Lecture notes on Ordinary Differential Equations
Annual Foundation School, IIT Kanpur,
Dec.3-28, 2007.

by
S. Sivaji Ganesh
Dept. of Mathematics, IIT Bombay,
Mumbai-76.
e-mail: sivaji.ganesh@gmail.com

Plan of lectures

(1) First order equations: Variable-Separable Method.

(2) Existence and uniqueness of solutions to initial value problems.

(3) Continuation of solutions, Saturated solutions, and Maximal interval of existence.

(4) Continuous dependence on data, Global existence theorem.

(5) Linear systems, Fundamental pairs of solutions, Wronskian.

References

1. V.I. Arnold, *Ordinary differential equations*, translated by Silverman, (Printice-Hall of India, 1998).
2. E.A. Coddington, *Introduction to ordinary differential equations*, (Prentice-Hall of India, 1974).
3. P. Hartman, *Ordinary differential equations*, (Wiley, 1964).
4. M.W. Hirsh, S. Smale and R.L. Devaney, *Differential equations, dynamical systems & Chaos*, (Academic press, 2004).
5. L.C. Piccinini, G. Stampacchia and G. Vidossich, *Ordinary differential equations in* $\mathbb{R}^N$*: problems and methods*, (Springer-Verlag, 1984).
6. M.R.M. Rao, *Ordinary differential equations: theory and applications*, (Affiliated East-West, 1980).
7. D.A. Sanchez, *Ordinary differential equations: A brief eclectic tour*, (Mathematical association of America, 2002).
8. I.I. Vrabie, *Differential equations*, (World scientific, 2004).
9. W.Wolfgang, *Ordinary differential equations*, (Springer-Verlag, 1998).

Lecture-1
First order equations: Basic concepts

We introduce basic concepts of theory of ordinary differential equations. A scalar ODE will be given geometric interpretation and thereby try to gain a geometric understanding of solution structure of ODE whose vector field has some invariance. This understanding is then used to solve equations of variable-separable type.

1.1 Basic concepts

We want to translate the feeling of what should be or what is an Ordinary Differential Equation (ODE) into mathematical terms. Defining some object like ODE, for which we have some rough feeling, in English words is not really useful unless we know how to put it mathematically. As such we can start by saying "Let us look at the following differential equation ...". But since many books give this definition, let us also have one such. The reader is referred to Remark 1.2 for an example of an "ODE" that we really do not want to be an ODE.
Let us start with

Hypothesis

Let $\Omega \subseteq \mathbb{R}^{n+1}$ be a domain and $\mathbb{I} \subseteq \mathbb{R}$ be an interval. Let $F : \mathbb{I} \times \Omega \to \mathbb{R}$ be a function defined by $(x, z, z_1, \ldots z_n) \mapsto F(x, z, z_1, \ldots z_n)$ such that F is not a constant function in the variable z_n.

With this notation and hypothesis on F we define the basic object in our study, namely, an Ordinary differential equation.

Definition 1.1 (ODE) *Assume the above hypothesis. An ordinary differential equation of order n is defined by the relation*

$$F\left(x, y, y^{(1)}, y^{(2)}, \ldots y^{(n)}\right) = 0, \tag{1.1}$$

where $y^{(n)}$ stands for n^{th} derivative of unknown function $x \mapsto y(x)$ with respect to the independent variable x.

Remark 1.2

1. *As we are going to deal with only one independent variable through out this course, we use the terminology 'differential equation' in place of 'ordinary differential equation' at times. Also we use the abbreviation ODE which stands for Ordinary Differential Equation(s). Wherever convenient, we use the notation prime $'$ to denote a derivative w.r.t. independent variable x; for example, y' is used to denote $y^{(1)}$).*

2. *Note that the highest order of derivative of unknown function y appearing in the relation* (1.1) *is called the order of the ordinary differential equation. Look at the carefully framed hypothesis above that makes sure the appearance of n^{th} derivative of y in* (1.1).

3. *(*Arnold*) If we define an ODE as a relation between an unknown function and its derivates, then the following equation will also be an ODE.*

$$\frac{dy}{dx}(x) = y \circ y(x). \tag{1.2}$$

However, note that our Defintion 1.1 does not admit (1.2) *as an ODE. Also, we do not like to admit* (1.2) *as an ODE since it is a non-local relation due to the presence of non-local operator 'composition". On the other hand recall that derivative is a local operator in the sense that derivative of a function at a point, depends only on the values of the function in a neighbourhood of the point.*

Having defined an ODE, we are interested in its solutions. This brings us to the question of existence of solutions and finding out all the solutions. We make clear what we mean by a solution of an ODE.

Definition 1.3 (Solution of an ODE) *A real valued function ϕ is said to be a solution of ODE (1.1) if $\phi \in \mathcal{C}^n(\mathbb{I})$ and*

$$F\left(x, \phi(x), \phi^{(1)}(x), \phi^{(2)}(x), \dots \phi^{(n)}(x)\right) = 0, \qquad \forall x \in \mathbb{I}. \tag{1.3}$$

Remark 1.4

(1) *There is no guarantee that an equation such as (1.1) will have a solution.*

(i) *The equation defined by $F(x, y, y') = (y')^2 + y^2 + 1 = 0$ has no solution. Thus we cannot hope to have a general theory for equations of type* (1.1). *Note that F is a smooth function of its arguments.*

(ii) *The equation*

$$y' = \begin{cases} 1 & \text{if} \quad x \geq 0 \\ -1 & \text{if} \quad x < 0, \end{cases}$$

does not have a solution on any interval containing 0. *This follows from Darboux's theorem about derivative functions.*

(2) *To convince ourselves that we do not expect every ODE to have a solution, let us recall the situation with other types of equations involving Polynomials, Systems of linear equations, Implicit functions. In each of these cases, existence of solutions was proved under some conditions. Some of those results also characterised equations that have solution(s), for example, for systems of linear equations the characterisation was in terms of ranks of matrix defining the linear system and the corresponding augmented matrix.*

(3) *In the context of ODE, there are two basic existence theorems that hold for equations in a special form called* **normal** *form. We state them in Section 3.1.*

As observed in the last remark, we need to work with a less general class of ODE if we expect them to have solutions. One such class is called ODE in normal form and is defined below.

Hypothesis (H)

Let $\Omega \subseteq \mathbb{R}^n$ be a domain and $\mathbb{I} \subseteq \mathbb{R}$ be an interval. Let $f : \mathbb{I} \times \Omega \to \mathbb{R}$ be a continuous function defined by $(x, z, z_1, \dots z_{n-1}) \mapsto f(x, z, z_1, \dots z_{n-1})$.

Definition 1.5 (ODE in Normal form) *Assume Hypothesis (H) on f. An ordinary differential equation of order n is said to be in normal form if*

$$y^{(n)} = f\left(x, y, y^{(1)}, y^{(2)}, \dots y^{(n-1)}\right). \tag{1.4}$$

Definition 1.6 (Solution of ODE in Normal form) *A function $\phi \in C^n(\mathbb{I}_0)$ where $\mathbb{I}_0 \subseteq \mathbb{I}$ is a subinterval is called a solution of ODE* (1.4) *if for every $x \in \mathbb{I}_0$, the $(n+1)$-tuple $\left(x, \phi(x), \phi^{(1)}(x), \phi^{(2)}(x), \dots \phi^{(n-1)}(x)\right) \in \mathbb{I} \times \Omega$ and*

$$\phi^{(n)}(x) = f\left(x, \phi(x), \phi^{(1)}(x), \phi^{(2)}(x), \dots \phi^{(n-1)}(x)\right), \quad \forall x \in \mathbb{I}_0. \tag{1.5}$$

Remark 1.7

1. *Observe that we want equation* (1.5) *to be satisfied "for all* $x \in \mathbb{I}_0$*" instead of "for all* $x \in \mathbb{I}$*". Compare now with defintion of solution given before in Definition 1.3 which is more stringent. We modified the concept of solution, by not requiring that the equation be satisfied by the solution on entire interval* $\mathbb{I}$*, due to various examples of ODEs that we shall see later which have solutions only on a subinterval of* $\mathbb{I}$*. We dont want to miss them!! Note that the equation*

$$y' = \begin{cases} 1 & \textit{if} \quad y \geq 0 \\ -1 & \textit{if} \quad y < 0, \end{cases}$$

does not admit a solution defined on $\mathbb{R}$*. However it has solutions defined on intervals* $(0, \infty)$, $(-\infty, 0)$*. (Find them!)*

2. *Compare Definition 1.5 with Definition 1.1. See the item (ii) of Remark 1.2, observe that we did not need any special effort in formulating Hypothesis (H) to ensure that* n^{th} *derivative makes an appearance in the equation* (1.4).

Convention From now onwards an ODE in normal form will simply be called ODE for brevity.

Hypothesis $(\mathbf{H_S})$

Let $\Omega \subseteq \mathbb{R}^n$ be a domain and $\mathbb{I} \subseteq \mathbb{R}$ be an interval. Let $\mathbf{f} : \mathbb{I} \times \Omega \to \mathbb{R}^n$ be a continuous function defined by $(x, \mathbf{z}) \mapsto \mathbf{f}(x, \mathbf{z})$ where $\mathbf{z} = (z_1, \ldots z_n)$.

Definition 1.8 (System of ODEs) *Assume Hypothesis* $(\mathrm{H_S})$ *on* $\mathbf{f}$*. A first order system of* n *ordinary differential equations is given by*

$$\mathbf{y}' = \mathbf{f}\,(x, \mathbf{y})\,. \tag{1.6}$$

The notion of solution for above system is defined analogous to Definition 1.5. A result due to D'Alembert enables us to restrict a general study of any ODE in normal form to that of a first order system in the sense of the following lemma.

Lemma 1.9 (D'Alembert) *An* n^{th} *order ODE* (1.4) *is equivalent to a system of* n *first order ODEs.*

Proof :

Introducing a transformation $\mathbf{z} = (z_1, z_2, \ldots, z_n) := \left(y, y^{(1)}, y^{(2)}, \ldots y^{(n-1)}\right)$, we see that $\mathbf{z}$ satisfies the linear system

$$\mathbf{z}' = (z_2, \ldots, z_n, f(x, \mathbf{z})) \tag{1.7}$$

Equivalence of (1.4) and (1.7) means starting from a solution of either of these ODE we can produce a solution of the other. This is a simple calculation and is left as an exercise. ■

Note that the first order system for $\mathbf{z}$ consists of n equations. This n is the order of (1.4).

Exercise 1.10 *Define higher order systems of ordinary differential equations and define corresponding notion of its solution. Reduce the higher order system to a first order system.*

1.2 Geometric interpretation of a first order ODE and its solution

We now define some terminology that we use while giving a geometric meaning of an ODE given by

$$\frac{dy}{dx} = f\,(x, y)\,. \tag{1.8}$$

We recall that f is defined on a domain D in $\mathbb{R}^2$. In fact, $D = \mathbb{I} \times \mathbb{J}$ where $\mathbb{I}$, $\mathbb{J}$ are sub-intervals of $\mathbb{R}$.

Definition 1.11 (Line element) *A line element associated to a point $(x, y) \in D$ is a line passing through the point (x, y) with slope p. We use the triple (x, y, p) to denote a line element.*

Definition 1.12 (Direction field/ Vector field) *A direction field (sometimes called vector field) associated to the ODE* (1.8) *is collection of all line elements in the domain D where slope of the line element associated to the point (x, y) has slope equal to $f(x, y)$. In other words, a direction field is the collection*

$$\left\{ (x, y, f(x, y)) \,:\, (x, y) \in D \right\}.$$

Remark 1.13 (Interpretations)

1. *The ODE* (1.8) *can be thought of prescribing line elements in the domain D.*
2. *Solving an ODE can be geometrically interpreted as finding curves in D that fit the direction field prescribed by the ODE. A solution (say ϕ) of the ODE passing through a point $(x_0, y_0) \in D$ (i.e., $\phi(x_0) = y_0$) must satisfy $\phi'(x_0) = f(x_0, y_0)$. In other words,*

$$(x_0, y_0, \phi'(x_0)) = (x_0, y_0, f(x_0, y_0)).$$

3. *That is, the ODE prescribes the slope of the tangent to the graph of any solution (which is equal to $\phi'(x_0)$). This can be seen by looking at the graph of a solution.*
4. *Drawing direction field corresponding to a given ODE and fitting some curve to it will end up in finding a solution, at least, graphically. However note that it may be possible to fit more than one curve passing through some points in D, which is the case where there are more than one solution to ODE around those points. Thus this activity (of drawing and fitting curves) helps to get a rough idea of nature of solutions of ODE.*
5. *A big challenge is to draw direction field for a given ODE. One good starting point is to identify all the points in domain D at which line element has the same slope and it is easy to draw all these lines. These are called* isoclines*; the word means "leaning equally".*

Exercise 1.14 *Draw the direction field prescribed by ODEs where $f(x, y) = 1$, $f(x, y) = x$, $f(x, y) = y$, $f(x, y) = y^2$, $f(x, y) = 3y^{2/3}$ and fit solution curves to them.*

Finding a solution of an ODE passing through a point in D is known as Initial value problem. We address this in the next section.

1.3 Initial Value Problems

We consider an Initial Value Problem (also called Cauchy problem) for an ODE (1.4). It consists of solving (1.4) subject to what are called *Initial conditions.* The two basic theorems we are going to present are concerning an IVP for a first order ODE.

Definition 1.15 (Initial Value Problem for an ODE) *Let $x_0 \in \mathbb{I}$ and $(y_1, y_2, \ldots, y_n) \in \Omega$ be given. An Initial Value Problem (IVP) for an ODE in normal form is a relation satisfied by an unknown function y given by*

$$y^{(n)} = f\left(x, y, y^{(1)}, y^{(2)}, \ldots y^{(n-1)}\right), \qquad y^{(i)}(x_0) = y_i,\ i = 0, \ldots, (n-1). \tag{1.9}$$

Definition 1.16 (Solution of an IVP for an ODE) *A solution ϕ of ODE* (1.4) *(see Definition 1.6) is said to be a solution of IVP if $x_0 \in \mathbb{I}_0$ and*

$$\phi^{(i)}(x_0) = y_i,\ i = 0, \ldots, (n-1).$$

This solution is denoted by $\phi(.; x_0, y_0, y_1, \ldots, y_{n-1})$ to remember the IVP solved by ϕ.

