Differential Equations

Kexing Ying

January 11, 2021

Contents

1 Introduction

While we have seen differential equations in year one, we have mostly focused on the different methods of solving specific differential equations. This cannot be expected for general differential equations and in this year, we will focus on existence and uniqueness of solutions to differential equations and develop qualitative tools to help us understand these solutions.

We recall that an algebraic equation is an equation of the form $f(x) = 0$ while a differential equation is an equation of the form $\dot{x} = f(x)$ for some function $f : \mathbb{R} \to \mathbb{R}$. That is, an algebraic equation has real numbers as solutions while an differential equation has functions as its solution.

As an example, let us consider the simple differentiable equation

$$\dot{x} = ax, \tag{1}$$

for some $a \in \mathbb{R}$. Then, a function $\lambda : I \to \mathbb{R}$ solves 1 if $\dot{\lambda} = a\lambda$ for all $t \in I$ where $I \subseteq \mathbb{R}$ is a interval. These types of differentiable equations occurs often in relation in growth and decay and one can easily see that the family of functions

$$\lambda_b : \mathbb{R} \to \mathbb{R} = t \mapsto be^{at}, \;\; b \in \mathbb{R},$$

are solutions to 1. Of course, we know this already, so an more interesting question would be whether or not this family contains all the solutions to 1. It turns out to be true, and to show this we will assume $\mu : I \to \mathbb{R}$ is a solution to $\dot{x} = ax$. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mu e^{-at}\right) = \dot{\mu}e^{-at} - a\mu e^{-at} = 0,$$

since $\dot{\mu} = a\mu$ and so, μe^{-at} is constant, i.e. there exists $b \in \mathbb{R}$ such that $\mu e^{-at} = b$ and hence,

$$\mu = be^{at}.$$

This demonstrates that all solutions to 1 are members of the aforementioned solution family and hence, we have found **all** of the solutions to 1.

With the above example, we see that rather than working with solutions that are in finite-dimensional vector spaces, our solution are in function spaces which are typically infinite-dimensional. This is studied in more detail in the next year's *functional analysis* course, and in general, infinite-dimensional spaces are more difficult to grasp. However, for the vast majority of materials in this course, a finite-dimensional thinking suffices while we will also cover some material from functional analysis to understand the differentiable equations as well.

1.1 Ordinary Differential Equations and Initial Value Problems

There are two types of differential equations – *autonomous differential equations* and *nonautonomous differentiable equations*. Autonomous differential equations are differentiable equations of the form $\dot{x} = f(x)$ such as equation 1 while nonautonomous differential equations are equations of the form $\dot{x} = f(t, x)$.

We note that this does not cover higher-order differential equations, but from last year, we recall that one may reduce a higher-order differential equations into a first-order differential equation in vector form and thus, the theories we develop within this course will also apply to higher-order differential equations.

Definition 1.1 (Ordinary Differential Equation)**.** Let $d \in \mathbb{N}$, $D \subseteq \mathbb{R} \times \mathbb{R}^d$ be open, and a function $f \to D \to \mathbb{R}^d$. Then, an equation of the form

$$\dot{x} = f(t, x)$$

is called a d-dimensional (first-order) ordinary differential equation.

A differentiable function $\lambda : I \to \mathbb{R}^d$ on some interval $I \subseteq \mathbb{R}$ is called a solution of the differential equation if and only if for all $(t, \lambda(t) \in D$, if $t \in I$ then,

$$\dot{\lambda}(t) = f(t, \lambda(t)).$$

We say that an ordinary differential equation is autonomous if f is independent of t and nonautonomous otherwise.

We will only consider ordinary differential equations (ODE) in this course while partial differential equations, that is differential equations which solutions are functions which depends on multiple variables are covered in the second year course **Partial Differential Equations in Action**.

Proposition 1 (Constant solutions to autonomous differential equations)**.** Let $D \subseteq \mathbb{R}^d$ be an open set and $f : D \to \mathbb{R}^d$ be a function where $d \in \mathbb{N}$. Then, there exists a constant solution $\lambda : \mathbb{R} \to \mathbb{R}^d : x \mapsto a$ to the autonomous differential

$$\dot{x} = f(x)$$

for some $a \in \mathbb{R}^d$ if and only if $f(a) = 0$.

Proof. ($\implies$) Suppose that $\lambda : I \to \mathbb{R}^d : x \mapsto a$ is a solution the $\dot{x} = f(x)$. Then

$$0 = \dot{\lambda}(t) = f(\lambda(t)) = f(a).$$

($\impliedby$) Suppose there exists some $a \in \mathbb{R}^d$ such that $f(a) = 0$, then verifying, we find $\lambda : \mathbb{R} \to \mathbb{R}^d : x \mapsto a$ is a solution to the differential equation. $\square$

This proposition allows us to find solutions to many autonomous ODEs as, indeed, if $f : D \to \mathbb{R}^d$ has a root $a \in \mathbb{R}^d$, the above proposition guarantees that $\lambda : \mathbb{R} \to \mathbb{R}^d : x \mapsto a$ is a solution to $\dot{x} = f(x)$.

Definition 1.2 (Initial Value Problem)**.** Let $d \in \mathbb{N}$, $D \subseteq \mathbb{R} \times \mathbb{R}^d$ be an open set, and $f : D \to \mathbb{R}^d$ be a function. The system of equations from combining the differential equation

$$\dot{x} = f(t, x),$$

with the initial condition

$$x(t_0) = x_0$$

where $(t_0, x_0) \in D$ is called an initial value problem.

A solution to the above initial value problem is a function $\lambda : I \to \mathbb{R}^d$ that is a solution to the differential equation $\dot{x} = f(t, x)$ and $\lambda(t_0) = x_0$.

While, previously, we have seen a differential equation which always has a solution. This, however, is not always the case.

Example 1. Consider the differential equation $\dot{x} = f(x)$, where

$$f(x) = \begin{cases} 1, & x < 0; \\ -1, & x \geq 0, \end{cases}$$

with the boundary condition $x(0) = 0$. As we will see on the problem sheet, this differential equation indeed does not have a solution.

Example 2. Consider the initial value problem $\dot{x} = f(x) = \sqrt{|x|}$ with the boundary condition $x(0) = 0$.

Since $f(0) = 0$ we have $x(t) = 0$ is a constant solution by proposition 1. Furthermore, by consider the function

$$\lambda_b(t) = \begin{cases} 0, & t \leq b; \\ \frac{1}{4}(t-b)^2, & t > b, \end{cases}$$

we find λ_b to also be a solution for any $b \in \mathbb{R}_0^+$. However, as for all $t \leq b$, $\lambda_b(t) = 0$, we see that x is not unique given values of x at t.

Before moving on, let us quickly recall the *separation of variables* procedure for solving differential equations. Suppose we are to solve a differential equation of the form,

$$\dot{x} = g(t)h(x),$$

with the boundary condition $x(t_0) = x_0$ where $g : I \to \mathbb{R}$ and $h : J \to \mathbb{R}$ are continuous functions. Then, we have

$$\int_{x_0}^{x} \frac{\mathrm{d}y}{h(y)} = \int_{t_0}^{t} g(s)\mathrm{d}s.$$

Thus, by evaluating the integral, we can express x in t. We see that this procedure does indeed provide us with a correct solution by simply applying FTC on both sides of the equation.

Example 3. Consider the initial value problem $\dot{x} = tx^2$ with the boundary condition $x(t_0) = x_0$ where $x_0 \neq 0$. By the separation of variables, we find

$$x = \frac{2x_0}{2 + x_0(t_0^2 - t^2)}.$$

However, we see that for this solution, x does not necessarily exist for all t. Indeed, if $t_0 = 0, x_0 = 1$, we find

$$x = \frac{2}{2 - t^2},$$

which does not have a solution for $t = \pm\sqrt{2}$, and so, the solution does not exist globally.

1.2 Visualisations

Visualisations are very important for differential equations as it provides us a mental image of how to think about differential equations. In general, there are two main methods to visualise differential equations:

- nonautonomous differential equations $\dot{x} = f(t, x)$ via the *solution portrait* in the extended phase space;
- autonomous differential equations $\dot{x} = f(x)$ via the *phase portrait* in the phase space.

Suppose we have the nonautonomous differential equations $\dot{x} = f(t, x)$ where $f : D \subseteq \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$, have a solution $\lambda : I \to \mathbb{R}^d$, i.e. for all $t \in I$, $\dot{\lambda}(t) = f(t, \lambda(t))$. Thus, we see that the vector $(1, f(t_0, \lambda(t_0))$ for some $t_0 \in I$ is tangential to the solution which passes through $(t_0, \lambda(t_0))$. This can be done for all $p \in \mathbb{R} \times \mathbb{R}^d$ and so, by drawing these vectors at a sufficient number of points, we can have a mental image of what the solution looks like. A plot of these vectors is refereed to as a **vector field**.

A **solution portrait** is given by a visualisation of several solution curves in the (t, x)-space, the so called *extended phase space*. This is called as such since the x-space is normally refereed as the phase space, and so, we are extending it by the time axis.

On the other hand, suppose we have the autonomous differential equation $\dot{x} = f(x)$ where $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^d$ has a solution $\lambda : I \to \mathbb{R}^d$. By considering that f is independent of t, we see that translating λ along t is also a valid solution. Indeed, at $x = \lambda(t_0)$, if λ' is another solution such that $\lambda'(t_1) = \lambda(t_0)$, then $\dot{\lambda}'(t_1) = \dot{\lambda}(t_0)$. This property is refereed to as *translation invariance* and all solutions of autonomous differential equations are translation invariant.

Proposition 2. let $\dot{x} = f(x)$ be an autonomous differential equation. Then, if $\lambda : I \to \mathbb{R}^d$ is a solution to this differential equation, so is

$$\mu : \bar{I} \to \mathbb{R}^d : t \mapsto \lambda(t + \tau),$$

where $\tau \in \mathbb{R}$, $\bar{I} := I + \tau$.

Proof. Follows straight away by chain rule. $\square$

Example 4. Consider the harmonic oscillator $\ddot{x} = -x$. By converting it into a first order differential equation, we have

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}.$$

By solving the system, we find

$$\lambda(t) = \begin{bmatrix} \cos(t) \sin(t) \\ -\sin(t) \cos(t) \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

is a solution. Indeed, this describes a oscillatory motion in the phase space where the position is circular.

With consideration with the example above, we see that we may project a solution portrait $(t, x) \mapsto (1, f(x))$ onto $x \mapsto f(x)$ resulting in a **phase portrait**.

2 Existence and Uniqueness

As we have seen, differential equations need not have unique solutions given an initial value. Indeed, we have also seen that it is not guaranteed to have a solution at all. We will in this chapter resolve the question on whether or not a solution exists by presenting a theory that guarantees existence and uniqueness for solutions to initial value problems.

2.1 Picard iterates

We observe that often times, to solve an differential equation, we need to reformulate it as an integral equation.

