MA2AA1 (ODE’s): Lecture Notes

Sebastian van Strien (Imperial College)

Spring 2018

Contents

1 Introduction

1.1 Practical Arrangement

- The lectures for this module will take place **Monday 10-12** and **Tuesday 3-4** in Clore.

- Each week I will hand out a sheet with problems. It is very important you go through these thoroughly, as these will give the required training for the exam and class tests.

- Support classes: **Tuesday 4-5, from January 26**.

- The **support classes** will be run rather differently from previous years. The objective is to make sure that you will get a lot out of these support classes.

Attend problem classes!

- The main way to revise for the tests and the exam is by doing the exercises.

- There will be two **class tests**. These will take place on Friday week 19 and and 25 (i.e. the 5th and 11th week of this term). Each of these count for 5% .

- **Questions are most welcome**, during or after lectures and during office hour.

- I aim to do ask questions and do problems in class, and use clickers on your mobile `https://www.menti.com` for getting feedback.

- My office hour is to be agreed with students reps. Office hour will in my office 6M36 Huxley Building.

1.2 Relevant material

- There are many books which can be used in conjunction to the module, but **none are required**.

- The **lecture notes** displayed during the lectures will be posted on blackboard.

- The lectures will also be recorded on panopto.

- There is absolutely no need to consult any book. However, recommended books are

  - Simmons + Krantz, *Differential Equations: Theory, Technique, and Practice*, about 40 pounds. This book covers a significant amount of the material we cover. Some students will love this text, others will find it a bit longwinded.
  - Agarwal + O'Regan, *An introduction to ordinary differential equations*.
  - Teschl, *Ordinary Differential Equations and Dynamical Systems*. These notes can be downloaded for free *from the authors webpage*.
  - Hirsch + Smale (or in more recent editions): Hirsch + Smale + Devaney, *Differential equations, dynamical systems, and an introduction to chaos*.
  - Arnold, *Ordinary differential equations*. This book is an absolute jewel and written by one of the masters of the subject. It is a bit more advanced than this course, but if you consider doing a PhD, then get this one. You will enjoy it.

  Quite a few additional exercises and lecture notes can be freely downloaded from the internet.

1.3 Notation and aim of this course

Notation:

- $\dot{x}$ will ALWAYS mean $\dfrac{dx}{dt}$
- y' usually means $\dfrac{dy}{dx}$ but also sometimes $\dfrac{dy}{dt}$; which one will *always* be clear from the context.

This course is about studying differential equations of the form

$$\dot{x} = f(x), \text{ resp. } \dot{y} = g(t, y),$$

- This is short for finding a function $t \mapsto x(t)$ resp. $t \mapsto y(t)$ so that

  $$\frac{dx}{dt} = f(x(t)) \text{ resp. } \frac{dy}{dt} = g(t, y(t)).$$

  In particular this means that (*in this course*) **we will assume that** $t \mapsto x(t)$ **differentiable**.

- In **ODE's** *the independent variable* (usually the t) is *one-dimensional.* In a *Partial Differential Equation* **(PDE)** such as

  $$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0$$

  the unknown function u is differentiated w.r.t. several variables.

Aim of this course is to find out when or whether such an equation has a solution and determine its properties.

1.4 Examples of differential equations

- The oldest example of a differential equation is the law of Newton: $m\ddot{x}(t) = F(x(t)) \quad \forall t$. Here F is the gravitational force. Using the gravitational force in the vicinity of the earth, we approximate this by

  $$m\ddot{x}_1 = 0, m\ddot{x}_2 = 0, m\ddot{x}_3 = -g.$$

  This has solution

  $$x(t) = x(0) + v(0)t - \frac{g}{2m}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} t^2.$$

- According to Newton's law, the gravitational pull between two particles of mass m and M is $F(x) = \gamma m M x/|x|^3$. This gives

  $$m\ddot{x}_i = -\frac{\gamma m M x_i}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} \text{ for } i = 1, 2, 3$$

- For three bodies we have something similar. Now it is no longer possible to explicitly solve this equation.

- Poincaré showed this in 1887 in his famous solution to a prize competition posed by the king of Sweden: `https://en.wikipedia.org/wiki/Henri_Poincare#Three-body_problem`).

- Nowadays we know that solutions can be chaotic.

- For a similar system, see also `https://www.youtube.com/watch?v=AwT0k09w-jw`.

- One needs some theory be sure that there are solutions and that they are unique.

Consider the **initial value problem** (I.V.P.):

$$\frac{dx}{dt} = f(t,x) \text{ and } x(0) = x_0$$

where $f\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$.
Does this I.V.P. have solutions? Are these solutions unique?

1.4.1 There might NOT exist any solution

Example 1.1

Consider $x' = f(x)$ *where*

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

Then there exists **no solution** *to the initial value*

$$\dot{x} = f(x), x(0) = 0.$$

Question: Is

$$x(t) = \begin{cases} t & \text{for } t \leq 0 \\ -t & \text{for } t > 0 \end{cases}$$

a solution?

1.4.2 Solutions may NOT be unique

Example 1.2

Take

$$\dot{x} = |x|^{1/2}, x(0) = 0.$$

Then $x(t)$ defined by

$$x(t) = \begin{cases} 0 & \text{for } t \leq 1 \\ (t-1)^2/4 & \text{for } t \geq 1 \end{cases}$$

is a solution.

Question: Can you think of other solutions? How many solutions has this I.V.P?

1.4.3 Physical meaning of non-uniqueness

Example 1.3

Consider a cylindrical bucket of cross sectional area A has a small hole of area a at the bottom. It is filled with water of height $h(t)$ which leaks out of the bucket under the effect of Earth's gravitational field, g. The equation becomes

$$\dot{h} = -\frac{a}{A}\sqrt{2gh}.$$

The solution of this equation is non-unique: this follows from the previous example.

Question: Why is the non-uniqueness not surprising: consider the problem/equation with backward time.

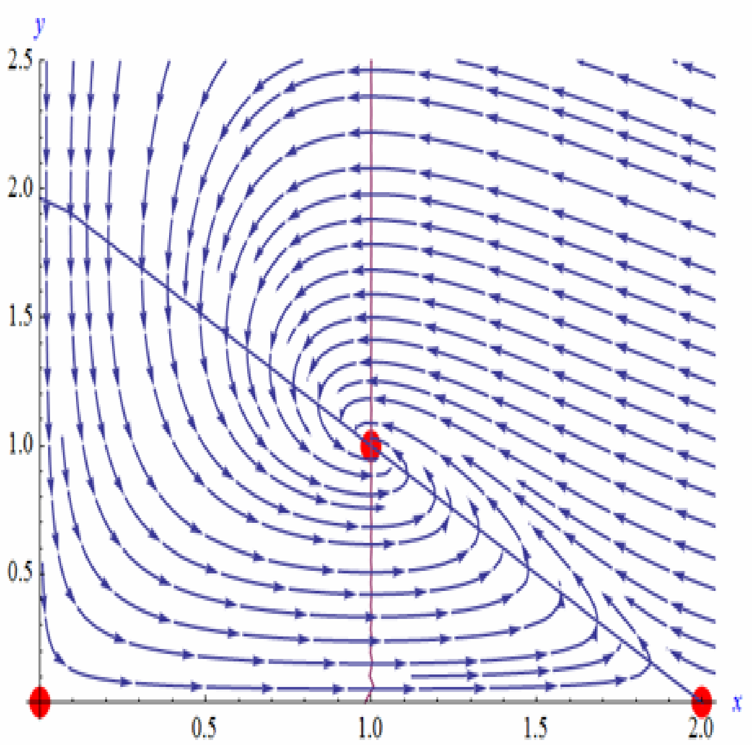
1.4.4 Uniqueness and determinism

Having a unique solution of the (I.V.P.)

$$\frac{dx}{dt} = f(t, x) \text{ and } x(0) = x_0$$

means that this problem is deterministic.

An example where f does not depend on t and for which $f : \mathbb{R}^2 \to \mathbb{R}^2$ is drawn below:



- An example of an ODE related to vibrations of bridges (or springs) is the following (see Appendix C, Subsection C.7):

$$Mx'' + cx' + kx = F_0 \cos(\omega t).$$

  One reason you should want to learn about ODE's is:

  - `http://www.ketchum.org/bridgecollapse.html`
  - `http://www.youtube.com/watch?v=3mclp9QmCGs`
  - `http://www.youtube.com/watch?v=gQK21572oSU`

- This is related to synchronisation phenomena, for example of a large number of metronomes, see for example `https://www.youtube.com/watch?v=5v5eBf2KwF8`

1.5 Issues which will be addressed in the course include:

- do solutions of ODE's exist?
- are they unique?
- most differential equations, cannot be solved explicitly. One aim of this course is to develop methods which makes it possible to obtain information on the behaviour of solutions anyway.

2 The Banach fixed point theorem

In this chapter we will present a general method for proving that solutions to certain equations exist. This method even gives a method for finding approximations of these solutions.

X will denote the space in which the solution is supposed to lie. Sometimes we will take $X = \mathbb{R}$ or $X = \mathbb{R}^n$, but in the next chapter we will apply this chapter to proving solutions of differential equations exist, and X will be a space of functions (so an infinite dimensional space).

2.1 Banach spaces

Banach spaces are just generalisations of $\mathbb{R}^n$.

- A **vector space** X is a space so that if $v_1, v_2 \in X$ then $c_1 v_1 + c_2 v_2 \in X$ for each $c_1, c_2 \in \mathbb{R}$ (or, more usually, for each $c_1, c_2 \in \mathbb{C}$).

- A **norm** on X is a map $||\cdot||\colon X \to [0,\infty)$ so that

  1. $||0|| = 0, ||x|| > 0\ \forall x \in X \setminus \{0\}$.
  2. $||cx|| = |c|||x||\ \forall c \in \mathbb{R}$ and $x \in X$
  3. $||x+y|| \leq ||x||+||y||\ \forall x, y \in X$ (triangle inequality).

- A **Cauchy sequence** in a vector space with a norm is a sequence $(x_n)_{n\geq 0} \in X$ so that for each $\epsilon > 0$ there exists N so that $||x_n - x_m|| \leq \epsilon$ whenever $n, m \geq N$.

- A vector space with a norm is **complete** if each Cauchy sequence $(x_n)_{n\geq 0}$ converges, i.e. there exists $x \in X$ so that $||x_n - x|| \to 0$ as $n \to \infty$.

- X is a **Banach space** if it is a vector space with a norm which is complete.

2.2 Metric spaces

Reminder from analysis II:

- A **metric space** X is a space with together with a function $d\colon X \times X \to \mathbb{R}^+$ (called metric) so that
  1. $d(x,x) = 0$ and $d(x,y) = 0$ implies $x = y$.
  2. $d(x,y) = d(y,x)$
  3. $d(x,z) \le d(x,y) + d(y,z)$ (triangle inequality).
- A sequence $(x_n)_{n\ge 0} \in X$ is called **Cauchy** if for each $\epsilon > 0$ there exists N so that $d(x_n, x_m) \le \epsilon$ whenever $n, m \ge N$.
- The metric space is **complete** if each Cauchy sequence $(x_n)_{n\ge 0}$ converges, i.e. there exists $x \in X$ so that $d(x_n, x) \to 0$ as $n \to \infty$.

2.3 Metric space versus Banach space

- Given a norm $||\cdot||$ on a vector space X one can also define the metric $d(x,y) = ||y - x||$ on X. So a Banach space is automatically a metric space.
- A metric space is **not** necessarily a Banach space. For example the set $S^2 = \{x \in \mathbb{R}^3; |x| = 1\}$, together with the Euclidean metric is a metric space, but not a vector space.

2.4 Examples

Example 2.1

Consider $\mathbb{R}$ *with the norm* $|x|$*. You have seen in Analysis I that this space is complete.*

In the next two examples we will consider $\mathbb{R}^n$ with two different norms. As is usual in year ≥ 2, we write $x \in \mathbb{R}^n$ rather than $\underline{x}$ for a vector.

Example 2.2

Consider the space $\mathbb{R}^n$ *and define* $|x| = \sqrt{\sum_{i=1}^n x_i^2}$ *where* x *is the vector* $(x_1, \dots, x_n)$*. It is easy to check that* $|x|$ *is a norm (the main point to check is the triangle inequality). This norm is usually referred to as the Euclidean norm (as* $d(x, y) = |x - y|$ *is the Euclidean distance).*

Example 2.3

Consider the space $\mathbb{R}^n$ *and the supremum norm* $|x|_\infty := \max_{i=1}^n |x_i|$ *(it is easy to check that this is a norm).*

Regardless which of these two norms we put on $\mathbb{R}^n$, in both cases the space we obtain becomes a complete metric space (this follows from Example 1).

Without saying this explicitly everywhere, in this course, we will always endow $\mathbb{R}^n$ with the Euclidean metric. In other courses, you will also come across other norms on $\mathbb{R}^n$ (for example the l^p norm $(\sum_{i=1}^n |x_i|^p)^{1/p}$, $p \geq 1$.

Example 2.4

One can define several norms on the space of $n \times n$ matrices. One, which is often used, is the **matrix norm**

$$||A|| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|Ax|}{|x|}$$

when A is a $n \times n$ matrix. Here x, Ax are vectors and $|Ax|, |x|$ are the Euclidean norms of these vectors. By linearity of A we have

$$\sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|Ax|}{|x|} = \sup_{x \in \mathbb{R}^n, |x|=1} |Ax| \tag{1}$$

and so the latter also defines $||A||$. In particular, since $x \mapsto Ax$ is continuous, $||A||$ is a finite real number.

Question: Why does $\leq$ and $\geq$ hold in (1)?
Question: Why is $A \mapsto ||A||$ a norm?

Now we will consider a compact interval I and the vector space $C(I, \mathbb{R})$ of continuous functions from I to $\mathbb{R}$. In the next two examples we will put two different norms on $C(I, \mathbb{R})$. In one case, the resulting vector is complete and in the other it is not.

Example 2.5

The set $C(I, \mathbb{R})$ endowed with the **supremum norm** $||x||_\infty = \sup_{t \in I} |x(t)|$ *is a Banach space. That $|| \cdot ||_\infty$ is a norm is easy to check, but the proof that $||x||_\infty$ is complete is more complicated and will not proved in this course (this result is shown in the metric spaces course).*

Example 2.6

The space $C([0,1], \mathbb{R})$ endowed with the L^1 norm $||x||_1 = \int_0^1 |x(s)|\, ds$ is **not** *complete.*

Remark: The previous two examples show that the same set can be complete w.r.t. one norm (or metric) and incomplete w.r.t. to another norm (or metric).

For completeness, let us prove the assertion in Example 2.6 that the norm $||x||_1$ is *not* complete.

This page will be skipped in the lectures, and is not examinable.

This proof is *non-examinable* and will *not be covered in class*).

To prove that the norm $||x||_1$ is not complete, we find a Cauchy sequence (w.r.t. this norm) which does not converge. Let us choose for this sequence the functions

$$x_n(s) = \begin{cases} \min(\sqrt{n}, 1/\sqrt{s}) & \text{for } s > 0 \\ \sqrt{n} & \text{for } s = 0. \end{cases}$$

This sequence of functions is Cauchy w.r.t. the $||\cdot||_1$ norm: take $m > n$ then $\int_0^1 |x_n(s) - x_m(s)|\, ds = \int_0^{1/m} |\sqrt{m} - \sqrt{n}|\, ds + \int_{1/m}^{1/n} |1/\sqrt{s} - \sqrt{n}|\, ds \le 1/\sqrt{m} + 2/\sqrt{n} \le 3/\sqrt{n} \to 0$.

Let us now show by contradiction that this Cauchy sequence does not converge: assume that there exists continuous function $x \in C([0,1], \mathbb{R})$ so that $||x - x_n||_1$ converges to zero. Since x is continuous, there exists k so that $|x(s)| \le \sqrt{k}$ for all s. Then it is easy to show that $||x_n - x||_1 \ge 1/(2\sqrt{k}) > 0$ when n is large:

Indeed, for $n \ge k$ and $s \in [0, 1/k)$, we have $x_n(s) - x(s) \ge x_n(s) - \sqrt{k} > 0$. Hence $||x_n - x||_1 \ge \int_0^{1/k} |x_n(s) - x(s)| ds \ge \int_0^{1/k} x_n(s) - (1/k)\sqrt{k} = \int_0^{1/n} x_n(s) + \int_{1/n}^{1/k} x_n(s) - (1/k)\sqrt{k} = (1/n)\sqrt{n} + (2/\sqrt{k} - 2/\sqrt{n}) - (1/k)\sqrt{k} \ge 1/(2\sqrt{k})$ when n is large.

Hence $||x_n - x||_1 \ge 1/(2\sqrt{k}) > 0$ and therefore the Cauchy sequence x_n does not converge.

2.5 Banach Fixed Point Theorem

Theorem 2.7 (*Banach Fixed Point Theorem*)

Let X be a complete metric space and consider $F\colon X \to X$ so that there exists $\lambda \in (0,1)$ so that

$$d(F(x),F(y)) \leq \lambda d(x,y) \text{ for all } x,y \in X$$

Then F has a unique fixed point p,

$$F(p) = p.$$

and for every $x_0 \in X$ the sequence defined by $x_{n+1} = F(x_n)$ converges to this p.

Proof: (**Existence**) Take $x_0 \in X$ and define $(x_n)_{n\geq 0}$ by $x_{n+1} = F(x_n)$. This is a Cauchy sequence:

$$d(x_{n+1},x_n) = d(F(x_n),F(x_{n-1})) \leq \lambda d(x_n,x_{n-1}).$$

Hence for each $n \geq 0$, $d(x_{n+1},x_n) \leq \lambda^n d(x_1,x_0)$. Therefore when $n \geq m$,

$$d(x_n,x_m) \leq d(x_n,x_{n-1}) + \cdots + d(x_{m+1},x_m) \leq$$

$$\leq (\lambda^{n-1} + \cdots + \lambda^m)d(x_1,x_0) \leq \lambda^m/(1-\lambda)d(x_1,x_0).$$

So $(x_n)_{n\geq 0}$ is a Cauchy sequence and has a limit p, i.e. $d(x_n,p) \to 0$.

By the triangle inequality and the fact that F is Lipschitz, $d(F(p),p) \leq d(F(p),F(x_n)) + d(F(x_n),p) \leq \lambda d(x_n,p) + d(x_{n+1},p) \to 0$. Hence $F(p) = p$.

(**Uniqueness**) If $F(p) = p$ and $F(q) = q$ then $d(p,q) = d(F(p),F(q)) \leq \lambda d(p,q)$. Since $\lambda \in (0,1)$, $\quad p = q$. ■

Remark: Since a Banach space is also a complete metric space, the previous theorem also holds for a Banach space.

Example 2.8

Let $g\colon [0,\infty) \to [0,\infty)$ *be defined by* $g(x) = (1/2)e^{-x}$*. Then* $|g'(x)| = |(1/2)e^{-x}| \le 1/2$ *for all* $x \ge 0$ *and so there exists a unique* $p \in \mathbb{R}$ *so that* $g(p) = p$*. (By the Mean Value Theorem* $\dfrac{g(x)-g(y)}{x-y} = g'(\zeta)$ *for some* ζ *between* x, y*. Since* $|g'(\zeta)| \le 1/2$ *for each* $\zeta \in [0,\infty)$ *this implies that* g *is a contraction. Also note that* $g(p) = p$ *means that the graph of* g *intersects the line* $y = x$ *at* (p,p)*.)*

Example 2.9

Find a sequence x_n which converges to $\sqrt{a}$ by Newton's method.

To do this, take $f(x) = x^2 - a$, take x_n close to $\sqrt{a}$ and choose x_{n+1} as the root of the linear approximation $L_n(x) = f(x_n) + f'(x_n)(x - x_n)$ of f at x_n.

Question: *draw the functions f and L_n.*

$L_n(x_{n+1}) = 0$ *gives*

$$x_{n+1} = x_n - [f'(x_n)]^{-1} f(x_n).$$

In this particular instance, this expression takes the form

$$x_{n+1} = x_n - \frac{x^2 - a}{2x_n} = \frac{1}{2}(x_n + \frac{a}{x_n}).$$

Is

$$T(x) := \frac{1}{2}(x + \frac{a}{x})$$

a contraction and on what space X?

Question: *Draw the graph of T and show that $T\colon [\sqrt{a}, \infty) \to [\sqrt{a}, \infty)$, so T maps $X = [\sqrt{a}, \infty)$ into itself.*

$T\colon [\sqrt{a}, \infty) \to [\sqrt{a}, \infty)$ is a contraction since

$$|T(x) - T(y)| = \frac{1}{2}|x + \frac{a}{x} - y - \frac{a}{y}| = \frac{1}{2}|1 - \frac{a}{xy}||x - y|$$

and since $0 \le 1 - \dfrac{a}{xy} \le 1$ for $x, y \ge \sqrt{a}$.

2.6 Lipschitz functions

Let X be a Metric space. Then we say that a function $f\colon X \to X$ is **Lipschitz** if there exists $K > 0$ so that

$$d(f(x), f(y)) < Kd(x, y).$$

Example 2.10

Let A be a $n \times n$ matrix. Then $\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^n$ is Lipschitz. Indeed, $|Ax - Ay| \leq K|x - y|$ where K is the **matrix norm** *of A defined by $||A|| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \dfrac{|Ax|}{|x|}$. Remember that $||A||$ is also equal to $\max_{x \in \mathbb{R}^n; |x|=1} |Ax|$.*

Example 2.11

The function $\mathbb{R} \ni x \mapsto x^2 \in \mathbb{R}$ is not Lipschitz: there exists no constant K so that $|x^2 - y^2| \leq K|x - y|$ for all $x, y \in \mathbb{R}$.

Example 2.12

On the other hand, the function $[0, 1] \ni x \mapsto x^2 \in [0, 1]$ is Lipschitz.

Example 2.13

The function $[0, 1] \ni x \mapsto \sqrt{x} \in [0, 1]$ is not Lipschitz.

2.7 The Multivariable Mean Value Theorem

To check that a function is Lipschitz the following theorem is often useful.

Theorem 2.14 (***Multivariable Mean Value Theorem***)

If

$$f\colon \mathbb{R}^n \to \mathbb{R}^m$$

is continuously differentiable then $\forall x, y \in \mathbb{R}^n$ *there exists* $\xi \in [x,y]$ *(where* $[x,y]$ *is the line connecting* x *and* y*) so that* $|f(x) - f(y)| \leq ||Df_\xi|| |x-y|$.

Proof : See Appendix A. ■

Here Df_ξ is the Jacobian matrix of f and ξ and $||Df_\xi||$ is the norm of this matrix.

Question: Why do you have $=$ in the above theorem when $n = 1$ and (in general) $\leq$ when $n > 1$?

This theorem implies in particular that if $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable, then for each $R > 0$ there exists K so that

$$|f(x) - f(y)| \leq K|x-y| \text{ for all } |x|, |y| \leq R.$$

This theorem also implies:

Corollary 2.15

Let U *be an open set in* $\mathbb{R}^n$ *and* $f\colon U \to \mathbb{R}$ *be continuously differentiable. Then* $f\colon C \to \mathbb{R}$ *is Lipschitz for any compact set* $C \subset U$*. When* $n = 1$ *this follows from the Mean Value Theorem, and for* $n > 1$ *this follows from the above theorem.*

2.8 The usefulness of the Banach contraction theorem

The Banach fixed point theorem is used in all branches of mathematics, both pure and applied. Many people refer to it as the Banach contraction theorem. In the next section we use it to prove an existence and uniqueness theorem for ODE's. In Appendix A it is used to proving the inverse function theorem in higher dimensions.

3 Existence and Uniqueness of solutions for ODEs

In this chapter we will prove a theorem which gives sufficient conditions for a differential equation to have solutions.

3.1 The Picard Theorem for ODE's (for functions which are globally Lipschitz)

In this section we will use the Banach fixed point theorem to show that many differential equations have solutions.

Theorem 3.1

Picard Theorem (global version).
Consider a continuous map $f\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ *which satisfies the Lipschitz inequality*
$|f(s,u) - f(s,v)| \le K|u - v|$ *for all* $s \in \mathbb{R}$, $u, v \in \mathbb{R}^n$. *Let* $h = \frac{1}{2K}$.

Then there exists a unique $x\colon (-h,h) \to \mathbb{R}^n$ *satisfying the initial value problem*

$$\frac{dx}{dt} = f(t,x) \text{ and } x(0) = x_0. \tag{2}$$

This theorem is also called Picard-Lindelöf theorem or Cauchy-Lipschitz theorem, and was developed by these mathematicians in the 19th century.

Question: Why does the non-uniqueness in Example 1.2 not contradict this theorem?

Remark 3.2

Later on we will discuss when one can take $h = \infty$.

Proof : By integration it follows that

$$\frac{dx}{dt} = f(t,x) \text{ and } x(0) = x_0. \tag{3}$$

is equivalent to

$$x(t) - x(0) \overset{\bullet}{=} \int_0^t f(s, x(s))\, ds. \tag{4}$$

It follows that the initial value problem is equivalent to finding a fixed point x of the operator $P\colon X \to X$ defined by

$$P(u)(t) := x_0 + \int_0^t f(s, u(s))\, ds$$

on the Banach space $X := C([-h,h], \mathbb{R}^n)$ with norm $||u|| = \max_{t\in[-h,h]} |u(t)|$.

The reason for the assumption that f is continuous is that it guarantees that this integral exists (of course weaker assumptions on f would suffice.)

Note that P assigns to a function $x \in X$ another function which we denote by $P(x)$. To define the function $P(x)$, we need to evaluate its vector value at some $t \in [-h,h]$. This is what $P(x)(t)$ means. So a solution of $P(x) = x$ is equivalent to finding a solution of (4) and therefore of (3).

Let us show that

$$P(x)(t) := x_0 + \int_0^t f(s, x(s))\, ds$$

is a contraction. Take $x, y\colon [-h, h] \to \mathbb{R}^n$. Then for all $t \in [-h, h]$ one has

Eqations (*)-(***) are clarified on the next page.

$$|P(x)(t) - P(y)(t)| = |\int_0^t (f(s, x(s)) - f(s, y(s)))\, ds| \overset{*}{\leq}$$

$$\int_0^t |f(s, x(s)) - f(s, y(s))|\, ds \overset{**}{\leq} K \textstyle\int_0^t |x(s) - y(s)|\, ds$$

$$\overset{***}{\leq} (hK)||x - y|| \leq (1/2)||x - y||.$$

So

$$||P(x) - P(y)|| = \sup_{t \in [-h,h]} |P(x)(t) - P(y)(t)|$$

$$\leq (1/2)||x - y||$$

and so P is a contraction on the Banach space X. By the previous theorem therefore P has a unique fixed point. ■

(*)

$$|\int_0^t (f(s,x(s)) - f(s,y(s)))\,ds| \overset{*}{\leq} \int_0^t |f(s,x(s)) - f(s,y(s))|\,ds.$$

holds because

Lemma 3.3

$|\int_0^t u(s)\,ds| \leq \int_0^t |u(s)|\,ds$ for any function $u \in X$. (If $t < 0$ the r.h.s. in fact should be $\int_t^0 |u(s)|\,ds$ or $|\int_0^t |u(s)|\,ds|$.)

Proof : Take $0 = t_0 < t_1 < \cdots < t_n = t$ with $|t_{i+1} - t_i| < \epsilon$. Then $\int_0^t u(s)\,ds$ is equal to the limit of the Riemann sum $\sum_{i=0}^{n-1}(t_{i+1} - t_i)u(\xi_i)$ where $\xi_i \in (t_i, t_{i+1})$ as $n \to \infty$ so as $\epsilon \to 0$. By the triangle inequality, $|\sum_{i=0}^{n-1}(t_{i+1} - t_i)u(\xi_i)| \leq \sum_{i=0}^{n-1}(t_{i+1} - t_i)|u(\xi_i)|$ and the right hand side is again a Riemann sum which in the limit (as $n \to \infty$) converges to $\int_0^t |u(s)|\,ds$. ■

(**) The Lipschitz assumption on f implies

$$\int_0^t |f(s,x(s)) - f(s,y(s))|\,ds \overset{**}{\leq} K \int_0^t |x(s) - y(s)|\,ds.$$

(***) Finally

$$K \int_0^t |x(s) - y(s)|\,ds \overset{***}{\leq} (hK)||x - y|| \leq (1/2)||x - y||$$

holds because $|x(s) - y(s)| \leq ||x - y||$ (because $||x - y|| = \sup_{s \in [-h,h]} |x(s) - y(s)|$). So $\int_0^t |x(s) - y(s)|\,ds \leq t \cdot ||x - y||$, and using $|t| \leq h$ inequality (***) follows.

3.2 Application to linear differential equations

Corollary 3.4

Consider the IVP

$$x' = Ax \text{ with } x(0) = x_0 \tag{5}$$

where A is a $n \times n$ matrix and $x(t) \in \mathbb{R}^n$.

1. *There exists $h > 0$ so that (5) has a unique solution $x\colon (-h, h) \to \mathbb{R}^n$. (Later on we shall show $h = \infty$.)*
2. *Write $x_0 = (c_0, \dots, c_n) \in \mathbb{R}^n$ and let $u_i(t)$ be the (unique) solution of $\dot{u}_i = Au_i$, $u_i(0) = e_i$ where e_i is the i-th basis vector. Then*

   $$x(t) = c_1 u_1(t) + \cdots + c_n u_n(t).$$

3. *$x(t) = e^{At} x_0$ where*

   $$e^{At} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}.$$

Proof : **(1)** Note that $|Ax - Ay| \leq K|x - y|$ where K is the **matrix norm** of A defined by $||A|| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \dfrac{|Ax|}{|x|}$. So the Picard Theorem implies that the initial value problem (5) has a unique solution $t \mapsto x(t)$ for $|t| < h$. It is important to remark that the Picard theorem states that there exists $h > 0$ (namely $h = 1/(2K)$) so that there exists a solution $x(t)$ for $|t| < h$. So **at this point** we cannot yet guarantee that there exists a solution all $t \in \mathbb{R}$.