Definition 1.17 (Local and Global solutions of an IVP) *Let ϕ be a solution of an IVP for ODE* (1.4) *according to Definition 1.16.*

1. *If $\mathbb{I}_0 \subset \mathbb{I}$, then ϕ is called a local solution of IVP.*
2. *If $\mathbb{I}_0 = \mathbb{I}$, then ϕ is called a global solution of IVP.*

Remark 1.18

1. *Note that in all our definitions of solutions, a solution always comes with its domain of definition. Sometimes it may be possible to extend the given solution to a bigger domain. We address this issue in the next lecture.*
2. *When $n = 1$, geometrically speaking, graph of solution of an IVP is a curve passing through the point (x_0, y_0).*

Exercise 1.19 *Define an IVP for a first order system. Reduce an IVP for an n^{th} order ODE to that of an equivalent first order system.*

Lecture-2
First order equations: Variable-Separable method

Hypothesis ($\mathbf{H_{VS}}$)

Let $\mathbb{I} \subseteq \mathbb{R}$ and $\mathbb{J} \subseteq \mathbb{R}$ be intervals. Let $g : \mathbb{I} \to \mathbb{R}$ and $h : \mathbb{J} \to \mathbb{R} \setminus \{0\}$ be continuous functions.

If domains of functions are not specified, then they are assumed to be their "natural domains".

We consider a first order ODE in **Variable-Separable form** given by

$$\frac{dy}{dx} = g(x)h(y) \tag{2.1}$$

We gain an understanding of the equation (2.1) and its solution in three simple steps.

2.1 Direction field independent of y

The equation (2.1) takes the form

$$\frac{dy}{dx} = g(x) \tag{2.2}$$

Geometric understanding: Observe that slope of line element associated to a point depends only on its x coordinate. Thus the direction field is defined on the strip $\mathbb{I} \times \mathbb{R}$ and is invariant under translation in the direction of $\mathbf{Y}$ axis. Therefore it is enough to draw line elements for points in the set $\mathbb{I} \times \{0\}$. This suggests that if we know a solution curve then its translates in the $\mathbf{Y}$ axis direction gives all solution curves.
Let $x_0 \in \mathbb{I}$ be fixed and let us define a primitive of g on $\mathbb{I}$ by

$$G(x) := \int_{x_0}^{x} g(s)\, ds, \tag{2.3}$$

with the understanding that, if $x < x_0$, define $G(x) := -\int_{x}^{x_0} g(s)\, ds$.

The function G is a solution of ODE (2.2) by fundamental theorem of integral calculus, since $G'(x) = g(x)$ on $\mathbb{I}$. Moreover, $G(x_0) = 0$.

A solution curve passing through an arbitrary point $(\xi, \eta) \in \mathbb{I} \times \mathbb{R}$ can be obtained from G and is given by

$$y(x; \xi, \eta) = G(x) + (\eta - G(\xi)), \tag{2.4}$$

which is of the form $G(x)+C$ where $C = \eta - G(\xi)$. Thus all solutions of ODE (2.2) are determined. Moreover, all the solutions are *global*.

Example 2.1 *Solve the ODE* (2.2) *with* g *given by (i).* $x^2 + \sin x$ *(ii).* $x\sqrt{1-x^2}$ *(iii).* $\sin x^2$ *(iv) Have you run out of patience? An explicit formula for the solution does not mean having it explicitly!*

2.2 Direction field independent of x

The equation (2.1) takes the form

$$\frac{dy}{dx} = h(y). \tag{2.5}$$

This type of equations are called “autonomous” since the RHS of the equation does not depend on the independent variable x.

Geometric understanding: Observe that slope of line element associated to a point depends only on its y coordinate. Thus the direction field is defined on the strip $\mathbb{R} \times \mathbb{J}$ and is invariant under translation in the direction of $\mathbf{X}$ axis. Therefore it is enough to draw line elements for points in the set $\{0\} \times \mathbb{I}$. This suggests that if we know a solution curve then its translates in the $\mathbf{X}$ axis direction gives solution curves. Note that this is one of the main features of autonomous equations.

Exercise 2.2 *Verify that x^3, $(x-c)^3)$ with $c \in \mathbb{R}$ arbitrary are solutions of the ODE $y' = 3y^{2/3}$. Verify that $\phi(x) \equiv 0$ is also a solution. However this ϕ can not be obtained by translating any of the other solutions. Does this contradict the observation preceding this exercise?*

Exercise 2.3 *Formulate the geometric observation “if we know a solution curve then its translates in the $\mathbf{X}$ axis direction gives solution curves” as a mathematical statement and prove it. Compare and contrast with a similar statement made in Step 1.*

This case is considerably different from that of Step 1 where solutions are defined on the entire interval $\mathbb{I}$. To expect the difficulties awaiting us in the present case, it is advisable to solve the following exercise.

Exercise 2.4 *Verify that $\phi(x) = \frac{1}{(x+c)}$ are solutions of $y' = -y^2$ on certain intervals. Graph the solutions for $c = 0, \pm 1, \pm 2$. Verify that $\phi(x) \equiv 0$ is also a solution on $\mathbb{R}$.*

Note from the last exercise that it has a constant solution (also called, rest/equilibrium point since it does not move!) defined globally on $\mathbb{R}$ and also non-constant solutions defined on only a subinterval of the real line that varies with c. There is no way we can get a non-zero solution from zero solution by translation. Thus it is a good time to review Exercise 2.3 and find out if there is any contradition.

Observation: If $\xi \in \mathbb{J}$ satisfies $h(\xi) = 0$, then ξ is a rest point and $\phi(x) = \xi$ is a solution of (2.5) for $x \in \mathbb{R}$.

Since the function h is continuous, it does not change sign. Therefore we may assume, without loss of generality (WLOG), that $h(y) > 0$ for $y \in \mathbb{J}$. There is no loss of generality because the other case, namely $h(y) < 0$ for $y \in \mathbb{J}$, can be disposed off in a similar manner.

Formally speaking, ODE (2.5) ‘may be’ written as

$$\frac{dx}{dy} = \frac{1}{h(y)}.$$

Thus all the conclusions of Step 1 can be translated to the current problem. But ‘inversion of roles of x and y’ needs to be justified!

Let $y_0 \in \mathbb{J}$ be fixed and let us define a primitive of $1/h$ on $\mathbb{J}$ by

$$H(y) := \int_{y_0}^{y} \frac{1}{h(s)}\, ds. \tag{2.6}$$

We record some properties of H below.

1. The function $H : \mathbb{J} \to \mathbb{R}$ is differentiable (follows from fundamental theorem of integral calculus).

2. Since $h > 0$ on $\mathbb{J}$, the function H is strictly monotonically increasing. Consequently, H is one-one.
3. The function $H : \mathbb{J} \to H(\mathbb{J})$ is invertible. By definition, $H^{-1} : H(\mathbb{J}) \to \mathbb{J}$ is onto, and is also differentiable.
4. A finer observation yields $H(\mathbb{J})$ is an interval containing 0.

We write

$$H \circ H^{-1}(x) = x$$

A particular solution

The function H gives rise to an implicit expression for a solution y of ODE (2.5) given by

$$H(y) = x \tag{2.7}$$

This assertion follows from the simple calculation

$$\frac{d}{dx}H(y) = \frac{dH}{dy}\frac{dy}{dx} = \frac{1}{h(y)}\frac{dy}{dx} = 1.$$

Note that $H(y_0) = 0$ and hence $x = 0$. This means that the graph of the solution $y(x) := H^{-1}(x)$ passes through the point $(0, y_0)$. Note that this solution is defined on the interval $H(\mathbb{J})$ which may not be equal to $\mathbb{R}$.

Exercise 2.5 *When will the equality $H(\mathbb{J}) = \mathbb{R}$ hold? Think about it. Try to find some examples of h for which the equality $H(\mathbb{J}) = \mathbb{R}$ holds.*

Some more solutions

Note that the function $H^{-1}(x - c)$ is also a solution on the interval $c + H(\mathbb{J})$ (verify). Moreover, *these are all* the solutions of ODE (2.5).

All solutions

Let z be any solution of ODE (2.5) defined on some interval $\mathbb{I}_z$. Then we have

$$\frac{dz}{dx} = z'(x) = h(z(x))$$

Let $x_0 \in \mathbb{I}_z$ and let $z(x_0) = z_0$. Since $h \neq 0$ on $\mathbb{J}$, integrating both sides from x_0 to x yields

$$\int_{x_0}^{x} \frac{z'(x)}{h(z(x))}\, dx = \int_{x_0}^{x} 1\, dx$$

which reduces to

$$\int_{x_0}^{x} \frac{z'(x)}{h(z(x))}\, dx = x - x_0$$

The last equality, in terms of H defined by equation (2.6), reduces to

$$H(z) - \int_{y_0}^{z_0} \frac{1}{h(s)}\, ds = x - x_0$$

$$i.e., \qquad H(z) - H(z_0) = x - x_0 \tag{2.8}$$

From the equation (2.8), we conclude that $x - x_0 + H(z_0) \in H(\mathbb{J})$.
Thus, for x belonging to the interval $x_0 - H(z_0) + H(\mathbb{J})$, we can write the solution as

$$z(x) = H^{-1}\big(x - x_0 + H(z_0)\big). \tag{2.9}$$

Note that a solution of ODE (2.5) are not global in general.

Exercise 2.6 *Give a precise statement of what we proved about solutions of ODE* (2.5).

Exercise 2.7 *Prove that a solution curve passes through every point of the domain* $\mathbb{R} \times \mathbb{J}$.

Exercise 2.8 *In view of what we have proved under the assumption of non-vanishing of* h, *revisit Exercise 2.4, Exercise 2.3 and comment.*

Exercise 2.9 *Solve the ODE* (2.5) *with* h *given by (i).* ky *(ii).* $\sqrt{|y|}$ *(iii).* $1+y^2$ *(iv).* $3y^{2/3}$ *(v).* y^4+4 *(vi).* $\frac{1}{y^4+4}$

2.3 General case

We justify the formal calculation used to obtain solutions of ODE (2.1) which we recall here for convenience.

$$\frac{dy}{dx} = g(x)h(y) \tag{2.10}$$

The formal calculation to solve (2.10) is

$$\int \frac{dy}{h(y)} = \int g(x)\,dx + C \tag{2.11}$$

We justify this formal calculation on a rigorous footing. We introduce two auxiliary ODEs:

$$\frac{dy}{du} = h(y) \quad \text{and} \quad \frac{du}{dx} = g(x). \tag{2.12}$$

Notation We refer to the first and second equations of (2.12) as equations (2.12a) and (2.12b) respectively. We did all the hard work in Steps 1-2 and we can easily deduce results concerning ODE (2.10) using the auxiliary equations (2.12) introduced above.
Claim Let $(x_1, y_1) \in \mathbb{I} \times \mathbb{J}$ be an arbitrary point. A solution y satisfying $y(x_1) = y_1$ is given by

$$y(x) = H^{-1}\big(G(x) - G(x_1) + H(y_1)\big). \tag{2.13}$$

PROOF :
Recalling from Step 2 that the function H^{-1} is onto $\mathbb{J}$, there exists a unique u_1 such that $H^{-1}(u_1) = y_1$. Now solution of (2.12a) satisfying $y(u_1) = y_1$ is given by (see, formula (2.9))

$$y(u) = H^{-1}\big(u - u_1 + H(y_1)\big) \tag{2.14}$$

and is defined on the interval $u_1 - H(y_1) + H(\mathbb{J})$.

Solution of (2.12b), defined on $\mathbb{I}$, satisfying $u(x_1) = u_1$ is given by (see, formula (2.4))

$$u(x; x_1, u_1) = G(x) + (u_1 - G(x_1)).G(x) \in G(x_1) - H(y_1) + H(\mathbb{J}) \tag{2.15}$$

Combining the two formulae (2.14)-(2.15), we get (2.13). This formula makes sense for $x \in \mathbb{I}$ such that $G(x) + (u_1 - G(x_1)) \in u_1 - H(y_1) + H(\mathbb{J})$. That is, $G(x) \in G(x_1) - H(y_1) + H(\mathbb{J})$.
Thus a general solution of (2.1) is given by

$$y(x) = H^{-1}\big(G(x) - c\big). \tag{2.16}$$

This ends the analysis of variable-separable equations modulo the following exercise. ■

Exercise 2.10 *Prove that the set* $\big\{x \in \mathbb{I} : G(x) \in G(x_1) - H(y_1) + H(\mathbb{J})\big\}$ *is non-empty and contains a non-trivial interval.*

Exercise 2.11 *Prove: Initial value problems for variable-separable equations under Hypothesis* $(\mathbf{H_{VS}})$ *have unique solutions. Why does it not contradict observations made regarding solutions of* $y' = 3y^{2/3}$ *satisfying* $y(0) = 0$*?*

Lecture-3
First order equations: Local existence & Uniqueness theory

We discuss the twin issues of existence and uniqueness for Initial value problems corresponding to first order systems of ODE. This discussion includes the case of scalar first order ODE and also general scalar ODE of higher order in view of Exercise 1.19. However certain properties like boundedness of solutions do not carry over under the equivalence of Exercise 1.19 which are used in the discussion of extensibility of local solutions to IVP.

We compliment the theory with examples from the class of first order scalar equations. We recall the basic setting of IVP for systems of ODE which is in force through our discussion. Later on only the additional hypotheses are mentioned if and when they are made.

> **Hypothesis** ($\mathbf{H_{IVPS}}$)
>
> Let $\Omega \subseteq \mathbb{R}^n$ be a domain and $\mathbb{I} \subseteq \mathbb{R}$ be an interval. Let $\mathbf{f} : \mathbb{I} \times \Omega \to \mathbb{R}^n$ be a continuous function defined by $(x, \mathbf{y}) \mapsto \mathbf{f}(x, \mathbf{y})$ where $\mathbf{y} = (y_1, \ldots y_n)$. Let $(x_0, \mathbf{y}_0) \in \mathbb{I} \times \Omega$ be an arbitrary point.

Definition 3.1 *Assume Hypothesis* ($\mathrm{H_{IVPS}}$) *on* $\mathbf{f}$*. An IVP for a first order system of n ordinary differential equations is given by*

$$\mathbf{y}' = \mathbf{f}\,(x, \mathbf{y})\,, \quad \mathbf{y}(x_0) = \mathbf{y}_0. \tag{3.1}$$

As we saw before (Exercise 2.4) we do not expect a solution to be defined *globally* on the entire interval $\mathbb{I}$. Recall Remark 1.7 in this context. This motivates the following definition of solution of an IVP for systems of ODE.