Proposition 3. Consider the initial value problem

$$\dot{x} = f(t, x); \quad x(t_0) = x_0,$$

where $f : D \subseteq \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is continuous. Let $\lambda : I \to \mathbb{R}^d$, λ solves the IVP if and only if λ is continuous and solves the integral equation[1]

$$\lambda(t) = x_0 + \int_{t_0}^{t} f(s, \lambda(s))\mathrm{d}s.$$

Proof. Follows from FTC. □

However, as f depends on λ, we in general cannot compute the integral on the right hand side. Nonetheless, this brings us closer to the formulation of the Picard iterates.

Proposition 4. Let f be a continuous function. Then by defining a_0 for some value and $a_{n+1} = f(a_n)$, if $(a_n)_{n=1}^{\infty}$ converges to some value a, then $a = f(a)$.

Proof. We see that if $(a_n)_{n=1}^{\infty}$ converges to some value a, $f(a_n) \to f(a)$ by sequential continuity. However, as $\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} a_{n+1} = a$, by uniqueness of limits on Hausdorff spaces, we have $a = f(a)$. □

With this proposition in mind, we may apply a similar iteration on the integral equation resulting in the Picard iterates.

Definition 2.1 (Picard Iterates)**.** Given an IVP, the Picard iterates is the sequence of functions $(\lambda_n : J \to \mathbb{R}^d)_{n=1}^{\infty}$ be defined such that $\lambda_0(t) = x_0$, and

$$\lambda_{n+1}(t) = x_0 + \int_{t_0}^{t} f(s, \lambda_n(s))\mathrm{d}s.$$

With connotation to the motivating proposition, we hope that (λ_n) converges to the solution λ_∞ in some notion. As we shall see, the Picard iterates needs to converge to λ_∞ uniformly in order for λ_∞ to be a solution.

Proposition 5. Given an IVP, if the Picard iterates $(\lambda_n)_{n=0}^{\infty}$ converges uniformly to λ_∞, then λ_∞ is a solution to the IVP.

[1]We recall that the integral of a vector valued function is simply the integral of the components.

Proof. Consider the following chain of equalities,

$$\lambda_\infty(t) = \lim_{n\to\infty} \lambda_{n+1}(t) = \lim_{n\to\infty} x_0 + \int_{t_0}^{t} f(s, \lambda_n(s))\mathrm{d}s$$
$$= x_0 + \int_{t_0}^{t} f(s, \lim_{n\to\infty} \lambda_n(s))\mathrm{d}s = x_0 + \int_{t_0}^{t} f(s, \lambda_\infty(s))\mathrm{d}s,$$

where the third equality is true as $\lambda_n \to \lambda_\infty$ uniformly (see first year analysis for proof). Thus, by proposition 3, λ_∞ solves the IVP. □

With this proposition in mind, it is *sometimes* possible to show that a particular Picard iterates converges uniformly towards some function resulting in a solution to the corresponding IVP.

2.2 Lipschitz Continuity

We recall from first year analysis the definition of Lipschitz continuity. We are interested in Lipschitz continuity since it is very helpful when showing the existence and uniqueness of solutions of IVPs. Indeed, by reformulating the Picard iterates as an operator P on the Banach space $\mathcal{C}^\circ(J, \mathbb{R}^d)$, we find that P is a contraction (i.e. P has Lipschitz constant < 1) if f satisfies certain Lipschitz condition, and hence, by Banach's fixed point theorem, has a fixed point.

We recall that a normed vector space is a vector space V equipped with a norm $\|\cdot\|$ such that the norm satisfies positive definiteness, absolute homogeneity and the triangle inequality. Indeed, we recall the chain of induced structure: Inner product space $\implies$ Normed vector space $\implies$ (Metric space $\implies$ Topological space) $\vee$ Vector space.

Definition 2.2 (Lipschitz Continuity)**.** Let $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ be two normed vector spaces and suppose $X \subset V$ and $Y \subseteq W$, then $f : X \to Y$ is Lipschitz continuous if there exists some $K > 0$ (a Lipschitz constant) such that for all $x, y \in X$,

$$\|f(x) - f(y)\|_W \leq K\|x - y\|_V.$$

We recall the sufficient conditions for which a function is Lipschitz continuous.

Proposition 6. Let $f : I \to \mathbb{R}$ be differentiable with bounded derivative on some interval I, then f is Lipschitz continuous.

Proof. By MVT, for all $x, y \in I$, there exists some $c \in (x, y)$ such that

$$f(x) - f(y) = f'(c)(x - y).$$

So by taking absolute value on both sides, and by choosing the Lipschitz constant as the supremum of $|f'|$ we have found some $K := \sup_{c\in I} \|f'(c)\|$, so

$$|f(x) - f(y)| \leq K\|x - y\|.$$

□

We see that, if f is continuously differentiable, f' is continuous on the compact set I, and so f' is uniformly continuous, and so is bounded. Thus, if $f : I \to \mathbb{R}$ is continuously differentiable, then it is Lipschitz continuous.

For higher dimensions, we require the mean value inequality for higher dimensions and so an similar result is achieved. To look at the mean value theorem for higher dimensions, let us first introduce the operator norm for linear maps.

Definition 2.3 (Operator Norm)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, then the operator norm of f is

$$\|f\|_{\text{op}} := \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|f(x)\|}{\|x\|} \equiv \sup_{x \in \mathbb{R}^n, \|x\|=1} \|f(x)\|,$$

where $\|\cdot\|$ are the Euclidean norms.

As we saw in linear algebra last term, the operator norm form norms on the space of linear maps as the name suggests. By considering that the set $S_1 := \{x \in \mathbb{R}^n, \|x\| = 1\}$ is closed and bounded, and hence compact by Heine-Borel, as linear maps are continuous, f attains its maximum as on S_1, so we may in fact write the operator norm as

$$\|f\|_{\text{op}} := \max_{x \in \mathbb{R}^n, \|x\|=1} \|f(x)\| < \infty.$$

Proposition 7. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map and $x \in \mathbb{R}^n$, then,

$$\|f(x)\| \leq \|f\|_{\text{op}} \|x\|.$$

Proof. If $x = 0$ then both sides are zero so suppose otherwise. If $x \neq 0$, then $\|x\| > 0$ and so it suffices to show

$$\frac{\|f(x)\|}{\|x\|} \leq \|f\|_{\text{op}}.$$

But, this follows directly from the definition so we are done. □

Proposition 8. If $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, then f is Lipschitz continuous with the Lipschitz constant $\|f\|_{\text{op}}$.

Proof. This follows as, for all $x, y \in \mathbb{R}^n$

$$\|f(x) - f(y)\| = \|f(x - y)\| \leq \|f\|_{\text{op}} \|x - y\|.$$

□

Definition 2.4 (Closed Line Segment)**.** For convenience, for all $x, y \in \mathbb{R}^n$, we define the closed line segment

$$[x, y] := \{\alpha x + (1 - \alpha) y \mid \alpha \in [0, 1]\}.$$

Theorem 1 (Mean Value Inequality)**.** Let $D \subseteq \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^m$ be a function which is continuously differentiable. Then, for all $x, y \in D$ with $[x, y] \subseteq D$, there exists some $\xi \in [x, y]$ such that

$$\|f(x) - f(y)\| \leq \|f'(\xi)\|_{\text{op}} \|x - y\|.$$

Proof. Let $g : [0,1] \to \mathbb{R}^m : \alpha \mapsto f(\alpha x + (1-\alpha)y)$ and we see that g is continuously differentiable as it is the composition of two continuous differentiable functions. Consider,

$$\|f(x) - f(y)\| = \|g(1) - g(0)\| = \left\| \int_0^1 g'(\alpha)\mathrm{d}\alpha \right\|.$$

By the chain rule, we have $g'(\alpha) = (x-y)f'(\alpha x + (1-\alpha)y)$ and so,

$$\begin{aligned}
\left\| \int_0^1 g'(\alpha)\mathrm{d}\alpha \right\| &= \left\| \int_0^1 f'(\alpha x + (1-\alpha)y)(x-y)\mathrm{d}\alpha \right\| \\
&= \int_0^1 \|f'(\alpha x + (1-\alpha)y)(x-y)\|\mathrm{d}\alpha \\
&\leq \|x-y\| \int_0^1 \|f'(\alpha x + (1-\alpha)y)\|_{\text{op}}\mathrm{d}\alpha \\
&\leq \|x-y\| \max_{\alpha \in [0,1]} \|f'(\alpha x + (1-\alpha)y)\|_{\text{op}}.
\end{aligned}$$

So, by defining $\xi = \alpha x + (1-\alpha)y$ where α maximises $\|f'(\alpha x + (1-\alpha)y)\|_{\text{op}}$, we have

$$\|f(x) - f(y)\| \leq \|f'(\xi)\|_{\text{op}}\|x-y\|.$$

□

We note that in the proof above, we assumed $\| \int f\| \leq \int \|f\|$. We shall prove this claim now.

Proposition 9. Let $I \subseteq \mathbb{R}$ be a interval and let $f : I \to \mathbb{R}^m$ be continuous. Then,

$$\left\| \int_{t_0}^{t_1} f(s)\mathrm{d}s \right\| \leq \left| \int_{t_0}^{t_1} \|f(s)\|\mathrm{d}s \right|,$$

for all $t_0, t_1 \in I$.

Proof. We use Riemann rather than Darboux sums for this occasion. Since, as we have shown in first year analysis that Riemann and Darboux sums are equivalent, this does not matter.

Wlog. assume $t_0 < t_1$, and let us consider the n-th Riemann sum of $\left\| \int_{t_0}^{t_1} f(s)\mathrm{d}s \right\|$,

$$\left\| \frac{t_1 - t_0}{n} \sum_{i=0}^{n-1} f\left(t_0 + \frac{i}{n}(t_1 - t_0)\right) \right\| \leq \frac{t_1 - t_0}{n} \left\| \sum_{i=0}^{n-1} f\left(t_0 + \frac{i}{n}(t_1 - t_0)\right) \right\|.$$

Thus, by taking $n \to \infty$, the inequality is achieved. □

Now, with the mean value inequality under our belt, we may generalise the result about Lipschitz continuity for higher dimensions.

Corollary 1.1. Let $U \subseteq \mathbb{R}^n$ be open, $f : U \to \mathbb{R}^m$ be continuously differentiable and $C \subseteq U$ be compact and convex. Then, $f|_C : C \to \mathbb{R}^m$ is Lipschitz continuous.

Proof. Follows straight away from the mean value inequality. □

We note the restriction to a convex subset C since, if otherwise, for all $x, y \in \mathbb{R}^n$, it is not necessarily true that $[x, y] \subseteq U$, and so, the mean value inequality does not apply.

2.3 Picard-Lindelöf Theorem

We have now come to a very important theorem in this course – the Picard-Lindelöf theorem. The Picard-Lindelöf theorem is a strong statement providing us the existence and uniqueness of solutions to particular IVPs.

As we have seen previously, we shall approach this using the Banach's fixed point theorem. By considering the Banach space $C^\circ(J, \mathbb{R}^d)$ equipped with the supremum norm, and by showing a particular mapping is a contraction, we may apply Banach's fixed point theorem resulting in the existence of a solution to the IVP.