(2) For each choice of $x_0 = (c_1, \dots, c_n) \in \mathbb{R}^n$ there exists a unique solution $x(t)$ (for $|t|$ small) of $\dot{x} = Ax$, $x(0) = x_0$. For each $i = 1, \dots, n$, let $u_i(t)$ be the (unique) solution of $\dot{u}_i = Au_i$, $u_i(0) = e_i$ where e_i is the i-th basis vector. Let $u(t) = c_1u_1(t) + \cdots + c_nu_n(t)$. Since linear combinations of solutions of $x' = Ax$ are also solutions and since $u(0) = x_0$, we get by uniqueness that $x(t) \equiv u(t)$. It follows that

$$u(t) = c_1u_1(t) + \cdots + c_nu_n(t)$$

is the general solution of $x' = Ax$.

(3) Take

$$x_0(t) :\equiv x_0$$

and define

$$x_{n+1}(t) = P(x_n)(t) := x_0 + \int_0^t Ax_n(s)\,ds.$$

Then

$$\begin{aligned} x_1(t) \quad &= x_0 + \textstyle\int_0^t Ax_0(s)\,ds = x_0 + tAx_0. \\ x_2(t) \quad &= x_0 + \textstyle\int_0^t Ax_1(s)\,ds = x_0 + tAx_0 + \frac{t^2}{2}A^2x_0 \end{aligned}$$

By induction

$$x_n(t) = x_0 + tAx_0 + \frac{t^2}{2}A^2x_0 + \cdots + \frac{t^n}{n!}A^nx_0 = \sum_{k=0}^{n} \frac{t^kA^k}{k!}x_0.$$

So the solution of (5) is

$$x(t) = e^{At}x_0 \text{ where we write } e^{At} = \sum_{k=0}^{\infty} \frac{t^kA^k}{k!}.$$

The proof of the Picard Theorem shows that this infinite sum exists (i.e. converges) when $|t| < h$. Later on we shall show that it exists for all t. ■

3.3 The Picard Theorem for functions which are locally Lipschitz

The previous theorem does not apply to many differential equations, such as $x' = x^2$ (because $\mathbb{R} \ni x \mapsto x^2$ is not Lipschitz). So let's state a 'local' version of this theorem which merely required that the r.h.s. of the differential equation is Lipschitz (in the state variable) on some open set U.

Theorem 3.5

Picard Theorem (local version).

Let U be an open subset of $\mathbb{R} \times \mathbb{R}^n$ containing $(0, x_0)$ and assume that

(a) $f\colon U \to \mathbb{R}^n$ *is continuous,*

(b) $|f| \leq M$

(c) $|f(t,u) - f(t,v)| \leq K|u-v|$ *for all* $(t,u), (t,v) \in U$

(d) $h \in (0, \frac{1}{2K}]$ *is chosen so that* $[-h,h] \times \{y; |y - x_0| \leq hM\} \subset U$ *(such a choice for h is possible since U open).*

Then there exists a **unique** *solution* $(-h,h) \ni t \mapsto x(t)$ *of the IVP:*

$$\frac{dx}{dt} = f(t,x) \text{ and } x(0) = x_0. \tag{6}$$

Remark: assumption (d) automatically holds for *some* $h > 0$.

Proof : Fix $h > 0$ as in the theorem, write $I = [-h, h]$, and let $B := \{y \in \mathbb{R}^n; |y - x_0| \le hM\}$. Next define $X = C(I, B)$ as the space of continuous functions $x \colon I \to B \subset \mathbb{R}^n$ and

$$P \colon C(I,B) \to C(I,B) \;\text{ by }\; P(x)(t) = x_0 + \int_0^t f(s, x(s))\, ds$$

Then the initial value problem (6) is equivalent to the fixed point problem

$$x = P(x).$$

Question: Is $P(x)$ a real number or is $P(x)$ a function?

Question: Why not write $P(x(t))$?

We need to show that P **is well-defined**, i.e. that the expression $P(x)(t) = x_0 + \int_0^t f(s, x(s))\, ds$ makes sense, and that when $x \in C(I, B)$ then $P(x) \in C(I, B)$.

Remember that

$$B := \{y \in \mathbb{R}^n; |y - x_0| \leq hM\}$$

and that (by the choice of $h > 0$)

$$f \colon U \to \mathbb{R}^n \text{ where } U \supset [-h, h] \times B.$$

As mentioned, we need to show that P **is well-defined**, i.e.

1. the expression $P(x)(t) = x_0 + \int_0^t f(s, x(s))\,ds$ makes sense for any $x \in B$, i.e. we need to show that $f(s, x(s))$ is actually defined.
2. $x \in C(I, B)$ implies $P(x) \in C(I, B)$.

This can be seen as follows:

1a. since $x \in C(I, B)$, for $t \in I = [-h, h]$, $(t, x(t)) \in I \times B \subset U$ and $f(t, x(t))$ is well-defined for all $t \in [-h, h]$;

1b. hence $x_0 + \int_0^t f(s, x(s))\,ds$ is well-defined;

2a. $|f| \leq M$ implies $[-h, h] \ni t \mapsto x_0 + \int_0^t f(s, x(s))\,ds$ is continuous;

2b. hence $t \mapsto P(x)(t)$ is a continuous map;

2c. finally, $|P(x)(t) - x_0| \leq \int_0^h |f(s, x(s))|\,ds \leq hM$. So $P(x)(t) \in B$ for all $t \in [-h, h]$.

2d. Hence $P(x) \in C(I, B)$.

Let us next show that

$$P\colon C(I,B) \to C(I,B) \;\text{ by }\; P(x)(t) = x_0 + \int_0^t f(s,x(s))\,ds$$

is a **contraction**: for each $t \in [-h,h]$,

$$|P(x)(t) - P(y)(t)| = \left|\int_0^t (f(s,x(s)) - f(t,y(s)))\,ds\right|$$

$$\le \int_0^t |f(s,x(s)) - f(s,y(s))|\,ds$$

$$\le K\int_0^t |x(s) - y(s)|\,ds \quad \text{(Lipschitz)}$$

$$\le Kt \max_{|s|\le t} |x(s) - y(s)|$$

$$\le Kh||x-y|| \le ||x-y||/2 \qquad \text{since } h \in (0, \tfrac{1}{2K}])$$

Since this holds for all $t \in [-h,h]$ we get $||P(x) - P(y)|| \le ||x-y||/2$. So P has a unique fixed point, and hence the integral equation, and therefore the ODE, has a unique solution. ■

3.3.1 Existence and uniqueness in the continuously differentiable case

If $f\colon U \to \mathbb{R}^n$ is continuously differentiable, then it is not necessary to explicitly check whether the Lipschitz property holds.

Theorem 3.6 (*Existence and uniqueness in the continuously differentiable case*)

Assume $V \subset \mathbb{R} \times \mathbb{R}^n$ is open and $f\colon V \to \mathbb{R}^n$ continuously differentiable. Then for each $(0, x_0) \in V$ there exists $h > 0$ and a unique solution $x\colon (-h, h) \to \mathbb{R}^n$ of $\dot{x} = f(t, x), x(0) = x_0$.

Proof : By the Mean Value Theorem 2.14, for each convex, compact subset $C \subset V$ with $(0, x_0) \in C$ there exists $K \in \mathbb{R}$ so that for all $(t, x), (t, y) \in C$,

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

Moreover, there exists M so that $|f| \leq M$ on C. Now apply the previous local Picard theorem to any subset $U \subset C$ containing $(0, x_0)$. ■

3.3.2 Existence and uniqueness theorem in the autonomous case

$x' = f(t,x)$ is **autonomous** if f does not depend on t, so $U = \mathbb{R} \times V$ and f is the form $f(t,x) = g(x)$ for all $(t,x) \in U$. So in this special setting Theorem 3.5 takes the following form:

Theorem 3.7

Picard Theorem (local autonomous version)
Let $V \subset \mathbb{R}^n$ be open and $g\colon V \to \mathbb{R}^n$ continuous, $|g| \le M$, $|g(u) - g(v)| \le K|u - v|$ for all $u, v \in V$. Assume that $x \in V$ and that h is chosen so that

$$0 < h < 1/(2K) \text{ and } \{y; |y - x_0| \le hM\} \subset V.$$

Then there is a unique solution $x \in (-h,h) \to \mathbb{R}^n$ of

$$x' = g(x), x(0) = x_0.$$

3.4 Some comments on the assumptions in the existence and uniqueness theorems

- To obtain existence from Theorem 3.5 it is enough to find *some* open set $U \ni (0, x_0)$ for which the assumptions hold.

- Often one can apply Theorem 3.5 or 3.6, but not Theorem 3.1. Take for example $x' = (1 + x^2), x(0) = 1$. Then the r.h.s. is **not Lipschitz** on all of $\mathbb{R}$. The function $x \mapsto 1 + x^2$ is locally Lipschitz though.

- It is not necessary to take the initial time to be $t = 0$. The Picard Theorem also gives that there exists $h > 0$ so that the initial value problem

  $$x' = f(t, x), x(t_0) = x_0$$

  has a solution $(t_0 - h, t_0 + h) \ni t \mapsto x(t) \in \mathbb{R}^n$.

- If $(t, x) \mapsto f(t, x)$ has additional smoothness, the solutions will be more smooth. For example, suppose that $f(t, x)$ is real analytic (i.e. $f(t, x)$ can be written as a convergent power series in t and x), then the solution $t \mapsto x(t)$ is also real analytic.

3.5 Solutions can be very complicated

The previous theorem implies that the Lorenz differential equation

$$\begin{aligned}\dot{x} &= \sigma(y - x)\\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz\end{aligned} \tag{7}$$

with, for example $\sigma = 10, r = 28, b = 8/3$, has solutions. However the solutions are very, very complicated and no explicit expression is known for them. Not surprisingly:

`http://www.youtube.com/watch?v=ByH8_nKD-ZM`

4 Linear systems in $\mathbb{R}^n$

In this section we consider

$$x' = Ax \text{ with } x(0) = x_0 \qquad (8)$$

where A is a $n \times n$ matrix and $\mathbb{R} \ni t \mapsto x(t) \in \mathbb{R}^n$.

In Example 3.4 we saw that

$$e^{tA} = \sum_{k \geq 0} \frac{1}{k!}(At)^k$$

is defined for $|t|$ small and that $x(t) = e^{tA}x_0$ is a solution of (8) for $|t|$ small. In this section we will show that e^{tA} is well-defined for all $t \in \mathbb{R}$ and show how to compute this matrix.

Example 4.1

Let $A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$. *Then one has inductively* $(tA)^k = \begin{pmatrix} (t\lambda)^k & 0 \\ 0 & (t\mu)^k \end{pmatrix}$. *So* $e^{tA} = \begin{pmatrix} e^{t\lambda} & 0 \\ 0 & e^{t\mu} \end{pmatrix}$.

Example 4.2

Let $A = \begin{pmatrix} \lambda & \epsilon \\ 0 & \lambda \end{pmatrix}$. *Then one has inductively* $(tA)^k = \begin{pmatrix} (t\lambda)^k & \epsilon k t^k \lambda^{k-1} \\ 0 & (t\lambda)^k \end{pmatrix}$. *By calculating the infinite sum of each entry we obtain* $e^{tA} = \begin{pmatrix} e^{t\lambda} & \epsilon t e^{t\lambda} \\ 0 & e^{t\lambda} \end{pmatrix}$.

Lemma 4.3

Let $A = (a_{ij})$ *be a* $n \times n$ *matrix. Then its exponential* $e^A := \sum_{k \geq 0} \frac{1}{k!} A^k$ *is also a well-defined* $n \times n$ *matrix.*

Proof : let $a_{ij}(k)$ be the matrix coefficients of A^k and define $a := ||A||_\infty := \max|a_{ij}|$. Then

$$\begin{aligned}
|a_{ij}(2)| &= \textstyle\sum_{k=1}^n |a_{ik}a_{kj}| \leq na^2 \leq (na)^2\\
|a_{ij}(3)| &= \textstyle\sum_{k,l}^n |a_{ik}a_{kl}a_{lj}| \leq n^2a^3 \leq (na)^3\\
&\vdots\\
|a_{ij}(k)| &= \textstyle\sum_{k_1,k_2,\ldots,k_n=1}^n |a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{n-1}k_n}| \leq n^{k-1}a^k \leq (na)^k
\end{aligned}$$

So $\sum_{k=0}^\infty \dfrac{|a_{ij}(k)|}{k!} \leq \sum_{k=0}^\infty \dfrac{(na)^k}{k!} = \exp(na)$ which means that the series $\sum_{k=0}^\infty \dfrac{a_{ij}(k)}{k!}$ converges absolutely by the comparison test. So e^A is well-defined. ■

Each coefficient of e^{tA} depends on t. So define $\dfrac{d}{dt}e^{tA}$ to be the matrix obtained by differentiating each coefficient.

Lemma 4.4

$t \mapsto x(t) := e^{tA}x_0$ *is the solution of the IVP* $x' = Ax, x(0) = x_0$.

Proof : Since $e^{0A}x_0 = Ix_0 = x_0$ it suffices to prove that $\dfrac{d}{dt}\exp(tA) = A\exp(tA) = \exp(tA)A$:

here $\overset{*}{=}$ follows from the definition or from Lemma 4.5(2)

$$\frac{d}{dt}\exp(tA) = \lim_{h\to 0}\frac{\exp((t+h)A) - \exp(tA)}{h} \overset{*}{=}$$

$$= \lim_{h\to 0}\frac{\exp(tA)\exp(hA) - \exp(tA)}{h} =$$

$$= \exp(tA)\lim_{h\to 0}\frac{\exp(hA) - I}{h} = \exp(tA)A.$$

Here the last equality follows from the definition of $\exp(hA) = I + hA + \dfrac{h^2}{2!}A^2 + \ldots.$ ■

4.1 Some properties of $\exp(A)$

Lemma 4.5

Let A, B, T be $n \times n$ matrices and T invertible. Then

1. *If $B = T^{-1}AT$ then $\exp(B) = T^{-1}\exp(A)T$;*
2. *If $AB = BA$ then $\exp(A+B) = \exp(A)\exp(B)$*
3. *$\exp(-A) = (\exp(A))^{-1}$*

Proof : (1) $T^{-1}(A+B)T = T^{-1}AT + T^{-1}BT$ and $(T^{-1}AT)^k = T^{-1}A^kT$. Therefore

$$T^{-1}(\sum_{k=0}^{n}\frac{A^k}{k!})T = \sum_{k=0}^{n}\frac{(T^{-1}AT)^k}{k!}.$$

(2) follows from:

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$$

$$e^B = I + B + \frac{B^2}{2!} + \frac{B^3}{3!} + \dots$$

$$e^Ae^B = I + A + B + \frac{A^2}{2!} + AB + \frac{B^2}{2!} + \frac{A^3}{3!} + \frac{A^2B}{2!} + \frac{AB^2}{2!} + \frac{B^3}{3!} + \dots$$

Since $AB = BA$ we have $(A+B)^2 = A^2 + 2AB + B^2$ etc.
So e^Ae^B is equal to

$$I + (A+B) + \frac{(A+B)^2}{2!} + \frac{(A+B)^3}{3!} + \dots = \exp(A+B)$$

(3) follows from (2) taking $B = -A$. ■

For general matrices $\exp(A+B) \neq \exp(A)\exp(B)$.

Example 4.6

Take $A = \begin{pmatrix} \lambda & \epsilon \\ 0 & \lambda \end{pmatrix}$ *and let us compute* e^{tA} *again, but now using the previous lemma.* $tA = t\Lambda + tN$ *where* $\Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ *and* $N = \begin{pmatrix} 0 & \epsilon \\ 0 & 0 \end{pmatrix}$. *Note that* $\Lambda N = \lambda N = N\Lambda$ *and that* $N^2 = 0$. *So*

$$e^{tN} = I + tN = \begin{pmatrix} 1 & t\epsilon \\ 0 & 1 \end{pmatrix}, e^{t\Lambda} = \begin{pmatrix} e^{t\lambda} & 0 \\ 0 & e^{t\lambda} \end{pmatrix}$$

and by the previous lemma

$$e^{tA} = e^{t\Lambda}e^{tN} = \begin{pmatrix} e^{t\lambda} & \epsilon t e^{t\lambda} \\ 0 & e^{t\lambda} \end{pmatrix}.$$

Example 4.7

Similarly, one can derive

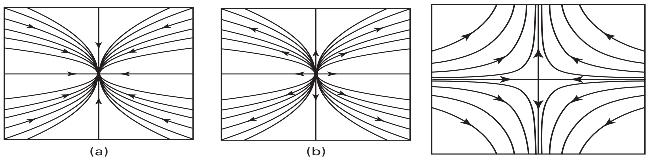
$$\exp\left(t\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\right) = \begin{pmatrix} e^{at}\cos(bt) & e^{at}\sin(bt) \\ -e^{at}\sin(bt) & e^{at}\cos(bt) \end{pmatrix}$$

using this lemma, see the first assignment on problem sheet 2. Another proof is given in Section 4.4.1, see the example below the proof of Proposition 4.11.

4.2 Solutions of 2×2 systems

Theorem 4.8

Let A *be a* 2×2 *matrix and let* λ, μ *be its eigenvalues. If*
(Case a) $\lambda, \mu < 0$ *(**sink**). Then* $x(t) \to 0$ *as* $t \to \infty$.
(Case b) If $\lambda, \mu > 0$ *(**source**). Then* $x(t) = e^{tA}x_0 \to \infty$ *as* $t \to \infty$ *for any* $x_0 \neq 0$.



*(Case c) $\lambda < 0 < \mu$ (**saddle**). Then $x(t) = e^{tA}x_0 \to 0$ as $t \to \infty$ if x_0 lies on the line spanned by the eigenvector corresponding to the eigenvalue $\lambda < 0$, and $x(t) \to \infty$ otherwise.*

Proof : By the Jordan theorem from linear algebra (which we discuss further later on in this chapter), for each 2×2 matrix A there exists T so that $T^{-1}AT$ takes the form of one of the previous examples. ■

Below we will prove this theorem in a more general form.

4.3 A general strategy for computing e^{tA}

In general it is not so easy to compute e^{tA} directly from the definition. For this reason we will discuss

- eigenvalues and eigenvectors;
- using eigenvectors to put a matrix in a new form;
- using eigenvectors and eigenvalues to obtain solutions directly.

Reminder: A vector $v \neq 0$ is an **eigenvector** if $Av = \rho v$ for some $\rho \in \mathbb{C}$ where ρ is called the corresponding **eigenvalue**. So, $(A - \rho I)v = 0$ and $\det(A - \rho I) = 0$. The equation $\det(A - \rho I) = 0$ is a polynomial of degree n in ρ and whose roots are the eigenvalues of A.

Theorem 4.9

If an $n \times n$ matrix A has a basis of eigenvectors $v_1, \ldots, v_n$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$ then
(a) the general solution of $x' = Ax$ is of the form

$$x(t) = c_1 v_1 e^{\lambda_1 t} + \cdots + c_n v_n e^{\lambda_n t}$$

(b) If we take T the matrix with columns $v_1, \ldots, v_n$ then

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ and } e^{tA} = T \begin{pmatrix} e^{t\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{t\lambda_n} \end{pmatrix} T^{-1}.$$

Proof : To see part (a), note that $x(t) = c_1 v_1 e^{\lambda_1 t} + \cdots + c_n v_n e^{\lambda_n t}$ obviously is a solution of $x' = Ax$. Since v_i form a basis, for each $x_0 \in \mathbb{R}^n$ one can find c_i so that $c_1 v_1 + \cdots + c_n v_n = x_0$ and so this is the general solution of the ODE.

Part (b) holds because $T^{-1}ATe_i = T^{-1}Av_i = T^{-1}\lambda_i v_i = \lambda_i e_i$ and by the 1st part of Lemma 4.5. ■

Note that by a lemma from Linear Algebra, if all eigenvalues λ_i of A are distinct then the corresponding eigenvectors $v_1, \ldots, v_n$ are linearly independent and span $\mathbb{R}^n$.

Theorem 4.9 gives *two* (closely related) methods for solving a linear differential equation. Let us illustrate both methods in one example.

Example 4.10

Take $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -2 \\ 0 & 2 & -2 \end{pmatrix}$. *Consider*

$$\det \begin{pmatrix} 1-\lambda & 2 & -1 \\ 0 & 3-\lambda & -2 \\ 0 & 2 & -2-\lambda \end{pmatrix} = -(-1+\lambda)(-2-\lambda+\lambda^2).$$

So A *has eigenvalues* $2, 1, -1$. *To compute the eigenvector w.r.t.* 2 *we need to find* v *with*

$$\begin{pmatrix} -1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & 2 & -4 \end{pmatrix} v = 0.$$

which gives $v = (3, 2, 1)$ *(or multiples). Computing the eigenvectors associated to the other eigenvalues, we find* A *has eigenvalues* $2, 1, -1$ *with eigenvectors* $(3, 2, 1), (1, 0, 0), (0, 1, 2)$.

First method: *The general solution is of the form*

$$c_1 e^{2t} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + c_2 e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_3 e^{-t} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \text{ where } c_i \in \mathbb{R} \text{ is arbitrary .}$$

Second method: *for each vector* $c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ *there exists* $x_0 \in \mathbb{R}^3$ *so that* $c = T^{-1}x_0$. *Hence*

$$\exp(tA)x_0 = T \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{-t} \end{pmatrix} T^{-1}x_0 = T \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{-t} \end{pmatrix} c =$$

$$T \begin{pmatrix} c_1 e^{2t} \\ c_2 e^t \\ c_3 e^{-t} \end{pmatrix} = c_1 e^{2t} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + c_2 e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_3 e^{-t} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

4.4 Special cases

Let us now deal with the case when a matrix has complex or repeated eigenvalues.

4.4.1 Complex eigenvectors

Let us explain what to do in the 2×2 case:

Proposition 4.11

If an 2×2 matrix A with complex eigenvalues $\lambda, \bar{\lambda}$ and eigenvector $v, \bar{v}$. $v_1 = \zeta_1 + i\zeta_2$, $v_2 = \zeta_1 - i\zeta_2$, $\lambda_1 = a + ib$ and $\lambda_2 = a - ib$ where ζ_1, ζ_2 are real vectors and a, b are real numbers. Then

(a) the general solution of $x' = Ax$ is of the form

$$x(t) = d_1 e^{at} \left(\cos(bt)\zeta_1 - \sin(bt)\zeta_2\right) + d_2 e^{at} \left(\sin(bt)\zeta_1 + \cos(bt)\zeta_2\right).$$

(b) If we take T the matrix with columns ζ_1, ζ_2 then $T^{-1}AT = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ and $e^{tA} = e^{ta} T \begin{pmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{pmatrix} T^{-1}$.

Proof: $A(\zeta_1 + i\zeta_2) = (a+bi)(\zeta_1 + i\zeta_2) = (a\zeta_1 - b\zeta_2) + i(a\zeta_2 + b\zeta_1)$. So $A(\zeta_1) = a\zeta_1 - b\zeta_2$ and $A(\zeta_2) = (a\zeta_2 + b\zeta_1)$. It follows that if T is the matrix consisting of columns ζ_1, ζ_2 then

$$T^{-1}AT = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Indeed, $AT(e_1) = A(\zeta_1) = a\zeta_1 - b\zeta_2 = aT(e_1) - bT(e_2)$ and so $T^{-1}AT(e_1) = ae_1 - be_2$. Similarly $T^{-1}AT(e_2) =$

$be_1 + ae_2$. So

$$\exp(tA) = T\exp(T^{-1}tAT)T^{-1} = T\begin{pmatrix} e^{at}\cos(bt) & e^{at}\sin(bt) \\ -e^{at}\sin(bt) & e^{at}\cos(bt) \end{pmatrix} T^{-1}.$$

Here we use Example 4.7. Now write $\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = T^{-1}x_0$ and check that

$$\exp(At)x_0 = T\begin{pmatrix} e^{at}\cos(bt) & e^{at}\sin(bt) \\ -e^{at}\sin(bt) & e^{at}\cos(bt) \end{pmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} =$$

$$d_1e^{at}\left(\cos(bt)\zeta_1 - \sin(bt)\zeta_2\right) + d_2e^{at}\left(\sin(bt)\zeta_1 + \cos(bt)\zeta_2\right). \blacksquare$$

Another (but more or less equivalent) way to deal with complex eigenvalues and eigenvectors is to diagonalise, and choose coefficients at the end which make the solutions real:

Example 4.12

Let $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. *Its eigenvalues are* $1+i$ *with e.v.* $\begin{pmatrix} 1 \\ i \end{pmatrix}$ *and* $1-i$ *with e.v.* $\begin{pmatrix} 1 \\ -i \end{pmatrix}$.

If we take $T = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$, *then* $T^{-1}AT = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}$. *So* $\exp(tT^{-1}AT) = \begin{pmatrix} \exp(t(1+i)) & 0 \\ 0 & \exp(t(1-i)) \end{pmatrix}$. *So*

$$\exp(tA)x_0 = T\begin{pmatrix} \exp(t(1+i)) & 0 \\ 0 & \exp(t(1-i)) \end{pmatrix} T^{-1}x_0.$$

Writing $T^{-1}x_0 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ *gives*

$$\exp(tA)x_0 = T\begin{pmatrix} c_1\exp(t(1+i)) \\ c_2\exp(t(1-i)) \end{pmatrix}$$

which is equal to

$$e^t \begin{pmatrix} (c_1 + c_2)\cos(t) + i(c_1 - c_2)\sin(t) \\ -(c_1 + c_2)\sin(t) + i(c_1 - c_2)\cos(t) \end{pmatrix} \tag{9}$$

For each choice of d_1, d_2 real, one can find c_1, c_2 (complex) so that $d_1 = c_1 + c_2$ and $d_2 = i(c_1 - c_2)$. (Note that nothing other than saying that $x_0 = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = T \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$. Therefore (9) becomes equal to

$$e^t \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}.$$

Together this gives

$$\exp(tA)x_0 = e^t \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}.$$

Plugging in $t = 0$, one again obtains that $x_0 = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$.

4.4.2 Repeated eigenvalues: the 2×2 case

Proposition 4.13

Let A have an eigenvalue ρ with double multiplicity. Let v be the eigenvalue w.r.t. ρ. Then there exists a vector v_2 so that $(A - \rho I)v_2 = v_1$. (The general procedure is explained in Appendix D.) and

(a) The general solution is of the form

$$x(t) = (c_1 e^{\rho t} + c_2 t e^{\rho t})v_1 + c_2 e^{\rho t} v_2$$

(b) If we let T be the matrix consisting of colums v_1 and v_2 then

$$T^{-1}AT = \begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix} \text{ and } e^{tA} = T \begin{pmatrix} e^{\rho t} & te^{\rho t} \\ 0 & e^{\rho t} \end{pmatrix} T^{-1}.$$

Proof: $Te_1 = v_1$ and $Te_2 = v_2$. So $ATe_1 = Av_1 = \rho v_1$ and $ATe_2 = Av_2 = \rho v_2 + v_1$. It follows that $T^{-1}AT = \begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix}$. So the 2nd part of (b) follows from Example 4.6. (a) follows. ■

Example 4.14

Take $A = \begin{pmatrix} 1 & 9 \\ -1 & -5 \end{pmatrix}$ and let us compute the solution of $x' = Ax$ with $x_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. $\det(A - \rho I) = \begin{pmatrix} 1-\rho & 9 \\ -1 & -5-\rho \end{pmatrix} =$ *$(\rho + 2)^2$ so the eigenvalue -2 appears with double multiplicity. $(A - \rho I)v = \begin{pmatrix} 3 & 9 \\ -1 & -3 \end{pmatrix} v = 0$ implies v is a multiple of $v_1 := \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ so there exists only one eigenvector. To find the 2nd 'generalised eigenvector' consider*

$(A-\rho I)v_2 = v_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ which gives $v_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ as a solution. Take T the matrix with columns v_1, v_2, i.e. $Te_i = v_i$. Then

$$T^{-1}AT = J := \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}$$

Hence

$$e^{tA}x_0 = Te^{tJ}T^{-1} = T\begin{pmatrix} e^{-2t} & te^{-2t} \\ 0 & e^{-2t} \end{pmatrix}T^{-1}x_0 =$$

$$= c_1\begin{pmatrix} 3 \\ -1 \end{pmatrix}e^{-2t} + c_2\left(t\begin{pmatrix} 3 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)e^{-2t}$$

where we take $T^{-1}x_0 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$. Of course for varying choice of c_1, c_2 this gives the general solution, and when we want that $x(0) = x_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ then $c_1 = 1$, $c_2 = -2$ solves the initial value problem.

4.4.3 A $n \times n$ matrix with n times repeated eigenvalue

If an $n \times n$ matrix A has **only one eigenvector** v (which implies that its eigenvalue λ appears with multiplicity n) then

- one can define inductively $v_1 = v$ and $(A - \lambda I)v_{i+1} = v_i$.
- $v_1, \dots, v_n$ are linearly independent and span $\mathbb{R}^n$.
- If we take T the matrix with columns $v_1, \dots, v_n$ then

$$T^{-1}AT = \begin{pmatrix} \lambda & 1 & & & 0 \\ 0 & \lambda & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix}.$$

- $e^{tA} = Te^{t\lambda}(I + tN + \cdots + \frac{t^{n-1}}{(n-1)!}N^{n-1})T^{-1}$ where $N = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}.$

4.5 Complex Jordan Normal Form (General Case)

Theorem 4.15

For each $n \times n$ matrix A there exists a (possibly complex) matrix T so that $T^{-1}AT$ takes the Jordan Normal Form:

$$T^{-1}AT = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{pmatrix} \text{ where}$$

$$J_j = \begin{pmatrix} \rho_j & 1 & 0 & & 0 \\ 0 & \rho_j & 1 & & \\ & & \ddots & & \\ 0 & & & \rho_j & 1 \\ 0 & & & & \rho_j \end{pmatrix}$$

and where ρ_j is an eigenvalue of A and where the the dimension of J_j is the smallest k_j so that $\dim \ker(A - \rho_j I)^{k_j+1} = \dim \ker(A - \rho_j I)^{k_j}$.

