18.152: Introduction to Partial Differential Equations

Thomas Walpuski
Massachusetts Institute of Technology

2016-02-29
About this course

What you can expect

I will give you an introduction to partial differential equations as best as I can. If there is any particular topic you are interested in and would like me to cover, let me know and I will make an effort to include it in one of the lectures.

There are going to be lectures notes available on my MIT webpage at https://math.mit.edu/~walpuski/18.152/.

If you find mistakes, be it typos or actual mathematical issues, please point them out to me and I will fix them.

My office hours are Wednesdays, 10am–12n in 2-231B. If these times are inconvenient for you, can always approach me after the lectures or email me to make an appointment.

If you have any special circumstances that I need to pay attention to please let me know as soon as possible and I will try to accommodate your request. Also, see below for more information.

What I expect from you

Exercises and homework

Every lecture contains a number of exercises. I expect you to attempt all of them. They are an important part of the class. Almost every week, I will prepare a problem set (mostly containing a selection of these exercises) which will be marked and will make up your homework grade. The problem sets can be found on my MIT webpage at https://math.mit.edu/~walpuski/18.152/.

I will drop the lowest homework grade. Reasonable requests for late submission will be accepted; in particular, I will accept any request if it is backed with a letter from S3.

I strongly encourage you to prepare your homework in \LaTeX. Efficient usage of \LaTeX is a crucial skill and this is a good a time as any to learn or to practice it. (If you need to learn \LaTeX first, then for the first three weeks of the term submitting two days late is deemed reasonable.)

I also strongly encourage you to work on the problem sets in groups. However, you need to write your solutions yourself and list all of your collaborators and sources in each assignment.
Reading
I do expect you to read the lecture notes. I will sometimes ask you to read up one a specific topic in preparation for the next lecture.

Midterm
There is going to be one midterm on March 15, 2016 in place of the usual class.

Final exam
The final exam is going to take place on May 19, 2016 from 9am to 12n in 4-153.

Final grade
For the final grade I will use the following weights:

- homework 40%
- midterm 20%
- final exam 40%

Special circumstances

Student Support Services
If you are dealing with a personal or medical issue that is impacting your ability to attend class, complete work, or take an exam, please discuss this with Student Support Services (S3). The deans in S3 will verify your situation, and then discuss with you how to address the missed work. Students will not be excused from coursework without verification from Student Support Services. You may consult with Student Support Services in 5-104 or at 617-253-4861. Also, S3 has walk-in hours Monday-Friday, 10-11am and 2-3pm.

Student Disability Services
MIT is committed to the principle of equal access. Students who need disability accommodations are encouraged to speak with Kathleen Monagle, Associate Dean, prior to or early in the semester so that accommodation requests can be evaluated and addressed in a timely fashion. Even if you are not planning to use accommodations, it is recommended that you meet with SDS staff to familiarise yourself with the services and resources of the office. You may also consult with Student Disability Services in 5-104 or at 617-253-1674. If you have already been approved for
accommodations, please contact me early in the semester so that we can work
together to get your accommodation logistics in place.
## Contents

1 **What are PDE and why should you care?** 6  
   1.1 What is a PDE? ................................................. 6  
   1.2 Some interesting examples ............................. 7  
   1.3 Questions about PDE ................................. 9  
   1.4 Linear PDE .................................................. 10  
   1.5 More stuff you should know ......................... 12  

2 **Ordinary Differential Equations** 13  
   2.1 Uniqueness of solutions ............................. 14  
   2.2 Local existence theorems ............................ 17  
   2.3 Dependence of solutions on initial conditions ... 18  

3 **First-order PDE** 19  
   3.1 Change of coordinates ............................... 19  
   3.2 Warm-up: the transport equation ................. 20  
   3.3 The method of characteristics ................. 22  

4 **The inviscid Burgers’ equation** 25  

5 **Introduction to the Heat Equation** 30  
   5.1 Boundary conditions ............................... 31  
   5.2 A toy model .............................................. 33  
   5.3 Solution of the heat equation with Dirichlet boundary conditions in dimension one ................................. 34  