Theorem 2 (Picard-Lindelöf Theorem (global version))**.** Consider a nonautonomous differential equation

$$\dot{x} = f(t, x),$$

where $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is continuous and satisfy the global Lipschitz condition, that is, there exists some $K \in \mathbb{R}^+$, such that,

$$\|f(t, x) - f(t, y)\| \leq K\|x - y\|,$$

for all $t \in \mathbb{R}$, $x, y \in \mathbb{R}^d$. Then, the IVP

$$\dot{x} = f(t, x); \;\; x(t_0) = x_0,$$

has a unique solution on $[t_0 - h, t_0 + h]$, given by $\lambda : [t_0 - h, t_0 + h] \to \mathbb{R}^d$, where $h := 1/2K$.

To prove this theorem, let us first prove the following lemma in which we shall simply use the same notations as established for convenience.

Lemma 2.1. Let X be the Banach space $C^\circ([t_0 - h, t_0 + h], \mathbb{R}^d)$ equipped with the supremum norm $\| \cdot \|_\infty$. Then the function,

$$P : X \to X : \lambda \mapsto \left(t \mapsto x_0 + \int_{t_0}^{t} f(s, \lambda(s))\mathrm{d}s \right)$$

is a contraction on X.

Proof. To show P is a contraction, it suffices to show that P is Lipschitz continuous with some Lipschitz constant $K \in \mathbb{R}^+_{<1}$; in this case, we shall show P has Lipschitz constant $1/2$.

Let $u_1, u_2 \in X$, by the definition of the suprmum norm, it suffices to show that for all $t \in [t_0 - h, t_0 + h]$,

$$\|P(u_1)(t) - P(u_2)(t)\| \leq \frac{1}{2}\|u_1 - u_2\|_\infty.$$

This follows as,

$$\begin{aligned}
\|P(u_1)(t) - P(u_2)(t)\| &= \left\| \int_{t_0}^{t} f(s, u_1(s))\mathrm{d}s - \int_{t_0}^{t} f(s, u_2(s))\mathrm{d}s \right\| \\
&= \left\| \int_{t_0}^{t} f(s, u_1(s)) - f(s, u_2(s))\mathrm{d}s \right\| \\
&\leq \left| \int_{t_0}^{t} \|f(s, u_1(s)) - f(s, u_2(s))\|\mathrm{d}s \right|
\end{aligned}$$

Since f satisfy the Lipschitz condition, $\|f(s,u_1(s)) - f(s,u_2(s))\| \le K\|u_1(s) - u_2(s)\| \le K\|u_1 - u_2\|_\infty$ and so,

$$\begin{aligned}\left|\int_{t_0}^t \|f(s,u_1(s)) - f(s,u_2(s))\|\mathrm{d}s\right| &\le K\|u_1 - u_2\|_\infty \left|\int_{t_0}^t \mathrm{d}s\right| \\ &\le Kh\|u_1 - u_2\|_\infty = \frac{1}{2}\|u_1 - u_2\|_\infty.\end{aligned}$$

□

With that, we can apply the Banach fixed point theorem.

Proof. (Picard-Lindelöf theorem). Let P to be the function as before, then by the Banach fixed point theorem, there exists a unique $\lambda \in X$ such that $P(\lambda) = \lambda$. So, λ is the unique local solution to our IVP. □

While the theorem only provides us with a local solution, the solution can be easily extended on the whole space by reapplying the theorem at the end points of the local unique solution (see problem sheet).

We remark that we have proven the global version of the Picard-Lindelöf theorem in contrast to the local version. The global version requires f is globally defined and is also the Lipschitz condition is held globally; these conditions will be relaxed for the local version. Indeed, it is not difficult to see that the global Lipschitz condition is too strong by considering the differential equations such as $\dot{x} = tx^2$.

Let us now differentiate between the functions that are globally and locally Lipschitz continuous.

Definition 2.5 (Globally Lipschitz Continous)**.** $f : D \subseteq \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is said to be globally Lipschitz continuous with respect to x if there exists some $K > 0$ such that for all $t \in \mathbb{R}$, $x, y \in \mathbb{R}^d$,

$$\|f(t,x) - f(t,y)\| \le K\|x - y\|.$$

Definition 2.6 (Locally Lipschitz continuous)**.** $f : D \subseteq \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is said to be locally Lipschitz continuous with respect to x if for all $(t_0, x_0) \in D$, there exists some neighbourhood U of (t_0, x_0) and some $K > 0$ such that

$$\|f(t,x) - f(t,y)\| \le K\|x - y\|,$$

for all $(t,x), (t,y) \in U$.

Theorem 3 (Picard-Lindelöf Theorem (local version))**.** Let $D \subseteq \mathbb{R} \times \mathbb{R}^d$ be open and $f : D \to \mathbb{R}^d$ be continuous and locally Lipschitz continuous. Then, for all $(t_0, x_0) \in D$, the IVP

$$\dot{x} = f(t,x); \quad x(t_0) = x_0$$

has a unique solution on an interval of the form $[t_0 - h, t_0 + h]$ where $h = h(t_0, x_0)$ (Qualitative version). Furthermore, if we denote U for the neighbourhood of (t_0, x_0) on which the Lipschitz condition is held with the Lipschitz constant $K > 0$, by defining

$$W^{\tau,\delta}(t_0, x_0) := [t_0 - \tau, t_0 + \tau] \times \overline{B_\delta(x_0)} \subseteq U,$$

for some sufficent $\tau, \delta > 0$, and $M :\geq \|f(t,x)\|$ on $W^{\tau,\delta}(t_0, x_0)$, there exists a unique solution on the interval $[t_0 - h, t_0 + h]$, where $h := \min\{\tau, 1/2K, \delta/M\}$ (Quantitative version).

Proof. See extra meterials. $\square$

Proposition 10. Let $\Omega \subseteq \mathbb{R}^d$ be open and $f : \Omega \to \mathbb{R}^d$ be continuously differentiable. Then f is locally Lipschitz continuous.

Proof. For al $x \in \Omega$, let $U \subseteq \Omega$ be a compact set containing x (U exists as Ω is open, there exists some $\delta > 0$ such that $B_\delta(x) \subseteq \Omega$, and so, we can simply take $U = \overline{B_{\delta/2}(x)}$). Then, since f' is continuous, it is bounded by some $M > 0$. Thus, for all $x, y \in U$, by the mean value inequality, there exists some $\xi \in [x, y]$ such that

$$\|f(x) - f(y)\| \leq \|f'(\xi)\|_{\mathrm{op}} \|x - y\| \leq M\|x - y\|.$$

Thus, f is locally Lipschitz at x on U with the Lipschitz constant M. $\square$

Lemma 2.2. Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ be continuous and locally Lipschitz where D is open. Then, given two solutions $\lambda : I \to \mathbb{R}^d$ and $\mu : J \to \mathbb{R}^d$ of the differential equation $\dot{x} = f(t, x)$, either $\lambda(t) = \mu(t)$ or $\lambda(t) \neq \mu(t)$ for all $t \in I \cap J$.

Proof. Suppose there exists some $t_0, t_1 \in I \cap J$ such that $\lambda(t_0) = \mu(t_0)$ and $\lambda(t_1) \neq \mu(t_1)$, and furthermore, Wlog. assume $t_1 > t_0$. Then, by defining

$$\tilde{t} := \sup\{t > t_0 \mid \lambda(t') = \mu(t'),\ \forall t' \in [t_0, t]\},$$

by the continuity of λ and μ, $\lambda - \mu$ is continuous, and so, by considering $0 = \lim_{n\to\infty}(\lambda - \mu)(\tilde{t} - 1/n) = (\lambda - \mu)(\tilde{t})$, we have $\lambda(\tilde{t}) = \mu(\tilde{t})$. Then, by applying the local version of the Picard-Lindelöf theorem with the initial value $(\tilde{t}, \lambda(\tilde{t}) = \mu(\tilde{t}))$, we see that there exists two different solutions #. $\square$

As we have seen, the local version of Picard-Lindelöf provides us with a unique solution on some interval around the initial value. We would now like to maximise the size of this interval around the initial value and such solutions are called maximal solutions.

Definition 2.7 (Maximal Existence Interval)**.** Given a initial value problem $\dot{x} = f(t, x)$, $x(t_0) = x_0$, we define

$$I_+(t_0, x_0) := \sup\{t \geq t_0 \mid \exists \text{ solution on } [t_0, t]\},$$

$$I_-(t_0, x_0) := \inf\{t \leq t_0 \mid \exists \text{ solution on } [t, t_0]\},$$

and the maximal existence interval,

$$I_{\max}(t_0, x_0) := (I_-(t_0, x_0), I_+(t_0, x_0)).$$

Theorem 4. There exists a maximal solution[2]

$$\lambda_{\max} : I_{\max}(t_0, x_0) \to \mathbb{R}^d$$

[2] Any solution is defined on a subset of $I_{\max}(t_0, x_0)$.

of the IVP $\dot{x} = f(t, x)$, $x(t_0) = x_0$ where f is continuous and locally Lipschitz. Furthermore, if $I_+(t_0, x_0) < \infty$, then either the maximal solution is unbounded for $t \supseteq t_0$, that is,

$$\sup_{t \geq t_0} \|\lambda_{\max}(t)\| = \infty,$$

or $\partial D \neq \varnothing$, and

$$\lim_{t \uparrow I_+(t_0, x_0)} \operatorname{dist}((t, \lambda_{\max}(t)), \partial D) = 0,$$

where $\operatorname{dist}(y, A) := \inf\{\|x - y\| \mid x \in A\}$. Analogous case for $I_-(t_0, x_0) > -\infty$.

Proof. Let $\bar{t} \in I_{\max}(t_0, x_0)$, then, by the definition of the maximal existence interval, there exists some solution $\mu : I \to \mathbb{R}^d$ such that $\bar{t}, t_0 \in I$. Now, by considering that the solutions cannot cross each other, we see that all solutions of the IVP on $[t_0, \bar{t}]$ must coincide as they crosses at (t_0, x_0). Thus, the definition

$$\lambda_{\max}(\bar{t}) := \mu(\bar{t})$$

is well-defined. Indeed, $\lambda_{\max}$ is a solution at $\bar{t}$ since

$$\dot{\lambda}_{\max}(\bar{t}) = \dot{\mu}(\bar{t}) = f(\bar{t}, \mu(\bar{t})) = f(\bar{t}, \lambda_{\max}(\bar{t})).$$

Now, by the construction of $I_{\max}$, the only possible case on which the the theorem does not hold is that $I_+(t_0, x_0) \in I_{\max}(t_0, x_0)$. So, for contradiction, assume that there exists some solution $\mu : [t, I_{\max}(t_0, x_0)] \to \mathbb{R}^d$ to the IVP. However, by the local version of Picard-Lindelöf, there exists some open U containing $I_{\max}(t_0, x_0)$ such that there exists a unique solution η of the IVP on the interval $(I_{\max}(t_0, x_0) - h, I_{\max}(t_0, x_0) + h)$, and hence, we may extend μ to be a solution on $[t, I_{\max}(t_0, x_0) + h]$ contradicting the maximum condition #.