If J_j is a 1×1 matrix, then $J_j = (\rho_j)$. Associated to each block J_j, there exists an eigenvector v_j (with eigenvalue ρ_j). The dimension of J_j is equal to the maximal integer k_j so that there exist vectors $w_j^1, w_j^2, \ldots, w_j^{k_j} \neq 0$ (where $w_j^1 = v_j$) inductively defined as $(A - \rho_j I)w_j^{i+1} = w_j^i$ for $i = 1, \ldots, k_j - 1$. The matrix T has columns $w_1^1, \ldots, w_1^{k_1}, \ldots, w_p^1, \ldots, w_p^{k_p}$.

In the computations in Subsection 4.4.3, we showed how to determine T so this holds. The proof in the general case is given in one of the appendices.

4.6 Real Jordan Normal Form

Splitting real and complex parts we obtain:

For each real $n \times n$ matrix A there exists a real $n \times n$ matrix T so that $T^{-1}AT$ takes the real Jordan Normal Form:

$T^{-1}AT = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{pmatrix}$ where J_j is either as in the complex Jordan Normal form when ρ_j real or if it is complex equal to

$$J_j = \begin{pmatrix} C_j & I & 0 & & 0 \\ 0 & C_j & I & & \\ & & \ddots & & \\ 0 & & & C_j & I \\ 0 & & & & C_j \end{pmatrix} \text{ where } C_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$$

where $\rho_j = a_j + ib_j$ and $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Proof: See appendix D.

If J_j is a 2×2 matrix, then $\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$.

5 Maximal solutions, the flow property and continuous dependence on initial conditions

Let us now get back to general (nonlinear) ODE's. In this chapter, when we say that $f\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ is so that the existence and uniqueness of solutions of $x' = f(t,x)$ holds, we mean that for each x_0 there exists $h>0$ so that $\dot{x} = f(t,x), x(0) = x_0$ has a **unique** solution $x\colon (-h,h) \to \mathbb{R}^n$.

5.1 Extending solutions

Lemma 5.1

Let $f\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ be so that the existence and uniqueness of solutions of $x' = f(t,x)$ holds. Assume

$$x_1\colon I_1 \to \mathbb{R}^n, x_2\colon I_2 \to \mathbb{R}^n$$

are both solutions of the initial value problem

$$\dot{x}(t) = f(t,x), x(0) = x_0.$$

Then

$$x_1(t) = x_2(t) \text{ for all } t \in I_1 \cap I_2.$$

Proof: Let $I = I_1 \cap I_2$. Note that the existence and uniqueness theorem implies that $x_1(t) = x_2(t)$ for all t in some interval $(-h,h)$, but here I could be much larger that $(-h,h)$.

We claim that $x_1(t) = x_2(t)$ for all $t \in I$ with $t \geq 0$. If not, then define

$$b := \inf\Big\{t > 0 : t \in I \text{ and } x_1(t) \neq x_2(t)\Big\}.$$

Therefore, we get that $x_1(t) = x_2(t)$ for all $t \in [0,b)$. Due to continuity of $x_1(t)$ and $x_2(t)$ in $t \in I$, we get that

$x_1(b) = x_2(b) = p \in W$. Using the existence and uniquenes theorem, there exists $h > 0$ such that the following IVP

$$x'(t) = f(t, x), \qquad x(b) = p$$

has a unique solution $x : (b-h, b+h) \to \mathbb{R}^n$. Consequently, $x_1(t) = x_2(t)$ for all $t \in (b-h, b+h)$ which contradicts the definition of b, unless b is the right endpoint of I, proving the claim.

Similarly, $x_1(t) = x_2(t)$ for all $t \in I$ with $t \leq 0$. ■

5.2 The maximal solution

Proposition 5.2

Let $f\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be so that the existence and uniqueness of solutions of $x' = f(t, x)$ holds. Then for each $x_0 \in \mathbb{R}^n$ there exists and interval $J(x_0) = (\alpha(x_0), \beta(x_0))$ where $\alpha(x_0) < 0 < \beta(x_0)$ so that the following two properties hold:

(i) there exists a solution $x\colon J(x_0) \to \mathbb{R}^n$ of the initial value problem $\dot{x} = f(t, x), x(0) = x_0$;

(ii) if $x\colon J \to \mathbb{R}^n$ is a solution of the value problem $\dot{x} = f(t, x), x(0) = x_0$ then $J \subset J(x_0)$.

Since there is no larger time interval on which there exists a solution, $x\colon J(x_0) \to \mathbb{R}^n$ is called the *maximal solution*.

Proof : Let $J(x_0)$ be the union of all intervals containing 0 on which there is a solution. By the previous lemma any two solutions agree on the intersections of the two intervals. Hence there exists a solution on the union of all such intervals.

In more detail: Let $J(x_0)$ denote the union of all open intervals J containing 0 such that there exists a solution x_J :

$J \to W$ of the following IVP

$$x_J' = f(x_J), \qquad x_J(0) = x_0. \tag{10}$$

Thus, $J(x_0)$ is also an open interval and of the form $(\alpha(x_0), \beta(x_0))$ for some $\alpha(x_0) < 0 < \beta(x_0)$. Now define $x\colon J(x_0) \to \mathbb{R}^n$ by

$x(t) = x_J(t)$ when $t \in J$ for some J from the def. of $J(x_0)$.

To show this makes sense (so different choices for J don't give different definitions), take two intervals J_1, J_2 from the definition of $J(x_0)$ with $t \in J_1 \cap J_2$. Since (by assumption) $x_{J_i}\colon J_i \to \mathbb{R}^n$ are both solutions of (10), the previous lemma gives $x_1(t) = x_2(t)$ for all $t \in J_1 \cap J_2$ and so the definitions coincide. Property (i) and (ii) follow automatically. ■

5.3 Property of the maximal solution

Theorem 5.3

Let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable. Let $x(t)$ be a solution of

$$x' = f(x), x(0) = x_0 \tag{11}$$

Let $0 \in I = (a, b)$ be the maximal interval for which $I \ni t \mapsto x(t)$ is well-defined (see the previous theorem). Then $b < \infty$ implies $|x(t)| \to \infty$ as $t \uparrow b$.

Proof: Step 1: Suppose that the conclusion of this theorem is wrong. Then there exists $R > 0$ and $t < b$ *arbitrarily* close to b so that $|x(t)| \le R$.

Step 2: Let K be the Lipschitz constant of f on the set $\{x \in \mathbb{R}^n; |x| \le R + 1\}$. (That K exists follows from the Mean Value Theorem, since the derivative $x \mapsto Df(x)$ is continuous and since the set $\{x \in \mathbb{R}^n; |x| \leq R+1\}$ is compact.) Moreover, let M be the maximum of $x \mapsto |f(x)|$ on $\{x \in \mathbb{R}^n; |x| \leq R+1\}$. Note that the unit ball around any $p \in \{\mathbb{R}^n; |x| \leq R\}$ is contained in $\{x \in \mathbb{R}^n; |x| \leq R+1\}$.

Step 3: Hence, by the local Picard theorem, for any $h > 0$ so that $hM < 1$ and $hK < 1/2$ and for *any* $p \in \{\mathbb{R}^n; |x| \leq R\}$ and *any* $t_0 \in \mathbb{R}$, the initial value problem

$$y' = f(y)\ , y(t_0) = p \tag{12}$$

has a solution $y\colon (t_0 - h, t_0 + h) \to \mathbb{R}^n$.

Step 4: Take $t_0 \in (a,b)$ so close to b so that $t_0 + h > b$ and so that, moreover, $|x(t_0)| \leq R$ (this is possible by Step 1).

Step 5: Choose $p = x(t_0)$ and let y be the solution of the IVP (12). Since x, y restricted to $t \in (t_0, b)$ both solve (12), $x(t) = y(t)$ for all $t \in (t_0, b)$.

Step 6: Extend $x\colon (a,b) \to \mathbb{R}^n$ to $\hat{x}\colon (a, t_0 + h) \to \mathbb{R}^n$ by defining $\hat{x}(t) = x(t)$ for $t \in (a,b)$ and $\hat{x}(t) = y(t)$ for $[b, t_0 + h)$. This is a solution of (11) defined on the interval $(a, t_0 + h)$. Since $t_0 + h > b$ this contradicts the maximality of (a,b). ■

5.4 Solutions of autonomous ODE's cannot cross and have the flow property

We say that $x' = f(t,x)$ is **autonomous** or **time-independent** if $f\colon U \to \mathbb{R}^n$ does not depend on t, i.e. if it is of the form $x' = f(x)$. In this case, we often denote by $\phi_t(x_0)$ the solution of

$$x' = f(x), x(0) = x_0 \text{ that is } \frac{d}{dt}\phi_t(x_0) = f(\phi_t(x_0)) \text{ and } \phi_0(x_0) = x_0.$$

So $\phi_t(x_0)$ is the solution going through x_0.

Theorem 5.4

Assume that $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable and consider the solution $\phi_t(x)$ of the autonomous ODE $x' = f(x)$. Then

(a) solutions cannot cross: if x_1 and x_2 are solutions and $x_1(t_1) = x_2(t_2)$ then $x_1(t+t_1) = x_2(t+t_2)$ for all $t \in \mathbb{R}$ for which this is defined.

(b) the flow property holds: $\phi_t(\phi_s(x)) = \phi_{t+s}(x)$.

Proof: To explain and prove this, let x_1, x_2 are solutions with $x_1(t_1) = x_2(t_2) = p \in V$ then

$$x_3(t) = x_1(t+t_1) \text{ and } x_4(t) = x_2(t+t_2)$$

are both solutions to $x' = f(x)$ with $x(0) = p$. So by uniqueness of solutions:

$$x_3 \equiv x_4 \text{ i.e. } x_1(t_1+t) = x_2(t_2+t)\forall t.$$

Property (b) will be proved in the assignments. ■

5.5 Higher order differential equations

Consider a higher order differential equation of the form

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_0(t)y = b(t) \tag{13}$$

where $y^{(i)}$ stands for the i-th derivative of y w.r.t. t.

- One can rewrite (13) as a first order ODE, by defining

  $$z_1 = y, z_2 = y^{(1)}, \ldots, z_n = y^{(n-1)}.$$

  The higher order differential equation (13) is equivalent to

  $$\frac{d}{dt}\begin{pmatrix} z_1 \\ \cdots \\ z_{n-1} \\ z_n \end{pmatrix} = \begin{pmatrix} z_2 \\ \cdots \\ z_n \\ b(t) - [a_{n-1}(t)z_n + \cdots + a_0(t)z_1] \end{pmatrix}$$

- Picard's theorem implies $\exists!$ solution of this ODE which satisfies $(z_1(0), \ldots, z_n(0)) = (y(0), \ldots, y^{(n-1)}(0))$ provided $a_0(t), \ldots, a_{n-1}(t), b(t)$ are all bounded by some constant M.

- One can rewrite the vectorial equation as

  $$\frac{d}{dt}\begin{pmatrix} z_1 \\ \cdots \\ z_{n-1} \\ z_n \end{pmatrix} = A(t)\begin{pmatrix} z_1 \\ \cdots \\ z_{n-1} \\ z_n \end{pmatrix} + \begin{pmatrix} 0 \\ \cdots \\ 0 \\ b(t) \end{pmatrix}$$

  where $A(t)$ is matrix with coefficients depending on t. Therefore, as in subsection 3.4, the general solution of the non-homogeneous ODE is of the form $c_1y_1 + \cdots + c_ny_n + p$ where p is a particular solution. There are at most n degrees of freedom.

5.6 Continuous dependence on initial conditions

Theorem 5.5

Continuous dependence on initial conditions *Let* $U \subset \mathbb{R} \times \mathbb{R}^n$ *be open,* $f, g\colon U \to R^n$ *be continuous and assume that*

$$K = \sup_{(t,u),(t,v)\in U} \frac{|f(t,u) - f(t,v)|}{|u-v|}, \; M = \sup_{(t,u)\in U} |f(t,u) - g(t,u)|$$

are finite. If $x(t)$ *and* $y(t)$ *are respective solutions of the IVP's*

$$\begin{cases} x' & = f(t,x) \\ x(0) & = x_0 \end{cases} \quad \text{and} \begin{cases} y' & = g(t,y) \\ y(0) & = y_0 \end{cases}$$

Then

$$|x(t) - y(t)| \le |x_0 - y_0| e^{K|t|} + \frac{M}{K}(e^{K|t|} - 1).$$

5.7 Gronwall Inequality

Proof:

$$|x(t) - y(t)| \le |x_0 - y_0| + \int_0^t |f(s, x(s)) - g(s, y(s))|\, ds.$$

Moreover,

$$|f(s,x(s)) - g(s,y(s))| \le$$
$$\le |f(s,x(s)) - f(s,y(s))| + |f(s,y(s)) - g(s,y(s))| \le$$
$$\le K|x(s) - y(s)| + M.$$

Hence, writing $u(t) := |x(t) - y(t)|$ we have

$$u(t) \le |x_0 - y_0| + \int_0^t (K|u(s)| + M)$$

and therefore the required inequality follows from the following lemma.

Lemma 5.6

Gronwall Inequality

$$u(t) \leq C_0 + \int_0^t (Ku(s) + M)\, ds \text{ for all } t \in [0,h] \implies$$

$$u(t) \leq C_0 e^{Kt} + \frac{M}{K}\left(e^{Kt} - 1\right) \text{ for all } t \in [0,h].$$

Proof: Let's only prove this only when $M = 0$. Define

$$U(t) = C_0 + \int_0^t (Ku(s))\, ds.$$

Then $u(t) \leq U(t)$. Differentiating, we obtain

$$U'(t) = Ku(t).$$

Hence

$$U'(t)/U(t) = Ku(t)/U(t) \leq K$$

and therefore

$$\frac{d}{dt}\log(U(t)) \leq K.$$

Since $U(0) = C_0$ this gives

$$u(t) \leq U(t) \leq C_0 e^{Kt}.$$

■

5.8 Consequences of Gronwall inequality

- Let us interpret the previous result for $f = g$. Then $M = 0$ and

  $$\left\{ \begin{array}{ll} x' & = f(t,x) \\ x(0) & = x_0 \end{array} \right. \quad \text{and} \left\{ \begin{array}{ll} y' & = f(t,y) \\ y(0) & = y_0 \end{array} \right. \quad implies$$

  $$|x(t) - y(t)| \le |x_0 - y_0| e^{K|t|}.$$

  In particular, uniqueness follows. (This of course we knew already: f satisfies the Lipschitz assumption.)

- The previous inequality states:

  $$|x(t) - y(t)| \le |x_0 - y_0| e^{K|t|} + 0.$$

  So in principle orbits can separate exponentially fast.

5.9 The Lorenz equations: the butterfly effect

If solutions indeed separate exponentially fast, then the differential equation is said to have *sensitive dependence on initial conditions*. (The flapping of a butterfly in the Amazon can cause a hurricane over the Atlantic.)

This sensitive dependence occurs in very simple differential equations, for example in the famous Lorenz differential equation

$$\begin{array}{ll} \dot{x} & = \sigma(y - x) \\ \dot{y} & = rx - y - xz \\ \dot{z} & = xy - bz \end{array} \tag{14}$$

with $\sigma = 10, r = 28, b = 8/3$.

This equation has solutions which are chaotic and have sensitive dependence.

`http://www.youtube.com/watch?v=ByH8_nKD-ZM`

5.10 Double pendulum

There are many physical system where sensitive dependence of initial conditions occurs. For example the **double pendulum**, see for example `https://www.youtube.com/watch?v=U39RMUzCjiU` or `https://www.youtube.com/watch?v=fPbExSYcQgY`.

6 Power Series Solutions

Theorem 6.1

If f is real analytic near $(x_0, 0)$, then $x' = f(t,x), x(0) = x_0$ has a real analytic solution, i.e. the solution $t \mapsto x(t)$ is a power series in t which converges for $|t| < h$.

To *prove* this theorem one considers in the differential equation $x' = f(t,x)$ the time t be complex! We will not pursue this here.

Note that in this chapter we obtain take the derivative w.r.t. x, so write instead $y' = f(x,y)$ and look for solutions $x \mapsto y(x)$.

In this chapter we will consider some examples. Typically, one the coefficients appearing in the power series expansions of the solutions can be found inductively as in the next examples.

Example 6.2

$y' = y$. Then substitute $y = \sum_{i \geq 0} a_i x^i$ and $y' = \sum_{j \geq 1} j a_j x^{j-1} = \sum_{i \geq 0} (i+1) a_{i+1} x^i$. Comparing powers gives $\sum_{i \geq 0} (a_i x^i - (i+1) a_{i+1} x^i) = 0$ and so $a_{i+1} = a_i/(i+1)$. So $a_n = C/n!$ which gives $y(x) = C \sum_{n \geq 0} x^n/n! = C \exp(x)$.

6.1 Legendre equation

Example 6.3

Consider the **Legendre** *equation at $x = 0$:*

$$(1-x^2)y'' - 2xy' + p(p+1)y = 0.$$

Write $y = \sum_{i \geq 0} a_i x^i$,

$$y' = \sum_{j \geq 1} j a_j x^{j-1} = \sum_{i \geq 0} (i+1) a_{i+1} x^i.$$

$$y'' = \sum_{j\geq 2} j(j-1)a_j x^{j-2} = \sum_{i\geq 0}(i+2)(i+1)a_{i+2}x^i.$$

We determine a_i as follows.

$$y'' - x^2y'' - 2xy' + p(p+1)y =$$

$$\sum_{i\geq 0}(i{+}2)(i{+}1)a_{i+2}x^i - \sum_{i\geq 2} i(i{-}1)a_i x^i - 2\sum_{i\geq 1} ia_i x^i + p(p{+}1)\sum_{i\geq 0} a_i x^i$$

$$= \sum_{i\geq 2}\left[(i+2)(i+1)a_{i+2} - i(i-1)a_i - 2ia_i + p(p+1)a_i\right]x^i +$$

$$+(2a_2 + 6xa_3) - 2a_1x + p(p+1)(a_0 + a_1x)$$

$$\sum_{i\geq 0}(i{+}2)(i{+}1)a_{i+2}x^i - \sum_{i\geq 2} i(i{-}1)a_i x^i - 2\sum_{i\geq 1} ia_i x^i + p(p{+}1)\sum_{i\geq 0} a_i x^i$$

$$= \sum_{i\geq 2}\left[(i+2)(i+1)a_{i+2} - i(i-1)a_i - 2ia_i + p(p+1)a_i\right]x^i +$$

$$+(2a_2 + 6xa_3) - 2a_1x + p(p+1)(a_0 + a_1x)$$

So collecting terms with the same power of x together gives
$a_2 = -\frac{p(p+1)}{2}a_0$ *and* $a_3 = \frac{(2-p(p+1))}{6}a_1$ *and*

$$a_{i+2} = \frac{[i(i-1) + 2i - p(p+1)]a_i}{(i+1)(i+2)} = -\frac{(p-i)(p+i+1)}{(i+2)(i+1)}a_i.$$

If p is an integer, $a_{p+2j} = 0$ for $j \geq 0$. Convergence of
$y = \sum_{i\geq 0} a_i x^i$ *for $|x| < 1$ follows from the ratio test.*

6.2 Second order equations with singular points

Sometimes one encounters a differential equation where the solutions are not analytic because the equation has a pole. For example

$$y'' + (1/x)y' - (1/x^2)y = 0.$$

Or more generally if the equation can be written in the form

$$y'' + p(x)y' - q(x)y = 0$$

where p has a pole of order 1 and q a pole of order 2. That is,

$$p(x) = \frac{a_{-1}}{x} + \sum_{n \geq 0} a_n x^n, q(x) = \frac{b_{-2}}{x^2} + \frac{b_{-1}}{x} + \sum_{n \geq 0} b_n x^n \quad (15)$$

and where the sums are convergent. Such systems are said to have a **regular singular point** at $x = 0$.

Even though the existence and uniqueness theorem from Chapter 3 no longer guarantees the existence of solutions, it turns out that a solution of the form $y = x^m \sum_{i \geq 0} a_i x^i$ exists. Here, in general, $m \in \mathbb{C}$ and $\sum a_i x^i$ converges near 0). For simplicity we always assume $a_0 \neq 0$.

It will turn out that m is necessarily a root of a quadratic equation (called the indicial equation), which therefore has two roots m_1, m_2. (In these notes we will only encounter cases when $m \in \mathbb{R}$.) The general form of $y'' + p(x)y' - q(x)y = 0$ will be then of the form

$$y(x) = Ax^{m_1} \sum_{i \geq 0} a_i x^i + Bx^{m_2} \sum_{i \geq 0} b_i x^i$$

where a_i and b_i satisfy some recursive equations.

Example 6.4

$$2x^2 y'' + x(2x+1)y' - y = 0.$$

Substitute $y = \sum_{i\geq 0} a_i x^{m+i}$ *where we CHOOSE* m *so that* $a_0 \neq 0$. *Then* $y' = \sum_{i\geq 0}(m+i)a_i x^{m+i-1}$ *and* $y'' = \sum_{i\geq 0}(m+i)(m+i-1)a_i x^{m+i-2}$. *Note that* m *may not be an integer so we always start with* $i = 0$. *Plugging this in gives*

$$2\sum_{i\geq 0}(m+i)(m+i-1)a_i x^{m+i} + 2\sum_{i\geq 0}(m+i)a_i x^{m+i+1} +$$

$$+\sum_{i\geq 0}(m+i)a_i x^{m+i} - \sum_{i\geq 0} a_i x^{m+i} = 0.$$

Collecting the coefficient in front of x^m *gives*

$$(2m(m-1) + m - 1)a_0 = 0.$$

Since we assume $a_0 \neq 0$ *we get the indicial equation*

$$2m(m-1) + m - 1 = 0$$

which gives $m = -1/2, 1$. *The coefficient in front of all the terms with* x^{m+i} *gives*

$$2(m+i)(m+i-1)a_i + 2(m+i-1)a_{i-1} + (m+i)a_i - a_i = 0, \text{ i.e.}$$

$$[2(m+i)(m+i-1) + (m+i) - 1]a_i = -2(m+i-1)a_{i-1}.$$

Taking $m = -1/2$ *this reduces to:* $a_j = \dfrac{3-2j}{-3j+2j^2}a_{j-1}$.

If $m = 1$ *then this gives* $a_j = \dfrac{-2j}{3j+2j^2}a_{j-1}$.

So the general solution is of the form:

$$y = Ax^{-1/2}\left(1 - x + (1/2)x^2 + \dots\right) + Bx\left(1 - (2/5)x + \dots\right).$$

The ratio test gives that $(1 - x + (1/2)x^2 + \dots)$ *and* $(1 - (2/5)x + \dots)$ *converge for all* $x \in \mathbb{R}$.

Remark: The equation required to have the lowest order term vanish is called the **indicial equation** which has two roots m_1, m_2 (possibly of double multiplicity).

Theorem 6.5

Consider a differential equation $y'' + p(x)y' - q(x)y = 0$ where p, q are as in equation (15). Then

- *If $m_1 - m_2$ is not an integer than we obtain two independent solutions of the form $y_1(x) = x^{m_1} \sum_{i \geq 0} a_i x^i$ and $y_2(x) = x^{m_2} \sum_{i \geq 0} a_i x^i$.*
- *If $m_1 - m_2$ is an integer than one either can find a 2nd solution in the above form, or - if that fails - a 2nd solution of the form $\log(x) y_1(x)$ where $y_1(x)$ is the first solution.*

Certain families of this kind of differential equation with regular singular points, appear frequently in mathematical physics.

- **Legendre equation**

$$y'' - \frac{2x}{1-x^2} y' + \frac{p(p+1)}{1-x^2} y = 0$$

- **Bessel equation**

$$x^2 y'' + xy' + (x^2 - p^2) y = 0$$

- **Gauss' Hypergeometric equation**

$$x(1-x)y'' + [c - (a+b+1)x]y' - aby = 0$$

For suitable choices of a, b solutions of this are the sine, cosine, arctan and log functions.

7 Boundary Value Problems, Sturm-Liouville Problems and Oscillatory equations

Instead of initial conditions, in this chapter we will consider differential equations satisfying boundary values. Examples:

- $y'' + y = 0, y(0) = 0, y(\pi) = 0$. The space of solutions is linear: $\{y_c; y_c(x) = c\sin(x), c \in \mathbb{R}\}$.
- $y'' + y = 0, y(0) = 0, y(\pi) = \epsilon \neq 0$ has no solutions: $y(x) = a\cos(x) + b\sin(x)$ and $y(0) = 0$ implies $a = 0$ and $y(\pi) = 0$ has no solutions.
- Clearly boundary problems are more subtle.
- We will concentrate on equations of the form $u'' + \lambda u = 0$ with **boundary conditions**, where λ is a free parameter.
- Such problems are relevant for a large class of physical problems: heat, wave and Schroedinger equations.
- This generalizes Fourier expansions.

7.1 Motivation: wave equation

Consider the **wave equation:**

$$\frac{\partial^2}{\partial t^2}u(x,t) = \frac{\partial^2}{\partial x^2}u(x,t) \tag{16}$$

where $x \in [0, \pi]$ and the end points are fixed:

$$u(0,t) = 0, u(\pi,t) = 0 \text{ for all } t\ , \tag{17}$$

$$u(x,0) = f(x), \frac{\partial}{\partial t}u(x,t)|_{t=0} = 0. \tag{18}$$

This is a model for a string of length π on a musical instrument such as a guitar; before the string is released the shape of the string is $f(x)$. It is also a model for how a bridge vibrates.

7.1.1 How to solve the wave equation

- As usual, one solves (16) by trying to find solutions of the form $u(x,t) = w(x) \cdot v(t)$. Substituting this expression into the wave equation gives $w''(x)/w(x) = v''(t)/v(t)$. Since the left hand does not depend on t and the right hand side not on x this expression is equal to some constant λ and we get

  $$w'' = \lambda w \text{ and } v'' = \lambda v.$$

  By analogy to the setting in linear algebra one often calls λ (or sometimes $-\lambda$) an **eigenvalue** and w an **eigenfunction**.

- We need to set $w(0) = w(\pi) = 0$ to satisfy the boundary conditions that $u(0,t) = u(\pi,t) = 0$ for all t.

- If $\lambda = 0$ then $w(x) = c_3 + c_4 x$ and because of the boundary conditions $c_3 = c_4 = 0$, and therefore $w \equiv 0$. Hence $u \equiv 0$, which is the trivial solution. So we may as well assume $\lambda \neq 0$. In this case it is convenient to write $\lambda = -\mu^2$ where μ is not necessarily real.

- Since $\lambda \neq 0$, $v'' = \lambda v$ has solution

  $$v(t) = c_1 \cos(\mu t) + c_2 \sin(\mu t).$$

- Since $\lambda \neq 0$, $w'' - \lambda w = 0$, has solution $w(x) = c_3 \cos(\mu x) + c_4 \sin(\mu x)$. Since $w(0) = w(\pi) = 0$, $c_3 = 0$ and $w(\pi) = c_4 \sin(\mu\pi) = 0$ implies $\mu = n \in \mathbb{N}$ (or $c_4 = 0$ which implies again $u \equiv 0$). So

  $$w(x) = c_4 \sin(nx) \text{ and } \lambda = -n^2 \text{ and } n \in \mathbb{N} \setminus \{0\}.$$

Check that μ is non-real $\implies \sin(\mu\pi) \neq 0$.

- So for any $n \in \mathbb{N}$ we obtain that

  $$u_n(x,t) = w_n(x) v_n(t) = (c_{1,n} \cos(nt) + c_{2,n} \sin(nt)) \sin(nx)$$

  is a solution of (16) and (17).

- The boundary condition $\frac{\partial}{\partial t}u(x,t)|_{t=0} \equiv 0$ gives

  $$\sum c_{2,n} n \sin(nx) \equiv 0 \implies c_{2,n} = 0 \text{ for all } n \geq 0.$$

  Since the problem is linear, we therefore obtain

  $$u(x,t) = \sum_{n=1}^{\infty} c_{1,n} \cos(nt) \sin(nx)$$

  is a solution.

- This expression shows that the string can only vibrate with frequencies which are of the form $n\pi$ where $n \in \mathbb{N}$.

7.1.2 The boundary condition $u(x, 0) = f(x)$: Fourier expansion

The final boundary condition is that

$$u(x,0) = \sum c_{1,n} \sin(nx) = f(x) \text{ for all } x \in [0,\pi]. \quad (19)$$

This looks like Fourier expansion: That only sin terms appear in this expansion is because $f(0) = f(\pi) = 0$, as explained in Lemma F.2 in Appendix F.

Theorem 7.1

L^2 Fourier Theorem. If $f\colon [0,2\pi] \to \mathbb{R}$ is continuous (or continuous except at a finite number of points) then we one can coefficients $c_{1,n}, c_{2,n}$ so that

$$f \sim \sum_{n=0}^{\infty} (c_{1,n} \cos(nx) + c_{2,n} \sin(nx))$$

where $\sim$ means that

$$\int_0^{2\pi} |f(x) - \sum_{n=0}^{N} (c_{1,n} \cos(nx) + c_{2,n} \sin(nx))|^2 \, dx \to 0$$

as $N \to \infty$. Moreover, if f' is differentiable then

$$f' \sim \sum_{n=0}^{\infty}(-nc_{1,n}\sin(nx) + nc_{2,n}\cos(nx))$$

7.1.3 C^2 solution of the wave equation

The L^2 Fourier Theorem does **not** claim that $f(x)$ is everywhere equal to $\sum_{n=0}^{\infty}(c_{1,n}\cos(nx) + c_{2,n}\sin(nx))$: the infinite sum in the right hand side may not be well-defined for all x. To obtain this we need the following:

Theorem 7.2

Fourier Theorem with **uniform convergence***. Assume $f\colon [0, 2\pi] \to \mathbb{R}$ is C^1 (continuously differentiable) then one can find $c_{1,n}, s_{1,n}$ so that*

$$\sum_{n=1}^{N}[c_{1,n}\cos(nx) + c_{2,n}\sin(nx)] \text{ converges \textbf{uniformly} to } f(x) \text{ as } N \to \infty.$$

Using this theorem one can prove that if f is C^3, $f(0) = f(\pi) = f''(0) = f''(\pi)$ then there exists a C^2 function u of the form $u(x,t) = \sum_{n=1}^{\infty} c_{1,n}\cos(nt)\sin(nx)$ which is a genuine solution of the wave equation. Some details for this are given in Appendix F (which is non-examinable).

