re^n} &= (f(t, \overline{x}(t)), \overline{x}(t)\ ) \\ &\leq ||f(t, \overline{x}(t))||\ ||\overline{x}(t)|| \\ &\leq |a(t)|\ ||\overline{x}(t)||^2 \text{ for } ||\overline{x}(t)|| \geq \rho\end{aligned}$$

This implies that

$$\frac{d}{dt}\left(||\overline{x}(t)||^2\right) \leq \frac{|a(t)|}{2}\ ||\overline{x}(t)||^2$$

and hence

$$||\overline{x}(t)||^2 \leq ||\overline{x}(0)||^2 e^{\int_{t_0}^{t} \frac{|a(t)|}{2} dt} \ \text{ for } \ ||\overline{x}(t)|| \geq \rho$$

If $||\overline{x}(t)|| \leq \rho$ for all t, then we apply the previous theorem to $K = \{\overline{x} : ||\overline{x}|| \leq \rho\}$ and get the result. If $||\overline{x}(t)|| > \rho$, we again apply the same theorem to
$K = \left\{\overline{x} : ||\overline{x}||^2 \leq ||\overline{x}(0)||^2 e^{\int_{t_0}^{\beta} \frac{|a(t)|}{2} dt}\right\}$. ▮

Corollary 3.2.2 *Let $\overline{x} \mapsto f(\ .,\ \overline{x})$ be Lipschitz continuous for all $\overline{x}(t) \in \Re^n$. Then Eq. (3.1.1) has a global solution.*

Proof : This follows from the previous corollary. ▮

Corollary 3.2.3 *Let $f(t,\overline{x}) : I \times \Re^n \to \Re^n$ be continuous with $\bar{x} \to -f(t,\bar{x})$ strongly monotone and $f(t,0) = 0$. Then the IVP given by Eq. (3.1.1) has a global solution.*

Proof : We have a solution $\overline{x}(t)$ of Eq. (3.1.1) satisfying

$$\frac{d\overline{x}}{dt} - f(t,\overline{x}(t)) = 0$$

Taking scalar product with $\overline{x}(t)$, we get

$$\left(\frac{d\overline{x}}{dt}, \overline{x}(t)\right) - (f(t,\overline{x}(t)), \overline{x}(t))) = 0$$

As $-f$ strongly monotone, this gives

$$\begin{aligned} \frac{d}{dt}\left(||x(t)||^2\right) &= 2\left(f(t,\overline{x}(t)), \overline{x}(t)\right) \\ &\leq -2c||\overline{x}(t)||^2 \end{aligned}$$

This implies that

$$||x(t)||^2 \leq ||x(0)||^2 e^{-2ct} \leq ||x(0)||^2 \ \ \forall \ t$$

This gives that $\overline{x}(t) \in K = \{\overline{x} \in \Re^n : ||\overline{x}|| \leq ||x(0)||\}$. Hence Theorem 3.2.4, implies that $\overline{x}(t)$ is a global solution of Eq. (3.1.1). ∎

For more details on this section, refer Perko[7].

3.3 Dependence on Initial Condition

In this section we investigate the dependence of the solution of the differential equation

$$\frac{d\bar{x}}{dt} = f(t,\bar{x}(t)) \tag{3.3.1}$$

on the initial condition $\bar{x}(t_0) \in \Re^n$. We first state and prove the following lemma called Gronwall's lemma.

Lemma 3.3.1 *Suppose that $g(t)$ is a continuous real valued function satisfying the conditions*

$$(i) \ \ g(t) \geq 0 \qquad (ii) \ \ g(t) \leq c + k\int_{t_0}^{t} g(s)ds \ \ \forall t \in [t_0, t_1]$$

(c, k are positive constants.)
Then, we have

$$g(t) \leq c\exp{(kt)} \ \ \forall \ \ t \in [t_0, t_1]$$

Proof : Let $G(t) = c + k\int_{t_0}^{t_1} g(s)ds, \quad t \in [t_0, t_1]$.

Then, $G(t) \geq g(t)$ and $G(t) > 0$ on $[t_0, t_1]$. By second fundamental theorem of calculus

$$\dot{G}(t) = kg(t)$$

and hence

$$\frac{\dot{G}(t)}{G(t)} = k\frac{g(t)}{G(t)} \leq k \ \ \forall\, t \in [t_0, t_1]$$

This gives

$$\frac{d}{dt}\left(\log G(t)\right) \leq k$$

Equivalently

$$G(t) \leq G(t_0)\exp\left(kt\right) \forall t \in [t_0, t_1]$$

which is the required result. ∎

Theorem 3.3.1 *Let $f(t, \bar{x})$ satisfy all conditions of Theorem 3.1.1. Then the solution $\bar{x}(t)$ of the differential equation (3.1.1) depends continuously on the initial condition $\bar{x}_0$.*

Proof : Let $\bar{x}(t)$, $\bar{y}(t)$ be unique solutions of Eq. (3.1.1), with respect to the initial conditions $\bar{x}_0$ and $\bar{y}_0$, respectively. Denote by J the interval of existence of solutions.
Using the corresponding integral equation analog we have

$$\begin{aligned}
\bar{x}(t) &= \bar{x}_0 + \int_{t_0}^{t} f(s, \bar{x}(s))\, ds & t \in J \\
\bar{y}(t) &= \bar{y}_0 + \int_{t_0}^{t} f(s, \bar{y}(s))\, ds & t \in J
\end{aligned}$$

Substracting the above equations and using triangle inequality we get

$$||\bar{x}(t) - \bar{y}(t)||_{\Re^n} \leq ||\bar{x}_0 - \bar{y}_0||_{\Re^n} + \int_{t_0}^{t} ||f(s, \bar{x}(s)) - f(s, \bar{y}(s))||_{\Re^n}\, ds$$

Borrowing our earlier notation of Theorem 3.1.1, we have $\bar{x}, \ \bar{y} \in X = C(J, \Omega_0), \quad \Omega_0 \subset \Omega$. Using Lipschitz continuity of f and Gronwall inequality, we get

$$||\bar{x}(t) - \bar{y}(t)||_{\Re^n} \leq ||\bar{x}_0 - \bar{y}_0||_{\Re^n} + M \int_{t_0}^{t} ||\bar{x}(s) - \bar{y}(s)||_{\Re^n}\, ds \tag{3.3.2}$$

and

$$||\bar{x}(t) - \bar{y}(t)||_{\Re^n} \leq ||\bar{x}_0 - \bar{y}_0||_{\Re^n} \exp\left(Mt\right) \ \ \forall t \in J = [t_0, t_f] \tag{3.3.3}$$

This gives that

$$||\bar{x} - \bar{y}|| = \sup_{t \in J} ||\bar{x}(t) - \bar{y}(t)||_{\Re^n} \leq ||\bar{x}_0 - \bar{y}_0||_{\Re^n} \ \exp(Mt_f) \tag{3.3.4}$$

This implies that the mapping $\bar{x}_0 \mapsto \bar{x}$ of $\Re^n$ into $X = C(J, \Omega_0)$ is continuous. ∎

We shall give an alternate proof of the above theorem by using the following lemma, using fixed points, rather than Gronwall's lemma.

Lemma 3.3.2 *Let (X, d) be a complete metric space and let F, F_k be contraction mappings on X with the same contraction constant $\alpha < 1$ and with the fixed points x, x_k respectively. Suppose that $\lim_{k \to \infty} F_k y = Fy \ \forall y \in X$. Then $\lim_{k \to \infty} x_k = x$.*

Proof : Using the error estimate given in Theorem 2.4.1 for F_k, we get

$$d(x_k, F_k^l y) \leq \frac{\alpha^l}{1-\alpha} d(F_k y, y), \quad \text{for any arbitrary } \ y \in X \ \text{ and for any } l.$$

Set $l = 0$ and $y = x$. We have

$$d(x_k, x) \leq \frac{1}{1-\alpha} d(F_k x, x) \to 0, \quad \text{as } k \to \infty$$

This implies that $x_k \to x$. ∎

Alternate Proof (of Theorem 3.3.1): Let the space X and constants M, m, δ, τ be as defined in Theorem 3.1.1 .
Let $\bar{x}_k(t), \ \bar{x}(t)$ be solutions of Eq. (3.3.1) with initial conditions $\bar{x}_k^{(0)}$ and $\bar{x}^{(0)}$ respectively and let $\bar{x}_k^{(0)} \to \bar{x}^{(0)}$.
Define $F_k, \ F : X \mapsto X$ as under

$$\begin{aligned} [F_k \bar{x}](t) &= \bar{x}_k^{(0)} + \int_{t_0}^{t} f(s, \bar{x}(s))\, ds, \quad \bar{x} \in X \\ [F \bar{x}](t) &= \bar{x}^{(0)} + \int_{t_0}^{t} f(s, \bar{x}(s))\, ds, \quad \bar{x} \in X \end{aligned}$$

We have

$$\begin{aligned} ||F_k \bar{x}(t) - \bar{x}^{(0)}|| &\leq ||\bar{x}_k^{(0)} - \bar{x}^{(0)}|| + m\tau \\ &< \delta \text{ for } k \text{ large (as } \bar{x}_k^{(0)} \to \bar{x}^{(0)}) \end{aligned}$$

Thus F_k maps X to itself and so is the case for F (already proved in section 1). Also $\bar{x}_k, \ \bar{x}$ are fixed points of F_k and F, respectively with the same contraction constant $\alpha = \tau m < 1$. Also, $F_k \bar{x} \to F \bar{x}$ for $\bar{x} \in X$, as $\bar{x}_k^{(0)} \to \bar{x}^{(0)}$ in X. Using Lemma 3.3.2, we get that the fixed points $\bar{x}_k$ of F_k and $\bar{x}$ of F also converge. This proves the theorem.

3.4 Cauchy-Peano Existence Theorem

We shall prove an existence theorem for IVP given by Eq. (3.1.1) with the assumptions of only continuity and boundedness of f.
For this purpose, we need Schauder's fixed point theorem - Theorem 2.4.2.
We also need the following theorem, called Ascoli-Arzela theorem (refer Rudin [8]), which gives the relative compactness for a class $\mathcal{F}$ of continuous functions.

Definition 3.4.1 *A class $\mathcal{F}$ of continuous functions is said to be equicontinous, if given $\epsilon > 0$, $\exists\, \delta > 0$ such that*

$$||f(t) - f(t')|| < \epsilon \text{ whenever } |t - t'| < \delta \text{ for all } f \in \mathcal{F}$$

(δ is independent of $f \in \mathcal{F}$).

Definition 3.4.2 *$\mathcal{F}$ is said to be uniformly bounded if $\exists\ M$ independent of $f \in \mathcal{F}$ such that*

$$||f(t)|| \le M, \ \forall\, t \ \text{ and } \forall\, f \in \mathcal{F}$$

Theorem 3.4.1 (Ascoli-Arzela theorem) *Let $\mathcal{F}$ be a class of continuous functions defined over some interval J. Then $\mathcal{F}$ is relatively compact iff $\mathcal{F}$ is equicontinous and uniformly bounded.*

Theorem 3.4.2 (Cauchy-Peano theorem) *Let E be as defined in Theorem 3.1.1 and $(t_0, x_0) \in E$. Assume that $f : E \to \Re^n$ is continuous. Then the initial value problem given by Eq. (3.1.1) has a solution.*

Proof : Let the space X and constants $m,\ \tau,\ \delta$, be as defined in Theorem 3.1.1
Recall that X is a complete metric space of all continuous mappings from $J = [t_0 - \tau, t_0 + \tau]$ into $\Omega_0 = \{x \in \Omega : ||\bar{x} - \bar{x}_0|| \le \delta\}$ and hence is a closed, convex bounded subset of the space $C(J)$.
As before, define $F : X \mapsto X$ as

$$[F\bar{x}(t)] = \bar{x}_0 + \int_{t_0}^{t} f(s, \bar{x}(s))\, ds, \ \ t \in J$$

Define $\mathcal{F} = \{F\bar{x} : \bar{x} \in X\}$
$\mathcal{F}$ is a collection of continuous functions defined on J. $\mathcal{F}$ is equicontinous, since

$$\begin{aligned} ||F\bar{x}(t_1) - F\bar{x}(t_2)||_{\Re^n} &\le \int_{t_1}^{t_2} ||f(s, \bar{x}(s))\,||_{\Re^n}\, ds \\ &\le m|t_1 - t_2| \ \ \forall\, \bar{x} \in X \end{aligned}$$

Further, $\mathcal{F}$ is uniformly bounded as

$$\begin{aligned} ||F\bar{x}(t)||_{\Re^n} &\le ||\bar{x}_0||_{\Re^n} + \int_{t_0}^{t} ||f(s, \bar{x}(s))||_{\Re^n}\, ds \\ &\le ||\bar{x}_0||_{\Re^n} + m\tau \ \ \forall\, \bar{x} \in X \end{aligned}$$

Hence by Ascoli-Arzela theorem, $\mathcal{F}$ is relatively compact.
To apply Schauder's fixed point theorem, it remains to be shown that $F : X \mapsto X$ is continuous.We first observe the following.
$f : J \times \Omega_0 \mapsto \Re^n$ is continuous on a compact set and hence it is uniformly continuous. This implies that given $\epsilon > 0$, $\exists\ \delta(\epsilon) > 0$ such that

$$||f(s, \bar{x}(s)) - f(s, \bar{y}(s))||_{\Re^n} < \frac{\epsilon}{\tau} \text{ whenever } ||x(s) - y(s)||_{\Re^n} < \delta$$

and this δ depends only on ϵ.
We have

$$||F\bar{x}(t) - F\bar{y}(t)||_{\Re^n} \le \int_{t_0}^{t} ||f(s, \bar{x}(s)) - f(s, \bar{y}(s))||_{\Re^n}\ ds$$

This gives

$$\begin{aligned} ||\bar{x} - \bar{y}|| < \delta \Rightarrow ||F\bar{x} - F\bar{y}|| &\le \int_{t_0}^{t_0+\tau} ||f(s, \bar{x}(s)) - f(s, \bar{y}(s))||_{\Re^n} ds \\ &\le \frac{\epsilon}{\tau}\tau = \epsilon \end{aligned}$$

That is, $F : X \mapsto X$ is continuous. Thus $F : X \mapsto X$ satisfies all assumptions and hence has a fixed point $\bar{x} \in X$. That is

$$\bar{x}(t) = \bar{x}_0 + \int_{t_0}^{t} f(s, \bar{x}(s))ds, \quad t \in J$$

This proves that Eq. (3.1.1) has a solution.

However, this solution need not be unique. In Example 1.2.1, the existence of a solution of the initial value problem

$$\begin{aligned} \frac{dx}{dt} &= f(x), \qquad \bar{x}(0) = 0 \\ f(x) &= \begin{cases} 2\sqrt{x}, & x \ge 0 \\ 0, & x \le 0 \end{cases} \end{aligned}$$

is guaranteed by Cauchy-Peano theorem. This solution is not unique, as already seen before. ∎

For more on existence and uniqueness of IVP refer Amann [1], Arnold [2], Hartman [3] and Mattheij and Molenaar [5].

3.5 Exercises

1. Find the regions (in $\Re^n \times \Re^n$) of existence and uniquness of solutions of the IVP associated with the following differential equations.

(a) $\frac{dx}{dt} = \sqrt[3]{xt}$

(b) $\frac{dx}{dt} = \sqrt{|x|}$

(c) $\frac{dx}{dt} = \begin{cases} \sqrt{x} & x \geq 0 \\ -\sqrt{-x} & x \leq 0 \end{cases}$

(d) $\frac{dx}{dt} = \begin{cases} \frac{2x}{t} & t \neq 0 \\ 0 & x = 0 \end{cases}$

2. (a) Let $x(t)$ be a solution of the differential equation

$$\frac{dx}{dt} = (t^2 - x^2)\sin x + x^2 \cos x$$

which vanishes for $t = t_0$. Show that $x(t) \equiv 0$.

(b) Find the solution of the IVP

$$\begin{aligned} \frac{dx}{dt} &= x|x| \\ x(t_0) &= 0 \end{aligned}$$

3. Show that the IVP

$$\begin{aligned} \frac{d^2x}{dt^2} &= f(t, x(t)) \\ x(t_0) &= x_{01}, \quad \dot{x}(t_0) = x_{02} \end{aligned}$$

is equivalent to the integral equation

$$x(t) = x_{01} + (t - t_0)x_{02} + \int_{t_0}^{t_1} (t - s)f(s, x(s))ds, \quad t \in J$$

where J is the interval of existence of solution.

4. Let $x(t)$ and $y(t)$ be two solutions of the nonhomogeneous linear differential equation

$$\frac{d^n x}{dt^n} + a_{n-1}(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1(t)\frac{dx}{dt} + a_0(t)x = b(t)$$

in the interval J around the initial point t_0. Show that

$$u(t_0)exp(-2k|t - t_0|) \leq u(t) \leq u(t_0)exp(k|t - t_0|)$$

where $u(t)$ is given by

$$u(t) = \sum_{j=0}^{n-1} \left(x^{(j)}(t) - y^{(j)}(t) \right)$$

$$x^{(j)}(t) = \frac{d^j x}{dt^j},\ y^{(j)} = \frac{d^j x}{dt^j} \text{ and } k = 1 + \sum_{j=1}^{n-1} \sup_{t\in J} |a_j(t)|.$$

5. (a) Show that $0 < x_0 < x(t) < \pi$ if $0 < x_0 < \pi$, where $x(t)$ is the solution of the IVP

$$\begin{aligned} \frac{dx}{dt} &= t \sin t \\ x(0) &= x_0 \end{aligned}$$

(b) Show that $x_0 < x(t) < 1$ for $t \geq 0$ and $0 < x(t) < x_1$ for $t \leq 0$, if $x(t)$ is the solution of the IVP

$$\begin{aligned} \frac{dx}{dt} &= x - x^2 \\ x(0) &= x_0, \quad 0 < x_0 < 1 \end{aligned}$$

6. Consider the IVP

$$\begin{aligned} \frac{dx}{dt} &= x^p \\ x(0) &= 1 \end{aligned}$$

Find the soultion of this problem for different values of p. Find how the maximum interval I_{max} of existence depends on p.

7. Let $x(t)$ be the solution of Eq. (3.1.1), where $x \to f(t,x)$ is Lipschitz continuous from $\Re$ to $\Re$ for all $t \in \Re$.

Let $x_1(t),\ x_2(t)$ be two continuous function on $\Re$, such that

$$\begin{aligned} \frac{dx_1}{dt} &\leq f(t, x_1(t)) \\ x_1(t_0) &\leq x_0 \end{aligned}$$

and

$$\begin{aligned} \frac{dx_2}{dt} &\geq f(t, x_2(t)) \\ x_2(t_0) &\geq x_0 \end{aligned}$$

Show that

$$x_1(t) \leq x(t) \leq x_2(t) \qquad \forall t \geq t_0$$

8. Find the maximal interval of existence of solution for the IVP

$$\begin{aligned} \frac{dx_1}{dt} &= -\frac{x_2}{x_3}, \quad \frac{dx_2}{dt} = -\frac{x_1}{x_3}, \ \frac{dx_3}{dt} = 1 \\ \bar{x}(\frac{1}{\pi}) &= (0, 1, \frac{1}{\pi}) \end{aligned}$$

9. Show that for the IVP

$$\frac{dx}{dt} = ax, \quad x(0) = 1$$

the Picard iterates $x_n(t)$ converge to the unique solution e^{at}, for more than one initial guess $x_0(t)$.

10. Consider the following perturbed IVP (corresponding to the above problem)

$$\frac{dy}{dt} = ax, \quad y(0) = 1 + \epsilon$$

Show that $|x(t) - y(t)| \leq |\epsilon| e^{|a|t}$

Graphically, show that the above estimate is very accurate for $a > 0$, but very inaccurate for $a < 0$.

References

[1] Herbet Amann, *Ordinary Differential Equations : Introduction to Nonlinear Analysis, Watter de Gruyter, 1990.*

[2] V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations, Springer-Verlag, 1983.*

[3] Philip Hartman, *Ordinary Differential Equations, Birkauser, 1982.*

[4] Mohan C Joshi and R. K. Bose, *Some Topics in Nonlinear Functional Analysis, Halsted Press, 1985.*

[5] R. M. M. Mattheij and J. Molenaar, *Ordinary Differential Equations in Theory and Practice, John Wiley, 1996.*

[6] A. Pazy, *Semigroup of Linear Operators and Applications to Partial Differential Equations, Springer - Verlag, 1983.*

[7] Lawrence Perko, *Differential Equations and Dynamical Systems, Springer - Verlag, 1991.*

[8] Walter Rudin, *Real and Complex Analysis, Tata McGraw Hill Publications, 1974.*

Chapter 4

Linear System And Transition Matrix

The focus of this chapter is on the linear system of differential equations, wherein the concept of transition matrix plays a crucial role.
We give the definition of transition matrix, its properties and the method of computation. Some of the illustrative examples, introduced earlier in Chapter 1, are recalled for demonstrating the methodology involved in computation.

4.1 Linear System

In this chapter we shall be concerned with the properties of the linear system of the form

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t) + \overline{b}(t) \tag{4.1.1}$$

This is a special case of Eq. (3.1.1) of Chapter 3.

Here $A(t)$ is an $n \times n$ matrix and $\overline{b}(t)$ is a vector in $\Re^n$, with entries $a_{ij}(t)$ and $b_i(t)$ respectively, as continuous functions of t. The existence and uniqueness of global solution of Eq. (4.1.1) with the initial condition

$$\overline{x}(t_0) = \overline{x}_0 \tag{4.1.1(a)}$$

follows from Theorem 3.1.1.
To study the solution of Eq. (4.1.1), it helps to study the homogeneous system

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t) \tag{4.1.2}$$

We shall first prove the existence of n linearly independent solutions of Eq. (4.1.2).

Definition 4.1.1 *Let $\overline{x}^{(1)}(t),\ \overline{x}^{(2)}(t),\ \cdots,\ \overline{x}^{(n)}(t)$ be n-vector valued functions. The Wronskian $W(t)$, is defined by*

$$W(t) = \begin{vmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \cdots & x_1^{(n)}(t) \\ x_2^{(1)}(t) & x_2^{(2)}(t) & \cdots & x_2^{(n)}(t) \\ \vdots & \cdots & \ddots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \cdots & x_n^{(n)}(t) \end{vmatrix}$$

Here $\overline{x}^{(i)} = \left(x_1^{(i)}, x_2^{(i)},\ \cdots, x_n^{(i)}\right),\quad 1 \le i \le n.$

Theorem 4.1.1 *Let $W(t_0) \neq 0$ for some $t_0 \in J \subseteq \Re$. Then $\overline{x}^{(1)}(t),\ \overline{x}^{(2)}(t),\ \cdots,\ \overline{x}^{(n)}(t)$ are linearly independent in the interval J.*

Proof : If possible, let $\left\{\overline{x}^{(1)}(t),\ \cdots,\ \overline{x}^{(n)}(t)\right\}$ be linearly dependent. This implies that there exist n constants $c_1,\ c_2,\ \cdots,\ c_n$ (not all zero) such that

$$c_1\overline{x}^{(1)}(t) + c_2\overline{x}_2(t) + \cdots + c_n\overline{x}^{(n)}(t) = \bar{0} \tag{4.1.3}$$

Eq. (4.1.3) implies that

$$\begin{array}{ccccccccc} c_1x_1^{(1)}(t_0) & + & c_2x_1^{(2)}(t_0) & + & \cdots & + & c_nx_1^{(n)}(t_0) & = & 0 \\ c_1x_2^{(1)}(t_0) & + & c_2x_2^{(2)}(t_0) & + & \cdots & + & c_nx_2^{(n)}(t_0) & = & 0 \\ \vdots & + & \cdots & + & \ddots & + & \vdots & & \vdots \\ c_1x_n^{(1)}(t_0) & + & c_2x_n^{(2)}(t_0) & + & \cdots & + & c_nx_n^{(n)}(t_0) & = & 0 \end{array} \tag{4.1.4}$$

Eq. (4.1.4) is a linear system of equations in n-unknowns $c_1,\ c_2,\ \cdots,\ c_n$. As the determinant $W(t_0)$ of this system is nonzero, it follows that $c_i = 0$ for $1 \le i \le n$ (as RHS of Eq. (4.1.4) is zero). This is a contradiction. Hence the set $\left\{\overline{x}^{(1)}(t),\ \overline{x}^{(2)}(t),\ \cdots,\ \overline{x}^{(n)}(t)\right\}$ is linearly independent. ∎

This gives us the following theorem.

Theorem 4.1.2 *There exists n linearly independent solutions of Eq. (4.1.2) in J.*

Proof : We consider n initial value problems

$$\begin{aligned} \frac{d\overline{x}}{dt} &= A(t)\overline{x}(t), & t \in J \\ \overline{x}(t_0) &= \overline{e}^{(i)}, & 1 \le i \le n \end{aligned}$$

where $\overline{e}^{(i)}$ are the unit vectors $(0,\ \cdots, 1\ , \cdots,\ 0)$.

By Theorem 3.1.1, the above set of n initial value problems has n unique solutions $\overline{x}^{(1)}(t),\ \overline{x}^{(2)}(t),\ \cdots,\ \overline{x}^{(n)}(t)$, with initial conditions $\overline{x}^{(i)}(t_0) = \overline{e}^{(i)}$, $(1 \le i \le n)$. Let

$$W(t) = \begin{vmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \cdots & x_1^{(n)}(t) \\ x_2^{(1)}(t) & x_2^{(2)}(t) & \cdots & x_2^{(n)}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \cdots & x_n^{(n)}(t) \end{vmatrix}$$

Then

$$W(t_0) = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \quad = \quad 1$$

By Theorem 4.1.1, $\overline{x}^{(1)}(t),\ \cdots,\ \overline{x}^{(n)}(t)$ are linearly independent on J. ∎

Let X be the vector space of solutions of Eq. (4.1.2). The following theorem gives the dimension of X.

Theorem 4.1.3 *The dimension of the solution space X of the linear homogeneous system given by Eq. (4.1.2) is n.*

Proof : We have already proved in Theorem 4.1.2 that Eq. (4.1.2) has n linearly independent solutions $\overline{x}^{(1)}(t),\ \cdots,\ \overline{x}^{(n)}(t)$ with $\overline{x}^{(i)}(t_0) = \overline{e}^{(i)}$.
Let $\overline{x}(t)$ be any general solution of Eq. (4.1.2). Let us assume that $\overline{x}(t_0) = \overline{x}^{(0)}$, with $\overline{x}^{(0)} = \left(x_1^{(0)},\ x_2^{(0)},\ \cdots,\ x_n^{(0)}\right)$. Consider the function $\overline{y}(t)$ defined as $\overline{y}(t) = x_1^{(0)}\overline{x}^{(1)}(t) + x_2^{(0)}\overline{x}^{(2)}(t) + \cdots + x_n^{(0)}\overline{x}^{(n)}(t)$. It is clear that

$$\frac{d\overline{y}}{dt} = A(t)\overline{y}(t)$$

Also,

$$\begin{aligned} \overline{y}(t_0) &= x_1^{(0)}\overline{x}^{(1)}(t_0) + x_2^{(0)}\overline{x}^{(2)}(t_0) + \cdots + x_n^{(0)}\overline{x}^{(n)}(t_0) \\ &= x_1^{(0)}\overline{e}^{(1)} + x_2^{(0)}\overline{e}^{(2)} + \cdots + x_n^{(0)}\overline{e}^{(n)} \\ &= \left(x_1^{(0)},\ \cdots,\ .x_n^{(0)}\right) = \overline{x}^{(0)} \end{aligned}$$

Thus $\overline{y}(t)$ is a solution of the initial value problem

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t),\ \overline{x}(t_0) = \overline{x}^{(0)}$$

By uniqueness theorem of IVP, it follows that

$$\overline{x}(t) = \overline{y}(t) = \sum_{i=1}^{n} x_i^{(0)} \overline{x}^{(i)}(t)$$

That is, $\bar{x}(t)$ is spanned by n linearly independent solutions $\bar{x}^{(1)}(t), \bar{x}^{(2)}(t), \cdots, \bar{x}^{(n)}(t)$ of Eq. (4.1.2) and hence X has dimension n. ∎

At times, it helps to know the relation between $W(t)$ and the matrix $A(t)$. The following theorem gives the relation.

Theorem 4.1.4 *Let $\overline{x}^{(1)}(t), \cdots, \overline{x}^{(n)}(t)$ be solutions of Eq. (4.1.2) and let $t_0 \in J$. Then we have the Abel's formula*

$$W(t) = W(t_0) \exp\left(\int_{t_0}^{t} Tr\,[A(s)]\, ds \right) \tag{4.1.5}$$

where $Tr\,[A(t)]$ is defined as

$$Tr[A(t)] = \sum_{i=1}^{n} a_{ii}(t)$$

Proof : We have

$$W(t) = \begin{vmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \cdots & x_1^{(n)}(t) \\ x_2^{(1)}(t) & x_2^{(2)}(t) & \cdots & x_2^{(n)}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \cdots & x_n^{(n)}(t) \end{vmatrix}$$

This gives

$$\frac{dW(t)}{dt} = \sum_{i=1}^{n} \begin{vmatrix} x_1^{(1)}(t) & \cdots & x_1^{(n)}(t) \\ x_1^{(1)}(t) & \cdots & x_1^{(n)}(t) \\ \vdots & \ddots & \vdots \\ x_{i-1}^{(1)}(t) & \cdots & x_{i-1}^{(n)}(t) \\ \dot{x}_i^{(1)}\ (t) & \cdots & \dot{x}_i^{(n)}\ (t) \\ x_{i+1}^{(1)}(t) & \cdots & x_{i+1}^{(n)}(t) \\ \vdots & \ddots & \vdots \\ x_n^{(1)}(t) & \cdots & x_n^{(n)}(t) \end{vmatrix} \tag{4.1.6}$$

In view of Eq. (4.1.2), we have

$$\dot{x}_i^{(k)}(t) = \frac{dx_i^{(k)}(t)}{dt} = \sum_{j=1}^{n} a_{ij}(t)x_j^{(k)}(t), \qquad 1 \le k \le n$$

and hence Eq. (4.1.6) gives

$$\frac{dW(t)}{dt} = \sum_{i=1}^{n} \begin{vmatrix} x_1^{(1)}(t) & \cdots & x_1^{(n)}(t) \\ \vdots & \vdots & \vdots \\ \sum_{j=1}^{n} a_{ij}(t)x_j^{(1)}(t) & \cdots & \sum_{j=1}^{n} a_{ij}(t)x_j^{(n)}(t) \\ \vdots & \vdots & \vdots \\ x_n^{(1)}(t) & \cdots & x_n^{(n)}(t) \end{vmatrix}$$

Multiply the first row by $a_{i1}(t)$, the second by $a_{i2}(t)$ and so on except the i^{th} row and substract their sum from the i^{th} row. We get

$$\frac{dW(t)}{dt} = \sum_{i=1}^{n} \begin{vmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \cdots & x_1^{(n)}(t) \\ \vdots & \vdots & \vdots & \vdots \\ a_{ii}(t)x_j^{(1)}(t) & a_{ii}(t)x_j^{(2)}(t) & \cdots & a_{ii}(t)x_j^{(n)}(t) \\ \vdots & \vdots & \vdots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \cdots & x_n^{(n)}(t) \end{vmatrix}$$

That is

$$\frac{dW}{dt} = \sum_{i=1}^{n} a_{ii}(t)W(t) = W(t)Tr[A(t)]$$

which gives

$$W(t) = W(t_0)\exp\left(\int_{t_0}^{t} Tr[A(t)]dt\right)$$

∎

Corollary 4.1.1 *If $\bar{x}^{(1)}(t), \cdots, \bar{x}^{(n)}(t)$ are linearly independent solutions of Eq. (4.1.2), then $W(t) \neq 0$ on J.*

Example 4.1.1 *Cosider the n^{th} order linear differential equation in $J \subseteq \Re$*

$$\frac{d^n x(t)}{dt^n} + a_{n-1}(t)\frac{d^{n-1}x(t)}{dt} + \cdots + a_1(t)\frac{dx(t)}{dt} + a_0 x(t) = 0 \qquad (4.1.7)$$

with $a_i(t),\ 0 \le i \le n-1$ as continuous functions of $t \in J$.

Proceeding as in Corollary 3.1.2, one can show that Eq. (4.1.7) is equivalent to the linear first order homogeneous system

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t) \tag{4.1.8}$$

where

$$A(t) = \begin{bmatrix} 0 & 1 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}, \quad \overline{x}(t) = \begin{bmatrix} x(t) \\ \frac{dx}{dt} \\ \vdots \\ \frac{d^{n-1}x(t)}{dt^{n-1}} \end{bmatrix}$$

Recall that $A(t)$ is the companion matrix of Eq. (4.1.7).
It is now clear from Theorem 4.1.3 that Eq. (4.1.8) has n linearly independent solutions $\bar{x}_1(t),\ \bar{x}_2(t),\ \cdots,\ \bar{x}_n(t)$ and its Wronskian is defined as

$$W(t) = \begin{vmatrix} x_1(t) & x_2(t) & \cdots & x_n(t) \\ x_1^{(1)}(t) & x_2^{(1)}(t) & \cdots & x_n^{(1)}(t) \\ x_1^{(2)}(t) & x_2^{(2)}(t) & \cdots & x_n^{(n)}(t) \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{(n)}(t) & x_2^{(n)}(t) & \cdots & x_n^{(n)}(t) \end{vmatrix}$$

Here $x_k^{(n)}(t)$ denotes $\dfrac{d^n x_k(t)}{dt^n}$.

Remark 4.1.1 *The converse of Theorem 4.1.1 is not true, as we see from the following example.*

$\overline{x}^{(1)}(t) = \begin{bmatrix} t \\ 1 \end{bmatrix}$, $\overline{x}^{(2)}(t) = \begin{bmatrix} t^2 \\ t \end{bmatrix}$ *are linearly independent functions defined on* $\Re$ *but* $W(t) \equiv 0$.

4.2 Transition Matrix

In the previous section we proved that the homogeneous linear system

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t) \tag{4.2.1}$$

has n linearly independent solutions $\overline{x}^{(1)},\ \overline{x}^{(2)},\ \cdots,\ \overline{x}^{(n)}$.
Define a nonsingular matrix $\Psi(t)$ as

$$\Psi(t) = \left[\overline{x}^{(1)}(t),\ \overline{x}^{(2)}(t),\ \cdots,\ \overline{x}^{(n)}(t)\right] \tag{4.2.2}$$

Then, we have

$$\begin{aligned} \frac{d\Psi(t)}{dt} &= \left[\frac{d\overline{x}^{(1)}(t)}{dt},\ \frac{d\overline{x}^{(2)}(t)}{dt},\ \cdots,\ \frac{d\overline{x}^{(n)}(t)}{dt}\right] \\ &= \left[A(t)\overline{x}^{(1)}(t),\ A(t)\overline{x}^{(2)}(t),\ \cdots,\ A(t)\overline{x}^{(n)}(t)\right] \end{aligned}$$

That is

$$\frac{d\Psi}{dt} = A(t)\Psi(t) \tag{4.2.3}$$

Definition 4.2.1 *A nonsingular matrix $\Psi(t)$ is called a fundamental matrix if it satisfies the matrix differential equation given by Eq. (4.2.3).*

Definition 4.2.2 *A nonsingular matrix $\Phi(t, t_0)$ is called principal fundamental matrix or transition matrix or evolution matrix if it satisfies the matrix differential equation given by Eq. (4.2.3) with the initial condition $\Phi(t_0, t_0) = I$.*

That is, transition matrix $\Phi(t, t_0)$, is the solution of the matrix initial value problem

$$\frac{d\Phi(t)}{dt} = A(t)\Phi(t), \qquad \Phi(t_o) = I \tag{4.2.4}$$

Here the solution $\Phi(t)$ of the above differential equation lies in the space $\mathcal{M}_{n\times n}$ of $n \times n$ matrices. Note that $\mathcal{M}_{n\times n}$ is a Hilbert space (finite dimensional) with inner product and norm defined as

$$(A, B)_{\mathcal{M}_{n\times n}} = \sum_{i=1}^{n} \left(\bar{a}^{(i)}, \bar{b}^{(i)}\right),\ \|A\|^2_{\mathcal{M}_{n\times n}} = \sum_{i=1}^{n} \|\bar{a}^{(i)}\|^2$$

where $A = \left[\bar{a}^{(1)}, \bar{a}^{(2)}, \cdots, \bar{a}^{(n)}\right]$ and $B = \left[\bar{b}^{(1)}, \bar{b}^{(2)}, \cdots, \bar{b}^{(n)}\right]$.

Proceeding as before, the solvability of Eq. (4.2.4) is equivalent to the solvability of the Volterra integral equation

$$\phi(t) = I + \int_{t_0}^{t} A(s)\phi(s)ds, \quad t \in J = (t_0, t_f) \subset (t_0, \infty) \tag{4.2.5}$$

in the space $C[J, \mathcal{M}_{n\times n}]$. Its unique solution $\phi(t)$ is given by

$$\phi(t) = I + \sum_{n=1}^{\infty} \int_{t_0}^{t} A_n(s)ds$$

where $A_n(t) = \int_{t_0}^{t}\int_{t_0}^{\sigma_1}\cdots\int_{t_0}^{\sigma_{n-1}} [A(\sigma_1)A(\sigma_2)\cdots A(\sigma_n)\, d\sigma_n \cdots d\sigma_1]$. Equivalentely, we have

$$\phi(t) = I + \int_{t_0}^{t} A(\sigma_1)d\sigma_1 + \int_{t_0}^{t}\int_{t_0}^{\sigma_1} A(\sigma_1)A(\sigma_2)\, d\sigma_2\, d\sigma_1 + \cdots \tag{4.2.6}$$

Eq. (4.2.5) or Eq. (4.2.6) is called the Peano-Baker series for the transition matrix $\Phi(t)$ of the linear differential equation given by Eq. (4.2.1) (refer Brockett [2]).

Example 4.2.1 *Let $A(t)$ be a constant matrix A. Then the transition matrix of the homogeneous system given by Eq. (4.2.1) is of the form*

$$\begin{aligned}\Phi(t, t_0) &= I + \int_{t_0}^{t} Ads + \int_{t_0}^{t}\int_{t_0}^{s} A^2 d\tau ds + \cdots \\ &= I + A(t-t_0) + A^2\frac{(t-t_0)^2}{2!} + \cdots \\ &= \exp\left(A(t-t_0)\right)\end{aligned}$$

Example 4.2.2 *Let* $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
$A^2 = -I,\ A^3 = -A,\ A^4 = I,\ A^5 = A, \cdots, A^{2k} = (-1)^k I, A^{2k+1} = (-1)^k A.$
This gives

$$\begin{aligned}\exp\left(A(t-t_0)\right) &= I + A(t-t_0) - I\frac{(t-t_0)^2}{2!} - \frac{A(t-t_0)^3}{3!} \\ &\quad + \frac{(t-t_0)^4}{4!} + A\frac{(t-t_0)^5}{5!} + \cdots \\ &= I\left[1 - \frac{(t-t_0)^2}{2!} + \frac{(t-t_0)^4}{4!} + \cdots\right] \\ &\quad + A\left[(t-t_0) - \frac{(t-t_0)^3}{3!} + \cdots\right] \\ &= \begin{bmatrix} \cos(t-t_0) & \sin(t-t_0) \\ -\sin(t-t_0) & \cos(t-t_0) \end{bmatrix}\end{aligned}$$

The following theorem implies that two different fundamental matrices corresponding to the same system differ by a multiple of a constant nonsingular matrix.

Theorem 4.2.1 *Let $\Psi_1(t)$ be a fundamental matrix of the system given by Eq. (4.2.1). Then $\Psi_2(t)$ is a fundamental matrix of Eq. (4.2.1) iff there exists a nonsingular constant matrix C such that $\Psi_2(t) = \Psi_1(t)C$.*

Proof : Let $\Psi_2(t) = \Psi_1(t)C$. Then

$$\frac{d\Psi_2(t)}{dt} = \frac{d\Psi_1(t)}{dt}\,C = A\Psi_1(t)\,C = A\Psi_2(t)$$

This implies that $\Psi_2(t)$ is a fundamental matrix of Eq. (4.2.1). Conversely, let $\Psi_2(t)$ be another fundamental matrix of Eq. (4.2.1).
Set

$$C(t) = \Psi_1^{-1}(t)\Psi_2(t)$$

This gives

$$\Psi_2(t) = \Psi_1(t)C(t)$$

and

$$\frac{d\Psi_2(t)}{dt} = \frac{d\Psi_1(t)}{dt}C(t) + \Psi_1(t)\frac{dC(t)}{dt}$$

Hence

$$\begin{aligned} A(t)\Psi_2(t) &= A(t)\Psi_1(t)C(t) + \Psi_1(t)\frac{dC(t)}{dt} \\ &= A(t)\Psi_2(t) + \Psi_1(t)\frac{dC(t)}{dt} \end{aligned}$$

This implies that

$$\Psi_1(t)\frac{dC(t)}{dt} = 0$$

and hence

$$\frac{dC(t)}{dt} = 0$$

Thus C is a constant matrix. Also C is invertible as $\Psi_1(t),\ \Psi_2(t)$ are invertible. ∎

Corollary 4.2.1 *If $\Psi(t)$ is a fundamental matrix of Eq. (4.2.1), the transition matrix $\Phi(t,t_0)$ of Eq. (4.2.1) is given by*

$$\Phi(t,t_0) = \Psi(t)\Psi^{-1}(t_0)$$

Corollary 4.2.2 *The initial value problem*

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t), \qquad \overline{x}(t_0) = \overline{x}_0$$

has a solution $x(t)$, given by

$$\bar{x}(t) = \Psi(t)\Psi^{-1}(t_0)\bar{x}_0$$

where $\Psi(t)$ is a fundamental matrix of Eq. (4.2.1). That is

$$\overline{x}(t) = \Phi(t,t_0)\overline{x}_0$$

Proof : We have

$$\frac{d\Psi(t)}{dt} = A(t)\Psi(t)$$

This gives

$$\left[\frac{d\overline{x}^{(1)}(t)}{dt}, \frac{d\overline{x}^{(2)}(t)}{dt} \cdots \frac{d\overline{x}^{(n)}(t)}{dt}\right] = \left[A\overline{x}^{(1)},\ A\overline{x}^{(2)}, \ldots A\overline{x}^{(n)}\right]$$

where

$$\Psi(t) = \left[\overline{x}^{(1)},\ \overline{x}^{(2)}, \ldots \overline{x}^{(n)}\right]$$

with $\{\overline{x}^{(i)}\}_{i=1}^{n}$ linearly independent.
This implies that

$$\frac{d\overline{x}^{i}}{dt} = A\overline{x}^{i}, \qquad 1 \leq i \leq n$$

Thus we have n linearly independent solutions of Eq. (4.2.1) and as its solution space is of dimension n, it follows that every solution $\overline{x}(t)$ of Eq. (4.2.1) is given by

$$\overline{x}(t) = c_1\overline{x}^{(1)}(t) + \cdots + c_n\overline{x}^{(n)}(t)$$

That is

$$\overline{x}(t) = \Psi(t)\overline{c}$$

where $\bar{c} = (c_1, c_2, \cdots, c_n)$.
Since $\overline{x}(t_0) = \bar{x}_0$, we get

$$\overline{c} = \Psi^{-1}(t_0)\overline{x}_0$$

This gives

$$\overline{x}(t) = \Psi(t)\Psi^{-1}(t_0)\overline{x}_0 = \Phi(t, t_0)\overline{x}_0$$

∎

We also note the following facts.

Remark 4.2.1 *If $\Psi_1(t)$ and $\Psi_2(t)$ are two fundamental matrices of the systems $\frac{d\overline{x}}{dt} = A_1(t)\overline{x}(t)$ and $\frac{d\overline{x}}{dt} = A_2(t)\overline{x}(t)$, respectively ($A_1 \neq A_2$). Then, $\Psi_1(t) \neq \Psi_2(t)$.*

This follows from the fact that $A_1(t) = \frac{d\Psi_1(t)}{dt}\Psi_1^{-1}(t)$ and
$A_2(t) = \frac{d\Psi_2(t)}{dt}\Psi_2^{-1}(t)$.

Remark 4.2.2 $\Phi(t, t_0)$ *satisfies the following properties*

(i) $\Phi(t, t_0) = \Phi(t, t_1)\Phi(t_1, t_0)$,

(ii) $\Phi^{-1}(t,t_0) = \Phi(t_0,t)$

(iii) $\Phi(t,t) = I$

for all $t, t_0, t_1 \in J$

To realise this, just observe that $\Phi(t,t_0) = \Psi(t)\Psi^{-1}(t_0)$.

We shall say a linear homogeneous differential equation in $\bar{p}(t)$ is adjoint equation associated with the given equation $\frac{d\bar{x}}{dt} = A(t)\bar{x}(t)$, provided that for any initial data, the scalar product $(\overline{x}(t), \overline{p}(t))$ is constant for all t.

This gives the following theorem.

Theorem 4.2.2 *The adjoint equation associated with*

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t)$$

is

$$\frac{d\bar{p}}{dt} = -A(t)^{\top}\bar{p}(t) \tag{4.2.7}$$

Proof : We have

$$\begin{aligned} \frac{d}{dt}\left[(\bar{x}(t),\ p(t))\right] &= \left(\frac{d\bar{x}}{dt},\ \bar{p}(t)\right) + \left(\bar{x}(t),\ \frac{d\bar{p}(t)}{dt}\right) \\ &= (A(t)\bar{x}(t), \bar{p}(t)) + \left(\bar{x}(t), -A(t)^{\top}\bar{p}(t)\right) \\ &= ([A(t) - A(t)]\,\bar{x}(t),\ \bar{p}(t)) \\ &= 0 \end{aligned}$$

This gives

$$(\bar{x}(t), \bar{p}(t)) = constant$$

That is Eq. (4.2.7) is the adjoint equation associated with Eq. (4.2.1). ∎

Theorem 4.2.3 *If* $\Psi(t)$ *is a fundamental matrix of the system given by Eq. (4.2.1), then* $\left[\Psi^{-1}(t)\right]^{\top}$ *is a fundamental matrix of the adjoint system.*

Proof : We have $\Psi(t)\Psi^{-1}(t) = I$. Differentiating this equality, we get

$$\frac{d\Psi(t)}{dt}\Psi^{-1}(t) + \Psi(t)\frac{d\Psi^{-1}(t)}{dt} = 0$$

and hence

$$A(t) + \Psi(t)\frac{d\Psi^{-1}(t)}{dt} = 0$$

This gives

$$\frac{d\Psi^{-1}(t)}{dt} = -\Psi^{-1}(t)A(t)$$

and hence

$$\left[\frac{d\Psi^{-1}(t)}{dt}\right]^\top = -A(t)^\top \left[\Psi^{-1}(t)\right]^\top$$

∎

Theorem 4.2.1 gives us the following corollary.

Corollary 4.2.3 *If $\Psi(t)$ is a fundamental matrix of Eq. (4.2.1) then $\chi(t)$ is fundamental matrix of its adjoint system given by Eq. (4.2.7) iff*

$$\chi(t)^\top \Psi(t) = C$$

where C is a constant nonsingular matrix.

Remark 4.2.3 *If $A(t)$ is skew symmetric, that is, $-[A(t)]^\top = A(t)$, then the transition matrix is orthogonal $\left([\Phi(t,t_0)]\left[\Phi^\top(t,t_0)\right] = I\right)$. This follows from the fact that*

(i) $\frac{d}{dt}\left[\phi^\top(t,t_0)\phi(t,t_0)\right] = \phi^\top(t,t_0)\left[A^\top(t) + A(t)\right]\phi(t,t_0) = 0$ *and*

(ii) $\phi(t_0,t_0) = I$.