For the second part of the theorem, assume $I_+(t_0, x_0) < \infty$. Then, we may find some sequence $(t_n)_{n=1}^{\infty} \subseteq D$ that converges to $I_-(t_0, x_0)$ and assume, for contradiction, there exists some $M > 0$, such that $\|\lambda_{\max}(t_n)\| \leq M$ and $\operatorname{dist}((t_n, \lambda_{\max}(t_n)), \partial D) \geq 1/M$ for all $n \in \mathbb{N}$.

By assumption, we have $(t_n, \lambda_{\max}(t_n))$ is a bounded sequence, and so, by Bolzano-Weierstrass, it has a convergent subsequence $(t_{n_i}, \lambda_{\max}(t_{n_i}))$ which converges to $(t^*, x^*) \in D$. Now, on the problem sheet, we shall see that, the solution is uniform on a neighbourhood of (t^*, x^*). But, this contradicts the maximal condition, so we are done! $\square$

2.4 General Solution and Flows

By the existence of the maximal solution, we find there to be a general solution for all initial pairs.

Proposition 11. Let $f : D \subseteq \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ be continuous and locally Lipschitz. Then there exists a maximal solution $\lambda_{\max}$ to the IVP

$$\dot{x} = f(t, x), \quad x(t_0) = x_0$$

for every initial pair $(t_0, x_0) \in D$.

Proof. Follows straight away from the existence of maximal solutions. □

In the case as described in the proposition, we can define $\lambda : D \times \mathbb{R} \to \mathbb{R}^d$ such that $\lambda((t_0, x_0), t)$ is the maximal solution to $\dot{x} = f(t, x)$ and $x(t_0) = x_0$. This λ is called the general solution. Commonly, we shall flatten the arguments of λ such that $\lambda = \lambda(t, t_0, x_0)$.

Proposition 12. Let $\dot{x} = f(t, x)$ be a differential equation with the general solution λ. Then, given $(t_0, x_0) \in D$,

- $I_{\max}(s, \lambda(s, t_0, x_0)) = I_{\max}(t_0, x_0)$;
- $\lambda(t_0, t_0, x_0) = x_0$;
- $\lambda(t, s, \lambda(s, t_0, x_0)) = \lambda(t, t_0, x_0)$.

Proof. The second statement is immediate since the general solution satisfies the IVP so let us consider the two other claims.

Consider the initial pair $(s.\lambda(s, t_0, x_0))$. Then, we see that both $\lambda(\cdot, t_0, x_0)$ and $\lambda(\cdot, s, \lambda(s, t_0, x_0))$ coincides at $(s, \lambda(s, t_0, x_0))$. Now, since the maximal solution is unique, the results follow. □

In the case that the differential equation is autonomous, we find the the general solution contains redundancy due to translational invariant. So, this leads us to the concept of *flow*.

Let $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^d$ be continuous and locally Lipschitz and suppose λ is a general solution to the differential equation $\dot{x} = f(x)$. Then, if $t_0 \in \mathbb{R}$ and $x_0 \in D$ is fixed, then $\lambda(\cdot, t_0, x_0)$ solves the IVP

$$\dot{x} = f(x), \ \ x(t_0) = x_0.$$

However, consider $\lambda(\cdot, 0, x_0)$ is a solution to the IVP is initial condition $x(0) = x_0$, we may translate the solution so that

$$\lambda^*(t) = \lambda(t - t_0, 0, x_0),$$

which is also a solution to the IVP with initial condition $x(t_0) = x_0$. So, since the solution is unique, we see that

$$\lambda(t - t_0, 0, x_0) = \lambda(t, t_0, x_0).$$

So, since the left hand equation contains only two variable, we conclude that the general solution has redundant information.

Definition 2.8 (Flow)**.** Let $\dot{x} = f(x)$ with $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^d$. For all $x_0 \in D$ let us define

$$J_{\max}(x_0) := I_{\max}(0, x_0),$$

and

$$\phi(t, x_0) := \lambda(t, 0, x_0),$$

then ϕ is called the flow of the autonomous differential equation $\cdot x = f(x)$.

Proposition 13. Let ϕ be a flow of $\dot{x} = f(x)$ where $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^d$ is continuous and locally Lipschitz. Then for $x \in D$, we have

- $J_{\max}(\phi(t, x)) = J_{\max}(x) - t$;

- $\phi(0, x) = x$;
- $\phi(t, \phi(s, x)) = \phi(t + s, x)$;
- $\phi(-t, \phi(t, x)) = x$.

Proof. Easy as before. $\square$

Definition 2.9 (Orbit)**.** Let ϕ be a flow of $\dot{x} = f(x)$ where $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^d$ is continuous and locally Lipschitz. Then, given $x \in D$, the orbit (or trajectory) through x is

$$\sigma(x) := \{\phi(t, x) \in D \mid t \in J_{\max}(x)\}.$$

Furthermore, the positive half-orbit and the negative half-orbit through x are

$$\sigma^+(x) := \{\phi(t, x) \in D \mid t \in J_{\max}(x) \cap \mathbb{R}^+\},$$

$$\sigma^-(x) := \{\phi(t, x) \in D \mid t \in J_{\max}(x) \cap \mathbb{R}^-\},$$

respectively.

We would like to classify the different possible orbits and it turns out there are three possible orbits.

- $\sigma(x)$ is a singleton. This occurs when $f(x) = 0$ and $J_{\max}(x) = \mathbb{R}$ and we call x an equilibrium of the equation;
- $\sigma(x)$ is a closed curve. That is there exists $t > 0$ such that $\phi(t, x) = x$ and $f(x) \neq 0$. In this case $J_{\max}(x) = \mathbb{R}$ and we call x is called a periodic orbit;
- $\sigma(x)$ is not a closed curve nor a singleton. In this case, the solution does not necessarily exist for all time, however, we see that the map $t \mapsto \phi(t, x)$ is injective.

Lastly, let us consider the continuity and differentiability of general solutions and flows. Indeed, by construction, the general solution and flow must be differentiable and so, is continuous, however, we would like to also consider continuity with respect to the other arguments. In this course, we will state without proof that both the general solution and flow are continuous functions under the standard assumptions while this will be further investigated in more advanced modules.

3 Linear Systems

Linear differential equations are important models for certain applications, but the theory of linear differential equations is of utmost importance also for non-linear systems.

Consider the differential equation

$$\dot{x} = f(t, x),$$

and assume $\mu : I \to \mathbb{R}^d$ is a solution. As in perturbation theory, or the real world, there might be some small deviations from the theoretical solution and we would like to examine what happens if the initial condition is a small step away. This can be studied through the variational equation

$$\dot{y} = \frac{\partial f}{\partial x}(t, \mu(t))y = A(t)y,$$

which is a nonautonomous linear system (see problem sheet). In general, the nonautonomous case if difficult to solve exactly while the situation is more simple if

- f is autonomous;
- μ is constant, i.e. $\mu = x_0$ and $f(x_0) = 0$.

Indeed, if the two conditions above are satisfied, then the variational equation becomes $\dot{y} = Ay$ where $A = f'(x_0)$.

3.1 Autonomous Linear Differential Equations

For a matrix $A \in \mathbb{R}^{d\times d}$, consider

$$\dot{x} = Ax.$$

Indeed, by considering, for all $x, y \in \mathbb{R}^d$,

$$\|Ax - Ay\| = \|A(x - y)\| \leq \|A\|_{\mathrm{op}}\|x - y\|,$$

we have $x \mapsto Ax$ is globally Lipschitz, and so, we may apply the global version of Picard-Lindelöf resulting in the unique existence of a globally defined flow $\phi : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$.

To find this solution, we will again use the Picard iterates. For all $t \in [-h, h]$ where $h = 1/2\|A\|_{\mathrm{op}}$, we define $\lambda_0(t) = x_0$ and for $n \in \mathbb{N}$,

$$\lambda_{n+1}(t) = x_0 + \int_0^t A\lambda_n(s)\mathrm{d}s.$$

Inductively, one finds that the closed form of the Picard iterates can be written as

$$\lambda_n(t) = \sum_{k=0}^{n} \frac{t^k A^k}{k!} x_0.$$

So, by taking $n \to \infty$, by the Picard-Lindelöf theorem, we have for all $t \in [-h, h]$, the flow ϕ of the differential equation is

$$\phi(t, x_0) = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} x_0 =: e^{At}$$

and we call it the matrix exponential function. However, we note that we have not justified that e^{At} exists and hence well-defined. We shall prove this now.

Lemma 3.1. For all $B, C \in \mathbb{R}^{d\times d}$, $\|BC\|_{\text{op}} \leq \|B\|_{\text{op}}\|C\|_{\text{op}}$.

Proof. Let $x \in \mathbb{R}^d \setminus \{0\}$, then

$$\|BCx\| \leq \|B\|_{\text{op}}\|Cx\| \leq \|B\|_{\text{op}}\|C\|_{\text{op}}\|x\|,$$

and so, we have

$$\frac{\|BCx\|}{\|x\|} \leq \|B\|_{\text{op}}\|C\|_{\text{op}}.$$

Thus, by taking the supremum on the left hand side, the inequality is established. $\square$

Proposition 14. For $B \in \mathbb{R}^{d\times d}$, the matrix exponential defined as

$$e^B := \sum_{k=0}^{\infty} \frac{1}{k!} B^k$$

exists.

Proof. By the above lemma, through induction, we have $\|B^k\|_{\text{op}} \leq \|B\|_{\text{op}}^k$ for all $k \in \mathbb{N}$. Then, consider

$$e^{\|B\|_{\text{op}}} = \sum_{k=0}^{\infty} \frac{\|B\|_{\text{op}}^k}{k!} \geq \sum_{k=0}^{\infty} \frac{\|B^k\|_{\text{op}}}{k!},$$

where the right hand side is a increasing bounded sequence. So,

$$\sum_{k=0}^{\infty} \frac{\|B^k\|_{\text{op}}}{k!}$$

exists. Now, by considering for all $n > m$, we have

$$\left\| \sum_{k=0}^{n} \frac{B^k}{k!} - \sum_{k=0}^{m} \frac{B^k}{k!} \right\|_{\text{op}} = \left\| \sum_{k=m+1}^{n} \frac{B^k}{k!} \right\|_{\text{op}} \leq \sum_{k=m+1}^{n} \frac{\|B^k\|_{\text{op}}}{k!} \to 0$$

by the trivial test. Thus, $\sum_{k=0}^{n} \frac{B^k}{k!}$ converges as required. $\square$

Theorem 5. Consider the linear differential equation

$$\dot{x} = Ax.$$

Then, the flow of this differential equation is give by

$$\phi(t, x) = e^{At}x,$$

for all $t \in \mathbb{R}, x \in \mathbb{R}^d$.