7.2 A typical Sturm-Liouville Problem

Let us consider another example:

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

- If $\lambda = 0$ then $y(x) = c_1 + c_2 x$ and the boundary conditions give $y(x) = 1 - x$ (or multiples of $y(x) = 1 - x$).

- If $\lambda \neq 0$ we write again $\lambda = \mu^2$. The equation $y'' + \lambda y = 0$ gives $y(x) = c_1 e^{i\mu x} + c_2 e^{-i\mu x}$. Plugging this expression in $y(0) + y'(0) = 0$, $y(1) = 0$ gives $(c_1 + c_2) + i\mu(c_1 - c_2) = 0$ and $c_1 e^{i\mu} + c_2 e^{-i\mu} = 0$. So $c_2 = -c_1 e^{2i\mu}$ and $c_1[(1 + i\mu) - (1 - i\mu)e^{2i\mu}] = 0$. Since we can assume $c_1 \neq 0$ (otherwise $y \equiv 0$), the last equation reduces to $[(1 + i\mu)e^{-i\mu} - (1 - i\mu)e^{i\mu}] = 0$ (*).

- If $\lambda < 0$ then μ is purely imaginary because $\lambda = \mu^2$ and so $\mu = iR$ for some real R. Then (*) corresponds to $(1-R)e^{+R} - (1+R)e^{-R} = 0$, i.e. $e^{2R} = (1+R)/(1-R)$, and it is easy to see that the only real solution of this is $R = 0$ and so $\lambda = 0$ (which was treated before).

- If $\lambda > 0$ and so μ is real then (*) implies $\tan \mu = \mu$ (see margin). This equation has infinitely many solutions $\mu_n \in [0, \infty)$, $n = 0, 1, \ldots$ $\mu_n \in ((2n - 1)\pi/2, (2n + 1\pi/2)$ when $n \geq 0$. In fact, $\mu_0 = 0$ and $\mu_n \approx (2n + 1)\pi/2$ when n is large.

Indeed, $0 = (1 + i\mu)(\cos \mu - i \sin \mu) - (1 - i\mu)(\cos \mu + i \sin \mu) = (2i(\mu \cos \mu - \sin \mu)$, so $\tan \mu = \mu$.

- $y(x) = [c_1 e^{i\mu x} + c_2 e^{-i\mu x}] = c_1[e^{i\mu x} - e^{2i\mu} e^{-i\mu x}] = c_1 e^{i\mu}[e^{-i\mu + i\mu x} - e^{i\mu - i\mu x}] = 2\tilde{c}_1 \sin(\mu x - \mu)$. Here $\tilde{c}_1$ is a new (complex) constant. So $y_n(x) = \sin(\mu_n(1 - x))$ is an eigenfunction.

- Summarising we get eigenvalues: $\lambda_n = \mu_n^2$ with $\mu_0 = 0$, $\mu_n \in ((2n - 1)\pi/2, (2n + 1\pi/2))$ for $n \geq 1$ and for n large $\lambda_n \approx (2n + 1)^2(\pi/2)^2$. and eigenfunctions: $y_0(x) = 1 - x$ and $y_n(x) = \sin(\sqrt{\lambda_n}(1 - x))$, $n \geq 1$.

7.3 The Sturm-Liouville Theorem

The previous example is a special case of the following problem: given functions $p, q, r\colon [a,b] \to \mathbb{R}$ find **both** $y\colon [a,b] \to \mathbb{R}$ **and** λ so that

$$(p(x)y')' + q(x)y + \lambda r(x)y = 0. \tag{20}$$

Theorem 7.3

Sturm-Liouville Theorem *Assume that $p, r > 0$ are continuous and p is C^1 on $[a,b]$. Then (20) with the boundary conditions (21)*

$$\alpha_0 y(a) + \alpha_1 y'(a) = 0, \beta_0 y(b) + \beta_1 y'(b) = 0. \tag{21}$$

(where α_i, β_i are assumed to be real and neither of the vectors $(\alpha_0, \alpha_1), (\beta_0, \beta_1)$ are allowed to be zero) has infinitely many solutions y_n and λ_n with the following properties:

1. *The numbers λ_n (which are usually called* eigenvalues*) are real, distinct and of single multiplicity;*

2. *The eigenvalues λ_n tend to infinity, so $\lambda_1 < \lambda_2 < \dots$ and $\lambda_n \to \infty$.*

3. *If $n \neq m$ then corresponding* eigenfunctions *y_n, y_m are real and orthogonal in the sense that*

$$\int_a^b y_m(x) y_n(x) r(x)\, dx = 0.$$

4. Each *continuous function f can be expanded in terms of the eigenfunctions, as in the Fourier case: one can find coefficients c_n, $n = 0, 1, 2, \dots$ so that f is the limit of the sequence of functions $\sum_{n=0}^{N} c_n y_n$.*

If f is merely continuous than this convergence is in the L^2 sense, while if f is C^2 then this convergence is uniform (this sentence is not examinable, and in this course we will not cover a proof of this sentence).

We will **not** be able to **prove** this theorem in this course, but for those who are interested there is a brief outline of the strategy of the proof below.

Remark 1: If y_n, y_m are solutions and we set

$$W(y_m, y_n)(x) := \det(\begin{matrix} y_m(x) & y_m'(x) \\ y_n(x) & y_n'(x) \end{matrix}) = y_m(x)y_n'(x) - y_n(x)y_m'(x)$$

then $W(a) = 0$ and $W(b) = 0$. To see that $W(a) = 0$, note that the first boundary condition in equation (21) implies

$$\begin{pmatrix} y_m(a) & y_m'(a) \\ y_n(a) & y_n'(a) \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = 0.$$

Since $(\alpha_0, \alpha_1) \neq (0, 0)$ the determinant of the matrix is zero.

Remark 2: For any function f there is an easy way to obtain the necessary conditions for a_n so that $f(x) = \sum_{n \geq 0} a_n y_n(x)$. Just define (v, w) is the **inner product**:

$$(v, w) = \int_a^b v(t)\overline{w}(t)\,dt$$

and take

$$\begin{aligned} (f, ry_k) &= (\textstyle\sum_{n \geq 0} a_n y_n(x), ry_k) = \sum_{n \geq 0} a_n (y_n, ry_k) \\ &= a_k(y_k, ry_k). \end{aligned}$$

Here we used in the last equality that $(y_n, ry_k) \neq 0$ implies $n = k$. Hence

$$a_k := \frac{(f, ry_k)}{(y_k, ry_k)}.$$

7.4 A glimpse into symmetric operators

Sturm-Liouville problems are solved using some operator theory: the 2nd order differential equation is equivalent to

$$Ly(x) = \lambda r(x)y(x) \text{ where } L = \left(-\frac{d}{dx}p(x)\frac{d}{dx} - q(x)\right).$$

Let H be the space of functions $y\colon [a,b] \to \mathbb{R}$ satisfying the boundary conditions

$$\alpha_0 y(a) + \alpha_1 y'(a) = 0, \beta_0 y(b) + \beta_1 y'(b) = 0. \tag{22}$$

H is an infinite dimensional linear space.

$L\colon H \to H$ turns out to be turns out to be a **symmetric operator** in the sense that $(Lv, w) = (v, Lw)$ where $(v, w) = \int_a^b v(x)\overline{w(x)}\,dx$ is as defined above.

The situation for analogous to the finite dimensional case: L is a symmetric (and satisfies some additional properties) $\implies$ its eigenvalues are real, and its eigenfunctions form a basis.

Lemma 7.4

$L\colon H \to H$ is symmetric (self-adjoint) in the sense that $(Lv, w) = (v, Lw)$ for all $v, w \in H$.

Proof: Let $Lu = -(pu')' - qu$ and $Lv = -(pv')' - qv$.

$\int_a^b L(u)\bar{v}\,dx = \int_a^b [-(pu')'\bar{v} - qu\bar{v}]\,dx$. $\int_a^b u\overline{L(v)}\,dx = \int_a^b [-u(p\overline{v}')' - qu\bar{v}]\,dx$.

$$\begin{aligned}\int_a^b & -(pu')'\bar{v}\,dx = -pu'\bar{v}|_a^b + \int_a^b pu'\bar{v}'\,dx \\ &= -pu'\bar{v}|_a^b + pu\bar{v}'|_a^b - \int_a^b u(p\bar{v}')'\,dx.\end{aligned}$$

Hence

$$\begin{aligned}\int_a^b [L(u)\bar{v} - & u\overline{L(v)}]\,dx = -p(x)[u'\bar{v} - u\bar{v}']\Big|_a^b \\ &= [p(b)W(u,\bar{v})(b) - p(a)W(u,\bar{v})(a)]\end{aligned}$$

If u, v satisfy the boundary conditions, then $\alpha_0 u(a)+\alpha_1 u'(a) = 0$ and $\alpha_0 v(a) + \alpha_1 v'(a) = 0$. Since α_0, α_1 are real, therefore $\alpha_0 \bar{v}(a) + \alpha_1 \bar{v}'(a) = 0$. Hence, $W(u, \bar{v})(a) = W(u, \bar{v})(b) = 0$, and therefore

$$\int_a^b [L(u)\bar{v} - u\overline{L(v)}]\, dx = 0.$$

This implies $(Lu, v) = (u, Lv)$. ■

Lemma 7.5

Eigenvalues of $L\colon H \to H$ are real and eigenfunctions are orthogonal to one another.

Proof: Define $(u, v) = \int_a^b u(x)\overline{v(x)}\, dx$. Then the previous lemma showed $(Lu, v) = (u, Lv)$.

- Suppose that $Ly = r\lambda y$. Then the **eigenvalue λ is real**: Indeed,

  $$\lambda(ry, y) = (\lambda ry, y) = (Ly, y) = (y, Ly) = \bar{\lambda}(y, ry) = \bar{\lambda}(ry, y)$$

  since r is real. Since $(ry, y) > 0$ it follows that $\lambda = \bar{\lambda}$.

- Suppose that $Ly = r\lambda y$ and $Lz = r\mu z$.
  $\lambda \neq \mu \implies \int_a^b r(x)y(x)\overline{z(x)}\, dx = (ry, z) = 0$.
  So the eigenfunctions y, z are **orthogonal**. Indeed,

  $$\begin{aligned} \lambda(ry, z) &= (\lambda ry, z) = (Ly, z) = (y, Lz) = (y, \mu rz) \\ &= \bar{\mu}(y, rz) = \bar{\mu}(ry, z) = \mu(ry, z). \end{aligned}$$

  where we have used that r and μ are real. Since $\lambda \neq \mu$ it follows that $(ry, z) = 0$. ■

7.5 The strategy for proving the Sturm-Liouville Theorem 7.3

In this subsection, which is **non-examinable** and which will **not be covered in class**, we explain some of the ideas used in the proof of Theorem 7.3.

- Define a Hilbert space H: this is a Banach space over the complex numbers with the inner product as above. Note $(v, w) = \overline{(w, v)}$ where $\overline{z}$ is complex conjugation. Define the norm $||v|| = \sqrt{(v, v)} = ||v|| = \sqrt{\int_a^b |v(x)|^2 \, dx}$, the so-called L^2 norm (generalizing the norm $||z|| = \sqrt{(z, z)} = \sqrt{z\overline{z}}$ on $\mathbb{C}$).

- Associate to each linear map $A\colon H \to H$, the operator norm $||A|| = \sup_{f \in H, ||f||=1} ||Af||$;

- to call a linear operator A is **compact** if for each sequence $||f_n|| \leq 1$, there exists a convergent subsequence of Af_n.

- $A\colon H \to H$ is called compact, if there exists a sequence of eigenvalues $\alpha_n \to 0$ and eigenfunctions u_n. These eigenvalues are all real and the eigenfunctions are orthogonal. If the closure of $A(H)$ is equal to H, then for each $f \in H$ then one can write $f = \sum_{j=0}^{\infty} (u_j, f) u_j$.

- The operator L in Sturm-Liouville problems is **not** compact, and that is why one considers some related operator (the resolvent) which is compact.

The Sturm-Liouville Theorem is fundamental in

- quantum mechanics;

- in large range of boundary value problems;

- and related to geometric problems describing properties of geodesics.

7.6 Oscillatory equations

Consider $(py')' + ry = 0$ where $p > 0$ and C^1 as before and r continuous.

Theorem 7.6

Let y_1, y_2 be solutions. Then the **Wronskian** $x \mapsto W(y_1, y_2)(x) := y_1(x)y_2'(x) - y_2(x)y_1'(x)$ *has constant sign. (We also write $W(x) = W(y_1, y_2)(x)$.)*

Proof : Now assume by contradiction that $x \mapsto W(x)$ changes sign. Then W would be zero as some point x. Let us show that this implies that $W \equiv 0$.

$(py_1')' + ry_1 = 0$ and $(py_2')' + ry_2 = 0$. Multiplying the first equation by y_2 and the second one by y_1 and subtract:

$$0 = y_2(py_1')' - y_1(py_2')' = y_2p'y_1' + y_2py_1'' - y_1p'y_2' - y_1py_2''.$$

Differentiating W and substituting the last equation in

$$pW' = p[y_1'y_2' + y_1y_2'' - y_2'y_1' - y_2y_1''] = py_1y_2'' - py_2y_1''$$

gives

$$pW' = -p'W. \tag{23}$$

This implies that if $W(x) = 0$ for some $x \in [a, b]$ then $W(z) = 0$ for all $z \in [a, b]$. Indeed, the differential equation (23) can be written as $W' = f(x, W)$ where $f(x, W) := (-p'(x)/p(x))W$. Since $W \mapsto f(x, W)$ is Lipschitz in W (it is linear in W), the *only* solution of the IVP $W' = f(x, W), W(x) = 0$ is $W \equiv 0$. ■

Lemma 7.7

$W(y_1, y_2) \equiv 0 \implies \exists c \in \mathbb{R}$ with $y_1 = cy_2$ (or $y_2 = 0$).

Proof : Since $W(y_1, y_2) = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$, the assumption $W(y_1, y_2) = 0$, $y_2 \neq 0$ implies that the columns of the matrix are linearly dependent, and so (y_1, y_1') is a multiple of (y_2, y_2'), i.e.

$$\begin{pmatrix} y_1(x) \\ y_1'(x) \end{pmatrix} = c(x) \begin{pmatrix} y_1(x) \\ y_1'(x) \end{pmatrix}.$$

Can $c(x)$ depend on x? No:

$$\begin{aligned} y_1(x) &= c(x)y_2(x) \text{ and } y_1'(x) = c(x)y_2'(x) \;\; \forall x \\ &\implies c(x)y_2'(x) = y_1'(x) = c'(x)y_2(x) + c(x)y_2'(x) \;\; \forall x. \end{aligned}$$

Hence $c' \equiv 0$. ■

Theorem 7.8

Sturm Separation Theorem *Consider $(py')' + ry = 0$ with p, r as above. Let y_1, y_2 be two solutions of which are independent (one is not a constant multiple of the other). Then zeros are interlaced: between consecutive zeros of y_1 there is a zero of y_2 and vice versa.*

Proof : Assume $y_1(a) = y_1(b) = 0$. $y_1'(a) \neq 0$ (otherwise $y_1 \equiv 0$) and $y_2'(b) \neq 0$. We may choose a, b so that $y_1(x) > 0$ for $x \in (a, b)$. Then $y_1'(a)y_1'(b) < 0$. (Draw a picture.) Also,

$$W(y_1, y_2)(a) = -y_2(a)y_1'(a) \text{ and } W(y_1, y_2)(b) = -y_2(b)y_1'(b).$$

Since $y_1'(a)y_1'(b) < 0$ and $W(y_1, y_2)(a)W(y_1, y_2)(b) > 0$ (W does not change sign), we get $y_2(a)y_2(b) < 0$, which implies that y_2 has a zero between a and b. ■

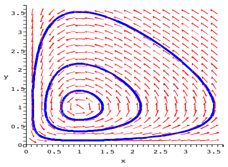
8 Nonlinear Theory

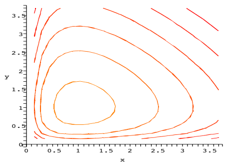
In the remainder of this course we will study initial value problems associated to autonomous differential equations

$$x' = f(x), x(0) = x_0 \tag{24}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is C^∞. We saw:

- There exists $\delta(x_0) > 0$ so that this has a unique solution $x\colon (-\delta, \delta) \to \mathbb{R}^n$;
- There exists a unique **maximal domain of existence** $I(x_0) = (\alpha(x_0), \beta(x_0))$ and a unique **maximal** solution $x\colon I(x_0) \to \mathbb{R}^n$.
- If $\beta(x_0) < \infty$ then $|x(t)| \to \infty$ when $t \uparrow \beta(x_0)$.
- If $\alpha(x_0) > -\infty$ then $|x(t)| \to \infty$ when $t \downarrow \alpha(x_0)$.
- The solution is often denoted by $\phi_t(x_0)$.
- The map $(t, x) \mapsto \phi_t(x)$ is **continuous** in (t, x). That it is continuous in x follows from Theorem 8.8 (taking $f = g$ and so $M = 0$ in that theorem). That it is jointly continuous in (t, x) then follows easily.
- One has the **flow property**: $\phi_{t+s}(x_0) = \phi_t\phi_s(x_0)$, $\phi_0(x_0) = x_0$.
- Solutions do not intersect. One way of making this precise goes as follows: $t > s$ and $\phi_t(x) = \phi_s(y)$ implies $\phi_{t-s}(x) = y$. (So $\phi_s(\phi_{t-s}(x)) = \phi_t(x) = \phi_s(y)$ implies $\phi_{t-s}(x) = y$.)





8.1 The orbits of a flow

Rather than studying each initial value problem separately, it makes sense to study the **flow** ϕ_t associated to $x' = f(x), x(0) = x_0$. The curves $t \mapsto \phi_t(x)$ are called the **orbits**. For example we will show that the flow of

$$\begin{aligned} \dot{x} &= Ax - Bxy \\ \dot{y} &= Cy + Dxy \end{aligned}$$

is equal to

8.2 Singularities

Consider $x' = f(x), x(0) = x_0$.

If $f(x_0) = 0$ then we say that x_0 is a *rest point* or *singularity*. In this case $x(t) \equiv x_0$ is a solution, and by uniqueness **the** solution. So $\phi_t(x_0) = x_0$ for all $t \in \mathbb{R}$.

This notion is so important that several alternative names are used for this: **rest point**, **fixed point**, **singular point** or **critical point**.

Near such points usually a linear analysis suffices.

Since $f(x_0) = 0$, and assuming that f is C^1 we obtain by Taylor's Theorem

$$f(x) = f(x_0) + A(x - x_0) + o(|x - x_0|^1) = A(x - x_0) + o(|x - x_0|)$$

where $o(|x - x_0|)$ is so that $o(|x - x_0|)/|x - x_0| \to 0$ as $x \to x_0$. (By the way, if f is C^2 we have $f(x) = A(x - x_0) + O(|x - x_0|^2)$.)

$A = Df(x_0)$ is called the **linear part** of f at x_0.

8.3 Stable and Unstable Manifold Theorem

A matrix A is called **hyperbolic** if its eigenvalues $\lambda_1, \ldots, \lambda_n$ have non-zero real part, i.e. satisfy $\Re(\lambda_i) \neq 0$, $i = 1, \ldots, n$. Order the eigenvalues so that

$\Re(\lambda_i) < 0$ for $i = 1, \ldots, s$ and $\Re(\lambda_i) > 0$ for $i = s+1, \ldots, n$.

Let E^s (resp. E^u) be the eigenspace associated to the eigenvalues $\lambda_1, \ldots, \lambda_s$ (resp. $\lambda_{s+1}, \ldots, \lambda_n$).

A singular point x_0 of f is called **hyperbolic** if the matrix $Df(x_0)$ is hyperbolic.

Theorem 8.1

Stable and Unstable Manifold Theorem *Let x_0 be a singularity of f and assume x_0 is hyperbolic. Then there exist a manifold $W^s(x_0)$ of dimension s and a manifold $W^u(x_0)$ of dimension $n - s$ both containing x_0 so that*

$$x \in W^s(x_0) \iff \phi_t(x) \to x_0 \text{ as } t \to \infty,$$

$$x \in W^u(x_0) \iff \phi_t(x) \to x_0 \text{ as } t \to -\infty.$$

$W^s(x_0), W^u(x_0)$ are tangent to x_0+E^s resp. x_0+E^u at x_0.

Remarks:

- Here we will not give the general definition of the notion of manifold, but simply define it near x_0 as the graph of a smooth function from the linear space E^s to E^u. Most of the time we will consider the case when E^s and E^u have dimension one (and then $W^s(x_0)$ and $W^u(x_0)$ is a **curve**). If E^s has dimension two, then the object we obtain is a **surface**.

- If $s = n$ then the singularity is called a **sink** and this case the above theorem asserts that $W^s(x_0)$ is a neighbourhood of x_0.

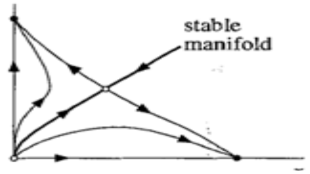


Figure 1: An example of a differential equation which will be studied later on in which there are several singularities: with a sink, source and saddle.

- If $1 \leq s < n$ then it is called a **saddle**.
- $s = 0$ then it called a **source**.
- $W^s(x_0)$ is called the **stable manifold**.
- $W^u(x_0)$ is called the **unstable manifold**.

Example 8.2

Take $x' = x + y^2, y' = -y + x^2$. *The linearisation of this system at* $(0,0)$ *is* $x' = x, y' = -y$. *So the system has eigenvalues* $1, -1$ *with corresponding eigenvectors* $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ *and* $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. *So* E^u *is equal to the* x*-axis and* E^s *is equal to the* y*-axis. Hence, by Theorem 8.1 there is supposed to an invariant manifold* $W^u(0)$ *(a curve) which is tangent to the* x*-axis.*

This means that there exists a function $g(x)$ *so that*

$$W^u(x) = \{(x, g(x)); x \in \mathbb{R}\}$$

near $(0,0)$ *with* $g(0) = 0$ *and* $g'(0) = 0$. *(That we insist on* $g'(0) = 0$ *is to ensure that the curve is tangent to the* x *axis.*

How to find the Taylor expansion $g(x) = a_2x^2 + a_3x^3 + \cdots + O(|x|^k)$ *of* g? *Take a point* $(x, y) \in W^u(0)$ *so that* $y = g(x)$. *We need that* $y(t) = g(x(t))$ *for all* t. *Since* $t \to -\infty$

we have $y(t) \to 0$, $x(t) \to 0$ and so $y(t) - g(x(t)) \to 0$ holds in the limit as $t \to -\infty$ (recall that $g(0) = 0$).

So $y(t)-g(x(t)) = 0$ for all t holds $\implies$ $y(t)-g(x(t))$ is constant $\iff$ $\frac{dy}{dt} - \frac{dg}{dx}\frac{dx}{dt} = 0$ $\iff$

$$g'(x) = (\frac{dy}{dt})/(\frac{dx}{dt}) = \frac{-y+x^2}{x+y^2}. \tag{25}$$

Substituting this in (25) gives

$$2a_2x + 3a_3x^2 + \cdots = \frac{-[a_2x^2 + a_3x^3 + \dots] + x^2}{x + [a_2x^2 + a_3x^3 + \dots]^2}.$$

Comparing terms of the same power, shows that $2a_2 = (1 - a_2)$ and so on. Thus we determine the power series expansion of $g(x)$.

Proof of Theorem 8.1. We will only prove this theorem in the case that $s = n$ and when the matrix $A = Df(x_0)$ has n real eigenvalues $\lambda_i < 0$ and n eigenvectors $v_1, \dots, v_n$. **For simplicity also assume** $x_0 = 0$. Consider x near $x_0 = 0$ and denote the orbit through x by $x(t)$.

Let T be the matrix consisting of the vectors $v_1, \dots, v_n$ (that is $Te_j = v_j$). Then $T^{-1}AT = \Lambda$ where Λ is a diagonal matrix (with $\lambda_1, \dots, \lambda_n$ on the diagonal).

You may want to assume that T is the identity matrix when you go through the first the proof for the first time

Let us show that $\lim_{t\to\infty} x(t) = 0$ provided x is close to 0. Let us write $y(t) = T^{-1}(x(t))$. It is sufficient to show that $y(t) \to 0$. Instead we will show

$$|y(t)|^2 = T^{-1}(x(t)) \cdot T^{-1}(x(t)) \to 0 \text{ as } t \to \infty.$$

To see this we will prove that there exists $\rho' > 0$ so that whenever $y(0)$ is close to zero, then

$$\frac{d|y(t)|^2}{dt} \le -\rho'|y(t)|^2. \tag{26}$$

This would be enough, because if we write $z(t) = |y(t)|^2$ then we get $z' \leq -\rho' z$ which means

$$|y(t)|^2 = z(t) \leq z(0)e^{-t\rho'} \leq |y(0)|^2 e^{-t\rho'}$$

which shows that $y(t) \to 0$ as $t \to \infty$ with rate $\rho'/2$.

So we need to estimate $\dfrac{d|y(t)|^2}{dt}$ from above:

$$\begin{aligned}\frac{d|y(t)|^2}{dt} &= \frac{d}{dt}\left(T^{-1}x(t) \cdot T^{-1}(x(t)\right) = 2T^{-1}x \cdot T^{-1}\dot{x} \\ &= 2T^{-1}x \cdot T^{-1}f(x) \\ &= 2T^{-1}x \cdot T^{-1}Ax + 2T^{-1}x \cdot T^{-1}[f(x) - Ax].\end{aligned}$$

Let us first estimate the **first term** in this sum under the assumption that all eigenvalues of A are real. Then

$$T^{-1}x \cdot T^{-1}Ax = y \cdot \Lambda y \leq -\rho|y|^2 \tag{27}$$

where $\rho = \min_{i=1,\dots,n} |\lambda_i|$. Here we use that Λ is diagonal with all eigenvalues real (and therefore the eigenvectors are real and so T and y are also real).

The **second term** can be estimated as follows: Since $f(x) - Ax = o(|x|)$ for any $\epsilon > 0$ there exists $\delta > 0$ so that $|f(x) - Ax| \leq \epsilon|x|$ provided $|x| \leq \delta$. Hence using the Cauchy inequality and the matrix norm we get

$$T^{-1}x \cdot T^{-1}[f(x) - Ax] \leq |y| \cdot |T^{-1}[f(x) - Ax]| \leq |y| \cdot ||T^{-1}|| \cdot \epsilon|x|.$$

provided $|x| \leq \delta$. Of course we have that $|x| = |TT^{-1}x| = |Ty| \leq ||T|| \cdot |y|$. Using this in the previous inequality gives

$$2T^{-1}x \cdot T^{-1}[f(x) - Ax] \leq 2\epsilon||T|| \cdot ||T^{-1}|| \cdot |y|^2. \tag{28}$$

Using (27) and (28) in the estimate for $\dfrac{d|y(t)|^2}{dt}$ gives

$$\frac{d|y(t)|^2}{dt} \leq -2\rho|y(t)|^2 + 2\epsilon||T|| \cdot ||T^{-1}|| \cdot |y(t)|^2 \leq -\rho'|y(t)|^2$$

where $\rho' = (2\rho - 2\epsilon||T|| \cdot ||T^{-1}||)$, Provided we take $\epsilon > 0$ sufficiently small we get that $\rho' > 0$. Thus we obtain the required inequality (26) and we are done.

If A is diagonalisable but the eigenvalues are no longer real, then the estimate in the inequality in (27) needs to be altered slightly. Let us explain the required change by considering an example. Take $A = \begin{pmatrix} -a & b \\ -b & -a \end{pmatrix}$. Note A has eigenvalues $-a \pm bi$ and that A is already in the real Jordan normal form. Moreover,

$$\begin{aligned} y \cdot Ay &= \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \cdot \begin{pmatrix} -ay_1 + by_2 \\ -by_1 - ay_2 \end{pmatrix} \\ &= -a\left([y_1(t)]^2 + [y_2(t)]^2\right) = -a|y|^2. \end{aligned}$$

so the argument goes through. Using the real Jordan normal form theorem, the same method applies as long as A has a basis of n eigenvectors. This concludes the proof of **Theorem 8.1** in this setting. We will skip the prove in the general setting, but the next example shows what happens if there is no basis of eigenvectors.

In fact, when we prove that $x(t)$ by showing that $|y(t)|^2$ tends to zero, we use the function $U(x) := |T^{-1}(x)|^2$. Later we will call this a **Lyapounov function**.