We shall now give the method of variation of parameter to solve the non-homogeneous system

$$\frac{d\overline{x}}{dt} = A(t)\overline{x}(t) + \overline{b}(t) \tag{4.2.8}$$

We set

$$\overline{x}(t) = \Phi(t,t_0)\overline{v}(t) \tag{4.2.9}$$

which satisfies Eq. (4.2.8), $\overline{v}(t)$ to be determined.
Plugging this representation of $\bar{x}(t)$ in Eq. (4.2.8), we get

$$\left[\frac{d}{dt}\Phi(t,t_0)\right]\overline{v}(t) + \Phi(t,t_0)\frac{d\overline{v}(t)}{dt} = A(t)\ \Phi(t,t_0)\ \overline{v}(t) + \overline{b}(t)$$

This gives

$$A(t)\Phi(t,t_0)\bar{v}(t) + \Phi(t,t_0)\frac{d\overline{v}(t)}{dt} = A(t)\Phi(t,t_0)\bar{v}(t) + \overline{b}(t)$$

and hence

$$\begin{aligned}\frac{d\overline{v}(t)}{dt} &= \Phi^{-1}(t,t_0)\overline{b}(t) \\ &= \Phi(t_0,t)\overline{b}(t)\end{aligned}$$

Integrating this equation, we get

$$\overline{v}(t) = \overline{v}(t_0) + \int_{t_0}^{t} \Phi(t_0, s)\overline{b}(s)ds$$

As $\overline{x}(t) = \Phi(t, t_0)\overline{v}(t)$, it follows $\overline{x}(t_0) = \overline{v}(t_0)$ and

$$\overline{x}(t) = \Phi(t, t_0)\overline{v}(t_0) + \int_{t_0}^{t} \Phi(t, t_0)\Phi(t_0, s)\overline{b}(s)ds$$

This gives us the variation of parameter formula

$$\overline{x}(t) = \Phi(t, t_0)\overline{x}_0 + \int_{t_0}^{t} \Phi(t, s)\overline{b}(s)ds$$

Example 4.2.3 *Solve*

$$\frac{d\overline{x}}{dt} = A\overline{x}(t) + \overline{b}(t)$$

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \overline{b}(t) = \begin{bmatrix} t \\ 0 \end{bmatrix}$$

The transition matrix for this system is given by

$$\Phi(t, t_0) = \begin{bmatrix} \cos(t - t_0) & \sin(t - t_0) \\ -\sin(t - t_0) & \cos(t - t_0) \end{bmatrix}$$

By the method of variation of parameter we have

$$\overline{x}(t) = \overline{x}^{(0)} + \int_{t_0}^{t} \Phi(t, s)\overline{b}(s)ds$$

where $x^{(0)} = \overline{x}(t_0)$.

$$\begin{aligned} \int_{t_0}^{t} \Phi(t, s)\overline{b}(s)ds &= \begin{bmatrix} \int_{t_0}^{t} s\cos(t - s)ds \\ \\ \int_{t_0}^{t} -s\sin(t - s)ds \end{bmatrix} \\ &= \begin{bmatrix} t_0 \sin(t - t_0) + 1 - \cos(t - t_0) \\ \\ t_0 \cos(t - t_0) - 1 + \sin(t - t_0) \end{bmatrix} \end{aligned}$$

This gives the solution of the non-homogeneous system as

$$\begin{bmatrix} x_1(t) \\ \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_1^{(0)} \\ \\ x_2^{(0)} \end{bmatrix} + \begin{bmatrix} t_0 \sin(t - t_0) + 1 - \cos(t - t_0) \\ \\ t_0 \cos(t - t_0) - 1 + \sin(t - t_0) \end{bmatrix}$$

For more on this section refer Agarwal and Gupta [1].

4.3 Periodic Linear System

Let us first consider a homogenous periodic system

$$\begin{aligned} \frac{d\bar{x}}{dt} &= A(t)\bar{x}(t), \quad t \in \Re && (4.3.1a) \\ A(t+T) &= A(t), \quad\quad t \in \Re && (4.3.1b) \end{aligned}$$

We shall make use of the following lemma, which we shall state without proof (refer Coddington and Levinson [3] for further details).

Lemma 4.3.1 *Let C be any nonsingular matrix, then $\exists$ a matrix R such that $C = e^R$.*

We observe that Eq. (4.3.1) is a non-autonomous system, yet it is possible to get a simpler representition of the transition matrix $\phi(t, t_0)$ in view of the periodicity of $A(t)$.

noindent The Peano-Baker series representation of the transition matrix $\phi(t, t_0)$ gives

$$\phi(t, t_0) = I + \int_{t_0}^{t} A(\sigma_1)\, d\sigma_1 + \int_{t_0}^{t} A(\sigma_1) \int_{t_0}^{\sigma_1} A(\sigma_2) d\sigma_2\, d\sigma_1 + \cdots$$

As $\int_{t_0+T}^{t+T} A(\sigma)\, d\sigma = \int_{t_0}^{t} A(\sigma)\, d\sigma$, it follows that $\phi(t, t_0)$ is also periodic of period T :

$$\phi(t + T, t_0 + T) = \phi(t, t_0).$$

Theorem 4.3.1 *(Floquet-Lyapunov). If $A(t + T) = A(t)$, the associated transition matrix $\phi(t, t_0)$ of the system given by Eq. (4.3.1) has the form*

$$\phi(t, t_0) = P^{-1}(t) e^{R(t-t_0)} P(t_0)$$

where $P(t)$ is a nonsingular periodic matrix function and R is a fixed matrix.

Proof : As $\phi(T, 0)$ is a nonsingular matrix, by Lemma 4.3.1, there exists a matrix R such that $\phi(T, 0) = e^{RT}$ (redefine R in Lemma 4.3.1).
We now define a nonsingular $P(t)$, through its inverse $P^{-1}(t)$ as

$$P^{-1}(t) = \phi(t, 0)e^{-Rt}$$

Then $P^{-1}(t + T)$ is satisfies the relationship

$$\begin{aligned} P^{-1}(t+T) &= \phi(t+T, 0)e^{-R(t+T)} \\ &= \phi(t+T, T)\phi(T, 0)e^{-RT}e^{-Rt} \\ &= \phi(t+T, T)e^{-Rt} \\ &= P^{-1}(t) \end{aligned}$$

That is $P(t)$ is periodic.
Also we have

$$\phi(t,0) = P^{-1}(t)e^{Rt}, \;\; \phi(0,t) = e^{-Rt}P(t) \tag{4.3.2}$$

Eq. (4.3.2) imply that

$$\begin{aligned} \phi(t,t_0) &= \phi(t,0)\phi(0,t_0) \\ &= P^{-1}(t)e^{Rt}e^{-Rt_0}P(t_0) \\ &= P^{-1}(t)e^{R(t-t_0)}P(t_0) \end{aligned}$$

This proves the theorem. ∎

It is a pleasant surprise to observe that if we make the transformation

$$\bar{x}(t) = P^{-1}(t)\bar{z}(t) \tag{4.3.3}$$

then $\bar{z}(t)$ satisfies the autonomous system

$$\frac{d\bar{z}}{dt} = Rz(t) \tag{4.3.4}$$

This can be seen as follows. Since $\frac{d\phi}{dt} = A(t)\phi(t)$, we have

$$\frac{d}{dt}\left[P^{-1}(t)e^{R(t-t_0)}P(t_0)\right] = A(t)\left[P^{-1}(t)e^{R(t-t_0)}P(t_0)\right]$$

Equivalently,

$$\begin{aligned} \left[\frac{d}{dt}P^{-1}(t)\right]\; e^{R(t-t_0)}\; P(t_0) + P^{-1}(t)\; Re^{R(t-t_0)}\; P(t_0) \\ = A(t)P^{-1}(t)\; e^{R(t-t_0)}P(t_0) \end{aligned}$$

which gives

$$P^{-1}(t)R = A(t)P^{-1}(t) - \frac{d}{dt}P^{-1}(t) \tag{4.3.5}$$

Operating Eq. (4.3.5) on $\bar{z}(t)$ we get

$$\begin{aligned} P^{-1}(t)R\bar{z}(t) &= A(t)P^{-1}(t)\bar{z}(t) - \left[\frac{d}{dt}P^{-1}(t)\right]\bar{z}(t) \\ &= A(t)\bar{x}(t) - \left[\frac{d}{dt}P^{-1}(t)\right]\bar{z}(t) \end{aligned} \tag{4.3.6}$$

Using transformation given by Eq. (4.3.3) we get

$$\frac{d\bar{x}}{dt} = \left[\frac{d}{dt}P^{-1}(t)\right]\bar{z}(t) + P^{-1}(t)\frac{d\bar{z}}{dt}$$

and hence Eq. (4.3.6) reduces to

$$\begin{aligned} P^{-1}(t)R\bar{z}(t) &= A(t)\bar{x}(t) - \frac{d\bar{x}}{dt} + P^{-1}(t)\frac{d\bar{z}}{dt} \\ &= P^{-1}(t)\frac{d\bar{z}}{dt} \end{aligned}$$

This gives

$$\frac{d\bar{z}}{dt} = R\bar{z}(t)$$

which is the desired result.

Example 4.3.1 *Consider the following periodic system*

$$\begin{aligned} \frac{d\bar{x}}{dt} &= A(t)\bar{x}(t) \\ A(t) &= \begin{bmatrix} -2\cos^2 t & -1-\sin 2t \\ 1-\sin 2t & -2\sin^2 t \end{bmatrix} \\ A(t+2\pi) &= A(t) \end{aligned}$$

One can verify that

$$\phi(t) = \begin{bmatrix} e^{-2t}\cos t & -\sin t \\ e^{-2t}\sin t & \cos t \end{bmatrix}$$

is a fundamental matrix of this system.
One observes that

$$\begin{aligned} \phi(t,0) &= \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} e^{-2t} & 0 \\ 0 & 1 \end{bmatrix} \\ &= P^{-1}(t)e^{Rt} \end{aligned}$$

where $P(t) = \phi(t,0) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$ *is periodic of period* 2π *and* $R = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$

Let us now examine the non-homogeneous initial value periodic system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + \bar{b}(t) \tag{4.3.7a}$$

$$\bar{x}(t_0) = \bar{x_0} \tag{4.3.7b}$$

where $A(t+T) = A(t),\ \bar{b}(t+T) = \bar{b}(t)$.
We state the following theorem regarding the solvability of Eq. (4.3.7), refer Brocket[2] for a proof of this theorem.

Theorem 4.3.2 *Let $\phi(t, t_0)$ be the transition matrix generated by $A(t)$. Then the solution of the non-homogeneous system given by Eq. (4.3.7) can be written as*

$$\bar{x}(t) = \bar{x_p}(t) + \phi(t, t_0)\left[\bar{x_0} - \bar{x_p}(t_0)\right] \tag{4.3.8}$$

iff $\bar{b}(t)$ is orthogonal to the solution $\bar{p}(t)$ of the adjoint equation

$$\frac{d\bar{p}}{dt} = -A^\top(t)\bar{p}(t) \tag{4.3.9}$$

($\bar{b}$ is orthogonal to $\bar{p}$ iff $\displaystyle\int_{t_0}^{t_0+T} \left(\bar{p}(\sigma), \bar{b}(\sigma)\right) d\sigma = 0.$)

Example 4.3.2 *Consider the harmonic oscillator problem with periodic external force $b(t)$ (refer Example 1.3.1)*

$$\frac{d^2x}{dt^2} + \omega^2 x(t) = b(t),\ \omega^2 = \frac{\lambda}{ma}$$

This is equivalent to

$$\begin{aligned} \begin{bmatrix} \dfrac{dx_1}{dt} \\ \\ \dfrac{dx_2}{dt} \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ b(t) \end{bmatrix} \\ \Longleftrightarrow \frac{d\bar{x}}{dt} &= A\bar{x}(t) + \bar{b}(t) \end{aligned} \tag{4.3.10}$$

with $x_1(t) = x(t),\ x_2(t) = \dot{x}(t)$.
The eigenvalues of A are $\pm\omega i$ and hence the transition matrix corresponding to A is given by

$$\begin{aligned} \phi(t, t_0) &= \begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix} \begin{bmatrix} \cos\omega(t-t_0) & \sin\omega(t-t_0) \\ \\ -\sin\omega(t-t_0) & \cos\omega(t-t_0) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \\ 0 & \dfrac{1}{\omega} \end{bmatrix} \\ &= \begin{bmatrix} \cos\omega(t-t_0) & \dfrac{1}{\omega}\sin\omega(t-t_0) \\ \\ -\omega\sin\omega(t-t_0) & \cos\omega(t-t_0) \end{bmatrix} \end{aligned}$$

Hence the solution of the homogeneous equation corresponding to Eq. (4.3.10) is given by

$$\begin{aligned} \overline{x}_H(t) = \Bigg(& x_0^{(1)}\cos\omega(t-t_0) + \frac{x_0^{(2)}}{\omega}\sin\omega(t-t_0)\ , \\ & -x_0^{(1)}\omega\sin\omega(t-t_0) + x_0^{(2)}\cos\omega(t-t_0) \Bigg) \end{aligned}$$

This gives

$$x_H(t) = x_H(0)\cos\omega(t-t_0) + \frac{\dot{x}_H(0)}{\omega}\sin\omega(t-t_0) \tag{4.3.11}$$

The adjoint system corresponding to Eq. (4.3.10) is given by

$$\frac{d\bar{p}}{dt} = \begin{bmatrix} 0 & \omega^2 \\ -1 & 0 \end{bmatrix}\bar{p}(t)$$

The transition matrix of the adjoint system is

$$\begin{bmatrix} \cos\omega(t-t_0) & \omega\sin\omega(t-t_0) \\ \\ -\frac{1}{\omega}\sin\omega(t-t_0) & \cos\omega(t-t_0) \end{bmatrix}$$

Hence the non-homogeneous equation will have a solution iff $\bar{p}(t)$ is orthogonal to $\bar{b}(t)$. That is, $b(t)$ satisfies the orthogonality condition

$$\int_{t_0}^{t_0+T}\begin{bmatrix} \cos\omega(t-t_0) & \omega\sin\omega(t-t_0) \\ \\ -\frac{1}{\omega}\sin\omega(t-t_0) & \cos\omega(t-t_0) \end{bmatrix}\begin{bmatrix} 0 \\ \\ b(t) \end{bmatrix} dt = 0$$

This implies that $b(t)$ satisfies the conditon

$$\int_{t_0}^{t_0+T}\sin\omega(t-t_0)\ b(t)d(t) = \int_{t_0}^{t_0+T}\cos\omega(t-t_0)\ b(t)d(t) = 0$$

where T is the time period $\dfrac{2\pi}{\omega}$.

4.4 Computation of Transition Matrix

We shall discuss some basic properties of the linear transformation $\exp A$ in order to facilitate its computaton. We note that $\exp(A(t-t_0))$ is the representation of the transition matrix of the system

$$\frac{d\overline{x}}{dt} = Ax(t) \tag{4.4.1}$$

where A is independent of t.
We note the following properties.

[1] If P and Q are linear transformation on $\Re^n$ and $S = PQP^{-1}$, then $\exp(S) = P\exp(Q)P^{-1}$.

This follows from the fact

$$\begin{aligned}\exp S = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{(PQP^{-1})^k}{k!} &= P\left(\lim_{n\to\infty}\sum_{k=0}^{n}\frac{Q^k}{k!}\right)P^{-1} \\ &= P\exp(Q)P^{-1}\end{aligned}$$

This implies that if $P^{-1}AP = \text{diag}[\lambda_j]$, then

$$\Phi(t,t_0) = \exp\left(A(t-t_0)\right) = P \ \text{diag}[\ \exp\left(\lambda_j(t-t_0)\right)\ ]\ P^{-1}$$

Here, $\text{diag}[\lambda_j]$ represents the diagonal matrix with entries λ_j.

[2] If P and Q are linear transformation on $\Re^n$ which commute then

$$\exp\left(P+Q\right) = \exp\left(P\right)\exp\left(Q\right)$$

Example 4.4.1 *If* $A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$, *then*

$$\exp\left(A(t-t_0)\right) = \exp\left(a(t-t_0)\right)\begin{bmatrix} 1 & b(t-t_0) \\ 0 & 1 \end{bmatrix}$$

This follows from the fact that

$$A = aI + B, \quad B = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$$

Also, B commutes with I and $B^2 = 0$. Using property[2], we get the required representation as follows,

$$\begin{aligned} \exp\left(A(t-t_0)\ \right) &= \exp\left(a(t-t_0)I\ \right)\exp B(t-t_0) \\ &= \exp\left(a(t-t_0)\ \right)[I + B(t-t_0)] \\ &= \exp\left(a(t-t_0)\ \right)\begin{bmatrix} 1 & b(t-t_0) \\ 0 & 1 \end{bmatrix} \end{aligned}$$

[3] If $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, then $\exp\left(A\right) = \exp\left(a\right)\begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}$

Denote $\lambda = a + ib$. By induction, we have

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}^k = \begin{bmatrix} Re(\lambda^k) & -Im(\lambda^k) \\ Im(\lambda^k) & Re(\lambda^k)) \end{bmatrix}$$

This gives us $\exp\left(A\right)$ as

$$\begin{aligned} \exp(A) &= \sum_{k=0}^{\infty}\begin{bmatrix} Re\left[\frac{\lambda^k}{k!}\right] & -Im\left[\frac{\lambda^k}{k!}\right] \\ Im\left[\frac{\lambda^k}{k!}\right] & Re\left[\frac{\lambda^k}{k!}\right] \end{bmatrix} \\ &= \begin{bmatrix} Re\left[\exp\left(\lambda\right)\right] & -Im\left[\exp\left(\lambda\right)\right] \\ Im\left[\exp\left(\lambda\right)\right] & Re\left[\exp\left(\lambda\right)\right] \end{bmatrix} \\ &= \exp\left(a\right)\begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix} \end{aligned}$$

In a more general setting, we have the following theorem.

Theorem 4.4.1 *If the set of eigenvalues* $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ *of an* $n \times n$ *matrix* A *is real and distinct, then any set of corresponding eigenvectors* $\{\overline{u}^1, \cdots, \overline{u}^n\}$ *forms a basis for* $\Re^n$. *The matrix* $P = \left[\overline{u}^{(1)}, \cdots, \overline{u}^{(n)}\right]$ *is invertible and* $P^{-1}AP = diag\left[\lambda_1, \lambda_2, \ldots, \lambda_n\right]$ *and hence the transition matrix* $\phi(t, t_0)$ *is given by*

$$\phi(t,t_0) = P \begin{bmatrix} \exp\left(\lambda_1(t-t_0)\right) & 0 & \cdots & 0 \\ 0 & \exp\left(\lambda_2(t-t_0)\right) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \exp\lambda_n(t-t_0) \end{bmatrix} P^{-1}$$

Proof : We have

$$A\bar{u}^{(i)} = \lambda_i \bar{u}^{(i)}, \ \ \bar{u}^{(i)} \neq \bar{0}, \ \lambda_i \neq \lambda_j$$

We claim that $\bar{u}^{(i)}, \cdots, \bar{u}^{(n)}$ are linearly independent. We proceed by induction. Obviously, the result is true for $k = 1$ as $\bar{u}^{(1)} \neq 0$ is linearly independent. So let the result be true for $k \leq n-1$. We shall show that it is also true for all $k = n$. Let

$$\alpha_1 \bar{u}^{(1)} + \alpha_2 \bar{u}^{(2)} + \cdots + \alpha_n \bar{u}^{(n)} = \bar{0}, \ \ \alpha_i \in \Re \tag{4.4.2}$$

Operating by A we get

$$\alpha_1 \lambda_1 \bar{u}^{(1)} + \alpha_2 \lambda_2 \bar{u}^{(2)} + \cdots + \alpha_n \lambda_n \bar{u}^{(n)} = \bar{0} \tag{4.4.3}$$

Multiply Eq. (4.4.2) by λ_n and subtract from Eq. (4.4.3), we get

$$\alpha_1(\lambda_1 - \lambda_n)\bar{u}^{(1)} + \alpha_2(\lambda_1 - \lambda_n)\bar{u}^{(2)} + \cdots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)\bar{u}^{(n-1)} = \bar{0}$$

By induction $\bar{u}^{(1)}, \cdots, \bar{u}^{(n-1)}$ are linear independent and hence $\alpha_i = 0, \ 1 \leq i \leq n-1$.
Putting these values in Eq. (4.4.2) we get $\alpha_n \lambda_n \bar{u}^{(n)} = 0$ which gives $\alpha_n = 0$ as $\lambda_n \bar{u}^{(n)} \neq 0$.
Thus $\bar{u}^{(1)} \cdots \bar{u}^{(n)}$ are linear independent by induction. Hence the matrix $P = \left[\bar{u}^{(1)}, \cdots, \bar{u}^{(n)}\right]$ is nonsingular. Further, we have

$$\begin{aligned} AP &= \left[A\bar{u}^{(1)}, \cdots, A\bar{u}^{(n)}\right] \\ &= \left[\lambda_1 \bar{u}^{(1)}, \cdots, \lambda_n \bar{u}^{(n)}\right] \\ &= P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = P \text{ diag } [\lambda_j] \end{aligned}$$

This implies that $P^{-1}AP = \text{ diag } [\lambda_j]$
Proceeding as before, the transition matrix $\phi(t, t_0)$ is given by

$$e^{A(t-t_0)} = P \exp(\text{ diag } [\lambda_j(t-t_0)]) P^{-1}$$

$$= P \begin{bmatrix} e^{\lambda_1(t-t_0)} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2(t-t_0)} & \cdots & 0 \\ \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & e^{\lambda_n(t-t_0)} \end{bmatrix} P^{-1}$$

∎

Example 4.4.2 $A = \begin{bmatrix} -1 & -3 \\ 0 & 2 \end{bmatrix}$
$\lambda_1 = -1, \;\; \lambda_2 = 2.$ *A pair of corresponding eigenvectors is* $(1, 0), \; (-1, 1)$.

$$P = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

This gives

$$\begin{aligned} \Phi(t-t_0) &= \exp(A(t-t_0)) \\ &= P \begin{bmatrix} \exp(-(t-t_0)) & 0 \\ 0 & \exp(-2(t-t_0)) \end{bmatrix} P^{-1} \\ &= \begin{bmatrix} \exp(-(t-t_0)) & \exp(-(t-t_0)) - \exp(2(t-t_0)) \\ 0 & \exp(2(t-t_0)) \end{bmatrix} \end{aligned}$$

If eigenvalues are complex, we state the following theorem concerning the computation of the transition matrix.

Theorem 4.4.2 *Let A be a $2n \times 2n$ matrix with $2n$ distinct complex eigenvalues $\lambda_j = a_j + ib_j, \; \overline{\lambda}_j = a_j - ib_j$ with corresponding eigenvectors $\overline{w}^{(j)} = \overline{u}^{(j)} \pm i \; \overline{v}^{(j)}, \;\; 1 \le j \le n$. Then $\{\overline{u}^{(1)}, \; \overline{v}^{(1)}, \cdots, \overline{u}^{(n)}, \; \overline{v}^{(n)}\}$ is a basis for $\Re^{2n}$ and the matrix $P = [\overline{u}^{(1)}, \; \overline{v}^{(1)}, \cdots, \; \overline{u}^{(n)}, \; \overline{v}^{(n)}]$ is invertible and*
$P^{-1}AP = \; diag \begin{bmatrix} a_j & b_j \\ -b_j & a_j \end{bmatrix}$ *with 2×2 blocks along the diagonal . The transiton matrix $\phi(t, t_0)$ has the representation.*

$$\phi(t, t_0) = P \exp\left(diag \begin{bmatrix} a_j & b_j \\ -b_j & a_j \end{bmatrix} \right) P^{-1}$$

$$= P \;\; diag \left((\exp(a_j(t-t_0))) \begin{bmatrix} \cos(b_j(t-t_0)) & \sin(b_j(t-t_0)) \\ -\sin(b_j(t-t_0)) & \cos(b_j(t-t_0)) \end{bmatrix} \right) P^{-1}$$

Example 4.4.3 $A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

A has eigenvalues $\lambda_{1,2} = 1 \pm i, \lambda_{3,4} = 2 \pm i$. A corresponding pair of eigenvectors is

$$\overline{w}^{(1)} = (i, \ 1, \ 0, \ 0)\,, \ \overline{w}^{(2)} = (0, \ 0, \ 1+i, \ 1)$$

$$\begin{aligned} P &= \begin{bmatrix} \bar{v}^{(1)} & \bar{u}^{(1)} & \bar{v}^{(2)} & \bar{u}^{(2)} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ P^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ P^{-1}AP &= \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \end{aligned}$$

$$\Phi(t, t_0) = P \begin{bmatrix} \exp(t)\cos t & -\exp(t)\sin t & 0 & 0 \\ \exp(t)\sin t & \exp(t)\cos t & 0 & 0 \\ 0 & 0 & \exp(2t)\cos t & -\exp(2t)\sin t \\ 0 & 0 & \exp(2t)\sin t & \exp(2t)\cos t \end{bmatrix} P^{-1}$$

$$= \begin{bmatrix} \exp(t)\cos t & -\exp(t)\sin t & 0 & 0 \\ \exp(t)\sin t & \exp(t)\cos t & 0 & 0 \\ 0 & 0 & \exp(2t)(\cos t + \sin t) & -2\exp(2t)\sin t \\ 0 & 0 & 2\exp(2t)\sin t & \exp(2t)(\cos t - \sin t) \end{bmatrix}$$

For multiple eigenvalues of A we proceed as follows.

Definition 4.4.1 *let λ be an eigenvalue of $n \times n$ matrix A of multiplicity $m \leq n$. Then for $k = 1 \ldots m$, any non zero solution $\overline{v}$ of*

$$[(A - \lambda I)^k]\overline{v} = 0$$

is called a **generalized eigenvector of** *A.*

Definition 4.4.2 *A matrix N is said to be* **nilpotent** *of order k if $N^{k-1} \neq 0$ and $N^k = 0$.*

Theorem 4.4.3 *Let A be a real $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ repeated according to their multiplicity. Then there exist a basis of generalized eigenvectors $\overline{v}^{(1)}, \overline{v}^{(2)}, \ldots, \overline{v}^{(n)}$ with $P = [\overline{v}^{(1)}, \overline{v}^{(2)}, \ldots, \overline{v}^{(n)}]$ invertible and $A = S + N$, where $P^{-1}SP = diag\,[\lambda_j]$. Further, the matrix $N = A - S$ is nilpotent of order $k \leq n$ and S and N commute. The transition matrix $\phi(t, t_0)$ is computed as follows*

$$\begin{aligned} \Phi(t, t_0) &= \left[P\, diag\, [\exp\left(\lambda_j(t - t_0)\right)]\, P^{-1}\right] \\ &\quad \times \left[1 + N(t - t_0) + \cdots + N^{k-1}\frac{(t - t_0)^{k-1}}{(k-1)!}\right] \end{aligned}$$

If λ is of multiplicity n, then $S = diag[\lambda]$.

For proofs of Theorem 4.4.2 and Theorem 4.4.3, refer Perko[4].

Example 4.4.4 $A = \begin{bmatrix} 0 & -2 & -1 & -1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$\lambda = 1$ is of multiplicity 4. $S = [I]_{4\times 4}$ and $N = A - S$.

$$\begin{aligned} N &= \begin{bmatrix} -1 & -2 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\ N^2 &= \begin{bmatrix} -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ N^3 &= 0 \end{aligned}$$

$$\Phi(t, t_0) = \exp\left(t - t_0\right)\left[I + N(t - t_0) + \frac{N^2(t - t_0)^2}{2!}\right]$$

$$=$$

$$\exp(t-t_0)\times$$

$$\begin{bmatrix} 1-[(t-t_0) + \frac{(t-t_0)^2}{2}] & -[2(t-t_0) + \frac{(t-t_0)^2}{2}] & -[(t-t_0) + \frac{(t-t_0)^2}{2}] & -[(t-t_0) + \frac{(t-t_0)^2}{2}] \\ (t-t_0) & 1+(t-t_0) & (t-t_0) & (t-t_0) \\ \frac{(t-t_0)^2}{2} & (t-t_0)+\frac{(t-t_0)^2}{2} & 1+\frac{(t-t_0)^2}{2} & \frac{(t-t_0)^2}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Example 4.4.5 $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}$

Here $\lambda_1 = 1, \lambda_2 = \lambda_3 = 2$
$\overline{v}^{(1)} = (1,\ 1,\ -2\ ),\ \ \overline{v}^{(2)} = (0,\ 0,\ 1\ ), \overline{v}^{(3)}$, *obtained by solving* $(A-2I)^2\overline{v} = 0$, *is given by*

$$0 \quad = \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ -2 & 0 & 0 \end{bmatrix} \overline{v}$$

That is $\overline{v}^{(3)} = (0,\ 1,\ 0)$
So, we get

$$P \quad = \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

$$N \quad = \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

$$\begin{aligned} \Phi(t,0) \quad &= \quad P \begin{bmatrix} \exp(t) & 0 & 0 \\ 0 & \exp(2t) & 0 \\ 0 & 0 & \exp(2t) \end{bmatrix} P^{-1}[1+Nt] \\ &= \begin{bmatrix} \exp(t) & 0 & 0 \\ \exp(t)-\exp(2t) & \exp(2t) & 0 \\ -\exp(2t)+(2-t)\exp(2t) & t\exp(2t) & \exp(2t) \end{bmatrix} \end{aligned}$$

Example 4.4.6 *Let us consider the problem of mechanical oscillations discussed in Example 1.3.1, which is modelled by the second order differential equaiton*

$$\frac{d^2x}{dt^2} + k\frac{dx}{dt} + \omega^2 x = F(t) \tag{4.4.4}$$

k is the resistance and $\omega^2 = \dfrac{\lambda}{ma}$.
This is equivalent to the following non-homogenous linear system

$$\begin{aligned} \frac{d\bar{x}}{dt} &= \begin{bmatrix} 0 & 1 \\ -\omega^2 & -k \end{bmatrix} \bar{x} + \begin{bmatrix} 0 \\ F(t) \end{bmatrix} \\ &= A\bar{x} + \bar{b}(t) \end{aligned}$$

$$A = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -k \end{bmatrix}, \quad \bar{b} = \begin{bmatrix} 0 \\ F(t) \end{bmatrix}$$

(i) *Eigenvalues of A are complex if $k^2 - 4\omega^2 < 0$ and are given by* $-\dfrac{k}{2} \pm \dfrac{i}{2}\sqrt{4\omega^2 - k^2} = a \pm bi \ \left(a = -\dfrac{k}{2} \text{ and } b = \sqrt{\omega^2 - (k^2/4)}\right)$.
Hence the transition matrix is given by

$$\begin{aligned} &\phi(t, t_0) \\ &= \\ &e^{a(t-t_0)} \begin{bmatrix} 1 & 0 \\ -\dfrac{k}{2} & b \end{bmatrix} \begin{bmatrix} \cos b(t-t_0) & \sin b(t-t_0) \\ -\sin b(t-t_0) & \cos b(t-t_0) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \dfrac{k}{2b} & \dfrac{1}{b} \end{bmatrix} \\ &= \\ &e^{a(t-t_0)} \begin{bmatrix} \cos b(t-t_0) + \dfrac{k}{2b}\sin b(t-t_0) & \dfrac{1}{b}\sin b(t-t_0) \\ -\dfrac{\omega^2}{b}\sin b(t-t_0) & \cos b(t-t_0) - \dfrac{k}{2b}\sin b(t-t_0) \end{bmatrix} \end{aligned}$$

This gives the solution of the homogeneous equation corresponding to Eq. (4.4.4), passing through the initial point $x(t_0) = x_0^{(1)}, \ \dot{x}(t_0) = x_2^{(2)}$ *as*

$$x(t) = e^{a(t-t_0)}\left[x(t_0)\left(\cos b(t-t_0) + \frac{k}{2b}\sin b(t-t_0)\right) + \frac{\dot{x}(t_0)}{b}\sin b(t-t_0)\right]$$

(ii) *Eigenvalues of A are real and distinct if $k^2 - 4\omega^2 > 0$ and are given by* $-\dfrac{k}{2} \pm \dfrac{1}{2}\sqrt{k^2 - 4\omega^2} = \dfrac{-k \pm c}{2}, \quad c = \sqrt{k^2 - 4\omega^2}$.

So, the transition matrix corresponding to A is given by

$$\phi(t,t_0)$$

$$= \begin{bmatrix} 1 & 1 \\ \frac{-k+c}{2} & \frac{-k-c}{2} \end{bmatrix} \begin{bmatrix} e^{\left(\frac{-k+c}{2}\right)(t-t_0)} & 0 \\ 0 & e^{\left(\frac{-k-c}{2}\right)(t-t_0)} \end{bmatrix}$$

$$\times \begin{bmatrix} \frac{k+c}{2c} & \frac{1}{c} \\ \frac{-k+c}{2c} & -\frac{1}{c} \end{bmatrix}$$

$$= \begin{bmatrix} e^{\left(\frac{-k+c}{2}\right)(t-t_0)} & e^{\left(\frac{-k-c}{2}\right)(t-t_0)} \\ \left(\frac{-k+c}{2}\right) e^{\left(\frac{-k+c}{2}\right)(t-t_0)} & \left(\frac{-k-c}{2}\right) e^{\left(\frac{-k-c}{2}\right)(t-t_0)} \end{bmatrix}$$

$$\times \begin{bmatrix} \frac{k+c}{2c} & \frac{1}{c} \\ \frac{-k+c}{2c} & -\frac{1}{c} \end{bmatrix}$$

$$= \begin{bmatrix} \left[\frac{k+c}{2c} e^{\left(\frac{-k+c}{2}\right)(t-t_0)} + \frac{-k+c}{2c} e^{\left(\frac{-k-c}{2}\right)(t-t_0)}\right] & \left[\frac{1}{c} e^{\left(\frac{-k+c}{2}\right)(t-t_0)} - \frac{1}{c} e^{\left(\frac{-k-c}{2}\right)(t-t_0)}\right] \\ \left[\frac{c^2-k^2}{4c} e^{\left(\frac{-k+c}{2}\right)(t-t_0)} + \frac{k^2-c^2}{4c} e^{\left(\frac{-k-c}{2}\right)(t-t_0)}\right] & \left[\frac{-k+c}{2c} e^{\left(\frac{-k+c}{2}\right)(t-t_0)} + \frac{k+c}{c} e^{\left(\frac{-k-c}{2}\right)(t-t_0)}\right] \end{bmatrix}$$

Hence the solution of the homogeneous equation passing through the initial point is given by

$$
\begin{aligned}
& x(t) = \\
& x(t_0)\left[\frac{k+c}{2c}e^{\left(\frac{-k+c}{2}\right)(t-t_0)} + \frac{-k+c}{2c}e^{\left(\frac{-k-c}{2}\right)(t-t_0)}\right] \\
& + \dot{x}(t_0)\left[\frac{1}{c}e^{\left(\frac{-k+c}{2}\right)(t-t_0)} - \frac{1}{c}e^{\left(\frac{-k-c}{2}\right)(t-t_0)}\right] \\
& = \\
& e^{-\frac{k}{2}(t-t_0)}\left[x(t_0)\left\{\frac{k+c}{2c}e^{\left(\frac{c}{2}\right)(t-t_0)} + \frac{-k+c}{2c}e^{\left(\frac{-c}{2}\right)(t-t_0)}\right\}\right] \\
& + e^{-\frac{k}{2}(t-t_0)}\left[\dot{x}(t_0)\left\{\frac{1}{c}e^{\frac{c}{2}(t-t_0)} - \frac{1}{c}e^{-\frac{c}{2}(t-t_0)}\right\}\right] \\
= \; & e^{-\frac{k}{2}(t-t_0)}\left[e^{\frac{c}{2}(t-t_0)}\left\{\frac{k+c}{2c}x(t_0) + \frac{1}{c}\dot{x}(t_0)\right\}\right] \\
& + e^{-\frac{k}{2}(t-t_0)}\left[e^{\frac{-c}{2}(t-t_0)}\left\{\frac{-k+c}{2c}x(t_0) - \frac{1}{c}\dot{x}(t_0)\right\}\right] \\
= \; & e^{-\frac{k}{2}(t-t_0)}\left[\alpha e^{\frac{c}{2}(t-t_0)} + \beta e^{\frac{-c}{2}(t-t_0)}\right], \\
& \alpha = \frac{k+c}{2c}x(t_0) + \frac{\dot{x}(t_0)}{c}, \quad \beta = \frac{-k+c}{2c}x(t_0) - \frac{\dot{x}(t_0)}{c}
\end{aligned}
$$

(iii) If $k^2 - 4\omega^2 = 0$*, then the eigenvalues of* A *are repeated* $\frac{-k}{2}, \frac{-k}{2}$*. So, the transition matrix is given by*

$$
\phi(t-t_0) = \begin{bmatrix} e^{\left(\frac{-k}{2}\right)(t-t_0)} & 0 \\ 0 & e^{\left(\frac{-k}{2}\right)(t-t_0)} \end{bmatrix}[I + N(t-t_0)]
$$

where N is the nilpotent matrix given by

$$\begin{aligned} N &= A - \begin{bmatrix} -\frac{k}{2} & 0 \\ 0 & -\frac{k}{2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{k}{2} & 1 \\ -\omega^2 & -\frac{k}{2} \end{bmatrix} \end{aligned}$$

This gives

$$\phi(t,t_0) = \begin{bmatrix} e^{\left(\frac{-k}{2}\right)(t-t_0)} & 0 \\ 0 & e^{\left(\frac{-k}{2}\right)(t-t_0)} \end{bmatrix} \times \begin{bmatrix} 1+\frac{k}{2}(t-t_0) & (t-t_0) \\ -\omega^2(t-t_0) & 1-\frac{k}{2}(t-t_0) \end{bmatrix}$$

$$= \begin{bmatrix} \left[1+\frac{k}{2}(t-t_0)\right] e^{\left(\frac{-k}{2}\right)(t-t_0)} & (t-t_0)e^{\left(\frac{-k}{2}\right)(t-t_0)} \\ -\omega^2(t-t_0)e^{\left(\frac{-k}{2}\right)(t-t_0)} & \left[1-\frac{k}{2}(t-t_0)\right] e^{\left(\frac{-k}{2}\right)(t-t_0)} \end{bmatrix}$$

Hence, the solution of the homogeneous system is given by

$$\begin{aligned} x(t) &= x(t_0)\left(1+\frac{k}{2}(t-t_0)\right) e^{\left(\frac{-k}{2}\right)(t-t_0)} \\ &\quad + \dot{x}(t_0) e^{\left(\frac{-k}{2}\right)(t-t_0)} (t-t_0) \end{aligned}$$

$$
\begin{aligned}
&= e^{\left(\frac{-k}{2}\right)(t-t_0)} x(t_0) + (t-t_0)e^{\left(\frac{-k}{2}\right)(t-t_0)} \\
&\qquad \left\{\frac{k}{2}x(t_0) + \dot{x}(t_0)\right\} \\
&= \alpha e^{\left(\frac{-k}{2}\right)(t-t_0)} + \beta(t-t_0)e^{\left(\frac{-k}{2}\right)(t-t_0)}
\end{aligned}
$$

where

$$\alpha = x(t_0), \qquad \beta = \frac{k}{2}x(t_0) + \dot{x}(t_0)$$

Example 4.4.7 *Let us consider the linearized satellite problem, discussed earlier in Chapters 1 and 2,*

$$\frac{d\bar{x}}{dt} = A\bar{x}(t) + B\bar{u}(t)$$

where $\bar{x} = (x_1, x_2, x_3, x_4)$, $\bar{u} = (u_1, u_2)$,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

To compute the transition matrix of the above problem, we note that A has eigenvalues $0, 0, \pm\omega i$.
We follow the procedure described in Theorem 4.4.3.
The generalized eigenvectors corresponding to the eigenvalue $\lambda = 0$ *are* $(1, 0, 0, -\frac{3}{2}\omega), (0, 0, 1, 0)$ *and the eigenvectors corresponding to the complex eigenvalues* $\lambda = \pm i\omega$ *are* $(1, 0, 0, 2\omega)$, $(0, \omega, 2, 0)$.
This gives

$$P = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & \omega \\ 0 & 1 & 0 & 2 \\ -\frac{3}{2}\omega & 0 & -2\omega & 0 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 4 & 0 & 0 & \frac{2}{\omega} \\ 0 & -\frac{2}{\omega} & 1 & 0 \\ -3 & 0 & 0 & -\frac{2}{\omega} \\ 0 & \frac{1}{\omega} & 0 & 0 \end{bmatrix}$$

and hence the nilpotent matrix $N = A - S$, *where*

$$
\begin{aligned}
S &= P \ diag[B_j] P^{-1} \\
&= \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & \omega \\ 0 & 1 & 0 & 2 \\ -\frac{3}{2}\omega & 0 & 2\omega & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega \\ 0 & 0 & -\omega & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 & \frac{2}{\omega} \\ 0 & -\frac{2}{\omega} & 1 & 0 \\ -3 & 0 & 0 & -\frac{2}{\omega} \\ 0 & \frac{1}{\omega} & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 6\omega & 0 & 0 & 4 \\ 0 & -2\omega & 0 & 0 \end{bmatrix}
\end{aligned}
$$

That is

$$N = A - S = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -6\omega & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

N is nilpotent matrix of order 2. So, the transition matrix $\phi(t, t_0)$ is given by

$$\phi(t, t_0)$$

$$=$$

$$\left\{ P \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\omega(t-t_0) & \sin\omega(t-t_0) \\ 0 & 0 & -\sin\omega(t-t_0) & \cos\omega(t-t_0) \end{bmatrix} P^{-1} \right\} [I + N(t - t_0)]$$

$$=$$

$$\begin{bmatrix} 4 - 3\cos\omega(t-t_0) & \dfrac{\sin\omega(t-t_0)}{\omega} & 0 & \frac{2}{\omega}(1 - \cos\omega(t-t_0)) \\ 3\omega\sin\omega(t-t_0) & \cos\omega(t-t_0) & 0 & 2\sin\omega(t-t_0) \\ 6\sin\omega(t-t_0) & \dfrac{-2}{\omega}(1-\cos\omega(t-t_0)) & 1 & \dfrac{4}{\omega}\sin\omega(t-t_0) \\ 6\omega(-1+\cos\omega(t-t_0)) & -2\sin\omega(t-t_0) & 0 & -3+4\cos\omega(t-t_0) \end{bmatrix}$$

$$\times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -6\omega(t-t_0) & 0 & 1 & -3(t-t_0) \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$
= \begin{bmatrix}
4-3\cos\omega(t-t_0) & \dfrac{\sin\omega(t-t_0)}{\omega} & 0 & \dfrac{2(1-\cos\omega(t-t_0))}{\omega} \\
3\omega\sin\omega(t-t_0) & \cos\omega(t-t_0) & 0 & 2\sin\omega(t-t_0) \\
[6(-\omega)(t-t_0) + 6\sin\omega(t-t_0))] & -\left[\dfrac{2}{\omega} - \dfrac{2}{\omega}\cos\omega(t-t_0)\right] & 1 & -[3(t-t_0) - \dfrac{4}{\omega}\sin\omega(t-t_0)] \\
[-6\omega + 6\omega\cos\omega(t-t_0)] & -2\sin\omega(t-t_0) & 0 & -[3 - 4\cos\omega(t-t_0)]
\end{bmatrix}
$$

4.5 Euler's Linear System

The equation of the form

$$\frac{d\bar{x}}{dt} = \frac{1}{t}A\bar{x} \tag{4.5.1}$$

where A is an autonomous matrix, is called Euler's system.
We make a change variable of the form $t = e^s$ in the independent variable, so that Eq. (4.5.1) reduces to a linear sytem of the form

$$\frac{d\bar{x}}{dt} = A\bar{x} \tag{4.5.2}$$

One can now make use of the linear system theory to compute the solution of Eq. (4.5.2) and hence that of Eq. (4.5.1)

Example 4.5.1 *Solve*

$$t^2\frac{d^2x}{dt^2} - 2t\frac{dx}{dt} + 2x = 0 \tag{4.5.3}$$

Using the notation $x_1 = x$ *and* $x_2 = t\dfrac{dx}{dt}$, *we get*

$$\begin{aligned}
\frac{dx_1}{dt} &= \frac{1}{t}x_2 \\
\frac{dx_2}{dt} &= -\frac{2}{t}x_1 + \frac{3}{t}x_2
\end{aligned}$$

So, we get the Euler's system

$$\frac{d\bar{x}}{dt} = \frac{1}{t}\begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}\bar{x} \tag{4.5.4}$$

Making the change of variable $t = e^s$ we get

$$\frac{d\bar{x}}{ds} = A\bar{x}(s) \tag{4.5.5}$$

where $A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$ *has eigenvalues 1,2 with eigenvectors* $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. *Hence a fundamental matrix of the system given by Eq. (4.5.5) is*

$$\begin{aligned} \phi(s) &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} e^s & 0 \\ 0 & e^{2s} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2e^s - e^{2s} & 2e^s - 2e^{2s} \\ -e^s + e^{2s} & -e^s + 2e^{2s} \end{pmatrix} \end{aligned}$$

Hence the solution of Eq. (4.5.4) is given by

$$\bar{x}(t) = \begin{pmatrix} c_0(2t - t^2) + c_1(2t - t^2) \\ \\ c_0(-t + t^2) + c_1(-t + t^2) \end{pmatrix}$$

Example 4.5.2 *Solve*

$$t^3 \frac{d^3x}{dt^3} + t\frac{dx}{dt} - x = 0 \tag{4.5.6}$$

Using the notation $x_1 = x$, $x_2 = t\dfrac{dx}{dt}$, $x_3 = t^2\dfrac{d^2x}{dt^2}$, *we get that Eq. (4.5.6) is equivalent to the Euler's system*

$$\frac{d\bar{x}}{dt} = \frac{1}{t} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix} \bar{x} \tag{4.5.7}$$

$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix}$ *has 1 as eigenvalues of multiplicity 3.*
We use Theorem 4.4.3 to compute a fundamental matrix for the system

$$\frac{d\bar{x}}{dt} = A\bar{x}(s), \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix} \tag{4.5.8}$$

$$S = I, \quad N = A - S = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

N is nilpotent of order 3 and hence

$$\begin{aligned}
\phi(s) &= exp(s)\left[I + Ns + N^2s^2\right] \\
&= exp(s)\left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + s\begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}\right. \\
&\quad \left. + s^2\begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\right] \\
&= \exp(s)\left[\begin{pmatrix} 1-s+s^2 & s-s^2 & s^2 \\ s^2 & 1-s^2 & s+s^2 \\ s & -s & 1+s \end{pmatrix}\right]
\end{aligned}$$

Hence the solution of the Euler's system Eq. (4.5.7) is given by

$$\bar{x}(t) = t\begin{bmatrix} 1 - lnt + (lnt)^2 & lnt - (lnt)^2 & (lnt)^2 \\ (lnt)^2 & 1 - (lnt)^2 & lnt + (lnt)^2 \\ lnt & -lnt & 1 + lnt \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

This gives the solution of Eq. (4.5.6) as

$$\begin{aligned}
x(t) &= c_1t\left(1 - lnt + (lnt)^2\right) + c_2\left(lnt - (lnt)^2\right) + c_3t(lnt)^2 \\
&= c_1t + tlnt(c_2 - c_1) + t(lnt)^2(c_1 - c_2 + c_3)
\end{aligned}$$

or

$$x(t) = d_1\, t + d_2\, t\, lnt + d_3\, t\, (lnt)^2$$

4.6 Exercises

1. If $A = \begin{bmatrix} 0 & 1 \\ -1 & -2\delta \end{bmatrix}$, show that

$$e^{At} = \begin{bmatrix} e^{-\delta t}\left(\cos\omega t + \dfrac{\delta}{\omega}\sin\omega t\right) & \dfrac{1}{\omega}e^{-\delta t}\sin\omega t \\ -\dfrac{1}{\omega}e^{-\delta t}\sin\omega t & e^{-\delta t}\left(\cos\omega t - \frac{\delta}{\omega}\sin\omega t\right) \end{bmatrix}$$

where $\omega = \sqrt{1-\delta^2}, \quad \delta \leq 1.$

2. Let A and B be two $n \times n$ matrices given by

$$A = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

   Show that

   (a) $B^n = 0$.