Proof. As shown above, there exists some $h > 0$ such that $\phi(t,x) = e^{At}x$ for all $x \in \mathbb{R}^d$ and $t \in [-h,h]$. Then, for all $t,s \in [-h/2,h/2]$, $e^{A(t+s)}x = \phi(t+s,x) = \phi(t,\phi(s,x)) = e^{At}e^{As}x$. Indeed, by algebraic manipulation, we see that this property holds for all $t,s \in \mathbb{R}$.

Now, let $t \in \mathbb{R}$ and some $N \in \mathbb{N}$ such that $t/N \in [-h,h]$. Furthermore, define

$$\phi_{t/N} : \mathbb{R}^d \to \mathbb{R}^d \ : x \mapsto \phi\left(\frac{t}{N}, x\right),$$

and so,

$$\phi(t,x) = (\phi_{t/N} \circ \cdots \circ \phi_{t/N})(x) = \prod_{i=1}^{N} e^{\frac{t}{N}A}x = e^{tA}x$$

which is exactly what we wanted to show. $\square$

Proposition 15. Let $B, C, T \in \mathbb{R}^{d\times d}$ and suppose T is invertible. Then,

- $C = T^{-1}BT \implies e^C = T^{-1}e^BT$;
- $e^{-B} = (e^B)^{-1}$;
- $BC = CB \implies e^{B+C} = e^Be^C$;
- $B = \bigoplus B_i \implies e^B = \bigoplus e^{B_i}$.

Proof. See problem sheet. $\square$

3.2 Planar Linear Systems

Consider the linear system $\dot{x} = Ax$. By decomposing A through Jordan normal decomposition, we have $A = T^{-1}JT$ where J is in Jordan normal form. Furthermore, we see that $e^{At} = Te^{Jt}T^{-1}$ and so, if we have a explicit form for the matrix exponential of matrices in Jordan normal form, we would also obtain a method for calculating e^A explicitly.

We shall in particular consider the case where $A \in \mathbb{R}^{2\times 2}$. In this case, there are 4 possible Jordan forms which are

- $J = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ if A has two distinct eigenvalue $a, b \in \mathbb{R}$;
- $J = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ if A has one eigenvalue $a \in \mathbb{R}$ and $\dim E_a = 2$;
- $J = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$ if A has one eigenvalue $a \in \mathbb{R}$ and $\dim E_a = 1$;
- $J = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ if A has one eigenvalue $a + ib \in \mathbb{C}$ and $b \neq 0$.

We remark the last case is the real-version of the Jordan form for complex eigenvalues. We shall justify this later on.

Suppose A is not singular, i.e. 0 is not an eigenvalue of A and we shall consider the four cases individually.

(1) $J = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ for some distinct $a, b \in \mathbb{R} \setminus \{0\}$. In this case, since J is diagonal, we see straight away that

$$e^{Jt} = \begin{pmatrix} e^{at} & 0 \\ 0 & e^{bt} \end{pmatrix}.$$

By considering the phase portrait (union of all orbits) of this system, we have

$$\gamma(x_0, y_0) := \{\phi(t, (x_0, y_0)) \mid t \in \mathbb{R}\} = e^{Jt}(x_0, y_0)^T = \{(x_0 e^{at}, y_0 e^{bt} \mid t \in \mathbb{R}\}.$$

By defining $x := x_0 e^{at}$ assuming $x_0 \neq 0$, we see that $t = \frac{\log(x/x_0)}{a}$ and so, the trajectory becomes,

$$\gamma(x_0, y_0) = \left\{ \left(x, y_0 \left(\frac{x}{x_0} \right)^{\frac{b}{a}} \right) \mid \frac{x}{x_0} > 0 \right\}.$$

(2) $J = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ for some $a \in \mathbb{R} \setminus \{0\}$. This case follows from the first resulting in

$$e^{Jt} = \begin{pmatrix} e^{at} & 0 \\ 0 & e^{at} \end{pmatrix},$$

and so the phase portrait is,

$$\gamma(x_0, y_0) = \left\{ \left(x, x\frac{y_0}{x_0} \right) \mid \frac{x}{x_0} > 0 \right\}.$$

(3) $J = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$ for some $a \in \mathbb{R} \setminus \{0\}$. One can check that in this case

$$e^{Jt} = \begin{pmatrix} e^{at} & te^{at} \\ 0 & e^{at} \end{pmatrix},$$

(see problem sheet). In this case, we have the orbit

$$\sigma(x_0, y_0) = \left\{ \left(\frac{x_0}{y_0} y + \frac{y}{a} \log \frac{y}{y_0}, y \right) \mid \frac{y}{y_0} > 0 \right\}.$$

(4) $J = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ where $a, b \in \mathbb{R} \setminus \{0\}$. In this case, one see that

$$e^{Jt} = e^{at} \begin{pmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{pmatrix},$$

(see problem sheet). In this case, the motion is circular and more details is provided on the lecture slides.

In the case that A is singular, we have $J = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ and so,

$$e^{Jt} = \begin{pmatrix} e^{at} & 0 \\ 0 & 1 \end{pmatrix},$$

resulting in no movements in the y direction.

By investigating the trajectories qualitatively, we see that for all linear system $\dot{x} = Ax$ where $\rho = a + ib$ is an eigenvalue of A, we have b characterises rotation and a characterises exponential growth with the growth rate obtained by

$$a = \lim_{t \to \infty} \frac{1}{t} \log \mu(t).$$

In this case where $\mu(t) = e^{at}$ we have what is known as a pure exponential growth. There are also non-pure exponential growths such as $\mu(t) = t^n e^{at}$, and in this case, we also have the growth rate a.

Definition 3.1 (Lyapunov Exponent)**.** Let $\lambda : \mathbb{R} \to \mathbb{R}^d$ be a non-zero solution to $\dot{x} = Ax$ where $A \in \mathbb{R}^{d \times d}$. Then the Lyapunov exponent of λ is given by

$$\sigma_{\text{Lyap}}(\lambda) := \lim_{t \to \infty} \frac{1}{t} \log \|\lambda(t)\|,$$

if the limit exists. Indeed, one can show that, for a linear system the limit will always exist and equal the real part of one of the eigenvalues.

We say that λ is exponentially decaying if $\sigma_{\text{Lyap}}(\lambda) < 0$ and exponentially growing if $\sigma_{\text{Lyap}}(\lambda) > 0$.

3.2.1 Real Jordan Normal Form for Complex Eigenvalues

As our aim is to find an explicit representation of the flow $\phi(t, x) = e^{At}x$. While, we know how to find Jordan normal forms for general matrices, we shall now establish a method for finding the real Jordan normal form for complex eigenvalues.

Suppose $A \in \mathbb{R}^{d \times d}$ with real-JCF J. Then we would like to find invertible $T \in \mathbb{R}^{d \times d}$ such that

$$J = T^{-1} A T.$$

This is useful for us since then

$$e^{At} = T e^{Jt} T^{-1},$$

where the right hand side is explicitly computable.

Definition 3.2 (Real Jordan Block)**.** A matrix J is a real Jordan block if either it is a Jordan block consisting of real entries of in the form

$$J = \begin{pmatrix} C_j & I_2 & \cdots & 0 \\ 0 & C_j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_2 \\ 0 & 0 & \cdots & C_j \end{pmatrix}$$

where C_j is in the form

$$C_j = \begin{pmatrix} a_j & b_J \\ -b_j & a_j \end{pmatrix}$$

for some real a_j, b_j.

We see that the real Jordan block corresponds directly to the case where a matrix contains complex eigenvalues. Indeed, if we write the imaginary numbers in complex form, we have the desired form.

Theorem 6. Let $A \in \mathbb{R}^{d\times d}$. Then, there exists some invertible T such that $J := T^{-1}AT$ is in the real Jordan normal form, i.e. there exists $(J_i)_{i=1}^r$ where J_i are real Jordan blocks such that

$$J = \bigoplus_{i=1}^r J_i.$$

We shall not prove this theorem here however, let us consider the case that $d = 2$. Suppose $A \in \mathbb{R}^{2\times 2}$ has eigenvalues $\rho = a + ib, \bar{\rho} \in \mathbb{C}$ such that $\mathrm{Im}\rho \neq 0$. Then, the corresponding Jordan form of A is

$$J = \begin{pmatrix} \rho & 0 \\ 0 & \bar{\rho} \end{pmatrix}.$$

Suppose we denote $u+iv$ as the eigenvector corresponding to ρ such that $A(u+iv) = \rho(u+iv)$. Now, if u, v are linearly dependent, then there exists some $\lambda \in \mathbb{R}$ such that $u = \lambda v$ and so $u + iv = (\lambda + i)v$, resulting in $\rho(\lambda + i)v = \rho(u + iv) = A(u + iv) = (\lambda + i)Av$ and so, $\rho v = Av$ implying either $\rho \in \mathbb{R}$ #, so u, v are linearly independent and hence, $B := \{u, v\}$ forms a basis of $\mathbb{R}^2$. Thus, choosing the transformation with respect to this basis, we have

$$[J]_B = \begin{pmatrix} a & b \\ -b & a \end{pmatrix},$$

as required.

This argument can be easily extended to higher dimensions and so, we have a method for finding the real Jordan normal forms and we may continue to find the explicit representation of the matrix exponential function.

By recalling that $e^{\bigoplus_i J_i} = \bigoplus_i e^{J_i}$, we may obtain the explicit form of any matrix exponential function e^{Jt} given the explicitly representation of the individual exponential blocks.

Proposition 16. Let J_i be a Jordan block to the corresponding eigenvalue ρ_i. If ρ_i is real, then

$$e^{J_i t} = e^{\rho_i t} \begin{pmatrix} 1 & t & \frac{1}{2}t^2 & \cdots & \frac{t^{d_i - 1}}{(d_i-1)!} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \frac{1}{2}t^2 \\ & & & \ddots & t \\ & & & & 1 \end{pmatrix}.$$

Proof. We approach this similarly to that we had done for linear algebra – writing the Jordan block as $J_i = D + N$ where $D = \mathrm{diag}(\rho_i, \cdots, \rho_i)$ and N is the matrix with 1 above the diagonal. Since

$$e^{J_1} = e^{D+N} = e^D e^N$$

(as $DN = ND$) it suffices to find e^D and e^N respectively. As D is a diagonal matrix, we have $e^D = \operatorname{diag}(e_i^\rho, \cdots, e_i^\rho)$. On the other hand, by noticing N is a nilpotent matrix, we have

$$e^N = \sum_{k=0}^{\infty} \frac{N^k}{k!} = \sum_{k=0}^{d-1} \frac{N^k}{k!}.$$

With that, by induction, the result follows. $\square$

Proposition 17. Let J_i be a real Jordan block to the corresponding eigenvalue ρ_i. If $\rho_i = a_i + ib_i$ is , then

$$e^{J_i t} = e^{\rho_i t} \begin{pmatrix} G(t) & tG(t) & \frac{1}{2}t^2 G(t) & \cdots & \frac{t^{d_i-1}}{(d_i-1)!}G(t) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \frac{1}{2}t^2 G(t) \\ & & & \ddots & tG(t) \\ & & & & G(t) \end{pmatrix},$$

where

$$G(t) = \begin{pmatrix} \cos(b_i t) & \sin(b_i t) \\ -\sin(b_i t) & \cos(b_i t) \end{pmatrix}.$$

Proof. Same argument as above. $\square$

Definition 3.3 (Invariant)**.** A subset $E \subseteq \mathbb{R}^2$ is called invariant if for all $x \in E$, the flow, $\phi(t, x) \in E$ for all $t \in \mathbb{R}$.