Example 8.3

Let us consider a situation when the matrix does not have a basis of eigenvectors. Let $A = \begin{pmatrix} -1 & Z \\ 0 & -1 \end{pmatrix}$ *where* $Z \in \mathbb{R}$. *This has eigenvalues* -1 *(with double multiplicity). Take* $U(x,y) = ax^2 + bxy + cy^2$. *Then*

$$\begin{aligned} \dot{U} &= 2ax\dot{x} + b\dot{x}y + bx\dot{y} + 2cy\dot{y} \\ &= 2ax(-x + Zy) + b(-x + Zy)y + bx(-y) + 2cy(-y) \\ &= -2ax^2 + (2Za - b - b)xy + (Zb - 2c)y^2. \end{aligned}$$

Case 1: *If $Z \approx 0$, then we can take $a = 1, b = 0, c = 1$ because then $\dot{U} = -2x^2+(2Z)xy-2y^2 \leq 0$ (since $Z \approx 0$).*

Case 2: *If Z is large and $a = 1, b = 0, c = 1$ then we definitely don't get $\dot{U} \leq 0$. However, in this case we can set $b = 0$, and write*

$$\begin{aligned} \dot{U} &= -2ax^2 + (2Za)xy - 2cy^2 \\ &= -2a[x - (Z/2)y]^2 + (aZ^2/2 - 2c)y^2 \\ &= -2[x - (Z/2)y]^2 - y^2 < 0. \end{aligned}$$

where in the last line we substitutes $a = 1$ and $c = 1/2 + Z^2/4$. Thus $U = c$ corresponds to a 'flat' ellipse when Z is large.

General case: *This all seems rather ad hoc, but the Jordan normal form suggests a general method. Indeed A has an eigenvector $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (i.e. $(A+I)v_1 = 0$) and we can choose a 2nd vector v_2 so that $(A + I)v_2 = \epsilon v_1$ where $\epsilon > 0$ is small. So $v_2 = \begin{pmatrix} 0 \\ \epsilon/Z \end{pmatrix}$. Taking $T = (v_1 v_2)$ gives $T^{-1}AT = \begin{pmatrix} -1 & \epsilon \\ 0 & -1 \end{pmatrix}$. In this new coordinates we are in the same position as if $Z \approx 0$. So we can argue as in the first case.*

8.4 Hartman-Grobman

Theorem 8.4

Hartman-Grobman *Let x_0 be a singularity and that $A = Df(x_0)$ is a hyperbolic matrix. Then there exists a continuous bijection (a homeomorphism) $h\colon \mathbb{R}^n \to \mathbb{R}^n$ so that $h(x_0) = 0$ and so that near x_0,*

h sends orbits of $x' = f(x)$ to orbits of $y' = Ay$.

In other words, if $x(t)$, is a solution of $x' = f(x)$ and $x(t)$ is close to x_0 for $t \in [a,b]$ then $y(t) = h(x(t))$ is a solution of $y' = Ay$ for $t \in [a,b]$.

Remark: In other words, there exists an open set $U \ni x_0$ so that $h \circ \phi_t(x) = \phi_t^A \circ h(x)$ for each x, t so that $\cup_{0 \le s \le t} \phi_s(x) \subset U$. Here ϕ_t^A is the flow associated to $y' = Ay$ and ϕ_t the flow for $x' = f(x)$.

Remark: A **homeomorphism** is a continuous bijection whose inverse is also continuous. In Euclidean space (and 'manifolds'), this is the same as saying that it is continuous bijection.

8.5 Lyapounov functions

Using a suitable function to measure the 'distance' to a singularity, as in Theorem 8.1, is a very common method.

Definition 8.5

Let $W \subset \mathbb{R}^n$ be an open set containing x_0. $V : W \to \mathbb{R}$ is a **Lyapounov** *function for x_0 if it is C^1 and*

- $V(x_0) = 0$, $V(x) > 0$ *for* $x \in W \setminus \{x_0\}$;
- $\dot{V}(x) \le 0$ *for* $x \in W$.

Here $\dot{V}(x) := \frac{dV(x(t))}{dt}\Big|_{t=0} = DV_{x(t)} \frac{dx}{dt}\Big|_{t=0} = DV_{x(0)} f(x(0))$ *where $x(t)$ is the solution of the IVP $\dot{x} = f(x), x(0) = x$. Note that x is used both for a point $x \in \mathbb{R}^n$ and for a curve $x(t) \in \mathbb{R}^n$.*

In actual fact, we will use the notion of Lyapounov function more loosely, and for example sometimes give the same name to a function for which merely $\dot{V} \le 0$ and then we need the ideas of the proofs of Lemma 8.8 rather then the statement of this lemma itself.

Remarks: V should be thought of as a way to measure the distance to x_0. That $\dot{V} \le 0$ means that this 'distance' is non-increasing. In quite a few textbooks a Lyapounov function is one which merely satisfies the first property; let us call such functions weak-Lyapounov functions.

Warning: In some cases one calls a function Lyapounov even if it does not satisfies all its properties.

Definition 8.6

- x_0 *is called* **stable** *if for each* $\epsilon > 0$ *there exists* $\delta > 0$ *so that if* $x \in B_\delta(x_0)$ *implies* $\phi_t(x) \in B_\epsilon(x_0)$ *for all* $t \geq 0$. *(So you nearby points don't go far.)*
- x_0 *is called* **asymptotically stable** *if, it is stable and if for each* x *near* x_0*, one has* $\phi_t(x) \to x_0$.

Remark 8.7

The following lemma implies that if there exists a Lyapunov function V *with* $\dot{V} < 0$ *(outside* x_0*)) then there exists no other Lyapunov function* U *so that* $\dot{U} \geq 0$*. It is however possible that there exists a Lyapunov function* U *so that* $\dot{U} \leq 0$*; then* V *will tell you that* x_0 *is asymptotically stable, while* U *would only tell you that* x_0 *is stable.*

Lemma 8.8

Lyapounov functions

1. *If* $\dot{V} \leq 0$ *then* x_0 *is stable. Moreover,* $\phi_t(x)$ *exists for all* $t \geq 0$ *provided* $d(x, x_0)$ *is small.*
2. *If* $\dot{V} < 0$ *for* $x \in W \setminus \{x_0\}$ *then* $\forall x$ *is close to* x_0 *one gets* $\phi_t(x) \to x_0$ *as* $t \to \infty$*, i.e.* x_0 *is asymptotically stable.*

Proof : (1) It is enough to assume to consider the case that $\epsilon > 0$

is small. So take $\epsilon > 0$ so that $B_{2\epsilon}(x_0) \subset W$. Let

$$\tau := \inf_{y \in \partial B_\epsilon(x_0)} V(y).$$

Since $V > 0$ except at x_0 we get $\tau > 0$. It follows that

$$V^{-1}[0,\tau) \cap \partial B_\epsilon(x_0) = \emptyset. \tag{29}$$

Take $x \in V^{-1}[0,\tau) \cap B_\epsilon(x_0)$ Since $\phi_0(x) = x$ and $t \to V(\phi_t(x))$ is non-increasing, $\phi_t(x) \in V^{-1}[0,\tau)$ for all $t \geq 0$. Since $t \to \phi_t(x)$ is continuous curve, $\phi_0(x) = x \in B_\epsilon(x_0)$ and (29), it follows that $\phi_t(x) \in B_\epsilon(x_0)$ for all $t \geq 0$. In particular $\phi_t(x)$ remains bounded, and so $\phi_t(x)$ exists $\forall t$.

Since $V(x_0) = 0$ there exists $\delta > 0$ so that $B_\delta(x_0) \subset V^{-1}[0,\tau) \cap B_\epsilon(x_0)$. So $x \in B_\delta(x_0) \implies \phi_t(x) \in B_\epsilon(x_0)$ for all $t \geq 0$.

(2) $\dot{V} < 0$ implies that $t \to V(\phi_t(x))$ is strictly decreasing. Take $x \in B_\delta(x_0)$ and suppose by contradiction that $V(\phi_t(x))$ does not tend to 0 as $t \to \infty$. Then, since $t \mapsto V(\phi_t(x))$ is decreasing, there exists $V_0 > 0$ so that $V(\rho_t(x)) \geq V_0 > 0$. Hence $\exists \rho > 0$ with $\phi_t(x) \notin B_\rho(x_0)$ $\forall t \geq 0$. Combining this with part (1) gives that

$$\phi_t(x) \in \overline{B_\epsilon(x_0)} \setminus B_\rho(x_0) \text{ for all } t \geq 0.$$

But $\dot{V} < 0$, $\dot{V}$ is only zero at x_0 and therefore $\dot{V}$ attains its maximum in a compact set $\overline{B_\epsilon(x_0)} \setminus B_\rho(x_0)$. Hence $\exists \kappa > 0$ so that

$$\dot{V} \leq -\kappa \text{ whenever } x(t) \in \overline{B_\epsilon(x_0)} \setminus B_\rho(x_0).$$

But since $x(t)$ is in this compact set for all $t \geq 0$,

$$\dot{V} \leq -\kappa, \forall t \geq 0.$$

Hence

$$V(\phi_t(x)) - V(x) \leq -\kappa t \to -\infty \text{ as } t \to \infty,$$

contradicting $V \geq 0$. ■

Example 8.9

$$\begin{aligned} x' &= 2y(z-1) \\ y' &= -x(z-1) \\ z' &= xy \end{aligned}$$

Its linearisation is $A := \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. *Note* A *has eigenvalues* $\pm\sqrt{2}i$ *and* 0. *So* A *is not hyperbolic, and theorem 8.1 does not apply.*

Take $V(x,y,z) = ax^2 + by^2 + cz^2$. *Then*

$$\dot{V} = 2(ax\dot{x}+by\dot{y}+cz\dot{z}) = 4axy(z-1)-2bxy(z-1)+2cxyz.$$

We want $V \geq 0$ *and* $\dot{V} \leq 0$. *We can achieve this by setting* $c = 0$, $2a = b$. *This makes* $\dot{V} = 0$. *It follows that solutions stay on level sets of the function* $V = x^2 + 2y^2$. $x_0 = (0,0,0)$ *is* **not** *asymptotically stable. Strictly speaking* V *is* **not** *a Lyapounov function because* $V(0,0,z) = 0$*: more work is needed to check if* x_0 *is stable.*

Example 8.10

Consider the system $x' = -y - xy^2, y' = x - yx^2$. *The only singularity of this system is at* $(0,0)$. *Indeed, if* $x' = 0$, *then either* $y = 0$ *or* $1 + xy = 0$; *if* $y = 0$ *then* $x(1-xy) = 0$ *implies* $x = 0$; *if* $1 + xy = 0$ *then* $0 = x(1-xy) = 2x$ *implies* $x = 0$ *which contradicts* $1 + xy = 0$.

Let us show that $(0,0)$ *is asymptotically stable. To do this, take the quadratic function* $V(x,y) = x^2 + y^2$. *Then* $\dot{V} = 2x\dot{x} + 2y\dot{y} = -4x^2y^2 \leq 0$, *so* $(0,0)$ *is stable. Since* V *is decreasing (non-increasing), this implies that there exists* $V_0 \geq 0$ *so that* $V(x(t),y(t)) \downarrow V_0$. *If* $V_0 = 0$ *then the solution converges to* $(0,0)$ *as claimed. If* $V_0 > 0$ *then the solution would converge to the circle* $\{(x,y); x^2+y^2 = V_0\}$ *and so this circle would be a periodic orbit. (This follows*

Note that $\dot{V}(x,y) = 0$ *along the two axis, so we do NOT have that* $\dot{V} < 0$ *outside* $(0,0)$. *Nevertheless we can conclude that* $(0,0)$ *is asymptotically stable, see the main text.*

from the arrows.)

Note that the set $V(x,y) = x^2+y^2 = V_0$ *does not contain singular points and so there exists* $\delta > 0$ *so that* $|\dot{x}| + |\dot{y}| \geq \delta > 0$ *along this set. It follows that* $(x(t), y(t))$ *does not converge to a point of the circle* $V(x,y) = x^2 + y^2 = V_0$. *But the orbit can also not tend to a periodic orbit, since* $\dot{V} < 0$ *except when* $x = 0$ *or* $y = 0$. *(By looking at the arrows, one concludes that the orbits are tangent to circles when* $x = 0$ *or when* $y = 0$ *but otherwise spiral inwards.) It follows that* $V_0 = 0$ *and so we are done.*

8.6 The pendulum

Consider a pendulum moving along a circle of radius l, with a mass m and friction k. Let $\theta(t)$ be the angle from the vertical at time t. The force tangential to the circle is $-(kl\frac{d\theta}{dt} + mg\sin(\theta))$. So Newton's law gives
$ml\theta'' = -kl\theta' - mg\sin\theta$ i.e. $\theta'' = -(k/m)\theta' - (g/l)\sin\theta$.

Taking $\omega = \theta'$ gives

$$\begin{aligned} \theta' &= \omega \\ \omega' &= \frac{-g}{l}\sin(\theta) - \frac{k}{m}\omega. \end{aligned}$$

Singularities are $(n\pi, 0)$ which corresponds to the pendulum being in vertical position (pointing up or down). Linearizing this at $(0,0)$ gives

$$\begin{pmatrix} 0 & 1 \\ -g/l & -k/m \end{pmatrix}$$

which gives eigenvalues $(-k/m \pm \sqrt{(k/m)^2 - 4g/l})/2$.

Note that, as $l > 0$, the real part of $(-k/m \pm \sqrt{(k/m)^2 - 4g/l})/2$ is negative. (If $(k/m)^2 - 4g/l < 0$ then both e.v. are complex and if $(k/m)^2 - 4g/l > 0$ then both e.v. are real and negative.)

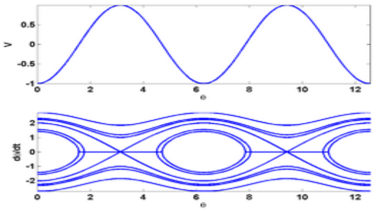


Figure 2: The phase portrait of the pendulum (no friction).

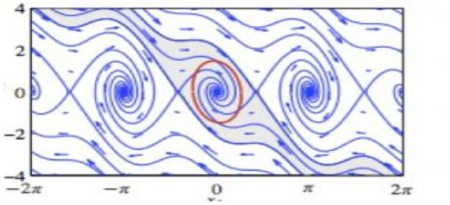


Figure 3: The phase portrait of the pendulum (with friction). The labels in the axis of this figure should have been $-4\pi, -2\pi, 0, 2\pi, 4\pi$.

Let us construct a Lyapounov function for this:

$$\begin{aligned} E &= \text{kinetic energy } + \text{ potential energy} \\ &= (1/2)mv^2 + mg(l - l\cos(\theta)) \\ &= (1/2)ml^2\omega^2 + mgl(1 - \cos(\theta)). \end{aligned}$$

Then $E \geq 0$ and $E = 0$ if and only if $\omega = 0$ and $\theta = n\pi$. Moreover,

$$\begin{aligned} \dot{E} &= ml(l\omega\omega' + g\theta' \sin\theta) \\ &= ml(l\omega(\frac{-g}{l}\sin(\theta) - \frac{k}{m}\omega) + g\omega\sin\theta) \quad \cdot \\ &= -kl^2\omega^2 \end{aligned}$$

If the friction $k > 0$ then $\dot{E} < 0$ except when $\omega = 0$. If the friction $k = 0$ then $\dot{E} = 0$ and so solutions stay on level sets of E.

8.7 Hamiltonian systems

When the friction $k = 0$ we obtain an example of a **Hamiltonian system**, i.e., a system for which there exists a function $H\colon \mathbb{R}^2 \to \mathbb{R}$ so that the equation of motion (i.e. the differential equation):

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial y}(x,y) \\ \dot{y} &= -\frac{\partial H}{\partial x}(x,y) \end{aligned}$$

H is the energy of the system, which is conserved over time:

$$\begin{aligned} \dot{H} &= \frac{\partial H}{\partial x}\dot{x} + \frac{\partial H}{\partial y}\dot{y} \\ &= \frac{\partial H}{\partial x}\frac{\partial H}{\partial y} + \frac{\partial H}{\partial y}(-\frac{\partial H}{\partial x}) \\ &= 0. \end{aligned}$$

8.8 Van der Pol's equation

In electrical engineering the following equation often arrises

$$\begin{aligned} \dot{x} &= y - x^3 + x \\ \dot{y} &= -x. \end{aligned} \tag{30}$$

This system has a singularity at $(x,y) = (0,0)$. Its linear part at $(0,0)$ is $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$. This has eigenvalues $(1 \pm \sqrt{3}i)/2$ and therefore $(0,0)$ is a source. What happens with other orbits?

Theorem 8.11

There is one periodic solution of this system and every non-equilibrium solution tends to this periodic solution.

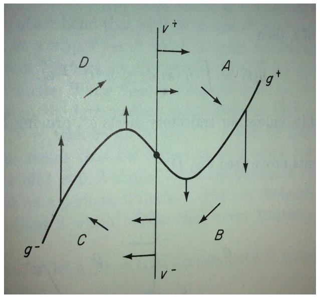


Figure 4: The phase portrait of the van der Pol equation.

The proof of this theorem will occupy the remainder of this section.

Define

$$v^{\pm} = \{(x,y); \pm y > 0, x = 0\} \text{ and } g^{\pm} = \{(x,y); \pm x > 0, y = x^3 - x\}.$$

This splits up $\mathbb{R}^2$ in regions A, B, C, D where horizontal and vertical speed is positive/negative.

$$\begin{aligned} \dot{x} &= y - x^3 + x \\ \dot{y} &= -x. \end{aligned}$$

Lemma 8.12

For any $p \in v^+$, $\exists t > 0$ with $\phi_t(p) \in g^+$.

Proof: Define $(x_t, y_t) = \phi_t(p)$.

- Since $x'(0) > 0$, $\phi_t(p) \in A$ for $t > 0$ small.
- $x' > 0, y' < 0$ in A. So the only way the curve $\phi_t(p)$ can leave the region $A \cap \{(x,y); y < y_0\}$ is via g^+.
- So $\phi_t(p)$ cannot go to infinity before hitting g^+.
- Hence $T = \inf\{t > 0; \phi_t(p) \in g^+\}$ is well-defined.
- We need to show $T < \infty$.

- Choose $t_0 \in (0,T)$ and let $a = x_{t_0}$. Then $a > 0$ and $x_t \geq a$ for $t \in [t_0, T]$.
- Hence $\dot{y} \leq -a$ for $t \in [t_0, T]$ and therefore $y(t) - y(t_0) \leq -a(t - t_0)$ for $t \in [t_0, T]$.
- $T = \infty \implies \lim_{t\to\infty} y(t) \to -\infty$ which gives a contradiction since $(x(t), y(t)) \in A$ for $t \in (0,T)$. ■

Similarly

Lemma 8.13

For any $p \in g^+,\ \exists t > 0$ *with* $\phi_t(p) \in v^-$.

For each $y > 0$ define $F(y) = \phi_t(0,y)$ where $t > 0$ is minimal so that $\phi_t(0,y) \in v^-$. Similarly, define for $y < 0$ define $F(y) = \phi_t(0,y)$ where $t > 0$ is minimal so that $\phi_t(0,y) \in v^+$. By symmetry $F(-y) = -F(y)$. Hence

$$P(p) = |F(|F(p)|)|.$$

Define the **Poincaré first return** map to v^+ as

$$P \colon v^+ \to v^+ \text{ by } (0,y) \mapsto (0, F^2(y)).$$

$P(p) = \phi_t(p)$ where $t > 0$ is minimal so that $\phi_t(p) \in v^+$.

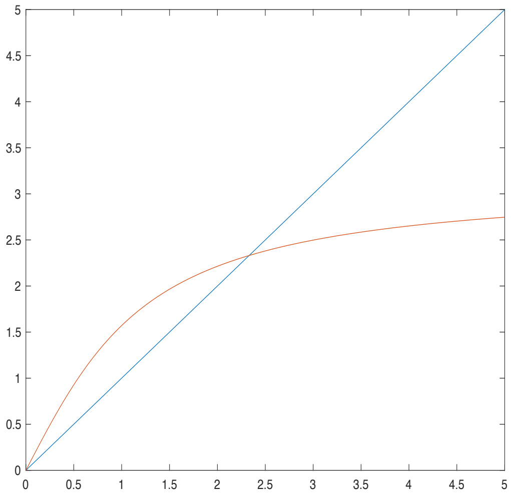


Figure 5: We will show that the graph of $x \mapsto |F(x)|$ looks like this. Since $G(x) = |F(|F(x)|)|$ the graph of G will look similar, and the unique fixed point q of F is necessarily the unique fixed point of G.

Define

$$p^* = (0, y^*) \in v^+ \quad \begin{array}{l}\text{so that } \exists t > 0 \text{ with } \phi_t(p^*) = (1, 0) \\ \text{and } \phi_s(p^*) \in A \text{ for } 0 < s < t.\end{array}$$

Lemma 8.14

1. *$P \colon v^+ \to v^+$ is continuous and increasing (here we order v^+ as $(0, y_1) < (0, y_2)$ when $y_1 < y_2$);*
2. *$P(p) > p$ when $p \in [0, p^*]$;*
3. *$P(p) < p$ when p is large;*
4. *$P \colon v^+ \to v^+$ has a unique attracting fixed point $q \in v^+ \in [p^*, \infty)$.*

Proof: The proof of (1): That P is continuous follows from the property that $(t, x) \mapsto \phi_t(x)$ is continuous and from a theorem we will only prove in the next chapter, specifically Theorem . Uniqueness of solns $\implies$ orbits don't cross $\implies$ P is increasing.

Instead of (2), (3) and (4) we shall prove the following statement:

$$[p^*, \infty) \ni \quad \begin{array}{c} p \mapsto \delta(p) := |F(p)|^2 - |p|^2 \\ \text{is strictly decreasing}\end{array} \tag{31}$$

$$\delta(p) > 0 \text{ for } p \in [0, p^*] \tag{32}$$

$$\delta(p) \to -\infty \text{ as } p \to \infty \tag{33}$$

Claim: This implies that $[p^*, \infty) \ni p \mapsto |P(p)|^2 - |p|^2$ is strictly decreasing.

Proof: Since $[0, \infty) \ni p \mapsto |F(p|$ is increasing, $|F(p^*)| > |p^*|$ and $[p^*, \infty) \ni p \mapsto |F(p)|^2 - |p|^2$ is strictly decreasing, it follows that $[p^*, \infty) \ni p \mapsto (|F|F(p)|)|^2 - |F(p)|^2$ is also strictly decreasing. Combining this with (31) proves the claim.

From (31)-(33) we obtain that the graph of $x \mapsto |F(x)|$ looks as in the above figure, and therefore imply assertions (2) and (3) and (4) in the lemma (see lecture). So it is enough to prove (31)-(33). Note that the fixed point q of F, is a zero of the function $p \mapsto \delta(p)$ and therefore $p > p^*$.

Step 1: A useful expression for $\delta(p)$. Define $U(x,y) = x^2 + y^2$. Pick $p \in v^+$ and let $\tau > 0$ be minimal so that $\phi_\tau(p) \in v^-$. (So $\phi_\tau(p) = F(p)$.) Hence

$$\begin{aligned} \delta(p): &= |F(p)|^2 - |p|^2 = U(\phi_\tau(p)) - U(\phi_0(p)) \\ &= \textstyle\int_0^\tau \dot{U}(\phi_t(p))\,dt. \end{aligned}$$

Note

$$\begin{aligned} \dot{U} &= 2x\dot{x} + 2y\dot{y} = \\ &= 2x(y - x^3 + x) + 2y(-x) = -2x(x^3 - x) = 2x^2(1 - x^2) \end{aligned}.$$

Hence

$$\delta(p) = 2\int_0^\tau [x(t)]^2(1 - [x(t)]^2)dt = 2\int_0^\tau x^2(1 - x^2)\,dt.$$

Here γ is the curve $[0,\tau] \ni t \to \phi_t(p)$.

Step 2: $\delta(p) > 0$ when $p \in (0,p^*]$. If $p \in (0,p^*]$ then $\delta(p) > 0$ because then $(1 - [x(t)]^2) \geq 0$ for all $t \in [0,\tau)$. So $|F(p)| > p$ and $P(p) > p$ for $p \in (0,p^*]$. So a periodic orbit **cannot** intersect the line segment $[0,p^*]$ in v^+.

Step 3: $\delta(p)$ when $p > p^*$. Choose $0 = \tau_0 < \tau_1 < \tau_2 < \tau_3 = \tau$ so that

- the curve $[\tau_1,\tau_2] \ni t \mapsto \gamma(t)$ has both endpoint on the line $x = 1$;

- the curve $[\tau_0, \tau_1] \ni t \mapsto \gamma(t)$ connects $p \in v^+$ to the line $x = 1$;
- the curve $[\tau_2, \tau_3] \ni t \mapsto \gamma(t)$ connects $F(p) \in v^-$ to the line $x = 1$.

Now consider

$$\delta_i(p) := 2 \int_{\tau_{i-1}}^{\tau_i} x^2(t)(1 - x^2(t))\, dt \text{ for } i = 1, 2, 3.$$

Note that $\delta(p) = \delta_1(p) + \delta_2(p) + \delta_3(p)$.

Step 4: $\delta_1(p)$ **is decreasing when** $p > p^*$**.**

- γ_1 is a curve which can be regarded as function of x.
- Hence we can write

  $$\int_0^{\tau_1} x^2(1-x^2)\, dt = \int_0^{\tau_1} \frac{x^2(1 - x^2)}{dx/dt}\, dx = \int_0^1 \frac{x^2(1 - x^2)}{y(x) - (x^3 - x)}\, dx$$

  where $y(x)$ is so that $(x, y(x))$ is the point on the curve γ_1 defined by $[\tau_0, \tau_1] \ni t \mapsto \gamma(t)$.
- As p moves up, the curve γ_1 (connecting $p \in v^+$ to a point on the line $x = 1$) moves up and so $y(x) - (x^3 - x)$ (along this curve) increases.
- Hence $p \to \delta_1(p) = 2 \int_0^{\tau_1} x^2(1 - x^2)\, dt$ decreases as p increases.

Step 5: $\delta_2(p)$ **is decreasing when** $p > p^*$**.**

- Along γ_2, the solution $x(t)$ is a function of $y \in [y_1, y_2]$ (where $(1, y_1)$, $y_1 > 0$ and $(1, y_2)$, $y_2 < 0$) are the intersections points of γ with the line $x = 1$.
- Since $-x = dy/dt$ we get

  $$\begin{aligned} \textstyle\int_{\tau_1}^{\tau_2} x^2(1 - x^2)\, dt &= \textstyle\int_{y_1}^{y_2} -x(y)(1 - [x(y)]^2)\, dy \\ &= \textstyle\int_{y_2}^{y_1} x(y)(1 - [x(y)]^2)\, dy \end{aligned}$$

in the 2nd integral one has $\int_{y_1}^{y_2}$ because that corresponds to the way the curve γ_1 is oriented.

- Since $x(y) \geq 1$ along γ_2 (and $y_2 < y_1$), this integral is negative.
- As p increases, the interval $[y_1, y_2]$ gets larger, and the curve γ_2 moves to the right and so $x(y)(1-[x(y)]^2)$ decreases. It follows that $\delta_2(p)$ decreases as p increases.
- It is not hard to show that $\delta_2(p) \to -\infty$ as $p \to \infty$, see lecture.

Exactly as for $\delta_1(p)$, one also gets that $\delta_3(p)$ decreases as p increases. This completes the proof of the equation (**??**) and therefore the proof of Lemma 8.14 and Theorem 8.11. ■

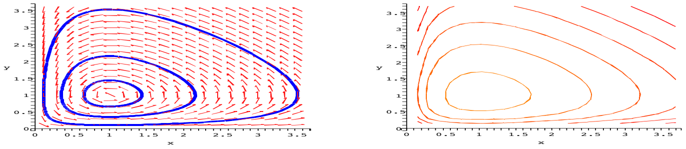
8.9 Population dynamics

A common predator-prey model is the equation

$$\begin{aligned} \dot{x} &= (A - By)x \\ \dot{y} &= (Cx - D)y. \end{aligned} \quad \text{where } A, B, C, D > 0$$

Here x are the number of rabbits and y the number of foxes. For example, $x' = Ax - Bxy$ expresses that rabbits grow with speed A but that the proportion that get eaten is a multiple of the number of foxes.

Let us show that the orbits look like the following diagram:



- Singularities are $(x, y) = (0, 0)$ and $(x, y) = (D/C, A/B)$.
- If p is on the axis, then $\phi_t(x)$ is on this axis for all $t \in \mathbb{R}$.
- At $(0, 0)$ the linearisation is $\begin{pmatrix} A & 0 \\ 0 & -D \end{pmatrix}$, so eigenvalues are $A, -D$ and $(0, 0)$ is a saddle point.
- At $(x, y) = (D/C, A/B)$ the linearisation is $\begin{pmatrix} A - By & -Bx \\ Cy & Cx - D \end{pmatrix} =$ $\begin{pmatrix} 0 & -BD/C \\ CA/B & 0 \end{pmatrix}$ which has eigenvalues $\pm ADi$ (purely imaginary).

$$\begin{aligned} \dot{x} &= (A - By)x \\ \dot{y} &= (Cx - D)y. \end{aligned} \quad \text{where } A, B, C, D > 0$$

- Analysing the direction field, suggests that orbits cycle around $(D/C, A/B)$ (see lecture).

- Try to find Lyapounov of the form $H(x, y) = F(x) + G(y)$.

- $\dot{H} = F'(x)\dot{x} + G'(y)\dot{y} = xF'(x)(A - By) + yG'(y)(Cx - D)$.

- If we set (that is, insist on) $\dot{H} = 0$ then we obtain
$$\frac{xF'}{Cx - D} = \frac{yG'}{By - A} \tag{34}$$

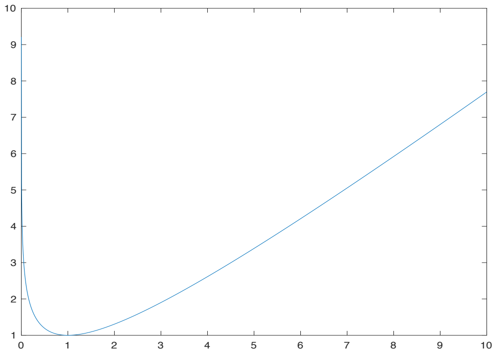
- LHS of (34) only depends on x and RHS only on y. So expression in (34) $= const$.

- We may as well set $const = 1$. This gives $F' = C - D/x$ and $G' = B - A/y$.