6 **Uniqueness for the Heat Equation** 37  
   6.1 Energy method ........................................ 38  
   6.2 Weak maximum principle ......................... 39  

7 **The Heat Kernel on $\mathbb{R}^n$** 41  

8 **Introduction to the Wave Equation** 45  
   8.1 A toy model .............................................. 45  
   8.2 The wave equation on $[0, 1]$ ..................... 46  
   8.3 d’Alembert’s formula .............................. 46
9 The wave equation in dimension three

9.1 Spherical means .................................................. 49
9.2 Kirchhoff’s formula ............................................. 50

A Divergence Theorem .............................................. 53

B Metric spaces ....................................................... 54

C Fourier Series on \([0, 1]\) ........................................ 58
LECTURE 1. What are PDE and why should you care?

1.1 What is a PDE?

Let $U$ be an open subset in $\mathbb{R}^n$ and $u: U \to \mathbb{R}$ be a function. We will typically work with the standard coordinate system $(x_1, \ldots, x_n)$ on $\mathbb{R}^n$ and think of $U$ as space and of $u$ as assigning values $u(x)$ to points $x = (x_1, \ldots, x_n)$ in this space (be it the temperature at this point, the height of a membrane lying over the space, . . . ). Sometimes, we want to single out a specific direction and think of it as time. In these cases we use coordinates $(t, x_1, \ldots, x_n)$ on $\mathbb{R}^{n+1}$.

I will assume that you are familiar and comfortable with partial derivatives. As you will know the notations for partial derivatives are a nightmare. The notations we use for partial derivatives in this class are of the following form:

$$\frac{\partial u}{\partial x_i} = \partial_i u, \quad \frac{\partial^2 u}{\partial x_i \partial x_j} = \partial_{ij} u, \quad \ldots .$$

Sometimes we want to encode all the first partial derivatives in one piece of data and write

$$\nabla u = \begin{pmatrix} \partial_1 u \\ \partial_2 u \\ \vdots \\ \partial_n u \end{pmatrix}.$$ 

This is a function with values in $\mathbb{R}^n$. To encode all the $k$–partial derivatives we write $\nabla^k u$. This is a function with values in $\mathbb{R}^{nk}$ since there are $n^k$ different $k$–th partial derivatives of $u$. (For $k = 2$ one can still try to use matrices as a book-keeping tool, but it is best to think of $\nabla^k u$ just as a collection of all the $k$–th partial derivatives.)

A partial differential equation is an equation we impose on the various partial derivatives of $u$ and these equations typically can be interpreted as saying that $u$ is a plausible model for something we care about or defines a mathematical object we are interested in. The rigorous definition we are going to use is the following:

**Definition 1.1.** A partial differential equation (PDE) of order $k$ in a single unknown $u: U \to \mathbb{R}$ is an equation of the form

$$F(x, u(x), \nabla u(x), \nabla^2 u(x), \ldots, \nabla^k u(x)) = 0 \quad \text{for all } x \in U$$

where $F$ is some function.

**Remark 1.3.** If $u$ takes values in $\mathbb{R}^m$ instead of just $m$, it still makes sense to write (1.2) for suitable functions $F$. To emphasise $u$ being vector valued one then
sometimes speaks of a system of PDE (if \( m \geq 2 \)). In this class, however, we will almost exclusively deal with the case \( m = 1 \).

**Example 1.4.** The equation
\[
\sin(\partial_t u) + \frac{1}{1 + (\partial_x^2 u)|\partial_y u|} = t.
\]
is a second order PDE.

As you can see the notion of PDE is extremely general and on this level of generality one can say almost nothing.

### 1.2 Some interesting examples

The preceding example was very artificial, and as far as I know it does not correspond to anything meaningful. Here I will give you a list of PDE, most of which we will discuss in greater detail later in this class.