   (b) $(\lambda I)B = B(\lambda)I$.

   (c) $e^{At} = e^{\lambda t}\left[I + tB + \frac{t^2}{2!}B^2 + \cdots + \frac{1}{n!}t^{n-1}B^{n-1}\right]$

3. Find two linearly independent solutions of the system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t)$$

   where

$$A(t) = \begin{bmatrix} 0 & 1 \\ -\frac{1}{t^2} & -\frac{1}{t} \end{bmatrix}, \quad t \neq 0$$

   Also, compute the Wronskian of the linearly independent solutions.

4. Compute the transition matrix of the system

$$\frac{d\bar{x}}{dt} = \begin{bmatrix} 0 & e^t \\ 0 & 0 \end{bmatrix} \bar{x}(t)$$

   by using Peano-Baker series.

5. Verify that

$$\phi(t) = \begin{bmatrix} e^{-2t}\cos t & -\sin t \\ e^{-2t}\sin t & \cos t \end{bmatrix}$$

   is a fundamental matrix of the system

$$\frac{d\bar{x}}{dt} = \begin{bmatrix} -2\cos^2 t & -1 - \sin 2t \\ 1 - \sin 2t & -2\sin t \end{bmatrix} \bar{x}(t)$$

6. Solve the non-homogeneous system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + \bar{b}(t), \quad \bar{x}(0) = 0$$

where $a(t)$ is an given in Problem 5 and $\bar{b}(t) = (1, e^{-2t})$.

7. Compute the transition matrix of the following system

$$\frac{d\bar{x}}{dt} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix} \bar{x}(t)$$

8. Solve the initial value problem

$$\begin{aligned} \frac{d\bar{x}}{dt} &= A\bar{x}(t) \\ \bar{x}(0) &= (1, t) \end{aligned}$$

where

(a) $A = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 2 & 0 & 1 & 0 \end{bmatrix}$

(b) $A = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

9. Show that the transition matrices corresponding to autonomous systems described by two similar matrices are necessarily similar.

10. Verify the following properties of the transition matrix $\phi(t, t_0)$ of the system described by Problem 7.

    (a) $[\phi(t, s)]^{-1} = \phi(s, t)$
    (b) $[\phi(t, t_0)\phi(t_0, s)] = \phi(t, s)$
    (c) $\phi(t, t) = I$

11. (a) Find two distinct fundamental matrices of the system

$$\frac{d\bar{x}}{dt} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \bar{x}(t)$$

(b) Find a fundamental matrix of the adjoint system of the above.

References

[1] Ravi P Agarwal and Ramesh C Gupta, *Essentials of Ordinary Differential Equatoins,McGraw Hill Book Company, 1991*

[2] Roger W Brockett, *Finite Dimensional Linear System, John Wiley and Sons, Inc, 1970*

[3] Earl A Coddington and Norman Levinson, *Theory of Ordinary Differential Equations, McGraw Hill, 1955*

[4] Lawrence Perko, *Differential Equations and Dynamical Systems, Springer Verlag, 1991*

Chapter 5

Stability Analysis

In this chapter we utilize the theory of transition matrix discussed in Chapter 4, to study boundedness property of solutions of both linear and nonlinear systems along with the stability of equilibrium points.
Various notions of stability including the one due to Lyapunov are introduced and examined leading to some interesting theorems.
Phase portraits, around the equilibrium points in two dimensions, are discussed in detail. This gives rise to the notion of saddle point, stable and unstable node, focus and centre. We also sketch phase portraits, vector fields and solution curves of some illustrative examples.

5.1 Asymptotic Behaviour of Linear System

Theorem 4.4.1 - 4.4.3 can be completely described in terms of the following theorem, known as Jordan canonical form of a matrix, which can be used in investigating the behaviour of solution of a linear system of the form given by Eq. (4.1.1) for large values of t.

Theorem 5.1.1 *(Jordan canonical form)*
Let A be a real matrix with real eigenvalues $\lambda_j,\ 1 \le j \le k$, and complex eigenvalues $\lambda_j = a_j \pm ib_j,\ k+1 \le j \le n$.
Then there exists a basis $\{\bar{v}_1, ..., \bar{v}_k, \bar{v}_{k+1}, \bar{u}_{k+1}, ..., \bar{v}_n, \bar{u}_n\}$ of $\Re^{2n-k}$ where $\bar{v}_j$ are generalized eigenvectors of A corresponding to real eigenvalues λ_j $(1 \le j \le k)$ and $\bar{w}_j$ are generalized eigenvectors of A corresponding to complex eigenvalues λ_j $(k+1 \le j \le n)$ with $\bar{u}_j = Re(\bar{w}_j)$ and $\bar{v}_j = Im(\bar{w}_j)$. Let $P = [\bar{v}_1, ..., \bar{v}_k, \bar{v}_{k+1}, \bar{u}_{k+1}, ..., \bar{v}_n, \bar{u}_n]$.

Then P is invertible and

$$P^{-1}AP = \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_r \end{bmatrix} \tag{5.1.1}$$

The elementary Jordan blocks $B = B_j,\ j = 1, 2, ..., r$ are either of the form

$$B = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ \vdots & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \qquad (*)$$

for λ as one of the real eigenvalues of A or of the form

$$B = \begin{bmatrix} D & I_2 & & 0 \\ & D & I_2 & \\ & & \ddots & I_2 \\ 0 & & & D \end{bmatrix} \qquad (**)$$

with

$$D = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \qquad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

for $\lambda = a \pm ib$ as a pair of complex eigenvalues of A.

This gives us the solvability of the linear system

$$\begin{aligned} \frac{d\bar{x}}{dt} &= A\bar{x}(t) \\ \bar{x}(t_0) &= \bar{x}_0 \end{aligned} \qquad (5.1.2)$$

Theorem 5.1.2 *The solution $\bar{x}(t)$ of the system given by Eq. (5.1.2) is of the form*

$$x(t) = \phi(t, t_0)\bar{x}_0$$

where $\phi(t, t_0) = \phi(t - t_0) = P\ diag\ [\exp B_j(t - t_0)]\ P^{-1}$ is the transition matrix of the system.

(a) If $B_j = B$ is an $m \times m$ matrix of the form (*), then $B = \lambda I + N$ and

$$\begin{aligned} \exp B(t - t_0) &= \exp \lambda(t - t_0) \exp N(t - t_0) \\ &= \exp \lambda(t - t_0) \times \\ & \begin{bmatrix} 1 & t - t_0 & \frac{(t-t_0)^2}{2!} & \cdots & \frac{(t-t_0)^{m-1}}{(m-1)!} \\ 0 & 1 & t - t_0 & \cdots & \frac{(t-t_0)^{m-2}}{(m-2)!} \\ \vdots & & \ddots & & \cdots \\ 0 & 0 & & \cdots & (t - t_0) \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix} \end{aligned}$$

(b) If $B_j = B$ is $2m \times 2m$ matrix of the form (**) and $\lambda = a \pm ib$ is a complex pair of eigenvalues, then

$$\exp B(t-t_0) = \exp a(t-t_0) \times \begin{bmatrix} R & R(t-t_0) & R\frac{(t-t_0)^2}{2!} & \cdots & R\frac{(t-t_0)^{m-1}}{(m-1)!} \\ 0 & R & R(t-t_0) & \cdots & R\frac{(t-t_0)^{m-2}}{(m-2)!} \\ \vdots & & \ddots & & \\ 0 & 0 & \cdots & R & R(t-t_0) \\ 0 & 0 & \cdots & 0 & R \end{bmatrix}$$

where $R = \begin{bmatrix} \cos b(t-t_0) & \sin b(t-t_0) \\ -\sin b(t-t_0) & \cos b(t-t_0) \end{bmatrix}$

Corollary 5.1.1 *Each component of the solution of the linear system Eq. (5.1.2) is a linear combination of functions of the form*

$$(t-t_0)^l \exp a(t-t_0) \cos b(t-t_0) \ \ or \ \ (t-t_0)^l \exp a(t-t_0) \sin b(t-t_0),$$

where $\lambda = a \pm ib$ is a pair of eigenvalues of the matrix A and l is smaller than the dimension of the space.

Let multiplicity of each eigenvalues λ_j be p_j where $\lambda_j = a_j + ib_j$. Let

$$a = \max_{1 \le j \le k} a_j \text{ and } p = \max_{1 \le j \le l} p_j$$

Then by Corollary (5.1.1), it follows that $\exists\, t_1 \ge 0$ such that

$$\|\phi(t)\| = \|\exp(At)\| \le c\ \exp(at)\ t^p \qquad \forall\, t \ge t_1$$

where c is some suitable constant.

Corollary 5.1.2 *Every solution of the linear system given by Eq. (5.1.2) tends to zero as $t \to \infty$ iff the real parts of the eigenvalues of the matrix A are negative.*

Corollary 5.1.3 *Every solution of Eq. (5.1.2) is bounded iff the real parts of the multiple eigenvalues of the matrix A are negative and real parts of the simple eigenvalues of A are non positive.*

Theorem 5.1.1 gives us the following well-known theorem called Cayley-Hamilton theorem.

Theorem 5.1.3 *Let $p(\lambda) = det(\lambda I - A)$, $A \in \Re^n$. Then $p(A) = 0$. That is, if $det(\lambda I - A) = \lambda^n + p_{n-1}\lambda^{n-1} + \cdots + p_0$, then $A^n + p_{n-1}A^{n-1} + \cdots + p_0 I = 0$.*

For the proof of the Theorem 5.1.1 and Theorem 5.1.3, refer Gantmacher[5]

We shall now investigate the following perturbed linear system

$$\frac{d\bar{x}}{dt} = A\bar{x} + B(t)\bar{x} \tag{5.1.3}$$

where $B(t)$ is a continuous matrix on the interval $[t_0, \infty)$. We wish to examine the relationship between the boundedness of solutions of Eq. (5.1.2) and Eq. (5.1.3).

Let $\phi(t, t_0) = \exp(A(t - t_0))$ be the transition matrix of the the linear system given by Eq. (5.1.2). By variation of parameter formula, the solvability of Eq. (5.1.3) is given by

$$\bar{x}(t) = \phi(t, t_0)\bar{x}_0 + \int_{t_0}^{t} \phi(t, s)B(s)\bar{x}(s)ds \tag{5.1.4}$$

We first prove the following variation of Gronwall's lemma.

Lemma 5.1.1 *(Generalized Gronwall's inequality)*
Let $u(t)$, $p(t)$, $q(t)$ be nonnegative continuous functions in the interval $I = [t_0, t_f]$ and let

$$u(t) \le p(t) + \int_{t_0}^{t} q(s)u(s)ds, \quad t \in I \tag{5.1.5}$$

Then we have

$$u(t) \le p(t) + \int_{t_0}^{t} p(\tau)q(\tau)\exp\left(\int_{\tau}^{t} q(s)ds\right)d\tau, \quad t \in I \tag{5.1.6}$$

Proof : Let

$$r(t) = \int_{t_0}^{t} q(s)u(s)ds$$

Then $r(t_0) = 0$ and $\dot{r}(t) = q(t)u(t)$.

By assumption, we have $u(t) \le p(t) + r(t)$ and hence

$$\dot{r}(t) = q(t)u(t) \le p(t)q(t) + r(t)q(t)$$

This gives

$$\dot{r}(t) - r(t)q(t) \le p(t)q(t)$$

and hence multiplying both side by $\exp\left(-\int_{t_0}^{t} q(s)ds\right)$, we get

$$\frac{d}{dt}\left[r(t)\exp\left(-\int_{t_0}^{t} q(s)ds\right)\right] \le \exp\left(-\int_{t_0}^{t} q(s)ds\right)p(t)q(t)$$

which implies that

$$\exp\left(-\int_{t_0}^{t} q(s)ds\right)r(t) \le \int_{t_0}^{t} p(\tau)q(\tau)\exp\left(-\int_{t_0}^{\tau} q(s)ds\right)d\tau$$

This gives

$$\begin{aligned}
r(t) &\le \int_{t_0}^{t} p(\tau)q(\tau)\left[\exp\left(\int_{\tau}^{t_0} q(s)ds + \int_{t_0}^{t} q(s)ds\right)\right] d\tau \\
&= \int_{t_0}^{t} q(\tau)p(\tau)\exp\left(\int_{\tau}^{t} q(s)ds\right) d\tau
\end{aligned}$$

and hence

$$u(t) \le p(t) + \int_{t_0}^{t} q(\tau)p(\tau)\exp\left(\int_{\tau}^{t} q(s)ds\right) d\tau$$

∎

Corollary 5.1.4 *If $p(t)$ is nondecreasing on I, then*

$$u(t) \le p(t)\exp\left(\int_{t_0}^{t} q(s)ds\right) \tag{5.1.7}$$

Proof : As $p(t)$ is nondecreasing, (5.1.6) gives

$$\begin{aligned}
u(t) &\le p(t)\left[1 + \int_{t_0}^{t} q(\tau)\exp\left(\int_{\tau}^{t} q(s)ds\right) d\tau\right] \\
&= p(t)\left[1 - \int_{t_0}^{t} \frac{d}{d\tau}\exp\left(\int_{\tau}^{t} q(s)ds\right) d\tau\right] \\
&= p(t)\left[\exp\left(\int_{t_0}^{t} q(s)ds\right)\right]
\end{aligned}$$

∎

We are now ready to state and prove the boundedness of solutions of the perturbed system given by Eq. (5.1.3).

Theorem 5.1.4 *Let all solutions of Eq. (5.1.2) be bounded on $[t_0, \infty)$. Then all solutions of Eq. (5.1.3) are also bounded provided*

$$\int_{t_0}^{\infty} \|B(t)\| \le \infty \tag{5.1.8}$$

Proof : As all solutions of Eq. (5.1.2) are bounded, it follows that $\exists\, c$ such that

$$\|\exp(At)\| \ \le\ c \quad \forall\, t \in\ I = [t_0, \infty).$$

From Eq. (5.1.4), we have

$$\begin{aligned}
\|\bar{x}(t)\| &\le \|\exp(A(t-t_0))\|\ \|\bar{x}_0\| + \int_{t_0}^{t} \|B(s)\exp(A(t-s))\|\ \|\bar{x}(s)\| ds \\
&\le c\|\bar{x}_0\| + c\int_{t_0}^{t} \|B(s)\|\ \|\bar{x}(s)\|\ ds \qquad (5.1.9)
\end{aligned}$$

Applying the generalized Gronwall's inequality given by Eq. (5.1.7) to the above inequality we have

$$\begin{aligned}\|\bar{x}(t)\| &\leq c\|\bar{x}_0\| \left[\exp\left(c\int_{t_0}^{t}\|B(s)\|ds\right)\right] \\ &\leq c\|\bar{x}_0\|\exp(M) \qquad \forall\, t \in I\end{aligned}$$

where

$$M = \exp\left(c\int_{t_0}^{\infty}\|B(s)\|ds\right) < \infty$$

This implies that Eq. (5.1.3) has bounded solutions. ∎

Remark 5.1.1 *We examine the boundedness of solutions of the non-homogeneous system*

$$\frac{d\bar{x}}{dt} = A\bar{x}(t) + \bar{b}(t) \tag{5.1.10}$$

under the assumption that

(a) Eq. (5.1.2) has bounded solutions and

(b) $\int_{t_0}^{\infty}\|\bar{b}(s)\|ds < \infty$.

One can show that the solutions of Eq. (5.1.10) remains bounded. The proof is similar to that of Theorem 5.1.3. The inequality given by Eq. (5.1.9) in the proof of Theorem 5.1.3 will be replaced by the following inequality

$$\|\bar{x}(t)\| \leq c\|\bar{x_0}\| + c\int_{t_0}^{t}\|\bar{b}(s)\|ds$$

The following theorem gives asymptotic convergence of solutions of Eq. (5.1.3) vis-a-vis asymptotic convergence of solutions of Eq. (5.1.2).

Theorem 5.1.5 *Let all solutions of Eq. (5.1.2) be such that they tend to zero as* $t \to \infty$. *Then all solutions of Eq. (5.1.3) are also asymptotically convergent to zero, provided*

$$\|B(t)\| \to 0 \text{ as } t \to \infty \tag{5.1.11}$$

Proof : By Corollary 5.1.2, it follows that all eigenvalues of A have negative real parts and hence there exist constants c and $\delta > 0$ such that

$$\|\exp(At)\| \leq c\exp(-\delta t) \quad \forall\, t \geq t_0$$

Using this inequality in Eq. (5.1.4) we get,

$$\|\bar{x}(t)\| \leq c\exp(-\delta(t-t_0))\,\|\bar{x}_0\| + c\int_{t_0}^{t}\exp(-\delta(t-s))\,\|B(s)\|\,\|\bar{x}(s)\|ds$$

In view of Eq. (5.1.11), given any $c_1 > 0$, $\exists\, t_1 \geq t_0 > 0$ such that

$$\|B(t)\| \leq c_1 \text{ for } t \geq t_1$$

and hence we get

$$\begin{aligned}
\|\bar{x}(t)\| \ \exp(\delta t) &\leq c\exp(\delta t_0)\ \|\bar{x}_0\| + c\int_{t_0}^{t_1} \exp(\delta s)\ \|B(s)\|\ \|\bar{x}(s)\|\ ds \\
&+ \int_{t_1}^{t} \|B(s)\| \exp(\delta s)\ \|\bar{x}(s)\|\ ds \\
&\leq c\exp(\delta t_0)\ \|\bar{x}_0\| + c\int_{t_0}^{t_1} \exp(\delta s)\ \|B(s)\|\ \|\bar{x}(s)\|\ ds \\
&+ c\int_{t_1}^{t} c_1 \exp(\delta s)\ \|\bar{x}(s)\|\ ds.
\end{aligned}$$

Let $u(t) = \exp(\delta t)\ ||\bar{x}(t)||$ and C denote the first two terms of RHS of the above inequality:

$$C = c\ \exp(\delta t_0)\ ||\bar{x}_0|| + c\int_{t_0}^{t_1} \exp(\delta s)\ ||B(s)||\ ||\bar{x}(s)||ds$$

Then we get,

$$u(t) \leq C + c\,c_1 \int_{t_1}^{t} u(s)ds\ , \quad t\ \in\ I$$

By Gronwall's lemma we get

$$\begin{aligned}
u(t) &\leq C\ \exp\Big(\int_{t_1}^{t} c\,c_1 ds\Big) \\
&= C\exp(c\,c_1(t-t_1)) \quad \forall\, t\ \in\ I
\end{aligned}$$

This gives

$$\begin{aligned}
||\bar{x}(t)|| &\leq \exp(-\delta t)C\ \exp(c\,c_1(t-t_1)) \\
&= C\ \exp(t(c\,c_1\ -\ \delta)\ -\ c\,c_1 t_1), \quad \forall\, t\ \in\ I
\end{aligned}$$

As c_1 is arbitrary, we choose c_1 such that $cc_1 < \delta$ and hence $\exp(t(c\,c_1\ -\ \delta)) \rightarrow 0$ as $t \rightarrow \infty$. This gives that $\|\bar{x}(t)\| \rightarrow 0$ as $t \rightarrow \infty$.
This proves the theorem. ∎

Remark 5.1.2 *If Eq. (5.1.2) has solutions asymptoticaly convergent to zero, then the solutions of non-homogeneous system given by Eq. (5.1.10) satisfies the same property if* $\|\bar{b}(t)\| \mapsto 0$ *as* $t \mapsto \infty$.

We now consider the boundedness and stability of solution of non-autonomous linear system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) \tag{5.1.12}$$

Observe that,

$$||\bar{x}(t)||^2 = (\bar{x}(t), \bar{x}(t))$$

Differentiating the above relation, we get

$$\begin{aligned}
\frac{d}{dt}\left(||\bar{x}(t)||^2\right) &= \left(\frac{d\bar{x}(t)}{dt}, \bar{x}(t)\right) + \left(\bar{x}(t), \frac{d\bar{x}(t)}{dt}\right) \\
&= (A(t)\bar{x}(t), \bar{x}(t)) + (\bar{x}(t), A(t)\bar{x}(t)) \\
&= \left((A(t) + [A(t)]^\top)\bar{x}(t), \bar{x}(t)\right)
\end{aligned}$$

$A(t) + [A(t)]^\top$ is a symmetric matrix and hence its eigenvalues are real. Let $M(t)$ be the largest eigenvalue of $A(t) + [A(t)]^\top$. We have the following theorem

Theorem 5.1.6 *(a) If $\int_{t_0}^{\infty} M(s)ds$ is bounded then Eq. (5.1.12) has bounded solutions.*

(b) If $\int_{t_0}^{\infty} M(s)ds = -\infty$, then Eq. (5.1.12) has solutions asymptotically converging to zero.

Proof : We have

$$\begin{aligned}
\frac{d}{dt}\left(||\bar{x}(t)||^2\right) &\leq \left((A(t) + [A(t)]^\top \bar{x}(t), \bar{x}(t)\right) \\
&\leq M(t)\ ||\bar{x}(t)||^2
\end{aligned}$$

This implies that

$$||\bar{x}(t)||^2 \leq ||\bar{x}(0)||^2 \exp\left(\int_{t_0}^{t} M(s)ds\right),\ t \in I$$

If $\int_{t_0}^{\infty} M(s)ds < \infty$, then we get **(a)**. If $\lim\limits_{t\to\infty} \left[\int_{t_0}^{t} M(s)ds\right] = -\infty$ then we get **(b)**. ∎

Example 5.1.1

$$\frac{d\bar{x}}{dt} = [A + B(t)]\,\bar{x}(t)$$

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \text{ and } B(t) = \begin{bmatrix} 0 & 0 \\ 0 & \dfrac{2a}{at+b} \end{bmatrix}$$

One can compute that the fundamental matrix of the unperturbed system is $\begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$ *which is bounded.*

The perturbed system has the fundamental matrix as

$$\begin{bmatrix} a\sin t - (at+b)\cos t & a\cos t + (at+b)\sin t \\ (at+b)\sin t & (at+b)\cos t \end{bmatrix}$$

which is unbounded. We note that $\int_0^t ||B(s)||ds \to \infty$ *as* $t \to \infty$.

Example 5.1.2

$$A(t) = \begin{bmatrix} \dfrac{1}{(1+t)^2} & t^2 \\ -t^2 & -1 \end{bmatrix} \text{ and } A(t) + A(t)^\top = \begin{bmatrix} \dfrac{2}{(1+t)^2} & 0 \\ 0 & -2 \end{bmatrix}$$

$M(t) = \dfrac{2}{(1+t)^2}$ *and this gives* $\int_0^\infty M(s)ds = 2$.

Hence by Theorem 5.1.5 the non-autonomous system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t)$$

has bounded solutions.

We shall now consider boundedness and asymptotic stability of the perturbed system of the form

$$\frac{d\bar{x}}{dt} = [A(t) + B(t)]\,\bar{x}(t) \tag{5.1.13}$$

corresponding to the non-autonomous system given by Eq. (5.1.12).

Assumptions 1 The matrix $A(t)$ corresponding to the system given by Eq. (5.1.12) is such that

(a)

$$\lim_{t \mapsto \infty} \inf \int_{t_0}^{t} TrA(s)ds > -\infty \text{ or } TrA(s) = 0 \tag{5.1.14}$$

(b) The perturbation matrix $B(t)$ satisfies the condition

$$\int_{t_0}^{\infty} ||B(t)||dt < \infty$$

Theorem 5.1.7 *Under* **Assumptions 1**, *if all solution of Eq. (5.1.12) are bounded, then all solutions of Eq. (5.1.13) are also bounded.*

Proof : Let $\psi(t)$ be a fundamental matrix of the system Eq. (5.1.12). Then, the boundedness of solutions implies that $\psi(t)$ is bounded. Also, by Abel's lemma (Theorem 4.1.4), we have

$$det[\psi(t)] = det[\psi(t_0)]exp(\int_{t_0}^{t} TrA(s)ds)$$

and hence

$$\psi^{-1}(t) = \frac{adj\psi(t)}{det\psi(t)} = \frac{adj\psi(t)}{det[\psi(t_0)]exp(\int_{t_0}^{t} TrA(s)ds)} \tag{5.1.15}$$

In view of Assumptions 1(a), it follows that $||\psi^{-1}(t)||$ is bounded.
We now apply the variation of parameter formula to the IVP corresponding to Eq. (5.1.13), to get the following integral represntation of the solution $\bar{x}(t)$.

$$\bar{x}(t) = \psi(t)\psi^{-1}(t_0)\bar{x_0} + \int_{t_0}^{t} \psi(t)\psi^{-1}(s)B(s)x(s)ds \tag{5.1.16}$$

Define

$$c = max\left\{sup_{t\geq t_0}||\psi(t)||, sup_{t\geq t_0}||\psi^{-1}(t)||\right\}, \quad c_0 = c||\bar{x_0}||.$$

Then Eq. (5.1.16) gives

$$||\bar{x}(t)|| \leq c_0 + c^2\int_{t_0}^{t} ||B(s)||||\bar{x}(s)||ds$$

and hence Gronwall's inequality gives

$$||\bar{x}(t)|| \leq c_0 exp\left(c^2\int_{t_0}^{t} ||B(s)||ds\right) < \infty$$

This proves the theorem. ∎

Theorem 5.1.8 *Under* **Assumptions 1** *, if all solution of system Eq. (5.1.12) are asymptotically convergent to zero, so is the case for all soluiton for Eq. (5.1.13).*

Proof : Asymptotic convergence of solution of Eq. (5.1.12) will imply that $||\psi(t)|| \longrightarrow 0$ as $t \longrightarrow \infty$. Using this fact in Eq. (5.1.15), we get that $||\psi^{-1}(t)|| \mapsto 0$ as $t \longrightarrow \infty$.
Let c be defined as before. Estimating Eq. (5.1.16), we get

$$||\bar{x}(t)|| \leq ||\psi(t)||\left[c_1 + c\int_{t_0}^{t} ||B(s)||||\bar{x}(s)||ds\right]$$

where $c_1 = ||\psi^{-1}(t_0)\bar{x_0}||$. The above inequality gives

$$\frac{||\bar{x}(t)||}{||\psi(t)||} \leq c_1 + c\int_{t_0}^{t} ||B(s)||||\psi(s)|| \left(\frac{||x(s)||}{||\psi(s)||}\right) ds$$

Put $||y(t)|| = \dfrac{||\bar{x}(t)||}{||\psi(t)||}$ and apply Gronwall's inequality to get

$$||y(t)|| \leq c_1 \left[exp \int_{t_0}^{t} c^2||B(s)||ds\right]$$

thereby implying that

$$||\bar{x}(t)|| \leq c_1||\psi(t)|| \left[exp \int_{t_0}^{t} c^2||B(s)||ds\right] \longrightarrow 0 \text{ as } t \longrightarrow \infty$$

∎

These theorems immediately lead to the following result corresponding to the asymptotic stability of the non-homogeneous system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + \bar{b}(t) \tag{5.1.17}$$

Assumptions 2

(a) Same as Assumption 1(a)

(b) The function $\bar{b}(t)$ is such that

$$\int_{t_0}^{t} ||\bar{b}(s)||ds < \infty$$

Theorem 5.1.9 *Let* **Assumptions 2** *hold. Then*

(a) boundedness of solutions of Eq. (5.1.12) will imply boundedness of solutions of Eq. (5.1.17) and

(b) asymptotic convergence of solutions of Eq. (5.1.12) to zero will imply asymptotic convergence of solutions to zero of the non-homogeneous equation Eq. (5.1.17).

If the non-autonomous system under consideration is periodic, we can dispatch with the Assumption 1(a). The boundedness and asymptotic convergence of the perturbed system Eq. (5.1.13) is preserved under Assumption 1(b).

Theorem 5.1.10 *Let $A(t)$ be periodic. Let* **Assumption 1(b)** *hold. Then*

(a) boundedness of solutions of Eq. (5.1.12) implies boundedness of solutions of Eq. (5.1.13) and

(b) *asymptotic convergence of solutions of Eq. (5.1.12) implies asymptotic convergence of solution of Eq. (5.1.13).*

Proof : By Floquet-Lyapunov theorem - Theorem 4.3.1, the transition matrix of the periodic system Eq. (5.1.12) is given by

$$\phi(t,t_0) = P^{-1}(t)e^{R(t-t_0)}P(t_0)$$

where P is periodic nonsingular matrix and R is a fixed martix.
By variation of parameter formula applied to the IVP corresponding to Eq. (5.1.13) we get

$$\begin{aligned}\bar{x}(t) &= P^{-1}(t)e^{R(t-t_0)}P(t_0)\bar{x_0} \\ &+ \int_{t_0}^{t} P^{-1}(t)e^{Rt}e^{-Rs}P(s)B(s)\bar{x}(s)ds\end{aligned}$$

This gives

$$\begin{aligned}\|\bar{x}(t)\| &\leq \|P^{-1}(t)\|\|e^{Rt}\|\|e^{-Rt_0}P(t_0)\bar{x}_0\| \\ &+ \int_{t_0}^{t}\|P^{-1}(t)\|\|e^{R(t-s)}\|\|\|P(s)\|\|\|B(s)\|\|\|\bar{x}(s)\|ds \qquad (5.1.18)\end{aligned}$$

$P(t)$ is periodic and nonsingular and hence $det\ P(t) \neq 0$ and hence both $||P(t)||$ and $||P^{-1}(t)||$ are bounded on $[t_0, \infty)$. Let

$$d = max\left\{\sup_{t\geq t_0}||P(t)||, \quad \sup_{t\geq t_0}||P^{-1}(t)||\right\}$$

Eq. (5.1.18) gives

$$\|\bar{x}(t)\| \leq d_1\|e^{Rt}\| + d^2\int_{t_0}^{t}\|e^{R(t-s)}\|\|\|B(s)\|\|x(s)\|ds \qquad (5.1.19)$$

where $d_1 = d||e^{-Rt_0}P(t_0)\bar{x_0}||$.
Boundedness of solutions of Eq. (5.1.12) implies that $||e^{Rt}|| \leq d_2$ and hence Eq. (5.1.19) gives

$$||\bar{x}(t)|| \leq d_1 d_2 + d^2 d_2\int_{t_0}^{t}||B(s)||||\bar{x}(s)||ds$$

and hence

$$||\bar{x}(t)|| \leq d_1 d_2\left[exp\left(d^2 d_2\int_{t_0}^{t}||B(s)||ds\right)\right]$$

This proves the boundedness of solutions of Eq. (5.1.13).
If solutions of Eq. (5.1.12) are asymptotically convergent to zero, we must have $||e^{Rt}|| \leq d_3 e^{-\alpha t}$ for some positive constant d_3 and α. This inequality immediately gives

$$||\bar{x}(t)|| \leq d_1 d_3 e^{-\alpha t} + d^2 d_3\int_{t_0}^{t}e^{-\alpha(t-s)}||B(s)||||\bar{x}(s)||ds$$

and hence

$$||\bar{x}(t)|| \leq d_1 d_3 e^{-\alpha t}\left[exp\left(d^2 d_3 \int_{t_0}^{t} ||B(s)||ds\right)\right] \longrightarrow 0 \text{ as } t \longrightarrow \infty$$

This proves the theorem. ∎

5.2 Stability Analysis - Formal Approach

We shall formally define various concepts of stability of solution of the initial value problem

$$\frac{d\bar{x}}{dt} = f(t, \bar{x}), \quad \bar{x}(t_0) = \bar{x}_0 \tag{5.2.1}$$

We assume that Eq. (5.2.1) has a global solution on $[t_0, \infty)$.

Definition 5.2.1 *Let $\bar{x}(t, t_0, \bar{x}_0)$ denote the solution of the initial value problem Eq. (5.2.1), indicating its dependence on t and also the initial point t_0 and initial value $\bar{x}_0$. This solution is said to be stable if for a given $\epsilon > 0$, $\exists\ \delta > 0$ such that*

$$||\Delta\bar{x}_0|| < \delta \Rightarrow ||\bar{x}(t, t_0, \bar{x}_0 + \Delta\bar{x}_0) - \bar{x}(t, t_0, \bar{x}_0)|| < \epsilon$$

Definition 5.2.2 *$\bar{x}(t, t_0, \bar{x}_0)$ is said to be asymptotically stable if it is stable and $\exists\ \delta_0 > 0$ such that*

$$||\Delta\bar{x}_0|| < \delta_0 \Rightarrow ||\bar{x}(t, t_0, \bar{x}_0 + \Delta\bar{x}_0) - \bar{x}(t, t_0, \bar{x}_0)|| \rightarrow 0 \ \textit{as} \ t \rightarrow \infty \tag{5.2.2}$$

To be more explicit, Eq. (5.2.2) can be restated as, given $\mu > 0$, there exist δ and $T(\mu)$ such that

$$||\Delta\bar{x}_0|| < \delta \Rightarrow ||\bar{x}(t, t_0, \bar{x}_0 + \Delta\bar{x}_0) - \bar{x}(t, t_0, \bar{x}_0)|| < \mu \ \ \forall\, t \geq t_0 + T(\mu)$$

Definition 5.2.3 *$\bar{x}(t, t_0, \bar{x}_0)$ is said to be uniformly stable if for any $\epsilon > 0$, $\exists\ \delta = \delta(\epsilon) > 0$ such that for any solution $\bar{x}^{(1)}(t, t_0, \bar{x}_1)$ of*

$$\frac{d\bar{x}}{dt} = f(t, \bar{x}), \quad \bar{x}(t_0) = \bar{x_1}$$

$$||\bar{x}(t_1) - \bar{x}^{(1)}(t_1)|| < \delta \ \Rightarrow ||\bar{x}(t) - \bar{x}^{(1)}(t)|| < \epsilon \ \ \forall\, t \geq t_1 \geq t_0$$

Theorem 5.2.1 *For the linear system*

$$\frac{d\bar{x}}{dt} = A(t)x(t) \tag{5.2.3}$$

stability is equivalent to the boundedness of the solution.

Proof : Let $\phi(t, t_0)$ be the transition matrix. Then,

$$||\bar{x}(t, t_0, \bar{x}_0 + \Delta\bar{x}_0) - \bar{x}(t, t_0, \bar{x}_0)|| = ||\phi(t, t_0)\Delta\bar{x}_0|| \le ||\phi(t, t_0)||\ ||\Delta\bar{x}_0||$$

If $\phi(t, t_0)$ is bounded, it follows that the solution $\bar{x}(t, t_0, \bar{x}_0)$ is stable.
Conversely, let the solution of the linear system be stable. Then, given $\epsilon > 0$, $\exists\ \delta > 0$ such that

$$||\bar{x}(t, t_0, \bar{x}_0 + \Delta\bar{x}_0) - \bar{x}(t, t_0, \bar{x}_0)|| = ||\phi(t, t_0)\Delta\bar{x}_0|| < \epsilon$$

whenever $||\Delta\bar{x}_0|| < \delta$.

Choose $\Delta\bar{x}_0 = \frac{\delta}{2}\bar{e}^{(j)}$. Then

$$||\phi(t, t_0)\Delta\bar{x}_0)|| = \frac{\delta}{2}\ ||\phi(t, t_0)\ \bar{e}^{(j)}||\ <\ \epsilon$$

But $\phi(t, t_0)\ \bar{e}^{(j)}$ is the j^{th} column vector $\phi^{(j)}$ of the matrix $\phi(t, t_0)$. Hence we have

$$||\phi(t, t_0)|| = \max_{1\ \le\ j\ \le\ n} ||\phi^{(j)}|| = \max_{1 \le j \le n} ||\phi(t, t_0)\bar{e}^{(j)}|| \le \frac{2\epsilon}{\delta}$$

This gives

$$||\bar{x}(t, t_0, \bar{x}_0)|| = ||\phi(t, t_0)\bar{x}_0||\ \le\ \frac{2\ \epsilon}{\delta}\ ||\bar{x}_0||$$

That is, the solution of the linear system given by Eq. (5.2.3) is bounded. ∎

Corollary 5.2.1 *Let $\phi(t)$ be a fundamental matrix of the system given by Eq. (5.2.3). Then Eq. (5.2.3) has asymptotically stable solutions iff $||\phi(t)|| \to 0$ as $t \to \infty$.*

For the uniform stability of the linear system we have the following theorem

Theorem 5.2.2 *Let $\phi(t)$ be a fundamental matrix of Eq. (5.2.3). Then all solutions of Eq. (5.2.3) are uniformly stable iff there exists a constant $c > 0$ such that*

$$||\phi(t)\phi^{-1}(t_1)|| \le\ c \qquad \forall\ t_0 \le t_1 \le t < \infty \tag{5.2.4}$$

Proof : Let $\bar{x}(t) = \bar{x}(t, t_0, \bar{x_0})$ be a solution of Eq. (5.2.3). Then for any $t_1 \ge t_0$ we have $\bar{x}(t) = \phi(t)\phi^{-1}(t_1)\bar{x}(t_1)$. Let $\bar{x}^{(1)}(t) = \phi(t)\phi^{-1}(t_1)\bar{x}^{(1)}(t_1)$ be any other solution. As Eq. (5.2.4) is satisfied, we have

$$\begin{aligned} ||\bar{x}^{(1)}(t) - \bar{x}(t)|| &\le ||\phi(t)\phi^{-1}(t_1)||||\bar{x}^{(1)}(t_1) - \bar{x}(t_1)|| \\ &\le c||\bar{x}^{(1)}(t_1) - \bar{x}(t_1)||\ \ \forall\ t_0 \le t_1 \le t < \infty \end{aligned}$$

Let $\epsilon > 0$ be given.

$$||\bar{x}^{(1)}(t_1) - \bar{x}(t_1)|| \le \frac{\epsilon}{c} = \delta(\epsilon) > 0 \text{ imply that } ||\bar{x}^{(1)}(t) - \bar{x}(t)|| < \epsilon$$

and hence $\bar{x}(t)$ is uniformly stable.

Conversley, let Eq. (5.2.3) uniformly stable and hence zero solution is uniformly stable. Hence given $\epsilon > 0$, $\exists\ \delta(\epsilon)$ such that $t_1 \geq t_0$ and

$$||\bar{x}^{(1)}(t)|| < \delta \Rightarrow ||\bar{x}^{(1)}(t)|| < \epsilon \ \ \forall\, t \geq t_1.$$

Thus we have $||\phi(t)\phi^{-1}(t_1)\bar{x}^{(1)}(t_1)|| < \epsilon \ \ \forall\, t \geq t_1$.
As in Theorem 5.2.1, we choose $\bar{x}^{(1)}(t_1) = \dfrac{\delta}{2}\bar{e}^{(j)}$. Then

$$||\phi(t)\phi^{-1}(t_1)\bar{x}^{(1)}(t_1)|| = ||\phi^{(j)}||\frac{\delta}{2} < \epsilon.$$

Here $\phi^{(j)}(t)$ is the j^{th} column of $\phi(t)\phi^{-1}(t_1)$. Hence, it follows that

$$||\phi(t)\phi^{-1}(t_1)|| = \max_{1\leq j\leq n} ||\phi^{(j)}(t)|| < \frac{2\epsilon}{\delta}, \ \ t \geq t_1$$

This proves Eq. (5.2.4) and the theorem. ∎

However, for nonlinear systems, boundedness and stability are distinct concepts.

Example 5.2.1 *Consider,*

$$\frac{dx}{dt} = t$$

This has a solution of the form $x(t) = x(t_0) + \dfrac{t^2}{2}$. *This is a stable solution but not bounded. Whereas for* $\dfrac{dx}{dt} = 0$, *all solutions are of the form* $x(t) = x(t_0)$ *which are uniformly stable but not asymptotically stable.*

Example 5.2.2 *Every solution of the system*

$$\frac{dx}{dt} \ = \ p(t)x(t)$$

is of the form

$$x(t) = x(t_0)\exp\left(\int_{t_0}^{t} p(s)ds\right)$$

Hence $x(t) \equiv 0$ *is asymptotically stable iff* $\int_{t_0}^{t} p(s)ds \to -\infty$ *as* $t \to \infty$.

For the nonlinear initial value problem given by Eq. (5.2.1), it is difficult to give the stability criterion. Instead, we examine a special form of Eq. (5.2.1), which is of the type

$$\frac{d\overline{x}}{dt} = A\overline{x}(t) + g(t, \overline{x}(t)) \tag{5.2.5}$$

$g(t, \overline{x})$ satisfies the following condition

$$\frac{\|g(t, \overline{x})\|}{\|\overline{x}\|} \to 0 \text{ as } \|\overline{x}\| \to 0 \tag{5.2.6}$$

In view of Eq. (5.2.6), it follows that $g(t, 0) = 0$ and hence 0 is a trivial solution of Eq. (5.2.5) with 0 initial condition.
We have the following stability theorem for such a kind of system.

Theorem 5.2.3 *Suppose the real part of eigenvalues of the matrix A are negative and the function $g(t, x)$ satisfies Eq. (5.2.6). Then the trivial solution of Eq. (5.2.5) is asympotically stable.*

Proof : The solvability of Eq. (5.2.5) with $\overline{x}(t_0) = \overline{x}_0$ is equivalent to the solvability of the following equation

$$\overline{x}(t) = \exp\left(A(t - t_0)\right) \overline{x}_0 + \int_{t_0}^{t} \exp\left(A(t - s)\right) g(s, \overline{x}(s)) ds \tag{5.2.7}$$

Since real(λ) < 0, (λ is any eigenvalue of A), it follows that there exists $c,\ \delta > 0$ such that $|| \exp(At)|| \le c \exp(-\delta t) \ \ \forall\ t \ge 0$. In view of this fact, Eq. (5.2.7) gives

$$\|\overline{x}(t)\| \le c\ \exp\left(-\delta(t - t_0)\right) \|\overline{x}_0\| + c \int_{t_0}^{t} \exp\left(-\delta(t - s)\right) \|g(s, \overline{x}(s))\|\ ds \tag{5.2.8}$$

Also the condition given by Eq. (5.2.6) on $g(t, \overline{x})$ implies that for given $m > 0$, $\exists\ d > 0$ such that

$$||g(t, \overline{v})|| \ \le\ m||\overline{v}|| \quad \text{whenever} \quad ||\overline{v}|| \ \le\ d \tag{5.2.9}$$

We assume that the initial condition $\overline{x}(t_0) = \overline{x}_0$ is such that $||\bar{x}_0|| < d$. Then in view of continuity of the function $\overline{x}(t)$, there exists $t_1 > t_0$ such that $||\overline{x}(t)|| < d$ for all $t \in [t_0,\ t_1] = I$.
Hence $t \in I$, implies that $||\overline{x}(t)|| < d$ and Eq. (5.2.9) gives that $||g(t, \overline{x}(t))|| \le m||\overline{x}(t))||$ for all $t \in I$. Using this inequality in Eq. (5.2.8), we get

$$\|\overline{x}(t)\| \exp(\delta t) \le c\ \exp(\delta t_0)\|x_0\| + c\ m \int_{t_0}^{t} \exp(\delta s)\ ||\overline{x}(s)||ds, \ \ t \in I$$

By Gronwall's lemma, we get that

$$||\overline{x}(t)|| \exp(\delta t) \le c\ ||\overline{x}_0||\ \exp(\delta t_0)\ \exp\left(c\ m(t - t_0)\right), \quad t \in I$$

That is

$$||\overline{x}(t)|| \le c\ ||\overline{x}_0|| \exp\left((c\ m - \delta)(t - t_0)\right), \quad t \in I \tag{5.2.10}$$

Since $\overline{x}_0$ and m are at our disposal, we choose m and $\bar{x}_0$ in such a way that $c\ m < \delta$ and $\|\overline{x}_0\| \ \le\ \min\left(d, \dfrac{\epsilon}{c}\right)$. Then Eq. (5.2.10) implies that $||\overline{x}(t)|| \ \le\ d$

for $t \in I$ and this bound d does not depend on t and hence by continuation theorem (Theorem 3.2.4) the solution $\overline{x}(t)$ can be extended for $t \in [t_0, \ \infty)$. Further Eq. (5.2.10) implies that given $\epsilon > 0$, $\exists \, \eta = min\left(d, \frac{\epsilon}{c}\right)$ such that

$$\|\overline{x}_0\| < \eta \ \ \Rightarrow \ \ \|\overline{x}(t, t_0, \overline{x}_0)\| < \epsilon \ \ \forall \, t \in [t_0, \ \infty)$$

and also $||\overline{x}(t)|| \ \to 0$ as $t \ \to \infty$.
This gives us the asymptotic stability of the zero solution of Eq. (5.2.5). ▮

Example 5.2.3 *The motion of a simple pendulum with damping is given by*

$$\ddot{\theta} + \frac{k}{m}\dot{\theta} + \frac{g}{l}\sin\theta = 0$$

Approximating $\sin\theta$ *by* $\theta - \frac{\theta^3}{3!} + \ldots$, *we get*

$$\sin\theta = \theta + f(\theta), \quad \frac{f(\theta)}{\theta} \to 0 \quad as \ \theta \to 0$$

This gives

$$\ddot{\theta} + \frac{k}{m}\dot{\theta} + \frac{g}{l}\theta + \frac{g}{l}f(\theta) = 0 \tag{5.2.11}$$

where $\frac{f(\theta)}{\theta} \to 0$ *as* $\theta \to 0$.

Eq. (5.2.11) is equivalent to

$$\frac{d\overline{x}}{dt} = A\overline{x} + F(\overline{x})$$

where

$$\overline{x} = (\theta, \ \dot{\theta}), \quad A = \begin{bmatrix} 0 & 1 \\ -\frac{g}{l} & -\frac{k}{m} \end{bmatrix} \ \ and \ F(\overline{x}) = F(\theta, \ \dot{\theta}) = \left(0, \ -\frac{g}{l}f(\theta)\right)$$

A has eigenvalues $-\frac{k}{2m} \pm \left(\frac{k^2}{4m^2} - \frac{g}{l}\right)^{\frac{1}{2}}$ *both of which have negative real parts if* k, m, g *and* l *are positive and*

$$\begin{aligned} \|F\overline{x}\| &= \frac{g}{l}\left|\frac{\theta^3}{3!} - \ldots\right| \\ &\leq M|\theta|^3 \to 0 \ \ as \ \ \|\overline{x}\| \to 0 \end{aligned}$$

Thus Theorem 5.2.3 is applicable and hence the main pendulum problem given by Eq. (5.2.11) is asymptotically stable.

In a similar way we have the following theorem for the unstability of zero solution of Eq. (5.2.5).

Theorem 5.2.4 *Assume that A possesses at least one eigenvalue with positive real part and $g(t, \overline{x})$ satisfies Eq. (5.2.6). Then the trivial solution of Eq. (5.2.5) is unstable.*

For details analysis concerning this section refer Agarwal and Gupta[1].

5.3 Phase Portrait Analysis

For the phase portrait, we shall concentrate on the autonomous system

$$\frac{d\overline{x}}{dt} = f(\overline{x}(t)) \tag{5.3.1}$$

Definition 5.3.1 *The set of solution curves in the phase space $\Re^n$ of the system given by Eq. (5.3.1) is called the phase portrait of the system.*

We shall be interested in the stability of the system around equilibrium points - those points where $f(\overline{x}) = 0$. Also, phase space will be assumed to be two dimensional so that its geometry can be studied clearly.
If $\bar{x_0}$ is an isolated equilibrium point of $f(\bar{x})$ and if f is differentiable at $\bar{x_0}$, then we have (by Eq. (2.3.4))

$$f(\bar{x}) = f(\bar{x_0}) + f'(x_0)(\bar{x} - \bar{x_0}) + g(\bar{x})$$

where $\dfrac{\|g(\bar{x})\|}{\|\bar{x} - \bar{x_0}\|} \to 0$ as $\|\bar{x} - \bar{x_0}\| \to 0$.