Definition 3.4 (Spectrum)**.** Suppose $\dot{x} = Ax$, then the spectrum of A is

$$\Sigma(A) := \{\operatorname{Re}(\rho_i) \mid \rho \text{ is a eigenvalue of } A\}.$$

By the spectral theorem from linear algebra, we see that we can decompose $\mathbb{R}^d$ into eigenspaces of A, i.e.

$$A = E_{\rho_1} \oplus E_{\rho_2} \oplus \cdots \oplus E_{\rho_q},$$

and furthermore, the Lyapunov exponents are the ρ_i.

Definition 3.5 (Semi-Simple)**.** An eigenvalue of a matrix A is called semi-simple if its algebraic multiplicity equals its geometric multiplicity.

Proposition 18. Consider $\dot{x} = Ax$ where $A \in \mathbb{R}^{d\times d}$. Then, there exists some $\gamma \in \mathbb{R}$ such that $\gamma > \max \Sigma(A)$ and there exists some $K > 0$ such that

$$\|e^{At}\| \le K e^{\gamma t},$$

for all $t \ge 0$. Furthermore, If all eigenvalues ρ with $\operatorname{Re}\rho = \max \Sigma(A)$ are semi-simple, then, we can choose $\gamma = \max \Sigma(A)$.

Proof. By writing A in Jordan normal form, we have, there exists some invertible $T \in \mathbb{R}^{d\times d}$ such that $J = T^{-1}AT$ is in Jordan normal form. Now, since all norms on finite dimensional spaces are equivalent, it suffices to show there exists some $\tilde{K} > 0$ such that

$$\|e^{Jt}\|_\infty \leq \tilde{K}e^{\gamma t},$$

where $\|\cdot\|_\infty$ is the supremum norm. Now, by recalling our classification of the matrix exponential by a Jordan block, we have write

$$e^{Jt} = g(t)t^n e^{\rho t},$$

where g is bounded where $n = 0$ if and only if A is semi-simple. Since, $e^{\gamma t}$ increases faster than t^n, the result follows. $\square$

3.3 Variation of Constants

We are interested in the general solution of the inhomogeneous linear differential equation

$$\dot{x} = Ax + g(t)$$

for some $g : I \to \mathbb{R}^d$ which is continuous.

Proposition 19. The general solution to $\dot{x} = Ax + g(t)$ is given by

$$\lambda(t, t_0, x_0) = e^{A(t-t_0)}x_0 + \int_{t_0}^t e^{A(t-s)}g(s)\mathrm{d}s,$$

for all $t, t_0 \in I$, $x_0 \in \mathbb{R}^d$.

Proof. Let $\mu_g(t) := \int_{t_0}^t e^{A(t-s)}g(s)\mathrm{d}s = e^{At}\int_{t_0}^t e^{-As}g(s)\mathrm{d}s$. Then, mu_g is a solution to $\dot{x} = Ax + g(t)$ since

$$\dot{\mu}_g(t) = Ae^{At}\int_{t_0}^t e^{-As}g(s)\mathrm{d}s + e^{At}\ e^{-At}g(t) = A\mu g + g(t).$$

Now, let $\lambda_h(t)$ be the general solution to the homogeneous system $\dot{x} = Ax$, and so $\lambda_h(t, t_0, x_0) = e^{A(t-t_0)}x_0$. Then the general solution ν for some fixed (t_0, x_0) is

$$\nu(t) = \lambda_h(t, t_0, x_0) + \mu_g(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^t e^{A(t-s)}g(s)\mathrm{d}s.$$

$\square$

The homogeneous linear system $\dot{x} = A(t)x$ where $A : I \to \mathbb{R}^{d\times d}$ is not solvable exactly in general. Thus, inhomogeneous systems in the form of

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

is also in general not solvable. However, solutions can be obtained for $d = 1$, with

$$\lambda(t, t_0, x_0) = e^{\int_{t_0}^t A(s)\mathrm{d}s}x_0 + \int_{t_0}^t e^{\int_s^t A(\tau)\mathrm{d}\tau}g(s)\mathrm{d}s.$$

4 Nonlinear Systems

4.1 Stability

Definition 4.1 (Types of Stability)**.** Consider the autonomous differential equation

$$\dot{x} = f(x)$$

when $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz and let $\phi(t, x)$ be the induced flow. Then, given the equilibrium x^*, i.e. $f(x^*) = 0$,

- x^* is stable if for all $\epsilon > 0$, there exists some $\delta > 0$, such that $\|\phi(t, x) - x^*\| < \epsilon$ for all $x \in B_\delta(x^*)$, $t \geq 0$;
- x^* is unstable if x^* is not stable;
- x^* is attractive if there exists some $\delta > 0$, such that

  $$\lim_{t \to \infty} \phi(t, x) = x^*,$$

  for all $x \in B_\delta(x^*)$;
- x^* is asymptotically stable if x^* is stable and attractive;
- x^* is exponentially stable if there exists some $\delta > 0$, $K \geq 1$, $\gamma < 0$ such that

  $$\|\phi(t, x) - x^*\| \leq K e^{\gamma t} \|x - x^*\|$$

  for all $x \in B_\delta(x^*)$ and $t \geq 0$.
- x^* is repulsive if there exists some $\delta > 0$ such that

  $$\lim_{t \to -\infty} \phi(t, x) = x^*,$$

  for all $x \in B_\delta(x^*)$.

We remark that the conditions for stable equilibrium is similar but stronger than to continuity. Indeed, we see that x^* is stable requires δ to be independent from t while continuity does not.

In general, stability and attractive are not related. To see this, consider the example $\dot{x} = 0$. In this case every point $x^* \in \mathbb{R}^d$ is a equilibrium that is stable since we can simply choose $\delta = \epsilon$. However, x^* is not attractive since there is no movement. Similarly, attractive does not imply stability (as we shall see on the problem sheet, attractive does in fact imply stability in one dimension) by the following example,

$$\begin{cases} \dot{x} = x + xy - (x + y)\sqrt{x^2 + y^2}, \\ \dot{y} = y - x^2 + (x - y)\sqrt{x^2 + y^2}. \end{cases}$$

This system written in polar coordinates becomes $\dot{r} = r(1 - r)$; $\dot{\phi} = r(1 - \cos\phi)$. By considering the equilibrium $x^* = (1, 0)$, we have a attractive equilibrium that is not stable.

Definition 4.2. Consider the differential equation $\dot{x} = f(x)$, with flow ϕ. The orbit $\sigma(x)$ is

- called homoclinic if there exists an equilibrium $x^* \in D \setminus \{x\}$ such that

$$\lim_{t\to\infty} \phi(t,x) = x^* \text{ and } \lim_{t\to-\infty} \phi(t,x) = x^*;$$

- called heteroclinic if there exists two equilibria $x_1^* \neq x_2^*$ such that

$$\lim_{t\to\infty} \phi(t,x) = x_1^* \text{ and } \lim_{t\to-\infty} \phi(t,x) = x_2^*.$$

Let us now consider the stability of linear systems.

Theorem 7 (Stability of Linear Systems)**.** Consider the autonomous linear system

$$\dot{x} = Ax,$$

where $A \in \mathbb{R}^{d\times d}$. Then the trivial equilibrium $x^* = 0$ of this system is

- stable if and only if the real part of all eigenvalues ρ of A is non-positive and if Re $\rho = 0$, then ρ is semi-simple;
- exponentially stable if and only if Re $\rho < 0$ for all eigenvalues ρ of A.

Proof. We shall prove the first part of the theorem with the second part found on the exercise sheet.

Let J be the Jordan normal form of A such that there exists some $T \in GL_d(\mathbb{R})$ such that $J = T^{-1}AT$ and so, the flow of the system is

$$\phi(t,x) = e^{At}x = Te^{Jt}T^{-1}x.$$

($\Longrightarrow$) We show the contrapositive. Suppose first that A has a eigenvalue ρ with a positive real part. Then, e^{Jt} is not bounded implying e^{At} is unbounded # Similar for the case where ρ is not semi-simple.

($\Longleftarrow$) If both assumption holds then there exists some $K > 0$ such that $\|e^{At}\| \leq K$ for all $t \geq 0$. So, for all $\epsilon > 0$, choose $\delta := \epsilon/K$. Then, for all $x \in B_\delta(0)$, we get

$$\|\phi(t,x)\| = \|e^{At}x\| \leq \|e^{At}\|\|x\| < K\delta = \epsilon.$$

Hence, $x^* = 0$ is stable. $\square$

Now that we understand the stability of linear systems, we would like to consider the same for nonlinear systems. As it is in general very difficult to solve a nonlinear system explicitly, we consider the system locally in a neighbour hood of a reference solution, e.g. given the nonlinear system

$$\dot{x} = f(x),$$

with the equilibrium x^*, we consider the linear system

$$\dot{y} = f'(x^*)y.$$

4.2 Linearised Stability

Definition 4.3 (Hyperbolicity)**.** A matrix $A \in \mathbb{R}^{d\times d}$ is called hyperbolic if all eigenvalues λ of A has a non-zero real part, i.e. $\Sigma(A) \cap \{0\} = \varnothing$. An equilibrium x^* of a differential equation

$$\dot{x} = f(x),$$

where $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable, is called hyperbolic if the matrix $f'(x^*) \in \mathbb{R}^{d\times d}$ is hyperbolic.

Hyperbolicity provide us with a powerful result about the neighbourhood of the equilibrium point of a nonlinear system. Indeed, by the Hartman – Grobman theorem, there exists a homomorphism between the system and the linearisation.

Theorem 8 (Hartman-Grobman Theorem)**.** In a neighbourhood of an hyperbolic equilibrium, there exists a homomorphism between the phase portraits of the nonlinear system and the linearisation.

Proof. Beyond this course. □

Instead of proving the whole theorem which is beyond this course, we shall consider a special case, where $\Sigma(f'(x^*)) \subseteq (0,\infty)$ or $\Sigma(f'(x^*)) \subseteq (-\infty, 0)$.

Lemma 4.1 (Gronwall Lemma)**.** Let $u : [a,b] \to \mathbb{R}$ be continuous and $c, d \geq 0$. Suppose

$$0 \leq u(t) \leq c + d\int_a^t u(s)\mathrm{d}s,$$

for all $t \in [a,b]$. Then, we have the explicit estimate

$$u(t) \leq ce^{d(t-a)},$$

for all $t \in [a,b]$.