**Example 1.5.** An equation of the form
\[
v(t, x) \cdot \nabla u(t, x) = 0
\]
with \( v: U \rightarrow \mathbb{R}^{n+1} \) some vector-field on \( \mathbb{R}^{n+1} \) is called a transport equation. Solutions of the transport equation can be used to model concentrations of substances in moving media (think air/water pollution). There is a fantastic theory for understanding this PDE—in fact, first order PDE in general, called the method of characteristics.

**Example 1.6.** The inviscid Burger’s equation is the PDE
\[
\partial_t u + u \partial_x u = 0.
\]
One can think of it as a non-linear transport equation. As we will see the non-linearity will cause singularities (shocks) in the solutions, and this PDE is one of the simplest to exhibit this phenomenon.

**Example 1.7.** The heat equation is the PDE
\[
\partial_t u + \Delta u = 0
\]
where
\[
\Delta u := -\sum_{i=1}^{n} \partial_i^2 u
\]
is the *Laplacian* of $u$. It models the time-evolution of a heat distribution on $U$ with $u(t, x)$ describing the temperature at time $t$ and the point $x$ in space. Later in this class I will derive a number of properties of solutions heat equation which are in accordance with this model and I might also explain how to arrive at this model.

The heat equation is also intimately related with Brownian motion in probability theory. If you are interested in this let me know, I will talk about this in some detail.

One can make sense of the heat equation on an arbitrary Riemannian manifold $(M, g)$ (don’t be scared if you don’t know what that is). It turns out that by studying the heat equation on $M$ one can learn a lot about the geometry of $M$. Richard Hamilton [Ham82] introduced a famous non-linear heat equation for the metric $g$ itself, called *Ricci flow*:

$$\partial_t g(t) = -2\text{Ric}_g(t) = -\Delta g(t) + \ldots$$

This flow was used by Grigori Perelman to solve the Poincaré conjecture in 2003, see, e.g., [MT07]. Another non-linear heat equation was used by Simon Donaldson in his PhD thesis to establish a tight relation between so called Hermitian–Yang–Mills metrics an holomorphic vector bundles [Don85].

**Example 1.8.** The *wave equation* is the PDE

$$\partial_t^2 u + \Delta u = 0.$$  

It models vibrating membranes (or strings). Although it looks somewhat similar to the heat equation it is remarkably different as we will see later.

The Maxwell’s equations in the theory of electro-magnetism can be thought of as a wave equation. Einstein’s equations of General Relativity are a system of non-linear wave equations.

**Example 1.9.** We will spend a lot of time in this class studying the *Laplace equation*

$$\Delta u = 0$$

and its inhomogeneous version, the *Poisson equation*

$$\Delta u = f.$$  

Solutions of the Laplace equation are static solutions to the heat and wave equation, so they model steady-state heat distributions and standing waves.

On a Riemannian manifold $(M, g)$ one can study various versions of the Laplace equation and it turns out that the solutions are intimately related with the topology of $M$: “you can use standing waves on $M$ to predict the number of holes in $M$, and vice versa”. This goes under the name of *Hodge theory*. 

8
Example 1.10. The minimal hypersurface equation

$$\sum_{i=1}^{n} \partial_i \left( \frac{\partial_i u}{\sqrt{1 + |\nabla u|^2}} \right) = 0$$

describes hypersurfaces in $\mathbb{R}^{n+1}$ of least volume (with prescribed boundary). A deep study of this equation is beyond the scope of this course (and this equation is still being very actively researched), but I will discuss some simple results later in this class.

You may know many, many more examples of PDE from whatever your background is. (As you might have noticed my bias is towards geometry.) If there is a particular PDE, which you are fond of and would like me to discuss, let me know and I will try to accommodate your request.

1.3 Questions about PDE

Many PDE arise as models of some, say, physical or geometric phenomenon. If you are coming from a modelling/physical perspective, then the first thing you might be interested in would be: Does this model describe what I want it to describe? Is it a good model? Is a certain effect/phenomenon that I observe in “the real world” a consequence of the equation or does it only occur for special parameters? Does the model make interesting predictions?