Definition 3.2 (Solution of an IVP for systems of ODE) *An n-tuple of functions* $\mathbf{u} = (u_1, \ldots u_n) \in \mathbf{C}^1(\mathbb{I}_0)$ *where* $\mathbb{I}_0 \subseteq \mathbb{I}$ *is a subinterval containing the point* $x_0 \in \mathbb{I}$ *is called a solution of IVP* (3.1) *if for every* $x \in \mathbb{I}_0$*, the* $(n+1)$*-tuple* $(x, u_1(x), u_2(x), \cdots, u_n(x)) \in \mathbb{I} \times \Omega$,

$$\mathbf{u}'(x) = \mathbf{f}\,(x, \mathbf{u}(x))\,, \forall x \in \mathbb{I}_0 \quad \textit{and} \quad \mathbf{u}(x_0) = \mathbf{y}_0. \tag{3.2}$$

We denote this solution by $\mathbf{u} = \mathbf{u}(x; \mathbf{f}, x_0, \mathbf{y}_0)$ *to remind us that the solution depends on* $\mathbf{f}, \mathbf{y}_0$ *and* $\mathbf{u}(x_0) = \mathbf{y}_0$.

The **boldface** notation is used to denote vector quantities and we drop boldface for scalar quantities.

Remark 3.3 *The IVP* (3.1) *involves an interval* $\mathbb{I}$*, a domain* Ω*, a continuous function* $\mathbf{f}$ *on* $\mathbb{I} \times \Omega$*,* $x_0 \in \mathbb{I}$*,* $\mathbf{y}_0 \in \Omega$*. Given* $\mathbb{I}$*,* $x_0 \in \mathbb{I}$ *and* Ω*, we may pose many IVPs by varying the data* $(\mathbf{f}, \mathbf{y}_0)$ *belonging to the set* $\mathbf{C}(\mathbb{I} \times \Omega) \times \Omega$*. .*

Basic questions

There are three basic questions associated to initial value problems. They are

(i) Given $(\mathbf{f}, \mathbf{y}_0) \in \mathbf{C}(\mathbb{I} \times \Omega) \times \Omega$, does the IVP (3.1) admit at least one solution?

(ii) Assuming that for a given $(\mathbf{f}, \mathbf{y}_0) \in \mathbf{C}(\mathbb{I} \times \Omega) \times \Omega$, the IVP (3.1) has a solution, is the solution unique?

(iii) Assuming that for each $(\mathbf{f}, \mathbf{y}_0) \in \mathbf{C}(\mathbb{I} \times \Omega) \times \Omega$ the IVP admits a unique solution $\mathbf{y}(x; \mathbf{f}, x_0, \mathbf{y}_0)$ on a common interval $\mathbb{I}_0$ containing x_0, what is the nature of the following function?

$$\mathcal{S} \,:\, \mathbf{C}(\mathbb{I} \times \Omega) \times \Omega \longrightarrow \mathbf{C}^1(\mathbb{I}_0) \tag{3.3}$$

defined by

$$(\mathbf{f}, \mathbf{y}_0) \longmapsto \mathbf{y}(x; \mathbf{f}, x_0, \mathbf{y}_0). \tag{3.4}$$

We address questions (i), (ii) and (iii) in Sections 3.1, 3.2 and 4.3 respectively. Note that we do not require a solution to IVP be defined on the entire interval $\mathbb{I}$ but only on a subinterval containing the point x_0 at which initial condition is prescribed. Thus it is interesting to find out if every solution can be extended to $\mathbb{I}$ and the possible obstructions for such an extension. We discuss this issue in the next lecture.

3.1 Existence of local solutions

There are two important results concerning existence of solutions for IVP (3.1). One of them is proved for any function $\mathbf{f}$ satisfying Hypothesis ($\mathrm{H_{IVPS}}$) and the second assumes *Lipschitz continuity* of the function $\mathbf{f}$ in addition. As we shall see in Section 3.2, this extra assumption on $\mathbf{f}$ gives rise not only to another proof of existence but also uniqueness of solutions.

Both proofs are based on equivalence of IVP (3.1) and an integral equation.

Lemma 3.4 *A continuous function* $\mathbf{y}$ *defined on an interval* $\mathbb{I}_0$ *containing the point* x_0 *is a solution of IVP* (3.1) *if and only if* $\mathbf{y}$ *satisfies the integral equation*

$$\mathbf{y}(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}(s))\, ds \quad \forall\, x \in \mathbb{I}_0. \tag{3.5}$$

PROOF :
If $\mathbf{y}$ is a solution of IVP (3.1), then by definition of solution we have

$$\mathbf{y}'(x) = \mathbf{f}(x, \mathbf{y}(x)). \tag{3.6}$$

Integrating the above equation from x_0 to x yields the integral equation (3.5).

On the other hand let $\mathbf{y}$ be a solution of integral equation (3.5). Observe that, due to continuity of the function $t \to \mathbf{y}(t)$, the function $t \to \mathbf{f}(t, \mathbf{y}(t))$ is continuous on $\mathbb{I}_0$. Thus RHS of (3.5) is a differentiable function w.r.to x by fundamental theorem of integral calculus and its derivative is given by the function $x \to \mathbf{f}(x, \mathbf{y}(x))$ which is a continuous function. Thus $\mathbf{y}$, being equal to a continuously differentiable function via equation (3.5), is also continuously differentiable. The function $\mathbf{y}$ is a solution of ODE (3.1) follows by differentiating the equation (3.5). Evaluating (3.5) at $x = x_0$ gives the initial condition $\mathbf{y}(x_0) = \mathbf{y}_0$. ■

Thus it is enough to prove the existence of a solution to the integral equation (3.5) for showing the existence of a solution to the IVP (3.1). Both the existence theorems that we are going to prove, besides making use of the above equivalence, are proved by an approximation procedure.

The following result asserts that joining two solution curves in the extended phase space $\mathbb{I} \times \Omega$ gives rise to another solution curve. This result is very useful in establishing existence of solutions to IVPs; one usually proves the existence of a solution to the right and left of the point x_0 at which initial condition is prescribed and then one gets a solution (which should be defined in an open interval containing x_0) by joining the right and left solutions at the point $(x_0, \mathbf{y}_0)$.

Lemma 3.5 (Concatenation of two solutions) *Assume Hypothesis* ($\mathrm{H_{IVPS}}$). *Let* $[a,\, b]$ *and* $[b,\, c]$ *be two subintervals of* $\mathbb{I}$. *Let* $\mathbf{u}$ *and* $\mathbf{w}$ *defined on intervals* $[a,\, b]$ *and* $[b,\, c]$ *respectively be solutions of IVP with initial data* $(a,\, \xi)$ *and* $(b,\, \mathbf{u}(b))$ *respectively. Then the concatenated function* $\mathbf{z}$ *defined on the interval* $[a,\, c]$ *by*

$$\mathbf{z}_(x) = \begin{cases} \mathbf{u}(x) & \textit{if} \quad x \in [a,\, b]. \\ \mathbf{w}(x) & \textit{if} \quad x \in (b,\, c]. \end{cases} \tag{3.7}$$

is a solution of IVP with initial data $(a,\, \xi)$.

PROOF :

It is easy to see that the function $\mathbf{z}$ is continuous on $[a,\, c]$. Therefore, by Lemma 3.4, it is enough to show that $\mathbf{z}$ satisfies the integral equation

$$\mathbf{z}(x) = \xi + \int_a^x \mathbf{f}(s, \mathbf{z}(s))\, ds \quad \forall\, x \in [a,\, c]. \tag{3.8}$$

Clearly the equation (3.8) is satisfied for $x \in [a,\, b]$, once again, by Lemma 3.4 since $\mathbf{u}$ solves IVP with initial data $(a,\, \xi)$ and $\mathbf{z}_(x) = \mathbf{u}(x)$ for $x \in [a,\, b]$. Thus it remains to prove (3.8) for $x \in (b,\, c]$. For $x \in (b,\, c]$, once again by Lemma 3.4, we get

$$\mathbf{z}(x) = \mathbf{w}(x) = \mathbf{u}(b) + \int_b^x \mathbf{f}(s, \mathbf{w}(s))\, ds = \mathbf{u}(b) + \int_b^x \mathbf{f}(s, \mathbf{z}(s))\, ds. \tag{3.9}$$

Since

$$\mathbf{u}(b) = \xi + \int_a^b \mathbf{f}(s, \mathbf{u}(s))\, ds = \xi + \int_a^b \mathbf{f}(s, \mathbf{z}(s))\, ds, \tag{3.10}$$

substituting for $\mathbf{u}(b)$ in (3.9) finishes the proof of lemma. ■

Rectangles

As $\mathbb{I} \times \Omega$ is an open set, we do not know if functions defined on this set are bounded; also we do not know the shape or size of Ω. If we are looking to solve IVP (3.1) (*i.e.,* to find a solution curve passing through the point $(x_0,\, \mathbf{y}_0)$), we must know how long a solution (if and when it exists) may live. In some sense this depends on size of a rectangle $R \subseteq \mathbb{I} \times \Omega$ centred at $(x_0,\, \mathbf{y}_0)$ defined by two positive real numbers a, b

$$R = \{x \,:\, |x - x_0| \leq a\} \times \{\mathbf{y} \,:\, \|\mathbf{y} - \mathbf{y}_0\| \leq b\} \tag{3.11}$$

Since $\mathbb{I} \times \Omega$ is an open set, we can find such an R (for some positive real numbers a, b) for each point $(x_0,\, \mathbf{y}_0) \in \mathbb{I} \times \Omega$. Let M be defined by $M = \sup_R \|\mathbf{f}(x, \mathbf{y}\|$.

Note that rectangle R in (3.11) is symmetric in the x space as well. However a solution may be defined on an interval that is not symmetric about x_0. Thus it looks restrictive to consider R as above. It is definitely the case when x_0 is very close to one of the end points of the interval $\mathbb{I}$. Indeed in results addressing the existence of solutions for IVP, separately on intervals left and right to x_0 consider rectangles $R^\dagger \subseteq \mathbb{I} \times \Omega$ of the form

$$R^\dagger = [x_0,\, x_0 + a] \times \{\mathbf{y} \,:\, \|\mathbf{y} - \mathbf{y}_0\| \leq b\} \tag{3.12}$$

3.1.1 Existence theorem of Peano

Theorem 3.6 (Peano) *Assume Hypothesis* $(\mathrm{H_{IVPS}})$. *Then the IVP* (3.1) *has at least one solution on the interval* $|x - x_0| \leq \delta$ *where* $\delta = \min\{a,\, \frac{b}{M}\}$. ■

3.1.2 Cauchy-Lipschitz-Picard existence theorem

From real analysis, we know that continuity of a function at a point is a local concept (as it involves values of the function in a neighbourhood of the point at which continuity of function is in question). We talk about uniform continuity of a function with respect to a domain. Similarly we can define Lipschitz continuity at a point and on a domain of a function defined on subsets of $\mathbb{R}^n$. For ODE purposes we need functions of $(n+1)$ variables and Lipschitz continuity w.r.t. the last n variables. Thus we straight away define concept of Lipschitz continuity for such functions.

Let $R \subseteq \mathbb{I} \times \Omega$ be a rectangle centred at $(x_0,\, \mathbf{y}_0)$ defined by two positive real numbers a, b (see equation (3.11)).

Definition 3.7 (Lipschitz continuity) *A function* $\mathbf{f}$ *is said to be Lipschitz continuous on a rectangle* R *with respect to the variable* y *if there exists a* $K > 0$ *such that*

$$\|\mathbf{f}(x,\mathbf{y}_1) - \mathbf{f}(x,\mathbf{y}_1)\| \leq K\|\mathbf{y}_1 - \mathbf{y}_2\| \quad \forall (x,\mathbf{y}_1),\, (x,\mathbf{y}_2) \in R. \tag{3.13}$$

Exercise 3.8 *1. Let* $n = 1$ *and* f *be differentiable w.r.t. the variable* y *with a continuous derivative defined on* $\mathbb{I}\times\Omega$*. Show that* f *is Lipschitz continuous on any rectangle* $R \subset \mathbb{I}\times\Omega$*.*

2. *If* f *is Lipschitz continuous on every rectangle* $R \subset \mathbb{I}\times\Omega$*, is* f *differentiable w.r.t. the variable* y*?*

3. *Prove that the function* h *defined by* $h(y) = y^{2/3}$ *on* $[0,\infty)$ *is not Lipschitz continuous on any interval containing* 0*.*

4. *Prove that the function* $f(x,y) = y^2$ *defined on doamin* $\mathbb{R}\times\mathbb{R}$ *is not Lipschitz continuous. (this gives yet another reason to define Lipschitz continuity on rectangles)*

We now state the existence theorem and the method of proof is different from that of Peano theorem and yields a bilateral interval containing x_0 on which existence of a solution is asserted.

Theorem 3.9 (Cauchy-Lipschitz-Picard) *Assume Hypothesis* $(\mathrm{H_{IVPS}})$*. Let* $\mathbf{f}$ *be Lipschitz continuous with respect to the variable* $\mathbf{y}$ *on* R*. Then the IVP* (3.1) *has at least one solution on the interval* $\mathbb{J}$ $|x - x_0| \leq \delta$ *where* $\delta = \min\{a, \frac{b}{M}\}$*.*

PROOF :

Step 1: Equivalent integral equation We recall (3.5), which is equivalent to the given IVP below.

$$\mathbf{y}(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s,\mathbf{y}(s))\,ds \quad \forall\, x \in \mathbb{I}. \tag{3.14}$$

By the equivalence of above integral equation with the IVP, it is enough to prove that the integral equation has a solution. This proof is accomplished by constructing, what are known as Picard approximations, a sequence of functions that converges to a solution of the integral equation (3.14).

Step 2: Construction of Picard approximations

Define the first function $\mathbf{y}_0(x)$, for $x \in \mathbb{I}$, by

$$\mathbf{y}_0(x) := \mathbf{y}_0 \tag{3.15}$$

Define $\mathbf{y}_1(x)$, for $x \in \mathbb{I}$, by

$$\mathbf{y}_1(x) := \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s,\mathbf{y}_0(s))\,ds. \tag{3.16}$$

Note that the function $\mathbf{y}_1(x)$ is well-defined for $x \in \mathbb{I}$. However, when we try to define the next member of the sequence, $\mathbf{y}_2(x)$, for $x \in \mathbb{I}$, by

$$\mathbf{y}_2(x) := \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s,\mathbf{y}_1(s))\,ds, \tag{3.17}$$

caution needs to exercised. This is because, we do not know about the values that the function $\mathbf{y}_1(x)$ assumes for $x \in \mathbb{I}$, there is no reason that those values are inside Ω. However, it happens that for $x \in \mathbb{J}$, where the interval $\mathbb{J}$ is as in the statement of the theorem, the expression on RHS of (3.17) which defined function $\mathbf{y}_2(x)$ is meaningful, and hence the function $\mathbf{y}_2(x)$ is well-defined for $x \in \mathbb{J}$. By restricting to the interval $\mathbb{J}$, we can prove that the following sequence of functions is well-defined: Define for $k \geq 1$, for $x \in \mathbb{J}$,

$$\mathbf{y}_k(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s,\mathbf{y}_{k-1}(s))\,ds. \tag{3.18}$$

Proving the well-definedness of Picard approximations is left as an exercise, by induction. In fact, the graphs of each Picard approximant lies inside the rectangle R (see statement of our theorem). That is,

$$\|\mathbf{y}_k(x) - \mathbf{y}_0\| \leq b, \quad \forall\, x \in [x_0 - \delta,\, x_0 + \delta]. \tag{3.19}$$

The proof is immediate from

$$\mathbf{y}_k(x) - \mathbf{y}_0 = \int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}_{k-1}(s))\, ds, \quad \forall\, x \in \mathbb{J}. \tag{3.20}$$

$$\text{Therefore,} \quad \|\mathbf{y}_k(x) - \mathbf{y}_0\| \leq M|x - x_0|, \quad \forall\, x \in \mathbb{J}. \tag{3.21}$$

and for $x \in \mathbb{J}$, we have $|x - x_0| \leq \delta$.