- So $F(x) = Cx - D \log x, G(y) = By - A \log y + F_0$ and $H(x, y) = Cx - D \log x + By - A \log y + G_0$ where F_0, G_0 are constants. Note that $F(x)$ and $G(y)$ both have unique minima at $x = D/C$ and $y = A/B$. Let us take F_0, G_0 so that the minima values so that $F(D/C) = 0$ and $G(A/B) = 0$. Then $F, G \geq 0$, F, G both have their minima values 0 and go to infinity as $x, y \to 0$ or $x, y \to \infty$, see the figure.

  From this one concludes that the level sets $\{(x, y) : H(x, y) = c\}$ are smooth closed curves when $c > 0$, a single point if $c = 0$ and empty when $c < 0$.

Summarising:



Theorem 8.15

Take $(x, y) \neq (D/C, A/B)$ with $x, y > 0$ and consider its orbits under

$$\begin{aligned} \dot{x} &= (A - By)x \\ \dot{y} &= (Cx - D)y. \end{aligned} \quad \text{where } A, B, C, D > 0.$$

Then $t \mapsto \phi_t(x, y)$ is periodic (i.e. is a closed curve).

Proof: Take $H_0 = H(x, y)$ and let $\Sigma = \{(u, v); H(u, v) = H_0\}$.

- The orbit $\phi_t(x, y)$ stays on the level set Σ of H.
- It moves with positive speed.
- So it returns in finite time.
- Orbits exist for all time, because it remains on Σ (and therefore cannot go to infinity). ■

9 Dynamical Systems

So far we saw:

- Most differential equations cannot be solved explicitly.
- Nevertheless in many instances one can still prove many properties of its solutions.
- The point of view taken in the field **dynamical systems** is to concentrate on
  - attractors and limit sets: what happens *eventually*;
  - statistical properties of orbits.

In this chapter we will discuss a result which describes the planar case (i.e. the two-dimensional case).

Throughout the remainder of these notes, we will tacitly assume the solution $\phi_t(x)$ through x exists for all $t \geq 0$.

9.1 Limit Sets

Let ϕ_t be the flow of a dynamical system and take a point x.

Definition 9.1

*Then the ω-**limit set of** x, denoted by $\omega(x)$, is defined as*

$$\{y; \exists t_n \to \infty \text{ so that } \phi_{t_n}(x) \to y\}.$$

So $\omega(x)$ describes where the point x eventually goes. It turns out that $\omega(x)$ is a closed set, see the next lemma, but it is possible that $\omega(x) = \emptyset$.

Lemma 9.2

$\omega(x)$ is closed.

Proof : Take a point $y \notin \omega(x)$. Then there exists no sequence $t_n \to \infty$ so that $\phi_{t_n}(x) \to y$. So there exists an open neighbourhood U of y so that $\phi_t(x) \cap U = \emptyset$ for all t large. But then each point in U is in the complement of $\omega(x)$. Hence the complement of $\omega(x)$ is open, proving the lemma. ■

We say that x lies on a **periodic orbit** if $\phi_T(x) = x$ for some $T > 0$. The smallest such $T > 0$ is called the period of x. Note that then

- $\gamma = \cup_{t \in [0,T)} \phi_t(x)$ is closed curve without self-intersections, and
- $\omega(x) = \gamma$.

9.2 Local sections

Definition 9.3

A hyperplane $S \ni p$ in $\mathbb{R}^n$ is a called a **local section at** *p for the autonomous differential equation $x' = f(x)$ if: $f(p) \neq 0$ and the vector $f(p)$ at p does not lie in the hyperplane.*

Definition 9.4

A subset S of a hyperplane in $\mathbb{R}^n$ is a called a **section** *for the autonomous differential equation $x' = f(x)$ if: for every $p \in S$, $f(p) \neq 0$ and the vector $f(p)$ at p does not lie in the hyperplane.*

Theorem 9.5 (Flow Box Theorem)

Assume S is a local section at p and assume q is that $\phi_{t_0}(q) = p$ for some minimal $t_0 > 0$. Then there exists a neighbourhood U of q and a smooth function $\tau\colon U \to \mathbb{R}$ so that $\tau(q) = t_0$ and so that $\phi_{\tau(x)}(x) \in S$ for each $x \in U$.

Remark 9.6

If $t_0 > 0$ is the **minimal** *positive time so that $\phi_{t_0}(q) \in S$ then $\tau(x) > 0$ will also be minimal so that $\phi_{\tau(x)}(x) \in S$. $\tau(x)$ is then called the* **first arrival time** *and the map $P(x) = \phi_{\tau(x)}(x)$ the Poincaré entry map to S.*

Proof : Let $g\colon \mathbb{R}^n \to \mathbb{R}$ be an affine function of the form $g(x) = a \cdot x + b$ which determines S, i.e. so that $S = \{x; g(x) = 0\}$. Define $G(x,t) = g(\phi_t(x))$. Then $G(q,t_0) = g(\phi_{t_0}(q)) = g(p) = 0$. Moreover,

$$\begin{aligned}\frac{\partial G}{\partial t}(q,t_0) &= Dg(\phi_{t_0}(q))\frac{\partial \phi_t}{\partial t}(q)\big|_{t=t_0} = Dg(p)f(\phi_{t_0}(q)) \\ &= Dg(p)f(p) = a \cdot f(p) \neq 0 \text{ (because } S \text{ is a section at } p).\end{aligned}$$

Hence by the implicit function theorem there exists a function $x \mapsto \tau(x)$ so that $G(x,\tau(x)) = 0$ for x near q. Hence $\phi_{\tau(x)} \in S$ for x near q. ■

Remark 9.7

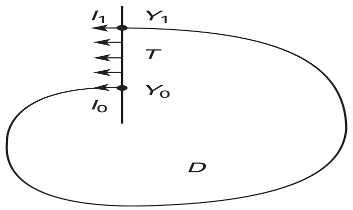
1. *If S is a section at p and x is close to p, then there exists t close to zero so that $\phi_t(x) \in S$ and so that $\phi_t(x)$ is still close to zero. This follows from the previous theorem (by taking $q = p$ and $t_0 = 0$).*

2. *From the precious remark it follows that if $\phi_{t_n}(x) \to p$ for some sequence $t_n \to \infty$ then there exists $t'_n \to \infty$ so that $\phi_{t'_n}(x) \to p$ and $\phi_{t'_n}(x) \in S$ and so that*

$|t_{n'} - t_n| \to 0$.

3. *If p lies on a periodic orbit with period T and S is a local section at p, then $\phi_T(p) = p$ and then there exists a neighbourhood U of p and a map $P\colon S \cap U \to S$ so that $P(p) = p$. This is called the* **Poincaré return map**.

4. *As in the example of the van der Pol equation, one can use the Poincaré map to check whether the periodic orbit is attracting.*



9.3 Planar Systems, i.e. ODE in $\mathbb{R}^2$

Theorem 9.8

Let S be a section for a planar differential equation, so S is a piece of a straight line. Let $\gamma = \cup_{t\geq 0}\phi_t(x)$ and let $y_0, y_1, y_2 \in S \cap \gamma$. Then y_0, y_1, y_2 lie ordered on γ if and only if they lie ordered on S.

Proof : Take $y_0, y_1, y_2 \in \gamma \cap c$. Assume that y_0, y_1, y_2 are *consecutive* points on γ, i.e. assume $y_2 = \phi_{t_2}(y_0)$, $y_1 = \phi_{t_1}(y_0)$ with $t_2 > t_1 > 0$. Let $\gamma' = \cup_{0\leq s\leq t_1}\phi_s(y_0)$ and consider the arc c in S connecting y_0 and y_1. Then

- $c \cup \gamma'$ is a closed curve which bounds a compact set D (here we use a special case of a deep result namely the Jordan theorem).
- Either all orbits enter D along c or they all leave D along c.
- Either way, since the orbit through y does not have self-intersections and because of the orientation of $x' = f(x)$ along S, $\phi_{t_2}(y_0)$ cannot intersect c, see figure. ■

In this chapter we tacitly assume that if γ is a closed curve in $\mathbb{R}^2$ without self-intersections, then the complement of γ has two connected components: one bounded one and the other unbounded. This result is called the *Jordan curve theorem* which looks obvious, but its proof is certainly not easy. It can be proved using algebraic topology.

Lemma 9.9

If $y \in \omega(x)$. Then the orbit through y intersects any section at most once.

Proof: Assume by contradiction that $y_1 = \phi_u(y)$ and $y_2 = \phi_v(y)$ (where $v > u \geq 0$) are contained on a local section S.

Since $y \in \omega(x)$ where exists $t_n \to \infty$ so that $\phi_{t_n}(x) \to y$. Hence $\phi_{t_n+u}(x) \to y_1$ and $\phi_{t_n+v}(x) \to y_2$. Because $y_1, y_2 \in S$, this implies that for n large there exists $u_n, v_n \to 0$ so that

$$\phi_{t_n+u+u_n}(x) \in S, \phi_{t_n+v+v_n}(x) \in S \text{ for all } n \geq 0,$$

$$\phi_{t_n+u+u_n}(x) \to y_1, \phi_{t_n+v+v_n}(x) \to y_2 \text{ as } n \to \infty.$$

Here we use Remark 9.7(2).

Take $n' > n$ so large that

$$t_n + u + u_n < t_n + v + v_n < t_{n'} + u + v_{n'}. \qquad (35)$$

Here the first inequality holds when n is sufficiently large because $u_n, v_n \to 0$ and $u < v$ and the second inequality holds when $n' >> n$ because then $t_{n'} - t_n$ is large and $v_n, v_{n'} \approx 0$.

Provided $n' > n$ are large, the three points

$$\phi_{t_n+u+u_n}(x), \phi_{t_n+v+v_n}(x), \phi_{t_{n'}+u+v_{n'}}(x)$$

do NOT lie ordered on S. Indeed, the first and last points are close to y_1 and the middle one is close to y_2.

This and (35) contradict the previous theorem. ■

9.4 Poincaré Bendixson

Theorem 9.10 (*Poincaré-Bendixson Theorem*)

Consider a planar differential equation, take $x \in \mathbb{R}^2$ and assume that $\omega := \omega(x)$ is non-empty, bounded and does not contain a singular point of the differential equation. Then ω is a periodic orbit.

That is, we have an autonomous differential equation in $\mathbb{R}^2$, $\dot{x} = f(x)$ with $x \in \mathbb{R}^2$.

Proof :

- Assume that ω does not contain a singular point.
- Take $y \in \omega$. Then there exists $s_m \to \infty$ so that $\phi_{s_m}(x) \to y$. Hence for each fixed $t > 0$, $\phi_{s_m+t}(x) \to \phi_t(y)$ as $m \to \infty$. It follows that the forward orbit $\gamma = \cup_{t\geq 0}\phi_t(y)$ is contained in ω. Since ω is compact, *any* sequence $\phi_{t_n}(y)$ has a convergent subsequence (which is contained in ω). Hence $\omega(y) \neq \emptyset$ and $\omega(y) \subset \omega$.
- Take $z \in \omega(y)$. Since z is not a singular point, there exists a section S containing z. Since $z \in \omega(y)$, there exists $t_n \to \infty$ so that $\phi_{t_n}(y) \to z$ and $\phi_{t_n}(y) \in S$.
- By the previous lemma, $\phi_{t_n}(y) = \phi_{t_{n'}}(y)$ for all n, n'. So $\exists T > 0$ so that $\phi_T(y) = y$ and y lies on a periodic orbit.
- We will skip the proof that ω is *equal* to the orbit through y. ■

9.5 Consequences of Poincaré-Bendixson

Definition 9.11

We say that A is a forward invariant domain in $\mathbb{R}^2$ if $x \in A$ implies that $\phi_t(x) \in A$ for $t > 0$.

Using the Lefschetz index formula (which we have not discuss in this course and is related to the Euler index) one can deduce the following:

Theorem 9.12

Let γ be a periodic orbit of a differential equation $x' = f(x)$ in the plane surrounding a region D. Then

- *D contains a singularity;*
- *if, moreover, all singularities of f are hyperbolic, then D contains a singularity which is either a sink or a source.*

9.6 Further Outlook

- The Poincaré Bendixson theorem implies that planar differential equations cannot have 'chaotic' behaviour.

- Differential equations in dimension ≥ 3 certainly can have chaotic behaviour, see the 3rd year course *dynamical systems* (M3PA23) and for example `http://www.youtube.com/watch?v=ByH8_nKD-ZM` and can undergo bifurcations (discussed in the 3rd year course *bifurcatiom theory M3PA24*).

- To describe their statistical behaviour one uses probabilistic arguments; this area of mathematics is called *ergodic theory*. This is a 4th year course (M4PA36). For more information see for example, `http://en.wikipedia.org/wiki/Ergodic_theory`

- Instead of differential equations one also studies discrete dynamical systems, $x_{n+1} = f(x_n)$. When $f\colon \mathbb{C} \to \mathbb{C}$ is a polynomial this leads to the study of *Julia sets* using tools from complex analysis. For more information, see `http://en.wikipedia.org/wiki/Julia_set`.

- Ideas from the field of dynamical systems are increasingly used in modern applications of mathematics in for example biology, economics, machine learning, game theory etc.

- Next year I will be teaching a course on this topic: *games and dynamics (M3PA48)*, which will cover various learning models. For example, I will discuss reinforcement learning, which is used by Artificial Intelligence companies such as Deep Mind.

9.7 Dynamical Systems

Dynamical systems is an extremely active area, and is both interesting for people focusing on pure as well as those more interested in applied mathematics.

For example, Fields Medalists whose work is in or related to this area, include: Avilla (2014, complex dynamics), Lindenstrauss (2010, ergodic theory), Smirnov (2010, part of his work relates to complex dynamics), Tao (2006, part of his work related to ergodic theory), McMullen (1998, complex dynamics), Yoccoz (1994, complex dynamics), Thurston (1982, a significant amount of work was about low and complex dynamics), Milnor (1962, his current work is in complex dynamics).

Applied dynamicists often aim to understand specific dynamical phenomena, related to for example biological systems, network dynamics, stability and bifurcation issues etc.

One of the appeals of dynamical systems that it uses mathematics from many branches of mathematics, but also that it is so relevant for applications.

Appendix A Multivariable calculus

Some of you did not do multivariable calculus. This note provides a crash course on this topic and includes some very important theorems about multivariable calculus which are not included in other 2nd year courses.

A.1 Jacobian

Suppose that $F\colon U \to V$ where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^p$. We say that F is differentiable at $x \in U$ if there exists a linear map $A\colon \mathbb{R}^n \to \mathbb{R}^m$ (i.e. a $m \times n$ matrix A)

$$\frac{|(F(x+u) - F(x)) - Au|}{|u|} \to 0$$

as $u \to 0$. In this case we define $DF_x = A$.

- In other words $F(x+u) = F(x) + Au + o(|u|)$. ($A$ is the linear part of the Taylor expansion of F).
- How to compute DF_x? This is just the Jacobian matrix, see below.
- If $f\colon \mathbb{R}^n \to \mathbb{R}$ then Df_x is a $1 \times n$ matrix which is also called grad(f) or $\nabla f(x)$.

Example A.1

Let $F(x,y) = \begin{pmatrix} x^2 + yx \\ xy - y \end{pmatrix}$ *then*

$$DF_{x,y} = \begin{pmatrix} 2x + y & x \\ y & x - 1 \end{pmatrix}.$$

Usually one denotes by $(Df_\xi)u$ is the directional derivative of f (in the direction u) at the point ξ.

Example A.2

If $F(x,y) = \begin{pmatrix} x^2+yx \\ xy-y \end{pmatrix}$ *and* $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ *then* $(DF_{x,y})e_1 = \begin{pmatrix} 2x+y & x \\ y & x-1 \end{pmatrix} e_1 = \begin{pmatrix} 2x+y \\ y \end{pmatrix}$. *This is what you get when you fix* y *and differentiate w.r.t.* x *in* $F(x,y)$.

For each fixed y *one has a curve* $x \mapsto F(x,y) = \begin{pmatrix} x^2+yx \\ xy-y \end{pmatrix}$ *and* $(DF_{x,y})e_1 = \begin{pmatrix} 2x+y \\ y \end{pmatrix}$ *gives its speed vector.*

Remark: Sometimes one writes $DF(x,y)u$ instead of $DF_{x,y}u$.

If u is the i-th unit vector e_i then one often writes $D_iF_{x,y}$ and if $i=1$ something like $D_xF(x,y)$.

Theorem A.3 (*Multivariable Mean Value Theorem*)

If $f\colon \mathbb{R} \to \mathbb{R}^m$ *is continuously differentiable then* $\forall x,y \in \mathbb{R}$ *there exists* $\xi \in [x,y]$ *so that* $|f(x)-f(y)| \leq |Df_\xi||x-y|$.

Proof: By the Main Theorem of integration, $f(y)-f(x) = \int_x^y Df_s\,ds$ (where Df_t is the $n\times 1$ matrix (i.e. vertical vector) of derivatives of each component of f. So

$$\begin{aligned} |f(x)-f(y)| &= |\textstyle\int_x^y Df_s\,ds| \leq \int_x^y |Df_s|\,ds \\ &\leq \max_{s\in(x,y)} |Df_s|\;|x-y| \leq |Df_\xi||x-y| \end{aligned}$$

for some $\xi \in [x,y]$. ■

Corollary: If $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable then for each $x,y \in \mathbb{R}^n$ there exists ξ in the arc $[x,y]$ connecting x and y so that $|f(x)-f(y)| \leq |Df_\xi(u)||x-y|$ where $u=(x-y)/|x-y|$. **Proof:** just consider f restricted to the line connecting x,y and apply the previous theorem.

A.2 The statement of the Inverse Function Theorem

Theorem A.4 (*The Inverse Function Theorem*)

Let $U \subset \mathbb{R}^n$ be open, $p \in U$ and $F\colon U \to \mathbb{R}^n$ be continuously differentiable and suppose that the matrix DF_p is invertible. Then there exist open sets $W \subset U$ and $V \subset \mathbb{R}^n$ with $p \in W$ and $F(p) \in V$, so that $F\colon W \to V$ is a bijection and so that its inverse $G\colon V \to W$ is also differentiable.

Definition A differentiable map $F\colon U \to V$ which has a differentiable inverse is called a **diffeomorphism**.

Proof: Without loss of generality we can assume that $p = 0 = F(p)$ (just apply a translation). By composing with a linear transformation we can even also assume $DF_0 = I$. Since we assume that $x \mapsto DF_x$ is continuous, there exists $\delta > 0$ so that

$$|| \, I - DF_x|| \leq 1/2 \text{ for all } x \in \mathbb{R}^n \text{ with } |x| \leq 2\delta. \qquad (36)$$

Here, as usual, we define the norm of a matrix A to be

$$||A|| = \sup\{|Ax|; |x| = 1\}.$$

Given y with $|y| \leq \delta/2$ define the transformation

$$T_y(x) = y + x - F(x).$$

Note that

$$T_y(x) = x \Longleftrightarrow F(x) = y.$$

So finding a fixed point of T_y gives us the point x for which $G(y) = x$, where G is the inverse of F that we are looking for.

We will find x using the Banach Contraction Mapping Theorem.

(**Step 1**) By (36) we had $|| \, I - DF_x|| \leq 1/2$ when $|x| \leq 2\delta$. Therefore, the Mean Value Theorem applied to $x \mapsto x - F(x)$ gives

$$|x - F(x) - (0 - F(0))| \leq \frac{1}{2}|x - 0| \text{ for } |x| \leq 2\delta$$

Therefore if $|x| \leq \delta$ (and since $|y| \leq \delta/2$),

$$|T_y(x)| \leq |y| + |x - F(x)| \leq \delta/2 + \delta/2 = \delta.$$

So T_y maps the closed ball $B := B_\delta(0)$ into itself.

(**Step 2**) $T_y \colon B \to B$ is a contraction since if $x, z \in B_\delta(0)$ then $|x-z| \leq 2\delta$ and so we obtain by the Mean Value Theorem again

$$|T_y(x) - T_y(z)| = |x - F(x) - (z - F(z))| \leq \frac{1}{2}|x - z|. \quad (37)$$

(**Step 3**) Since $B_\delta(0)$ is a complete metric space, there exists a unique $x \in B_\delta(0)$ with $T_y(x) = x$. That is, we find a unique x with $F(x) = y$.

(**Step 4**) The upshot is that for each $y \in B_{\delta/2}(0)$ there is precisely one solution $x \in B_\delta(0)$ of the equation $F(x) = y$. Hence there exists $W \subset B_\delta(0)$ so that the map

$$F \colon W \to V := B_{\delta/2}(0)$$

is a bijection. So $F \colon W \to V$ has an inverse, which we denote by G.

(**Step 5**) G **is continuous:** Set $u = F(x)$ and $v = F(z)$. Applying the triangle inequality in the first inequality and equation (37) in the 2nd inequality we obtain,

$$|x - z| = |(x - z) - (F(x) - F(z)) + (F(x) - F(z))| \leq$$

$$\leq |(x - z) - (F(x) - F(z))| + |F(x) - F(z)| \leq$$

$$\leq \frac{1}{2}|x - z| + |F(x) - F(z)|.$$

So $|G(u) - G(v)| = |x - z| \leq 2|F(x) - F(z)| = 2|u - v|$.

(Step 6) G **is differentiable:**

$|(G(u)-G(v))-(DF_z)^{-1}(u-v)| = |x-z-(DF_z)^{-1}(F(x)-F(z))| \leq$

$||(DF_z)^{-1}||\cdot|DF_z(x-z)-(F(x)-F(z))| = o(|x-z|) = 2o(|u-v|)$.

as $||(DF_z)^{-1}||$ is bounded, using the definition and the last inequality in step 5. Hence

$$|G(u) - G(v) - (DF_z)^{-1}(u - v)| = o(|u - v|)$$

proving that G is differentiable and that $DG_v = (DF_z)^{-1}$.

Example A.5

Consider the set of equations

$$\frac{x^2 + y^2}{x} = u, \sin(x) + \cos(y) = v.$$

Given (u, v) *near* $(u_0, v_0) = (2, \cos(1) + \sin(1))$ *is it possible to find a unique* (x, y) *near to* $(x_0, y_0) = (1, 1)$ *satisfying this set of equations? To check this, we define*

$$F(x, y) = \begin{pmatrix} \frac{x^2+y^2}{x} \\ \sin(x) + \cos(y) \end{pmatrix}.$$

The Jacobian matrix is

$$\begin{pmatrix} \frac{x^2-y^2}{x^2} & \frac{2y}{x} \\ \cos(x) & -\sin(y) \end{pmatrix}.$$

The determinant of this is $\frac{y^2-x^2}{x^2}\sin(y) - \frac{2y}{x}\cos(x)$ *which is non-zero near* $(1, 1)$. *So* F *is invertible near* $(1, 1)$ *and for every* (u, v) *sufficiently close to* (u_0, v_0) *one can find a unique solution near to* (x_0, y_0) *to this set of equations. Near* $(\pi/2, \pi/2)$ *the map* F *is probably not invertible.*

A.3 The Implicit Function Theorem

Theorem A.6 (*Implicit Function Theorem*)

Let $F\colon \mathbb{R}^p \times \mathbb{R}^n \to \mathbb{R}^n$ be differentiable and assume that $F(0,0) = 0$. Moreover, assume that $n \times n$ matrix obtained by deleting the first p columns of the matrix $DF_{0,0}$ is invertible. Then there exists a function $G\colon \mathbb{R}^p \to \mathbb{R}^n$ so that for all (x,y) near $(0,0)$

$$y = G(x) \Longleftrightarrow F(x,y) = 0.$$

The proof is a fairly simple application of the inverse function theorem, and won't be given here. The $\mathbb{R}^p$ part in $\mathbb{R}^p \times \mathbb{R}^n$ can be thought as parameters.

Example A.7

Let $f(x,y) = x^2 + y^2 - 1$. Then one can consider this as locally as a function $y(x)$ when $\partial f/\partial y = 2y \neq 0$.

Example A.8

Consider the following equations:

$$\begin{aligned} x^2 - y^2 - u^3 + v^2 + 4 &= 0, \\ 2xy + y^2 - 2u^2 + 3v^4 + 8 &= 0. \end{aligned}$$

Can one write u, v as a function of x, y in a neighbourhood of the solution $(x,y,y,v) = (2,-1,2,1)$? To see this, define

$$F(x,y,u,v) = (x^2 - y^2 - u^3 + v^2 + 4, 2xy + y^2 - 2u^2 + 3v^4 + 8).$$

We have to consider the part of the Jacobian matrix which concerns the derivatives w.r.t. u, v at this point. That is

$$\left.\begin{pmatrix} -3u^2 & 2v \\ -4u & 12v^3 \end{pmatrix}\right|_{(2,-1,2,1)} = \begin{pmatrix} -12 & 2 \\ -8 & 12 \end{pmatrix}$$

which is an invertible matrix.

So locally, near $(2,-1,2,1)$ *one can write*

$$(u,v) = G(x,y) \text{ that is } F(x,y,G_1(x,y),G_2(x,y)) = 0.$$

It is even possible to determine $\partial G_1/\partial x$ (i.e. $\partial u/\partial x$). Indeed, writing $u = G_1(x,y)$ and $v = G_2(x,y)$ and differentiate:

$$\begin{aligned} x^2 - y^2 - u^3 + v^2 + 4 &= 0, \\ 2xy + y^2 - 2u^2 + 3v^4 + 8 &= 0, \end{aligned}$$

with respect to x. This gives

$$\begin{aligned} 2x - 3u^2 \tfrac{\partial u}{\partial x} + 2v \tfrac{\partial v}{\partial x} &= 0, \\ 2y - 4u \tfrac{\partial u}{\partial x} + 12v^3 \tfrac{\partial v}{\partial x} &= 0. \end{aligned}$$

So

$$\begin{aligned} \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial x} \end{pmatrix} &= \begin{pmatrix} 3u^2 & -2v \\ 4u & -12v^3 \end{pmatrix}^{-1} \begin{pmatrix} 2x \\ 2y \end{pmatrix} \\ &= \frac{1}{8uv - 36u^2v^2} \begin{pmatrix} -12v^3 & 2v \\ -4u & 3u^2 \end{pmatrix} \begin{pmatrix} 2x \\ 2y \end{pmatrix} \end{aligned}$$

Hence $\dfrac{\partial u}{\partial x} = \dfrac{(-24xv^3 + 4vy)}{8uv - 36u^2v^2}$.

Appendix B Prerequisites

B.1 Function spaces

1. Let $f\colon [0,1] \to \mathbb{R}$ be a function and $f_n\colon [0,1] \to \mathbb{R}$ be a sequence functions. Define what it means to say that $f_n \to f$ **uniformly**.

   Answer: for all $\epsilon > 0$ there exists n_0 so that for all $n \geq n_0$ and all $x \in [0,1]$ one has $|f_n(x) - f(x)| < \epsilon$.

   Answer 2: $||f_n - f||_\infty \to 0$ as $n \to \infty$ where $||f_n - f||_\infty = \sup_{x\in[0,1]} |f_n(x) - f(x)|$.

2. Let $f\colon [0,1] \to \mathbb{R}$ be a function and $f_n\colon [0,1] \to \mathbb{R}$ be a sequence functions. Define what it means to say that $f_n \to f$ **pointwise**.

   Answer: for all $\epsilon > 0$ and all $x \in [0,1]$ there exists n_0 so that for all $n \geq n_0$ one has $|f_n(x) - f(x)| < \epsilon$.

3. Let $f\colon [0,1] \to \mathbb{R}$ be a function and $f_n\colon [0,1] \to \mathbb{R}$ be a sequence functions. Assume that $f_n \to f$ uniformly and that f_n is continuous. Show that f is continuous.

   Answer: Take $\epsilon > 0$, $x \in [0,1]$. Choose n_0 so that $|f_n - f|_\infty < \epsilon/3$ for $n \geq n_0$ and pick $\delta > 0$ so that $|f_{n_0}(x) - f_{n_0}(y)| < \epsilon/3$ for all y with $|y - x| < \delta$. Then for all y with $|y - x| < \delta$, $|f(x) - f(y)| < |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(y)| + f_{n_0}(y) - f(y)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$.

4. Let $f\colon [0,1] \to \mathbb{R}$ be a function and $f_n\colon [0,1] \to \mathbb{R}$ be a sequence functions. Assume that $f_n \to f$ pointwise and that f_n is continuous. Show that f is not necessarily continuous.

   Answer: Take $f_n(x) = (1 - nx)$ for $x \in [0, 1/n]$ and $f_n(x) = 0$ elsewhere. Then $f_n \to f$ pointwise, where

$f(0) = 1$ and $f(x) = 0$ for $x \in (0, 1]$.

Appendix C Explicit methods for solving ODE's

This Appendix summarises explicit methods for solving ODE's. Since most of the material is already covered in first year material, it will not be covered in the lectures.

C.1 State independent

- This section summarises techniques for **solving** ODE's.
- The first subsections are about finding $x \colon \mathbb{R} \to \mathbb{R}$ so that $x' = f(x, t)$ and $x(0) = x_0$ where $f \colon \mathbb{R}^2 \to \mathbb{R}$.
- So the issue is to find curves with prescribed tangents.
- Let us first review methods for explicitly solving such equations (in part reviewing what you already know).

C.2 State independent $\dot{x} = f(t)$.

In this case, each solution is of the form $x(t) = \int_0^t f(s)\, ds + x(0)$.

Example C.1

Assume the graph $t \mapsto (t, x(t))$ has tangent vector $(1, \sin(t))$ at t. Then $x'(t) = \sin(t)$ and so $x(t) = -\cos(t) + c$. So the solution of the ODE $x'(t) = \sin(t)$ finds a curve which is tangent to the arrows of the vector field.