In pure mathematics one typically disconnects from reality declares that “the model is the actual thing”. Whether you do this or not, given a PDE, typical questions one wants to answer are:

- Does the PDE have a solution?
- If it has solutions, then how many are there? Are there finitely many or maybe even a unique one?
- Can I prescribe more data (initial data, boundary data, ...) to pin down a unique solution?
- If there are infinitely many, is the space of solutions finite or infinite dimensional? Does it have any interesting structure?
- What are the basic properties of solutions? Can they grow in space/time? Do they decay?
- Can solutions become singular? If so, how do they become singular?
• How do the solutions change when one varies the data? Is the change continuous? Are there special solutions at which the change is discontinuous?

• Can we make quantitative statements about the solution given the data? How “large” can a solution be for data of a given “size”? What are good measures of “size”?

You may think that the best way to answer these questions is to find formulae for all possible solutions and then just compute the answers. In some cases this is what we will do, indeed. Often, however, one cannot expect to find explicit formulae, and even if we can find the closed form formulae they are often very very complicated and computing anything from them can be a very tedious task. Instead, I will show you many cases in which you just use the PDE itself to derive consequences (a theory if you wish) for the solutions of the PDE. This is often as good (or even better) than having a very complicated solution formula.

1.4 Linear PDE

As I mentioned earlier, the notion of PDE is too general to say much. An important class of PDE for which a good amount of theory exist is that of linear PDE. Let me introduce some definitions.

Definition 1.11. A linear differential operator of order \( k \) is map \( \mathcal{L} : C^k(U) \to C^0(U) \) of the form

\[
(\mathcal{L} u)(x) = \sum_{i=0}^{k} a_i(x) [\nabla^i u(x)]
\]

where \( a_k(x) : \mathbb{R}^n \to \mathbb{R} \) is linear and depends continuously on \( x \).

If the \( a_i \) do not depend on \( x \) at all, we say \( \mathcal{L} \) is a constant coefficient linear differential operator.

Example 1.12. The Laplacian \( \Delta \) is a constant coefficient linear differential operator, and so are the heat operator and the wave operator

\[
\partial_t + \Delta \quad \text{and} \quad \partial_t^2 + \Delta.
\]

These differential operators are archetypal; we will spend a lot of time studying them.

Example 1.13. The map

\[
u \mapsto \mathcal{L} u := (t + x^2) \Delta u
\]

is a linear differential operator, but its coefficients are not constant.
Example 1.14. The equation
\[ \partial_t^2 u + \cos(\partial_x u) \partial_x^5 u = 0 \]
is a non-linear PDE.

Remark 1.15. The linear in linear differential operator comes from the fact that
\[ \mathcal{L}(u + v) = \mathcal{L}u + \mathcal{L}v \quad \text{and} \quad \mathcal{L}(\lambda u) = \lambda \mathcal{L}u \]
for any \( u, v \in C^k(U) \) and \( \lambda \in \mathbb{R} \). (If this is not clear to you, you should sit down and verify this formally.)

This means that \( \mathcal{L} \) is a linear map between infinite-dimensional vector spaces. Functional analysis is a theory of such infinite-dimensional vector spaces and many results from it can be used to (and in fact have been developed to) understand (certain) linear differential operators extremely well.

Definition 1.16. A (constant coefficient) linear PDE is a PDE of the form
\[ \mathcal{L}u = f \]
with \( \mathcal{L} \) a (constant coefficient) linear differential operator and \( f \) a function on \( U \).

If \( f = 0 \), we say that the linear PDE is homogeneous.

The linearity of \( \mathcal{L} \) has important consequences, often summarised under the heading of “superposition principle”.

Proposition 1.17. If \( v, w \) are solutions of the homogeneous linear PDE
\[ \mathcal{L}u = 0, \]
then so is \( v + \lambda w \) for any \( \lambda \in \mathbb{R} \).

That is, the space of solutions
\[ \ker \mathcal{L} = \{ u : \mathcal{L}u = 0 \} \]
is a vector space.