As $f(\bar{x_0}) = 0$, we have $f(\bar{x}) = f'(x_0)(\bar{x} - \bar{x}_0) + g(\bar{x})$. Translating the origin to $\bar{x_0}$, we get that Eq. (5.3.1) is equivalent to the system

$$\frac{d\bar{x}}{dt} = A\bar{x} + g(\bar{x}) \tag{5.3.2}$$

with $\dfrac{\|g(\bar{x})\|}{\|\bar{x}\|} \to 0$ as $\|\bar{x}\| \to 0$ and $A = f'(x_0)$.
We shall now study Eq. (5.3.2) through the linear system

$$\frac{d\bar{x}}{dt} = A\bar{x} \tag{5.3.3}$$

Theorem 5.2.3 and 5.2.4 give the relationship between the stability of the trivial solution of the nonlinear system of the form Eq. (5.3.2) with that of the linear system given by Eq. (5.3.3). We restate this relationship in terms of the following theorem.

Theorem 5.3.1 *(a) If the equilibrium point zero of the linear system given by Eq. (5.3.3) is asymptotically stable, so is the equilibrium point zero of the nonlinear system given by Eq. (5.3.2).*

(b) *If the equilibrium point zero of the Eq. (5.3.3) is unstable, so is the unstability of the equilibrium point zero of Eq. (5.3.2).*

(c) *If the equilibrium point zero Eq. (5.3.3) is stable then the zero solution of Eq. (5.3.2) can be stable, asymptotically stable or unstable.*

Example 5.3.1 *We investigate the phase portrait of the linear system given by Eq. (5.3.3) around the origin with* $A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$. *The solution of the linear system is given by*

$$\overline{x}(t) = \begin{bmatrix} \exp(-t) & 0 \\ 0 & \exp(2t) \end{bmatrix} \overline{c}, \quad \overline{c} = \overline{x}_0 = \overline{x}(t_0).$$

The solution curves of the above system are of the form

$$x_2 = \frac{k}{{x_1}^2}, \quad k = {c_1}^2 c_2 \; (\overline{c} = (c_1,\ c_2)).$$

The solution $\overline{x}(t)$ *defines a motion along these curves, that is each point* $\overline{c} \in \Re^2$ *moves to the point* $\overline{x}(t) \in \Re^2$ *after a time* t. *Hence the phase portrait in* $\Re^2$ *is as under.*

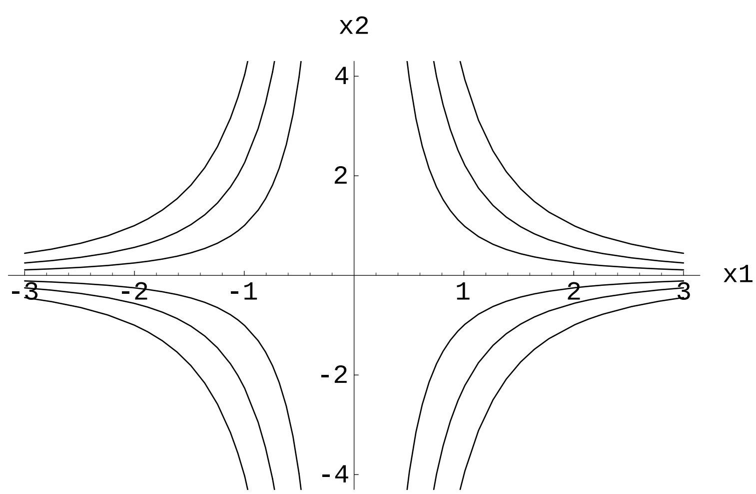

Figure 5.3.1: Phase portrait of Example 5.3.1

The function $f(\overline{x}) = A\overline{x}$ *defines a mapping on* $\Re^2$. *Recall that this is called a vector field. Since* $\frac{d\overline{x}}{dt} = f(\overline{x}) = A\overline{x}$, *at each point* $\overline{x} \in \Re^2$, *the solution curve* $\overline{x}(t)$ *is tangent to the vectors in the vector field* $f(\overline{x}) = A\overline{x}$. *So the vector field is as given below.*

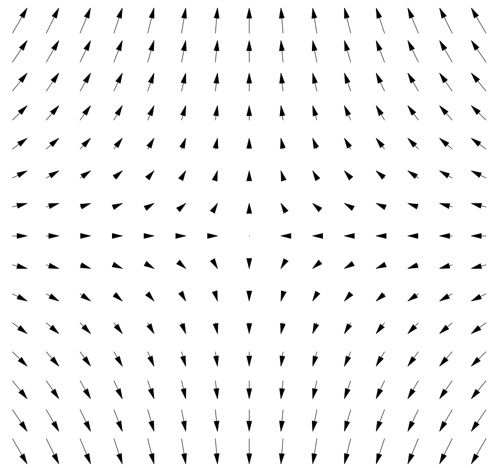

Figure 5.3.2: Vector field of Example 5.3.1

Example 5.3.2 *Consider now the coupled system*

$$\frac{d\overline{x}}{dt} = \begin{bmatrix} -1 & -3 \\ 0 & 2 \end{bmatrix} \overline{x} = A\overline{x}$$

The eigenvalues of the matrix A are -1 and 2 and the pair of corresponding eigenvectors is $(1,\ 0),\ \ (-1,\ 1)$.

$$P = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad P^{-1}AP = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} = D$$

So, under the transformation $\overline{y} = P^{-1}\bar{x}$, we obtain the decoupled system

$$\frac{d\overline{y}}{dt} = D\overline{y}$$

The phase portraits in $y_1 - y_2$ plane and $x_1 - x_2$ plane are as under.

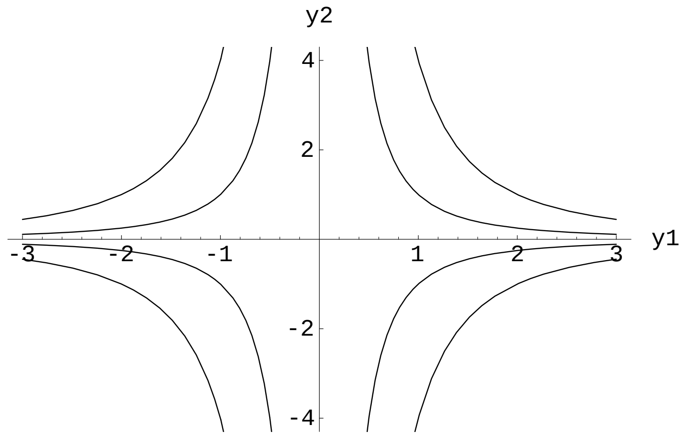

Figure 5.3.3: Phase portrait in $y_1 - y_2$ plane

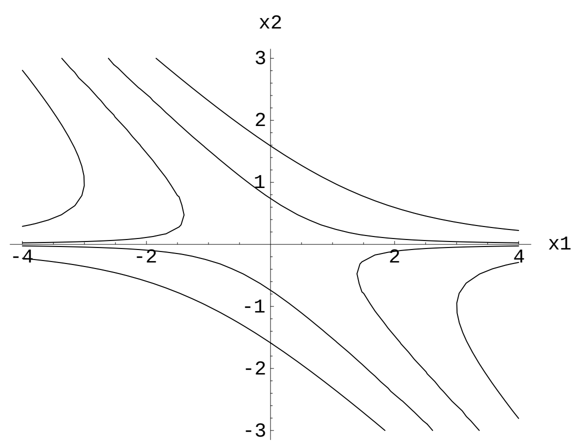

Figure 5.3.4: Phase portrait in $x_1 - x_2$ plane

In general, the phase portrait of a linear system given by Eq. (5.3.3) is best studied by an equivalent linear system

$$\frac{d\overline{x}}{dt} = B\overline{x} \tag{5.3.4}$$

where the matrix $B = P^{-1}AP$ *has one of the following form (by appealing to Theorem 5.1.1)*

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad or \quad B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

The solution of the initial value problem given by Eq. (5.3.4) with $\overline{x}(0) = \overline{x}_0$ *is*

$$\overline{x}(t) = \begin{bmatrix} \exp(\lambda t) & 0 \\ 0 & \exp(\mu t) \end{bmatrix} \overline{x}_0, \quad \overline{x}(t) = \exp(\lambda t) \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \overline{x}_0$$

or

$$\overline{x}(t) = \exp(at) \begin{bmatrix} \cos bt & -\sin bt \\ \sin bt & \cos bt \end{bmatrix} \overline{x}_0$$

Different cases will have different phase portraits.

Case 1: $B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad \lambda < 0 < \mu$

The phase portrait is of the form given in Figure 5.3.3.

In this case the system given by Eq. (5.3.4) is said to have a **saddle point at** $\overline{0}$. *If* $\lambda > 0 > \mu$, *then the arrows in the vector field are reversed as we see in the following Figure.*

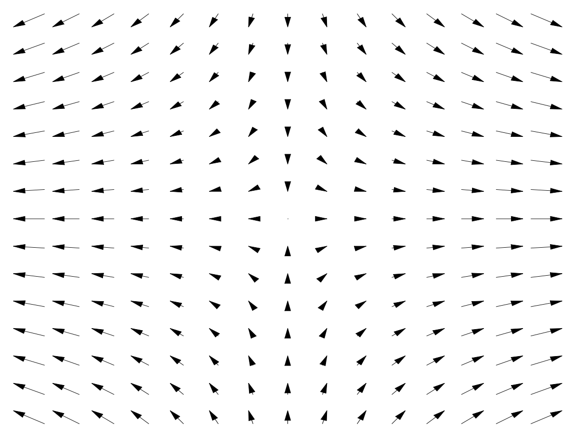

Figure 5.3.5: Vector field with arrows reversed

Case 2: $B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad \lambda \le \mu < 0 \quad or \quad B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad \lambda < 0$

In this case the origin is referred to as a **stable node** *(unstable if* $\lambda \ge \mu > 0$ *). The phase portraits are as under:*

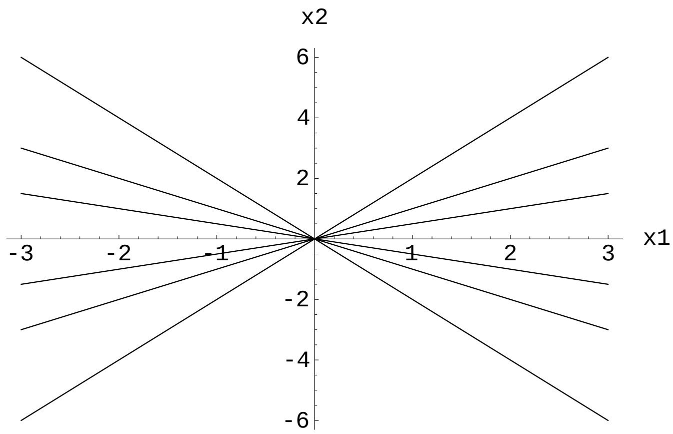

Figure 5.3.6: Phase portrait with $\lambda = \mu < 0$ (stable node)

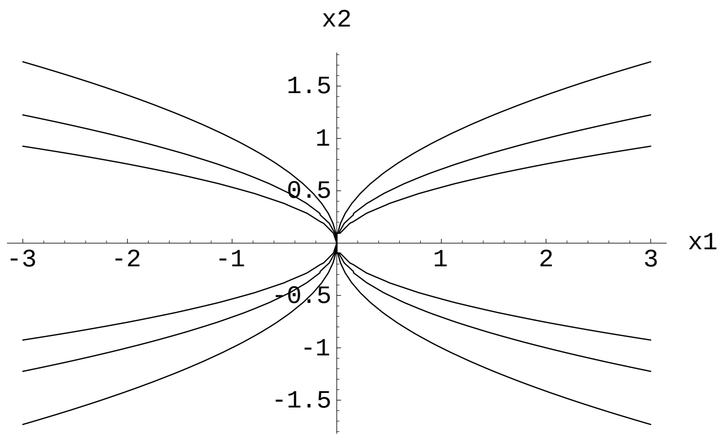

Figure 5.3.7: Phase portrait with $\lambda < \mu < 0$ (stable node)

When $B = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, $\lambda < 0$, *the phase portrait is as under.*

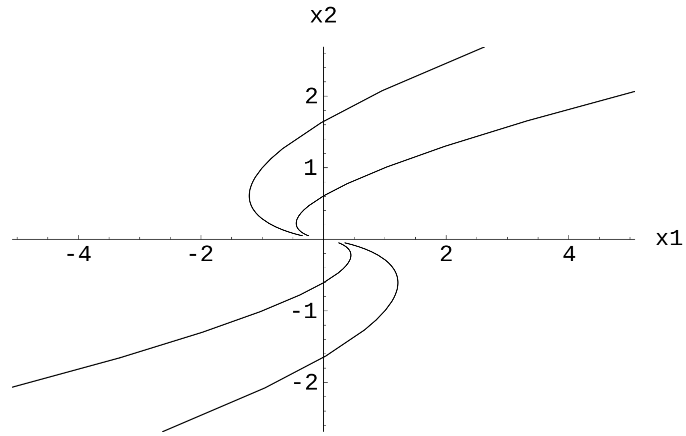

Figure 5.3.8: Origin as stable node

Case 3: $B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

If $a < 0$, the origin is a **stable focus** *with trajectories as spiral converging to zero. If $a > 0$, we have unstable focus.*

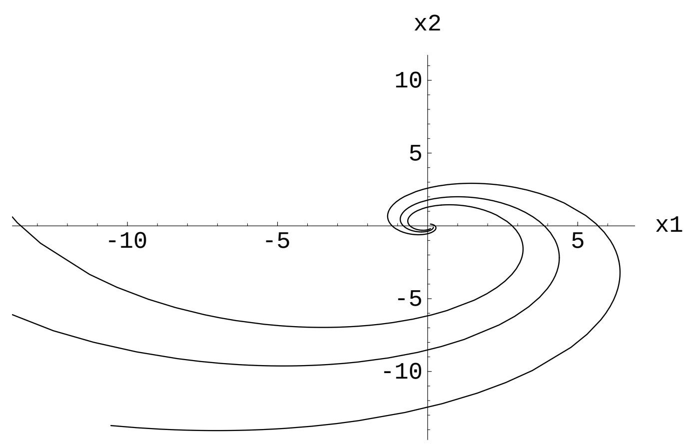

Figure 5.3.9: Stable focus

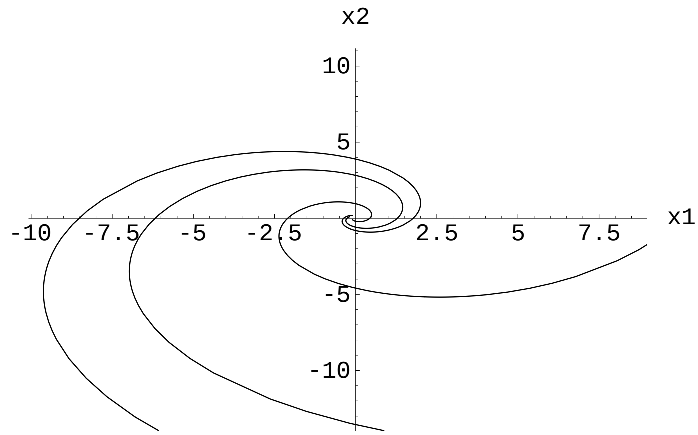

Figure 5.3.10: Unstable focus

Case 4: $B = \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}$

In this case the system given by Eq. (5.3.4) is said to have a **centre** *at the origin. The phase portrait is as under.*

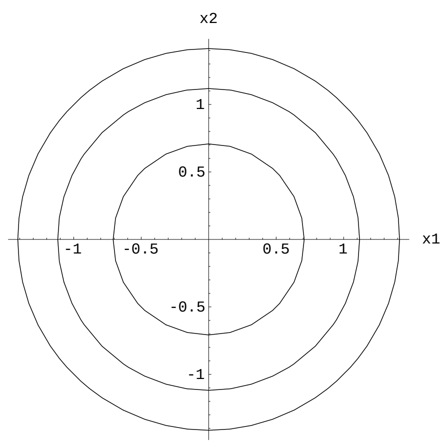

Figure 5.3.11: Origin as centre

Example 5.3.3

$$\frac{d\bar{x}}{dt} = A\bar{x}, \quad A = \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix}$$

A has eigenvalues $\lambda = \pm 2i$.

$$P = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}, \quad B = P^{-1}AP = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$$

$$\begin{aligned} \bar{x}(t) &= P \begin{bmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{bmatrix} P^{-1}\bar{c} \\ &= \begin{bmatrix} \cos 2t & -2\sin 2t \\ \frac{1}{2}\sin 2t & \cos 2t \end{bmatrix} \bar{c} \end{aligned}$$

$\bar{c} = \bar{x}(0)$.
It can be seen that ${x_1}^2 + 4{x_2}^2 = {c_1}^2 + 4{c_2}^2$.

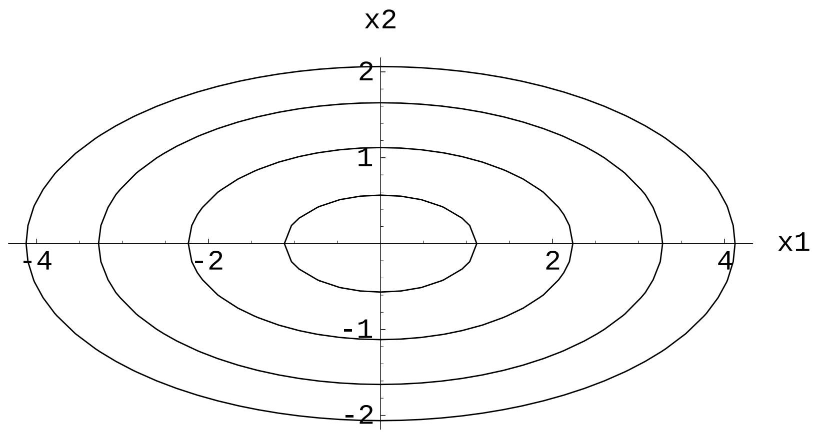

Figure 5.3.12: Phase portrait with origin as centre

The above discussion can be described in the following theorem, giving the saddle point, node, focus or centre at the origin of the linear system. We assume that *det* $A \neq 0$, so that $\bar{0}$ is the only equilibrium point.

Theorem 5.3.2 *Let* $\delta = detA$ *and* $\tau = traceA$ *and consider the linear system given by Eq. (5.3.3).*

(a) If $\delta < 0$*, then Eq. (5.3.3) has a saddle point at the origin.*

(b) If $\delta > 0$ *and* $\tau^2 - 4\delta \geq 0$*, then Eq. (5.3.3) has a node at the origin. it is stable if* $\tau < 0$ *and unstable if* $\tau > 0$*.*

(c) If $\delta > 0$ *and* $\tau^2 - 4\delta < 0$ *and* $\tau \neq 0$*, then Eq. (5.3.3) has a focus at the origin. It is stable if* $\tau < 0$ *and unstable if* $\tau > 0$*.*

(d) If $\delta > 0$ *and* $\tau = 0$ *then Eq. (5.3.3) has a centre at the origin.*

Example 5.3.4 *Consider the following linearized Predator-prey model (linearized around the equilibrium point* $(\frac{k}{\lambda}, \frac{a}{c})$*), given by Eq. (1.2.5)*

$$\frac{d\bar{x}}{dt} = A\bar{x}(t), \quad A = \begin{bmatrix} 0 & -\frac{kc}{\lambda} \\ \frac{a\lambda}{c} & 0 \end{bmatrix}$$

where $\bar{x} = (F, S)$*, with* F *representing fish population and* S *shark population. Using the notation of the above theorem (Theorem 5.3.2), we get* $\delta = det\ A = ak > 0$ *and* $\tau = trace\ A = 0$*. As* $\delta > 0$ *and* $\tau = 0$*, it implies that the equilibrium point* $(\frac{k}{\lambda}, \frac{a}{c})$ *is a centre.*

Example 5.3.5 *Consider the problem of mechanical oscillations discussed in Example 1.3.1*

$$\frac{d^2x}{dt^2} + k\frac{dx}{dt} + w^2x = 0$$

k *is the resistance and* $w^2 = \dfrac{\lambda}{ma}$*.*
This is equivalent to the following system

$$\frac{d\bar{x}}{dt} = A\bar{x}$$

where $\bar{x} = \left(x, \dfrac{dx}{dt}\right)$ *and* $A = \begin{bmatrix} 0 & 1 \\ -w^2 & -k \end{bmatrix}$.
$\delta = detA = w^2 > 0$ *and* $\tau = traceA = -k < 0$.

(i) If $\tau^2 - 4\delta = k^2 - 4w^2 < 0$*, we have origin as a stable focus.*

(ii) If $k^2 - 4w^2 \geq 0$*, we have origin as a stable node.*

For more on phase portrait analysis refer Perko [9].

5.4 Lyapunov Stability

The principle idea in the Lyapunov method is the following physical reasoning. If the rate of change $\frac{dE(\bar{x})}{dt}$ of the energy of a physical system is negative for every possible state $\bar{x}$ except for a single equilibrium state $\bar{x_e}$, then the energy will continually decrease until it finally assumes its minimum. From mathematical perspective, we look for a scalar valued Lyapunov function $V(\bar{x})$ of the state which is **(a)** positive **(b)** with $\dot{V}(\bar{x}) < 0 (\bar{x} \neq \bar{x_e})$ and **(c)** $V(\bar{x}) = \dot{V}(\bar{x}) = 0$ for $\bar{x} = \bar{x_e}$. One of the attractions of the Lyapunov method is its appeal to geometric intution as we see in the following two examples.

Example 5.4.1 *Let us recall Example 1.3.1, case 1 of a harmonic oscillator given by a second order differential equation.*

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \quad \omega^2 = \frac{\lambda}{ma}$$

In terms of a first order system we have

$$\frac{d\bar{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \bar{x}, \quad \bar{x} = (x_1, x_2)$$

The trajectories of the above system in the phase-plane are ellipses about the origin. The energy of the system is given by

$$E(\bar{x}) = E(x_1, x_2) = \omega^2 x_1^2 + x_2^2 = V(\bar{x})$$

The derivative $\dot{V}$ and V along the trajectory of the above system is given by

$$\begin{aligned} \dot{V}(\bar{x}) = \frac{dV\bar{x}(t)}{dt} &= (\bar{x}(t), \bigtriangledown V(\bar{x})) \\ &= \dot{x_1}\frac{\partial V}{\partial x_1} + \dot{x_2}\frac{\partial V}{\partial x_2} \\ &= 2\omega^2 x_1 x_2 + 2x_2\dot{x_2} \\ &= 0 \end{aligned}$$

So energy remains constant and hence the system is conservative.

Example 5.4.2 *Consider a slight modification of the earlier example*

$$\begin{aligned} \dot{x_1} &= x_2 - ax_1 \\ \dot{x_2} &= -x_1 - ax_2 \end{aligned}$$

with $a > 0$.
We define $V(\bar{x}) = (x_1^2 + x_2^2)$. Then

$$\begin{aligned} \dot{V}(\bar{x}(t)) &= (\dot{x}(t), \bigtriangledown V(\bar{x})) \\ &= -2a(x_1^2 + x_2^2)^2 \\ &= -2aV(\bar{x}) \\ &< 0 \ \ if \ \bar{x} \neq 0 \\ &= 0 \ \ if \ \bar{x} = 0 \end{aligned}$$

This implies that $V(\bar{x})$ is constantly decreasing along any solution of the above system. The trajectory of the above system is given by

$$R(t) = R_0 e^{-(t-t_0)}, \ \ R(t) = x_1^2(t) + x_2^2(t)$$

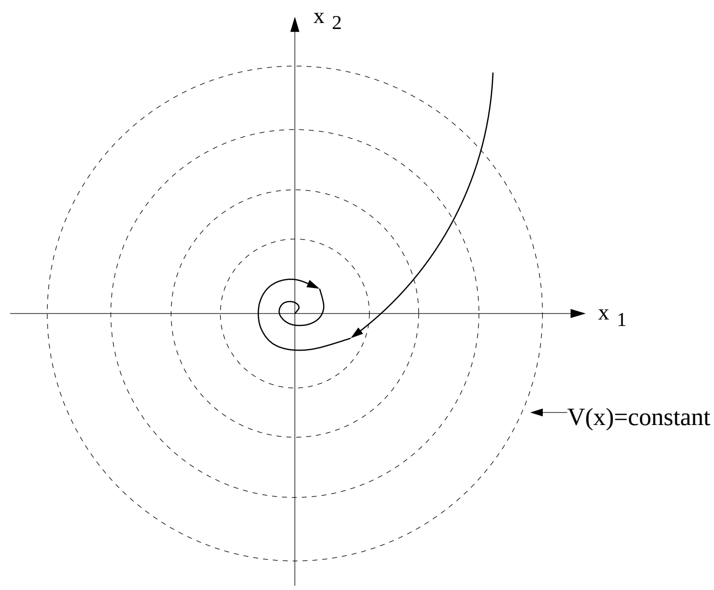

Figure 5.4.1: Trajectory along $V(\bar{x}) = \text{constant}$

Hence the trajectory of the above system cross the boundary of every region $V(\bar{x}) = $ constant from the outside towards inside. This leads to the following intersting geometric interpretation of the Lyapunov function. Let $V(\bar{x})$ be a measure of the distance of the state $\bar{x}$ from the orgin in the state space resulting in to $V(\bar{x}) > 0$ for $\bar{x} \neq 0$ and $V(0) = 0$. Suppose this distance is continuously decreasing as $t \to \infty$, that is, $\dot{V}(\bar{x}(t)) < 0$. Then $\bar{x}(t) \to 0$.

We now investigate the role of Lyapunov function to ensure the stability of the equilibrium state of the non-autonomous nonlinear system.

$$\frac{d\bar{x}}{dt} = f(t, \bar{x}(t)), \quad (t, \bar{x}) \in I \times \Re^n \tag{5.4.1}$$

where $f : I \times \Re^n \to \Re^n$ with origin as an equilibrium point ($f(t, 0) = 0$).
Assumptions 1: There exists a scalar valued function $V(t, \bar{x}) : I \times \Re^n \to \Re^n$, called Lyapunov function, such that it is continuous and has continuous partial derivatives w.r.t $\bar{x}$ and t with $V(t, 0) = 0$ and

(i) $V(t, \bar{x})$ is positive definite, that is, there exist a continuous, non-decreasing function α such that $\alpha(0) = 0$ and for all t and all $\bar{x} \neq 0$, we have

$$0 < \alpha(||\bar{x}||) \leq V(t, \bar{x}) \tag{5.4.2}$$

(ii) there exists a continuous scalar valued function r such that $r(0) = 0$ and the derivative $\dot{V}(\bar{x})$ of $V(\bar{x})$ along the motion $\bar{x}(t)$ satisfies the following

$$\dot{V}(t, \bar{x}(t)) \leq -r(||\bar{x}||) < 0 \;\; \forall \;\; (t, \bar{x}) \in I \times \Re^n \tag{5.4.3}$$

(ii) (a)

$$\dot{V}(t, \bar{x}(t)) \leq 0 \ \ \forall \ \ (t, \bar{x}) \in I \times \Re^n \tag{5.4.4}$$

(iii) there exists a continuous, non-decreasing function β such that $\beta(0) = 0$ and

$$V(t, \bar{x}(t)) \leq \beta(||\bar{x}||) \ \ \forall \ \ t \in I \tag{5.4.5}$$

Theorem 5.4.1 *Let $V(t, \bar{x})$ be a Lyapunov function of the system (5.4.1). Then the equilibrium state $\bar{0}$ is*

(a) stable if Assumptions (i) and (ii) hold

(b) uniformly stable if Assumptions (i),(ii)(a) and (iii) hold

(c) asymptotically stable if (i),(ii) and (iii) hold.

Proof :

(a) Let $\beta^*(t, ||\bar{x}||)$ be the maximum of $V(t, \bar{y})$ for $||\bar{y}|| \leq ||\bar{x}||$. Let $\epsilon > 0$ be arbitary. As $\beta^*(t_0, ||\bar{x}||)$ is a continuous function of $\bar{x}$, we can choose $\delta(t_0, \epsilon)$ such that

$$\beta^*(t_0, ||\bar{x_0}||) < \alpha(\epsilon) \ \text{ if } \ ||\bar{x_0}|| < \delta \tag{5.4.6}$$

Using (5.4.2), (5.4.4) and (5.4.6), we get

$$\alpha(||\bar{x}||) \leq V(t, \bar{x}(t, t_0, \bar{x_0})) \leq V(t_0, \bar{x_0}) \quad \leq \quad \beta^*(t_0, \bar{x_0}) < \alpha(\epsilon)$$

forall $t \geq t_0$ and $||\bar{x_0}|| < \delta$.

Since α is non-decreasing and positive, the above inequality implies that

$$||\bar{x}(t, t_0, \bar{x_0})|| < \epsilon \ \text{ whenever } \ ||\bar{x_0}|| < \delta \ \text{ for } \ t \geq t_0.$$

This proves the stability of the equilibrium state $\bar{0}$.

(b) Let (i), (ii)(a) and (iii) hold and assume that $\epsilon > 0$ is arbitary. We get $\delta(\epsilon)$ (independent of t_0) such that $\beta(\delta) < \alpha(\epsilon)$. Let $||\bar{x_0}|| < \delta$. Then proceeding as in part(a), we have

$$\alpha(\epsilon) > \beta(\delta) \geq V(t_0, \bar{x_0}) \geq V(t, \bar{x}(t, t_0, \bar{x_0})) \geq \alpha(||\bar{x}(t, t_0, \bar{x_0})||)$$

Here, for the second inequality, we have used Assumption 1(iii).

The above inequality gives

$$||\bar{x}(t, t_0, \bar{x_0})|| < \epsilon \ \text{ for } t \geq t_0 \ \text{ and } \ ||\bar{x_0}|| < \delta$$

where δ is independent of t_0.

This implies the uniform stability of the zero solution the equilibriun point.

(c) For asymptotic stability of the zero solution, we only need to show that $\exists\, r > 0$ such that

$$||\bar{x_0}|| \leq r \Rightarrow ||\bar{x}(t, t_0, \bar{x_0})|| \to 0 \text{ uniformly with } t \to \infty.$$

Equivallently, given $\mu > 0$, $\exists\ T(\mu, r)$ such that

$$||\bar{x_0}|| \leq r \Rightarrow ||\bar{x}(t, t_0, \bar{x_0})|| < \mu \text{ for } t \geq t_0 + T$$

Let functions $\alpha(\mu)$ and $\beta(\mu)$ be as defined in Eq. (5.4.3) and Eq. (5.4.5). Take any positive constant c_1 and find $r > 0$ such that $\beta(r) < \alpha(c_1)$. Let $||\bar{x_0}|| \leq r$. Then by part (b), we have

$$||\bar{x_0}|| \leq r \Rightarrow ||\bar{x}(t, t_0, \bar{x_0})|| < c_1 \ \forall \ t \geq t_0$$

Let $0 < \mu \leq ||\bar{x_0}||$. As $\beta(\mu)$ is continuous at 0, we find $\lambda(\mu) > 0$ such that $\beta(\lambda) < \alpha(\mu)$. Also $r(||\bar{x}||)$ is a continuous function and hence it assumes its minimum on a compact set $\lambda \leq ||\bar{x}|| \leq c_1(r)$ and let this minimum be $c_2(\mu, r)$.

Define $T(\mu, r) = \dfrac{\beta(r)}{c_2(\mu, r)} > 0$.

We first claim that $||\bar{x}(t, t_0, \bar{x_0})|| = \lambda$ for some $t = t_2$ in $[t_0, t_1]$, $t_1 = t_0 + T$. If not, then $||\bar{x}(t, t_0, \bar{x_0})|| > \lambda$ on $[t_0, t_1]$. Hence Eq. (5.4.2) gives

$$\begin{aligned}
0 < \lambda &\leq V(t_1, \bar{x}(t, t_0, \bar{x_0})) \\
&= V(t_0, x_0) + \int_{t_0}^{t_1} \dot{V}(\bar{x}(\tau, t_0, \bar{x_0}))d\tau \\
&\leq V(t_0, x_0) - \int_{t_0}^{t_1} r(||\bar{x}(\tau)||)d\tau \\
&\leq V(t_0, x_0) - (t_1, t_0)c_2 \\
&\leq \beta(r) - Tc_2 \\
&= 0
\end{aligned}$$

which is a contradiction.

Hence, we have

$$||\bar{x_2}|| = ||x(t_2, t_0, \bar{x_0})|| = \lambda$$

Therefore

$$\begin{aligned} \alpha(||\bar{x}(t,t_0,\bar{x_0})||) &\le V(t,\bar{x}(t,t_2,x_2)) \\ &\le V(t_2,x_2) \\ &\le \beta(\lambda) \\ &< \alpha(\mu) \quad \forall \ t \ge t_0 \end{aligned}$$

This implies that

$$||\bar{x}(t,t_0,\bar{x_0})|| < \mu \quad \forall \ t \ge t_0 + T(\mu,r) \ge 0$$

∎

The real weakness of Lyapunov method is that no general technique is known for the construction of the Lyapunov function $V(t,\bar{x})$. However, we shall describe the method of construction of Lyapunov function for asymptotically stable linear autonomuous system

$$\frac{d\bar{x}}{dt} = A\bar{x}$$

We assume that A has k distinct real eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_k$ and the remaining are complex eigenvalues $\lambda_j = a_j \pm ib_j, \quad k+1 \le j \le n$. Then by Theorem 5.1.1, there exists a basis $\{\bar{v_1}, \bar{v_2}, \cdots, \bar{v_k}, \bar{v_{k+1}}, \bar{v_{k+2}}, \cdots, \bar{v_n}, \bar{u_n}\}$ of $\Re^{2n-k}$ where $\bar{v_j}(1 \le j \le k)$ are eigenvectors corresponding real eigenvalues and $u_j \pm iv_j(k+1 \le j \le n)$ are the complex eigenvectors.
If $P = [\bar{v}_1, \bar{v}_2, \cdots, \bar{v}_k, \bar{v}_{k+1}, \bar{v}_{k+2}, \cdots, \bar{v}_n, \bar{u}_n]$, then P is nonsingular and $P^{-1}AP = D$, where D is block diagonal matrix of the form

$$D = \begin{bmatrix} \lambda_1 & & & & & & \\ & \lambda_2 & & & & & \\ & & \ddots & & & & \\ & & & \lambda_k & & & \\ & & & & B_{k+1} & & \\ & & & & & \ddots & \\ & & & & & & B_n \end{bmatrix}$$

with $B_i = \begin{bmatrix} a_i & -b_i \\ b_i & a_i \end{bmatrix}$ as a Jordan block corresponding to the eigenvalue $a_i \pm ib_i$. Using the transformation $\bar{x} = P\bar{y}$, we get the reduced system

$$\frac{d\bar{y}}{dt} = P^{-1}AP\bar{y} = D\bar{y}$$

For asymptotic stability, we require all diagonal elements of D to be negative. We seek Lyapunov function V of the forrm

$$V(\bar{y}) = (B\bar{y}, \bar{y}) \tag{5.4.7}$$

where B is a suitable positive definite matrix. Differentiating Eq. (5.4.7), we get

$$\begin{aligned}
\frac{dV}{dt} &= \left(B\frac{d\bar{y}}{dt}, \bar{y}(t)\right) + \left(B\bar{y}(t), \frac{dy}{dt}\right) \\
&= (BD\bar{y}(t), \bar{y}(t)) + (B\bar{y}(t), D\bar{y}(t)) \\
&= (BD\bar{y}(t), \bar{y}(t)) + \left(D^\top B\bar{y}(t), \bar{y}(t)\right) \\
&= \left(\left[BD + D^\top B\right] \bar{y}(t), \bar{y}(t)\right) \qquad (5.4.8)
\end{aligned}$$

In order to ensure that $\frac{dV}{dt}$ is negative definite, we shall require that

$$\frac{dV}{dt} = -\left(Q\bar{y}(t), \bar{y}(t)\right) \tag{5.4.9}$$

where Q is a given positive definite matrix. Eq. (5.4.9) will hold, if we assume that

$$BD + D^\top B = -Q \tag{5.4.10}$$

Thus, to get a Lyapunov function which is positive definite with $\frac{dV}{dt}$ as negative definite, we need to solve the matrix equation given by Eq. (5.4.10) for B with a given positive definite matrix Q (say $Q = I$).

Example 5.4.3 *Construct a Lyapunov function for the system*

$$\frac{d\bar{x}}{dt} = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -3 & -2 \\ 0 & 0 & 1 & -1 \end{bmatrix} \bar{x}$$

The eigenvalues of this system are $-1 \pm i, \ -2 \pm 2i$.
One gets

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \textit{with} \ P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 \end{bmatrix}$$

Then

$$P^{-1}AP = D = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & -2 & -2 \end{bmatrix}$$

We make the transformation $\bar{x} = P\bar{y}$ to get the reduced system

$$\frac{d\bar{y}}{dt} = D\bar{y}$$

Take $Q = I$ and solve the matrix equation

$$BD + D^T B = -I$$

to get

$$B = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{pmatrix}$$

This gives the Lyapunov function $V(\bar{y})$ as

$$V(\bar{y}) = \frac{1}{2}y_1^2 + \frac{1}{2}y_2^2 + \frac{1}{4}y_3^2 + \frac{1}{4}y_4^2$$

where

$$\bar{y} = P^{-1}\bar{x} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 \end{pmatrix}$$

We have

$$y_1 = x_1, y_2 = x_2, y_3 = x_4, y_4 = -x_3 - x_4$$

Hence

$$V(\bar{x}) = \frac{x_1^2}{2} + \frac{x_2^2}{4} + \frac{(x_3 + x_4)^2}{4} + \frac{x_4^2}{4}$$

Example 5.4.4 *Consider the nonlinear system of the form*

$$\frac{d\bar{x}}{dt} = A\bar{x} + g(\bar{x}) \tag{5.4.11}$$

A has eigenvalues with negative real part and $g(\bar{x})$ satisfies the growth condition of the form

$$||g(\bar{x})|| \leq k||\bar{x}|| \quad \forall \ \bar{x}$$

with small k.

We construct a Lyapunov function for Eq. (5.4.11) in the following way. $V(\bar{x}) = (B\bar{x}, \bar{x})$, B a positive definite matrix. Solve B to satisfy the matrix equation

$$A^\top B + B^\top A = -I$$

We have

$$\begin{aligned}\frac{dV}{dt} &= \dot{\bar{x}}^\top B\bar{x} + \bar{x}^\top B\dot{\bar{x}} \\ &= \left(\bar{x}^\top A^\top + [g(\bar{x})]^\top\right) B\bar{x} + \bar{x}^\top B(A\bar{x} + g(\bar{x})) \\ &= \bar{x}^\top (A^\top B + BA)\bar{x} + \left([g(\bar{x})]^\top B\bar{x} + \bar{x}^\top B[g(\bar{x})]\right) \\ &= -\bar{x}^\top \bar{x} + \left([g(x)]^\top B\bar{x} + \bar{x}^\top B[g(\bar{x})]\right)\end{aligned}$$

Estimating the second term within the square bracket, we get

$$\begin{aligned}\|[g(\bar{x})]^\top B\bar{x} + \bar{x}^\top B[g(\bar{x})]\| &\le 2\|g(\bar{x})\|\|B\|\|\bar{x}\| \\ &\le 2k\|B\|\|\bar{x}\|^2\end{aligned}$$

Hence $\frac{dV}{dt} \le -\|\bar{x}\|^2[1-2k\|B\|] = -\alpha(\|x\|)$ *where* $\alpha(r) = \alpha r^2,\ \ \alpha - 2k\|B\| > 0.$ *Thus the Lyapunov function* $V(\bar{x})$ *is such that* V *is positive definite with* $\frac{dV}{dt}$ *negative definite and hence the system Eq. (5.4.11) is asymptotically stable.*

For more on stability results refer Aggarwal and Vidyasagar [2], Amann [3], Borckett [4], Lakshmikantham et al[6,7] and Mattheij and Molenaar [8].

5.5 Exercises

1. Show that all the solution of the linear system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}$$

where $A(t) = \begin{bmatrix} -t & 0 & 0 \\ 0 & -t^2 & 0 \\ 0 & 0 & -t^2 \end{bmatrix}, \quad \begin{bmatrix} e^{-t} & -1 & -\cos t \\ 1 & -e^{-2t} & t^2 \\ \cos t & -t^2 & e^{-3t} \end{bmatrix}$

tend to zero as $t \to \infty$.

2. By using the Cayley-Hamilton theorem (Theorem 5.1.3), show that for $A \in \Re^{n\times n}$, there exist functions $\alpha_k(t),\ 1 \le k \le n-1$ such that $e^{At} = \sum_{k=0}^{n-1} \alpha_k(t) A^k$.

3. Show that all solutions of the following differential equations are bounded in $[t_0, \infty)$

(i) $\frac{d^2x}{dt^2} + c\frac{dx}{dt} + \left[1 + \frac{1}{1+t^2}\right] x = 0, \quad c > 0.$

(ii) $\frac{d^2x}{dt^2} + c\frac{dx}{dt} + \left[1 + \frac{1}{1+t^4}\right] x = 0, \quad c > 0.$

4. Test the stability, asymptotic stability or unstability for the trivial solution of the following linear systems.

(i) $\frac{d\bar{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \bar{x}$

(ii) $\frac{d\bar{x}}{dt} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -6 & -5 \end{bmatrix} \bar{x}$

(iii) $\frac{d\bar{x}}{dt} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 3 & 1 \end{bmatrix} \bar{x}$

(iv) $\frac{d\bar{x}}{dt} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -3 \\ 1 & -5 & -3 \end{bmatrix} \bar{x}$

5. Test the stability, asymptotic stability or unstability of trivial solution of each of the following system

(i)

$$\begin{aligned} \frac{dx_1}{dt} &= -2x_1 + x_2 + 3x_3 + 8x_1^2 + x_2^3 \\ \frac{dx_2}{dt} &= -6x_2 - 5x_3 + 7x_3^4 \\ \frac{dx_3}{dt} &= -2x_3 + x_1^4 + x_2^2 + x_3^3 \end{aligned}$$

(ii)

$$\frac{dx_1}{dt} = 2x_1 + x_2 - x_1^2 - x_2^2$$
$$\frac{dx_2}{dt} = x_1 + 3x_2 - x_1^3 \sin x_3$$
$$\frac{dx_3}{dt} = x_2 + 2x_3 + x_1^2 + x_2^2$$

6. Discuss the stability or unstability of the origin of the following linear sytem and sketch the phase portraits.

   (i) $\frac{dx_1}{dt} = -2x_1 + x_2$
   $\frac{dx_2}{dt} = -5x_1 - 6x_2$

   (ii) $\frac{dx_1}{dt} = 4x_1 + x_2$
   $\frac{dx_2}{dt} = 3x_1 + 6x_2$

   (iii) $\frac{dx_1}{dt} = x_2$
   $\frac{dx_2}{dt} = 2x_1 - x_2$

   (iv) $\frac{dx_1}{dt} = x_1 + x_2$
   $\frac{dx_2}{dt} = 3x_1 - x_2$

7. Consider the nonlinear system

$$\frac{dx}{dt} = a(t)x^3$$

   Find its solution.

   (i) Show that origin is uniformly asymptotically stable if $a(t) < 0$.

   (ii) If $a(t) = -\frac{1}{t+1}$, show that origin is asymptotically stable but not uniformly stable.

   (iii) If $a(t) = -\frac{1}{(t+1)^2}$, show that the solution of the corresponding IVP approaches a constant value as $t \to \infty$ but origin is uniformly stable.

8. Consider the nonlinear system

$$\frac{dx}{dt} = -x^p, \quad p > 1 \text{ and } p \text{ is integer}$$

   Show that

(i) If p is odd, then origin is uniformly stable and

(ii) If p is even, then origin is unstable.

9. Find Lyapunov functions around the origin for the following system.

(i) $\frac{dx}{dt} = -x$

(ii) $\frac{dx}{dt} = -x(1-x)$

(iii) $\frac{d\bar{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -1 & -a \end{bmatrix} \bar{x}, \quad a \geq 0$

10. Show that the origin is unstable equilibrium point of the system

$$\frac{d^2x}{dt^2} - \left(\frac{dx}{dt}\right)^2 sgn(\dot{x}) + x = 0$$

References

[1] Ravi Agarwal and Ramesh C Gupta, *Essentials of Ordinary Differential Equations, McGraw Hill Book Company, 1991*

[2] J K Aggarwal and M. Vidyasagar, (Editors) *Nonlinear Systems, Stability Analysis, Dowden, Hutchinson and Ross, Inc., 1977*

[3] Herbert Aman, *Ordinary Differential Equations: Introduction to Nonlinear Analysis, Walter de Gruyter, 1990*

[4] Roser W Brockett, *Finite Dimensional Linear Systems, John Wiley and Sons, Inc., 1970*

[5] F. R. Gantmacher, *The Theory of Matrices Vols. I and II, Chelsea Publishing Company, 1960*

[6] V. Lakshimantham and S. G. Deo, *Method of Variation of Parameters for Dynamic Sytems, Gorden and Breach Science Pub. , 1998*

[7] V. Lakshmikantham, S. Leela and A.A Martynyuk, *Stability Analysis of Nonlinear Systems, Marcel Dekker, 1989*

[8] R M M Mattheij and J Molenaar, *Ordinary Differential Equations in Theory and Practice, John Wiley, 1996*

[9] Lawrence Perko, *Differential Equations and Dynamical Systems, Springer-Verlag, 1991*

Chapter 6

Series Solution

This chapter lays emphasis on the classical theory of series solution for ODE by using power series expansion around ordinary and singular points.
We mainly focus on Legendre, Hermite and Bessel differential equations. We also deduce various interesting properties of Legendre and Hermite polynomials as well as Bessel functions.

6.1 Preliminaries

A power series in powers of $(t - t_0)$ is an infinite series of the form

$$\sum_{k=0}^{\infty} c_k(t-t_0)^k = c_0 + c_1(t-t_0) + c_2(t-t_0)^2 + \cdots \tag{6.1.1}$$

where t is a variable and $c_0, c_1, c_2, \cdots$ are constants. Recall that a series $\sum_{k=0}^{\infty} c_k(t-t_0)^k$ is convergent at a point t if the limit of the partial sums $s_n(t) = \sum_{k=0}^{n} c_k(t-t_0)^k$ exists. This limit $f(t)$ is denoted as the sum of the series at the point t. A series which does not converge is said to be a divergent series.

Example 6.1.1 *Consider the geometric power series* $\sum_{k=0}^{\infty} t^k = 1 + t + t^2 + \cdots$
The partial sums $s_n(t) = 1 + t + \cdots + t^n$ *satisfy the relation*

$$ts_n(t) = t + t^2 + \cdots + t^n + t^{n+1}$$

and hence $s_n(t) = \dfrac{1 - t^{n+1}}{1-t}$.

This gives $f(t) = \lim s_n(t) = \dfrac{1}{1-t}$ *for* $|t| < 1$.

Definition 6.1.1 *A power series* $\sum_{k=0}^{\infty} c_k(t-t_0)^k$ *is said to converge absolutely at a point* t *if the series* $\sum_{k=0}^{\infty} |c_k(t-t_0)^k|$ *converges.*

One can show that if the series converges absolutely, then it also converges. However, the converse is not true.

We have the following tests for checking the convergence or divergence of a series of real numbers.