Step 3: Convergence of Picard approximations

We prove the uniform convergence of sequence of Picard approximations $\mathbf{y}_k$ on the interval $\mathbb{J}$, by proving that this sequence corresponds to the partial sums of a uniformly convergent series of functions, and the series is given by

$$\mathbf{y}_0 + \sum_{l=0}^{\infty} \big[\mathbf{y}_{l+1}(x) - \mathbf{y}_l(x)\big], \tag{3.22}$$

Note that the sequence $\mathbf{y}_{k+1}$ corresponds to partial sums of series (3.22). That is,

$$\mathbf{y}_{k+1}(x) = \mathbf{y}_0 + \sum_{l=0}^{k} \big[\mathbf{y}_{l+1}(x) - \mathbf{y}_l(x)\big]. \tag{3.23}$$

Step 3A: Unifrom convergence of series (3.22) on $x \in \mathbb{J}$

We are going to compare series (3.22) with a convergence series of real numbers, uniformly in $x \in \mathbb{J}$, and thereby proving uniform convergence of the series. From the expression

$$\mathbf{y}_{l+1}(x) - \mathbf{y}_l(x) = \int_{x_0}^{x} \big\{\mathbf{f}(s, \mathbf{y}_l(s)) - \mathbf{f}(s, \mathbf{y}_{l-1}(s))\big\}\, ds, \quad \forall\, x \in \mathbb{J}, \tag{3.24}$$

we can prove by induction the estimate (this is left an exercise):

$$\|\mathbf{y}_{l+1}(x) - \mathbf{y}_l(x)\| \leq ML^l \frac{|x - x_0|^{l+1}}{(l+1)!} \leq \frac{M}{L} L^{l+1} \frac{\delta^{l+1}}{(l+1)!} \tag{3.25}$$

We conclude that the series (3.22), and hence the sequence of Picard iterates, converge uniformly on $\mathbb{J}$. This is because the above estimate (3.25) says that general term of series (3.22) is uniformly smaller than that of a convergent series, namely, for the function $e^{\delta L}$ times a constant.

Let $\mathbf{y}(x)$ denote the uniform limit of the sequence of Picard iterates $\mathbf{y}_k(x)$ on $\mathbb{J}$.

Step 4: The limit function $\mathbf{y}(x)$ solves integral equation (3.14)

We want to take limit as $k \to \infty$ in

$$\mathbf{y}_k(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}_{k-1}(s))\, ds. \tag{3.26}$$

Taking limit on LHS of (3.26) is trivial. Therefore, for $x \in \mathbb{J}$, if we prove that

$$\int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}_{k-1}(s))\, ds \longrightarrow \int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}(s))\, ds, \tag{3.27}$$

then we would obtain, for $x \in \mathbb{J}$,

$$\mathbf{y}(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}(s))\, ds, \tag{3.28}$$

and this finishes the proof. Therefore, it remains to prove (3.27). Let us estimate, for $x \in \mathbb{J}$, the quantity

$$\int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}_{k-1}(s))\, ds - \int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}(s))\, ds = \int_{x_0}^{x} \big\{ \mathbf{f}(s, \mathbf{y}_{k-1}(s)) - \mathbf{f}(s, \mathbf{y}(s)) \big\}\, ds \tag{3.29}$$

Since the graphs of $\mathbf{y}_k$ lie inside rectangle R, so does the graph of $\mathbf{y}$. This is because rectangle R is closed. Now we use that the vector field $\mathbf{f}$ is Lipschitz continuous (with Lipschitz constant L) in the variable $\mathbf{y}$ on R, we get

$$\left\| \int_{x_0}^{x} \big\{ \mathbf{f}(s, \mathbf{y}_{k-1}(s)) - \mathbf{f}(s, \mathbf{y}(s)) \big\}\, ds \right\| \leq \left| \int_{x_0}^{x} \big\| \big\{ \mathbf{f}(s, \mathbf{y}_{k-1}(s)) - \mathbf{f}(s, \mathbf{y}(s)) \big\} \big\|\, ds \right| \tag{3.30}$$

$$\leq L \left| \int_{x_0}^{x} \| \mathbf{y}_{k-1}(s) - \mathbf{y}(s) \|\, ds \right| \leq L |x - x_0| \sup_{\mathbb{J}} \| \mathbf{y}_{k-1}(x) - \mathbf{y}(x) \| \tag{3.31}$$

$$\leq L\delta \sup_{\mathbb{J}} \| \mathbf{y}_{k-1}(x) - \mathbf{y}(x) \|. \tag{3.32}$$

This estimate finishes the proof of (3.27), since $\mathbf{y}(x)$ is the uniform limit of the sequence of Picard iterates $\mathbf{y}_k(x)$ on $\mathbb{J}$, and hence for sufficiently large k, the quantity $\sup_{\mathbb{J}} \| \mathbf{y}_{k-1}(x) - \mathbf{y}(x) \|$ can be made arbitrarily small. ■

Some comments on existence theorems

Remark 3.10 (i) *Note that the interval of existence depends only on the bound M and not on the specific function.*

(ii) *If $\mathbf{g}$ is any Lipschitz continuous function (with respect to the variable $\mathbf{y}$ on R) in a α neighbourhood of $\mathbf{f}$ and ζ is any vector in an β neighbourhood of $\mathbf{y}_0$, then solution to IVP with data $(\mathbf{g}, \zeta)$ exists on the interval $\delta = \min\{a, \frac{b-\beta}{M+\alpha}\}$. Note that this interval depends on data $(\mathbf{g}, \zeta)$ only in terms of its distance to $(\mathbf{f}, \mathbf{y}_0)$.*

Exercise 3.11 *Prove that Picard's iterates need not converge if the vector field does is not locally Lipschitz. Compute successive approximations for the IVP*

$$y' = 2x - x\sqrt{y_+}, \quad y(0) = 0, \quad \textit{with} \quad y_+ = \max\{y, 0\},$$

and show that they do not converge, and also show that IVP has a unique solution. (Hint: $y_{2n} = 0$, $y_{2n+1} = x^2$, $n \in \mathbb{N}$).

3.2 Uniqueness

Recalling the definition of a solution, we note that if $\mathbf{u}$ solves IVP (3.1) on an interval $\mathbb{I}_0$ then $\mathbf{w} \stackrel{\text{def}}{=} \mathbf{u}|_{\mathbb{I}_1}$ is also a solution to the same IVP where $\mathbb{I}_1$ is any subinterval of $\mathbb{I}_0$ containing the point x_0. In principle we do not want to consider the latter as a different solution. Thus we are led to define a concept of equivalence of solutions of an IVP that does not distinguish $\mathbf{w}$ from $\mathbf{u}$ near the point x_0. Roughly speaking, two solutions of IVP are said to be equivalent if they agree on some interval containing x_0 (neighbourhood of x_0). This neighbourhood itself may depend on the given two solutions.

Definition 3.12 (local uniqueness) *An IVP is said to have local uniqueness property if for each $(x_0, \mathbf{y}_0) \in \mathbb{I} \times \Omega$ and for any two solutions $\mathbf{y}_1$ and $\mathbf{y}_2$ of IVP (3.1) defined on intervals $\mathbb{I}_1$ and $\mathbb{I}_2$ respectively, there exists an open interval $\mathbb{I}_{\delta lr} := (x_0 - \delta_l, x_0 + \delta_r)$ containing the point x_0 such that $\mathbf{y}_1(x) = \mathbf{y}_2(x)$ for all $x \in \mathbb{I}_{\delta lr}$.*

Definition 3.13 (global uniqueness) *An IVP is said to have global uniqueness property if for each* $(x_0, \mathbf{y}_0) \in \mathbb{I} \times \Omega$ *and for any two solutions* $\mathbf{y}_1$ *and* $\mathbf{y}_2$ *of IVP* (3.1) *defined on intervals* $\mathbb{I}_1$ *and* $\mathbb{I}_2$ *respectively, the equality* $\mathbf{y}_1(x) = \mathbf{y}_2(x)$ *holds for all* $x \in \mathbb{I}_1 \cap \mathbb{I}_2$.

Remark 3.14

1. *It is easy to understand the presence of adjectives local and global in Definition 3.12 and Definition 3.12 respectively.*
2. *There is no loss of generality in assuming that the interval appearing in Definition 3.12, namely* $\mathbb{I}_{\delta lr}$, *is of the form* $\mathbb{J}_{\delta lr} := [x_0 - \delta_l,\, x_0 + \delta_r]$. *This is because in any open interval containing a point* x_0, *there is a closed interval containing the same point* x_0 *and vice versa.*

Though it may appear that local and global uniqueness properties are quite different from each other, indeed they are the same. This is the content of the next result.

Lemma 3.15 *The following are equivalent.*

1. *An IVP has local uniqueness property.*
2. *An IVP has global uniqueness property.*

Proof :

From definitions, clearly (2) $\Longrightarrow$ (1). We turn to the proof of (1) $\Longrightarrow$ (2).

Let $\mathbf{y}_1$ and $\mathbf{y}_2$ be two solutions of IVP (3.1) defined on intervals $\mathbb{I}_1$ and $\mathbb{I}_2$ respectively. We prove that $\mathbf{y}_1 = \mathbf{y}_2$ on the interval $\mathbb{I}_1 \cap \mathbb{I}_2$; we split its proof in to two parts. We first prove the equality to the right of x_0, *i.e.*, on the interval $[x_0, \sup(\mathbb{I}_1 \cap \mathbb{I}_2))$ and proving the equality to the left of x_0 (*i.e.*, on the interval $(\inf(\mathbb{I}_1 \cap \mathbb{I}_2),\, x_0]$) follows a canonically modified argument. Let us consider the following set

$$\mathcal{K}_r = \{t \in \mathbb{I}_1 \cap \mathbb{I}_2 \,:\, \mathbf{y}_1(x) = \mathbf{y}_2(x) \quad \forall x \in [x_0,\, t]\,\}\,. \tag{3.33}$$

The set $\mathcal{K}_r$ has the following properties:

1. The set $\mathcal{K}_r$ is non-empty. This follows by applying local uniqueness property of the IVP with initial data $(x_0, \mathbf{y}_0)$.
2. The equality $\sup \mathcal{K}_r = \sup(\mathbb{I}_1 \cap \mathbb{I}_2)$ holds.
   Proof :

   Note that infimum and supremum of an open interval equals the left and right end points of its closure (the closed interval) whenever they are finite. Observe that $\sup \mathcal{K}_r \leq \sup(\mathbb{I}_1 \cap \mathbb{I}_2)$ since $\mathcal{K}_r \subseteq \mathbb{I}_1 \cap \mathbb{I}_2$. Thus it is enough to prove that strict inequality can not hold. On the contrary, let us assume that $a_r := \sup \mathcal{K}_r < \sup(\mathbb{I}_1 \cap \mathbb{I}_2)$. This means that $a_r \in \mathbb{I}_1 \cap \mathbb{I}_2$ and hence $\mathbf{y}_1(a_r)$, $\mathbf{y}_2(a_r)$ are defined. Since a_r is the supremum (in particular, a limit point) of the set $\mathcal{K}_r$ on which $\mathbf{y}_1$ and $\mathbf{y}_2$ coincide, by continuity of functions $\mathbf{y}_1$ and $\mathbf{y}_2$ on $\mathbb{I}_1 \cap \mathbb{I}_2$, we get $\mathbf{y}_1(a_r) = \mathbf{y}_2(a_r)$.

   Thus we find that the functions $\mathbf{y}_1$ and $\mathbf{y}_2$ are still solutions of ODE on the interval $\mathbb{I}_1 \cap \mathbb{I}_2$, and also that $\mathbf{y}_1(a_r) = \mathbf{y}_2(a_r)$. Thus applying once again local uniquness property of IVP but with initial data $(a_r, \mathbf{y}_1(a_r))$, we conclude that $\mathbf{y}_1 = \mathbf{y}_2$ on an interval $\mathbb{J}_{\delta lr} := [a_r - \delta_l,\, a_r + \delta_r]$ (see Remark 3.14). Thus combining with arguments of previous paragraph we obtain the equality of functions $\mathbf{y}_1 = \mathbf{y}_2$ on the interval $[x_0,\, a_r + \delta_r]$. This means that $a_r + \delta_r \in \mathcal{K}_r$ and thus a_r is not an upper bound for $\mathcal{K}_r$. This contradiction to the definition of a_r finishes the proof of 2.

As mentioned at the beginning of the proof, similar statements to the left of x_0 follow. This finishes the proof of lemma. ∎

Remark 3.16 1. *By Lemma 3.15 we can use either of the two definitions Definition 3.12 and Definition 3.12 while dealing with questions of uniqueness. Henceforth we use the word uniqueness instead of using adjectives local or global since both of them are equivalent.*

2. *One may wonder then, why there is a need to define both local and global uniqueness properties. The reason is that it is easy to prove local uniqueness compared to proving global uniqueness and at the same retaining what we intuitively feel about uniqueness.*

Example 3.17 (Peano) *The initial value problem*

$$y' = 3\,y^{2/3}, \quad y(0) = 0. \tag{3.34}$$

has infinitely many solutions. ■

However there are sufficient conditions on $\mathbf{f}$ so that the corresponding IVP has a unique solution. One such condition is that of Lipschitz continuity w.r.t. variable $\mathbf{y}$.