Proposition 1.18. If \( v, w \) are solutions to the inhomogeneous linear PDE
\[ \mathcal{L}u = f, \]
then \( v - w \) solves the homogeneous linear PDE. Conversely, if \( u_p \) is a solution to the inhomogeneous linear PDE and \( u_h \) is a solution to the homogeneous linear PDE, then \( u_p + u_h \) is a solution to the inhomogeneous linear PDE as well.

This means that the space of solutions to \( \mathcal{L}u = f \) is an affine space modelled on \( \ker \mathcal{L} \):
\[ \{ u : \mathcal{L}u = f \} = u_p + \ker \mathcal{L}. \]
1.5 More stuff you should know

For the rest of the class the material covered in Appendix A and Appendix B will be important. If I don’t manage to talk about this during Lecture 1, please consider reading these appendices to be homework due before Lecture 2.
LECTURE 2. Ordinary Differential Equations

We begin this PDE class with reviewing ODE. The purpose if this is two-fold. First of all, in the next lecture we will see a class of PDE that can be reduced to ODE. Second, some of the ideas used to prove Picard–Lindelöf’s theorem on the existence and uniqueness for ODEs can be applied to certain PDE as well.

Definition 2.1. A (system of) ordinary differential equations (ODEs) of first order is an equation of the form

\[ \dot{x}(t) = F(t, x(t)) \]

with \( F: U \to \mathbb{R}^n \) and \( U \subset \mathbb{R} \times \mathbb{R}^n \) open.

We also call (2.2) a dynamical system and call a time-dependent vector field.

Remark 2.3. Any system of ODEs of any order can be converted to a system of ODEs of first order by introducing auxiliary variables.

Exercise 2.4. If you unfamiliar with the idea of Remark 2.3, please, read up on this.

Let us introduce some language to talk about solutions of (2.2).

Definition 2.5. A solution of (2.2) is a differentiable map \( \phi: I \to \mathbb{R}^n \) defined on an interval \( I \subset \mathbb{R} \) such that for each \( t \in I \) we have

\[ (t, \phi(t)) \in U \quad \text{and} \quad \dot{\phi}(t) = F(t, \phi(t)). \]

We also say that \( \phi: I \to \mathbb{R}^n \) is an integral curve of the time-dependent vector field \( F \).

Definition 2.6. A system of ODE of first order (2.2) together with an equation of the form

\[ x(t_0) = x_0 \]

with \( (t_0, x_0) \in U \) is called an initial value problem (IVP).

A solution \( \phi: I \to \mathbb{R}^n \) to (2.2) is called a solution to the IVP (2.2) and (2.7) if \( t_0 \in I \) and \( \phi(t_0) = x_0 \).

It is often useful to rewrite an ODE as an equivalent integral equation.

Proposition 2.8. Suppose \( F \) is continuous. A function \( \phi: I \to \mathbb{R}^n \) is a solution to the IVP (2.2) and (2.7) if and only if for all \( t \in I \) we have

\[ \phi(t) = \phi(t_0) + \int_{t_0}^{t} F(s, \phi(s)) \, ds. \]
Proof. If $F$ is continuous, then any solution $\phi$ to (2.2) is $C^1$ and, in particular, continuous. Consequently $s \mapsto F(s, \phi(s))$ is continuous and thus integrable. The assertion now follows from the fundamental theorem of calculus.

2.1 Uniqueness of solutions

Example 2.10 (Failure of Uniqueness). Imagine an infinitely high\(^1\) cylindrical container filled with water and a drain at the bottom. Denote by $x$ the water level (in some suitable units). If the container is empty, then it stays empty and nothing changes; if the water is at level $x \geq 0$, then it drains at rate $2\sqrt{x}$. That is, $x$ is governed by the IVP

$$\dot{x}(t) = F(x(t)) \quad \text{and} \quad x(0) = 0$$

with

$$F(x) := \begin{cases} -2\sqrt{x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$  

For each $0 \leq T \leq \infty$, the function $\phi_T: \mathbb{R} \to \mathbb{R}$ defined by

$$\phi_T(t) = \begin{cases} (t + T)^2 & t \leq -T \\ 0 & t > -T \end{cases}$$

is a solution of this IVP; it describes the situation where the container has been draining from $t = -\infty$ until it is empty at $t = -T$.