(i) **Comparision test**

(a) Let a series $\sum_{k=0}^{\infty} a_k$ of real numbers be given and let there exists a convergent series $\sum_{k=0}^{\infty} b_k$ of nonnegative real numbers such that

$$|a_k| \leq b_k, \qquad k \geq 1$$

Then the orginal series $\sum_{k=0}^{\infty} a_k$ converges.

(b) Let a series $\sum_{k=0}^{\infty} a_k$ of real numbers be given and let there exists a divergent series $\sum_{k=0}^{\infty} d_k$ of nonnegative real numbers such that

$$|a_k| \geq d_k, \qquad k \geq 1$$

Then the orginal series $\sum_{k=0}^{\infty} a_k$ diverges.

(ii) **Ratio test**

Let the series $\sum_{k=0}^{\infty} a_k(\text{ with } a_n \neq 0)$ be such that $\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = L$.

(a) The series $\sum_{k=0}^{\infty} a_k$ converges absolutely if $L < 1$.

(b) The series $\sum_{k=0}^{\infty} a_k$ diverges if $L > 1$.

(c) No conclusion if $L = 1$.

(iii) **Root test**

If the series $\sum_{k=0}^{\infty} a_k$ is such that $\lim_{k\to\infty}\left((|a_k|)^{\frac{1}{k}}\right) = L$.

(a) The series $\sum_{k=0}^{\infty} a_k$ converges absolutely of $L < 1$.

(b) The series $\sum_{k=0}^{\infty} a_k$ diverges if $L > 1$.

(c) No conclusion if $L = 1$.

Theorem 6.1.1 *If the power series given by Eq. (6.1.1) converges at a point $t = t_1$, then it converges absolutely for every t for which $|t - t_0| < |t_1 - t_0|$.*

Proof : Since the series given by Eq. (6.1.1) converges for $t = t_1$, it follows that the partial sums $s_n(t_1)$ converge and hence $s_n(t_1)$ is Cauchy. This implies that

$$s_{n+1}(t_1) - s_n(t_1) \to 0 \text{ as } n \to \infty$$

This, in turn, implies that $c_k(t_1 - t_0)^k \longrightarrow 0$ as $k \to \infty$. Hence the elements $c_k(t_1 - t_0)^k$ of the series given by Eq. (6.1.1) are bounded at $t = t_1$. That is,

$$|c_k(t_1 - t_0)^k| \le M \text{ for all } k \ge 1 \tag{6.1.2}$$

Eq. (6.1.2) implies that

$$\begin{aligned} |c_k(t - t_0)^k| &= \left| c_k(t_1 - t_0)^k \left(\frac{t - t_0}{t_1 - t_0}\right)^k \right| \\ &\le M \left| \frac{t - t_0}{t_1 - t_0} \right|^k \end{aligned} \tag{6.1.3}$$

If $|t - t_0| < |t_1 - t_0|$, the series $\sum_{k=0}^{\infty} \left|\frac{t - t_0}{t_1 - t_0}\right|^k$ converges and hence by comparision test the series $\sum_{k=0}^{\infty} c_k(t - t_0)^k$ converges absolutely. ▮

Definition 6.1.2 *If the series $\sum_{k=0}^{\infty} c_k(t - t_0)^k$ converges absolutely for $|t - t_0| < r$ and diverges for $|t - t_0| > r$, then r is called the radius of convergence.*

The power series $\sum_{k=1}^{\infty} \frac{t^k}{k}$ converges for $|t| < 1$ and diverges for $|t| > 1$. At $t = 1$, it diverges and $t = -1$, it converges. Thus, the radius of convergence of this series is 1.

The radius of convergence of the power series given by Eq. (6.1.1) may be determined from the coefficients of the series as follows.

Theorem 6.1.2 *(Radius of convergence)*

(i) Suppose that the sequence $\{|c_k|^{\frac{1}{k}}\}$ *converges . Let* L *denotes its limit. Then, if* $L \neq 0$*, the radius of convergence* r *of the power series is* $\frac{1}{L}$.

(ii) If $L = 0$*, then* $r = \infty$ *and hence the series Eq. (6.1.1) converges for all* t.

(iii) If $\{|c_k|^{\frac{1}{k}}\}$ *does not converge, but it is bounded, then* $r = \frac{1}{l}$ *where* $l = sup\{|c_k|^{\frac{1}{k}}\}$. *If this sequence is not bounded, then* $r = 0$ *and hence the series is convergent only for* $t = t_0$.

Let $\sum_{k=0}^{\infty} c_k(t-t_0)^k$ be a power series with non zero radius of convergence r. Then the sum of the series is a function $f(t)$ of t and we write

$$f(t) = \sum_{k=0}^{\infty} c_k(t-t_0)^k \tag{6.1.4}$$

One can easily show the following.

(i) The function $f(t)$ in Eq. (6.1.4) is continuous at $t = t_0$.

(ii) The same function $f(t)$ can not be represented by two different power series with the same centre. That is, if

$$f(t) = \sum_{k=0}^{\infty} c_k(t-t_0)^k = \sum_{k=0}^{\infty} d_k(t-t_0)^k$$

in a disk: $|t-t_0| < r$, then $c_k = d_k$ for all $k \geq 0$.

We can carry out the standard operations on power series with ease - addition and subtraction, multiplication, term by term differentiation and integration.

(i) **Addition and Subtraction**

Two power series $\sum_{k=0}^{\infty} c_k(t-t_0)^k$ and $\sum_{k=0}^{\infty} d_k(t-t_0)^k$ can be added and subtracted in the common radius of convergence.

$$\begin{aligned} \text{If } f(t) &= \sum_{k=0}^{\infty} c_k(t-t_0)^k \text{ in } |t-t_0| < r_1 \quad \text{and} \\ g(t) &= \sum_{k=0}^{\infty} d_k(t-t_0)^k \text{ in } |t-t_0| < r_2 \end{aligned}$$

Then $f(t) \pm g(t) = \sum_{k=0}^{\infty}(c_k \pm d_k)(t-t_0)^k$ in $|t-t_0| < r = min(r_1, r_2)$.

(ii) **Multiplication**

$$\begin{aligned}\text{If } f(t) &= \sum_{k=0}^{\infty} c_k(t-t_0)^k \text{ in } |t-t_0| < r_1 \quad \text{and} \\ g(t) &= \sum_{k=0}^{\infty} d_k(t-t_0)^k \text{ in } |t-t_0| < r_2\end{aligned}$$

Then $h(t) = f(t)g(t)$ defined within the radius of convergence of each series and

$$h(t) = \sum_{k=0}^{\infty} c_k(t-t_0)^k$$

where $c_k = \sum_{m=0}^{k} c_m d_{k-m}$. That is

$$h(t) = c_0d_0 + (c_0d_1 + c_1d_0)\, t + (c_0d_2 + c_1d_1 + c_2d_0)\, t^2 + \cdots$$

The series converges absolutely within the radius of convergence of each series.

(iii) **Differentiation**

$$\text{Let } f(t) = \sum_{k=0}^{\infty} c_k(t-t_0)^k \text{ in } |t-t_0| < r$$

Then

$$\frac{df}{dt} = \dot{f} = \sum_{k=1}^{\infty} c_k k(t-t_0)^{k-1} \text{ in } |t-t_0| < r$$

(iv) **Integration**

$$\text{Let } f(t) = \sum_{k=0}^{\infty} c_k(t-t_0)^k \text{ in } |t-t_0| < r$$

Then

$$\int f(t)dt = \sum_{k=0}^{\infty} \frac{c_k}{k+1}(t-t_0)^{k+1} \text{ in } |t-t_0| < r$$

Definition 6.1.3 *We shall say that a function $f(t)$ is analytic at $t = t_0$ if it can be expanded as a sum of a power series of the form $\sum_{k=0}^{\infty} c_k(t-t_0)^k$ with a radius of convergence r. It is clear that if $f(t)$ is analytic at t_0, then $c_k = \frac{f^{(k)}(t_0)}{k!}$, $k = 0, 1, 2, \cdots$ ($f^k(t_0)$ denotes the k^{th} derivative of f at t_0).*

Example 6.1.2 *We have the binomial expansion for $(1-t)^{-k}$ for a fixed positive number k to give us*

$$(1-t)^{-k} = 1 + kt + \frac{k(k+1)}{2!}t^2 + \frac{k(k+1)...(k+r-1)}{r!}t^r + \cdots \quad for \ |t| < 1$$

We denote by

$$u_r = \frac{k(k+1)...(k+r-1)}{r!}$$

Then

$$(1-t)^{-k} = \sum_{r=0}^{\infty} u_r t^r, \quad |t| < 1$$

As a power series can be differentiated term by term in its interval of convergence, we have

$$k(1-t)^{-k-1} = \sum_{r=1}^{\infty} r u_r t^{r-1}, \quad |t| < 1$$

This gives

$$k(1-t)^{-k}\frac{1}{(1-t)} = \sum_{r=1}^{\infty} r u_r t^{r-1}, \quad |t| < 1$$

Again, using the power series expansion for $\frac{1}{1-t} = 1 + t + t^2 + \cdots$, *we get*

$$k\left(\sum_{r=0}^{\infty} t^r\right)\left(\sum_{r=0}^{\infty} u_r t^r\right) = \sum_{r=1}^{\infty} r u_r t^{r-1}$$

Using the product formula for LHS, we get

$$k\sum_{r=0}^{\infty} t^r \sum_{l=0}^{r} u_l = \sum_{r=0}^{\infty}(r+1)u_{r+1}t^r$$

Uniqueness of power series representation gives

$$(r+1)u_{r+1} = k\sum_{l=0}^{r} u_l \tag{6.1.5a}$$

where

$$u_r = \frac{k(k+1).....(k+r-1)}{r!}, \quad k \text{ is a fixed integer} \tag{6.1.5b}$$

We now make an attempt to define the concept of uniform convergence. To define this concept, we consider the following series whose terms are functions $\{f_k(t)\}_{k=0}^{\infty}$

$$\sum_{k=0}^{\infty} f_k(t) = f_0(t) + f_1(t) + f_2(t) + \cdots \tag{6.1.6}$$

Note that for $f_k(t) = c_k(t-t_0)^k$, we get the power series.

Definition 6.1.4 *We shall say that the series given by Eq. (6.1.6) with sum $f(t)$ in an interval $I \subset \Re$ is uniformly convergent if for every $\epsilon > 0$, we can find $N = N(t)$, not depending on t, such that*

$$|f(t) - s_n(t)| < \epsilon \ \text{for all}\, n \geq N(\epsilon)$$

where $s_n(t) = f_0(t) + f_1(t) + \cdots + f_n(t)$.

Theorem 6.1.3 *A power series $\sum_{k=0}^{\infty} c_k(t-t_0)^k$ with nonzero radius of convergennt is uniformly convergent in every closed interval $|t - t_0| \leq p$ of radius $p < r$.*

Proof : For $|t - t_0| \leq p$ and any positive integers n and l we have

$$\left|c_{n+1}(t-t_0)^{n+1} + \cdots + c_{n+l}(t-t_0)^{n+l}\right| \leq |c_{n+1}|\, p^{n+1} + \cdots + |c_{n+l}|\, p^{n+l} \qquad (6.1.7)$$

The series $\sum_{k=0}^{\infty} c_k(t-t_0)^k$ converges absolutely if $|t - t_0| \leq p < r$ (by Theorem 6.1.1) and hence by Cauchy convergence, given $\epsilon > 0$, $\exists\, N(\epsilon)$ such that

$$|c_{n+1}|\, p^{n+1} + \cdots + |c_{n+p}|\, p^{n+l} < \epsilon \quad \text{for } n \geq N(\epsilon), \ \ l = 1, 2, ..$$

Eq. (6.1.7) gives

$$\begin{aligned}&\left|c_{n+1}(t-t_0)^{n+1} + \cdots + c_{n+p}(t-t_0)^{n+l}\right| \\ &\leq |c_{n+1}|\, p^{n+1} + \cdots + |c_{n+p}|\, p^{n+l} < \epsilon \quad \text{for } n \geq N(\epsilon), \quad |t - t_0| \leq p < r\end{aligned}$$

This implies that, given $\epsilon > 0, \exists\, N(\epsilon)$ such that $|f(t) - s_n(t)| < \epsilon$ for $n \geq N(\epsilon)$ and all t where $s_n(t) = \sum_{k=0}^{n} c_k(t-t_0)^k$. This proves the uniform convergence of the power series inside the interval $|t - t_0| \leq p < r$. ∎

Example 6.1.3 *The geometric series $1 + t + t^2 + \cdots$ is uniformly convergent in the interval $|t| \leq p < 1$. It is not uniformly convergent in the whole interval $|t| < 1$.*

6.2 Linear System with Analytic Coefficients

We now revisit the non-autonomous system in $\Re^n$

$$\begin{aligned}\frac{d\bar{x}}{dt} &= A(t)\bar{x} + g(t) \qquad (6.2.1)\\ \bar{x}(0) &= \bar{x}_0\end{aligned}$$

where the matrix $A(t)$ is analytic at $t = 0$ and hence has the power series representation

$$A(t) = \sum_{k=0}^{\infty} A_k t^k \tag{6.2.2}$$

in its interval of convergence $|t| < r$. Here each A_k is $n \times n$ matrix.

So, the homogeneous system corresponding to Eq. (6.2.1) is given by

$$\begin{aligned} \frac{d\bar{x}}{dt} &= \sum_{k=0}^{\infty} A_k t^k \bar{x} \\ \bar{x}(0) &= \bar{x}_0 \end{aligned} \tag{6.2.3}$$

We shall look for analytic solution of Eq. (6.2.3), which is of the form $\bar{x}(t) = \sum_{k=0}^{\infty} \bar{c}_k t^k$. The vector coefficients $\bar{c}_k$ are to be determined. The point $t = 0$ is called ordinary point of the above system.

The following theorem gives the analytic solution of the system given by Eq. (6.2.3)

Theorem 6.2.1 *The homogeneous system given by Eq. (6.2.3) has analytic solution* $\bar{x}(t) = \sum_{k=0}^{\infty} \bar{c}_k t^k$ *in the interval of convergence* $|t| < r$. *This solution* $\bar{x}(t)$ *is uniquely determined by the initial vector* $\bar{x}_0$.

Proof : Let $\bar{x}(t) = \sum_{k=0}^{\infty} \bar{c}_k t^k$, where the vector coefficient $\bar{c}_k$ are yet to determined. Plugging this representation in Eq. (6.2.3) we get

$$\begin{aligned} \sum_{k=1}^{\infty} k\bar{c}_k t^{k-1} &= \left(\sum_{k=0}^{\infty} A_k t^k\right)\left(\sum_{k=0}^{\infty} \bar{c}_k t^k\right) \\ &= \sum_{k=0}^{\infty}\left(\sum_{l=0}^{k} A_{k-l}\bar{c}_l\right) t^k \end{aligned}$$

Equivalently, we have

$$\sum_{k=0}^{\infty}(k+1)\bar{c}_{k+1} t^k = \sum_{k=0}^{\infty}\left(\sum_{l=0}^{k} A_{k-l}\bar{c}_l\right) t^k \tag{6.2.4}$$

Uniqeness fo power series in the interval $|t| < r$, gives

$$(k+1)\bar{c}_{k+1} = \sum_{l=0}^{k} A_{k-l}\bar{c}_l \tag{6.2.5}$$

and hence

$$(k+1)\,\|\bar{c}_{k+1}\| \le \sum_{l=0}^{k} \|A_{k-l}\|\ \|\bar{c}_l\| \tag{6.2.6}$$

By Theorem 6.1.3, the power series $\sum_{k=0}^{\infty} A_k t^k$ converges absolutely and uniformly in the interval $|t| \le p < r$ and hence the terms $A_k p^k$ must be uniformly bounded. That is, $\exists\ M$ such that

$$\|A_k\| p^k \le M, \quad k \ge 0 \tag{6.2.7}$$

Using the above inequality in Eq. (6.2.6), we get

$$(k+1)\|\bar{c}_{k+1}\| \le \sum_{l=0}^{k} M \frac{\|\bar{c}_l\|}{p^{k-l}}$$

Put $d_l = p^l \|\bar{c}_l\|$, then the above inequality becomes

$$(k+1) d_{k+1} \le Mp \sum_{l=0}^{k} d_l \tag{6.2.8}$$

Using Eq. (6.2.8) inductively, we get

$$\begin{aligned} d_1 &\le Mpd_0 \\ d_2 &\le \frac{Mp}{2}(d_0 + d_1) \\ &\le \frac{1}{2}\left(Mp + M^2p^2\right) d_0 \\ &\vdots \\ d_k &\le \frac{Mp(Mp+1)(Mp+2) + \cdots + (Mp+k-1)}{k!} d_0 \end{aligned} \tag{6.2.9}$$

To claim that Eq. (6.2.8) holds for all k, we need to prove this inequality by induction. So, let this inequality be true for all $r < k$. By Eq. (6.2.8), we have

$$\begin{aligned} (r+1) d_{r+1} &\le Mp\|\bar{c}_0\| \sum_{l=0}^{r} d_l \\ &\le Mp\|\bar{c}_0\| \sum_{l=0}^{r} \frac{Mp(Mp+1)(Mp+2) + \cdots + (Mp+l-1)}{l!} \end{aligned}$$

Using the notation of Example 6.1.2, set

$$u_l = \frac{Mp(Mp+1)(Mp+2) + \cdots + (Mp+l-1)}{l!}$$

Hence using Eq. (6.1.5), we get

$$\begin{aligned}(r+1)d_{r+1} &\leq Mp\|\bar{c}_0\|\sum_{l=0}^{r}u_l \\ &= \|\bar{c}_0\|(r+1)u_{r+1}\end{aligned}$$

This gives

$$(r+1)d_{r+1} \leq (r+1)\frac{Mp(Mp+1)(Mp+2)+\cdots+(Mp+r)}{(r+1)!}\|\bar{c}_0\|$$

That is

$$d_{r+1} \leq \frac{Mp(Mp+1)(Mp+2)+\cdots+(Mp+r)}{(r+1)!}\|\bar{c}_0\|$$

This proves the induction.
Hence, it follows that

$$\|\bar{c}_k\| = \frac{d_k}{p^k} \leq \frac{Mp(Mp+1)(Mp+2)+\cdots+(Mp+k-1)}{k!p^k}\|\bar{c}_0\|$$

This gives

$$\begin{aligned}\|\bar{x}\| &= \left\|\sum_{k=0}^{\infty}\bar{c}_k t^k\right\| \\ &\leq \|\bar{c}_0\| \times \left(\sum_{k=0}^{\infty}\frac{Mp(Mp+1)(Mp+2)+\cdots+(Mp+k-1)}{k!}\left(\frac{|t|}{p}\right)^k\right)\end{aligned}$$

That is

$$\|\bar{x}\| \leq \frac{\|\bar{c}_0\|}{\left(1-\frac{|t|}{p}\right)^{Mp}} \tag{6.2.10}$$

provided $|t| \leq p < r$.
This proves the existence of analytic solution of the system given by Eq. (6.2.3), This solution is uniquely determined by the initial value $\bar{x}_0$. For if $\bar{x}(t)$ and $\bar{y}(t)$ are two different solutions of the initial value problem given by Eq. (6.2.3), then $\bar{z}(t) = \bar{x}(t) - \bar{y}(t)$ is the solution of the initial value problem

$$\begin{aligned}\frac{d\bar{z}}{dt} &= A(t)\bar{z}(t) \\ \bar{z}(0) &= \bar{0}\end{aligned}$$

Since $\bar{c}_0 = 0$, it follows that $\bar{z} = \bar{0}$, thereby implying that $\bar{x} = \bar{y}$. ▮

Example 6.2.1 *(Legendre Differential Equation)*

$$(1-t^2)\frac{d^2x}{dt^2} - 2t\frac{dx}{dt} + n(n+1)x = 0 \tag{6.2.11}$$

By substituting $x_1 = x,\ x_2 = \dfrac{dx}{dt}$, *this differential equation is equivalent to the following linear homogeneous system*

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t)$$

where

$$A(t) = \begin{pmatrix} 0 & 1 \\ \dfrac{-n(n+1)}{1-t^2} & \dfrac{2t}{1-t^2} \end{pmatrix}$$

As $A(t)$ is analytic in the interval $|t| < 1$, it follows by Theorem 6.2.1 that Eq. (6.2.11) has a unique analytic solution of the form $\sum\limits_{k=0}^{\infty} c_k t^k$. We shall now determine the coefficients c_k. We substitute

$$x(t) = \sum_{k=0}^{\infty} c_k t^k, \quad \frac{dx}{dt} = \sum_{k=1}^{\infty} k c_k t^{k-1}, \quad \frac{d^2x}{dt^2} = \sum_{k=2}^{\infty} k(k-1)c_k t^{k-2}$$

in Eq. (6.2.11) to get

$$(1-t^2)\sum_{k=2}^{\infty} k(k-1)c_k t^{k-2} - 2t\sum_{k=1}^{\infty} k c_k t^{k-1} + n(n+1)\sum_{k=1}^{\infty} c_k t^k = 0$$

Equivalently, we have

$$\sum_{k=0}^{\infty} \left[(k+2)(k+1)c_{k+2} + c_k\left[n(n+1) - (k-1)k - 2k\right]\right] t^k = 0$$

The uniqueness of series representaion in $|t| < 1$ gives

$$\begin{aligned} c_{k+2} &= -\frac{[n(n+1) - (k-1)k - 2k]}{(k+2)(k+1)} c_k \\ &= -\frac{(n-k)(n+k+1)}{(k+2)(k+1)} c_k, \qquad k = 0, 1, 2, \cdots \end{aligned} \tag{6.2.12}$$

This gives us

$$\begin{aligned} c_2 &= -\frac{n(n+1)}{2!}c_0, \quad c_4 = -\frac{(n+3)(n-2)}{4.3}c_2 = \frac{(n+3)(n+1)n(n-2)}{4!}c_0 \\ c_3 &= -\frac{(n+2)(n-1)}{3!}c_1, \quad c_5 = -\frac{(n+4)(n-3)}{5.4}c_3 \\ &\qquad\qquad = \frac{(n+4)(n+2)(n-3)(n-1)}{5!}c_1 \end{aligned}$$

and hence the coefficients c_k's are inductively defined.
So, we have

$$\begin{aligned} x(t) &= c_0 x_1(t) + c_1 x_2(t) \\ x_1(t) &= 1 - \frac{n(n+1)}{2!}t^2 + \frac{(n-2)n(n+1)(n+3)}{4!}t^4 - \cdots \\ x_2(t) &= t - \frac{(n-1)(n+2)}{3!}t^3 + \frac{(n-3)(n-1)(n+2)(n+4)}{5!}t^5 - \cdots \end{aligned}$$

Since RHS of Eq. (6.2.12) is zero for $k = n$, we have $0 = c_{n+2} = c_{n+4} = \cdots$ and hence if n is even, $x_1(t)$ reduces to a polynomial of degree n and if n is odd, the same is true for $x_2(t)$. These polynomials multiplied by some constants, are called Legendre polynomials.

In Eq. (6.2.12), writing c_k in terms of c_{k+2}, we get

$$c_k = -\frac{(k+2)(k+1)}{(n-k)(n+k+1)}c_{k+2} \quad (k \le n-2)$$

This gives

$$c_{n-2} = -\frac{n(n-1)}{2(2n-1)}c_n$$

We fix c_n and then compute the lower order coefficients. It is standard to choose $c_n = 1$ when $n = 0$ and $\dfrac{2n!}{2^n(n!)^2} = \dfrac{1.3.5...(2n-1)}{n!}$ for $n \neq 0$. This defines

$$\begin{aligned} c_{n-2} &= -\frac{n(n-1)}{2(2n-1)}\frac{(2n)!}{2^n(n!)^2} \\ &= -\frac{n(n-1)2n(2n-1)(2n-2)!}{2(2n-1)2^n(n-1)!n(n-1)(n-2)!} \\ &= -\frac{(2n-2)!}{2^n(n-1)!(n-2)!} \end{aligned}$$

In general

$$c_{n-2m} = \frac{(-1)^m(2n-2m)!}{2^n m!(n-m)!(n-2m)!}$$

So we get a solution of Legendre differential equation given by Eq. (6.2.11) as

$$P_n(t) = \sum_{m=0}^{M} \frac{(-1)^m(2n-2m)!}{2^n m!(n-m)!(n-2m)!}t^{n-2m}$$

where $M = \dfrac{n}{2}$ or $\dfrac{n-1}{2}$ whichever is an integer. Some of the Legendre polynomials are given as below

$$\begin{aligned} P_0(t) &= 1,\ P_1(t) = t,\ P_2(t) = \frac{1}{2}(3t^2-1),\ P_3(t) = \frac{1}{2}(5t^3-3t) \\ P_4(t) &= \frac{1}{8}(35t^4-30t^2+3),\ P_5(t) = \frac{1}{8}(63t^5-70t^3+15t) \end{aligned}$$

We shall discuss the properties of Legendre differential equation in section 4.

Example 6.2.2 *(Hermite Differential Equation)*
In the study of the linear harmonic oscillator in quantum mechanics, one encounters the following deifferential equation

$$\frac{d^2u}{dt^2} + \left[\lambda - t^2\right] u(t) = 0 \tag{6.2.13}$$

over the interval $(-\infty,\ \infty)$. *As we look for the solution* $u \in L_2(-\infty,\ \infty)$, *we put* $u(t) = e^{\left(\frac{-t^2}{2}\right)} x(t)$. *So Eq. (6.2.13) becomes*

$$\frac{d^2x}{dt^2} - 2t\frac{dx}{dt} + (\lambda - 1)\,x = 0$$

We take $\lambda = 2n+1$ *and get the equation*

$$\frac{d^2x}{dt^2} - 2t\frac{dx}{dt} + 2nx = 0 \tag{6.2.14}$$

Eq. (6.3.13) (or alternatively Eq. (6.2.14)) is called Hermite differential equation.

As the coefficient of the differential equation given by Eq. (6.2.14) are analytic in $(-\infty, \infty)$, we get a series solution of the form

$$x(t) = \sum_{k=0}^{\infty} c_k t^k \quad \text{for} \quad -\infty < t < \infty \tag{6.2.15}$$

Plugging the representation for $x(t)$, $\dfrac{dx}{dt}$ and $\dfrac{dx^2}{dt^2}$ in Eq. (6.2.14) we see that c_k satisfy the recurrence relation

$$\begin{aligned} c_{k+2} &= \frac{2k+1-(2n+1)}{(k+2)(k+1)}c_k \\ &= \frac{2(k-n)}{(k+2)(k+1)}c_k \end{aligned} \tag{6.2.16}$$

It is clear that $c_{n+2} = 0$ and hence the solution given by Eq. (6.2.15) is a polynomial. As in the previous example, we normalize c_n and put it equal to 2^n and then compute the coefficients in descending order. This gives us the solution, which is denoted by $H_n(t)$, the Hermite polynomial. It is given by

$$H_n(t) = \sum_{m=0}^{M} \frac{(-1)^m n!}{m!(n-2m)!}(2t)^{n-2m}$$

We shall discuss the properties of the Hermite polynomials in section 4.

6.3 Linear System with Regular Singularities

A system of the form

$$\frac{d\bar{x}}{dt} = \frac{1}{t - t_0} A(t)\bar{x} \tag{6.3.1}$$

where $A(t)$ is analytic at t_0, is said to have a regular singularity at t_0.

As before, we can assume that $t_0 = 0$ and $A(t)$ has series representation of the form

$$A(t) = \sum_{k=0}^{\infty} A_k t^k \tag{6.3.2}$$

in the interval of the convergence $|t| < r$.
Eq. (6.3.1) then reduces to the equation of the form

$$\frac{d\bar{x}}{dt} = \frac{1}{t}\left(\sum_{k=0}^{\infty} A_k t^k\right)\bar{x} \tag{6.3.3}$$

We prove the following theorem, giving the existence of a series solution of the equation given by Eq. (6.3.3)

Theorem 6.3.1 *The differential equation given by Eq. (6.3.3) has a series solution of the form* $\bar{x}(t) = t^{\mu}\sum_{k=0}^{\infty} \bar{c}_k t^k$ *in the interval* $|t| < r$*, provided* μ *is an eigenvalue of* A_0 *and no other eigenvalue of the form* $\mu + n$ *exists for* A_0*, where* n *is a positive integer.*

Proof : We introduce a new dependent variable $\bar{y}(t) = t^{\mu}\bar{x}(t)$. This gives

$$\begin{aligned}
\frac{d\bar{y}}{dt} &= \frac{1}{t}\left[A(t) - \mu I\right]\bar{y}(t) \\
&= \frac{1}{t}\left[\sum_{k=0}^{\infty} A_k t^k - \mu I\right]\bar{y}(t)
\end{aligned}$$

Assuming that $\bar{y}(t) = \sum_{k=0}^{\infty} \bar{c}_k t^k$, we get

$\sum_{k=1}^{\infty} \bar{c}_k k t^k = \sum_{k=0}^{\infty} \left[\left(\sum_{l=0}^{k} A_{k-l}\bar{c}_l \right) - \mu \bar{c}_k \right] t^k$. This gives us the recurrence relations

$$\begin{aligned} (A_0 - \mu I)\bar{c}_0 &= 0 \\ (A_0 - (\mu + I))\,\bar{c}_1 &= -A_1\bar{c}_0 \\ &\vdots \\ (A_0 - (\mu + n)I)\bar{c}_n &= -\sum_{l=0}^{n-1} A_{n-l}\bar{c}_l \\ &\vdots \end{aligned} \tag{6.3.4}$$

Since $|(A_0 - (\mu+n)I)\bar{c}_n| \neq 0$ for all $n \geq 1$, the above relations iteratively define $\bar{c}_0, \bar{c}_1, \cdots, \bar{c}_n, \cdots$ We shall now show that the series $x(t) = t^{\mu}\sum_{k=0}^{\infty} \bar{c}_k t^k$ converges.

As in Section 6.2, for some positive numbers M, and $p < r$ we have

$$\|A_k\|p^k \leq M, \quad k = 0, 1, 2, \cdots \tag{6.3.5}$$

We first observe the following

(i) $|A_0 - (\mu + k)I| \neq 0$ for all positive integer k

(ii) $\lim_{k\to\infty} \left| A_0 - \frac{(\mu+k)}{k} I \right|$
$= \lim_{k\to\infty} \left| \frac{(A_0 - \mu I)}{k} - \frac{k}{k} I \right|$
$= n$

In view of property (ii), it follows that there exists a positive quantity $\epsilon > 0$ such that

$$\frac{\|A_0 - (\mu + k)I\|}{k} > \epsilon, \quad k = 1, 2, 3, \cdots$$

So, the recurrence relation given by Eq. (6.3.4) provides

$$(k+1)\bar{c}_{k+1} = [A_0 - (\mu + k + 1)I]^{-1}(k+1)\sum_{l=0}^{k} A_{k+1-l}\bar{c}_l$$

Using Eq. (6.3.5) and (ii) we get

$$(k+1)\|\bar{c}_{k+1}\| \leq \frac{M(k+1)}{\epsilon}\sum_{l=0}^{k} \frac{\|\bar{c}_l\|}{p^{k+1-l}}$$

Equivalently,

$$(k+1)\|\bar{c}_{k+1}\|\, p^{k+1} \leq \frac{M}{\epsilon}(k+1)\sum_{l=0}^{k} \|\bar{c}_l\| p^l$$

Proceeding as in Theorem 6.2.1, we have the relation

$$\|\bar{c}_k\| \leq \frac{\left(\frac{M}{\epsilon}\right)\left(\frac{M}{\epsilon}+1\right)\cdots\left(\frac{M}{\epsilon}+k-1\right)}{k!p^k}\|\bar{c}_0\|$$

for all $k \geq 1$.

Applying this estimate to the series expansion for $\bar{x}(t)$, we get

$$\begin{aligned}
\|\bar{x}(t)\| &\leq \left\| t^{\mu}\sum_{k=0}^{\infty}\bar{c}_k t^k \right\| \\
&\leq |t|^{\mu}\|\bar{c}_0\|\sum_{k=0}^{\infty}\frac{\left(\frac{M}{\epsilon}\right)\left(\frac{M}{\epsilon}+1\right)\cdots\left(\frac{M}{\epsilon}+k-1\right)}{k!}\left(\frac{|t|}{p}\right)^k \\
&= \frac{|t|^{\mu}\|\bar{c}_0\|}{\left(1-\frac{|t|}{p}\right)^{\frac{M}{\epsilon}}}
\end{aligned}$$

in the interval of uniform convergence $|t| \leq p < r$.
This proves the theorem. ∎

Example 6.3.1 *(Bessel's Equation)*

$$\begin{aligned}
t^2\frac{d^2x}{dt^2}+t\frac{dx}{dt}+(t^2-\mu^2)x(t) &= 0 \qquad (6.3.6)\\
\Longleftrightarrow \frac{d^2x}{dt^2}+\frac{1}{t}\frac{dx}{dt}+(1-\frac{\mu^2}{t^2})x(t) &= 0
\end{aligned}$$

We use the substitution $x_1 = x(t),\ \ x_2 = t\dot{x}(t)$ *to get*

$$\begin{aligned}
\frac{dx_1}{dt} &= \frac{1}{t}x_2 \\
\frac{dx_2}{dt} &= \frac{dx_1}{dt}+t\frac{d^2x_1}{dt^2} \\
&= \frac{dx_1}{dt}-\frac{dx_1}{dt}-(t-\frac{\mu^2}{t})x_1(t) \\
&= -\frac{1}{t}(t^2-\mu^2)x_1(t)
\end{aligned}$$

This reduces Eq. (6.3.6) to

$$\begin{aligned}
\frac{d\bar{x}}{dt} &= \frac{1}{t}\begin{pmatrix} 0 & 1 \\ \mu^2-t^2 & 0 \end{pmatrix}\bar{x} \\
&= \frac{1}{t}\left[A_0+A_2t^2\right]\bar{x}
\end{aligned}$$

where $A_0 = \begin{pmatrix} 0 & 1 \\ \mu^2 & 0 \end{pmatrix}$, $\quad A_2 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}$. *It is clear that* A *is analytic in* $(-\infty,\ \infty)$ *and* A_0 *has eigenvalues* $\pm\mu$.

Case 1: μ *is not an integer*

Then Eq. (6.3.6) has solution of the form $x(t) = t^{\mu}\sum_{k=0}^{\infty}\bar{c}_k t^k$ *in* $(-\infty,\ \infty)$.

For the eigenvalue $\mu > 0$, *plugging the representation*

$$\begin{aligned} x(t) &= \sum_{m=0}^{\infty} c_m x^{m+\mu}, \quad \frac{dx}{dt} = \sum_{m=0}^{\infty} c_m (m+\mu) t^{m+\mu-1} \\ \frac{d^2x}{dt^2} &= \sum_{m=0}^{\infty} c_m (m+\mu)(m+\mu-1) t^{m+\mu-2} \end{aligned}$$

in Eq. (6.3.6) we get the recurrence relation

$$\begin{aligned} c_{2m} &= -\frac{1}{2^2 m(\mu+m)} c_{2m-2},\ m = 1, 2\cdots \\ c_{2m+1} &= 0, \ \ m = 0, 1, 2, \cdots \end{aligned}$$

This gives

$$\begin{aligned} c_2 &= -\frac{c_0}{2^2(\mu+1)} \\ c_4 &= -\frac{c_2}{2^2 2(\mu+2)} = \frac{c_0}{2^4 2!(\mu+1)(\mu+2)} \\ &\vdots \\ c_{2m} &= \frac{(-1)^m c_0}{2^{2m} m!(\mu+1)(\mu+2)\cdots(\mu+m)}, \\ &\qquad\qquad m = 1, 2, \cdots \\ c_{2m+1} &= 0, \ \ m = 0, 1, 2, 3, \cdots \end{aligned} \tag{6.3.7}$$

Now define the gamma function $\Gamma(\mu)$ *as*

$$\Gamma(\mu) = \int_0^{\infty} e^{-t} t^{\mu-1} dt, \ \ \mu > 0$$

which is a convergent integral. $\Gamma(\mu)$ *satisfies the property that* $\Gamma(\mu+1) = \mu\Gamma(\mu)$ *and* $\Gamma(n+1) = n!$ *If we put* $c_0 = \frac{1}{2^{\mu}}\Gamma(\mu+1)$, *then* c_{2m} *is given by*

$$\begin{aligned} c_{2m} &= \frac{(-1)^m}{2^{2m+\mu} m!(\mu+1)(\mu+2)\cdots(\mu+m)\Gamma(\mu+1)} \\ &= \frac{(-1)^m}{2^{2m+\mu} m!\Gamma(\mu+m+1)} \end{aligned}$$

This gives us the first solution $J_\mu(t)$, called Bessl's function of order μ of the differential equation given by Eq. (6.3.6), as

$$J_\mu(t) = t^\mu \sum_{m=0}^{\infty} \frac{(-1)^m t^{2m}}{2^{2m+\mu} m! \Gamma(\mu + m + 1)}$$

Extending gamma function for $\mu < 0$, we get the representation for the second solution $J_{-\mu}(t)$ as

$$J_{-\mu}(t) = t^{-\mu} \sum_{m=0}^{\infty} \frac{(-1)^m t^{2m}}{2^{2m-\mu} m! \Gamma(-\mu + m + 1)}$$

$J_\mu(t)$ and $J_\mu(t)$ are linearly independent.

Case 2: *$\mu = n$ integer*
We have

$$J_n(t) = t^n \sum_{m=0}^{\infty} \frac{(-1)^m t^{2m}}{2^{2m+n} m! (n + m)}$$

One can easily derive that

$$J_{-n}(t) = (-1)^n J_n(t)$$

This gives

$$\begin{aligned} J_0(t) &= \sum_{m=0}^{\infty} \frac{(-1)^m t^{2m}}{2^{2m}(m!)^2} \\ &= 1 - \frac{t^2}{2^2(1!)^2} + \frac{t^4}{2^4(2!)^2} - \frac{t^6}{2^6(3!)^2} + \cdots \\ J_1(t) &= \sum_{m=0}^{\infty} \frac{(-1)^m t^{2m+1}}{2^{2m+1}(k!)(k+1)!} \\ &= t - \frac{t^3}{2^3(1!)(2!)} + \frac{t^5}{2^5(2!)(3!)} - \frac{t^7}{2^7(3!)(4!)} + \cdots \end{aligned}$$

J_0 looks similar to cosine function and J_1 looks similar to sine function. The zeros of these functions are not completely regularly spaced and also oscillations are damped as we see in the following graphics.

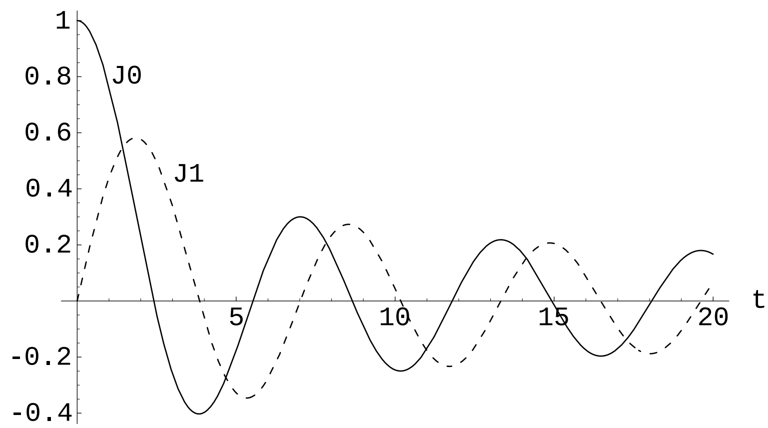

Figure 6.3.1: Sketch of J_0 and J_1

Example 6.3.2 *(Bessel's equation of order zero)*
Consider the equation

$$x\frac{d^2x}{dt^2} + \frac{dx}{dt} + x = 0$$

This is equivalent to the first order system of the form

$$\frac{d\bar{x}}{dt} = -\frac{1}{t}\begin{pmatrix} 0 & 1 \\ -t^2 & 0 \end{pmatrix}\bar{x} \tag{6.3.8}$$

where $x_1 = x,\ x_2 = t\dfrac{dx}{dt}$.

As we have seen before, a solution of this system is given by

$$x_1(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{4^k (k!)^2} \tag{6.3.9}$$

$$x_2(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{4^k (k!)^2} \tag{6.3.10}$$

To get another solution $\bar{y}(t)$ *linearly independent of* $\bar{x}(t)$ *and satisfying Eq. (6.3.8), we set*

$$\bar{y}(t) = \psi(t)\bar{z}(t) \ \ \textit{where} \ \ \psi(t) = \begin{pmatrix} 1 & x_1 \\ 0 & x_2 \end{pmatrix}.$$

This gives

$$\frac{d\bar{y}}{dt} = \frac{d\phi}{dt}\bar{z}(t) + \psi(t)\frac{d\bar{z}}{dt} = A(t)\psi(t)\bar{x}(t)$$

That is

$$\begin{pmatrix} 0 & \frac{1}{t}x_2 \\ 0 & -tx_1 \end{pmatrix}\bar{z} + \begin{pmatrix} 1 & x_1 \\ 0 & x_2 \end{pmatrix}\frac{dz}{dt} = \begin{pmatrix} 0 & \frac{1}{t}x_2 \\ -t & -tx_1 \end{pmatrix}\bar{z}(t)$$

Solving for $\frac{dz}{dt}$ *we get*

$$\begin{pmatrix} 1 & x_1 \\ 0 & x_2 \end{pmatrix}\frac{dz}{dt} = \begin{pmatrix} 0 & 0 \\ -t & 0 \end{pmatrix}\bar{z}(t)$$

and hence

$$\frac{dz}{dt} = \begin{pmatrix} t\frac{x_1}{x_2} & 0 \\ -\frac{t}{x_2} & 0 \end{pmatrix}\bar{z}(t)$$

So, we need to solve

$$\frac{dz_1}{dt} = t\frac{x_1}{x_2}z_1 = -\frac{\dot{x_2}}{x_2}z_1 \tag{6.3.11a}$$

$$\frac{dz_2}{dt} = -\frac{t}{x_2}z_1 \tag{6.3.11b}$$

Integrating Eq. (6.3.11(a)) - Eq. (6.3.11(b)) we get

$$z_1 = \frac{c_1}{x_2}, \quad z_2 = -c_1\int \frac{t}{x_2^2(t)}dt$$

Hence $\bar{y}(t) = \psi(t)\bar{z}(t)$ *is given by*

$$\bar{y}(t) = \begin{pmatrix} \frac{c_1}{x_2} - c_1x_1\int \frac{tdt}{x_2^2(t)} \\ -c_1x_2\int \frac{t}{x_2^2(t)}dt \end{pmatrix}$$

That is

$$\begin{aligned} y_1(t) &= \frac{c_1}{x_2} - c_1x_1\int \left[\frac{-\dot{x_2}/x_1}{x_2^2}\right]dt \\ &= -c_1x_1\int \frac{\dot{x_2}}{x_2x_1^2}dt \\ &= -c_1x_1\int \frac{1}{tx_1^2}dt \end{aligned}$$

$$y_2(t) = -c_1 x_2 \int \frac{t}{x_2^2} dt$$

Using the represntations of Eq. (6.3.9) - Eq. (6.3.10), we get

$$\begin{aligned} x_1(t) &= 1 - \frac{t^2}{4} + \cdots \\ x_2(t) &= -\frac{t^2}{2} + \frac{t^4}{16} + \cdots \end{aligned}$$

This gives

$$\begin{aligned} x_1 \int \frac{dt}{t x_1^2} &= x_1 \int \frac{1 + \frac{t^2}{2} + \cdots}{t} dt \\ &= x_1 lnt + x_1 (\frac{t^2}{4} + \cdots) \end{aligned}$$

and

$$\begin{aligned} x_2 \int \frac{t}{x_2^2} dt &= 4x_2 \int \frac{t\left(1 + \frac{t^2}{4} + \cdots\right)}{t^4} dt \\ &= \frac{-2x_2}{t^2} + x_2 lnt + \cdots \end{aligned}$$

Hence, it follows that the second solution $\bar{y}(t)$*, linearly independent of* $x(t)$*, is given by*

$$y(t) = x(t) lnt + x(t) \left(1 + \frac{t^2}{4} + \cdots\right)$$

To be more precise, one can show that

$$y(t) = J_0(t) lnt + \sum_{m=1}^{\infty} \frac{(-1)^m h_m t^{2m}}{2^{2m} (m!)^2}$$

where $h_m = \left(1 + \frac{1}{2} + \cdots + \frac{1}{m}\right)$.

The function $Y_0(t)$ *defined as*

$$Y_0(t) = a y(t) + b J_0(t), a = \frac{2}{\pi},\ b = r - ln2$$

where r *is the Euler's constant* $\lim_{n \to \infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{m} - lnn\right)$ *is called the Bessel's function of the second kind (order zero). Hence, it is given by*

$$Y_0(t) = \frac{2}{\pi} \left[J_0(t) \left(ln\frac{t}{2} + r \right) + \sum_{m=1}^{\infty} \frac{(-1)^m h_m t^{2m}}{2^{2m} (m!)^2} \right]$$

6.4 Properties of Legendre, Hermite and Bessel Functions

Legendre polynomial $P_n(t)$ is a solution of the differential equation

$$(1-t^2)\frac{d^2x}{dt^2} - 2t\frac{dx}{dt} + n(n+1)x(t) = 0 \tag{6.4.1}$$

In terms of the notation of Chapter 6, this is eigenvalue problem of the form

$$Lx = \lambda x \tag{6.4.2}$$

where L is the second order differential operator

$$Lx = (1-t^2)\frac{d^2x}{dt^2} - 2t\frac{dx}{dt} = \frac{d}{dt}\left((1-t^2)\frac{dx}{dt}\right) \tag{6.4.3}$$

We can think of L as an operator defined on $L_2[-1,1]$ with $D(L)$ as the set of all functions having second order derivatives.

Then Green's formula (to be done in Section 7.1) gives

$$\int_{-1}^{1} (yLx - xL^*y)\,dt = J(x,y)\Big|_{-1}^{1}$$

where $J(x,y)|_{-1}^{1} = a_0\left(y\frac{dx}{dt} - x\frac{dy}{dt}\right)\Big|_{-1}^{1}$ with $a_0(t) = (1-t^2)$.

As $a_0(1) = a_0(-1) = 0$, it follows that $J(x,y)|_{-1}^{1} = 0$ and hence

$$\int_{-1}^{1} (yLx - xL^*y)\,dt = 0 \quad \forall x,\ y \in D(L) = D(L^*)$$

That is $L^* = L$ and $D(L^*) = D(L)$.

We appeal to Theorem 7.5.1, of Chapter 7 to claim that the eigenfunctions of L form a complete orthonormal set. We have already proved in Example 6.2.1 that Eq. (6.4.1) has sequence of eigenvalues $\{n(n+1)\}$ with $\{P_n(t)\}$ corresponding eigenfunctions and hence they are orthogonal. However, for the sake of completeness we prove this result directly.