C.3 Separation of variables

Separation of variables: $\dot{x} = f(t)g(x)$. Then one can find solutions as follows.

$$\int_{x(0)}^{x(T)} \frac{dy}{g(y)} = \int_0^T \frac{1}{g(x(t))}\frac{dx}{dt}dt = \int_0^T f(t)\,dt.$$

Here the first equality follows from the substitution rule (taking $y = x(t)$) and 2nd from $\frac{1}{g(x(t))}\frac{dx}{dt} = f(t)$.

Example C.2

$\frac{dx}{dt} = ax + b, x(t) = x_0$. Then $\frac{dx}{ax+b} = dt, x(0) = x_0$ which gives, when $a \neq 0$,

$$(1/a)[\log(ax+b)]_{x_0}^{x(T)} = T,$$

$$\log((ax(T)+b)/(ax_0+b)) = aT$$

and therefore

$$x(T) = x_0 e^{aT} + \frac{e^{aT}-1}{a}b \text{ for } T \in (-\infty, \infty)$$

Example C.3

$\frac{dx}{dt} = x^2, x(0) = x_0$. Then $\frac{dx}{x^2} = dt, x(0) = x_0$. Hence $[-1/x]_{x_0}^{x(t)} = t$ and so $x(t) = \frac{1}{1/x_0 - t}$. Note that $x(t)$ is well-defined for $t \in (-\infty, 1/x_0)$ but that $x(t) \to \infty$ as $t \uparrow 1/x_0$. The solution goes to infinity in finite time.

Example C.4

$\frac{dx}{dt} = \sqrt{|x|}, x(0) = x_0$. If $x_0 > 0$ and $x(t) > 0$ then we obtain $\frac{dx}{\sqrt{x}} = dt, x(0) = x_0$ and so $2\sqrt{x(t)} - 2\sqrt{x_0} = t$.

Thus $x(t) = (\sqrt{x_0} + t/2)^2$ for $t \in (-2\sqrt{x_0}, \infty)$. When $t = -2\sqrt{x_0}$ then $x(t) = 0$, so need to analyse this directly.

*When $x_0 = 0$ then there are many solutions (*non-uniqueness*). For any $-\infty \le t_0 \le 0 \le t_1 \le \infty$*

$$x(t) = \begin{cases} -(t-t_0)^2/4 & \text{for } t \in (-\infty, t_0) \\ 0 & \text{for } t \in [t_0, t_1] \\ (t-t_1)^2/4 & \text{for } t \in (t_1, \infty) \end{cases}$$

is a solution.

So, without imposing some assumptions, solutions **need not be unique**.

C.4 Linear equations $x' + a(t)x = b(t)$.

To solve this, first consider the **homogeneous case** $x' + a(t)x = 0$. This can be solved by separation of variables: dx/x=-a(t)dt and so $x(t) = x_0 \exp[-\int_0^t a(s)\,ds]$.

To find the solution of the ODE, apply the *variation of variables 'trick'*: substitute $x(t) = c(t)\exp[-\int_0^t a(s)\,ds]$ in the equation and obtain an equation for $c(t)$.

Example C.5

$x' + 2tx = t$. The homogeneous equation $x' + 2tx = 0$ has solution $x(t) = ce^{-t^2}$.

Substituting $x(t) = c(t)e^{-t^2}$ into $x' + 2tx = t$ gives $c'(t)e^{-t^2} + c(t)(-2t)e^{-t^2} + 2tc(t)e^{-2t^2} = t$, i.e. $c'(t) = te^{t^2}$. Hence $c(t) = c_0 + (1/2)e^{t^2}$ and therefore $x(t) = c_0e^{-t^2} + (1/2)$. That the equation is of the form

$$c_0 \cdot \textit{solution of hom.eq} + \textit{ special solution}$$

is due to the fact that the space of solutions $x' + 2tx = 0$ is linear (linear combination of solutions are again solutions).

C.5 Exact equations $M(x, y)dx + N(x, y)dy = 0$ when $\partial M/\partial y = \partial N/\partial x$.

Suppose $f(x, y) \equiv c$ is a solution. Then $df = (\partial f/\partial x)dx + (\partial f/\partial y)dy = 0$ and this corresponds to the ODE if $\partial f/\partial x = M$ and $\partial f/\partial y = N$. But if f is twice differentiable we have

$$\partial M/\partial y = \partial^2 f/\partial x \partial y = \partial^2 f/\partial y \partial x = \partial N/\partial x.$$

It turns out that this necessary condition for 'exactness' is also sufficient if the domain we consider has no holes (is simply connected).

Example C.6

$(y - x^3)dx + (x + y^2)dy = 0$. *The exactness condition is satisfied (check!). How to find* f *with* $\partial f/\partial x = y - x^3$ *and* $\partial f/\partial y = x + y^2$*? The first equation gives* $f(x, y) = yx - (1/4)x^4 + c(y)$. *The second equation then gives* $x + c'(y) = \partial f/\partial y = x + y^2$. *Hence* $c(y) = y^3/3 + c_0$ *and* $f(x) = yx - (1/4)x^4 + y^3/3 + c_0$ *is a solution.*

Sometimes you can rewrite the equation to make it exact.

Example C.7

$ydx + (x^2y - x)dy = 0$. *This equation is* **not** *exact (indeed,* $\dfrac{\partial y}{\partial y} \neq \dfrac{\partial(x^2y - x)}{\partial x}$*). If we rewrite the equation as* $y/x^2dx + (y - 1/x)dy = 0$ *then it becomes exact.*

Clearly this was a lucky guess. Sometimes one can guess that by multiplying by a function of (for example) x the ODE becomes exact.

Example C.8

The equation $(xy - 1)dx + (x^2 - xy)dy = 0$ *is not exact.*

Let us consider the equation $\mu(x)(xy-1)dx + \mu(x)(x^2 - xy)dy = 0$. *The exactness condition is* $\mu x = \mu'(x^2 - xy) + \mu(2x - y)$. *Rewriting this gives* $\mu'(x)x(x-y) + \mu(x)(x - y) = 0$, *and so* $x\mu' + \mu = 0$ *implies the exactness condition. So we can take* $\mu(x) = 1/x$. *So instead of the original ODE we solve* $(y - 1/x)dx + (x - y)dy = 0$ *as in the previous example.*

C.6 Substitutions

- Sometimes one can simplify the ODE by a substitution.
- One instance of this method, is when the ODE is of the form $M(x,y)dx + N(x,y)dy = 0$ where M, N are homogeneous polynomials of the same degree.

  In this case we can simplify by substituting $z = y/x$.

Example C.9

$(x^2 - 2y^2)dx + xydy = 0$. *Rewrite this as* $\frac{dy}{dx} = \frac{-x^2+2y^2}{xy}$. *Substituting* $z = y/x$, *i.e.* $y(x) = z(x)x$ *gives*

$$x\frac{dz}{dx} + z = \frac{dy}{dx} = \frac{-1+2z^2}{z}.$$

Hence

$$\frac{dz}{dx} = \frac{-1}{z} + z.$$

This can be solved by separation of variables.

C.7 Higher order linear ODE's with constant coefficients

Note that each y_1 and y_2 are solutions of

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = 0 \qquad (38)$$

then linear combinations of y_1 and y_2 are also solutions.

Substituting $y(x) = e^{rx}$ in this equation gives:

$$e^{rn}\left(r^n + a_{n-1}r^{n-1} + \cdots + a_0\right) = 0.$$

Of course the polynomial equation $r^n + a_{n-1}r^{n-1} + \cdots + a_0 = 0$ has n solutions $r_1, \ldots, r_n \in \mathbb{C}$.

Case 1: If these r_i's are all different (i.e. occur with single multiplicity), then we obtain as a solution:

$$y(x) = c_1 e^{r_1 x} + \cdots + c_n e^{r_n x}.$$

Case 2: What if, say, r_1 is complex? Then $\bar{r}_1$ is also a root, so we may (by renumbering) assume $r_2 = \bar{r}_1$ and write $r_1 = \alpha + \beta i$ and $r_2 = \alpha - \beta i$ with $\alpha, \beta \in \mathbb{R}$. So

$$e^{r_1 x} = e^{\alpha x}(\cos(\beta x) + \sin(\beta x)i), e^{r_2 x} = e^{\alpha x}(\cos(\beta x) - \sin(\beta x)i),$$

and $c_1 e^{r_1 x} + c_2 e^{r_2 x} = (c_1 + c_2)e^{\alpha x}\cos(\beta x) + (c_1 - c_2)ie^{\alpha x}\sin(\beta)$. Taking $c_1 = c_2 = A/2 \in \mathbb{R} \implies c_1 e^{r_1 x} + c_2 e^{r_2 x} = Ae^{\alpha x}\cos(\beta x)$ On the other hand, taking $c_1 = -(B/2)i = -c_2 \implies c_1 e^{r_1 x} + c_2 e^{r_2 x} = Be^{\alpha x}\sin(\beta x)$ (nothing prevents us choosing c_i non-real!!).

So if $r_1 = \bar{r}_2$ is non-real, we obtain as a general solution

$$y(x) = Ae^{\alpha x}\cos(\beta x) + Be^{\alpha x}\sin(\beta x) + c_3 e^{r_3 x} + \cdots + c_n e^{r_n x}.$$

Case 3: Repeated roots: If $r_1 = r_2 = \cdots = r_k$ then one can check that $c_1 e^{r_1 x} + c_2 x e^{r_2 x} + \cdots + c_k x^k e^{r_1 x}$ is a solution.

Case 4: Repeated complex roots: If $r_1 = r_2 = \cdots = r_k = \alpha + \beta i$ are non-real, then we have corresponding roots $r_{k+1} = r_{k+2} = \cdots = r_{2k} = \alpha - \beta i$ and we obtain as solution

$$c_1 e^{\alpha x}\cos(\beta x) + \cdots + c_k x^k e^{\alpha x}\cos(\beta x) +$$

$$+ c_{k+1} e^{\alpha x}\sin(\beta x) + \cdots + c_{2k} x^k e^{\alpha x}\sin(\beta x).$$

Example C.10

Vibrations and oscillations of a spring

One can model an object attached to a spring by $Mx'' = F_s + F_d$ *where* F_d *is a damping force and* F_s *a spring force. Usually one assumes* $F_d = -cx'$ *and* $F_s = -kx$*. So*

$$Mx'' + cx' + kx = 0 \text{ or } x'' + 2bx' + a^2x = 0$$

where $a = \sqrt{k/M} > 0$ *and* $b = c/(2M) > 0$*.*

Using the previous approach we solve $r^2 + 2br + a^2$*, i.e.* $r_1, r_2 = \frac{-2b \pm \sqrt{4b^2 - 4a^2}}{2} = -b \pm \sqrt{b^2 - a^2}$*.*

Case 1: *If* $b^2 - a^2 > 0$ *then both roots are real and negative. So* $x(t) = x_0(e^{r_1t} + Be^{r_2t})$ *is a solution and as* $t \to \infty$ *we get* $x(t) \to 0$*.*

Case 2: If $b^2 - a^2 = 0$ then we obtain $r_1 = r_2 = -a$ and $x(t) = Ae^{-at} + Bte^{-at}$. So $x(t)$ still goes to zero as $\to \infty$, but when B is large, $x(t)$ can still grow for t not too large.

Case 3: If $b^2 - a^2 < 0$. Then $x(t) = e^{-bt}(A\cos(\alpha t) + B\sin(\alpha t))$ is a solution. Solutions go to zero as $t \to \infty$ but oscillate.

Example C.11

Vibrations and oscillations of a spring with forcing

Suppose one has external forcing

$$Mx'' + cx' + kx = F_0 \cos(\omega t).$$

If $b^2 - a^2 < 0$ *(using the notation of the previous example) then*

$$e^{-bt}(A\cos(\alpha t) + B\sin(\alpha t))$$

is still the solution of the homogeneous part and one can check

$$\frac{F_0}{\sqrt{(k-\omega^2 M)^2 + \omega^2 c^2}}(\omega c \sin(\omega t) + (k - \omega^2 M)\cos(\omega t) =$$

$$\frac{F_0}{\sqrt{(k-\omega^2M)^2+\omega^2c^2}}\cos(\omega t-\phi)$$

is a particular solution where $\omega = \arctan(\omega c/(k-\omega^2 M))$.

$$\frac{F_0}{\sqrt{(k-\omega^2M)^2+\omega^2c^2}}\cos(\omega t-\phi)$$

is a particular solution where $\omega = \arctan(\omega c/(k-\omega^2 M))$.

Here c *is the damping,* M *is the mass and* k *is the spring constant.*

- *If damping* $c \approx 0$ *and* $\omega \approx k/M$ *then the denominator is large, and the oscillation has large amplitude.*
- $(k-\omega^2M)^2+\omega^2c^2$ *is minimal for* $\omega = \sqrt{\frac{k}{M}-\frac{c^2}{2M^2}}$ *and so this is the natural frequency (or eigen-frequency).*
- *This is important for bridge designs (etc), see*
  - *http://www.ketchum.org/bridgecollapse.html*
  - *http://www.youtube.com/watch?v=3mclp9QmCGs*
  - *http://www.youtube.com/watch?v=gQK21572oSU*

C.8 Solving ODE's with maple

Example C.12

```
> ode1 := diff(x(t), t) = x(t)^2;
                                d                2
                               --- x(t) = x(t)
                                dt
> dsolve(ode1);
```

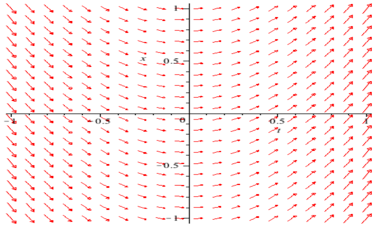


Figure 6: The vector field $(1, \sin(t))$ drawn with the Maple command: with(plots):fieldplot([1, sin(t)], t = -1 .. 1, x = -1 .. 1, grid = [20, 20], color = red, arrows = SLIM);

```
                                  1
                        x(t) = ---------
                               -t + _C1
> dsolve({ode1, x(0) = 1});
                                    1
                         x(t) = - -----
                                  t - 1
```

Example C.13

Example: $y'' + 1 = 0$.

```
> ode5 := diff(y(x), x, x)+1 = 0;
                    / d  / d      \\
                    |--- |--- y(x)|| + 1 = 0
                    \ dx \ dx     //
> dsolve(ode5);
                           1  2
                  y(x) = - - x  + _C1 x + _C2
                           2
```

C.9 Solvable ODE's are rare

It is not that often that one can solve an ODE explicitly. What then?

- Use approximation methods.
- Use topological and qualitative methods.
- Use numerical methods.

This module will explore all of these methods.

In fact, we need to investigate whether we can even speak about solutions. Do solutions exist? Are they unique? Did we find all solutions in the previous subsections?

C.10 Chaotic ODE's

Very simple differential equations can have complicated dynamics (and clearly cannot be solved analytically). For example the famous Lorenz differential equation

$$\begin{aligned} \dot{x} &= \sigma(y-x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz \end{aligned} \tag{39}$$

with $\sigma = 10, r = 28, b = 8/3$.

has solutions which are chaotic and have sensitive dependence (the *butterfly effect*).

`http://www.youtube.com/watch?v=ByH8_nKD-ZM`

Appendix D A proof of the Jordan normal form theorem

In this section we will give a proof of the Jordan normal form theorem.

Lemma D.1

Let $L_1, L_2 \colon V \to V$ where V is a finite dimensional vector space. Assume $L_1 L_2 = 0$ and $\ker(L_1) \cap \ker(L_2) = \{0\}$. Then $V = \ker(L_1) \oplus \ker(L_2)$.

Proof : Let n be the dimension of V and let $\Im(L_2)$ stands for the range of L_2. Note that $\dim\ker(L_2) + \dim\Im(L_2) = n$, Since $L_1 L_2 = 0$ it follows that $\ker(L_1) \supset \Im(L_2)$ and therefore $\dim\ker(L_1) \geq \dim\Im(L_2) = n - \dim\ker(L_2)$. As $\ker(L_1) \cap \ker(L_2) = \{0\}$, equality holds in $\dim\ker(L_1) + \dim\ker(L_2) \geq n$ and the lemma follows. ■

Proposition D.2

Let $L \colon V \to V$ where V is a finite dimensional vector space. Let $\lambda_1, \dots, \lambda_s$ be its eigenvalues with (algebraic) multiplicity m_i. Then one can write $V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$ where $V_i = \ker((L - \lambda_i I)^{m_i})$ and so $L(V_i) \subset V_i$.

Proof : Consider the polynomial $p(t) = \det(tI - L)$. This is a polynomial of degree n, where n is the dimension of the vector space and with leading term t^n. By the Cayley-Hamilton theorem one has $p(L) = 0$ and of course $p(L)$ is also of the form $L^n + c_1 L^{n-1} + \cdots + c_n = 0$. This can be factorised as

$$(L - \lambda_1 I)^{m_1} (L - \lambda_2 I)^{m_2} \cdots (L - \lambda_s I)^{m_s} = 0,$$

where all λ_i's are distinct - here we use that the factors $(L - \lambda_i I)$ commute.

We claim that $\ker((L - \lambda_i I)^{m_i}) \cap \ker(L - \lambda_j I)^{m_j} = 0$. Indeed, if not then we can take a vector $v \neq 0$ which is in the intersection. We may assume $m_i \geq m_j$. Choose $1 \leq m_j' \leq m_j$ minimal so that $(L - \lambda_j I)^{m_j'} v = 0$ and $(L - \lambda_j I)^{m_j'-1} v \neq 0$. Since $v \in \ker((L - \lambda_i I)^{m_i})$ we have that $w := (L - \lambda_j I)^{m_j'-1}(L - \lambda_i I)^{m_i} v$ is equal to 0, but on the other hand w is equal to $(L-\lambda_j I)^{m_j'-1}((L-\lambda_j I)+(\lambda_j - \lambda_i))^{m_i} v$ which, by expanding the latter expression (and using that $v \in \ker((L - \lambda_i I)^{m_i})$) is equal to $(L - \lambda_j I)^{m_j'-1}(\lambda_j - \lambda_i)^{m_i} v \neq 0$. This contradiction proves the claim.

This means that we can apply the previous lemma inductively to the factors $(L - \lambda_i I)^{m_i}$, and thus obtain the proposition. ■

It follows that if we choose T so that it sends the decomposition $\mathbb{R}^{n_1} \oplus \dots \mathbb{R}^{n_k}$, where $n_i = \dim V_i$, to $V_1 \oplus \cdots \oplus V_k$ then $T^{-1}LT$ is of the form $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_p \end{pmatrix}$ where A_i are square matrices corresponding to V_i (and the remaining entries are zero). The next theorem gives a way to find a more precise description for a linear transformation T so that $T^{-1}LT$ takes the Jordan form. Indeed, we apply the next theorem to each matrix A_i separately. In other words, for each choice of i, we take $W = V_i$, $A = (L - \lambda_i I)|V_i$ and $m = m_i$ in the theorem below.

Theorem D.3

Let $A\colon W \to W$ be a linear transformation of a finite dimensional vector space so that $A^m = 0$ for some $m \geq 1$. Then there exists a basis W of the form

$$u_1, Au_1, \dots, A^{a_1-1}u_1, \dots, u_s, \dots, A^{a_s-1}(u_s)$$

where $a_i \geq 1$ and $A^{a_i}(u_i) = 0$ for $1 \leq i \leq s$.

Remark: Note that $A^{a_j-1}(u_j) = (T - \lambda_i)^{a_j-1}(u_j)$ is in the kernel of $A = T - \lambda_i I$, so is an eigenvector of A corresponding to eigenvalues 0 (i.e. an eigenvector of T corresponding to eigenvalue λ_i. The vector $w_j^1 = A^{a_j-2}(u_j) = (T-\lambda_i I)^{a_j-2}(u_j)$ corresponds to a vector so that $Aw_j^1 = w_j$, so $Tw_j^1 = \lambda w_j^1 + w_j$, and so on. So as in Chapter 4, if we take the matrix T with columns

$$A^{a_1-1}u_1, \dots, u_1, A^{a_2-1}u_2, \dots, u_2, A^{a_s-1}u_s, \dots, u_s$$

then $T^{-1}LT$ will have the required Jordan form with λ on the diagonal, and 1's in the off-diagonal except in columns $a_1, a_1 + a_2, \dots, a_1 + a_2 + \cdots + a_s$.

Proof: The proof given below goes by induction with respect to the dimension of W. When $\dim W = 0$ the statement is obvious. Assume that the statement holds for dimensions $< \dim(W)$. Note that $A(W) \neq W$ since otherwise $AW = W$ and therefore $A^m(W) = A^{m-1}(W) = \cdots = W$ which is a contradiction. So $\dim A(W) < \dim W$ and by induction there exists $v_1, \dots, v_l \in A(W)$ so that

$$v_1, Av_1, \dots, A^{b_1-1}(v_1), \dots, v_l, Av_l, \dots, A^{b_l-1}v_l \qquad (40)$$

is a basis for $A(W)$ and $A^{b_i}(v_i) = 0$ for $1 \leq i \leq l$. Since $v_i \in A(W)$ one can choose u_i so that $Au_i = v_i$. The vectors $A^{b_1-1}v_1, \dots, A^{b_l-1}v_l$ are linearly independent and are contained in $\ker(A)$ and so we can find vectors $u_{l+1}, \dots, u_m$ so that

$$A^{b_1-1}v_1, \dots, A^{b_l-1}v_l, u_{l+1}, \dots, u_m \qquad (41)$$

forms a basis of $\ker(A)$. But then

$$u_1, Au_1, \dots, A^{b_1}(u_1), \dots, u_l, \dots, A^{b_l}u_l, u_{l+1}, \dots, u_m \qquad (42)$$

is the required basis of W. Indeed, consider a linear combination of vectors from (42) and apply A. Then, because $v_i = Au_i$, we obtain a linear combination of the vectors from (40) and so the corresponding coefficients are zero. The remaining vectors are in the kernel of A and are linearly independent because they correspond to (41). This proves the linear independence of (42). That (42) spans W holds, because the number of vectors appearing in (42) is equal to $\dim\ker(A) + \dim AW$. Indeed, $A^{b_1}(u_1), \ldots, A^{b_l}u_l, u_{l+1}, \ldots, u_m$ are all in $\ker(A)$ (they are the same vectors as the vectors appearing in (41)). The remaining number of vectors is $b_1 + \cdots + b_l$ which is the same as the dimension of AW, as (40) forms a basis of this space. It follows that the total number of vectors in (42) is the same as $\dim(\ker(A)) + \dim(AW)$ and so together with their linear independence this implies that (42) forms a basis of W. ■

Appendix E Calculus of Variations

Many problems result in differential equations. In this chapter we will consider the situation where these arise from a minimisation (variational) problem. Specifically, the problems we will consider are of the type

- Minimize
$$I[y] = \int_0^1 f(x, y(x), y'(x))\, dx \qquad (43)$$
where f is some function and y is an unknown function.

- Minimize (43) conditional to some restriction of the type $J[y] = \int_0^1 f(x, y(x), y'(x))\, dx = 1$.

E.1 Examples (the problems we will solve in this chapter):

Example E.1

Let $A = (0, 0)$ *and* $B = (1, 0)$ *with* $l, b > 0$ *and consider a path of the form* $[0, 1] \ni\mapsto c(t) = (c_1(t), c_2(t))$, *connecting* A *and* B. *What is the shortest path?*

Task: *Choose* $[0, 1] \ni t \mapsto c(t) = (c_1(t), c_2(t))$ *with* $c(0) = (0, 0)$ *and* $c(1) = (1, 0)$ *which minimises*

$$L[c] = \int_0^1 \sqrt{c_1'(t)^2 + c_2'(t)^2}\, dt.$$

Of course this is a line segment, but how to make this precise?

If we are not in a plane, but in a surface or a higher dimensional set, these shortest curves are called **geodesics**, and these are studied extensively in mathematics.

Example E.2

Let $A = (0,0)$ and $B = (l,-b)$ with $l, b > 0$ and consider a path of the form $(x, y(x))$, $x \in [0,l]$, connecting A and B. Take a ball starting at A and rolling along this path under the influence of gravity to B. Let T be the time this ball will take. Which function $x \mapsto y(x)$ which will minimise T?

The sum of kinetic and potential energy is constant

$$(1/2)mv^2 + mgh = const.$$

Since the ball rolls along $(x, y(x))$ we have $v(x) = \sqrt{-2gy(x)}$. Let $s(t)$ be the length travelled at time t. Then $v = ds/dt$. Hence $dt = ds/v$ or

$$T[y] := \int_0^l \frac{\sqrt{1 + y'(x)^2}}{\sqrt{-2gy(x)}}\, dx.$$

Task*: minimise $T[y]$ within the space of functions $x \mapsto y(x)$ for which y and y' continuous and $y(0) = 0$ and $y(l) = -b$. This is called the* **Brachisotochrome***, going back to Bernouilli in 1696.*

Example E.3

Take a closed curve in the plane without self-intersections and of length one. What is the curve c which maximises the area D it encloses? Again, let $[0,1] \ni \mapsto c(t) = (c_1(t), c_2(t))$ with $c(0) = c(1)$ and so that $s, t \in [0,1)$ and $s \neq t$ implies $c(s) \neq c(t)$.

The length of the curve is again $L[c] = \int_0^1 \sqrt{c_1'(t)^2 + c_2'(t)^2}\, dt$. To compute the area of D we use the Green theorem:

$$\int_c P dx + Q dy = \int\int_D (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}) dx dy$$

Take $P \equiv 0$ and $Q = x$. Then

$$\int\int_D (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}) dxdy = \int\int_D 1\, dxdy = \text{ area of } D.$$

So

$$A[c] = \int\int_D 1\, dxdy = \int_c xdy = \int_0^1 c_1(t)c_2'(t)\, dt.$$

This is an isoperimetric problem: *find the supremum of* $A[c]$ *given* $L[c] = 1$.

E.2 Extrema in the finite dimensional case

We say that $F\colon \mathbb{R}^n \to \mathbb{R}$ take a local minimum at $\tilde{x}$ if there exists $\delta > 0$ so that

$$F(x) \geq F(\tilde{x}) \text{ for all } x \text{ with } |x - \tilde{x}| < \delta.$$

Theorem E.4

Assume that F *is differentiable at* a *and also has a minimum at* $\tilde{x}$ *then* $DF(\tilde{x}) = 0$.

Proof : Let us first assume that $n = 1$. Then that f has a minimum means that $F(\tilde{x} + h) - F(\tilde{x}) \geq 0$ for all h near zero. Hence

$$\frac{F(\tilde{x} + h) - F(\tilde{x})}{h} \geq 0 \text{ for } h > 0 \text{ near zero and}$$

$$\frac{F(\tilde{x} + h) - F(\tilde{x})}{h} \leq 0 \text{ for } h < 0 \text{ near zero.}$$

Therefore

$$F'(\tilde{x}) = \lim_{h \to 0} \frac{F(\tilde{x} + h) - F(\tilde{x})}{h} = 0$$

Let us consider the case that $n > 1$ and reduce to the case that $n = 1$. So take a vector v at $\tilde{x}$, define $l(t) = \tilde{x} + tv$ and $g(t) := F \circ l(t)$. So we can use the first part of the proof and thus we get $g'(0) = 0$. Applying the chain rule $0 = g'(0) = Dg(0) = DF(l(0))Dl(0) = DF(\tilde{x})v$ and so

$$\frac{\partial F}{\partial x_1}(\tilde{x})v_1 + \cdots + \frac{\partial F}{\partial x_n}(\tilde{x})v_n = 0.$$

Hence $DF(\tilde{x})v = 0$ where $DF(\tilde{x})$ is the Jacobian matrix at $\tilde{x}$. Since this holds for all v, we get $DF(\tilde{x}) = 0$. ■

Remember we also wrote sometimes $DF_{\tilde{x}}$ for the matrix $DF(\tilde{x})$ and that $DF(\tilde{x})v$ is the *directional derivative* of f at $\tilde{x}$ in the direction v.

E.3 The Euler-Lagrange equation

The infinite dimensional case: the Euler-Lagrange equation

- In the infinite dimensional case, we will take $F\colon H \to \mathbb{R}$ where H is some function space. The purpose of this chapter is to generalise the previous result to this setting, and show that the solutions of this problem gives rise to differential equations.

- Mostly the function space is the space $C^1[a, b]$ of C^1 functions $y\colon [a, b] \to \mathbb{R}^n$. This space is an infinite dimensional vector space (in fact, a Banach space) with norm $|y|_{C^1} = \sup_{x\in[a,b]}(|y(x)|, |Dy(x)|)$.

- Choose some function $f\colon [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Take $(x, y, y') \in [a, b] \times \mathbb{R}^n \times \mathbb{R}^n$ **denote by f_y, $f_{y'}$ the corresponding partial derivatives**. So $f_y(x, y, y')$ and $f_{y'}(x, y, y')$ vectors. Attention: here y and y' are just the names of vectors in $\mathbb{R}^n$ (and not - yet - functions or derivatives of functions).

- Here f_y is the part of the $1 \times (1+n+n)$ vector Df which concerns the y derivatives.

Assume $f\colon [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with $f_y, f_{y'}$ continuous and define $I\colon C^1[a,b] \to \mathbb{R}$ by,

$$I[y] = \int_a^b f(x, y(x), y'(x))\, dx.$$

Given $\tilde{y}\colon [a,b] \to \mathbb{R}^n$, let's denote $f[\tilde{y}](x) = f(x, \tilde{y}(x), \tilde{y}'(x))$,

$f_y[\tilde{y}](x) = f_y(x, \tilde{y}(x), \tilde{y}'(x))$ and $f_{y'}[\tilde{y}](x) = f_{y'}(x, \tilde{y}(x), \tilde{y}'(x))$

where $f_y, f_{y'}$ are the corresponding partial derivatives of f. Fix $y_a, y_b \in \mathbb{R}^n$ and define

$$\mathcal{A} = \{y;\ y\colon [a,b] \to \mathbb{R}^n \text{ is } C^1 \text{ and } y(a) = y_a, y(b) = y_b\}.$$

Theorem E.5

If $\mathcal{A} \ni y \mapsto I[y]$ has a minimum at $\tilde{y}$ then

1. *for every $v \in C^1[a,b]$ with $v(a) = v(b) = 0$ we get $\int_a^b (f_y[\tilde{y}] \cdot v + f_{y'}[\tilde{y}]v')\, dx = 0$.*
2. *$f_{y'}[\tilde{y}]$ exists, is continuous on $[a,b]$ and*

$$\frac{d}{dx} f_{y'}[\tilde{y}] = f_y[\tilde{y}].$$

Proof : Remember that

$$\mathcal{A} = \{y;\ y\colon [a,b] \to \mathbb{R}^n \text{ is } C^1 \text{ and } y(a) = y_a, y(b) = y_b\}.$$

Hence $v \in C^1[a,b]$ with $v(a) = v(b) = 0$, then $y + hv \in \mathcal{A}$ for each h. So the space $\mathcal{A}$ is affine.