The above describes a physical situation in which uniqueness cannot possibly hold. It turns out, however, that solutions to ODE are unique under very mild assumptions on $F$.

Definition 2.11. A function $F: U \to \mathbb{R}^m$ is called Lipschitz continuous if there exist a constant $L > 0$ such that for all $x, y \in U$

$$\frac{|F(x) - F(y)|}{|x - y|} \leq L.$$  

We call

$$\text{Lip}(F) := \sup \left\{ \frac{|F(x) - F(y)|}{|x - y|} : x, y \in U \right\} \in [0, \infty]$$

the Lipschitz constant of $F$. Clearly, $F$ is Lipschitz continuous if and only if $\text{Lip}(F) < \infty$.

\(^1\)This is ridiculous of course, but the model becomes easier to describe assuming this.
Remark 2.12. Lipschitz continuity is a rather strong form of continuity. It is closely related with differentiably: Rademacher’s Theorem asserts that Lipschitz functions are almost everywhere differentiable. To know what “almost everywhere” means and understand the proof you would have to know a bit of measure theory (which I don’t expect you to know for this course).

Hypothesis 2.13. Suppose that \( F: U \to \mathbb{R}^n, (t, x) \mapsto F(t, x) \) is continuous and for each \((t, x) \in U\) there exists a neighbourhood \( K \) of \( t \in \mathbb{R} \) and \( V \) of \( x \in \mathbb{R}^n \) such that \( K \times V \subset U \) and there exists a constant \( L > 0 \) such that for all \( s \in K \) and \( y, z \in V \)

\[
\left| \frac{F(s, y) - F(s, z)}{|y - z|} \right| \leq L.
\]

Remark 2.15. This might appear like a very technical hypothesis, but it is exactly what is needed to make the proof work. If you don’t like this hypothesis, assume instead that \( F \) is \( C^1 \).

Theorem 2.16 (Uniqueness Theorem). Assume Hypothesis 2.13. If \( \phi_1, \phi_2: I \to \mathbb{R}^n \) are solutions to (2.2) and for some \( t_0 \in I \) we have

\[ \phi_1(t_0) = \phi_2(t_0), \]

then

\[ \phi_1 = \phi_2. \]

Remark 2.17. Note that \( F \) in Example 2.10 is not Lipschitz in any neighbourhood of 0.

For the proof we need the following very simple but tremendously useful lemma.

Lemma 2.18 (Grönwall’s Lemma). Let \( g: I \to \mathbb{R} \) be a continuous function with \( g \geq 0 \) and \( t \in I \). If \( A, B \geq 0 \) are constants such that for all \( t \in I \)

\[ g(t) \leq A \left| \int_{t_0}^{t} g(s) \, ds \right| + B, \]

then for all \( t \in I \)

\[ g(t) \leq Be^{A|t-t_0|}. \]

Proof. We prove this for \( t \geq t_0 \) only. The case \( t < t_0 \) is similar. Please, make sure you understand how the following argument needs to be adapted in this case.

The function defined by

\[ G(t) := A \int_{t_0}^{t} g(s) \, ds + B \]
satisfies
\[ \dot{G}(t) \leq Ag(t) \leq AG(t). \]
Here the last inequality is where the hypotheses on \( g \) are used. It follows that
\[ G(t) \leq G(t_0)e^{A(t-t_0)} = Be^{A(t-t_0)}. \]
This completes the proof because \( g(t) \leq G(t) \).

\[ \square \]

**Proof of Theorem 2.16.** Define
\[ J := \{ t \in I : \phi_1(t) = \phi_2(t) \}. \]
We will show that \( J \) is a non-empty, closed and open subset of \( I \); hence, it must be the whole interval. (Do you know why this is true? This is a very simple topological fact, but it is very useful.)