Theorem 6.4.1 *The set of Legendre polymonials $\{P_n(t)\}$ satisfying Eq. (6.4.1) are orthogonal set of polynomials in $L_2[-1,1]$.*

Proof : We have

$$\frac{d}{dt}\left((1-t^2)\frac{dP_n}{dt}\right) = -n(n+1)P_n(t) \tag{6.4.4}$$

$$\frac{d}{dt}\left((1-t^2)\frac{dP_m}{dt}\right) = -m(m+1)P_m(t) \tag{6.4.5}$$

Multiplying Eq. (6.4.4) by $P_m(t)$ and Eq. (6.4.5) by $P_n(t)$ and substracting the equations and integrating, we get

$$\begin{aligned}
& \int_{-1}^{1} [m(m+1) - n(n+1)] P_m(t)P_n(t)dt \\
= & \int_{-1}^{1} \left[\frac{d}{dt}\left((1-t^2)\frac{dP_n}{dt}\right)P_m(t) - \frac{d}{dt}\left((1-t^2)\frac{dP_m}{dt}P_n(t)\right)\right] dt \\
= & \left[(1-t^2)\left[\dot{P}_n(t)P_m(t) - \dot{P}_m(t)P_n(t)\right]\right]_{-1}^{1} \\
& - \int_{-1}^{1} \left[(1-t^2)\left[\dot{P}_n(t)\dot{P}_m(t) - \dot{P}_m(t)\dot{P}_n(t)\right]\right] \\
= & \ 0
\end{aligned}$$

Hence $(P_m, P_n) = \int_{-1}^{1} P_m(t)P_n(t)dt = 0$. That is $P_m \perp P_n$. ∎

Theorem 6.4.2 *The Legendre polynomials $P_n(t)$ satisfy the Rodrigues formula*

$$P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}\left[(t^2-1)^n\right], \quad n = 0, 1, 2, \cdots \tag{6.4.6}$$

Proof : We have

$$P_n(t) = \frac{1}{2^n n!}\sum_{k=0}^{M} \frac{(-1)^k n!}{k!(n-k)!}\frac{(2n-2k)!}{(n-2k)!}t^{n-2k}$$

Since $\frac{d^n}{dt^n}\left[t^{2n-2k}\right] = \frac{(2n-2k)!}{(n-2k)!}t^{n-2k}$, it follows that

$$P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}\left[\sum_{k=0}^{M} \frac{(-1)^k n!}{k!(n-k)!}t^{2n-2k}\right]$$

It is now clear that the sum $\left[\sum_{k=0}^{M} \frac{(-1)^k n!}{k!(n-k)!}t^{2n-2k}\right]$ is the binomial expansion of $(t^2-1)^n$ and hence

$$P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2-1)^n$$

∎

Definition 6.4.1 *Let $\{f_n(t)\}$ be a sequence of functions in an interval I. A function $F(t,u)$ is said to be a generating function of this sequence if*

$$F(t,u) = \sum_{n=0}^{\infty} f_n(t)u^n \tag{6.4.7}$$

We have the following theorem giving the generating function for the sequence $\{P_n(t)\}$ of Legendre polynomials.

Theorem 6.4.3 *For Legendre polynomials $\{P_n(t)\}$, we have*

$$(1-2tu+u^2)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} P_n(t)u^n \tag{6.4.8}$$

Proof : Let $|t| \le r$ (arbitary positive number) and $|u| < (1+r^2)^{\frac{1}{2}} - r$. Then

$$\begin{aligned} |2tu - u^2| &\le 2|t||u| + |u^2| \\ &\le 2r(1+r^2)^{\frac{1}{2}} - 2r^2 + (1+r^2) + r^2 - 2r(1+r^2)^{\frac{1}{2}} \\ &= 1 \end{aligned}$$

So expanding $(1-2tu+u^2)^{-\frac{1}{2}}$ in a binomial series, we get

$$\begin{aligned} [1-u(2t-u)]^{-\frac{1}{2}} &= 1 + \frac{1}{2}u(2t-u) + \frac{1}{2}\frac{3}{4}u^2(2t-u)^2 + \cdots \\ &\quad + \frac{1.3......(2n-1)}{1.2.....(2n)}u^n(2t-u)^n + \cdots \end{aligned}$$

The coefficient of u^n in this expression is

$$\begin{aligned} & \frac{1.3......(2n-1)}{2.4....(2n)}(2t)^n - \frac{1.3......(2n-3)}{2.4....(2n-2)}(2t)^{n-2} + \cdots \\ = & \frac{1.3......(2n-1)}{n!}\left[t^n - \frac{n(n-1)}{(2n-1)^2}t^{n-2} + \frac{n(n-1)(n-2)(n-3)}{(2n-1)(2n-3)2.4}t^{n-4} + \cdots\right] \\ = & P_n(t) \end{aligned}$$

∎

Theorem 6.4.4 *The Legendre polynomials satisfy the following recurrence relation*

$$(n+1)P_{n+1}(t) = (2n+1)tP_n(t) - nP_{n-1}(t), \quad n = 1, 2, ... \tag{6.4.9}$$

Proof : Differentiating the relation given by Eq. (6.4.8) with respect to u, we get

$$(t-u)(1-2tu+u^2)^{-\frac{3}{2}} = \sum_{n=1}^{\infty} nP_n(t)u^{n-1}$$

This gives

$$(t-u)(1-2tu+u^2)^{-\frac{1}{2}} = (1-2tu+u^2)\sum_{n=0}^{\infty} nP_n(t)u^{n-1}$$

That is

$$(t-u)\sum_{n=1}^{\infty} P_n(t)u^n = (1-2tu+u^2)\sum_{n=0}^{\infty} nP_n(t)u^{n-1}$$

Equivalently,

$$\begin{aligned} & \sum_{n=0}^{\infty} tP_n(t)u^n - \sum_{n=0}^{\infty} P_n(t)u^{n+1} \\ = & \sum_{n=1}^{\infty} nP_n(t)u^{n-1} - 2\sum_{n=1}^{\infty} ntP_n(t)u^n \\ & + \sum_{n=0}^{\infty} nP_n(t)u^{n+1} \end{aligned}$$

Rewriting, we get

$$\begin{aligned} & \sum_{n=1}^{\infty} tP_n(t)u^n - \sum_{n=1}^{\infty} P_{n-1}(t)u^n \\ = & \sum_{n=1}^{\infty}(n+1)P_{n+1}(t)u^n - 2\sum_{n=1}^{\infty} ntP_n(t)u^n \\ & + \sum_{n=1}^{\infty}(n-1)P_{n-1}(t)u^n \end{aligned}$$

Comparing the coefficient of u^n, we get

$$(n+1)P_{n+1}(t) = (2n+1)tP_n(t) - nP_{n-1}(t)$$

∎

Corollary 6.4.1

$$\|P_n\|^2 = \int_{-1}^{1} P_n^2(t)dt = \frac{2}{2n+1} \tag{6.4.10}$$

Proof : Recurrence relation gives

$$\begin{aligned} (2n+1)tP_n(t) &= (n+1)P_{n+1}(t) + nP_{n-1}(t) \\ (2n-1)tP_{n-1}(t) &= nP_n(t) + (n-1)P_{n-2}(t) \end{aligned}$$

These two equations give

$$\begin{aligned} 0 &= \int_{-1}^{1} \left[tP_n(t)P_{n-1}(t) - tP_{n-1}(t)P_n(t)\right]dt \\ &= \frac{(n+1)}{(2n+1)}\int_{-1}^{1} P_{n+1}(t)P_{n-1}(t) + \frac{n}{(2n+1)}\int_{-1}^{1} P_{n-1}^2(t)dt \\ & -\frac{(n-1)}{(2n-1)}\int_{-1}^{1} P_{n-2}(t)P_n(t) - \frac{n}{(2n-1)}\int_{-1}^{1} P_n^2(t)dt \end{aligned}$$

This implies that

$$\int_{-1}^{1} P_n^2(t)dt = \frac{(2n-1)}{(2n+1)} \int_{-1}^{1} P_{n-1}^2(t)dt$$

and hence

$$\begin{aligned} \int_{-1}^{1} P_n^2(t)dt &= \frac{(2n-1)}{(2n+1)} \frac{(2n-3)}{(2n-1)} \cdots \frac{2}{1} \int_{-1}^{1} P_0^2(t)dt \\ &= \frac{2}{(2n+1)} \end{aligned}$$

▮

We now examine the properties of the Hermite polynomials. Recall that Hermite polynomial $H_n(t)$ satisfy the differential equation

$$\frac{d^2x}{dt^2} - 2t\frac{dx}{dt} + 2nx = 0$$

and is given by

$$H_n(t) = \sum_{k=0}^{\infty} \frac{(-1)^n n!}{k!(n-2k)!}(2t)^{n-2k}$$

We have the following theorem concerning the orthogonality of $\{H_n(t)\}$ in $L_2(-\infty, \infty)$ with respect to the weight function e^{-t^2}.

Theorem 6.4.5 *The Hermite Polynomials $H_n(t)$ are orthogonal set of polynomials in the space $L_2(-\infty, \infty)$ with respect to the weight function e^{-t^2}.*

Theorem 6.4.6 *For Hermite polynomials $H_n(t)$, we have the following formulae.*

(i) Rodgriues formula

$$H_n(t) = (-1)e^{t^2} \frac{d^n}{dt^n} e^{-t^2}$$

(ii) Generating function

$$e^{2tu-u^2} = \sum_{n=0}^{\infty} H_n(t) \frac{u^n}{n!}$$

(iii) Recurrence relation

$$H_{n+1}(t) = 2tH_n(t) - 2nH_{n-1}(t), \quad H_0 = 1, H_1 = 2t$$

We now enunciate the properties of the Bessel's function $J_n(t)$.

Theorem 6.4.7 *For each fixed nonnegative integer n, the sequence of Bessel functions $J_n(k_m t)$, where k_m are the zeros of $J_n(k)$, form an orthogonal set in $L_2[0,1]$ with respect to the weight function t. That is*

$$\int_0^1 tJ_n(k_l t)J_n(k_m t)dt = 0, \quad l \neq m \tag{6.4.11}$$

Proof : The Bessel function $J_n(t)$ satisfies the differential equation (refer Eq. (6.3.6))

$$t^2 \ddot{J}_n(t) + t\dot{J}_n(t) + (t^2 - n^2)J_n(t) = 0$$

Set $t = ks$, then the above equation reduces to the differential equation

$$\frac{d}{ds}\left[s\dot{J}_n(ks)\right] + \left(-\frac{n^2}{s} + k^2 s\right) J_n(ks) = 0 \tag{6.4.12}$$

Let k_m be the zeros of the Bessel function $J_n(k)$, then we have

$$\frac{d}{ds}\left[s\dot{J}_n(k_l s)\right] + \left(-\frac{n^2}{s} + k_l^2 s\right) J_n(k_l s) \quad = \quad 0 \tag{6.4.13}$$

and

$$\frac{d}{ds}\left[s\dot{J}_n(k_m s)\right] + \left(-\frac{n^2}{s} + k_m^2 s\right) J_n(k_m s) = 0 \tag{6.4.14}$$

Eq. (6.4.13) - Eq. (6.4.14) give

$$\begin{aligned} & (k_l^2 - k_m^2) \int_0^1 sJ_n(k_l s)J_n(k_m s)ds \\ = & \int_0^1 \left[\frac{d}{ds}\left[s\dot{J}_n(k_l s)\right] J_n(k_m s) - \frac{d}{ds}\left[s\dot{J}_n(k_m s)\right] J_n(k_l s)\right] ds \\ = & \ 0, \qquad k_l \neq k_m \end{aligned}$$

using the fact that $J_n(k_l) = 0 = J_n(k_m)$. ∎

Theorem 6.4.8 *The generating function for the sequence $\{J_n(t)\}$ of Bessel functions is given by*

$$exp\left(\frac{1}{2}t\left(u - \frac{1}{u}\right)\right) = \sum_{n=-\infty}^{\infty} J_n(t)u^n \tag{6.4.15}$$

and hence $J_{-n}(t) = (-1)^n J_n(t)$.

Proof : Expanding $exp\left(\frac{1}{2}t\left(u-\frac{1}{u}\right)\right)$ in powers of u, we get

$$\begin{aligned} exp\left(\frac{1}{2}t\left(u-\frac{1}{u}\right)\right) &= \left(exp\frac{1}{2}tu\right)\left(exp\frac{1}{2}\left(\frac{-t}{u}\right)\right) \\ &= \left(\sum_{k=0}^{\infty}\frac{(tu)^k}{2^k k!}\right)\left(\sum_{l=0}^{\infty}\frac{(-t)^l}{2^l l!\ u^l}\right) \\ &= \sum_{-\infty}^{\infty} c_n(t)u^n \end{aligned}$$

The coefficient $c_n(t)$ of u^n in the above expression is $\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{t}{2}\right)^{2k+n}}{k!(k+n)!} = J_n(t)$ and hence we get the generating function relation given by Eq. (6.4.15) for $J_n(t)$.

If we replace u by $-\frac{1}{v}$ in Eq. (6.4.15), we get

$$\begin{aligned} exp(\frac{1}{2}(v-\frac{1}{v})t) &= \sum_{n=-\infty}^{\infty} J_n(t)(-1)^n v^{-n} \\ &= \sum_{n=-\infty}^{\infty} J_{-n}(t)v^n \end{aligned}$$

Hence, it follows that $\quad J_{-n}(t) = (-1)^n J_n(t)$

∎

Theorem 6.4.9 *The Bessel functions $J_\mu(t)$ satisfy the following recurrence relations*

$$J_{\mu-1}(t) + J_{\mu+1}(t) = \frac{2\mu}{t}J_\mu(t) \tag{6.4.16a}$$

$$J_{\mu-1}(t) - J_{\mu+1}(t) = 2J_\mu(t) \tag{6.4.16b}$$

Proof : We have

$$J_\mu(t) = t^\mu\left(\sum_{k=0}^{\infty}\frac{(-1)^k t^{2k}}{2^{2k+\mu}(k!)\Gamma(\mu+k+1)}\right)$$

Multiplying $J_\mu(t)$ by t^μ and pulling $t^{2\mu}$ under the summation, we have

$$t^\mu J_\mu(t) = \sum_{k=0}^{\infty}\frac{(-1)^k t^{2k+2\mu}}{2^{2k+\mu}(k!)\Gamma(\mu+k+1)}$$

Differentiating the above expression we have

$$\begin{aligned} \frac{d}{dt}\left(t^{\mu} J_{\mu}(t)\right) &= \sum_{k=0}^{\infty} \frac{(-1)^k 2(k+\mu) t^{2k+2\mu-1}}{2^{2k+\mu}(k!)\Gamma(\mu+k+1)} \\ &= t^{\mu} t^{\mu-1} \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{2^{2k+\mu-1}(k!)\Gamma(\mu+k)} \\ &= t^{\mu} J_{\mu-1}(t) \end{aligned}$$

Thus, we have

$$\frac{d}{dt}\left(t^{\mu} J_{\mu}(t)\right) = t^{\mu} J_{\mu-1}(t)$$

and

$$\frac{d}{dt}\left(t^{-\mu} J_{\mu}(t)\right) = -t^{-\mu} J_{\mu+1}(t)$$

Expanding the LHS of the above expressions, we have

$$\begin{aligned} \mu t^{\mu-1} J_{\mu} + t^{\mu} \dot{J}_{\mu} &= t^{\mu} J_{\mu-1} \\ -\mu t^{\mu-1} J_{\mu} + t^{\mu} \dot{J}_{\mu} &= -t^{\mu} J_{\mu+1} \end{aligned}$$

Adding and substracting the above relations, we get

$$\begin{aligned} J_{\mu-1} + J_{\mu+1} &= \frac{2\mu}{t} J_{\mu}(t) \\ J_{\mu-1} - J_{\mu+1} &= 2\dot{J}_{\mu}(t) \end{aligned}$$

▮

Corollary 6.4.2 *From the definition of $J_{\mu}(t)$, we have*

$$J_{\frac{1}{2}}(t) = \sqrt{\frac{2}{\pi t}} \sin t, \quad J_{-\frac{1}{2}}(t) = \sqrt{\frac{2}{\pi t}} \cos t$$

and hence the above recurrence relations give

$$J_{\frac{3}{2}}(t) = \sqrt{\frac{2}{\pi t}} \left(\frac{\sin t}{t} - \cos t\right)$$

$$J_{-\frac{3}{2}}(t) = \sqrt{\frac{2}{\pi t}} \left(-\frac{\cos t}{t} - \sin t\right)$$

We have the following graphics for the above functions.

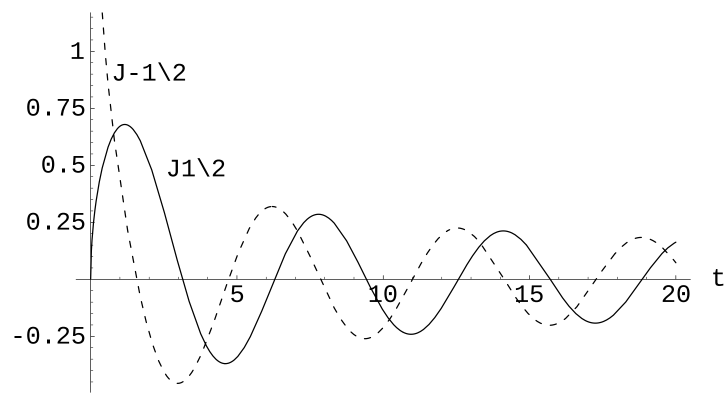

Figure 6.4.2: Sketch of $J_{-1/2}(t)$ and $J_{1/2}(t)$

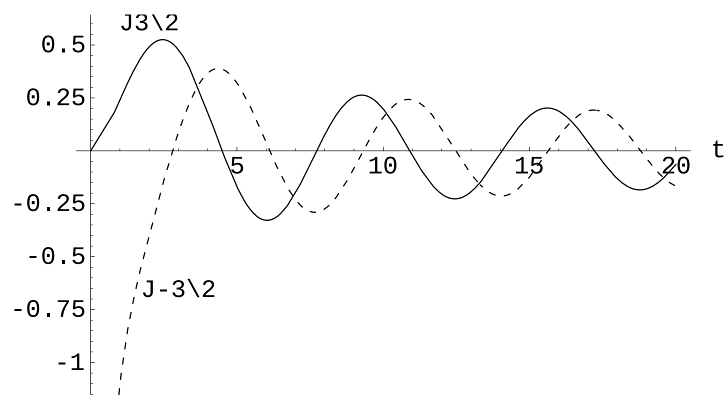

Figure 6.4.3: Sketch of $J_{3/2}(t)$ and $J_{-3/2}(t)$

For more details on varioius topics in this chapter, refer Agarwal and Gupta [1], Brown and Churchill [2], Hochstadt [3] and Kreyzig [4].

6.5 Exercises

1. Let the functions $f, g : \Re \to \Re$ be defined as follows

$$f(t) = \begin{cases} \dfrac{\sin t}{t}, & t \neq 0 \\ 1, & t = 0 \end{cases} \quad , \quad g(t) = \begin{cases} \dfrac{1-\cos t}{t^2}, & t \neq 0 \\ \dfrac{1}{2}, & t = 0 \end{cases}$$

Show that f and g are analytic at $t = 0$.

2. Find the series solution of the following IVPs

(i) $\ddot{x} + t\dot{x} - 2x = 0, \;\; x(0) = 1, \;\; \dot{x}(0) = 0$

(ii) $t(2-t)\ddot{x} - 6(t-1)\dot{x} - 4x = 0, \;\; x(1) = 1, \;\; \dot{x}(1) = 0$

(iii) $\ddot{x} + e^t\dot{x} + (1+t^2)x = 0, \;\; x(0) = 1, \;\; \dot{x}(0) = 0$

(iv) $\ddot{x} - (\sin t)x = 0, \;\; x(\pi) = 1, \;\; \dot{x}(\pi) = 0$

3. Show that the general solution of the Chebyshev differential equation

$$(1-t^2)\ddot{x} - t\dot{x} + a^2 x = 0$$

is given by

$$\begin{aligned} x(t) &= c_0\left[1 + \sum_{n=1}^{\infty} \frac{(-a)^2(2^2-a^2)......((2n-2)^2-a^2)}{(2n)!} t^{2n}\right] \\ &+ c_1\left[t + \sum_{n=1}^{\infty} \frac{(1-a^2)(3^2-a^2)......((2n-1)^2-a^2)}{(2n-1)!} t^{2n+1}\right] \end{aligned}$$

4. Show that the two linearly independent solutions of the following differential equation

$$t^2\ddot{x} + t(t-\frac{1}{2})\dot{x} + \frac{1}{2}x = 0$$

are given by

$$\begin{aligned} x_1(t) &= |t| \sum_{n=0}^{\infty} (-1)^n \frac{(2t)^n}{(2n+1)(2n-1)......3.1} \\ x_2(t) &= |t|^{\frac{1}{2}} \sum_{n=0}^{\infty} (-1)^n \frac{t^n}{n!} \end{aligned}$$

5. Show that the two linearly independent solutions of the differential equation

$$t(1-t)\ddot{x} + (1-t)\dot{x} - x = 0$$

are given by

$$
\begin{aligned}
x_1(t) &= 1+\sum_{n=1}^{\infty}\frac{1.2.5......((n-1)^2+1)}{(n!)^2}t^n \\
x_2(t) &= x_1(t)ln|t|+2\left(\sum_{n=1}^{\infty}\frac{1.2.5.....((n-1)^2+1)}{(n!)^2}\right) \\
&\quad \times\left(\sum_{k=1}^{n}\frac{k-2}{k((k-1)^2+1)}\right)t^n
\end{aligned}
$$

6. Using Rolle's theorem show that between two consucative zeros of $J_n(t)$, there exists precisely one zero of $J_{n+1}(t)$.

7. Prove the following identities for Bessel functions $J_\mu(t),\ \mu\in\Re$.

   (i) $\int t^\mu J_{\mu-1}(t)dt = t^\mu J_\mu(t)+c$

   (ii) $\int t^{-\mu}J_{\mu+1}(t)dt = -t^{-\mu}J_\mu(t)+c$

   (iii) $\int J_{\mu+1}(t)dt = \int J_{\mu-1}(t)dt - 2J_\mu(t)$

8. Show that

$$\|J_n(k_{lm}t)\|^2 = \int_0^1 tJ_n^2(k_{lm}t)dt = \frac{1}{2}J_{n+1}^2(k_{lm})$$

9. Using Rodrigues formula, show that

$$P_n(0) = \begin{cases} 0, & n \text{ is odd} \\ (-1)^{n/2}\dfrac{1.3....n-1}{2.4.....n}, & n \text{ is even} \end{cases}$$

10. Let $p(t)$ be a polynomial of degree n. Show that $p(t)$ is orthogonal to all Legendre polynomials of degree strictly less than n.

References

[1] Ravi P. Agarwal and Ramesh C. Gupta, *Essentials of Ordinary differential Equations, McGraw Hill Book Company, 1993*

[2] James Ward Brown and Rulev Churchill, *Fourier Series and Boundary Value Problems, McGraw Hill, Inc., 1993*

[3] Harry Hochstadt, *Differential Equations, A Modern Approach, Dover Publications, Inc., 1964*

[4] Erwin Kreyszig, *Advanced Engineering Mathematics, John Wiley and Sons , Inc., 1999*

Chapter 7

Boundary Value Problems

In this chapter we mainly deal with boundary value problems involving second order differential operators. The main tool for solvability analysis of such problems is the concept of Green's function. This transforms the differential equation in to integral equation, which, at times, is more informative.
Hence there is a detailed discussion involving the definition of Green's function through Dirac-delta function, its properties and construction. Numerous examples are given for illustration.
Towards the end, we give eigenfunction expansion technique for computing the solution of boundary value problems. Nonlinear boundary value problem is also the topic of our investigation. We use both fixed point theory and monotone operators for this analysis.

7.1 Introduction

We shall be concerned with the solution of the ordinary differential equation

$$Lx = f \tag{7.1.1}$$

over the interval $a \leq t \leq b$, subject to certain boundary conditions. Here L is a second order linear differential operator of the form

$$L = a_0(t)\frac{d^2}{dt^2} + a_1(t)\frac{d}{dt} + a_2(t), \quad a_0(t) \neq 0 \text{ on } [a,b]$$

Since L is of second order, there will be two boundary conditions of the form

$$B_i(x) = C_i, \quad i = 1, 2 \tag{7.1.2}$$

where C_i's are given constants, B_i's are functions of the unknown variable x. Throughout this chapter, we shall limit ourselves to those B_i's which are linear combinations of x and its derivative. So the most general form of the boundary condition Eq. (7.1.2) will be written as

$$\begin{aligned} B_1(x) &\equiv \alpha_{11}x(a)+\alpha_{12}\dot{x}(a)+\beta_{11}x(b)+\beta_{12}\dot{x}(b)=C_1 \\ B_2(x) &\equiv \alpha_{21}x(a)+\alpha_{22}\dot{x}(a)+\beta_{21}x(b)+\beta_{22}\dot{x}(b)=C_2 \end{aligned} \tag{7.1.3}$$

where the coefficients $\alpha_{ij},\ \beta_{ij}$ are given real numbers. To ensure that Eq. (7.1.3) gives two distinct boundary conditions we assume that the row vectors $(\alpha_{11},\ \alpha_{12},\ \beta_{11},\ \beta_{12})$ and $(\alpha_{21},\ \alpha_{22},\ \beta_{21},\ \beta_{22})$ are linearly independent .

If $C_1 = C_2 = 0$, we say that boundary conditions are homogeneous,

If $\beta_{11} = \beta_{12} = \alpha_{21} = \alpha_{22} = 0$, the boundary conditions are called unmixed. Eq. (7.1.3) then reduces to

$$\begin{aligned} \alpha_{11}x(a)+\alpha_{12}\dot{x}(a) &= C_1 \\ \beta_{21}x(b)+\beta_{22}\dot{x}(b) &= C_2 \end{aligned} \tag{7.1.4}$$

If $\alpha_{12} = \beta_{11} = \beta_{12} = \alpha_{21} = \beta_{21} = \beta_{22} = 0$, we have initial conditions

$$\begin{aligned} x(a) &= C_1 \\ \dot{x}(a) &= C_2 \end{aligned}$$

The boundary conditions are periodic if they are of the form

$$x(a) = x(b), \quad \dot{x}(a) = \dot{x}(b) \tag{7.1.5}$$

Example 7.1.1

$$\begin{aligned} \frac{d^2x}{dt^2} &= f(t) \\ x(0) &= x(1) = 0 \end{aligned} \tag{7.1.6}$$

This denotes the deflection of a string stretched under a unit tension at fixed end points 0 *and* 1 *and subjected to force distribution* $f(t)$*. Writing Eq. (7.1.6) in the form Eq. (7.1.1), we have*

$$Lx = f, \quad B_1(x) = 0 = B_2(x)$$

where $L \equiv \dfrac{d^2}{dt^2},\ B_1(x) \equiv x(0), \quad B_2(x) \equiv x(1).$

It is clear that we have homogeneous unmixed boundary conditions.

Example 7.1.2 *Consider a rod of length unity and unit cross-section. We assume that the rod is homogeneous and the flow of heat is in the direction of* t *only. Suppose there is a heat source of density* $f(t)$ *in the interior, then the temperature distribution* $x(t)$ *satisfies*

$$-\frac{d}{dt}\left(K\frac{dx}{dt}\right) = f(t)$$

Here, K is the thermal conductivity of the rod, which is a function of t.

If the end points of the rod are kept at zero temperature we have the boundary value problem

$$Lx = f, \quad B_1(x) = 0 = B_2(x)$$

where $L \equiv -\dfrac{d}{dt}\left(K\dfrac{d}{dt}\right), \quad B_1(x) \equiv x(0), \quad B_2(x) \equiv x(1).$

Here again, we have homogeneous unmixed boundary conditions.

Example 7.1.3 *Consider the motion of a particle along a straight line. The force on the particle is along the line and the particle starts from rest at the origin. Then the displacement $x(t)$ satisfies Newton's law*

$$m\frac{d^2x}{dt^2} = f(t)$$

with initial conditions $x(0) = 0, \quad \dfrac{dx}{dt}(0) = 0.$

We can write this in our standard form as

$$Lx = f, \quad B_1(x) = 0 = B_2(x)$$

where $L \equiv m\dfrac{d^2}{dt^2}, \quad B_1(x) \equiv x(0), \quad B_2(x) \equiv \dot{x}(0).$

This is an initial-value problem.

7.2 Adjoint of a Differential Operator

As before, we shall denote by $L_2[a,b]$ the Hillbert space of all square integrable real valued functions on $[a,b]$.

Definition 7.2.1 *Let $L : L_2[a,b] \to L_2[a,b]$ be a second order differential operator. By domain $D(L)$, we mean a subspace of $L_2[a,b]$ which consists of functions which have piecewise-continuous derivatives of second order and which satisfy homogeneous boundary conditions*

$$B_1(x) = 0 = B_2(x)$$

Definition 7.2.2 *Let L be the differential operator acting on the Hilbert space $L_2[a,b]$ with domain D. Then, using Definition 2.2.6 the adjoint operator $L^* : L_2[a,b] \to L_2[a,b]$ is defined implicitly by the equation*

$$(y, Lx) = (L^*y, x); \quad x \in D(L) \text{ and } y \in D(L^*) \tag{7.2.1}$$

$D(L^*)$ will be obtained in the following steps:

$$\begin{aligned}
(y, Lx) &= \int_a^b yLx dt \\
&= \int_a^b y(a_0\ddot{x} + a_1\dot{x} + a_2x)\, dt
\end{aligned}$$

On the RHS, integrating by parts we get

$$\begin{aligned}(y, Lx) &= [a_0(y\dot{x} - x\dot{y}) + xy(a_1 - \dot{a}_0)]_a^b \\ &+ \int_a^b x\left[\frac{d^2}{dt^2}(a_0 y) - \frac{d}{dt}(a_1 y) + a_2 y\right] dt\end{aligned}$$

Applying the definition of L^* we see that

$$\begin{aligned}(L^* y, x) &= \int_a^b x\left[\frac{d^2}{dt^2}(a_0 y) - \frac{d}{dt}(a_1 y) + a_2 y\right] dt \\ &+ [a_0(y\dot{x} - x\dot{y}) + xy(a_1 - \dot{a}_0)]_a^b \qquad (7.2.2)\end{aligned}$$

Thus we see that L^* consists of two parts, a second order differential operator

$$\frac{d^2}{dt^2}(a_0 y) - \frac{d}{dt}(a_1 y) + a_2 y \qquad (7.2.3)$$

appearing in the integrand and some boundary terms.

Definition 7.2.3 *The operator given by Eq. (7.2.3) is called the formal adjoint of L and is denoted by L^*. So the formal adjoint of the operator*

$$L = a_0 \frac{d^2}{dt^2} + a_1 \frac{d}{dt} + a_2$$

is

$$L^* = a_0 \frac{d^2}{dt^2} + (2\dot{a}_0 - a_1)\frac{d}{dt} + (\ddot{a}_0 - \dot{a}_1 + a_2)$$

Definition 7.2.4 *The expression $J(x, y) = a_0(y\dot{x} - x\dot{y}) + (a_1 - \dot{a}_0)xy$ is called the bilinear concomitant of x and y. From Eq. (7.2.2) we get the Green's formula*

$$\int_a^b (yLx - xL^* y)dt = J(x, y)\Big|_a^b$$

Definition 7.2.5 *L is said to be formally self-adjoint if $L^* = L$.*

It is clear that L is formally self-adjoint if $\dot{a}_0(t) = a_1(t)$. So, a formally self-adjoint operator L can be written in the form

$$L = \frac{d}{dt}\left(a_0(t)\frac{d}{dt}\right) + a_2(t)$$

$J(x, y)$ is then given by

$$a_0\,(y\dot{x} - \dot{y}x)$$

and the Green's formula reduces to

$$\int_a^b (yLx - xLy)\,dt = J(x, y) = a_0\,(y\dot{x} - \dot{y}x)\Big|_a^b$$

We say that $y \in D^*$ (domain of L^*) if it is twice differentiable and if $J(x,y)|_a^b = 0$ for every $x \in D$. From the definition of $J(x,y)$, it can be seen that $J(x,y)|_a^b = 0$ for all $x \in D$ only if y satisfies two homogeneous boundary conditions $B_1^*(y) = B_2^*(y) = 0$. These are called adjoint boundary conditions.

Definition 7.2.6 *Consider the boundary value problem*

$$\begin{aligned} Lx = f; \;\; a \le t \le b \\ B_1(x) = 0 = B_2(x) \end{aligned} \tag{7.2.4}$$

Its adjoint boundary value problem is

$$\begin{aligned} L^*x &= f \\ B_1^*(x) &= 0 = B_2^*(x) \end{aligned} \tag{7.2.5}$$

Boundary value problem given by Eq. (7.2.4) is said to be self-adjoint, if

$$L = L^* \;\; and \;\; D = D^*$$

Remark 7.2.1 *If we consider L^* as an operator on the Hilbert space $L_2[a,b]$ with domain D^*, one can show that L is self-adjoint as an unbounded operator if the boundary value problem given by Eq. (7.2.4) is self-adjoint in the sense of the above definition (Refer Dunford and Schwartz [3]).*

Example 7.2.1 $L = \dfrac{d^2}{dt^2}$ *with boundary conditions*

$$\begin{aligned} B_1(x) \equiv x(0) = 0 \\ B_2(x) \equiv \dot{x}(0) = 0 \end{aligned}$$

L is formally self-adjoint and so $L^* = \dfrac{d^2}{dt^2}$.

To determine adjoint boundary conditions, we need to find $J(x,y)|_0^1$.

$$\begin{aligned} J(x,y) &= (y\dot{x} - \dot{y}x)\,|_0^1 \\ &= \dot{x}(1)y(1) - \dot{y}(1)x(1) - \dot{x}(0)y(0) + \dot{y}(0)x(0) \\ &= \dot{x}(1)y(1) - x(1)\dot{y}(1) \end{aligned}$$

So $J(x,y)|_0^1 = 0$ for all $x \in D$ iff $B_1^*(y) \equiv y(1) = 0, \;\; B_2^*(y) \equiv \dot{y}(1) = 0$.

Thus we see that though the operator is formally self-adjoint. $D^* \ne D$. Hence the boundary value problem is not self-adjoint.

Example 7.2.2 *Consider the second order differential operator*

$$L = a_0\frac{d^2}{dt^2} + a_1(t)\frac{d}{dt} + a_2(t)$$

with boundary conditions $B_1(x) \equiv x(0) = 0, \;\; B_2(x) \equiv \dot{x}(1) = 0$. We know that

$$L^* = a_0\frac{d^2}{dt^2} + (2\dot{a}_0 - a_1)\frac{d}{dt} + (\ddot{a}_0 - \dot{a}_1 + a_2)$$

L will be formally self-adjoint iff $\dot{a}_0 = a_1$.

To determine the adjoint boundary conditions we first determine $J(x,y)|_0^1$.

$$\begin{aligned}
J(x,y)|_0^1 &= a_0(y\dot{x} - x\dot{y}) + (a_1 - \dot{a}_0)xy|_0^1 \\
&= a_0(1)\left[y(1)\dot{x}(1) - x(1)\dot{y}(1)\right] + \left[a_1(1) - \dot{a}_0(1)\right]x(1)y(1) \\
&\quad -a_0(0)\left[y(0)\dot{x}(0) - x(0)\dot{y}(0)\right] + [a_1(0) - \dot{a}_0(0)]x(0)y(0) \\
&= [a_1(1) - \dot{a}_0(1)]x(1)y(1) - a_0(1)\dot{y}(1)x(1) - a_0(0)y(0)\dot{x}(0) \\
J(x,y)|_0^1 &= 0 \quad \textit{for all} \;\; x \in D \;\; \textit{iff} \\
B_1^* &\equiv y(0) = 0 \\
B_2^* &\equiv \dot{y}(1) - \frac{[a_1(1) - \dot{a}_0(1)]y(1)}{a_0(1)} = 0
\end{aligned}$$

Thus, we see that the adjoint boundary conditions differ from the original one and $D^* \neq D$. However, if L is formally delf-adjoint, then $D^* = D$. That is, the boundary value problem is self-adjoint.

7.3 Green's Function

Let the boundary value problem be given by

$$\begin{aligned}
Lx &= f \\
B_1(x) &= 0 = B_2(x)
\end{aligned} \tag{7.3.1}$$

We want to find L^{-1}, the inverse of L, and use it to solve the above equation. As we will see later, this inverse operator is an integral operator. The kernel of this operator will turn out to be the Green's function of the system Eq. (7.3.1), which we now define.

First we define fundamental solution for the operator L.

Definition 7.3.1 *Any solution of the equation*

$$LT = \delta(t-s) \tag{7.3.2}$$

is said to be a fundamental solution for the operator L.

Since RHS of Eq. (7.3.2) is a distribution, the solution T of Eq. (7.3.2) must be interpreted in the sense of a distribution. But we will show that the solution of Eq. (7.3.2) is a weak solution.

We prove this assertion by writing Eq. (7.3.2) formally as

$$LT = 0, \quad t \neq s$$

Let $x(t)$ be an arbitrary solution of the homogeneous equation and let $y_s(t)$ be any other solution which is to be determined.

We get a particular solution of Eq. (7.3.2) in the following way:

$$T = \begin{cases} y_s(t), & t > s \\ \\ x(t), & t < s \end{cases}$$

We will determine $y_s(t)$ by appropriate matching conditions at $t = s$. Assume that T is continuous at $t = s$, then integrating Eq.(7.3.2) we get

$$\lim_{\epsilon \to 0} \int_{s-\epsilon}^{s+\epsilon} a_0(t) \frac{d^2T}{dt^2} dt = 1,$$

where a_0 is the coefficient of $\dfrac{d^2T}{dt^2}$ occuring on the expression for LT. That is,

$$a_0(s) \left[\left.\frac{dT}{dt}\right|_{t=s+} - \left.\frac{dT}{dt}\right|_{t=s-} \right] = 1$$

or $\dot{y}_s(s) = \dot{x}(s) + \dfrac{1}{a_0(s)}$

Thus a solution T of (7.3.2) is given by

$$T(t, s) = \begin{cases} y_s(t), & t > s \\ \\ x(t), & t < s \end{cases} \tag{7.3.3}$$

where $x(t)$ is an arbitrary solution of the homogeneous equation and $y_s(t)$ is another solution of the homogeneous equation satisfying

$$\begin{aligned} y_s(s) &= x(s) \\ \dot{y}_s(s) &= \dot{x}(s) + \frac{1}{a_0(s)} \end{aligned}$$

By the theory of differential equaitons, such a solution $y_s(t)$ exists and is determined by $x(t)$ and $a_0(t)$.

We now show that $T(t, s)$ defined by Eq. (7.3.3) is a weak solution of Eq. (7.3.2).

That is $(T, L^*\phi) = (\delta, \phi) = \phi(s)$ for every test function ϕ.

By definition

$$L^*\phi = \sum_{k=0}^{2} (-1)^k \frac{d^k}{dt^k}(a_{n-k}\phi)$$

$$(T, L^*\phi) = \int_{-\infty}^{\infty} T(t) L^*\phi(t) dt$$

Integrating by parts twice and using boundary conditions, we get

$$(T, L^*\phi) = \phi(s) + \int_{-\infty}^{s-} \phi Lx dt + \int_{s+}^{\infty} \phi L y_s dt$$

Since both x and y_s satisfy the homogeneous equation $LT = 0$, we get

$$(T, L^*\phi) = \phi(s) = (\delta(t-s),\ \phi)$$

That is

$$LT = \delta(t-s)$$

We are now in a position to define the Green's function of Eq. (7.3.1).

Definition 7.3.2 *The Green's function $g(t,s)$ of Eq.(7.3.1) is a fundamental solution of*

$$Lg = \delta(t-s)$$

which satisfies additional boundary condition $B_1(g) = 0 = B_2(g)$. In view of what we have just proved, alternatively we can define the Green's function $g(t,s)$ as a solution of the boundary value problem

$$\begin{aligned} Lg &= 0, \quad a \le t < s, \quad s < t \le b \\ B_1(g) &= 0 = B_2(g) \end{aligned} \tag{7.3.4}$$

g is continuous at $t = s$ and $\left.\frac{dg}{dt}\right|_{t=s+} - \left.\frac{dg}{dt}\right|_{t=s-} = \frac{1}{a_0(s)}$.

Example 7.3.1 *Consider the string problem*

$$\begin{aligned} \frac{d^2x}{dt^2} &= f(t) \\ x(0) &= x(1) = 0 \end{aligned}$$

Green's function $g(t,s)$ for the above sustem is the solution of the auxiliary problem

$$\begin{aligned} \frac{d^2g}{dt^2} &= \delta(t-s) \\ g(0,s) &= 0 = g(1,s) \end{aligned}$$

That is

$$\begin{aligned} \frac{d^2g}{dt^2} &= 0, \quad 0 \le t < s, \ \ s < t \le 1 \\ g(0,s) &= 0 = g(1,s) \end{aligned}$$

$g(t,s)$ is continuous at $t = s$, $\left.\dfrac{dg}{dt}\right|_{t=s+} - \left.\dfrac{dg}{dt}\right|_{t=s-} = 1.$

Integrating the above equation and imposing boundary conditions at 0 *and* 1*, we get*

$$g(t,s) = \begin{cases} At, & 0 \le t < s \\ C(t-1), & s < t \le 1 \end{cases}$$

The continuity of $g(t,s)$ at $t = s$ gives

$$g(t,s) = \begin{cases} At, & 0 \le t \le s \\ \dfrac{A(1-t)s}{(1-s)}, & s \le t \le 1 \end{cases}$$

Finally, the jump condition on the derivatives $\dfrac{dg}{dt}$ at $t = s$ gives

$$\frac{As}{s-1} - A = 1, \quad A = s-1$$

So

$$g(t,s) = \begin{cases} t(s-1), & 0 \le t \le s \\ s(t-1), & s \le t \le 1 \end{cases}$$

Example 7.3.2 *Consider the initial-value problem representing motion of a particle in a straight line*

$$\begin{aligned} m\frac{d^2x}{dt^2} &= f(t), \quad 0 \le t < \infty \\ x(0) &= \dot{x}(0) = 0 \end{aligned}$$

Green's function for the above system is given by

$$\begin{aligned} m\frac{d^2g}{dt^2} &= 0, \quad 0 \le t < s, \quad s < t < \infty \\ g(0,s) &= 0 = \dot{g}(0,s) \end{aligned}$$

g is continuous at $t = s$ and $\left.\dfrac{dg}{dt}\right|_{t=s+} - \left.\dfrac{dg}{dt}\right|_{=s-} = \dfrac{1}{m}.$

Integrating and using the boundary conditions we get

$$g(t,s) = \begin{cases} 0, & t \le s \\ At + B, & s \le t \end{cases}$$

Using continuity of g at $t = s$ we get $B = -As$.
So

$$g(t,s) = \begin{cases} 0, & t \le s \\ A(t-s), & s \le t \end{cases}$$

Finally the jump condition for the derivative gives $A = \dfrac{1}{m}$.
Hence

$$g(t,s) = \begin{cases} 0, & t \le s \\ \dfrac{(t-s)}{m}, & s \le t \end{cases}$$

Definition 7.3.3 *The adjoint Green's function $h(t,s)$ is defined as the solution of the system*

$$\begin{aligned} L^*h &= \delta(t-s) \\ B_1^*(h) &= 0 = B_2^*(h) \end{aligned} \tag{7.3.5}$$

Theorem 7.3.1 *Theadjoint Green's function $h(t,s)$ and the Green's function $g(t,s)$ satisfy the relation $h(t,s) = g(s,t)$. In particular, if the original boundary value problem given by Eq. (7.3.1) is self-adjoint, then the Green's function $g(t,s)$ is symmetric.*

Proof : We have

$$\begin{aligned} Lg &= \delta(t-s) \\ B_1(g) &= 0 = B_2(g) \end{aligned} \tag{7.3.6}$$

and

$$\begin{aligned} L^*h &= \delta(t-\eta) \\ B_1^*(h) &= 0 = B_2^*(h) \end{aligned} \tag{7.3.7}$$

Multiply Eq. (7.3.6) by $h(t,\eta)$, Eq. (7.3.7) by $g(t,s)$, subtract and integrate from $t = a$ to $t = b$. Then we get

$$\int_a^b (hLg - gL^*h)dt = \int_a^b h(t,\eta)\delta(t-s)dt - \int_a^b g(t,s)\delta(t-\eta)dt$$

Using Green's formula $\int_a^b (hLg - gL^*h) = J(g,h)\Big|_a^b$ and substituting this in the left hand side of the above equation we get

$$J(g,h)|_a^b = h(s,\eta) - g(\eta,s)$$

Since h satisfies the adjoint boundary condition $J(g,h)|_a^b = 0$ and hence $h(s,\eta) = g(\eta,s)$ for all s,η. If the system is self-adjoint, then we know that $h = g$ and hence

$$g(s,\eta) = g(\eta,s) \quad \text{for all} \quad s,\eta$$

That is, g is symmetric. ∎

Now we shall show how Green's function can be used to obtain the solution of the boundary value problem by constructing the inverse of the corresponding differential operator.

Theorem 7.3.2 *Suppose $g(t,s)$ is a Green's function for the boundary value problem $Lx = f,\ \ a \le t \le b,\ B_1(x) = 0 = B_2(x)$. Then it has a solution given by*

$$x(t) = \int_a^b g(t,s)f(s)ds \tag{7.3.8}$$

We shall use the following lemma which is stated without proof.

Lemma 7.3.1 (Leibnitz) *Let $f(t,\eta)$ be a continuous function of two variables in a region $\alpha(t) \le \eta \le \beta(t),\ \ c \le t \le d$, where $\alpha(t)$ and $\beta(t)$ are continuous functions. Then the integral*

$$g(t) = \int_{\alpha(t)}^{\beta(t)} f(t,\eta)d\eta; \quad c \le t \le d \tag{7.3.9}$$

represents a continuous function of t. If the derivatives $\dot{\alpha}(t)$ and $\dot{\beta}(t)$ exist, and $\dfrac{\partial f}{\partial t}$ as well as f are continuous, then the derivative of the integral in Eq. (7.3.9) is given by the generalized Leibnitz formula

$$\dot{g}(t) = \int_{\alpha(t)}^{\beta(t)} \frac{\partial}{\partial t} f(t,\eta)d\eta + f(t,\beta(t))\dot{\beta}(t) - f(t,\alpha(t))\dot{\alpha}(t)$$

Proof (Proof of the main theorem): Let $g(t,s)$ be the Green's function for the boundary value problem given by Eq. (7.3.1). We can write Eq. (7.3.8) as

$$x(t) = \int_a^b g(t,s)f(s)ds = \int_a^{t-} g(t,s)f(s)ds + \int_{t+}^b g(t,s)f(s)ds$$

Then, using the above mentioned lemma, we have

$$\begin{aligned}\dot{x}(t) &= \int_a^{t-} g_t(t,s)f(s)ds + \int_{t+}^b g_t(t,s)f(s)ds \\ &+ [g(t,t-)f(t) - g(t,t+)f(t)]\end{aligned}$$

Due to the continuity of $g(t,s)$, the last bracketed expression vanishes. Also

$$\begin{aligned}\ddot{x}(t) &= \int_a^{t-} g_{tt}(t,s)f(s)ds + \int_{t+}^b g_{tt}(t,s)f(s)ds \\ &+ g_t(t,t-)f(t) - g_t(t,t+)f(t)\end{aligned}$$

We note that $g_t(t,t+) = g_t(t-,t)$ and $g_t(t,t-) = g_t(t+,t)$ due to the continuity of $g(t,s)$ in the triangular regions ABC and ABD respectively (Refer Fig. 7.2.1).