Assume that $I\colon\ C^1[a,b] \to \mathbb{R}$ has a minimum at $\tilde{y}$, which means that

$$I[\tilde{y}+hv] \geq I[\tilde{y}]\ \forall v \in C^1[a,b], v(a)=v(b)=0\ \forall h \in \mathbb{R}$$

$$I[\tilde{y}+hv]-I[\tilde{y}] =$$

$$= \int_a^b f(x,(\tilde{y}+hv)(x),(\tilde{y}+hv)'(x))-f(x,\tilde{y}(x),\tilde{y}'(x))dx.$$

By Taylor's Theorem,

$$f(x,(\tilde{y}+hv)(x),(\tilde{y}+hv)'(x))-f(x,\tilde{y}(x),\tilde{y}'(x)) =$$

$$f_y[\tilde{y}]hv + f_{y'}[\tilde{y}]hv' + o(h).$$

So

$$I[\tilde{y}+hv]-I[\tilde{y}] = h\cdot\left[\int_a^b \left[f_y[\tilde{y}]v + f_{y'}[\tilde{y}]v'\right]\ dx\right] + o(h).$$

Hence a necessary condition for $\tilde{y}$ to be a minimum of I is

$$\int_a^b \left[f_y[\tilde{y}]v + f_{y'}[\tilde{y}]v'\right]\ dx = 0$$

for each $v \in C^1[a,b]$ with $v(a)=v(b)=0$.

Partial integration gives

$$\int_a^b f_{y'}[\tilde{y}]v'\,dx = \left(f_{y'}[\tilde{y}]v\right)\Big|_a^b - \int_a^b \frac{d}{dx}f_{y'}[\tilde{y}]v\,dx.$$

Remember $v(a)=v(b)=0$, so $\left(f_{y'}[\tilde{y}]v\right)\Big|_a^b = 0$. Therefore a necessary condition for $\tilde{y}$ to be a minimum of I is:

$$v \in C^1[a,b] \quad \text{with } v(a)=v(b)=0 \implies$$
$$\int_a^b \left[f_y[\tilde{y}] - \frac{d}{dx}f_{y'}[\tilde{y}]\right] v\,dx = 0.$$

This prove first assertion of Theorem and also the 2nd assertion because of the following lemma: ■

Lemma E.6

If $G\colon [a,b] \to \mathbb{R}$ is continuous and $\int_a^b Gv\,dx = 0$ for each $v \in C^1[a,b]$ with $v(a) = v(b) = 0$, then $G \equiv 0$.

Proof : If $G(x_0) > 0$ then $\exists \delta > 0$ so that $G(x) > 0$, $\forall x$ with $|x - x_0| < \delta$. Choose $v \in C^1[a,b]$ with $v(a) = v(b) = 0$, so that $v > 0$ on $x \in (x_0 - \delta, x_0 + \delta) \cap (a,b)$ and zero outside. Then $\int_a^b G(x)v(x)\,dx > 0$. ■

Quite often x does not appear in f. Then it is usually more convenient to rewrite the Euler-Lagrange equation:

Lemma E.7

If x does not appear explicitly in f, then $\dfrac{d}{dx} f_{y'}[\tilde{y}] = f_y[\tilde{y}]$ implies $f_{y'}[\tilde{y}]\tilde{y}' - f[\tilde{y}] = C$.

Proof :

$$\begin{aligned}
\frac{d}{dx}(f_{y'}[\tilde{y}]\tilde{y}' - f[\tilde{y}]) &= (\frac{d}{dx}f_{y'}[\tilde{y}])\tilde{y}' + f_{y'}[\tilde{y}]\tilde{y}'' \\
&\quad -(f_x[\tilde{y}] + f_y[\tilde{y}]\tilde{y}' + f_{y'}[\tilde{y}]\tilde{y}'') \\
&= y'\left\{\frac{d}{dx}f_{y'}[\tilde{y}] - f_y[\tilde{y}]\right\} - f_x[\tilde{y}].
\end{aligned}$$

Since $f_x = 0$, and by the E-L equation, the term $\{\cdot\} = 0$ this gives the required result. ■

Example E.8

Shortest curve connecting two points $(0,0)$ and $(1,0)$. Let us consider curves of the form $x \mapsto (x, y(x))$ and minimise the length: $I[c] = \int_a^b \sqrt{1 + y'(x)^2}\,dx$. The Euler-

Lagrange equation is $\frac{d}{dx} f_{y'}[\tilde{y}] = f_y[\tilde{y}] = 0$. *Note* $f_{y'} = \frac{y'}{\sqrt{1+(y')^2}}$, *so* $\frac{d}{dx}\frac{y'}{\sqrt{1+(y')^2}} = 0$. *Hence the EL equation gives* $\frac{\tilde{y}'}{\sqrt{1+(\tilde{y}')^2}} = C$. *This means that* $\tilde{y}' = C_1$. *Hence* $\tilde{y}(x) = C_1 x + C_2$. *With the boundary conditions this gives* $\tilde{y}(x) = 0$.

E.4 The brachistochrone problem

Example E.9

(See Example E.2) The curve $x \to (x, y(x))$ *connecting* $(0,0)$ *to* $(l, -b)$ *with the shortest travel time* **brachistochrone**. *Then* $f(x,y,y') = \frac{\sqrt{1+(y')^2}}{\sqrt{-2gy}}$. *Since* $y < 0$, *we orient the vertical axis downwards, that is we write* $z = -y$ *and* $z' = -y'$, *i.e. take* $f(x,z,z') = \frac{\sqrt{1+(z')^2}}{\sqrt{2gz}}$. *Note that*

$$f_{z'} = (1/2)\frac{1}{\sqrt{1+(z')^2}\sqrt{2gz}} 2z'.$$

The EL equation from the previous lemma gives $f'_{z'}[\tilde{z}]\tilde{z}' - f[\tilde{z}] = const$, *i.e. (writing* z *instead of* $\tilde{z}$*):*

$$\frac{(z')^2}{\sqrt{1+(z')^2}\sqrt{z}} - \frac{\sqrt{1+(z')^2}}{\sqrt{z}} = const.$$

Rewriting this gives

$$z[1+(z')^2] = const.$$

Rewriting this again gives the differential equation

$$\frac{dz}{dx} = \sqrt{\frac{C-z}{z}} \text{ or } \frac{dx}{dz} = \sqrt{\frac{z}{C-z}}$$

with $C > 0$. As usual we solve this by writing $dx = \sqrt{\frac{z}{C-z}} dz$ and so

$$x = \int \sqrt{\frac{z}{C-z}}\, dz.$$

Substituting $z = C\sin^2(s)$, where $s \in [0, \pi]$, gives

$$x = \int \sqrt{\frac{\sin^2(s)}{1-\sin^2(s)}} (2C)\sin(s)\cos(s)\, ds =$$

$$2C \int \sin^2(s)\, dt = C \int (1-\cos(2s))\, dt = (C/2)(2s - \sin(2s)) + A$$

Since the curve starts at $(0, 0)$ we have $A = 0$.

So we get

$$\begin{aligned} x(s) &= \frac{C}{2}(2s - \sin(2s)), \\ z(s) &= C\sin^2(s) = \frac{C}{2}(1-\cos(2s)). \end{aligned} \tag{44}$$

Here we choose C so that $z = b$ when $x = L$. This is called a cycloid, an evolute of the circle. This is the path of a fixed point on a bicycle wheel, as the bicycle is moving forward.

Substituting $2s$ to ϕ and taking $a = C/2$ we get

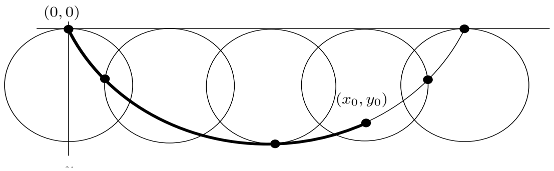
$$\begin{aligned} x(\phi) &= a(\phi - \sin(\phi)), \\ z(\phi) &= a(1-\cos(\phi)). \end{aligned} \tag{45}$$

What is a? Given $L = x_0$ and $-b = y_0$ we need to choose a, ϕ so that $x(\phi) = L$ and $z(\phi) = b$. This amounts two equations and two unknowns.

Two special cases:

- The right endpoint is $(L, 0)$, the top of the curve: then take $\phi = 2\pi$ and we have $x(2\pi) = 2\pi a$ and $z(2\pi) = a$.

- The right endpoint is $(L, 2a)$ and this is the bottom of the curve: then $\phi = \pi$ and $x(\pi) = a\pi$ and $y(\pi) = 2a$.



A remarkable property of the brachistochrone: Take an initial point $(\hat{x}, \hat{y})$ on this curve, and release it from rest. Then the time to hit the lower point of the curve is independent of the choice of the initial point!!!

Theorem E.10

For any initial point $(\hat{x}, \hat{y})$ (i.e. for any initial $\hat{\phi}$)

$$T = \int_{\hat{x}}^{L} \sqrt{\frac{1 + (z')^2}{2g(z - z_0)}}\, dx = \sqrt{\frac{a}{g}} \int_{\phi=\hat{\phi}}^{\pi} \sqrt{\frac{1 - \cos(\phi)}{\cos(\hat{\phi}) - \cos(\phi)}}\, d\phi$$

is equal to $= \pi\sqrt{a/g}$*. Wow!*

Proof : *Not examinable* Let us first show the integrals are equal:

$$x(\phi) = a(\phi - \sin(\phi)), z(\phi) = a(1 - \cos(\phi)) \implies$$

$$z' = \frac{dz}{dx} = \frac{\dfrac{dz}{d\phi}}{\dfrac{dx}{d\phi}} = \frac{a\sin(\phi)}{a(1 - \cos(\phi))} \implies$$

$$\sqrt{1 + (z')^2} = \sqrt{\frac{(1 - \cos(\phi))^2 + \sin^2(\phi)}{(1 - \cos(\phi))^2}} = \sqrt{\frac{2(1 - \cos(\phi))}{(1 - \cos(\phi))^2}}.\blacksquare$$

$$x(\phi) = a(\phi - \sin(\phi)), z(\phi) = a(1 - \cos(\phi)) \implies$$

$$z' = \frac{dz}{dx} = \frac{\frac{dz}{dt}}{\frac{dx}{dt}} = \frac{a\sin(\phi)}{a(1-\cos(\phi))} \implies$$

$$\sqrt{1+(z')^2} = \sqrt{\frac{(1-\cos(\phi))^2 + \sin^2(\phi)}{(1-\cos(\phi))^2}} = \sqrt{\frac{2(1-\cos(\phi))}{(1-\cos(\phi))^2}}.$$

Since $dx = a(1 - \cos(\phi))\, d\phi$ this gives

$$\sqrt{\frac{1+(z')^2}{2g(z-z_0)}}dx = \frac{\sqrt{a}}{\sqrt{g}}\sqrt{\frac{1-\cos(\phi)}{\cos(\hat{\phi}) - \cos(\phi)}}\, d\phi.$$

Showing the two integrals the same.

Claim: the following integral does not depend on $\hat{\phi}$:

$$\int_{\phi=\hat{\phi}}^{\pi} \sqrt{\frac{1-\cos(\phi)}{\cos(\hat{\phi}) - \cos(\phi)}}\, d\phi$$

Substitute $\sin(\phi/2) = \sqrt{1-\cos\phi}/\sqrt{2}$ and $\cos\phi = 2\cos^2(\phi/2) - 1$ gives:

$$\sqrt{\frac{1-\cos(\phi)}{\cos(\hat{\phi}) - \cos(\phi)}} = \sqrt{2}\frac{\sin(\phi/2)}{\sqrt{2[\cos^2(\hat{\phi}/2) - \cos^2(\phi/2)]}}$$

Substitute $u = \cos(\phi/2)/\cos(\hat{\phi}/2)$, then as ϕ varies between $[\hat{\phi}, \pi]$ then u varies from 1 to 0.

$$\int_{\hat{\phi}}^{\pi} \frac{\sin(\phi/2)}{\sqrt{\cos^2(\hat{\phi}/2) - \cos^2(\phi/2)}}\, d\phi$$

Substitute $u = \cos(\phi/2)/\cos(\hat{\phi}/2)$ gives

$$\frac{\sin(\phi/2)}{\sqrt{\cos^2(\hat{\phi}/2) - \cos^2(\phi/2)}} = \frac{\sin(\phi/2)}{\cos(\hat{\phi}/2)\sqrt{1-u^2}}.$$

Since $du = -(1/2)\dfrac{\sin(\phi/2)}{\cos(\hat{\phi}/2)}\, d\phi$ and since u varies from 1 to 0 the integral is equal to

$$\int_0^1 \frac{2}{\sqrt{1-u^2}}\, du = 2\arcsin(u)\Big|_0^1 = \pi$$

So the time to decent from any point is $\pi\sqrt{a/g}$.
For history and some movies about this problem:

- `http://www.sewanee.edu/physics/TAAPT/TAAPTTALK.html`
- `http://www-history.mcs.st-and.ac.uk/HistTopics/Brachistochrone.html`
- `http://www.youtube.com/watch?v=li-an5VUrIA`
- `http://www.youtube.com/watch?v=gb81TxF2R_4&hl=ja&gl=JP`
- `http://www.youtube.com/watch?v=k6vXtjne5-c`
- Check out this book: Nahin: When Least Is Best. Great book!
- A student sent me the following link: `https://www.youtube.com/watch?v=Cld0p3a43fU`

E.5 Are the critical points of the functional I minima?

Are the critical points of I minima?

- In general we **cannot** guarantee that the solutions of the Euler-Lagrange equation gives a minimum.
- This is of course is not surprising: a minimum $\tilde{x}$ of $F\colon \mathbb{R}^n \to \mathbb{R}$ satisfies $DF(\tilde{x}) = 0$, but the latter condition is not enough to guarantee that $\tilde{x}$ is a minimum.
- It is also not always the case that a functional of the form $I[y] = \int_a^b f(x, y(x), y'(x))\, dx$ over the set $\mathcal{A} = \{y;\ y\colon [a,b] \to \mathbb{R}^n$ is C^1 and $y(a) = y_a, y(b) = y_b\}$ does have a minimum.
- Additional considerations are often required.

E.6 Constrains in finite dimensions

Often one considers problems where one has a constraint. Let us first consider this situation in finite dimensions:

E.6.1 Curves, surfaces and manifolds

Definition: We define a subset M of $\mathbb{R}^n$ to be a **manifold** (of codimension k) if $M = \{x \in \mathbb{R}^n; g(x) = 0\}$ where $g\colon \mathbb{R}^n \to \mathbb{R}^k$ and $k < n$ where the matrix $Dg(x)$ has rank k for each $x \in M$.

Remark: There are other, equivalent, definitions of manifolds and also some more general definitions of the notion of a manifold, but this goes outside the scope of this course.

Theorem E.11

Let $M \subset \mathbb{R}^n$ be a manifold of codimension k. Then near *every $x \in M$ one can write M as the graph of a function of $(n-k)$ of its coordinates.*

Examples. $M = \{(x,y) \in \mathbb{R}^2; x^2+y^2=1\}$ can be described locally in the form $x \mapsto (x, y(x))$ or in the form $y \mapsto (x(y), y)$.

Proof: Consider $x_0 \in M$ and for simplicity assume that the last k columns of the $k \times n$ matrix $Dg(x)$ are linearly independent. Then the $k \times k$ matrix made up of the last k columns of the $k \times n$ matrix $Dg(x)$ is invertible. This puts us in the position of the Implicit Function Theorem. Indeed, write $x = (u, v) \in \mathbb{R}^{n-k} \oplus \mathbb{R}^k$. The Implicit Function Theorem implies that there exists a function $G\colon \mathbb{R}^{n-k} \to \mathbb{R}^k$ so that

$$g(u,v) = 0 \iff v = G(u).$$

So M is locally a graph of a function G: the set is locally of the form $M = \{(u, G(u)); u \in \mathbb{R}^{n-k}\}$. (If some other combination of columns of $Dg(x)$ are linearly independent then we argue similarly.) ■

Examples:

- Assume that $g\colon \mathbb{R}^3 \to \mathbb{R}$ and consider $M = \{x \in \mathbb{R}^n; g(x) = 0\}$. Moreover assume that $Dg(x) \neq 0$ for each $x \in M$. Then M is a surface. Any (orientable) surface can be written in this form.
- The set $x^2+2y^2=1$, $x^2+y^4+z^6=1$ is a codimension-two manifold (i.e. a curve) in $\mathbb{R}^3$.

Definition: The **tangent plane** at $\hat{x} \in M$ is defined as the collection of vectors $v \in \mathbb{R}^n$ (based at $\hat{x}$) so that $Dg_{\hat{x}}(v) = 0$.

To motivate this definition consider a C^1 curve $\gamma\colon [0,1] \to M \subset \mathbb{R}^n$ with $\gamma(0) = \hat{x}$. Since $\gamma(t) \in M$, it follows that

$g \circ \gamma(t) = 0$ for all t and therefore

$$\frac{\partial g}{\partial x_1}(\hat{x})\gamma_1'(0) + \cdots + \frac{\partial g}{\partial x_n}(\hat{x})\gamma_n'(0) = 0.$$

This if we write $v = \gamma'(0)$ then $Dg(\hat{x})v = 0$. Hence $0 = Dg(\hat{x})v = \nabla g(\hat{x}) \cdot v$ where $\cdot$ is the usual dot product in $\mathbb{R}^n$. So the vector $\nabla g(\hat{x})$ is orthogonal to $v := \gamma'(0)$ for **each** such curve γ.

E.6.2 Minima of functions on constraints (manifolds)

Suppose $\tilde{x}$ is minimum of $F\colon M \to \mathbb{R}$ where $M = \{x \in \mathbb{R}^n; g(x) = 0\}$ and $g\colon \mathbb{R}^n \to \mathbb{R}$. What does this imply? Write $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_{n-1}, \tilde{x}_n)$.

Theorem E.12

*(**Lagrange multiplier**) If $Dg(\tilde{x}) \neq 0$ and $\tilde{x}$ is minimum of $F\colon M \to \mathbb{R}$, then $\exists \lambda \in \mathbb{R}$ with $DF(\tilde{x}) = \lambda Dg(\tilde{x})$.*

Proof: Since $Dg(\tilde{x}) \neq 0$, we get that $\frac{\partial g}{\partial x_i}(\tilde{x}) \neq 0$ for some $i = 1, \ldots, n$. In order to be definite assume $\frac{\partial g}{\partial x_n}(\tilde{x}) \neq 0$ and write $\tilde{w} = (\tilde{x}_1, \ldots, \tilde{x}_{n-1})$. By the Implicit Function Theorem, locally near $\tilde{w}$ there exits h so that $g(x) = 0 \iff x_n = h(x_1, \ldots, x_{n-1})$. So $\tilde{w}$ is minimum of $(x_1, \ldots, x_{n-1}) \mapsto F \circ (x_1, x_2, \ldots, x_{n-1}, h(x_1, \ldots, x_{n-1}))$. This means for all $i = 1, \ldots, n-1$:

$$\frac{\partial F}{\partial x_i}(\tilde{x}) + \frac{\partial F}{\partial x_n}(\tilde{x})\frac{\partial h}{\partial x_i}(\tilde{w}) = 0.$$

Since $g(x_1, \ldots, x_{n-1}, h(x_1, \ldots, x_{n-1})) = 0$ we also get

$$\frac{\partial g}{\partial x_i}(\tilde{x}) + \frac{\partial g}{\partial x_n}(\tilde{x})\frac{\partial h}{\partial x_i}(\tilde{w}) = 0 \ \forall i = 1, \ldots, n-1.$$

Substituting these into the previous equation and writing

$$\lambda = \frac{\dfrac{\partial F}{\partial x_n}(\tilde{x})}{\dfrac{\partial g}{\partial x_n}(\tilde{x})}$$

gives

$$\frac{\partial F}{\partial x_i}(\tilde{x}) - \lambda \frac{\partial g}{\partial x_i}(\tilde{x}) = 0 \ \forall i = 1, \ldots, n-1.$$

(For $i = n$ the last equation also holds, by definition.) ■

E.7 Constrained Euler-Lagrange Equations

Let $I[y] = \int_a^b f(x, y(x), y'(x))\, dx$ and $J[y] = \int_a^b g(x, y(x), y'(x))\, dx$ be functionals on

$$\mathcal{A} = \{y;\ y\colon [a,b] \to \mathbb{R}^n \text{ is } C^1 \text{ and } y(a) = y_a, y(b) = y_b\}.$$

as before. Define

$$M = \{y;\ y \in \mathcal{A} \text{ with } J[y] = 0\}.$$

Theorem E.13

If $M \ni y \mapsto I[y]$ has a minimum at $\tilde{y}$ then there exists $\lambda \in \mathbb{R}$ so that the E-L condition hold for $F = f - \lambda g$. That is,

$$\frac{d}{dx} F_{y'}[\tilde{y}] = F_y[\tilde{y}].$$

The idea of the proof combines the Lagrange multiplier approach with the proof of the previous Euler Lagrange theorem.

Example E.14

Maximize the area bounded between the graph of y and the *line segment* $[-1,1]\times\{0\}$, *conditional on the length of the arc being* L. *(This is a special case of* **Dido's problem**.*)*

Let $\mathcal{A}$ *be the set of* C^1 *functions* $y\colon[-1,1]\to\mathbb{R}$ *with* $y(-1)=y(1)=0$. *Fix* $L>0$ and let

$$I[y]=\int_1^1 y(x)\,dx \text{ and } J[y]=\int_{-1}^1\sqrt{1+(y')^2}\,dx-L=0.$$

Write

$$f=y, g=\sqrt{1+(y')^2}, F=f-\lambda g=y-\lambda\sqrt{1+(y')^2}.$$

The Euler Lagrange equation in the version of Lemma E.7 gives $F_{y'}[\tilde{y}]\tilde{y}'-F[\tilde{y}[=C$ *which amounts to (writing* y *instead of* $\tilde{y}$*):*

$$\frac{-\lambda(y')^2}{\sqrt{1+(y')^2}}-[y-\lambda\sqrt{1+(y')^2}]=C.$$

Rewriting this gives

$$1=\frac{(y+C)^2}{\lambda^2}(1+(y')^2).$$

Substituting $y+C=\lambda\cos\theta$ gives $y'=-\lambda\sin\theta\dfrac{d\theta}{dx}$. Substituting this in the previous equation gives

$$1=\cos^2\theta\left(1+\lambda^2\sin^2\theta\left(\frac{d\theta}{dx}\right)^2\right).$$

Since $\cos^2\theta+\sin^2\theta=1$, this implies

$$\lambda\cos\theta\frac{d\theta}{dx}=\pm1, \text{ i.e. } \frac{dx}{d\theta}=\pm\lambda\cos\theta$$

which means $x=\pm\lambda\sin\theta$ and $y+C=\lambda\cos\theta$: a circle segment!

Appendix F Some results on Fourier series

Lemma F.1

$\sum n^2|c_{1,n}| < \infty$ and $\sum n^2|c_{2,n}| < \infty \implies$

$$u(x,t) = \sum_{n\geq 1}(c_{1,n}\cos(nt) + c_{2,n}\sin(nt))\sin(nx)$$

is C^2. (C^2 means that the function is twice differentiable and that the 2nd derivatives are continuous.)

Proof : That $\sum_{n=1}^{N}(c_{1,n}\cos(nt)+c_{2,n}\sin(nt))\sin(nx)$ converges in this case follows from

Weierstrass test: if $M_n \geq 0$, $\sum M_n < \infty$ and $u_n\colon [a,b] \to \mathbb{R}$ is continuous with $\sup_{x\in[a,b]}|u_n(x)| \leq M_n$ then $\sum u_n$ converges uniformly on $[a,b]$ (and so the limit is continuous too!).

To get that u is one differentiable, we need to consider for example the convergence (as $N \to \infty$) of the d/dx derivative of

$$\sum_{n=1}^{N}(c_{1,n}\cos(nt) + c_{2,n}\sin(nt))\sin(nx)$$

which is equal to $\sum_{n=1}^{N}(c_{1,n}\cos(nt)+c_{2,n}\sin(nt))n\cos(nx)$, and since $\sum n|c_{1,n}|, \sum n|c_{2,n}| < \infty$ the latter converges. In this way we obtain that $u(x,t)$ is differentiable w.r.t. x.

To check that $u(x,t)$ is twice differentiable, we differentiate the sum term by term once more, and to apply Weierstrass again we need that $\sum n^2|c_{1,n}|, \sum n^2|c_{2,n}| < \infty$. ■

Next we need to make sure that the boundary conditions are satisfied.

Lemma F.2

If $f\colon [0,\pi] \to \mathbb{R}$ and $f(0) = f(\pi) = 0$ then f has a Fourier expansion of the form $f(x) \sim \sum_{n=1}^{\infty} s_{1,n} \sin(nx)$.

Proof: Define $g\colon [0,2\pi] \to \mathbb{R}$ so that $g(x) = f(x)$ for $x \in [0,\pi]$ and $g(x) = f(2\pi - x)$ for $x \in [\pi, 2\pi]$. It follows that $g(\pi - x) = -g(\pi + x)$ for $x \in [0,\pi]$. So this means that $\int_0^{2\pi} g(x) \cos(nx)\, dx = 0$ and therefore in the Fourier expansion of g the cosine terms vanish, and we have $g(x) = \sum_{n=1}^{\infty} s_{1,n} \sin(nx)$. In particular $f(x) = \sum_{n=1}^{\infty} s_{1,n} \sin(nx)$. ■

Theorem F.3

If f is C^3 and $f(0) = f(\pi) = f''(0) = f''(\pi) = 0$, then there are coefficients s_n, c_n so that

$$f(x) = \sum_{n\geq 1} s_n \sin(nx) \text{ and } f'(x) = \sum c_n \cos(nx)$$

and $\sum n^2 s_n^2 < \infty$ and $\sum n^2 c_n^2 < \infty$.

Proof: The proof below is in sketchy form only. Let us assume that f is C^2, $f(0) = f(\pi) = 0$. According to the Fourier Theorem 7.2 one can write $f(x) = \sum_{n\geq 1} s_n \sin(nx)$. Let us now show that if f is C^3 and $f(0) = f(\pi) = f''(0) = f''(\pi) = 0$ the assumptions in Lemma F.1 are satisfied, i.e. that f and f' can be written in the form $f(x) = \sum s_n \sin(nx)$ and $f'(x) = \sum c_n \cos(nx)$ and that $\sum n^2 s_n^2 < \infty$ and $\sum n^2 c_n^2 < \infty$. (We change the notation from the coefficients c_n to s_n in the main text since the new notation is more natural here.) Let us prove that $\sum |s_n| < \infty$.

First choose constants s_n and c_n so that $f(x) = \sum s_n \sin(nx)$ and $f'(x) = \sum c_n \cos(nx)$ (by the Fourier theorem one can write f' in this way since is C^2 and since $f''(0) = f''(\pi) = 0$).

Step 1:

$$(f', f') = \sum_{n,m\geq 0} c_n c_m(f') \int_0^\pi \cos(nx)\cos(mx) =$$

$$= (\pi/2)\sum_{n\geq 1} |c_n|^2 + \pi|c_0|^2.$$

It follows that $\sum_{n\geq 0}|c_n|^2 < \infty$.

Step 2: for $n \geq 1$ we have

$$s_n = (2/\pi)\int_0^\pi f(x)\sin(nx)\,dx$$

and

$$c_n = (2/\pi)\int_0^\pi f'(x)\cos(nx)\,dx.$$

Using partial integration on the last expression, and using that $f(0) = f(\pi) = 0$ gives for $n \geq 1$,

$$c_n = (2/\pi)\int_0^\pi f'(x)\cos(nx)\,dx = (2/\pi)[f(x)\cos(nx)]_0^\pi +$$

$$+n(2/\pi)\int_0^\pi f(x)\sin(nx)\,dx = ns_n.$$

It follows from this, $f(0) = f(\pi) = 0$ and Step 1 that $\sum n^2|s_n|^2 < \infty$.

Step 3: Now we use the Cauchy inequality $\sum a_n b_n \leq \sum a_n^2 \sum b_n^2$. Taking $a_n = 1/n$ and $b_n = n|s_n|$ we get that $\sum|s_n| = \sum a_n b_n \leq \sum a_n^2 \sum b_n^2$. By Step 2, $\sum b_n^2 < \infty$ and since $\sum 1/n^2 < \infty$, it follows that $\sum|s_n| < \infty$.

In the same way, we can prove that **if** f is C^3 and $f(0) = f(\pi) = f'(0) = f'(\pi) = f''(0) = f''(\pi) = 0$ then $\sum n^2|s_n| < \infty$. Therefore the assumptions in Lemma F.1 are satisfied.

If we assume $f(0) = f(\pi) = f''(0) = f''(\pi) = 0$ and consider $g(x) = f(x) - a_1\sin x - a_2\sin 2x$ with a_1, a_2 so that $g'(0) = g'(\pi) = 0$ then we can apply the previous paragraph to g. It follows that $\sum n^2|s_n(g)| < \infty$. This also implies $\sum n^2|s_n| < \infty$. This concludes the explanation of item 2 above Theorem 7.2. ■