Lemma 3.18 *Assume Hypothesis* ($\mathrm{H}_{\mathrm{IVPS}}$)*. If* $\mathbf{f}$ *is Lipschitz continuous on every rectangle* R *contained in* $\mathbb{I} \times \Omega$*, then we have global uniqueness.*

Proof :

Let $\mathbf{y}_1$ and $\mathbf{y}_2$ be two solutions of IVP (3.1) defined on intervals $\mathbb{I}_1$ and $\mathbb{I}_2$ respectively. By Lemma 3.4, we have

$$\mathbf{y}_i(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s, \mathbf{y}_i(s))\,ds, \quad \forall x \in \mathbb{I}_i \quad \text{and} \quad i = 1, 2. \tag{3.35}$$

Subtracting one equation from another we get

$$\mathbf{y}_1(x) - \mathbf{y}_2(x) = \int_{x_0}^{x} \big\{\mathbf{f}(s, \mathbf{y}_1(s)) - \mathbf{f}(s, \mathbf{y}_2(s))\big\}\,ds, \quad \forall x \in \mathbb{I}_1 \cap \mathbb{I}_2. \tag{3.36}$$

Applying norm on both sides yields, for $x \in \mathbb{I}_1 \cap \mathbb{I}_2$

$$\begin{aligned} \|\mathbf{y}_1(x) - \mathbf{y}_2(x)\| &= \left\| \int_{x_0}^{x} \big\{\mathbf{f}(s, \mathbf{y}_1(s)) - \mathbf{f}(s, \mathbf{y}_2(s))\big\}\,ds \right\| & (3.37) \\ &\le \int_{x_0}^{x} \big\|\mathbf{f}(s, \mathbf{y}_1(s)) - \mathbf{f}(s, \mathbf{y}_2(s))\big\|\,ds & (3.38) \end{aligned}$$

Choose δ such that $\|\mathbf{y}_1(s) - \mathbf{y}_0\| \le b$ and $\|\mathbf{y}_2(s) - \mathbf{y}_0\| \le b$, since we know that $\mathbf{f}$ is locally Lipschitz, it will be Lipschitz on the rectangle R with Lipschitz constant $L > 0$. As a consequence we get

$$\begin{aligned} \|\mathbf{y}_1(x) - \mathbf{y}_2(x)\| &\le L \sup_{|x - x_0| \le \delta} \|\mathbf{y}_1(x) - \mathbf{y}_2(x)\|\,|x - x_0| & (3.39) \\ &\le L\,\delta \sup_{|x - x_0| \le \delta} \|\mathbf{y}_1(x) - \mathbf{y}_2(x)\| & (3.40) \end{aligned}$$

It is possible to arrange δ such that $L\delta < 1$. From here we conclude that

$$\sup_{|x - x_0| \le \delta} \|\mathbf{y}_1(x) - \mathbf{y}_2(x)\| < \sup_{|x - x_0| \le \delta} \|\mathbf{y}_1(x) - \mathbf{y}_2(x)\| \tag{3.41}$$

Thus we conclude $\sup_{|x - x_0| \le \delta} \|\mathbf{y}_1(x) - \mathbf{y}_2(x)\| = 0$. This establishes local uniqueness and global uniqueness follows from their equivalence. ■

Example 3.19 *The IVP*

$$y' = \begin{cases} y \sin \frac{1}{y} & \text{if} \quad y \neq 0 \\ 0 & \text{if} \quad y = 0, \end{cases} \quad y(0) = 0.$$

has unique solution, despite the RHS not being Lipschitz continuous w.r.t. variable y *on any rectangle containing* $(0, 0)$*.*

Lecture-4
Saturated solutions & Maximal interval of existence

4.1 Continuation

We answer the following two questions in this section.

1. When can a given solution be continued?
2. If a solution can not be continued to a bigger interval, what prevents a continuation?

The existential results of Section 3.1 provide us with an interval containing the initial point on which solution for IVP exists and the length of the interval depends on the data of the problem as can be seen from expression for its length. However this does not rule out solutions of IVP defined on bigger intervals if not whole of $\mathbb{I}$. In this section we address the issue of extending a given solution to a bigger interval and the difficulties that arise in extending.

Intuitive idea for extension Take a local solution $\mathbf{u}$ defined on an interval $\mathbb{I}_0 = (x_1\, x_2)$ containing the point x_0. An intuitive idea to extend $\mathbf{u}$ is to take the value of $\mathbf{u}$ at the point $x = x_2$ and consider a new IVP for the same ODE by posing the initial condition prescribed at the point $x = x_2$ to be equal to $\mathbf{u}(x_2)$. Consider this IVP on a rectangle containing the point $(x_2, \mathbf{u}(x_2))$. Now apply any of the existence theorems of Section 3.1 and conclude the existence of solution on a bigger interval. Repete the previous steps and obtain a solution on the entire interval $\mathbb{I}$.

However note that this intuitive idea may fail for any one of the following reasons. They are (i). limit of function $\mathbf{u}$ does not exist as $x \to x_2$ (ii). limit in (i) may exist but there may not be any rectangle around the point $(x_2, \mathbf{u}(x_2))$ as required by existence theorems. In fact these are the principle difficulties in extending a solution to a bigger interval. In any case we show that for any solution there is a biggest interval beyond which it can not be extended as a solution of IVP and is called *maximal interval of existence* via Zorn's lemma.

We start with a few definitions.

Definition 4.1 (Continuation, Saturated solutions) *Let the function* $\mathbf{u}$ *defined on an interval* $\mathbb{I}_0$ *be a solution of IVP* (3.1). *Then*

1. *The solution* $\mathbf{u}$ *is called continuable at the right if there exists a solution of IVP* $\mathbf{w}$ *defined on interval* $\mathbb{J}_0$ *satisfying* $\sup \mathbb{I}_0 \leq \sup \mathbb{J}_0$ *and the equality* $\mathbf{u}(x) = \mathbf{w}(x)$ *holds for* $x \in \mathbb{I}_0 \cap \mathbb{J}_0$. *Any such* $\mathbf{w}$ *is called an extension of* $\mathbf{u}$. *Further if* $\sup \mathbb{I}_0 < \sup \mathbb{J}_0$, *then* $\mathbf{w}$ *is called a non-trivial extension of* $\mathbf{u}$. *The solution* $\mathbf{u}$ *is called saturated at the right if there is no non-trivial right-extension.*
2. *The solution* $\mathbf{u}$ *is called continuable at the left if there exists a solution of IVP* $\mathbf{z}$ *defined on interval* $\mathbb{K}_0$ *satisfying* $\inf \mathbb{K}_0 < \inf \mathbb{I}_0$ *and the equality* $\mathbf{u}(x) = \mathbf{z}(x)$ *holds for* $x \in \mathbb{I}_0 \cap \mathbb{K}_0$. *The solution* $\mathbf{u}$ *is called saturated at the left if it is not continuable at the left.*
3. *The solution* $\mathbf{u}$ *is called global at the right if* $\mathbb{I}_0 \supseteq \{x \in \mathbb{I} : x \geq x_0\}$. *Similarly,* $\mathbf{u}$ *is called global at the left if* $\mathbb{I}_0 \supseteq \{x \in \mathbb{I} : x \leq x_0\}$.

Remark 4.2 *If the solution* $\mathbf{u}$ *is continuable at the right, then by concactenating the solutions* $\mathbf{u}$ *and* $\mathbf{w}$ *we obtain a solution of IVP on a bigger interval* $(\inf \mathbb{I}_0, \sup \mathbb{J}_0)$, *where* $\mathbf{w}$ *defined on* $\mathbb{J}_0$ *is a right extension of* $\mathbf{u}$. *A similar statement holds if* $\mathbf{u}$ *is continuable at the left.*

Let us define the notion of a right (left) solution to the IVP (3.1).

Definition 4.3 *A function* $\mathbf{u}$ *defined on an interval* $[x_0, b)$ *(respectively, on* $(a, x_0]$*) is said to be a right solution (respectively, a left solution) if* $\mathbf{u}$ *is a solution of ODE* $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$ *on* (x_0, b) *(respectively, on* (a, x_0)*) and* $\mathbf{u}(x_0) = \mathbf{y}_0$.

Note that if $\mathbf{u}$ is a solution of IVP (3.1) defined on an interval (a, b), then $\mathbf{u}$ restricted to $[x_0, b)$ (respectively, to $(a, x_0]$) is a right solution (respectively, a left solution) to IVP.

For right and left solutions, the notions of continuation and saturated solutions become

Definition 4.4 (Continuation, Saturated solutions) *Let the function* $\mathbf{u}$ *defined on an interval* $[x_0, b)$ *be a right solution of IVP and let* $\mathbf{v}$ *defined on an interval* $(a, x_0]$ *be a left solution of IVP* (3.1)*. Then*

1. *The solution* $\mathbf{u}$ *is called continuable at the right if there exists a right solution of IVP* $\mathbf{w}$ *defined on interval* $[x_0, d)$ *satisfying* $b < d$ *and the equality* $\mathbf{u}(x) = \mathbf{w}(x)$ *holds for* $x \in [x_0, b)$*. Any such* $\mathbf{w}$ *is called a right extension of* $\mathbf{u}$*. The right solution* $\mathbf{u}$ *is called saturated at the right if it is not continuable at the right.*

2. *The solution* $\mathbf{v}$ *is called continuable at the left if there exists a solution of IVP* $\mathbf{z}$ *defined on interval* $(c, x_0]$ *satisfying* $c < a$ *and the equality* $\mathbf{u}(x) = \mathbf{z}(x)$ *holds for* $x \in (a, x_0]$*. The solution* $\mathbf{u}$ *is called saturated at the left if it is not continuable at the left.*

3. *The solution* $\mathbf{u}$ *is called global at the right if* $[x_0, b) = \{x \in \mathbb{I} \,:\, x \geq x_0\}$*. Similarly,* $\mathbf{u}$ *is called global at the left if* $(a, x_0] = \{x \in \mathbb{I} \,:\, x \leq x_0\}$*.*

The rest of the discussion in this section is devoted to analysing "continuability at the right", "saturated at the right" as the analysis for the corresponding notions "at the left" is similar. We drop suffixing "at the right" from now on to save space to notions of continuability and saturation of a solution.

4.1.1 Characterisation of continuable solutions

Lemma 4.5 *Assume Hypothesis* $(\mathrm{H_{IVPS}})$*. Let* $\mathbf{u} : \mathbb{I}_0 \to \mathbb{R}^n$ *be a right solution of IVP* (3.1) *defined on the interval* $[x_0, d)$*. Then the following statements are equivalent.*

(1) *The solution* $\mathbf{u}$ *is continuable.*

(2) (i) $d < \sup \mathbb{I}$ *and there exists*

(ii) $\mathbf{y}^* = \lim_{x \to d_-} \mathbf{y}(x)$ *and* $\mathbf{y}^* \in \Omega$.

(3) *The graph of* $\mathbf{u}$ *i.e.,*

$$\mathit{graph}\, \mathbf{u} = \big\{(x, \mathbf{u}(x)) \,:\, x \in [x_0, d)\big\} \tag{4.1}$$

is contained in a compact subset of $\mathbb{I} \times \Omega$.

PROOF :

We prove $(1) \Longrightarrow (2) \Longrightarrow (3) \Longrightarrow (2) \Longrightarrow (1)$. The implication $(1) \Longrightarrow (2)$ is obvious.

Proof of $(2) \Longrightarrow (3)$

In view of (2), we can extend the function $\mathbf{u}$ to the interval $[x_0, d]$ and let us call this extended function $\tilde{\mathbf{u}}$. Note that the function $x \longmapsto (x, \tilde{\mathbf{u}}(x)$ is continuous on the interval $[x_0, d]$ and the image of $[x_0, d]$ under this map is graph of $\tilde{\mathbf{u}}$, denoted by $\mathit{graph}\, \tilde{\mathbf{u}}$, is compact. But $\mathit{graph}\, \mathbf{u} \subset \mathit{graph}\, \tilde{\mathbf{u}} \subset \mathbb{I} \times \Omega$. Thus (3) is proved.

Proof of $(3) \Longrightarrow (2)$

Assume that $\mathit{graph}\, \mathbf{u}$ is contained in a compact subset of $\mathbb{I} \times \Omega$. As a consequence, owing to continuity of the function $\mathbf{f}$ on $\mathbb{I} \times \Omega$, there exists $M > 0$ such that $\|\mathbf{f}(x, \mathbf{u}(x))\| < M$ for all $x \in [x_0, d)$. Also, since $\mathbb{I}$ is an open interval, necessarily $d < \sup \mathbb{I}$. We will now prove that the limit in (2)(ii) exists.

Since $\mathbf{u}$ is a solution of IVP (3.1), by Lemma 3.4, we have

$$\mathbf{u}(x) = \mathbf{y}_0 + \int_{x_0}^{x} \mathbf{f}(s, \mathbf{u}(s))\, ds \quad \forall\, x \in [x_0, d). \tag{4.2}$$

Thus for $\xi, \eta \in [x_0,\, d)$, we get

$$\|\mathbf{u}(\xi) - \mathbf{u}(\eta)\| \leq \int_{\xi}^{\eta} \|\mathbf{f}(s, \mathbf{u}(s))\|\, ds \leq M|\xi - \eta|. \tag{4.3}$$

Thus $\mathbf{u}$ satisfies the hypothesis of Cauchy test on the existence of finite limit at d. Indeed, the inequality (4.3) says that $\mathbf{u}$ is uniformly continuous on $[x_0,\, d)$ and hence limit of $\mathbf{u}(x)$ is finite as $x \to d$. This follows from a property of uniformly continuous functions, namely they map Cauchy sequences to Cauchy sequences. Let us denote the limit by y^*. In principle being a limit, $y^* \in \overline{\Omega}$. To complete the proof we need to show that $y^* \in \Omega$. This is a consequence of the hypothesis that graph of $\mathbf{u}$ is contained in a compact subset of $\mathbb{I} \times \Omega$ and the fact that $\mathbb{I}$ and Ω are open sets.

Proof of (2) $\Longrightarrow$ (1)

As we shall see, the implication (2) $\Longrightarrow$ (1) is a consequence of existence theorem for IVP (Theorem 3.6) and concactenation lemma (Lemma 3.5).