Proof. Since u is continuous on a compact interval, it is bounded by some $M > 0$ such that $u(t) \leq M$ for all $t \in [a,b]$. Then,

$$u(t) \leq c + d\int_a^t M\mathrm{d}s = c + dM(t-a).$$

Iterating this, we obtain the sequence

$$u(t) \leq s_n := c\sum_{k=0}^{n-1} \frac{d^k(t-a)^k}{k!} + \frac{Md^n(t-a)^n}{n!},$$

for all $n \in \mathbb{N}$. Thus, by taking $n \to \infty$, the second term converges to 0 and we have

$$u(t) \leq ce^{d(t-a)},$$

for all $t \in [a,b]$. □

Theorem 9 (Linearised Stability)**.** Suppose $f : D \to \mathbb{R}^d$ is continuously differentiable where $D \subseteq \mathbb{R}^d$ is open. Consider the differential equation

$$\dot{x} = f(x),$$

with an equilibrium at $x^* \in D$ such that $\Sigma(f'(x^*)) \subseteq (-\infty, 0)$. Then the equilibrium is exponentially stable.

Proof. By changing the variable to $x - x^*$, we may Wlog. assume that $x^* = 0$.

Consider

$$\dot{x} = f(x) = f'(0)x + r(x),$$

where

$$r(x) := f(x) - f'(0)x = o(\|x\|),$$

that is, we separate x into a linear and a nonlinear term. Clearly, we have $r(0) = 0 = r'(0)$. Suppose we now denote the flow of the differential equation by ϕ. Now, since the real part of the eigenvalues of $f'(0)$ are negative, there exists some $K > 0$ and $\gamma < 0$ such that

$$\|e^{f'(0)t}\| \leq Ke^{\gamma t},$$

for all $t \geq 0$.

Choose $M \in (0, -\gamma/K)$ and we estimate r. Since r is continuously differentiable, there exists $\rho > 0$ such that $\|r'(x)\| \leq M$ for all $x \in \overline{B_\rho(0)}$. Then, by the mean value inequality, we have

$$\|r(x)\| \leq M\|x\|,$$

for all $x \in \overline{B_\rho(0)}$.

Furthermore, we define for each initial value $x \in B_\rho(0)$ the escape time,

$$T_e(x) := \sup\{T > 0 \mid \|\phi(t, x)\| \leq \rho, t \in [0, T)\}.$$

We note that T_e can be ∞ implying the trajectory never escapes the neighbourhood. This is desirable and as we shall see, this is true for small enough x.

Now, I claim that for $x_0 \in B_\rho(0)$,

$$\|\phi(t, x_0)\| \leq Ke^{(KM+\gamma)t}\|x_0\|,$$

for all $t \in [0, T_e(x_0))$. Indeed, since $t \mapsto \phi(t, x_0)$ is a solution to the differentiable equation, we have

$$\dot{x} = f'(0)x + r(\phi(t, x_0)).$$

Then, by the variation of constant formula, we have

$$\phi(t, x_0) = e^{f'(0)t}x_0 + \int_0^t e^{f'(0)(t-s)}r(\phi(s, x_0))\mathrm{d}s.$$

Hence, for all $t \in [0, T_e(x_0))$,

$$\begin{aligned}\|\phi(t, x_0)\| &\leq \|e^{f'(0)t}\|\|x_0\| + \int_0^t \|e^{f'(0)(t-s)}\|\|r(\phi(s, x_0))\|\mathrm{d}s \\ &\leq Ke^{\gamma t}\|x_0\| + \int_0^t Ke^{\gamma t(t-s)}\|r(\phi(s, x_0))\|\mathrm{d}s \\ &\leq Ke^{\gamma t}\|x_0\| + \int_0^t Ke^{\gamma t(t-s)}M\|\phi(s, x_0)\|\mathrm{d}s.\end{aligned}$$

Thus, by defining $u(t) := e^{-\gamma t}\|\phi(s, x_0)\|$, we apply the Gronwall lemma, resulting in the bound

$$u(t) \leq K\|x_0\| + KM\int_0^t u(s)\mathrm{d}s \implies u(t) = e^{-\gamma t}\|\phi(s, x_0)\| \leq K\|x_0\|e^{KMt},$$

for all $t \in [0, T_e(x_0))$ as claimed.

With this inequality at our disposal, by considering the fact that $t \mapsto e^{(KM+\gamma)t}$ is monotonically decreasing, we get

$$\|\phi(t, x_0)\| \leq Ke^{(KM+\gamma)t}\|x_0\| \leq K\|x_0\| \leq \rho,$$

for all $t \geq 0$, and $x_0 \in B_{\rho/K}(0)$. With that, we conclude $T_e(x_0) = \infty$ and the differential equation is exponentially stable. $\square$

Definition 4.4 (Stable and Unstable Sets)**.** Consider the differential equation $\dot{x} = f(x)$ with the associated flow ϕ and let x^* be an equilibrium. Then we define the stable set of x^* as

$$W^s(x^*) := \left\{x \in D \mid \lim_{t\to\infty} \phi(t, x) = x^*\right\},$$

and the unstable set of x^* as

$$W^u(x^*) := \left\{x \in D \mid \lim_{t\to-\infty} \phi(t, x) = x^*\right\}.$$

Due to Hartman-Grobman, we see that in many cases, the resulting stable and unstable sets are locally Euclidean, and hence, they are also refereed to as stable and unstable manifolds.

In the situation that x^* is an attractive equilibrium, $W^s(x^*)$ is called the domain of attraction and in this case $W^s(x^*)$ is open (see problem sheet).

Consider the linear system $\dot{x} = Ax$ where A is hyperbolic. Then, assuming A is diagonalisable, we have

$$\mathbb{R}^d = E_1 \oplus \cdots \oplus E_q,$$

where E_i are the eigenspace corresponding to the eigenvalue s_i. Then, we find

$$W^s(0) = \bigoplus_{i=1,\cdots q, s_i<0} E_i \quad \text{and} \quad W^u(0) = \bigoplus_{i=1,\cdots q, s_i>0} E_i,$$

and hence, we have $\mathbb{R}^d = W^s(0) \oplus W^u(0)$.

Definition 4.5 (Invariant)**.** A set $M \subseteq D$ is called

- positively invariant if for all $x \in M$, $\sigma^+(x) \subseteq M$;
- negatively invariant if for all $x \in M$, $\sigma^-(x) \subseteq M$;
- invariant if M is both positively and negatively invariant.

Straight away, we see that the stable and unstable manifolds are positively and negatively invariant respectively. Other examples of invariant manifolds are the singleton set $\{x^*\}$ where x^* is an equilibrium; periodic orbits; homoclinic/heteroclinic orbits, or just orbits in general.

4.3 Limit Sets

Consider the differential equation

$$\dot{x} = f(x),$$

where $f : D \to \mathbb{R}^d$ is locally Lipschitz continuous. Let us denote its flow by ϕ.

Definition 4.6 (Limit Set)**.** Fix $x \in D$, a point $x_\omega \in D$ is called an ω-limit point of x if there exists a sequence $(t_n)_{n \in \mathbb{N}} \to \infty$ such that

$$x_\omega = \lim_{n \to \infty} \phi(t_n, x),$$

and we call $\omega(x)$ for the set of all ω-limit points.

On the other hand, a point $x_\alpha \in D$ is called an α-limit point of x if there exists a sequence $(t_n)_{n \in \mathbb{N}} \to -\infty$ such that

$$x_\alpha = \lim_{n \to \infty} \phi(t_n, x),$$

and we call $\alpha(x)$ for the set of all α-limit points.

Straight away, if D is bounded, we have $\alpha(x) = \omega(x) = \varnothing$ for all $x \in D$.

Proposition 20. For $x \in D$, we have

$$\omega(x) = \bigcap_{t \geq 0} \overline{\sigma^+(\phi(t, x))} \text{ and } \alpha(x) = \bigcap_{t \leq 0} \overline{\sigma^-(\phi(t, x))}.$$

Intuitively, this characterisation of the limit set is the intersection of a sequence of nested orbits. Indeed, of an element to be an ω-limit point, it must remain in the orbit regardless of the starting time. However, as we may approach a limit point without ever reaching it, we take the closure of each orbit.

Proof. By symmetry, it suffices to prove the first statement.

Let $y \in \omega(x)$, then there exists some $t_n \to \infty$ such that $\lim_{n \to \infty} \phi(t_n, x) = y$. Since $t_n \to \infty$, for all $t \geq 0$, we have t_n restricted onto $[t, \infty)$ is a sequence such that $t_n \mid_{[t,\infty)} \to \infty$. Thus, as $\lim_{n \to \infty} \phi(t_n \mid_{[t,\infty)}, x) = y$, we have $y \in \overline{\sigma^+(\phi(t, x))}$ for all $t \geq 0$. Thus, $y \in \bigcap_{t \geq 0} \overline{\sigma^+(\phi(t, x))}$ and $\omega(x) \subseteq \bigcap_{t \geq 0} \overline{\sigma^+(\phi(t, x))}$.

On the other hand, if $y \in \bigcap_{t\geq 0} \overline{\sigma^+(\phi(t,x))}$, for all $t \geq 0$, $y \in \overline{\sigma^+(\phi(t,x))}$, and so, there exists some sequence $s_{t,n} \geq t$ such that $\lim_{n\to\infty} \phi(s_{t,n}, x) = y$. Thus, by choosing $t := n$, $s_n := s_{n,n} \to \infty$ since $s_n \geq n$, and $\lim_{n\to\infty} s_n = y$. With that, the equality is established. $\square$

Proposition 21. Fix $x \in D$, then

- the ω-limit set $\omega(x)$ is invariant. In addition, if $\sigma^+(x)$ is bounded and $\overline{\sigma^+(x)} \subseteq D$, then $\omega(x)$ is non-empty and compact.
- the α-limit set $\alpha(x)$ is invariant. In addition, if $\sigma^+(x)$ is bounded and $\overline{\sigma^-(x)} \subseteq D$, then $\alpha(x)$ is non-empty and compact.

Proof. By symmetry, it suffices to prove the first statement.

Suppose $\sigma^+(x)$ is bounded and $\overline{\sigma^+(x))} \subseteq D$, then, any sequence $(\phi(n,x))$ is bounded, and so, by Bolzanno-Weierstrass, there exists some (n_k), such that $\phi(n_k, x) \to x_\omega$ as $k \to \infty$. Hence, $x_\omega \in \omega(x)$ and thus, its nonempty.

By the above proposition, we have

$$\omega(x) = \bigcap_{t\geq 0} \overline{\sigma^+(\phi(t,x))},$$

which is closed, and so, by Heine-Borel, it suffices to show boundedness. However, as we have assumed $\sigma^+(x)$ is bounded, we have $\overline{\sigma^+(x)}$ is bounded and thus $\omega(x)$ is bounded and hence compact.