**Step 1.** By assumption \( t_0 \in J \); hence, \( J \) is non-empty.

**Step 2.** \( J \) is closed.

Since both \( \phi_1 \) and \( \phi_2 \) are differentiable, they are continuous and so is
\[ \delta := \phi_1 - \phi_2. \]
Thus \( J \) is closed because it can be written as
\[ J = \delta^{-1}(0). \]

**Step 3.** \( J \) is open.

Suppose we are given a point in \( J \). We may as well denote this point by \( t_0 \). We will prove that a small neighbourhood of \( t_0 \) in \( I \) is also contained in \( J \). Choose neighbourhoods \( K \) of \( t \in \mathbb{R} \) and \( V \subset \mathbb{R}^n \) of \( x_0 = \phi_1(t_0) = \phi_2(t_0) \in \mathbb{R}^n \) such that \( K \times V \subset U \) and (2.14) holds for all \( t \in K \) and \( y,z \in U \). Using the integral form of the ODE (2.9), for \( t \in K \cap I \), we can write
\[ |\delta(t)| \leq \left| \int_{t_0}^{t} |F(s, \phi_1(s)) - F(s, \phi_2(s))| \, ds \right| \leq L \left| \int_{t_0}^{t} |\delta(s)| \, ds \right|. \]
By Lemma 2.18 with \( B = 0, \delta = 0 \) in \( K \cap I \); hence, \( J \supset K \cap I \).
2.2 Local existence theorems

Exercise 2.19. Find a function $F: \mathbb{R} \to \mathbb{R}$ such that there is no solution to the IVP
\[
\dot{x}(t) = F(t) \quad \text{and} \quad x(0) = 0.
\]

Theorem 2.20 (Picard–Lindelöf). Assume Hypothesis 2.13. For every $(t_0, x_0) \in U$ there is an interval $I$ containing $t_0$ on which the IVP
\[
(2.21) \quad \dot{x} = F(t, x) \quad \text{and} \quad x(t_0) = x_0
\]
has a solution.

Note how the proof is constructive and may (in principle) be used to compute solutions. The relevant iteration is sometimes called Picard–Lindelöf iteration.

Proof. The actual proof may appear a bit technical, but the idea is very simple: Note that the integral form
\[
\phi(t) = x_0 + \int_{t_0}^{t} F(s, \phi(s)) \, ds
\]
has the shape of a fixed point equation
\[
\phi = T\phi
\]
where $T: X \to X$ is defined by
\[
(2.22) \quad T(\phi)(t) := x_0 + \int_{t_0}^{t} F(s, \phi(s)) \, ds
\]
and $X$ is a suitable space of functions. If we arrange things carefully $T$ will be a contraction and we can deduce the existence of a fixed point from Theorem B.17.

Now, let’s roll our sleeves up and prove this make this rigorous.

Step 1. Because of Hypothesis 2.13, we can fix $\delta, \varepsilon, L > 0$ such that (2.14) holds for all $s \in [t_0 - \delta, t_0 + \delta]$ and $y, z \in \overline{B}_{2\varepsilon}(x_0)$. We can also assume, by making $\varepsilon$ and $\delta$ smaller, that
\[
\gamma := \delta L
\]
and
\[
\delta(L\varepsilon + M) \leq \varepsilon
\]
where $M := \sup \{|F(s, x_0)| : s \in [t_0 - \delta, t_0 + \delta]\}$.
Step 2. Set $I := [t_0 - \delta, t_0 + \delta]$ and

$$X := C^0(I; \bar{B}_\varepsilon(x_0)) = \{ \phi : I \to \bar{B}_\varepsilon(x_0) \text{ continuous} \}$$

The formula (2.22) defines a map $T : X \to X$.

(2.22) certainly defines a map $X \to C^0(I; \mathbb{R}^n)$. To see that it maps into $X$, note that

$$|T\phi(t) - x_0| \leq \left| \int_{t_0}^t |F(s, \phi(s)) - F(s, x_0)| \, ds \right| + |F(s, x_0)| \leq \delta(L\varepsilon + M) \leq \varepsilon.$$

Step 3. $C^0(I; \bar{B}_\varepsilon(x_0))$ is a complete metric space.