Let $\mathbf{w}$ be a solution to IVP corresponding to the initial data $(d, y^*) \in \mathbb{I} \times \Omega$ defined on an interval $(e,\, f)$ containing the point d. Let $\mathbf{w}|_{[d,\, f)}$ be the restriction of $\mathbf{w}$ to the interval $[d,\, f) \subseteq \mathbb{I}$. Let $\tilde{\mathbf{u}}$ be defined as the continuous extension of $\mathbf{u}$ to the interval $(c,\, d])$ which makes sense due to the existence of the limit in (2)(ii). Concactenating $\tilde{\mathbf{u}}$ and $\mathbf{w}|_{[d,\, f)}$ yields a solution of the original IVP (3.1) that is defined on the interval $(c,\, f)$ and $d < f$. ■

Remark 4.6 *The important message of the above result is that a solution can be extended to a bigger interval provided the solution curve remains "well within" the domain $\mathbb{I} \times \Omega$ i.e., its right end-point lies in Ω.*

4.1.2 Existence and Classification of saturated solutions

The following result is concerning the existence of saturted solutions for an IVP. Once again we study "saturated at the right" and corresponding results for "saturated at the left" can be obtained by similar arguments. Thus for this discussion we always consider a soluton as defined on interval of the form $[x_0, d)$

Theorem 4.7 (Existence of saturated solutions) *If $\mathbf{u}$ defined on an interval $[x_0, d)$ is a right solution of IVP* (3.1)*, then either $\mathbf{u}$ is saturated, or $\mathbf{u}$ can be continued up to a saturated one.*

PROOF :

If $\mathbf{u}$ is saturated, then there is nothing to prove. Therefore, we assume $\mathbf{u}$ is not saturated. By definition of saturatedness of a solution, $\mathbf{u}$ is continuable. Thus the set $\mathcal{S}$, defined below, is non-empty.

$$\mathcal{S} = \text{Set of all solutions of IVP (3.1) which extend}\, \mathbf{u}. \tag{4.4}$$

We define a relation $\preceq$ on the set $\mathcal{S}$, called *partial order* as follows. For $\mathbf{w}, \mathbf{z} \in \mathcal{S}$ defined on intervals $[x_0,\, d_{\mathbf{w}})$ and $[x_0,\, d_{\mathbf{z}})$ respectively, we say that $\mathbf{w} \preceq \mathbf{z}$ if $\mathbf{z}$ is a continuation of $\mathbf{w}$.
Roughly speaking, if we take the largest (w.r.t. order $\preceq$) element of $\mathcal{S}$ then by it can not be further continued. To implement this idea, we need to apply *Zorn's lemma.* Zorn's lemma is equivalent to axiom of choice (see the book on Topology by JL Kelley for more) and helps in asserting existence of "maximal elements" provided the totally ordered subsets of $\mathcal{S}$ have an upper bound (upper bound for a subset $\mathcal{T} \subseteq \mathcal{S}$ is an element $\mathbf{h} \in \mathcal{S}$ such that $\mathbf{w} \preceq \mathbf{h}$ for all $\mathbf{w} \in \mathcal{S}$).

Exercise 4.8 *Show that the relation $\preceq$ defines a partial order on the set $\mathcal{S}$. Prove that each totally ordered subset of $\mathcal{S}$ has an upper bound.*

By Zorn's lemma, there exists a maximal element $\mathbf{q}$ in $\mathcal{S}$. Note that this maximal solution is saturated in view of the definition of $\preceq$ and maximality $\mathbf{q}$. ■

Remark 4.9 *Under the hypothesis of previous theorem, if a solution* $\mathbf{u}$ *of IVP* (3.1) *is continuable, then there may be more than one saturated solution extending* $\mathbf{u}$. *This possibility is due to non-uniqueness of solutions to IVP* (3.1). *The following exercise is concerned with this phenomenon. Further note that if solution* $\mathbf{u}$ *to IVP* (3.1) *is unique, then there will be a unique saturated solution extending it.*

Exercise 4.10 *Let* $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be defined by* $f(x, y) = 3\,y^{2/3}$. *Show that the solution* $y : [-1, 0] \to \mathbb{R}$ *defined by* $y(x) = 0$ *for all* $x \in [-1, 0]$ *of IVP satisfying the initial condition* $y(-1) = 0$ *has at least two saturated solutions extending it.*

The longevity of a solution (the extent to which a solution can be extended) is independent of the smoothness of vector field, where as existence and uniqueness of solutions are guaranteed for smooth vector fields.

Theorem 4.11 (Classification of saturated solutions) *Let* $\mathbf{u}$ *be a saturated right solution of IVP* (3.1), *and its domain of definition be the interval* $[x_0, d)$. *Then one of the following alternatives holds.*

(1) *The function* $\mathbf{u}$ *is unbounded on the interval* $[x_0, d)$.

(2) *The function* $\mathbf{u}$ *is bounded on the interval* $[x_0, d)$, *and* $\mathbf{u}$ *is global i.e.,* $d = \sup \mathbb{I}$.

(3) *The function* $\mathbf{u}$ *is bounded on the interval* $[x_0, d)$, *and* $\mathbf{u}$ *is not global i.e.,* $d < \sup \mathbb{I}$ *and each limit point of* $\mathbf{u}$ *as* $x \to d_-$ *lies on the boundary of* Ω.

PROOF :

If (1) is not true, then definitely (2) or (3) will hold. Therefore we assume that both (1) and (2) do not hold. Thus we assume that $\mathbf{u}$ is bounded on the interval $[x_0, d)$ and $d < \sup \mathbb{I}$. We need to show that each limit point of $\mathbf{u}$ as $x \to d_-$ lies on the boundary of Ω.

Our proof is by method of contradiction. We assume that there exists a limit point $\mathbf{u}^*$ of $\mathbf{u}$ as $x \to d_-$ in Ω. We are going to prove that $\lim\limits_{x \to d_-} \mathbf{u}(x)$ exists. Note that, once the limit exists it must be equal to $\mathbf{u}^*$ which is one of its limit points. Now applying Lemma 4.5, we infer that the solution $\mathbf{u}$ is continuable and thus contradicting the hypothesis that $\mathbf{u}$ is a saturated solution.

Thus it remains to prove that $\lim\limits_{x \to d_-} \mathbf{u}(x) = \mathbf{u}^*$ *i.e.,* $\|\mathbf{u}(x) - \mathbf{u}^*\|$ can be made arbitrarily small for x near $x = d$.

Since Ω is an open set and $\mathbf{u}^* \in \Omega$, there exists $r > 0$ such that $B[\mathbf{u}^*, r] \subset \Omega$. As a consequence, $B[\mathbf{u}^*, \epsilon] \subset \Omega$ for every $\epsilon < r$. Thus on the rectangle $R \subset \mathbb{I} \times \Omega$ defined by

$$R = [x_0,\, d] \times \{\mathbf{y} \,:\, \|\mathbf{y} - \mathbf{u}^*\| \leq r\}, \tag{4.5}$$

$\|\mathbf{f}(x, \mathbf{y})\| \leq M$ for some $M > 0$ since R is a compact set and $\mathbf{f}$ is continuous.
Since $\mathbf{u}^*$ is a limit point of $\mathbf{u}$ as $x \to d_-$, there exists a sequence (x_m) in $[x_0,\, d)$ such that $x_m \to d$ and $\mathbf{u}(x_m) \to \mathbf{u}^*$. As a consequence of definition of limit, we can find a $k \in \mathbb{N}$ such that

$$|x_k - d| < \min\left\{\frac{\epsilon}{2M}, \frac{\epsilon}{2}\right\} \quad \text{and} \quad \|\mathbf{u}(x_k) - \mathbf{u}^*\| < \min\left\{\frac{\epsilon}{2M}, \frac{\epsilon}{2}\right\}. \tag{4.6}$$

Claim: $\{(x,\, \mathbf{u}(x)) \,:\, x \in [x_k,\, d)\} \subset \mathbb{I} \times B[\mathbf{u}^*, \epsilon]$.

PROOF OF CLAIM: If the inclusion in the claim were false, then there would exist a point on the graph of $\mathbf{u}$ (on the interval $[x_k,\, d)$) lying outside the set $\mathbb{I} \times B[\mathbf{u}^*, \epsilon]$. Owing to the continuity of $\mathbf{u}$, the graph must meet the boundary of $B[\mathbf{u}^*, \epsilon]$. Let $x^* > x_k$ be the first instance at which the trajectory touches the boundary of $B[\mathbf{u}^*, \epsilon]$. That is, $\epsilon = \|\mathbf{u}(x^*) - \mathbf{u}^*\|$ and $\|\mathbf{u}(x) - \mathbf{u}^*\| < \epsilon$ for $x_k \leq x < x^*$. Thus

$$\epsilon = \|\mathbf{u}(x^*) - \mathbf{u}^*\| \;\leq\; \|\mathbf{u}(x^*) - \mathbf{u}(x_k)\| + \|\mathbf{u}(x_k) - \mathbf{u}^*\| < \int_{x_k}^{x^*} \|\mathbf{f}(s, \mathbf{u}(s))\|\, ds + \frac{\epsilon}{2} \tag{4.7}$$

$$< \; M(x^* - x_k) + \frac{\epsilon}{2} \;<\; M(d - x_k) + \frac{\epsilon}{2} \;<\; \epsilon \tag{4.8}$$

This contradiction finishes the proof of Claim.

Therefore, limit of $\mathbf{u}(x)$ as $x \to d$ exists. As noted at the beginning of this proof, it follows that $\mathbf{u}$ is continuable. This finishes the proof of the theorem. ■

Theorem 4.12 **(Classification of saturated solutions)** *Let $\mathbf{f} : \mathbb{I} \times \Omega \to \mathbb{R}^n$ be continuous on $\mathbb{I} \times \Omega$ and assume that it maps bounded subsets in $\mathbb{I} \times \Omega$ into bounded subsets in $\mathbb{R}^n$. Let $\mathbf{u} : [x_0, d) \to \mathbb{R}^n$ be a saturated right solution of IVP. Then one of the following alternatives holds.*

(1)′ The function $\mathbf{u}$ is unbounded on the interval $[x_0, d)$. If $d < \infty$ there exists $\lim_{x \to d_-} \|\mathbf{u}(x)\| = \infty$.

(2) The function $\mathbf{u}$ is bounded on the interval $[x_0, d)$, and $\mathbf{u}$ is global i.e., $d = \sup \mathbb{I}$.

(3)′ The function $\mathbf{u}$ is bounded on the interval $[x_0, d)$, and $\mathbf{u}$ is not global i.e., $d < \sup \mathbb{I}$ and limit of $\mathbf{u}$ as $x \to d_-$ exists and lies on the boundary of Ω. ■

Corollary 4.13 *Let $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be continuous. Let $\mathbf{u} : [x_0, d) \to \mathbb{R}^n$ be a saturated right solution of IVP. Use the previous exercise and conclude that one of the following two alternatives holds.*

(1) *The function $\mathbf{u}$ is global i.e., $d = \infty$.*

(2) *The function $\mathbf{u}$ is not global i.e., $d < \infty$ and $\lim_{x \to d_-} \|\mathbf{u}(x)\| = \infty$. This phenomenon is often referred to as $\mathbf{u}$* **blows up in finite time**. ■

4.2 Global Existence theorem

In this section we give some sufficient conditions under which every local solution of an IVP is global. One of them is the growth of $\mathbf{f}$ w.r.t. $\mathbf{y}$. If the growth is at most linear, then we have a global solution.

Theorem 4.14 *Let $\mathbf{f} : \mathbb{I} \times \mathbb{R}^n \to \mathbb{R}^n$ be continuous. Assume that there exist two continuous functions h, $k : \mathbb{I} \to \mathbb{R}_+$ (non-negative real-valued) such that*

$$\|\mathbf{f}(x, \mathbf{y})\| \le k(x)\|\mathbf{y}\| + h(x), \quad \forall\, (x, \mathbf{y}) \in \mathbb{I} \times \mathbb{R}^n. \tag{4.9}$$

Then for every initial data $(x_0, \mathbf{y}_0) \in \mathbb{I} \times \mathbb{R}^n$, IVP has at least one global solution.

Exercise 4.15 *Using Global existence theorem, prove that any IVP corresponding to a first order linear system has a unique global solution.*

4.3 Continuous dependence

In situations where a physical process is described (modelled) by an initial value problem for a system of ODEs, then it is desirable that any errors made in the measurement of either initial data or the vector field, do not influence the solution very much. In mathematical terms, this is known as continuous dependence of solution of an IVP, on the data present in the problem. In fact, the following result asserts that solution to an IVP has not only continuous dependence on initial data but also on the vector field $\mathbf{f}$.

Exercise 4.16 *Try to carefully formulate a mathematical statement on continuous dependence of solution of an IVP, on initial conditions and vector fields.*

An honest effort to answer the above exercise would make us understand the difficulty in formulating such a statement. In fact, many introductory books on ODEs do not address this subtle issue, and rather give a weak version of it without warning the reader about the principal difficulties. See, however, the books of Wolfgang [33], Piccinini *et. al.* [23]. We now state a result on continuous dependence, following Wolfgang [33].

Theorem 4.17 (Continuous dependece) *Let $\Omega \subseteq \mathbb{R}^n$ be a domain and $\mathbb{I} \subseteq \mathbb{R}$ be an interval containing the point x_0. Let $\mathbb{J}$ be a closed and bounded subinterval of $\mathbb{I}$, such that $x_0 \in \mathbb{J}$. Let $\mathbf{f} : \mathbb{I} \times \Omega \to \mathbb{R}^n$ be a continuous function. Let $\mathbf{y}(x; x_0, \mathbf{y}_0)$ be a solution on $\mathbb{J}$ of the initial value problem*

$$\mathbf{y}' = \mathbf{f}\,(x, \mathbf{y})\,, \quad \mathbf{y}(x_0) = \mathbf{y}_0. \tag{4.10}$$

Let S_α denote the α-neighbourhood of graph of $\mathbf{y}$, i.e.,

$$S_\alpha := \big\{(x, \mathbf{y}) \ : \ \|\mathbf{y} - \mathbf{y}(x; x_0, \mathbf{y}_0)\| \le \alpha,\, x \in \mathbb{J}\big\}. \tag{4.11}$$

Suppose that there exists an $\alpha > 0$ such that $\mathbf{f}$ satisfies Lipschitz condition w.r.t. variable $\mathbf{y}$ on S_α. Then the solution $\mathbf{y}(x; x_0, \mathbf{y}_0)$ depends continuously on the initial values and on the vector field $\mathbf{f}$.