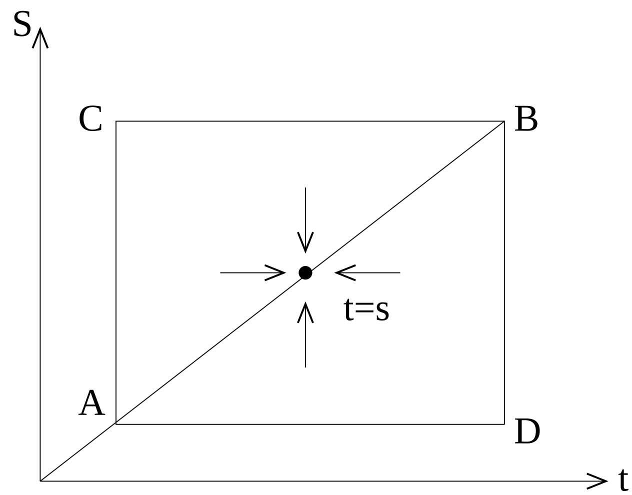

Figure 7.3.1: Integration in a triangular region

So, substituting the expressioins for $\ddot{x}(t)$ and $\dot{x}(t)$ in $Lx \equiv a_0(t)\ddot{x} + a_1(t)\dot{x} + a_2(t)x$, we have

$$\begin{aligned}
Lx &= \int_a^{t-} \left[a_0(t)g_{tt}(t,s) + a_1(t)g_t(t,s) + a_2(t)g(t,s)\right] ds \\
&+ \int_{t+}^{b} \left[a_0(t)g_{tt}(t,s) + a_1(t)g_t(t,s) + a_2(t)g(t,s)\right] ds \\
&+ f(t)\left[g_t(t,t-) - g_t(t,t+)\right] a_0(t) \\
&= 0 + 0 + f(t)\left[\frac{1}{a_0(t)} a_0(t)\right] \\
&= f(t)
\end{aligned}$$

Hence $x(t) = \int_a^b g(t,s)f(s)ds$ satisfies the given differential equation. In exactly the same manner one can show that x given by Eq. (7.3.8) satisfies $B_1(x) = 0 = B_2(x)$. Thus, $x(t)$ given by Eq. (7.3.8) is the solution of the boundary value problem given by Eq. (7.3.1). ■

Theorem 7.3.3 *The operator L and the integral operator G defined as*

$$Gf = \int_a^b g(t,s)f(s)ds$$

are inverses of each other.

Proof : It is clear from the definition of L and G that $LGf = f$ and $GLx = x$. Hence the result. ∎

Green's function also enables us to get a solution of the boundary value problem with non-homogeneous boundary conditions:

$$\begin{aligned} Lx &= f \\ B_1(x) &= \alpha, \quad B_2(x) = \beta \end{aligned} \tag{7.3.10}$$

Theorem 7.3.4 *If $g(t,s)$ is a Green's function of Eq. (7.3.10), then it has a solution $x(t)$ given by*

$$x(t) = \int_a^b g(t,s)f(s)ds - J(x(s), g(t,s))\Big|_a^b$$

Proof : We recall that the adjoint Green's function $h(t,s)$ satisfies the system

$$\begin{aligned} L^*h &= \delta(t-s) \\ B_1^*(h) &= 0 = B_2^*(h) \end{aligned} \tag{7.3.11}$$

Multiply Eq. (7.3.10) by $h(t,s)$, Eq. (7.3.11) by $x(t)$, subtract and integrate from $t = a$ to $t = b$. to get

$$\int_a^b (hLx - xL^*h)dt = \int_a^b f(t)h(t,s)dt - \int_a^b \delta(t-s)x(t)dt$$

Using Green's formula in the LHS, we have

$$J(x,h)|_a^b = \int_a^b h(t,s)f(t)dt - x(s)$$

That is

$$x(s) = \int_a^b h(t,s)f(t)dt - J(x,h)\Big|_a^b$$

Using the fact that $h(t,s) = g(s,t)$ we obtain

$$x(s) = \int_a^b g(s,t)f(t)dt - J(x(t), g(s,t))\Big|_a^b$$

or

$$x(t) = \int_a^b g(t,s)f(s)ds - J(x(s), g(t,s))\Big|_a^b$$

$J(x(s), g(t,s))|_a^b$ can be calculated explicitly by the given boundary conditions. ∎

An alternative method of solution of Eq. (7.3.10) is as follows.

Let $x_1(t)$ and $x_2(t)$ be non-trivial solutions of the homogeneous system, corresponding to Eq. (7.3.10) satisfying $B_1(x_1) = 0$ and $B_2(x_2) = 0$, respectively. Since a completely homogeneous system has only the trivial solution, we must have $B_1(x_2) \neq 0$. We write the solution of the system as

$$x(t) = \int_a^b g(t,s)f(s)ds + c_1 x_1(t) + c_2 x_2(t)$$

$g(t,s)$ is the Green's function for the system with homogeneous boundary conditions and c_1 and c_2 are constants to be determined. Applying boundary conditions, we have

$$\begin{aligned} c_2 B_1(x_2) + c_1 B_1(x_1) + \int_a^b B_1(g(t,s))f(s)ds &= \alpha \\ c_1 B_2(x_1) + c_2 B_2(x_2) + \int_a^b B_2(g(t,s))f(s)ds &= \beta \end{aligned}$$

Therefore,

$$x(t) = \int_a^b g(t,s)f(s)ds + \frac{\beta}{B_2(x_1)}x_1(t) + \frac{\alpha}{B_1(x_2)}x_2(t)$$

It is clear that this solution depends continuously on f, α and β and is unique.

Example 7.3.3 *Let us compute the solution of the non-homogeneous boundary value problem*

$$\begin{aligned} \frac{d^2x}{dt^2} &= f(t) \\ x(0) &= \alpha, \quad x(1) = \beta \end{aligned}$$

As in Example 7.3.1, its Green's function $g(t,s)$ is given by

$$g(t,s) = \begin{cases} t(s-1), & t \leq s \\ s(t-1), & t \geq s \end{cases}$$

It is a symmetric function and satisfies the boundary conditions $g(t,s)|_{t=0} = 0 = g(t,s)|_{t=1}$.

From the above

$$\dot{g}(t,s) = \begin{cases} s-1, & t \leq s \\ s, & t \geq s \end{cases}$$

$$\dot{g}(t,s)|_{t=0} = (s-1), \quad \dot{g}(t,s)|_{t=1} = s$$

Using the formula for $J(x, g)$ we get

$$\begin{aligned} J(x,g)|_0^1 &= [\dot{x}g - x\dot{g}]_0^1 \\ &= [\dot{x}(1)g(1) - x(1)\dot{g}(1)] - [\dot{x}(0)g(0) - x(0)\dot{g}(0)] \\ &= x(0)\dot{g}(0) - x(1)\dot{g}(1) \\ &= \alpha\dot{g}(0) - \beta\dot{g}(1) \\ &= \alpha(s-1) - \beta s \end{aligned}$$

Hence, by Theorem 7.3.4, the solution $x(s)$ of the above boundary value problem is given by

$$\begin{aligned} x(s) &= \int_0^1 g(t,s)f(t)dt - J(x,g)|_0^1 \\ &= \int_0^1 g(t,s)f(t)dt - \alpha(s-1) + \beta s \\ &= \int_0^s t(s-1)f(t)dt + \int_s^1 s(t-1)f(t)dt + \alpha(1-s) + \beta s \end{aligned}$$

or

$$x(t) = \int_0^t s(t-1)f(s)ds + \int_t^1 t(s-1)f(s)ds + \alpha(1-t) + \beta t$$

7.4 Construction of Green's Function and Modified Green's Function

We shall consider the boundary value problem

$$\begin{aligned} Lx &= f \\ B_1(x) &= \alpha, \ \ B_2(x) = \beta \end{aligned}$$

We prove the existence of Green's function of the above system where the boundary conditions are of unmixed and initial type only, by construction.

Theorem 7.4.1 *For the restricted case of unmixed and initial boundary conditions of the completely homogeneous system*

$$Lx = 0, \quad B_1(x) = 0 = B_2(x) \tag{7.4.1}$$

which has only the trivial solution, the Green's function for the boundary value problem exists and is unique.

Proof : Uniqueness of the solution is immediate. For, if there are two Green's functions g_1 and g_2 then $(g_1 - g_2)$ is a non-trivial solution of the completely homogeneous equation Eq. (7.4.1) which is a contradiction to the hypothesis. We need to prove the existence in two parts, first for unmixed boundary conditions and then for initial conditions.

Case 1 Unmixed boundary conditions are given by

$$\begin{aligned} B_1(x) &= \alpha_{11}x(a) + \alpha_{21}\dot{x}(a) = 0 \\ B_2(x) &= \beta_{21}x(b) + \beta_{22}\dot{x}(b) = 0 \end{aligned}$$

Let $x_1(t)$ be a non-trivial solution of $Lx = 0$ satisfying boundary condition at $t = a$ and $x_2(t)$ be a non-trivial solution satisfying boundary condition at $t = b$. Such a solution exists, for one can take $x_1(t)$ to be the unique solution of $Lx = 0,\ x(a) = \alpha_{12},\ \dot{x}(a) = -\alpha_{11}$. This follows from the theory of differential equation (Refer Coddington and Levinson[1].) Similarly, we have the existence of $x_2(t)$. x_1 and x_2 are linearly independent, otherwise the completely homogeneous system will have a non-trivial solution. The Green's function is then of the form

$$\begin{cases} Ax_1(t), & a \le t < s \\ \\ Bx_2(t), & s < t \le b \end{cases}$$

Continuity and jump conditions of $g(t,s)$ give

$$\begin{aligned} Ax_1(s) - Bx_2(s) &= 0 \\ -A\dot{x}_1(s) - B\dot{x}_2(s) &= \frac{1}{a_0(s)} \end{aligned}$$

Solving the above of equations for A and B we get

$$\begin{aligned} A &= \frac{x_2(s)}{a_0(s)W(x_1, x_2, s)} \\ B &= \frac{x_1(s)}{a_0(s)W(x_1, x_2, s)} \end{aligned}$$

where $W(x_1, x_2, s)$ is the Wronskian of x_1 and x_2 and is non-zero since x_1 and x_2 are linearly independent.

So

$$g(t,s) = \begin{cases} \dfrac{x_1(t)x_2(s)}{a_0(s)W(x_1, x_2, s)}, & a \le t < s \\ \\ \dfrac{x_1(s)x_2(t)}{a_0(s)W(x_1, x_2, s)}, & s < t \le b \end{cases}$$

This proves the existence of Green's function for the boundary value problem.

Case 2 Boundary conditions are of initial type

$$x(a) = 0 = \dot{x}(a)$$

Green's function is of the form

$$g(t,s) = \begin{cases} 0, & a \le t < s \\ \\ Ax_1(t) + Bx_2(t), & t > s \end{cases}$$

where x_1, x_2 are two linearly independent solutions of $Lx = 0$.
Continuity and jump conditions on $g(t,s)$ give

$$\begin{aligned} Ax_1(s) + Bx_2(s) &= 0 \\ A\dot{x}_1(s) + B\dot{x}_2(s) &= \frac{1}{a_0(s)} \end{aligned}$$

Solving for A and B, we get

$$\begin{aligned} A &= \frac{x_2(s)}{a_0(s)W(x_1, x_2, s)} \\ B &= \frac{x_1(s)}{a_0(s)W(x_1, x_2, s)} \end{aligned}$$

This gives

$$g(t,s) = \begin{cases} 0, & a \le t \le s \\ \\ \dfrac{x_2(t)x_1(s) - x_1(t)x_2 s}{a_0(s)W(x_1, x_2, s)}, & t \ge s \end{cases}$$

∎

Example 7.4.1 *Let us compute the Green's function for the string problem*

$$\begin{aligned} \frac{d^2x}{dt^2} &= f(t) \\ x(0) &= 0 = x(1) \end{aligned}$$

We take $x_1(t) = t \;\; x_2(t) = 1 - t$. *Then*

$$\begin{aligned} W(x_1, x_2, s) &= \begin{vmatrix} s & 1-s \\ \\ 1 & -1 \end{vmatrix} = -1 \\ &= \begin{vmatrix} x_1(s) & x_2(s) \\ \\ \dot{x}_1(s) & \dot{x}_2(s) \end{vmatrix} \\ A &= \frac{x_2(s)}{a_0(s)W(x_1, x_2, s)} = \frac{1-s}{-1} \\ B &= \frac{x_1(s)}{a_0(s)W(x_1, x_2, s)} = \frac{s}{-1} \end{aligned}$$

So

$$g(t,s) = \begin{cases} t(s-1), & t \le s \\ s(t-1), & t \ge s \end{cases}$$

Example 7.4.2 *Consider the initial-value problem describing the motion of a particle in a straight line*

$$\begin{aligned} m\frac{d^2x}{dt^2} &= f(t) \\ x(0) &= 0 = \dot{x}(0) \end{aligned}$$

We take $x_1(t) = 1, \quad x_2(t) = t$ *as two linearly independent solutions of the homogeneous differential equation* $\frac{d^2x}{dt^2} = 0$.
Then

$$W(x_1, x_2, s) = \begin{vmatrix} 1 & s \\ 0 & 1 \end{vmatrix} = 1$$

Hence

$$g(t,s) = \begin{cases} 0, & t \le s \\ \dfrac{(t-s)}{m}, & t \ge s \end{cases}$$

Example 7.4.3 *Consider the problem of heat distribution in a rod of unit length and unit cross-section and with thermal conductivity* $k(t)$. *If the heat source has density* $f(t)$, *then the governing equation is*

$$-\frac{d}{dt}\left(k\frac{dx}{dt}\right) = f(t)$$

Assume that the ends of the rod are kept at $0°$ *temperature*

$$x(0) = 0 = x(1)$$

We look for two linearly independent solutions $x_1(t)$ *and* $x_2(t)$ *of the homogeneous equation*

$$-\frac{d}{dt}\left(k\frac{dx}{dt}\right) = 0$$

which satisfy the boundary conditions at 0 and 1 respectively. So we write

$$\begin{aligned} x_1(t) &= \int_0^t \frac{ds}{k(s)}, \quad x_2(t) = \int_1^t \frac{dx}{k(s)} \\ W(x_1, x_2, s) &= \begin{vmatrix} x_1(s) & x_2(s) \\ \dfrac{1}{k(s)} & -\dfrac{1}{k(s)} \end{vmatrix} \end{aligned}$$

$$= -\frac{x_1(s) + x_2(s)}{k(s)} = -\frac{\displaystyle\int_0^1 k^{-1}(s)ds}{k(s)}$$

Hence

$$\begin{aligned} g(t,s) &= \begin{cases} \dfrac{x_1(t)x_2(s)}{k(s)} \left[\dfrac{k(s)}{\int_0^1 k^{-1}(t)dt}\right], & t \le s \\[2ex] \dfrac{x_1(s)x_2(t)}{k(s)} \left[\dfrac{k(s)}{\int_0^1 k^{-1}(t)dt}\right], & t \ge s \end{cases} \\ &= \left[\int_0^1 k^{-1}(t)dt\right]^{-1} \begin{cases} x_1(t)x_2(s), & t \le s \\[1ex] x_2(t)x_1(s), & t \ge s \end{cases} \end{aligned}$$

Example 7.4.4 *In the study of heat flow in a cylindrical shell, one comes across the Bessel's equation of the type*

$$\frac{d}{dt}\left(t\frac{dg}{dt}\right) + tg = \delta(t-s); \quad a \le t \le b$$

with boundary conditions

$$g = 0 \ \ at \ \ t = a \ \ and \ \ t = b$$

As we know, the homogeneous Bessel's equation

$$\frac{d}{dt}\left(t\frac{dg}{dt}\right) + tg = 0$$

has two linearly independent solutions $J_0(t)$ and $Y_0(t)$

$$J_0(t) = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{2^{2n}(n!)^2}$$

and

$$\begin{aligned} Y_0(t) &= \frac{2}{\pi}\left[\gamma + \frac{lnt}{2}\right] J_0(t) + \frac{2}{\pi}\frac{(t/2)^2}{(1!)^2} - \frac{2}{\pi}\frac{(t/2)^4}{(2!)^2}(1 + 1/2) \\ &\quad + \frac{2}{\pi}\frac{(t/2)^6}{(3!)^2}(1 + 1/2 + 1/3) - \cdots \end{aligned}$$

where $\gamma \cong 0.5772$ is the so called Euler constant which is the limit of $1 + \frac{1}{2} + \cdots + \frac{1}{5} - lns$ as s approaches infinity. But $J_0(t)$ and $Y_0(t)$ fail to satisfy the boundary conditions at a and b. So we define two other linearly independent solutions

$$\begin{aligned} Z_0^1(t) &= J_0(t)Y_0(a) - Y_0(t)J_0(a) \\ Z_0^2(t) &= J_0(t)Y_0(b) - Y_0(t)J_0(b) \end{aligned}$$

Obviously now $\quad Z_0^1(a) = 0 = Z_0^2(b).$

It can be shown that

$$\begin{aligned} W(Z_0^1, Z_0^2; t) &= W[J_0, Y_0; t][J_0(a)Y_0(b) - Y_0(a)J_0(b)] \\ &= \frac{2}{\pi t}[J_0(a)Y_0(b) - Y_0(a)J_0(b)] \end{aligned}$$

So

$$g(t,s) = \begin{cases} \dfrac{Z_0^1(t)Z_0^2(s)}{C}, & t \le s \\ \\ \dfrac{Z_0^1(s)Z_0^2(t)}{C}, & t \ge s \end{cases}$$

where $$C = \frac{2}{\pi}[J_0(a)Y_0(b) - Y_0(a)J_0(b).$$

Example 7.4.5 *Consider the boundary value problem*

$$\begin{aligned} -\frac{d^2x}{dt^2} - \lambda x &= f(t) \\ x(0) &= 0 = x(1) \end{aligned}$$

To find the Green's function for this problem we first look for two linearly independent solutions of the completely homogeneous problem $-\dfrac{d^2x}{dt^2} - \lambda x = 0$, *which satisfy the boundary condition at 0 and 1, respectively.*

$$x_1(t) = \sin\sqrt{\lambda}t, \;\; x_2(t) = \sin\sqrt{\lambda}(1-t)$$

$$\begin{aligned} W(x_1, x_2; t) &= \begin{vmatrix} \sin\sqrt{\lambda}t & \sin\sqrt{\lambda}(1-t) \\ \\ \sqrt{\lambda}\cos\sqrt{\lambda}t & -\sqrt{\lambda}\cos\sqrt{\lambda}(1-t) \end{vmatrix} \\ &= \sqrt{\lambda}\sin\sqrt{\lambda}(1-t)\cos\sqrt{\lambda}t \\ &\quad + \sin\sqrt{\lambda}t\cos\sqrt{\lambda}(1-t) \\ \\ &= \sqrt{\lambda}\sin\sqrt{\lambda}(1-t+t) \\ \\ &= \sqrt{\lambda}\sin\sqrt{\lambda} \end{aligned}$$

So

$$g(t,s) = \begin{cases} \dfrac{\sin\sqrt{\lambda}t\sin\sqrt{\lambda}(1-s)}{\sqrt{\lambda}\sin\sqrt{\lambda}}, & t \le s \\ \\ \dfrac{\sin\sqrt{\lambda}(1-t)\sin\sqrt{\lambda}s}{\sqrt{\lambda}\sin\sqrt{\lambda}}, & t \ge s \end{cases}$$

In general, where the boundary conditions are mixed, the Green's function may be found in a straignt forward way.

Let $x_1(t)$ *and* $x_2(t)$ *be any two linearly independent solutions of* $Lx = 0$. *Then write*

$$g(t,s) = Ax_1(t) + Bx_2(t) + \frac{x_1(t)x_2(s)H(s-t) + x_2(t)x_1(s)H(t-s)}{W(x_1, x_2; s)a_0(s)}$$

where A *and* B *are constants which will be determined so that* $g(t,s)$ *satisfies the given boundary conditions.*

It is clear that $g(t,s)$ *is continuous and its derivative has a jump of magnitude* $\frac{1}{a_0(s)}$. *Now consider the boundary conditions. Since they are linear and homogeneous, we have*

$$B_1(g) \equiv AB_1(x_1) + BB_1(x_2) + B_1(r) = 0 \qquad (7.4.2)$$
$$B_1(g) \equiv AB_1(x_1) + BB_1(x_2) + B_1(r) = 0 \qquad (7.4.3)$$

where $r(x)$ *stands for* $\frac{x_1(t)x_2(s)H(s-t) + x_2(t)x_1(s)H(t-s)}{W(x_1, x_2; s)a_0(s)}$

The boundary conditions give us two linear equations which will determine the values of A *and* B. *These equations will have solutions if the determinant*

$$\begin{bmatrix} B_1(x_1) & B_1(x_2) \\ \\ B_2(x_1) & B_2(x_2) \end{bmatrix}$$

does not vanish.

If the determinant vanishes, either $B_1(x_1) = B_2(x_1) = 0$, *or* $B_1(x_2) = B_2(x_2) = 0$ *which implies that either* $x_1(t)$ *or* $x_2(t)$ *is a non-trivial solution of the completely homogeneous system which is a contradiction. The other possiblility for the determinant to vanish is that there exists a constant* c *such that*

$$\begin{aligned} B_1(x_2) + cB_1(x_1) &= 0 \\ B_2(x_2) + cB_2(x_1) &= 0 \end{aligned}$$

These equations imply that $B_1(x_2 + cx_1) = B_2(x_2 + cx_1) = 0$, *which is not possible as discussed earlier. So by solving Eq. (7.4.2)- Eq. (7.4.3) for* A *and* B, *we get the desired Green's function.*

So far, we have proved the existence of Green's function for the system

$$Lx = f, \quad B_1(x) = 0 = B_2(x)$$

under the assumption that the completely homogeneous system $Lx = 0$, $B_1(x) = 0 = B_2(x)$ has only the trivial solution. What happens when the completely homogeneous system has a non-trivial solution ?

We state the following Alternative theorem which answers the question raised above

Consider the related systems:

(i) The compeletely homogeneous system

$$Lx = 0; \;\; a \le t \le b; \qquad B_1(x) = 0 = B_2(x)$$

(ii) The inhomogeneous system

$$Ly = f; \;\; a \le t \le b; \qquad B_1(y) = 0 = B_2(y)$$

(iii) The adjoint homogeneous system

$$L^*z = 0; \;\; a \le t \le b; \qquad B_1^*(z) = 0 = B_2^*(z)$$

Theorem 7.4.2 *(a) If (i) has only the trivial solution $x(t) = 0$ in $a \le t \le b$, then (iii) has also the trivial solution and (ii) has one and only one solution.*

(b) If (i) has a non-trivial solution, (iii) will also have non-trivial solution and (ii) has solution iff
$\int_a^b f(t)z(t) = 0$ *for every z which is a solution of (iii).*

For a proof of this, refer Stakgold [5].

Example 7.4.6 *Consider the self-adjoint system*

$$\begin{aligned} \frac{d^2g}{dt^2} &= \delta(t-s) \\ g(0) &= 0, \quad \dot{g}(1) = g(1) \end{aligned} \tag{7.4.4}$$

The completely homogeneous system

$$\frac{d^2g}{dt^2} = 0, \qquad g(0) = 0, \;\; \dot{g}(1) = g(1)$$

has a non-trivial solution $g = t$.

So it follows by alternative theorem that Eq. (7.4.4) will have a solution iff $\int_0^1 \delta(t-s)dt = 0$. Since $\int_0^1 \delta(t-s)dt \neq 0$, Eq. (7.4.4) will have no solution. That is, the ordinary Green's function does not exist.

Example 7.4.7 *Consider the self-adjoint system*

$$\begin{aligned} \frac{d^2g}{dt^2} &= \delta(t-s), \quad 0 \le t \le 1 \\ g(0) &= 0 = \dot{g}(1) \end{aligned} \tag{7.4.5}$$

The completely homogeneous system

$$\frac{d^2g}{dt^2} = 0, \quad \dot{g}(0) = 0 = \dot{g}(1)$$

has a non-trivial solution $g =$ const. Hence Eq. (7.4.5) has no solution, for $\int_0^1 \delta(t-s)ds \neq 0.$
That is, the ordinary Green's function can not be constructed for Eq. (7.4.5).

Modified Green's Function
Assume that the self-adjoint system

$$Lx = f; \;\; a \leq t \leq b; \qquad B_1(x) = 0 = B_2(x)$$

is such that the completely homogeneous system

$$Lx = 0; \;\; a \leq t \leq b; \qquad B_1(x) = 0 = B_2(x)$$

has non-trivial solution of the form $x_1(t)$, where $x_1(t)$ is normalized $\left(\int_a^b x_1^2(t)dt = 1\right)$.

It now follows that the Green's function which is a solution of the system

$$Lg = \delta(t-s); \qquad B_1(g) = 0 = B_2(g)$$

will not exist for $\int_a^b x_1(t)\delta(t-s)dt \neq 0$. *However, in order to compute the Green's function one has to resort to modified Green's function by the addition of new source density of strength $x_1(s)$.*

Definition 7.4.1 *Modified Green's function is the solution of the system*

$$\begin{aligned} Lg_M(t,s) &= \delta(t-s) - x_1(t)x_1(s) \\ B_1(g_M) &= 0 = B_2(g_M) \end{aligned}$$

It is clear that the solution of the above system exists for

$$\int_a^b \left[\delta(t-s) - x_1(t)x_1(s)\right] x_1(t)dt = 0$$

The method of construction of g_M is now similar to that of ordinary Green's function which we have already discussed.

Theorem 7.4.3 *The modified Green's function $g_M(t,s)$ is symmetric if*

$$\int_a^b g_M(t,s)x_1(t)dt = 0 \;\; \textit{for every } s$$

Proof : Consider $g_M(t,s_1)$ and $g_M(t,s_2)$ wich satisfy the systems:

$$\begin{aligned} Lg_M(t,s_1) &= \delta(t-s_1) - x_1(t)x_1(s_1) \\ B_1(g_M(t,s_1)) &= 0 = B_2(g_M(t,s_1)) \end{aligned} \tag{7.4.6}$$

and

$$\begin{aligned} Lg_M(t,s_2) &= \delta(t-s_2) - x_1(t)x_1(s_2) \\ B_1(g_M(t,s_2)) &= 0 = B_2(g_M(t,s_2)) \end{aligned} \tag{7.4.7}$$

Multiply Eq. (7.4.6) by $g_M(t,s_2)$ and Eq. (7.4.7) by $g_M(t,s_1)$ and subtract and integrate from a to b. Then we get

$$\begin{aligned} & \int_a^b \left[g_M(t,s_2)(Lg_M(t,s_1)) - g_M(t,s_1)(Lg_M(t,s_2))\right] dt \\ = & \int_a^b g_M(t,s_2)\delta(t-s_1)dt - \left(\int_a^b x_1(t)g_M(t,s_2)dt\right) x_1(s_1) \\ - & \int_a^b g_M(t,s_1)\delta(t-s_2)dt + \left(\int_a^b x_1(t)g_M(t,s_1)dt\right) x_1(s_2) \end{aligned}$$

Using Green's formula and boundary conditions for $g_M(t,s_1)$, $g_M(t,s_2)$, we have

$$\begin{aligned} 0 &= g_M(s_1,s_2) - g_M(s_2,s_1) - x_1(s_1)\left(\int_a^b x_1(t)g_M(t,s_2)dt\right) \\ & \quad + x_1(s_2)\left(\int_a^b x_1(t)g_M(t,s_1)dt\right) \end{aligned}$$

But $\int_a^b x_1(t)g_M(t,s)dt = 0$ for every s. Hence

$$g_M(s_1,s_2) = g_M(s_2,s_1)$$

∎

On imposing the conditions of symmetry on $g_M(t,s)$ we can alternatively define the modified Green's function as the solution of the system:

$$Lg_M(t,s) = \delta(t-s) - x_1(t)x_1(s); \quad a \le t,\ s \le b$$

satisfying

(i) $B_1(g_M(t,s)) = 0 = B_2(g_M(t,s))$

(ii) $g_M(t,s)$ is continuous at $t = s$

(iii) $\frac{dg_M}{dt}|_{t=s+} - \frac{dg_M}{dt}|_{t=s-} = \frac{1}{a_0(s)}$

and

(iv) $\int_a^b x_1(t)g_M(t,s)dt = 0$ for every s

For computational purposes, the above defining property of $g_M(t;s)$ will be used. We finally have the following theorem which gives a solution of the self-adjoint system by using the modified Green's function.

Theorem 7.4.4 *Let $g_M(t,s)$ be the modified Green's function of the self-adjoint system*

$$Lx = f; \quad a \le t \le b; \qquad B_1(x) = 0 = B_2(x) \tag{7.4.8}$$

having $x_1(t)$ as the only normalized solution of the completely homogeneous system.
If $\int_a^b f(t)x_1(t)dt = 0$, then Eq. (7.4.8) has a solution $x(t)$ given by

$$x(t) = \int_a^b g_M(t,s)f(s)ds + Cx_1(t).$$

Proof : The existence of a solution of Eq. (7.4.8) is clear by alternative theorem, for $\int_a^b f(t)x_1(t)dt = 0$.

We have

$$Lx = f; \qquad B_1(x) = 0 = B_2(x)$$

and

$$\begin{aligned} Lg_M &= \delta(t-s) - x_1(t)x_1(s) \\ B_1(g_M) &= 0 = B_2(g_M) \end{aligned}$$

By our standard procedure we get

$$\begin{aligned} \int_a^b (g_M Lx - xLg_M) &= \int_a^b g_M(t,s)f(t)dt - \int_a^b x(t)\delta(t-s)dt \\ &\quad + x_1(s)\int_a^b x_1(t)x(t)dt \end{aligned}$$

Using Green's formula, it reduces to

$$J(x, g_M) = \int_a^b g_M(t,s)f(t)dt - x(s) + Cx_1(s)$$

The LHS vanishes, fo both x and g_M satisfy zero boundary conditions and there fore we have

$$x(s) = \int_a^b g_M(t,s)f(t)dt + Cx_1(s)$$

or

$$x(t) = \int_a^b g_M(t,s) f(s) ds + C x_1(t)$$

∎

Example 7.4.8 *Consider the system*

$$\begin{aligned} \frac{d^2x}{dt^2} &= \sin 2\pi t; \quad 0 \le t \le 1 \\ \dot{x}(0) &= 0 = \dot{x}(1) \end{aligned} \tag{7.4.9}$$

The completely homogeneous system has a non-trivial solution $x =$ const.

Since $\int_0^1 \sin 2\pi t dt = 0$, *by Theorem 7.4.4, the above system has a solution $x(t)$ given by*

$$x(t) = \int_a^b g_M(t,s) \sin 2\pi s ds + C \tag{7.4.10}$$

To compute $g_M(t,s)$ we solve the system

$$\begin{aligned} \frac{d^2 g_M}{dt^2} &= \delta(t-s) - 1 \\ \dot{g_M}(0) &= 0 = \dot{g_M}(1) \\ g_M(t,s) \ \ &\text{is continuous at} \ \ t = s \\ \frac{dg_M}{dt}|_{t=s+} - \frac{dg_M}{dt}|_{t=s-} &= 1 \\ \int_0^1 g_M(t,s) dt &= 0 \ \ \text{for every} \ \ s \end{aligned} \tag{7.4.11}$$

Integrating Eq.(7.4.11) and using boundary conditions, we get

$$g_M(t,s) = \begin{cases} A - \frac{t^2}{2}, & t < s \\ C + t - \frac{t^2}{2}, & t > s \end{cases}$$

Continuity at $t = s$ implies that

$$A - \frac{s^2}{2} = C + s - \frac{s^2}{2}$$

or

$$C = A - s$$

So

$$g_M(t,s) = \begin{cases} A - \frac{t^2}{2}, & t \le s \\ A - s + t - \frac{t^2}{2}, & t \ge s \end{cases}$$

The last defining condition gives

$$\int_0^s (A - \frac{t^2}{2})dt + \int_s^1 (A - s + t - \frac{t^2}{2})dt = 0$$

That is, $As - \frac{s}{6} + \left[A - s + \frac{1}{2} - \frac{1}{6}\right] - \left[(A-s)s + \frac{s^2}{2} - \frac{s^3}{3}\right] = 0$, *which gives*

$$A = \left[s - \frac{s^2}{2} - \frac{1}{3}\right]$$

Hence, we have the following symmetric modified Green's function

$$\begin{aligned} g_M(t,s) &= \begin{cases} \left[s - \frac{s^2}{2} - \frac{1}{3}\right] - \frac{t^2}{2}, & t \le s \\ \left[s - \frac{s^2}{2} - \frac{1}{3}\right] - s + t - \frac{t^2}{2}, & t \ge s \end{cases} \\ &= \begin{cases} s - \frac{1}{3} - \frac{s^2+t^2}{2}, & t \le s \\ t - \frac{1}{3} - \frac{s^2+t^2}{2}, & t \ge s \end{cases} \end{aligned}$$

7.5 Eigenfunction Expansion - Sturm-Liouville System

Consider the self-adjoint system

$$\begin{aligned} Lx - \lambda x &= f, \quad 0 \le t \le b \\ B_1(x) &= 0 = B_2(x) \end{aligned} \tag{7.5.1}$$

where a and b are finite constants and $f \in L_2[a,b]$. Such a self-adjoint system is called Sturm-Liouville system.

We assume that the operator L acts on $L_2[a,b]$ with domain D define as

$$D = \{x \in L_2[a,b] : \ \dot{x}, \ddot{x} \in L_2[a,b], \ B_1(x) = 0 = B_2(x)\}$$

We consider the related eigenvalue problem

$$L\phi = \lambda\phi, \ \ B_1\phi = 0 = B_2\phi \tag{7.5.2}$$

with the assumption that $\lambda = 0$ is not an eigenvalue of L. We get the following important theorem, giving eigenfunctions of L as a basis for $L_2[a,b]$.

Theorem 7.5.1 *The self-adjoint system Eq.(7.5.2) has discrete eigenvalues* $\{\lambda_1, \lambda_2, \cdots, \lambda_n, \cdots\}$ *and the corresponding eigenfunctions* $\{\phi_1, \phi_2, \cdots, \phi_n, \cdots\}$ *form a compelete orthonormal set in* $L_2[a, b]$.

Proof : As zero is not the eigenvalue of L, there exists Green's function g such that the solvability of Eq. (7.5.2) is equivalent to the solvability of the integral equation

$$x(t) = \lambda \int_a^b g(t, s)x(s)ds \tag{7.5.3}$$

in the space $L_2[a, b]$. Let K be the integral operator generated by $g(t, s)$

$$[Kx](t) = \int_a^b g(t, s)x(s)ds$$

Then Eq. (7.5.3) is equivalent to the following operator equation

$$[Kx] = \frac{1}{\lambda}x = \mu x, \quad \mu = \frac{1}{\lambda} \tag{7.5.4}$$

in the space $L_2[a, b]$.

We note that $\mu = 0$ is not an eigenvalue of K. For, if $x \neq 0$ exists satisfying $\int_a^b g(t, s)x(s)ds = 0$, then the boundary value problem

$$Lw = x, \quad B_1(w) = 0 = B_2(w)$$

has a unique solution $w = \int_a^b g(t, s)x(s)ds = 0$. This implies that $x = Lw = 0$, a contradiction.

We observe that $\int_a^b \int_a^b g^2(t, s)dtds < \infty$ and hence the operator K generated by $g(t, s)$ is a compact operator on $L_2[a, b]$. K is also self-adjoint as the system Eq. (7.5.2) is self-adjoint. So, it follows from Theorem 2.2.6 that eigenvalues of K are discrete and the eigenfunctions of K form a basis for $L_2[a, b]$. As K and L are inverses of each other, it follows that eigenvalues of L are discrete and the corresponding eigenfunctions $\{\phi_1, \phi_2, \cdots, \phi_n, \cdots\}$ form a basis for $L_2[a, b]$. ∎

Theorem 7.5.2 *If* λ *is not any one of the eigenvalues* $\{\lambda_1, \lambda_2, \cdots\}$ *of* L*, then the self-adjoint system Eq. (7.5.1) has one and only one solution* $x(t) \in L_2[a, b]$ *given by*

$$x(t) = \sum_{n=1}^{\infty} \frac{f_n \phi_n(t)}{\lambda_n - \lambda}$$

where f_n *is given by* $f_n = (f, \phi_n)$*. Here (., .) is the innerproduct in* $L_2(a, b]$.

Proof : Since λ is not any one of the eigenvalues, uniqueness of the solution of Eq. (7.5.1) follows by Theorem 7.4.2. $\{\phi_1, \phi_2, \cdots\}$ is a complete orthonormal set in $L_2[a,b]$ and hence we can write

$$f(t) = \sum_{n=1}^{\infty} f_n \phi_n; \quad f_n = (f, \phi_n) = \int_a^b f(t)\phi_n(t)\, dt \tag{7.5.5}$$

and

$$x(t) = \sum_{n=1}^{\infty} x_n \phi_n; \quad x_n = (x, \phi_n) = \int_a^b x(t)\phi_n(t)\, dt \tag{7.5.6}$$

Operating on Eq. (7.5.6) by $(L - \lambda I)$ we get

$$(L - \lambda I)x(t) = \sum_{n=1}^{\infty} x_n (L - \lambda I)\phi_n(t) = \sum_{n=1}^{\infty} x_n(\lambda_n - \lambda)\phi_n(t)$$

Since $(L - \lambda I)x = f$, we get

$$\sum_{n=1}^{\infty} x_n(\lambda_n - \lambda)\phi_n(t) = \sum_{n=1}^{\infty} f_n \phi_n(t)$$

That is $\sum_{n=1}^{\infty} [f_n - x_n(\lambda_n - \lambda)]\, \phi_n(t) = 0$. Completeness of $\{\phi_1, \phi_2, \cdots\}$ implies that $f_n - x_n(\lambda_n - \lambda) = 0$ and hence $x_n = f_n/(\lambda_n - \lambda)$.

Plugging in the value of x_n in Eq. (7.5.6) we get

$$x(t) = \sum_{n=1}^{\infty} \frac{f_n}{(\lambda_n - \lambda)} \phi_n(t)$$

∎

Remark 7.5.1 *If $\lambda = \lambda_j$ for some j, Eq. (7.5.1) has infinitely many solutions if $f_j = 0$ and no solution if $f_j \neq 0$. This remark follows by Theorem 7.4.2.*

Corollary 7.5.1 *If zero is not one of the eigenvalues of the boundary value problem*

$$Lx = f; \quad B_1(x) = 0 = B_2(x) \tag{7.5.7}$$

then the above system has a unique solution $x(t)$ given by

$$x(t) = \sum_{n=1}^{\infty} \frac{f_n}{\lambda_n} \phi_n(t) = \sum_{n=1}^{\infty} \left(\int_a^b f(y)\phi_n(y)dy \right) \frac{\phi_n(t)}{\lambda_n}$$

Example 7.5.1 *Consider the self adjoint-system*

$$-\frac{d^2x}{dt^2} = f(t); \quad x(0) = 0, \quad \dot{x}(1) = \frac{1}{2}x(1)$$

The associated eigenvalue problem is

$$-\frac{d^2\phi}{dt^2} = \lambda\phi; \quad \phi(0) = 0, \quad \dot{\phi}(1) = \frac{1}{2}\phi(1)$$

We take up different cases, depending upon the algebraic sign of λ. It can be easily seen that there are no negative eigenvalues. For $\lambda > 0$, $\phi(t) = c_1 \sin\sqrt{\lambda}t$ $+c_2\cos\sqrt{\lambda}t$. Boundary condition $\phi(0) = 0$ gives $c_2 = 0$ and hence $\phi(t) = c_1\sin\sqrt{\lambda}t$.

The second boundary condition implies that λ satisfies the equation

$$\tan\sqrt{\lambda} = 2\sqrt{\lambda}$$

Let $\lambda = k^2$, then the above equation becomes $\tan k = 2k$. By considering the curves $y = \tan x$ and $y = 2x$, we see that the above equation has infinitely many solutions. So eigenvalues are given by $\lambda = k_n^2$ where k_n is a non-zero solution of $\tan k = 2k$. The corresponding normalized eigenfunctions are $-\dfrac{2}{\cos 2k_n}\sin k_n t$. These eigenfunctions are orthogonal and form a complete set. Hence, by Theorem 7.5.2, the solution $x(t)$ of the above system is given by

$$x(t) = \sum_{n=1}^{\infty} \frac{4\sin k_n t}{k_n^2(\cos 2k_n)^2}\left[\int_0^1 f(y)\sin k_n y\, dy\right]$$

Example 7.5.2 *Consider the boundary value problem*

$$\begin{aligned} -\frac{1}{t}\frac{d}{dt}\left(t\frac{dt}{dt}\right) &= f(t) \\ \lim_{t\to 0} t\frac{dx}{dt} = 0; \quad x(1) &= 0 \end{aligned}$$

We first make the above system self-adjoint by defining a new inner product $(x, y) = \displaystyle\int_0^1 tx(t)y(t)\, dt$ and hence its eigenvalues are real. The associated eigenvalue problem is

$$-\frac{1}{t}\frac{d}{dt}\left(t\frac{d\phi}{dt}\right) + \lambda\phi = 0$$

That is,

$$\frac{d\phi}{dt} + \frac{1}{t}\frac{d\phi}{dt} + \lambda\phi = 0$$

This is a Bessel's equation of zero order in the variable $\sqrt{\lambda}t$ and has the solution $\phi(t) = AJ_0(\sqrt{\lambda}t) + BY_0(\sqrt{\lambda}t)$. But the first boundary condition gives $B = 0$.

Hence, $\phi(t) = AJ_0(\sqrt{\lambda}t)$. The second boundary condition gives $J_0(\sqrt{\lambda}) = 0$. Thus, the zeros of the Bessel's function $J_0(t)$ gives the eigenvalues of the above boundary value problem. The corresponding eigenfunctions are $\phi_n = J_0(\sqrt{\lambda_n}t)$. $\{J_0(\sqrt{\lambda_n}t)\}$ forms a complete set and hence by Theorem 7.5.2

$$x(t) = \sum_{n=1}^{\infty} \frac{f_n \phi_n(t)}{\lambda_n} = \sum_{n=1}^{\infty} \frac{J_0(\sqrt{\lambda_n}t)}{\lambda_n \|J_0(\sqrt{\lambda_n}t)\|^2} \left[\int_0^1 f(y) J_0(\sqrt{\lambda_n}y)\, dy\right]$$

Theorem 7.5.3 *Let $\{\lambda_1, \lambda_2, \cdots\}$ be a discrete set of eigenvlaues of L with eigenfunctions $\{\phi_1, \phi_2, \cdots\}$ forming a complete set in $L_2[a, b]$. Then the Green's function for Eq. (7.5.1) is given by $g(t, s, \lambda) = \displaystyle\sum_{n=1}^{\infty} \frac{\phi_n(t)\phi_n(s)}{\lambda_n - \lambda}$ provided λ is not one of the eigenvalues.*

Proof : Green's function $g(t, s, \lambda)$ is the solution of the system $Lg - \lambda g = \delta(t - s)$, $B_1(g) = 0 = B_2(g)$. Applying Theorem 7.5.2 for $f(t) = \delta(t - s)$ (formally) we get

$$g(t, s, \lambda) = \sum_{n=1}^{\infty} \frac{\delta_n \phi_n(t)}{\lambda_n - \lambda}, \quad \text{where} \quad \delta_n = \int_a^b \delta(t - s)\phi_n(t)\, dt = \phi_n(t)$$

That is

$$g(t, s, \lambda) = \sum_{n=1}^{\infty} \frac{\phi_n(t)\phi_n(s)}{\lambda_n - \lambda}$$

∎

For more on this section refer Bose and Joshi [2].

7.6 Nonlinear Boundary Value Problem

In this section we shall be concerned with two point boundary value problem of the form

$$Lx = f(t, x(t)), \qquad t \in [a, b] \tag{7.6.1}$$

$$B_1(x) = \alpha, \quad B_2(x) = \beta \tag{7.6.2}$$

where L is the second order differential operator as defined before and $B_1(x)$, $B_2(x)$ are boundary conditions on x.

Here $f : [a, b] \times \Re \to \Re$ is nonlinear function satisfying Caratheodory conditons:

(i) $x \to f(t, x)$ is continuous for almost all $t \in [a, b]$.

(ii) $t \to f(t, x)$ is measurable for all values of $x \in \Re$.

Let $g(t,s)$ be the Green's function, corresponding to the operator L with given zero boundary conditions. Then the solvability of the nonlinear boundary value problem given by Eq. (7.6.1) is equivalent to the solvability of the integral equation

$$x(t) = \int_a^b g(t,s)f(s,x(s))ds + y(t) \tag{7.6.3}$$

where $y(t)$ is a known function given by $y(t) = -J(x(s), g(t,s))|_a^b$, J being the bilinear concomitant of x and g.

We shall prove the existence of solution of Eq. (7.6.3) in the space $X = L_2[a,b]$ by using operator theoretic approach in terms oprators K and N, defined as below

$$[Kx](t) = \int_a^b g(t,s)x(s)ds \tag{7.6.4}$$

$$[Nx](t) = f(t,x(t)) \tag{7.6.5}$$

Assumptions I

(i) $g(t,s)$ is a Hilbert Schmidt kernel . That is

$$k^2 = \int_a^b \int_a^b g^2(t,s)dtds < \infty \tag{7.6.6}$$

(ii) $f(t,x)$ satisfies growth condition of the form

$$|f(t,x)| \le a(t) + b|x|, \quad a \in L_2(a,b), \quad b > 0 \tag{7.6.7}$$

(iii) $x \to f(t,x)$ is Lipschitz continuous - there exists constant $\gamma > 0$ such that

$$|f(t,x_1) - f(t,x_2)| \le \gamma|x_1 - x_2| \quad \forall t \in [a,b], \quad \forall x_1, x_2 \in \Re \tag{7.6.8}$$

Theorem 7.6.1 *Let* **Assumptions I** *hold and let* $k\gamma < 1$. *Then Eq. (7.6.1) has a unique solution* $x^*(t)$ *which is obtained as a limit of the Picard iterates* $x_n(t)$, *defined iteratively as*

$$x_n(t) = Tx_{n-1}(t) = \int_0^1 g(t,s)f(s,x_{n-1}(s))ds + y(s) \tag{7.6.9}$$

with $x_0(t) = y(t)$.

Proof : Solvability of Eq. (7.6.1) or Eq. (7.6.3) is equivalent to the solvability of the operator equation

$$x = KNx + y \tag{7.6.10}$$

where $K,\ N$ are as defined by Eq. (7.6.4), Eq. (7.6.5), respectively.

We have already proved in Example 2.2.8 that K is a compact operator on X. Also, inview of **Assumption I(ii)**, it follows that N is a bounded continuous operator on X (refer Joshi and Bose [4]). Further **Assumption I(iii)** implies that N is Lipschitz continuous, that is,

$$\|Nx_1 - Nx_2\|_X \leq \gamma \|x_1 - x_2\|_X \quad \forall x_1, x_2 \in X$$

Let

$$Tx = KNx + y, \quad \forall x \in X \; (y \in X \;\text{ fixed})$$

Then

$$\begin{aligned} \|Tx_1 - Tx_2\| &= \|K(Nx_1 - Nx_2)\| \\ &\leq \|K\| \; \|Nx_1 - Nx_2\| \\ &\leq k\gamma \|x_1 - x_2\| \;\; \forall x_1, x_2 \in X \end{aligned}$$

If $\alpha = k\gamma < 1$, it follows that T is a contraction on the space X and hence it has a fixed point x^* and this point is approximated by the iterates $x_n = Tx_{n-1}, \quad x_0 = y$. This proves that $x^*(t)$ is a solution of Eq. (7.6.1) where $x^*(t) = \lim\limits_{n\to\infty} x_n(t)$. Thus, we have

$$x_n(t) \;\; = \;\; \left[y(t) + \int_a^b g(t,s) f(s, x_{n-1}(s)) ds \right]$$

and

$$x_0(t) \;\; = \;\; y(t)$$

This proves the theorem. ∎

Example 7.6.1 *Let us consider the nonlinear boundary value problem*

$$\frac{d^2x}{dt^2} = \gamma \sin x(t) \tag{7.6.11}$$

$$x(0) = \alpha, \quad x(1) = \beta \tag{7.6.12}$$

Following Example 7.3.1, we find that the Green's function $g(t,s)$ for the above problem is given by

$$g(t,s) = \begin{cases} t(s-1), & t \leq s \\ s(t-1), & t \geq s \end{cases}$$

This gives $\int_0^1 \int_0^1 g^2(t,s) dt ds = \dfrac{1}{90}$ *and hence* $k = \|K\| = \dfrac{1}{3(\sqrt{10})}$.