Lastly, we show that $\omega(x)$ is invariant. Let $x_0 \in \omega(x)$, and it suffices to prove $\sigma(x_0) \subseteq \omega(x)$. Since $x_0 \in \omega(x)$, there exists some $t_n \to \infty$ such that $\phi(t_n, x) \to x_0$ as $n \to \infty$. Then, for all $\tau \in J_{\max}(x_0)$, we can define $s_n = t_n + \tau$ and so,

$$\begin{aligned}\lim_{n\to\infty} \phi(x_n, a) = \lim_{n\to\infty} \phi(t_n + \tau, x) &= \lim_{n\to\infty} \phi(\tau, \phi(t_n, x)) \\ &= \phi(\tau, \lim_{n\to\infty} \phi(t_n, x)) = \phi(\tau, x_0),\end{aligned}$$

and hence, $\omega(x)$ is invariant. $\square$

Theorem 10 (Poincaré-Bendixson Theorem)**.** Consider the two-dimensional differential equation $\dot{x} = f(x)$, where $f : D \to \mathbb{R}^2$ is continuously differentiable on the open domain D. Now, if there exists some $x \in D$ such that $\sigma^+(x) \subseteq K \subseteq D$ where K is compact which contains finitely many equilibrium. Then, one of the following three statements hold of the ω-limit set $\omega(x)$:

- $\omega(x)$ is a singleton consisting of an equilibrium;
- $\omega(x)$ is a periodic orbit;
- $\omega(x)$ consists of equilibria and non-closed orbits. In particular, the non-closed orbits in $\omega(x)$ converge forward and backward in time to equilibria in $\omega(x)$, i.e. they are either homoclinic or heteroclinic orbits.

Proof. See extra material. □

We note that an analogous statement is also true for α-limit sets. The Poincaré-Bendixson theorem tells us that the limit sets are regular in two-dimensional differential equations (also regular in the one-dimensional case). This is not the case for higher dimensions, and in fact, for three-dimensions, chaotic differential equations can occur, e.g. the Lorenz system.

The following corollary is often used to prove the existence of a periodic orbit.

Corollary 10.1. Suppose that for some $x \in D$, the positive half-orbit $\sigma^+(x)$ lies in a compact subset $K \subseteq D$ that does not contain an equilibrium. Then $\omega(x)$ is a periodic orbit.

4.4 Lyapunov Functions

As we have previously seen, we may analyse the properties of an equilibrium without solving the differential equation. Indeed, if the equilibrium x^* is hyperbolic, then we understand local stability of x^*. However, we might like to understand a more global property, motivating us to study the Lyapunov functions.

Consider the differential equation $\dot{x} = f(x)$ and $f : D \to \mathbb{R}^d$ is locally Lipschitz (note we are not assuming continuous differentiability). Let $V : D \to \mathbb{R}$ be continuously differentiable and consider a solution $\mu : I \to D$ of $\dot{x} = f(x)$, then the derivative of V along μ is defined to be

$$\frac{\mathrm{d}}{\mathrm{d}t} V(\mu(t)) = V'(\mu(t)) \cdot \dot{\mu}(t) = V'(\mu(t)) \cdot f(\mu(t)).$$

Definition 4.7 (Orbital Derivative)**.** The orbital derivative of V is given by

$$\dot{V}(x) = V'(x) \cdot f(x) = \sum_{i=1}^{d} \frac{\partial V}{\partial x_i} f_i(x).$$

Here, the row vector $V'(x) \in \mathbb{R}^{1\times d}$ is the gradient of V at x.

Definition 4.8 (Lyapunov Function)**.** Consider the differential equation $\dot{x} = f(x)$ where $f : D \to \mathbb{R}^d$ is locally Lipschitz and let $V : D \to \mathbb{R}$ be a continuous differentiable function. Then, V is called a Lyapunov function if

$$\dot{V}(x) \leq 0, \text{ for all } x \in D.$$

We see that any Lyapunov function decreases along solutions, i.e. $V(\phi(t, x)) \leq V(x)$, for all $t \in [0, \sup J_{\max}(x))$. Indeed, for all $t \in [0, \sup J_{\max}(x))$, we have

$$V(\phi(t, x)) - V(\phi(0, x)) = \int_0^t \dot{V}(\phi(s, x)) \mathrm{d}s.$$

Since $\dot{V} \leq 0$, the integral is non-positive and so V is decreasing on ϕ.

Proposition 22. If V is a Lyapunov function, then the sublevel set of the form

$$S_c := \{x \in D \mid V(x) \leq c\}$$

is positively invariant.

Proof. Assume S_c is not positively invariant. Then there exists an $x \in S_c$ and $t > 0$ such that $\phi(t, s) \notin S_c \implies V(\phi(t, x)) > c$. However, by the fact that the Lyapunov function decreases along functions, and $x \in S_c$, $c \geq V(x) \geq V(\phi(t, x))$ #. $\square$

Theorem 11 (Lyapunov's Direct Method for Stability)**.** Consider an equilibrium $x^* \in D$ and let $V : D \to \mathbb{R}$ be a Lyapunov function. If $V(x^*) = 0$ and $V(x) > 0$ for all $x \in D \setminus \{x^*\}$, then, x^* is stable.

Proof. Let $\epsilon > 0$, then we will find some $\delta > 0$ such that for all $x_0 \in B_\delta(x^*)$, $\phi(t, x_0) \in B_\epsilon(x^*)$ for all $t \geq 0$. Wlog. since D is open, we may assume $\overline{B_\epsilon(x^*)} \subseteq D$. Then, we may define

$$m := \min\{V(x) \mid \|x - x^*\| = \epsilon\} > 0,$$

which is positive since the set $S := \{x \mid \|x - x^*\| = \epsilon\}$ is compact, and so, since V is continuous it achieves a minimum on S which is positive. Furthermore, by continuity of V, there exists a $\delta \in (0, \epsilon)$ such that

$$0 \leq V(x) \leq \frac{m}{2},$$

for all $x \in B_\delta(x^*)$. Since, a Lyapunov function is invariant forward in time, we have,

$$V(\phi(t, x)) \leq V(x) \leq \frac{m}{2},$$

for all $x \in B_\delta(x^*)$ and $t \geq 0$. But this implies that the orbit forward in time will never cross S since if that is the case, there exists some $t \geq 0$, $x \in B_\delta(x^*)$, such that $V(\phi(t, x)) \geq m$. So, we have established the stability of x^*. $\square$

The existence of Lyapunov functions also provides us with information about the location of ω-limit sets.

Theorem 12 (La Salle's Invariance Principle)**.** Let $V : D \to \mathbb{R}$ be a Lyapunov function. Then,

$$\omega(x) \subseteq \{y \in D \mid \dot{V}(y) = 0\},$$

for all $x \in D$.

Intuitively, since $\dot{V} \leq 0$, if $y \in \omega(x)$ and $\dot{V}(y) < 0$, then once we leave y, we may never approach it again. However, $y \in \omega(x)$ implies that there is a sequence of times such that the orbit approaches y, hence, $\dot{V}(y)$ must equal 0.

Proof. Assume to the contrary that $z \in \omega(x)$ such that $\dot{V}(z) < 0$. Then for some $\tau > 0$, we have $V(\phi(\tau, z)) < V(z)$. Since $z \in \omega(x)$, there exists a sequence $(t_n)_{n=1}^\infty \to \infty$ such that

$$\lim_{n \to \infty} \phi(t_n, x) = z.$$

Wlog. since $t_n \to \infty$ we may assume that, for all n, $t_{n+1} - t_n > \tau$. Due to the fact that $\dot{V}(y) \leq 0$, we have

$$V(\phi(t_{n+1}, x)) \leq V(\phi(t_n + \tau, x)) = V(\phi(\tau, \phi(t_n, x))),$$

for all $n \in \mathbb{N}$. Hence, taking $n \to \infty$ we have $V(z) \leq V(\phi(\tau, z))$ which contradicts $V(\phi(\tau, z)) < V(z)$ and hence, our claim holdes. $\square$

To see why this theorem is called an invariance principle, consider the following reformulation.

Corollary 12.1. Let $V : D \to \mathbb{R}$ be a Lyapunov function. Thus, for any $x \in D$, $\omega(x)$ is contained in the largest invariant subset of $\{y \in D \; mid \dot{V}(y) = 0\}$. Here the largest invariant subset is obtained by taking the union of all invariant subsets.

Proof. Follows straight away since $\omega(x)$ is invariant. □

Theorem 13 (Lyapunov's Direct Method for Asymptotic Stability)**.** Consider the differential equation with an equilibrium $x^* \in D$ and let $V : D \to \mathbb{R}$ be a Lyapunov function such That

- $V(x^*) = 0$ and $V(x) > 0$ for all $x \in D \setminus \{x^*\}$;
- $\dot{V}(x^*) = 0$ and $\dot{V}(x) < 0$ for all $x \in D \setminus \{x^*\}$.

Then, the equilibrium x^* is asymptotically stable.

Proof. The first condition ensures x^* is stable and so, if $\epsilon > 0$ such that $B_{2\epsilon}(x^*) \subseteq D$, there exists some $\delta > 0$ such that for all $x \in B_\delta(x^*)$ we have $\phi(t, x) \in B_\epsilon(x^*)$ where $t \geq 0$. Let us now fix $x \in B_\delta(x^*)$, and since $\phi(t, x) \in \overline{B_\epsilon(x^*)} \subseteq D$ for all $t \geq 0$, we obtain $\overline{\sigma^+(x)} \subseteq D$. Thus, by the properties of ω-limit sets, $\omega(x)$ is compact and non-empty. Now, by La Salle's principle, we have $\omega(x) \subseteq \{x^*\}$ and hence, $\omega(x) = \{x^*\}$.

We will now show attractivity in the δ-neighbourhood of x^*. Suppose otherwise, then there exists some $x \in B_\delta(x^*)$, such that $\phi(t, x) \not\to x^*$ as $t \to \infty$. Then, there exists some $\eta > 0$, $(t_n)_{n \in \mathbb{N}} \to \infty$, such that $\|\phi(t_n, x) - x^*\| \geq \eta$ for all $n \in \mathbb{N}$. Clearly, since $\phi(t_n, x)$ bounded insider ∂D, it contains a limit point in D. This limit point is thus an ω-limit point of x which contradicts the fact that $\omega(x) = \{x^*\}$. □

Corollary 13.1. Assuming the conditions for the above theorem is satisfied for the Lyapunov function V. Considering the sublevel sets of the Lyapunov function, which are in the form

$$S_c := \{x \in D \mid V(x) \leq c\},$$

where $c > 0$. Then, S_c is a subset of the domain of attraction $W^s(x^*)$ if S_c is compact.

Proof. Let S_c be compact and $x \in S_c$ and we need to show that $\lim_{t \to \infty} \phi(t, x) = x^*$. Since S_c is compact, we have $\overline{\sigma^+(x)} \subseteq D$ and $\omega(x) \neq \varnothing$. By La Salle's principle, we have $\omega(x) \subseteq \{x \in D \mid \dot{V}(x) = 0\} = \{x^*\}$, and so, $\omega(x) = \{x^*\}$. Then, the argument in the second part of the above theorem finishes the proof. □