That is: Given $\epsilon > 0$, there exists a $\delta > 0$ such that if $\mathbf{g}$ is continuous on S_α and the inequalities

$$\|\mathbf{g}\,(x, \mathbf{y}) - \mathbf{f}\,(x, \mathbf{y})\,\| \le \delta \text{ on } S_\alpha, \quad \|\zeta - \mathbf{y}_0\| \le \delta \tag{4.12}$$

are satisfied, then every solution $\mathbf{z}(x; x_0, \zeta)$ of the IVP

$$\mathbf{z}' = \mathbf{g}\,(x, \mathbf{z})\,, \quad \mathbf{z}(x_0) = \zeta. \tag{4.13}$$

exists on all of $\mathbb{J}$, and satisfies the inequality

$$\|\mathbf{z}(x; x_0, \zeta) - \mathbf{y}(x; x_0, \mathbf{y}_0)\| \le \epsilon, \quad x \in \mathbb{J}.\blacksquare \tag{4.14}$$

Remark 4.18 (i) *In words, the above theorem says : Any solution corresponding to an IVP where the vector field $\mathbf{g}$ near a Lipschitz continuous vector field $\mathbf{f}$ and initial data $(x_0,\, \zeta)$ near-by $(x_0,\, \mathbf{y}_0)$, stays near the unique solution of IVP with vector field $\mathbf{f}$ and initial data $(x_0,\, \mathbf{y}_0)$.*

(ii) *Note that, under the hypothesis of the theorem, any IVP with a vector field $\mathbf{g}$ which is only continuous, also has a solution defined on $\mathbb{J}$.*

(iii) *Note that the above theorem does not answer the third question we posed at the beginning of this chapter. The above theorem does not say anything about the function in* (3.3).

4.4 Well-posed problems

A mathematical problem is said to be well-posed (or, properly posed) if it has the **EUC property**.

(1) **E**xistence: The problem should have at least one solution.

(2) **U**niqueness: The problem has at most one solution.

(3) **C**ontinuous dependence: The solution depends continuously on the data that are present in the problem.

Theorem 4.19 *Initial value problem for an ODE $\mathbf{y}' = \mathbf{f}(x,\, \mathbf{y})$, where $\mathbf{f}$ is Lipschitz continuuos on a rectangle containining the initial data $(x,\, \mathbf{y}_0)$, is well-posed.* $\blacksquare$

Example 4.20 *"Solving $A\mathbf{x} = \mathbf{b}$ is not well-posed". Think why such a statement could be true.*

Lecture-5
2×2 Linear systems & Fundamental pairs of solutions

5.1 The Linear System

We consider the two-dimensional, linear homogeneous first order system of ODE

$$\begin{pmatrix} y_1' \\ \\ y_2' \end{pmatrix} = \begin{pmatrix} a(x) & b(x) \\ \\ c(x) & d(x) \end{pmatrix} \begin{pmatrix} y_1 \\ \\ y_2 \end{pmatrix} \tag{5.1}$$

The above system (5.1), with obvious notations, can be written as

$$\mathbf{y}' = \mathbf{A}(x)\mathbf{y} \tag{5.2}$$

An IVP is given by

$$\mathbf{y}' = \mathbf{A}(x)\mathbf{y}, \quad \mathbf{y}(x_0) = \begin{pmatrix} y_{10} \\ y_{20} \end{pmatrix} \tag{5.3}$$

1. Linear combinations of solutions of ODE (5.1) are also solutions of ODE (5.1).
2. The only solution of IVP (5.3) satisfying $\mathbf{y}(x_0) = 0$ is the trivial solution.
3. Given two solutions $\mathbf{w}_1$ and $\mathbf{w}_2$ of system (5.1), can the solution $\mathbf{y}$ of IVP (5.3) be obtained as a linear combination of $\mathbf{w}_1$ and $\mathbf{w}_2$? Supposing that the answer to this question is positive, we will find out the properties that the two solutions $\mathbf{w}_1$ and $\mathbf{w}_2$ must satisfy. Therefore assume that there exist C_1, $C_2 \in \mathbb{R}$ such that

   $$\mathbf{y}(x) = C_1\mathbf{w}_1(x) + C_2\mathbf{w}_2(x) \tag{5.4}$$

   The function $\mathbf{y}$ solves the IVP if and only if

   $$\begin{pmatrix} y_{10} \\ y_{20} \end{pmatrix} = \mathbf{y}(x_0) = C_1\mathbf{w}_1(x_0) + C_2\mathbf{w}_2(x_0) = C_1 \begin{pmatrix} w_{11}(x_0) \\ w_{12}(x_0) \end{pmatrix} + C_2 \begin{pmatrix} w_{21}(x_0) \\ w_{22}(x_0) \end{pmatrix} \tag{5.5}$$

   that is,

   $$\begin{pmatrix} y_{10} \\ y_{20} \end{pmatrix} = \begin{pmatrix} w_{11}(x_0) & w_{21}(x_0) \\ w_{12}(x_0) & w_{22}(x_0) \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \tag{5.6}$$

   The matrix equation (5.6) is solvable for arbitrary initial values $\begin{pmatrix} y_{10} \\ y_{20} \end{pmatrix}$ if and only if $\det \begin{pmatrix} w_{11}(x_0) & w_{21}(x_0) \\ w_{12}(x_0) & w_{22}(x_0) \end{pmatrix}$ is non-zero. This happens if and only if the columns $\begin{pmatrix} w_{11}(x_0) \\ w_{12}(x_0) \end{pmatrix}$ and $\begin{pmatrix} w_{21}(x_0) \\ w_{22}(x_0) \end{pmatrix}$ are linearly independent. In other words, the vectors $\mathbf{w}_1(x_0)$ and $\mathbf{w}_2(x_0)$ are linearly independent.

4. Suppose we want to solve arbitrary initial value problem, that is, $(x_0, \mathbf{y}_0)$ is arbitrary. This requires $\mathcal{W}(x; \mathbf{w}_1, \mathbf{w}_2) = \det \begin{pmatrix} w_{11}(x) & w_{21}(x) \\ w_{12}(x) & w_{22}(x) \end{pmatrix}$ to be non-zero for all $x \in (a, b)$. It seems that we are asking too much on solutions $\mathbf{w}_1$ and $\mathbf{w}_2$. However, it turns out that if $\det \begin{pmatrix} w_{11}(x) & w_{21}(x) \\ w_{12}(x) & w_{22}(x) \end{pmatrix}$ is non-zero at some x, then it will be non-zero for all $a < x < b$.

   Lemma 5.1 *The wronskian $\mathcal{W}(x; \mathbf{w}_1, \mathbf{w}_2)$ satisfies the ODE*

   $$\mathcal{W}'(x; \mathbf{w}_1, \mathbf{w}_2) = \mathit{trace}A(x)\ \mathcal{W}(x; \mathbf{w}_1, \mathbf{w}_2), \quad a < x < b. \tag{5.7}$$

   As a consequence, the function $\mathcal{W}(x; \mathbf{w}_1, \mathbf{w}_2)$ is either non-vanishing or identically zero on $a < x < b$.

PROOF :
From the definition of $\mathcal{W}(x;\mathbf{w}_1,\mathbf{w}_2)$, we get

$$\begin{aligned}\mathcal{W}'(x;\mathbf{w}_1,\mathbf{w}_2) &= (w_{11}(x)w_{22}(x)-w_{21}(x)w_{12}(x))' & (5.8)\\ &= w_{11}'(x)w_{22}(x)+w_{11}w_{22}'(x)-w_{21}'(x)w_{12}(x)-w_{21}(x)w_{12}'(x) & (5.9)\end{aligned}$$

Substituting the values of $w_{11}', w_{22}', w_{21}', w_{12}'$ from the ODE, we get

$$\begin{aligned}\mathcal{W}'(x;\mathbf{w}_1,\mathbf{w}_2) &= (a(x)+d(x))\mathcal{W}(x;\mathbf{w}_1,\mathbf{w}_2) & (5.10)\\ &= \text{trace}A(x)\ \mathcal{W}(x;\mathbf{w}_1,\mathbf{w}_2) & (5.11)\end{aligned}$$

Its general solution is given by

$$\mathcal{W}(x;\mathbf{w}_1,\mathbf{w}_2)=\mathcal{W}(x_0;\mathbf{w}_1,\mathbf{w}_2)\exp\left(\int_{x_0}^{x}\text{trace}A(s)\,ds\right) \tag{5.12}$$

This finishes the proof of lemma. ∎

Definition 5.2 *A pair of solutions* $\mathbf{w}_1$ *and* $\mathbf{w}_2$ *of system* (5.1) *is called a fundamental pair of solutions if any solution* $\mathbf{y}$ *of IVP* (5.3) *can be obtained as a linear combination of* $\mathbf{w}_1$ *and* $\mathbf{w}_2$*. That is, there exist* $C_1, C_2 \in \mathbb{R}$ *such that*

$$\mathbf{y}(x)=C_1\mathbf{w}_1(x)+C_2\mathbf{w}_2(x). \tag{5.13}$$

We answer the following questions concerning fundamental pairs of solutions.

1. Does there exist at least one fundamental pair of solutions?
2. Can we describe all fundamental pairs?

Answers

1. Existence of a fundamental pair follows by taking two solutions w_1 and w_2 of ODE satisfying IVP with initial data $\begin{pmatrix} w_{11}(x_0)\\ w_{12}(x_0)\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix} w_{21}(x_0)\\ w_{22}(x_0)\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$ respectively.
2. Let $P=\begin{pmatrix}p_{11} & p_{12}\\ p_{21} & p_{22}\end{pmatrix}$ be a non-singular matrix. Then $p_{11}\mathbf{w}_1+p_{21}\mathbf{w}_2,\ p_{12}\mathbf{w}_1+p_{22}\mathbf{w}_2$ is another fundamental pair. In fact this is the general form of a fundamental pair of solutions.

Special case of Second order linear equation

Let us now specialise to the second order homogeneous ODE (5.15) which is equivalent to the system

$$y''+p(x)y'+q(x)y=r(x). \tag{5.14}$$

$$y''+p(x)y'+q(x)y=0, \tag{5.15}$$

$$\begin{pmatrix} y_1'\\ \\ y_2'\end{pmatrix}=\begin{pmatrix}0 & 1\\ \\ -q(x) & -p(x)\end{pmatrix}\begin{pmatrix}y_1\\ \\ y_2\end{pmatrix} \tag{5.16}$$

Let us relate a fundamental pair of solutions of second order scalar equation with those of the equivalent system.

1. Let w_1, w_2 be a fundamental pair of solutions corresponding to ODE (5.15), then a fundamental pair of solutions for the system (5.16) is given by

$$\mathbf{z}_1(x)=\begin{pmatrix}w_1(x)\\ \\ w_1'(x)\end{pmatrix},\quad \mathbf{z}_2(x)=\begin{pmatrix}w_2(x)\\ \\ w_2'(x)\end{pmatrix} \tag{5.17}$$

Verification of the above assertion is easy.

2. Similarly, given a fundamental pair of solutions for the system $\mathbf{z}_1$, $\mathbf{z}_2$ one can produce a fundamental pair of solutions for the scalar equation by simply taking the first components of the vector valued functions $\mathbf{z}_1$, $\mathbf{z}_2$.

5.2 Nonhomogeneous equation

In this section we discuss on methods of solving nonhomogeneous equation. We follow the classical Variation of Parameters (VoP) method to find solutions of the nonhomogeneous equation starting from a general solution of the corresponding homogeneous equation. As already noted in the discussion of first order ODEs, VoP method involves computation of integrals that need lot of hard work even for simpler right hand side functions r. At least for second order equation with constant coefficients, as a second method, we prescribe the method of undetermined coefficients that works only for a restricted class of second order linear ODE with constant coefficients, and when the function r in (5.15) is of a special type involving only polynomials, exponentials.

5.2.1 Variation of Parameters method - Revisited.

We are going to exclusively work with the equivalent first order system of the second order nonhomogeneous ODE (5.14) given by

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -q(x) & -p(x) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ r(x) \end{pmatrix} \tag{5.18}$$

1. The starting point of VoP method is to have a general solution of homogeneous system. Let $\mathbf{z}_1(x) = \begin{pmatrix} w_1(x) \\ w_1'(x) \end{pmatrix}, \quad \mathbf{z}_2(x) = \begin{pmatrix} w_2(x) \\ w_2'(x) \end{pmatrix}$ be a fundamental pair of solutions corresponding to the equivalent system (5.16). Then a general solution $\mathbf{y}$ of homogeneous system is given by

$$\mathbf{y}(x) = C_1 \begin{pmatrix} w_1(x) \\ w_1'(x) \end{pmatrix} + C_2 \begin{pmatrix} w_2(x) \\ w_2'(x) \end{pmatrix} \tag{5.19}$$

   Re-writing the above equality,

$$\mathbf{y}(x) = \begin{pmatrix} w_1(x) & w_2(x) \\ w_1'(x) & w_2'(x) \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \Phi(x)\mathbf{C}(x) \tag{5.20}$$

   We assume that $\Phi(x_0)$ is identity matrix and there is no loss of generality in assuming this.

2. The second step is to allow $\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$ to be a function of x, and search for a particular solution $\mathbf{y}_p$ of the nonhomogeneous system in the form

$$\mathbf{y}_p(x) = \Phi(x)\mathbf{C}(x) \tag{5.21}$$

   Substitute the ansatz for $\mathbf{y}_p$ in (5.18) to get

$$\mathbf{y}_p'(x) = \Phi'(x)\mathbf{C}(x) + \Phi(x)\mathbf{C}'(x) = \begin{pmatrix} 0 & 1 \\ -q(x) & -p(x) \end{pmatrix} \mathbf{y}_p(x) + \begin{pmatrix} 0 \\ r(x) \end{pmatrix} \tag{5.22}$$

   It is easy to verify that the matrix valued function Φ satisfies

$$\Phi'(x) = \begin{pmatrix} 0 & 1 \\ -q(x) & -p(x) \end{pmatrix} \Phi(x) \tag{5.23}$$

In view of (5.23), the equation (5.22) yields

$$\mathbf{C}'(x) = (\Phi(x))^{-1} \begin{pmatrix} 0 \\ r(x) \end{pmatrix} \tag{5.24}$$

Integrating the above equation from x_0 to x, we get

$$\mathbf{C}(x) = \mathbf{C}(x_0) + \int_{x_0}^{x} (\Phi(s))^{-1} \begin{pmatrix} 0 \\ r(s) \end{pmatrix} ds \tag{5.25}$$

Substituting the value of $(\Phi(s))^{-1}$ in (5.25), we get

$$\mathbf{C}(x) = \mathbf{C}(x_0) + \int_{x_0}^{x} \frac{1}{\mathcal{W}(s; \mathbf{w}_1, \mathbf{w}_2)} \begin{pmatrix} w_2'(s) & -w_2(s) \\ -w_1'(s) & w_1(s) \end{pmatrix} \begin{pmatrix} 0 \\ r(s) \end{pmatrix} ds \tag{5.26}$$

That is,

$$\mathbf{C}(x) = \mathbf{C}(x_0) + \int_{x_0}^{x} \frac{1}{\mathcal{W}(s; \mathbf{w}_1, \mathbf{w}_2)} \begin{pmatrix} -w_2(s)r(s) \\ w_1(s)r(s) \end{pmatrix} ds \tag{5.27}$$

Thus, a particular solution is given by

$$\mathbf{y}_p(x) = \Phi(x)\mathbf{C}(x) = \Phi(x) \left[\mathbf{C}(x_0) + \int_{x_0}^{x} \frac{1}{\mathcal{W}(s; \mathbf{w}_1, \mathbf{w}_2)} \begin{pmatrix} -w_2(s)r(s) \\ w_1(s)r(s) \end{pmatrix} ds \right] \tag{5.28}$$

We will take $\mathbf{C}(x_0) = 0$ since the term corresponding to $\mathbf{C}(x_0)$ contributes only a solution to the homogeneous system and we are looking for a particular solution. Thus we get a solution of IVP for the nonhomogeneous system and is given by

$$\mathbf{y}(x) = \begin{pmatrix} w_1(x) & w_2(x) \\ w_1'(x) & w_2'(x) \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \Phi(x) \left[\int_{x_0}^{x} \frac{1}{\mathcal{W}(s; \mathbf{w}_1, \mathbf{w}_2)} \begin{pmatrix} -w_2(s)r(s) \\ w_1(s)r(s) \end{pmatrix} ds \right] \tag{5.29}$$

We substitute the value of $\Phi(x)$ in (5.29), and take only the first component of the resulting equation which will be a solution to the IVP for the corresponding second order equation and we denote it by y. Thus we have the following formula for y.

$$y(x) = \xi w_1(x) + \eta w_2(x) - w_1(x) \int_{x_0}^{x} \frac{w_2(s)r(s)}{\mathcal{W}(s; \mathbf{w}_1, \mathbf{w}_2)} ds + w_2(x) \int_{x_0}^{x} \frac{w_1(s)r(s)}{\mathcal{W}(s; \mathbf{w}_1, \mathbf{w}_2)} ds \tag{5.30}$$

Putting it in a more appealing format,

$$y(x) = \xi w_1(x) + \eta w_2(x) + \int_{x_0}^{x} \frac{\{w_2(x)w_1(s) - w_1(x)w_2(s)\}r(s)}{\mathcal{W}(s; \mathbf{w}_1, \mathbf{w}_2)} ds \tag{5.31}$$

We will come back to the above formula while dealing with what are called *boundary value problems*.