Solvability of Eq. (7.6.11) is equivalent to the solvability of the integral equation

$$x(t) = \gamma \int_0^1 g(t,s) \sin x(s) ds + y(t) \tag{7.6.13}$$

where $y(t) = \alpha(1-t) + \beta t$.

It is clear that $f(x) = \gamma \sin x$ *which satisfies* **Assumptions I(i) and (ii)**. *If we assume that* $\gamma \left(\frac{1}{3(\sqrt{10})} \right) < 1$, *it follows that Eq. (7.6.11) has a unique solution* $x^*(t)$, *which is the limit of* $x_n(t)$, *defined, iteratively as*

$$x_n(t) = \alpha(1-t) + \beta t + \gamma \int_0^1 g(t,s) \sin x_{n-1}(s) ds$$

with $x_0(t) = \alpha(1-t) + \beta t$.

If f is monotone increasing, we may resort to the concept of monotone operators (see Section 2.4) to get a solvability results for Eq. (7.6.1).

Assumptions II

(i) $[g(t,s)]$ is symmetric positive semidefinite (K is monotone) and Hilbert - Schmidt.

(ii) Assumption I(ii) holds.

(iii) $x \to f(t,x)$ is strongly monotone: $\exists\, c > 0$ such that

$$\left(f(t,x_1) - f(t,x_2)\right)(x_1 - x_2) \geq c(x_1 - x_2)^2 \quad x_1, x_2 \in \Re$$

Theorem 7.6.2 *Under* **Assumptions II**, *Eq. (7.6.1) has a unique solution.*

Refer Joshi and Bose [4] for the proof of this theorem.

Example 7.6.2 *Consider the following boundary value problem:*

$$\frac{d^2x}{dt^2} + \gamma x + \sin x = 0 \tag{7.6.14}$$

$$x(0) = \alpha, \quad x(1) = \beta \tag{7.6.15}$$

Assume that $\gamma > 1$.

Solvability of Eq. (7.6.14) is equivalent to the solvability of the integral equation

$$x(t) + \int_0^1 g(t,s)[\gamma x(s) + \sin x(s)] ds = \alpha(1-t) + \beta t$$

Observe that $[Kx](t) = \int_0^1 g(t,s) x(s) ds$ *is positive semi definite as* K *is the inverse of the differential operator* $Lx = -\frac{d^2x}{dt^2}$ *which is positive semi definite (monotone).*

As $f(t,x) = \gamma x + \sin x$, *we get* $\dfrac{\partial f(t,x)}{\partial x} = \gamma + \cos x \geq (\gamma - 1) > 0$ *for all* $x \in \Re$. *Hence,* $x \to f(t,x)$ *is a strongly monotone function:*

$$(f(t,x_1) - f(t,x_2))(x_1 - x_2) \geq (\alpha - 1)(x_1 - x_2)^2 \quad \forall x_1, x_2 \in \Re$$

As g *and* f *satisfy all* **Assumptions II**, *it follows that Eq. (7.6.14) has a unique solution.*

Remark 7.6.1 *If* γ *is such that* $\gamma > \left[3(\sqrt{10}) - 1\right] > 1$ *then Eq. (7.6.14) is not solvable through Theorem 7.6.1 whereas Theorem 7.6.2 is applicable.*

7.7 Exercises

1. Consider the operator L with boundary conditions $\dot{x}(0) = x(1), \;\; \dot{x}(1) = 0$. Find L^* and adjoint boundary conditions.

2. Find the Green's function for the operator L given by

   (a) $L = (t+1)\dfrac{d^2}{dt^2}$ with boundary conditions $x(0) = 0 = x(1)$.

   (b) $L = \dfrac{d^2}{dt^2}$ with boundary conditions $\dot{x}(0) = x(1), \;\; \dot{x}(1) = 0$.

3. Find the modified symmetric Green's function for the system $L = \dfrac{d^2}{dt^2}$, $-1 \leq t \leq 1$ with boundary conditions $x(-1) = x(1)$ and $\dot{x}(-1) = \dot{x}(1)$.

4. Show that the differential operator

$$Lx = -\frac{1}{w(t)}(a_0(t)\dot{x}) + a_2(t)x, \quad w(t) > 0$$

   is formally self adjoint if the inner product is defined by

$$(x,y) = \int_0^1 w(t)x(t)y(t)dt$$

5. Prove that the solution $x(t) = \int_0^1 g(t,s)f(s)ds$ of the system

$$Lx = f, \quad B_1(x) = 0 = B_2(x)$$

   depends continuously on f.

6. Using Green's function technique, show that the solution $x(t)$ of the boundary value problem

$$\frac{d^2x}{dt^2} + tx = 1, \quad x(0) = \dot{x}(0) = 0$$

satisfies the Volterra integral equation

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

Prove that the converse of the above is also true. Compare this approach with that of the initial value problem.

7. Obtain the solution of the following boundary value problems through eigenfunction expansion.

   (a) $\frac{d^2x}{dt^2} + x = t, \;\; x(0) = 0, \;\; \dot{x}(1) = 0$

   (b) $\frac{d^2x}{dt^2} + x = \sin t, \;\; x(1) = 2, \;\; \dot{x}(0) = 1$

   (c) $\frac{d^2x}{dt^2} + x = e^t, \;\; x(0) = x(1), \;\; \dot{x}(0) = \dot{x}(1)$

8. Compute the first few Picard iterates for the following nonlinear boundary value problems

   (a) $\frac{d^2x}{dt^2} = x + x^3, \;\; x(0) = 0 = x(1)$

   (b) $\frac{d^2x}{dt^2} = x + \sin x, \;\; \dot{x}(0) = 0 = \dot{x}(1)$

References

[1] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations, McGraw Hill Book Co., 1955*

[2] R. K. Bose and M. C. Joshi, *Methods of Mathematical Physics, Tata McGraw Hill Publ., co., Ltd., 1984*

[3] N. Dunford and J.T. Schwartz, *Linear Operators, Parts I and II, Interscinece Publ., Inc., 1966*

[4] M. C. Joshi and R. K. Bose, *Some Topic in Nonlinear Functional Analysis, Halsted Press, 1968*

[5] I. Stakgold, *Boundary Vlaue Problems of Mathematical Physics, Vol I and II, The Macmillan Co. 1968*

Chapter 8

Control Theory Of Linear Systems

Control theory deals with the maneuver of the state or trajectory of the system, modelled by ODE. Some of the important properties of such systems are controllability, observability and optimality.
In this chapter we show how operator theoretic approach in function spaces is useful to discuss the above mentioned properties. We substantiate our theory by concrete examples.

8.1 Controllability

We shall be interested in the linear system of the form

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + B(t)\bar{u}(t) \quad (8.1.1a)$$

$$\bar{x}(t) = \bar{x}_0 \quad (8.1.1b)$$

where $A(t)$ is $n \times n$ matrix, $B(t)$ is $n \times m$ matrix, $\bar{u}(t) \in \Re^m$ and $\bar{x}(t) \in \Re^n$. $\bar{u}(t)$ is called control or input vector and $\bar{x}(t)$ the corresponding trajectory or state of the system.

The typical controllability problem involves the determination of the control vector $\bar{u}(t)$ such that the state vector $\bar{x}(t)$ has the desired properties. We assume that the entries of the matrices $A(t)$, $B(t)$ are continuous so that the above system has a unique solution $\bar{x}(t)$ for a given input $\bar{u}(t)$.

Definition 8.1.1 *The linear system given by Eq. (8.1.1) is said to be controllable if given any initial state $\bar{x}_0$ and any final state $\bar{x}_f$ in $\Re^n$, there exist a control $\bar{u}(t)$ so that the correspondoing trajectory $\bar{x}(t)$ of Eq. (8.1.1) satisfies the condition*

$$\bar{x}(t_0) = \bar{x}_0, \quad \bar{x}(t_f) = \bar{x}_f \quad (8.1.2)$$

The control $\bar{u}(t)$ said to steer the trajectory from the initial state $\bar{x}_0$ to the final state $\bar{x}_f$.

Assume that $\phi(t, t_0)$ is the transition matrix of the above system. Then by the variation of parameter formula, Eq. (8.1.1) is equivalent to the following integral equation

$$\bar{x}(t) = \phi(t, t_0)\bar{x}_0 + \int_{t_0}^{t} \phi(t, \tau)B(\tau)\bar{u}(\tau)d\tau \tag{8.1.3}$$

As $(t, \tau) \to \phi(t, \tau)$ is continuous, it follows that $\|\phi(t, \tau)\| \leq M$ for all $t, \tau \in [t_0,\ t_f]$.

So controllability of Eq. (8.1.1) is equivalent to finding $\bar{u}(t)$ such that

$$\bar{x}_f = \bar{x}(t_f) = \phi(t_f, t_0)\bar{x}_0 + \int_{t_0}^{t_f} \phi(t_f, \tau)B(\tau)\bar{u}(\tau)d\tau$$

Equivalently,

$$\bar{x}(t_f) - \phi(t_f, t_0)\bar{x}_0 = \int_{t_0}^{t_f} \phi(t_f, \tau)B(\tau)\bar{u}(\tau)d\tau \tag{8.1.4}$$

Let us define a linear operator $L : X = L_2\left([t_0, t_f],\ \Re^m\right)$ by

$$[L\bar{u}] = \int_{t_0}^{t_f} \phi(t_f, \tau)B(\tau)\bar{u}(\tau)d\tau \tag{8.1.5}$$

Then, in view of (8.1.4), the controllability problem reduces to showing that operator L is surjective.

Theorem 8.1.1 *The system given by Eq. (8.1.1) is controllable iff the controllability Grammian*

$$W(t_0, t_f) = \int_{t_0}^{t_f} \left[\phi(t_f, \tau)B(\tau)B^{\top}(\tau)\phi^{\top}(t, t_f)\right] d\tau \tag{8.1.6}$$

is nonsingular. A control $\bar{u}(t)$ steering the system from the initial state $\bar{x}_0$ to the final state $\bar{x}_f$ is given by

$$\bar{u}(t) = B^{\top}(t)\phi^{\top}(t_f, t)\left[W(t_0, t_f)\right]^{-1}\left[\bar{x}_f - \phi(t_f, t_0)\bar{x}_0\right]$$

Proof : Let $(\bar{x}_0, \bar{x}_f)$ be any pair of vectors in $\Re^n \times \Re^n$. The controllability problem of Eq. (8.1.1) reduces to the determination of $\bar{u} \in X$ such that

$$\int_{t_0}^{t_f} \phi(t_f, \tau)B(\tau)\bar{u}(\tau)d\tau = \bar{z}_f = \bar{x}_f - \phi(t_f, t_0)\bar{x}_0$$

This is equivalent to solving

$$L\bar{u} = \bar{z}_f$$

in the space X for a given $\bar{z}_f$, where L is given by Eq. (8.1.5).

We first note that L is a bounded linear operator as $\phi(t,\tau)$ is bounded. Let $L^* : \Re^n \to X$ be its adjoint. This is determined as follows.
For $\bar{u} \in X$ and $\bar{\alpha} \in \Re^n$, we have

$$\begin{aligned} (L\bar{u}, \bar{\alpha})_{\Re^n} &= \left(\bar{\alpha}, \int_{t_0}^{t_f} \phi(t_f,\tau)B(\tau)\bar{u}(\tau)d\tau\right)_{\Re^n} \\ &= \int_{t_0}^{t_f} (\bar{\alpha}, \phi(t_f,\tau)B(\tau)\bar{u}(\tau))_{\Re^n}\, d\tau \\ &= \int_{t_0}^{t_f} \left(B^\top(\tau)\phi^\top(t_f,\tau)\bar{\alpha}, \bar{u}(\tau)\right)_{\Re^n} d\tau \\ &= (\bar{u}, L^*\bar{\alpha})_X \end{aligned}$$

This implies that

$$[L^*\bar{\alpha}](t) = B^\top(t)\phi^\top(t_f,t)\bar{\alpha} \tag{8.1.7}$$

To claim that L is onto, it is sufficient to prove that $LL^* : \Re^n \to \Re^n$ is onto (equivalently nonsingular). That is, given any $\bar{z}_f \in \Re^n$, find $\bar{\alpha} \in \Re^n$ such that

$$LL^*\bar{\alpha} = \bar{z}_f \tag{8.1.8}$$

If Eq. (8.1.8) is solvable, then a control $\bar{u}$ which solves $L\bar{u} = \bar{z}_f$ is given by $\bar{u} = L^*\bar{\alpha}$.
Computing LL^* we get

$$\begin{aligned} [LL^*](\bar{\alpha}) &= L\left(B^\top(t)\phi^\top(t_f,t)\bar{\alpha}\right) \\ &= \left[\int_{t_0}^{t_f} \left[\phi(t_f,t)B(t)B^\top(t)\phi^\top(t_f,t)\right] dt\right]\bar{\alpha} \\ &= W(t_0,t_f)\bar{\alpha} \end{aligned}$$

If the controllability Grammian $W(t_0,t_f)$ is nonsingular, we have a control $\bar{u}(t)$ given by

$$\begin{aligned} \bar{u}(t) = L^*\alpha &= L^*(LL^*)^{-1}\bar{z}_f \\ &= L^*\left[W(t_0,t_f)\right]^{-1}\bar{z}_f \\ &= B^\top(t)\phi^\top(t_f,t)\left[W(t_0,t_f)\right]^{-1}\left[\bar{x}_f - \phi(t_f,t_0)\bar{x}_0\right] \end{aligned}$$

which will steer the state from the initial state $\bar{x}_0$ to the final state $\bar{x}_f$.

Conversely, let the Eq. (8.1.1) be controllable. That is, L is onto. Equivalently, $R(L) = \Re^n$. Then, by Theorem 2.2.1, $N(L^*) = [R(L)]^\perp = [\Re^n]^\perp = \{0\}$. Hence L^* is 1-1. This implies LL^* is also 1-1, which is proved as under. Assume that $LL^*\bar{\alpha} = 0$. This gives $0 = (LL^*\bar{\alpha}, \bar{\alpha}) = (L^*\bar{\alpha}, L^*\alpha) = \|L^*\alpha\|^2$ and hence $L^*\bar{\alpha} = \bar{0}$. L^* 1-1 implies that $\bar{\alpha} = 0$. Thus $N(LL^*) = \{\bar{0}\}$.

Thus $W(t_0, t_f) = LL^* : \Re^n \to \Re^n$ is 1-1 and hence nonsingular . This proves the result. ∎

The controllability Grammian $W(t_0, t_f)$ has some intersting properties, which we now describe.

Theorem 8.1.2 *(i) $W(t_0, t_f)$ is symmetric and positive semidefinite.*
(ii) $W(t_0, t_f)$ satisfies the linear differential equation

$$\begin{aligned} \frac{d}{dt}[W(t_0, t)] &= A(t)W(t_0, t) + W(t_0, t)A^\top(t) + B(t)B^\top(t) \\ W(t_0, t_0) &= 0 \end{aligned}$$

(iii) $W(t_0, t_f)$ satisfies functional equation

$$W(t_0, t_f) = W(t_0, t) + \phi(t_f, t)W(t, t_f)\phi^\top(t_f, t)$$

Proof : We have

$$W(t_0, t_f) = \int_{t_0}^{t_f} \left[\phi(t_f, \tau)B(\tau)B^\top(\tau)\phi^\top(t_f, \tau)\right] d\tau$$

(i) It is clear that $W^\top(t_0, t_f) = W(t_0, t_f)$. Further,

$$\begin{aligned} (W(t_0, t_f)\bar{\alpha}, \bar{\alpha})_{\Re^n} &= \int_{t_0}^{t_f} \left(\phi(t_f, \tau)B(\tau)B^\top(\tau)\phi^\top(t_f, \tau)\bar{\alpha},\ \bar{\alpha}\right)_{\Re^n} d\tau \\ &= \int_{t_0}^{t_f} \|B^\top(\tau)\phi^\top(t_f, \tau)\bar{\alpha}\|^2 \\ &\geq 0 \end{aligned}$$

for all $\bar{\alpha} \in \Re^n$.

(ii) Using Leibnitz rule, we get

$$\begin{aligned} \frac{d}{dt}(W(t_0, t)) &= \frac{d}{dt}\left[\int_{t_0}^{t} \phi(t_f, \tau)B(\tau)B^\top(\tau)\phi(t, \tau)d\tau\right] \\ &= B(t)B^\top(t) + \left[\int_{t_0}^{t} \frac{d}{dt}\left[\phi(t_f, \tau)B(\tau)B^\top(\tau)\phi^\top(t, \tau)\right] d\tau\right] \\ &= B(t)B^\top(t) + \left[\int_{t_0}^{t} A(t)\phi(t_f, \tau)B(\tau)B^\top(\tau)\phi(t, \tau)d\tau\right] \\ &\quad + \left[\int_{t_0}^{t} \phi(t_f, \tau)B(\tau)B^\top(\tau)\phi^\top(t, \tau)A^\top(t)d\tau\right] \\ &= A(t)W(t_0, t) + W(t_0, t)A^\top(t) + B(t)B^\top(t) \end{aligned}$$

Bounadry condition $W(t_0, t_0) = 0$ follows from the definition.

(iii) To get the functional equation we express $W(t_0, t_f)$ as follows

$$\begin{aligned} W(t_0, t_f) &= \int_{t_0}^{t} \phi(t_f, \tau)B(\tau)B^\top(\tau)\phi^\top(t_f, \tau)d\tau \\ &\quad + \int_{t}^{t_f} \phi(t_f, \tau)B(\tau)B^\top(\tau)\phi^\top(t_f, \tau)d\tau \\ &= W(t_0, t) + \phi(t_f, t)\left(\int_{t}^{t_f} \phi(t, \tau)B(\tau)B^\top(\tau)\phi^\top(t, \tau)d\tau\right)\phi^\top(t_f, t) \\ &= W(t_0, t) + \phi(t_f, t)W(t, t_f)\phi^\top(t_f, t) \end{aligned}$$

∎

Remark 8.1.1 *Theorem 8.1.1 implies that to check the controllability of the linear system given by Eq. (8.1.1), one needs to verify the invertibility of the Grammian matrix $W(t_0, t_f)$. This is a very tedious task. However, if $A(t)$ and $B(t)$ are time invariant matrics A, B, then the controllability of the linear system is obtained in terms of the rank of the following controllability matrix*

$$C = \left[B, AB, \cdots, A^{n-1}B\right] \tag{8.1.9}$$

The next theorem, in this direction, is due to Kalman [4].

Theorem 8.1.3 *The linear autonomous system*

$$\frac{d\bar{x}}{dt} = A\bar{x}(t) + B\bar{u}(t) \tag{8.1.10a}$$

$$\bar{x}(t_0) = \bar{x}_0 \tag{8.1.10b}$$

is controllable iff the controllability matrix C given by Eq. (8.1.9) has rank n.

Proof : Assume the rank of $C = n$. It follows that $N(C^\top) = \{\bar{0}\}$. We shall prove that $N(W(t_0, t_f)) = \{\bar{0}\}$. Let $\bar{\alpha} \in \Re^n$ be such that $W(t_0, t_f)\bar{\alpha} = 0$. This implies that

$$\begin{aligned} 0 = \|W(t_0, t_f)\bar{\alpha}\|^2 &= \int_{t_0}^{t_f} \|B^\top\phi^\top(t_f, \tau)\bar{\alpha}\|^2 d\tau \\ &= \int_{t_0}^{t_f} \|B^\top e^{A^\top(t_f-\tau)}\bar{\alpha}\|^2 d\tau \end{aligned}$$

This implies that

$$B^\top e^{A^\top(t_f-\tau)}\bar{\alpha} = 0 \text{ for all } t \in [t_0, t_f] \tag{8.1.11}$$

Expanding the LHS of Eq. (8.1.11) in a Taylor series arround $t = t_f$, we get that

$$0 = B^\top\bar{\alpha} = B^\top(A^\top)\bar{\alpha} = B^\top(A^\top)^2\bar{\alpha} = \cdots B^\top(A^\top)^{n-1}\bar{\alpha}$$

This implies that $\bar{\alpha} \in N(C^\top) = \{\bar{0}\}$.

Conversely, let $N(W(t_0, t_f)) = \{\bar{0}\}$. We shall show that $N(C^\top) = \{\bar{0}\}$. Assume that $\bar{\alpha} \in N(C^\top)$. This implies that

$$B^\top (A^\top)^k \bar{\alpha} = 0, \quad 1 \leq k \leq n-1 \tag{8.1.12}$$

We have

$$W(t_0, t_f)\bar{\alpha} = \int_{t_0}^{t_f} \left[\phi(t_f, \tau) B B^\top \phi^\top(t, \tau)\bar{\alpha}\right] d\tau$$

As $\phi(t_f, \tau) = e^{A(t_f - \tau)}$, it follows by Cayley-Hamilton theorem (refer problem number 4 in Section 5.5), that there exist functions $\alpha_k(t_f - \tau)$ such that $\phi(t_f, \tau) = \sum_{k=0}^{n-1} \alpha_k(t_f - \tau) A^k$ and hence $W(t_0, t_f)\bar{\alpha}$ is given by

$$W(t_0, t_f)\bar{\alpha} = \int_{t_0}^{t_f} \phi(t_f, \tau) B(\tau) \left[\sum_{k=0}^{n-1} \alpha_k(t_f - \tau) B^\top (A^T)^k \bar{\alpha}\right] d\tau \tag{8.1.13}$$

In view of Eq. (8.1.12), we have $\sum_{k=0}^{n-1} \alpha_k(t_f - \tau) B^\top (A^\top)^k \bar{\alpha} = 0$. This implies that $\bar{\alpha} \in N(W(t_0, t_f)) = \{\bar{0}\}$.
This proves the theorem. ∎

Example 8.1.1 *Let us recall Example 2.3.5 which gives the linearized motion of a satellite of unit mass orbiting around the earth. This is given by the control system of the form given by Eq. (8.1.10), where*

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3w^2 & 0 & 0 & 2w \\ 0 & 0 & 0 & 1 \\ 0 & -2w & 0 & 0 \end{bmatrix}$$

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

For this system, one can easily compute the controllability matrix, $C = \left[B, AB, A^2B, A^3B\right]$. *It is given by*

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 2w & -w^2 & 0 \\ 1 & 0 & 0 & 2w & -w^2 & 0 & 0 & -2w^3 \\ 0 & 0 & 0 & 1 & -2w & 0 & 0 & -4w^2 \\ 0 & 1 & -2w & 0 & 0 & -4w^2 & 2w^3 & 0 \end{bmatrix}$$

One can verify that rank of C is 4 and hence the linearized motion of the satellite is controllable. It is intersting to ask the following question.
What happens when one of the controls or thrusts becomes inoperative ?

For this purpose set $u_2 = 0$ and hence B reduces to $B_1 = [0\ 1\ 0\ 0]^\top$.

So, the controllability matrix $C_1 = [B_1, AB_1, A^2B_1, A^3B_1]$, is given by

$$C_1 = \begin{bmatrix} 0 & 1 & 0 & -w^2 \\ 1 & 0 & -w^2 & 0 \\ 0 & 0 & -2w & 0 \\ 0 & -2w & 0 & 2w^3 \end{bmatrix}$$

C_1 has rank 3.

On the other hand $u_1 = 0$ reduces B to $B_2 = [0\ 0\ 0\ 1]^\top$ and this gives controllability matrix $C_2 = [B_2, AB_2, A^2B_2, A^3B_2]$, as

$$C_2 = \begin{bmatrix} 0 & 0 & 2w & 0 \\ 0 & 2w & 0 & -2w^3 \\ 0 & 1 & 0 & -4w^2 \\ 1 & 0 & -4w^2 & 0 \end{bmatrix}$$

C_2 has rank 4.

Since u_1 was radial thrust and u_2 was trangential thrust, we see that the loss of radial thrust does not destroy controllability where as loss of tangential thrust does. In terms of practical importance of satellite in motion, the above analysis means that we can maneuver the satellite just with radial rocket thrust.

8.2 Observability

Consider the input-output system of the form

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + B(t)\bar{x}(t) \tag{8.2.1a}$$

$$\bar{y}(t) = C(t)\bar{x}(t) \tag{8.2.1b}$$

The questions concerning observability relate to the problem of determining the values of the state vector $\bar{x}(t)$, knowing only the output vector $\bar{y}(t)$ over some interval $I = [t_0, t_f]$ of time.

As in the previous section, $A(t) \in \Re^{n\times n}$, $B(t) \in \Re^{n\times m}$ are assumed to be continuous functions of t. Let $\phi(t, t_0)$ be the transition matrix of the above system. Then the output vector $\bar{y}(t)$ can be expresssed as

$$\bar{y}(t) = C(t)\phi(t, t_0)\bar{x}_(t_0) + \bar{y}_1(t) \tag{8.2.2}$$

where $\bar{y}_1(t)$ is known quantity of the form

$$y_1(t) = \int_{t_0}^{t_f} C(t)\phi(t, \tau)B(\tau)u(\tau)d\tau$$

Thus, from Eq. (8.2.2) it follows that if we are concerned about the determination of $\bar{x}(t_0)$, based on the output $\bar{y}(t)$, we need only the homogeneous system

$$\frac{d\bar{x}}{dt} = A\bar{x}(t) \tag{8.2.3a}$$

$$\bar{y}(t) = C(t)\bar{x}(t) \tag{8.2.3b}$$

in place of Eq. (8.2.1).

Definition 8.2.1 *We shall say that Eq. (8.2.3) is observable on $I = [t_0, t_f]$ if $\bar{y}(t) = 0$ on I implies that $\bar{x}(t) = 0$ on I.*

We define a linear operator $L : \Re^n \mapsto X = L_2\left([t_0, t_f],\ \Re^n\right)$ as

$$[L\bar{x}_0](t) = C(t)\phi(t, t_0)\bar{x}_0 = H(t)\bar{x}_0 \tag{8.2.4}$$

where $H : t \to H(t)$ is a matrix function which is continuous in t. Observability of the system Eq. (8.2.3), then, reduces to proving that L is one-one.

The following theorem gives the observability of the system given by Eq. (8.2.3), in terms of the nonsingularity of the matrix $M(t_0, t_f)$.

Theorem 8.2.1 *For the homogeneous system given by Eq. (8.2.3), it is possible to determine the initial state $\bar{x}(t_0)$ within an additive constant vector which lies with null space of $M(t_0, t_f)$ which is defined by*

$$M(t_0, t_f) = \int_{t_0}^{t_f} \phi^\top(t, t_0)C^\top(t)C(t)\phi(t, t_0)dt \tag{8.2.5}$$

Hence $\bar{x}(t_0)$ is uniquely determined if $M(t_0, t_f)$ is nonsingular. That is, Eq. (8.2.3) is observable ifff $M(t_0, t_f)$ is nonsingular.

Proof : If the output vector $\bar{y}(t)$ is known, solving $\bar{y}(t) = C(t)\phi(t, t_0)\bar{x}(t_0)$ for $\bar{x}(t_0)$ is equivalent to solving the operator equation

$$L\bar{x}(t_0) = \bar{y} \tag{8.2.6}$$

where L is defined by Eq. (8.2.4). Premultiplying Eq. (8.2.6) by L^*, we get

$$[L^*L]\,\bar{x}(t_0) = L^*\bar{y} \tag{8.2.7}$$

where $L^* : X \to \Re^n$ is given by

$$[L^*u] = \int_{t_0}^{t_f} \phi^\top(t, t_0)C^\top(t)u(t)dt \tag{8.2.8}$$

Observe that $M(t_0, t_f) = L^*L$. If RHS of Eq. (8.2.7) is known, $\bar{x}(t_0)$ is determined up to an additive constant lying in the null space of $L^*L = M(t_0, t_f)$. If L^*L is invertible, then it is clear that L is one-one and hence the system given by Eq. (8.2.3) is observable.

Conversely, if L is one-one then so is L^*L and hence $R(L^*L) = [N(L^*L)]^\perp = \Re^n$. This implies that L^*L is nonsingular. For such an observable system, the initial state $\bar{x}(t_0)$ is given by

$$\bar{x}(t_0) = (L^*L)^{-1}L^*y = [M(t_0, t_f)]^{-1}\left[\int_{t_0}^{t_f} \phi^\top(t, t_0)C^\top(t)y(t)dt\right] \tag{8.2.9}$$

This proves the theorem. ∎

$M(t_0, t_f)$ defined by Eq. (8.2.5) is called the observability Grammian.

Theorem 8.2.2 *The observability Grammian $M(t_0, t_f)$ satisfies the following properties*

(i) $M(t_0, t_f)$ is symmetric and positive semidefinite.

(ii) $M(t_0, t_f)$ satisfies the following matrix differential equation

$$\begin{aligned} \frac{d}{dt}[M(t, t_1)] &= -A^\top(t)M(t, t_1) - M(t, t_1)A(t) - C^\top(t)C(t) \\ M(t_1, t_1) &= 0 \end{aligned}$$

(iii) $M(t_0, t_f)$ satisfies the functional equation

$$M(t_0, t_f) = M(t_0, t) + \phi^\top(t, t_0)M(t, t_1)\phi(t, t_0)$$

The proof of the above theorem follows along the lines of the proof of the Theorem 8.1.2. See Brocket [1] for more details.

We also have the following theorem giving the necessary and sufficient contition which is easily verifiable for observability.

Let us denote by O the observability matrix

$$O = [C, CA, \cdots, CA^{n-1}] \tag{8.2.10}$$

Theorem 8.2.3 *The system given by Eq. (8.2.3) is observable iff the observable matrix O given by Eq. (8.2.10) is of rank n.*

Example 8.2.1 *As a continuation of Example 8.1.1, consider the satellite orbiting around the Earth.*
We assume that we can measure only the distance r of the satellite from the centre of the force and the angle θ. So, the defining state-output equation of the satellite is given by

$$\begin{aligned} \frac{d\bar{x}}{dt} &= A\bar{x}(t) \\ \bar{y}(t) &= Cx(t) \end{aligned}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3w^2 & 0 & 0 & 2w \\ 0 & 0 & 0 & 1 \\ 0 & -2w & 0 & 0 \end{bmatrix}$$

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

$\bar{x} = (x_1, x_2, x_3, x_4)$, $\bar{y} = (y_1, y_2)$; $y_1 = r$, $y_2 = \theta$ *being the radial and angle measuments. So, the observability matrix O is given by*

$$\begin{aligned} O &= [C, CA, CA^2, CA^3] \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 3w^2 & 0 & 0 & 2w \\ 0 & -2w & 0 & 0 \\ 0 & -w^2 & 0 & 0 \\ -6w^3 & 0 & 0 & -4w^2 \end{bmatrix} \end{aligned}$$

Rank of C is 4 and hence the above state -output system is observable.

To minimize the measurement, we might be tempted to measure y_1 only not y_2. This gives

$$
\begin{aligned}
C_1 &= [1\ 0\ 0\ 0\ 0]^\top \quad \textit{and} \\
O_1 &= [C_1, C_1A, C_1A^2, C_1A^3] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3w^2 & 0 & 0 & 2w \\ 0 & -w^2 & 0 & 0 \end{bmatrix}
\end{aligned}
$$

which is of rank 3.
However, if y_1 is not measured, we get

$$
\begin{aligned}
C_2 &= [0\ 0\ 1\ 0\ 0]^\top \quad \textit{and} \\
O_2 &= [C_2, C_2A, C_2A^2, C_2A^3] \\
&= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -2w & 0 & 0 \\ -6w^3 & 0 & 0 & -4w^2 \end{bmatrix}
\end{aligned}
$$

which is of rank 4.
Thus, the state of the satellite is known from angle measurements alone but this is not so for radial measurements.

8.3 Optimal Control

In this section we shall be concerned with the optimization problem of two types.

(i) Unconstrained optimal control problem:
Find the control $\bar{u}^* \in U = L_2([t_0, t_f], \Re^m)$ that minimizes the cost functional $J(u)$ which is given by

$$J(\bar{u}) = \int_{t_0}^{t_f} [(\bar{u}(t), \bar{v}(t)) + (R\bar{x}(t), \bar{x}(t))]\, dt \tag{8.3.1}$$

where $\bar{x} \in X$ is the state of the dynamical system

$$
\begin{aligned}
\frac{d\bar{x}}{dt} &= A(t)\bar{x}(t) + B(t)\bar{u}(t) && (8.3.2a) \\
\bar{x}(t_0) &= \bar{x}_0 && (8.3.2b)
\end{aligned}
$$

corresponding to the given control $\bar{u}(t)$. R is a positive definite symmetric matrix and $X = L_2([t_0, t_f],\ \Re^n)$.

(ii) Constrained optimal control system:

Find the control $\bar{u}^*$ that minimizes $J(\bar{u})$ over the constrained set $U_c \subseteq U$, where $J(\bar{u})$ and U_c are given as follows

$$J(\bar{u}) = \int_{t_0}^{t_f} (\bar{u}(t), \bar{u}(t))\, dt \tag{8.3.3}$$

$U_c = \{\bar{u} \in U :\ \bar{u}$ steers the state $\bar{x}(t)$ of Eq. (8.3.1) from the initial state $\bar{x}_0$ to the final state $\bar{x}_f\}$.

We shall show that unconstrained problem is a problem of minimization of a strictly convex functional defined on the Hilbert space U.

Using the transition matrix $\phi(t, t_0)$, we obtain the state $\bar{x}(t)$ of the linear system given by Eq. (8.3.2) through the integral equation.

$$\bar{x}(t) = \phi(t, t_0)\bar{x}_0 + \int_{t_0}^{t} \phi(t, \tau)B(\tau)u(\tau)d\tau \tag{8.3.4}$$

We now define a bounded linear operator K from U to X as follows.

$$[K\bar{u}](t) = \int_{t_0}^{t} \phi(t, \tau)B(\tau)\bar{u}(\tau)d\tau$$

Set $\bar{y}_0(t) = \phi(t, t_0)\bar{x}_0$, then in abstract notation, $\bar{x} \in X$ is given by

$$\bar{x} = \bar{y}_0 + K\bar{u} \tag{8.3.5}$$

So, the cost functional J on U, determined by Eq. (8.3.1), is given by

$$J(\bar{u}) = (\bar{u}, \bar{u})_U + (T\bar{x}, \bar{x})_X$$

where T is a positive definite operator induced by R on X.

Using the concept of the derivative of a functional (refer section 2.3), we obtain

$$J'(\bar{u}) = \bar{u} + K^*TK\bar{u} + K^*TK\bar{y}_0 \tag{8.3.6}$$

and

$$J''(u) = I + K^*TK \ \ \forall \ \ \bar{u} \in U \tag{8.3.7}$$

Theorem 8.3.1 *J is a strictly convex functional on the space U and hence it has a unique minimizer $\bar{u}^* \in U$ of $J(u)$ and is given by*

$$\bar{u}^* = -\left[I + K^*TK\right]^{-1}\left[K^*TK^*\bar{y}_0\right] \tag{8.3.8}$$

Proof : We have

$$\begin{aligned} J'(\bar{u}) &= \bar{u} + K^*TK\bar{u} + K^*TK\bar{y}_0 \\ J''(\bar{u}) &= I + K^*TK \end{aligned}$$

for all $\bar{u} \in U$.
Note that $J''(u) : U \to U$ is positive definite for all $u \in U$, as

$$\begin{aligned} (J''(u)\bar{v}, \bar{v}) &= (\bar{v}, \bar{v})_U + (K^*TK\bar{v}, \bar{v})_U \\ &= \|\bar{v}\|^2 + (TK\bar{v}, K\bar{v})_X \\ &> \|\bar{v}\|^2 \ \ (\text{for } \ \bar{v} \neq 0) \end{aligned}$$

in view of positive definiteness of T. As $J''(u)$ is positive definite, it follows by Remark 2.3.2 that J is strictly and hence it follows that J has unique minimizer $\bar{u}^* \in U$ and is given by the critical point of J (refer Joshi and Kannan [5] Chapter1). Hence, we have

$$\bar{u}^* + K^*TK\bar{u}^* = -K^*TK\bar{y}_0$$

Also, the invertibility of $[I + K^*TK]$ (refer Joshi and Bose [3]), gives

$$\bar{u}^* = -[I + K^*TK]^{-1}(K^*TK\bar{y}_0)$$

∎

Remark 8.3.1 *One can obtain integro-differential equation analog of Eq. (8.3.8) which is equivalent to the equation*

$$\bar{u}^*(t) + \int_t^{t_f} \phi(\tau, t)B(TK\bar{u}^*)(\tau)d\tau + \int_t^{t_f} (TK)(\bar{y}_0(\tau))d\tau = 0 \qquad (8.3.9)$$

Differentiating Eq. (8.3.9) we get

$$\frac{d\bar{u}^*}{dt} - BTKu^*(t) + \int_t^{t_f} \frac{d}{dt}[\phi(\tau, t)]\, BTKu^*(\tau)d\tau = TK\bar{y}_0(t)$$

As $\frac{d}{dt}[\phi(\tau, t)] = -\phi(\tau, t)A$, *we get the equivalent integro-differential equation*

$$\begin{aligned} \frac{d\bar{u}^*}{dt} = -BTKu^*(t) - \int_t^{t_f} \phi(\tau, t)ABTKu^*(\tau)d\tau \qquad & (8.3.10a) \\ = TK\bar{y}_0(t) \qquad & (8.3.10b) \end{aligned}$$

with $u^*(t_f) = 0$.

In the above equation we assume that the system Eq.(8.3.2) is autonomous.

We now address ourself to second problem pertaining to constrained optimal control. We have the following theorem in this direction.

Theorem 8.3.2 *Let $W(t_0, t_f)$ be the controllability Grammian of Eq. (8.3.2). If $\bar{u}^*$ is any control of the form*

$$\bar{u}^*(t) = B^\top(t)\phi^\top(t_f, t)\bar{\alpha} \tag{8.3.11}$$

where $\bar{\alpha}$ satisfies the equation

$$W(t_0, t_f)\bar{\alpha} = \bar{x}_f - \phi(t_f, t)\bar{x}_0 \tag{8.3.12}$$

Then $\bar{u}^* \in U_c$. Further, if $\bar{u}$ is any, controllability vector belonging to U_c, then

$$\int_{t_0}^{t_f} \|u^*(t)\|^2 dt \leq \int_{t_0}^{t_f} \|\bar{u}(t)\|^2 dt \tag{8.3.13}$$

Moreover, if $W(t_0, t_f)$ is nonsingular, then

$$\int_{t_0}^{t_f} \|\bar{u}^*(t)\|^2 dt = \left(W^{-1}(t_0, t_f)\left[\bar{x}_f - \phi(t_f, t_0)\right]\bar{x}_0, \left[\bar{x}_f - \phi(t_f, t_0)\right]\bar{x}_0\right)$$

Proof : We have

$$\begin{aligned} \bar{z}_f &= \bar{x}_f - \phi(t_f, t_0)\bar{x}_0 = \int_{t_0}^{t_f} \phi(t_f, \tau)B(\tau)\bar{u}(\tau)d\tau = L\bar{u} \\ \bar{z}_f &= \bar{x}_f - \phi(t_f, t_0)\bar{x}_0 = \int_{t_0}^{t_f} \phi(t_f, \tau)B(\tau)u^*(\tau)d\tau = L\bar{u}^* \end{aligned}$$

where L is defined by Eq. (8.1.5).
Subtracting the above two equations, we have

$$L\bar{u} - L\bar{u}^* = 0$$

This gives

$$\left(L\bar{u}^*, \bar{\alpha}\right)_U = \left(L\bar{u}, \bar{\alpha}\right)_U \tag{8.3.14}$$

where $\bar{\alpha}$ solves Eq. (8.3.12). Solvability of Eq. (8.3.12) is equivalent to the solvability of the following equation

$$LL^*\bar{\alpha} = \bar{z}_f$$

Also Eq. (8.3.11) is equivalent to defining $\bar{u}^* = L^*\bar{\alpha}$. Hence, Eq. (8.3.14) gives

$$\begin{aligned} \|\bar{u}^*\|_U^2 &= (\bar{u}^*, \bar{u}^*)_U \\ &= (\bar{u}^*, L^*\bar{\alpha})_U = (L\bar{u}, \alpha)_U = (\bar{u}, L^*\alpha)_U = (\bar{u}, \bar{u}^*)_U \end{aligned} \tag{8.3.15}$$

Eq. (8.3.15) gives

$$\begin{aligned} 0 \leq \|\bar{u} - \bar{u}^*\|_U^2 &= (\bar{u} - \bar{u}^*, \bar{u} - \bar{u}^*)_U \\ &= \|\bar{u}\|_U^2 - 2(\bar{u}, \bar{u}^*)_U + \|\bar{u}^*\|_U^2 \\ &= \|\bar{u}\|_U^2 - \|\bar{u}^*\|_U^2 \end{aligned}$$

That is

$$\|\bar{u}^*\|_U^2 \leq \|\bar{u}\|_U^2$$

or

$$\int_{t_0}^{t_f} (\bar{u}^*(t), \bar{u}^*(t))_{\Re^m}\, dt \leq \int_{t_0}^{t_f} (\bar{u}(t), \bar{u}(t))_{\Re^m}\, dt$$

Further, if $LL^* = W(t_0, t_f)$ is invertible, we get

$$\begin{aligned} \|\bar{u}^*\|^2 = \|L^*\bar{\alpha}\|^2 &= (L^*\bar{\alpha}, L^*\bar{\alpha}) \\ &= (LL^*\bar{\alpha}, \bar{\alpha}) = (W(t_0, t_f)\bar{\alpha}, \bar{\alpha}) \\ &= (\bar{x}_f - \phi(t_f, t)x_0,\ W^{-1}(t_0, t_f)\,[\bar{x}_f - \phi(t_f, t)\bar{x}_0]) \end{aligned}$$

∎

Example 8.3.1 *Consider the following electrical network, given by Figure 8.3.1.*

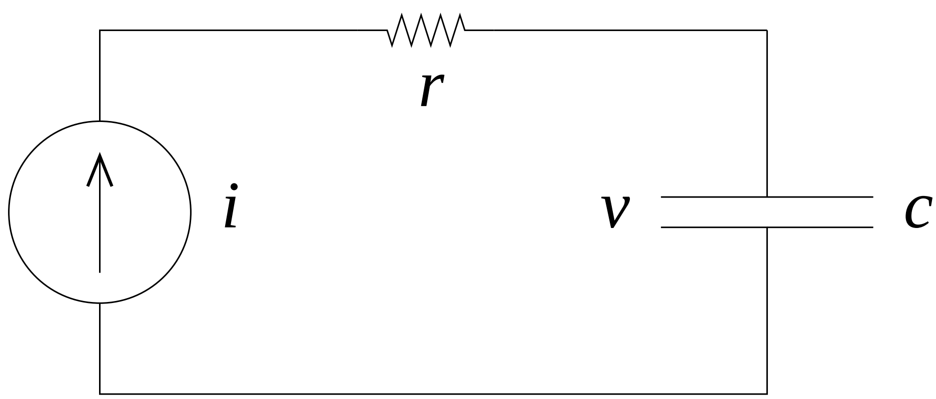

Figure 8.3.1: Electrical network

Its equation is given by

$$\frac{d}{dt}(cv(t)) = i(t), c = \text{ Capacitance} \tag{8.3.16}$$

i is the current out of the source. The energy dissipated in the resister in the time interval $[0, t_f]$ is $\int_0^{t_f} ri^2(t)dt$.

We wish to find the current $i(t)$ such that the voltage $v(0) = v_0$ and $v(t_f) = v_1$ and the energy disspipated is minimum.

Thus, we want to minimize $\int_0^{t_f} i^2(t)dt$ subject to the output voltage satisfying the equation Eq. (8.3.16) with the constraint that $v(0) = v_0$ and $v(t_f) = v_1$.

For this problem, the controllability Grammian $W(t_0, t_f) = \frac{t_f}{c^2}$ and hence the optimal current $i^*(t) = \frac{c}{t_f}(v_1 - v_0)$ and hence is constant on the interval $[0, t_f]$. For more on control theory refer Brockett [1] and Gopal[3].

8.4 Exercises

1. Show that the system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + B\bar{u}(t)$$

where $A(t) = \begin{bmatrix} 1 & \bar{e}^t \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

is controllable. Find the control $\bar{u}^*(t)$ with minimum L_2 - norm, which will steer the trajectory from the initial state $\bar{0}$ at $t = 0$ to the final state $(1, 1)$ at time $t_f = 1$.

2. Show that the linear system given by the pair $\{A, \bar{b}\}$ is controllable for all values of α_i's arising the definition of A.

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_1 & -\alpha_2 & -\alpha_3 & \cdots & -\alpha_n \end{bmatrix}, \quad \bar{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

3. Prove that the systems

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + B\bar{u}(t) \qquad (*)$$

and

$$\frac{d\bar{z}}{dt} = -A^\top(t)\bar{z}(t) \qquad (**)$$

$$\bar{w}(t) = B^\top \bar{z}(t)$$

are dual of each other. That is (*) is controllable iff (**) is observable.

4. Show that, if the time invariant system

$$\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + B\bar{u}(t)$$

is controllable, then there exists a matrix C such that

$$\frac{d\bar{x}}{dt} = (A + BC)\bar{x}(t) + \bar{b}_i v(t)$$

is also controllable where $\bar{b}_i$ is any nonzero column vector of B.

5. Consider the system

$$\frac{d\bar{x}}{dt} = g(t)\left[A\bar{x}(t) + B\bar{u}(t)\right]$$

with g continuous and bounded $(0 < \alpha \leq g(t) \leq \beta < \infty)$.

Show that if $C = \left[B, AB, \cdots, A^{n-1}B\right]$ has rank n, then the above system is controllable.

6. Obtain controllable conditions analogous to Theorem 8.1.1 for the following matrix differential equation

$$\frac{dX(t)}{dt} = A(t)X(t) + X(t)B(t) + C(t)U(t)D(t)$$

7. Show that the differential equation

$$\frac{d^2y}{dt^2} + u(t)\frac{dy}{dt} + y(t) = 0$$

is controllable (in the sense that given any $y_0, \dot{y}_0$ and $y_1, \dot{y}_1$, there exists a control $u(t)$ such that $y(t_0) = y_0,\ \ y(t_f) = y_1,\ \ \dot{y}(t_0) = \dot{y}_0,\ \ \dot{y}(t_f) = \dot{y}_1$), provided $y_0^2 + \dot{y}_0^2$ and $y_1 + \dot{y}_1^2$ are nonzero.

8. Show that there exists an initial state for the adjoint system

$$\frac{d\bar{p}}{dt} = -A^{\top}(t)\bar{p}(t)$$

such that the control $\bar{u}(t)$, which minimizes the functional

$$\phi(\bar{u}) = \int_{t_0}^{t_f} (\bar{u}(t), \bar{u}(t))\, dt$$

for steering the initial state $\bar{x}_0$ to the final state $\bar{x}_f$, is given by

$$\bar{u}(t) = -B^{\top}(t)\bar{p}(t)$$

References

[1] Roger W. Brockett, *Finite Dimensional Linear Systems, John Wiley and Sons, Inc, 1970*

[2]M. Gopal, *Modern Control System Theory, Wiley Eastern Ltd., 1989*

[3] Mohan C. Joshi and Ramendra K. Bose, *Some topic in Nonlinear Functional Analysis, Halsted Press, 1985*

[4] R. E. Kalman, *Mathematical Description of Linear Dynamical Systems, SIAM J Control, Vol I, pp152-192, 1963*

[5] Mohan C. Joshi and Kannan M. Moudgalya, *Optimization -Theory and Practice, Narosa Publishing House, Pvt. Ltd., 2004*

Index