5.3 Solving linear planar systems with constant coefficients

Consider the system of ODE

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =: A \begin{pmatrix} x \\ y \end{pmatrix}. \tag{5.32}$$

5.3.1 Fundamental matrix

Definition 5.3 (Fundamental matrix) *A matrix valued function Φ whose columns are solutions of the system of ODE* (5.32) *is called a solution matrix. A solution matrix Φ is called a fundamental matrix if the columns of Φ form a fundamental pair of solutions for the system* (5.32). *A fundamental matrix Φ is called the standard fundamental matrix if $\Phi(0)$ is the identity matrix.*

Remark 5.4 *Since the columns of a solution matrixΦ are solutions of* (5.32), *the matrix valued function Φ satisfies the system of ODE*

$$\Phi' = A\Phi. \tag{5.33}$$

Exercise 5.5 *A solution matrix is a fundamental matrix if and only if its determinant is not zero.*

Exercise 5.6 *Prove that if Ψ is a fundamental matrix then ΨC is also a fundamental matrix for every constant invertible matrix C. Prove that all fundamental matrices occur this way.*

Computation of fundamental matrix

By definition of a fundamental matrix, finding a fundamental pair of solutions to system of ODE (5.32) is equivalent to finding a fundamental matrix. In view of Exercise 5.6, fundamental matrix is not unique but the standard fundamental matrix is unique.

(1) Observe that $e^{\lambda t}\begin{pmatrix} a \\ b \end{pmatrix}$ is a non-trivial solution of (5.32) if and only if

$$\begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad A\begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix}. \tag{5.34}$$

That is, λ is an eigenvalue and $\begin{pmatrix} a \\ b \end{pmatrix}$ is an eigenvector corresponding to λ.

(2) **Question** Is it possible to find a fundamental pair, both of which are of the form $e^{\lambda t}\begin{pmatrix} a \\ b \end{pmatrix}$?

Answer Supposing that $\phi_1(t) = e^{\lambda_1 t}\begin{pmatrix} a \\ b \end{pmatrix}$ and $\phi_2(t) = e^{\lambda_2 t}\begin{pmatrix} c \\ d \end{pmatrix}$ are two solutions of (5.32), ϕ_1, ϕ_2 form a fundamental pair if and only if

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} \neq 0, \tag{5.35}$$

since the above determinant is the Wronskian of ϕ_1, ϕ_2 at $t = 0$.
That is, the matrix A should have two linearly independent eigenvectors. Note that this is equivalent to saying that A is diagonalisable.

(3) **Question** What if the matrix A does not have two linearly independent eigenvectors? This can happen when A has only one eigenvalue of multiplicity two. Inspired by a similar situation in the context of constant coefficient second order linear ODE, we are tempted to try $\phi_1(t) = e^{\lambda_1 t}\begin{pmatrix} a \\ b \end{pmatrix}$ and $\phi_2(t) = te^{\lambda_1 t}\begin{pmatrix} a \\ b \end{pmatrix}$ as a fundamental pair. But note that ϕ_1, ϕ_2 does not form a fundamental pair since Wronskian at $t = 0$ will be zero, also note that ϕ_2 is not even a solution of the linear system (5.32). Nevertheless, we can find a solution having the form of ϕ_1. Therefore, we try a variant of above suggestion to find another solution that together ϕ_1 constitutes a fundamental pair. Let

$$\phi_2(t) = te^{\lambda_1 t}\begin{pmatrix} a \\ b \end{pmatrix} + e^{\lambda_1 t}\begin{pmatrix} c \\ d \end{pmatrix} \tag{5.36}$$

Then $\phi_2(t)$ solves the system (5.32) if and only if

$$(A - \lambda_1 I)\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}. \tag{5.37}$$

One can easily verify that $\begin{pmatrix} a \\ b \end{pmatrix}$, $\begin{pmatrix} c \\ d \end{pmatrix}$ are linearly independent. Thus, ϕ_1, ϕ_2 defined by

$$\phi_1(t) = e^{\lambda_1 t}\begin{pmatrix} a \\ b \end{pmatrix}, \quad \phi_2(t) = te^{\lambda_1 t}\begin{pmatrix} a \\ b \end{pmatrix} + e^{\lambda_1 t}\begin{pmatrix} c \\ d \end{pmatrix} \tag{5.38}$$

is a fundamental pair, where $\begin{pmatrix} a \\ b \end{pmatrix}$, $\begin{pmatrix} c \\ d \end{pmatrix}$ are related by the equation (5.37).

(4) In case the matrix A does not have real eigenvalues, then eigenvalues are complex conjugates of each other. In this case,

$$(\lambda, \mathbf{v}) \quad \text{is an eigen pair if and only if} \quad (\overline{\lambda}, \overline{\mathbf{v}}) \quad \text{is also an eigen pair for } A. \tag{5.39}$$

Denoting $\lambda = r + iq$ (note $q \neq 0$), $\mathbf{v} = \begin{pmatrix} \alpha + i\beta \\ \gamma + i\delta \end{pmatrix}$, define

$$\phi_1(t) = e^{rt}\begin{pmatrix} \alpha\cos qt - \beta\sin qt \\ \gamma\cos qt - \delta\sin qt \end{pmatrix}, \quad \phi_2(t) = e^{rt}\begin{pmatrix} \alpha\sin qt + \beta\cos qt \\ \gamma\sin qt + \delta\cos qt \end{pmatrix}. \tag{5.40}$$

Verifying that the pair of functions defined above constitute a real-valued fundamental pair of solutions is left as an exercise.

5.3.2 Matrix exponentials

Inspired by the formula for solutions to linear scalar equation $y' = ay$, given by $y(t) = ce^{at}$, where $y(0) = c$, we would like to say that the system (5.32) has solution given by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \exp(tA)\begin{pmatrix} a \\ b \end{pmatrix}, \quad \text{where} \quad \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x(0) \\ y(0) \end{pmatrix}. \tag{5.41}$$

First of all, we must give a meaning for the symbol $\exp(A)$ as a matrix, which is called the exponental of matrix B. Then we must verify that the proposed formula for solution, in (5.41), is indeed a solution of (5.32). We do this next.

Definition 5.7 (Exponential of a matrix) *If A is a 2×2 matrix, then exponential of A (denoted by e^A, or $\exp(A)$, is defined by*

$$e^A = I + \sum_{k=1}^{\infty} \frac{A^k}{k!} \tag{5.42}$$

We record below, without proof, some properties of matrix exponentials.

Lemma 5.8 *Let A be a 2×2 matrix.*

(1) *The series in (5.42) converges.*

(2) $e^0 = I$.

(3) $(e^A)^{-1} = e^{-A}$

(4) $\exp(A+B) = \exp(A)\exp(B)$ *if the matrices A and B satisfy $AB = BA$.*

(5) *If $J = P^{-1}AP$, then $\exp(J) = P^{-1}\exp(A)P$.*

(6) $\frac{d}{dt}e^{tA} = Ae^{tA}$. ■

Theorem 5.9 *If A is a 2×2 matrix, then $\Phi(t) = e^{tA}$ is a fundamental matrix of the system (5.32). If $(x(t), y(t))$ is a solution of the system (5.32) with $(x(t_0), y(t_0)) = (x_0, y_0)$, then $(x(t), y(t))$ is given by*

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{(t-t_0)A} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \blacksquare \tag{5.43}$$

Remark 5.10 *Thanks to the above theorem, we do not need to struggle to find a fundamental matrix as we did earlier. We need to to take just the exponential of the matrix tA and that would give very easily a fundamental matrix. But it is not as simple as it seems to be. In fact, summing up the series for exponential of a matrix is not easy at all, even for relatively simpler matrices. To realise this, solve the exercise following this remark.*

Exercise 5.11 *Find exponential matrix for*

$$A_1 = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix},\ A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},\ A_3 = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}. \tag{5.44}$$

Remark 5.12 *After solving the above exercise, one would realise that it is easier to find fundamental matrices by finding fundamental pairs of solutions instead of summing up a series! There is an alternate method to calculate exponential matrix for tA, via fundamental pairs, by observing that exponential matrix is nothing but the standard fundamental matrix. So, find a fundamental matrix Φ, then exponential matrix e^{tA} is given by*

$$e^{tA} = [\Phi(0)]^{-1}\Phi(t). \tag{5.45}$$

Topics for further study

So far we studied the basic theory of ODE, and I followed mainly two books [24, 32].

Topological properties of trajectories of solutions (the values that solutions of ODE take in the space Ω, called Phase space) are very interesting. Another object associated to ODE is finding the subsets of Ω which are invariant under flows defined by solutions of ODE. These studies go under the name of **Dynamical systems**.

Given that ODEs arise as models of physical phenomena, their understanding helps in understanding the systems that ODE model.

Another class of important ODEs are those with periodic coefficients.

One can look up books [1, 6, 14, 17, 20, 22] for further study.

Bibliography

[1] V.I. Arnold, *Ordinary differential equations*, translated by Silverman, (Printice-Hall of India, 1998).

[2] V.I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, (Springer, 1988).

[3] G. Birkhoff and G.-C. Rota, *Ordinary differential equations*, (John wiley, 1978).

[4] W.E. Boyce and R.C. Diprima, *Elementary differential equations and boundary value problems*, (Johy wiley, 2003).

[5] F. Brauer and J.A. Nohel, *The qualitative theory of Ordinary differential equations: an introduction*, (Dover, 1989).

[6] M. Braun, *Differential equations and their applications : an introduction to applied mathematics*, (Springer-Verlag, 1993).

[7] L. Cesari, *Asymptotic behavior and stability problems in ordinary differential equations*, 2nd ed., (Academic Press, 1963).

[8] E.A. Coddington, *Introduction to ordinary differential equations*, (Prentice-Hall of India, 1974).

[9] E.A. Coddington and N. Levinson, *Theory of ordinary differential equations*, (McGraw-Hill, 1955).

[10] J. Cronin, *Differential equations: Introduction and qualitative theory*, (Marcel Dekker, 1980).

[11] S.G. Deo and V. Raghavendra, *Ordinary differential equations and stability theory*, (Tata McGraw-Hill, 1980).

[12] M.S.P. Eastham, *Theory of ordinary differential equations*, (Van Nostrand Reinhold, 1970).

[13] N.P. Erugin , *Linear systems of ordinary differential equations with periodic and quasi-periodic coefficients translation edited by Richard Bellman*, (Academic Press, 1966).

[14] K.O. Friedrichs, *Lecture notes on advanced ordinary differential equations*, (Gordon & Breach, 1965).

[15] M. Golomb and M. Shanks, *Elements of ordinary differential equations*, (McGraw-Hill, 1965).

[16] P. Hartman, *Ordinary differential equations*, (Wiley, 1964).

[17] M.W. Hirsh, S. Smale and R.L. Devaney, *Differential equations, dynamical systems & Chaos*, (Academic press, 2004).

[18] W. Hurewicz, *Lectures on ordinary differential equations*, (Cambridge : MIT Press, 1958).

[19] E.L. Ince, *Ordinary differential equations*, (Dover, 1956).

[20] D.W. Jordan and P. Smith, *Nonlinear ordinary differential equations : an introduction to dynamical systems*, (Oxford, 1987).

[21] T. Myint-U, *Ordinary differential equations*, (North-Holland, 1978).

[22] L. Perko, *Differential equations and dynamical systems* (Springer-Verlag, 1991).

[23] L.C. Piccinini, G. Stampacchia and G. Vidossich, *Ordinary differential equations in* $\mathbb{R}^N$*: problems and methods*, (Springer-Verlag, 1984).

[24] L.S. Pontriagin , *Ordinary differential equations*, (Addison-Wesley, 1962).

[25] A.L. Rabenstein, *Introduction to ordinary differential equations*, (Academic Press, 1966).

[26] E.D. Rainville and P.E. Bedient, *A short course in differential equations*, (IBH, 1973; Macmillan, 1969).

[27] M.R.M. Rao, *Ordinary differential equations: theory and applications*, (Affiliated East-West, 1980).

[28] W.T. Reid, *Sturmian theory for ordinary differential equations*, (Springer-Verlag, 1980).

[29] S.L. Ross, *Introduction to ordinary differential equations*, (Waltham : Blaisdell Pub., 1966).

[30] D.A. Sanchez, *Ordinary differential equations and stability theory*, (Dover, 1979).

[31] D.A. Sanchez, *Ordinary differential equations: A brief eclectic tour*, (Mathematical association of America, 2002).

[32] I.I. Vrabie, *Differential equations*, (World scientific, 2004).

[33] W.Wolfgang, *Ordinary differential equations*, (Springer-Verlag, 1998).