This is a question on problem set #1.

Step 4. $T : X \to X$ is a contraction.

It follows from (2.14) that

$$d(T\phi_1, T\phi_2) = \sup_I \left| \int_{t_0}^t F(s, \phi_1(s)) - F(s, \phi_2(s)) \, ds \right|$$

$$\leq \sup_I \left| \int_{t_0}^t \left| F(s, \phi_1(s)) - F(s, \phi_2(s)) \right| \, ds \right|$$

$$\leq \sup_I \int_{t_0}^t L|\phi_1(s) - \phi_2(s)| \, ds$$

$$\leq \delta Ld(\phi_1, \phi_2) = \gamma d(\phi_1, \phi_2).$$

This completes the proof. \qed

Remark 2.23. The conclusion of Theorem 2.20 holds under weaker assumptions: Peano’s theorem asserts that it suffices for $F$ to be continuous.

2.3 Dependence of solutions on initial conditions

Theorem 2.24. Suppose $F$ is $C^1$. If $\varepsilon > 0$ and $V \subset \mathbb{R}^n$ is a bounded open subset, such that $[-\varepsilon, \varepsilon] \times V \subset U$, then there is a $\delta \in (0, \varepsilon)$ and a $C^1$ map

$$\Phi : (-\delta, \delta) \times V \to U$$

such that

$$\Phi(0, x_0) = x_0$$

and for each $x_0 \in V$, $t \mapsto \Phi(t, x_0)$ is an integral curve of (2.2).

Although we are going to use this theorem later, we will not give a proof here. You might want to try to prove this yourself. If you struggle, it will help to consult [DK00, Lemma B.4].
LECTURE 3. First-order PDE

In this lecture I will explain one approach to first order quasi-linear PDE, called the method of characteristics. This approach can be extended to fully-nonlinear first order PDE. A beautiful exposition can be found in Arnold’s book [Arn04, Lectures 1 and 2].

Definition 3.1. A first order quasi-linear PDE is a PDE of the form

\[
  b(x, u(x)) \cdot \nabla u(x) + c(x, u(x)) u(x) = 0
\]

for a function \( u : U \to \mathbb{R} \) with \( U \subset \mathbb{R}^n \) an open set, \( b : U \times \mathbb{R}^n \to \mathbb{R}^n \) and \( c : U \times \mathbb{R}^n \to \mathbb{R} \) smooth.

Remark 3.3. Note that if \( b, c \) do not depend on the second variable, then (3.2) is a first order linear PDE.

Remark 3.4. We think of \( x \) as a “space variable”. Sometimes, however, we work on space-time. In these cases we replace \( \mathbb{R}^n \) above by \( \mathbb{R}^{n+1} \) and \( x \) by \( (t, x) \).

3.1 Change of coordinates

The method of characteristics is based on the observation that the notion of first order quasi-linear PDE is stable under coordinate change and the hope that one can find a good coordinate system in which (3.2) becomes very simple.

Proposition 3.5. Let \( \Psi : V \to U \) be a \( C^1 \)-diffeomorphism and \( u : U \to \mathbb{R} \). Define \( \tilde{u} : V \to \mathbb{R} \) by

\[
  \tilde{u}(x) := u(\Psi(x)).
\]

The function \( u \) satisfies (3.2) if and only if the function \( \tilde{u} \) satisfies the first order quasi-linear PDE

\[
  \tilde{b}(x, \tilde{u}(x)) \cdot \nabla \tilde{u}(x) + \tilde{c}(x, \tilde{u}(x)) \cdot \tilde{u}(x) = 0.
\]

with

\[
  \tilde{b}(x, y) := (d\Psi(x))^{-1} b(\Psi(x), y)
\]

\[
  \tilde{c}(x, y) := c(\Psi(x), y).
\]

Proof. By the chain rule

\[
  \nabla \tilde{u}(x) = d\Psi(x)^t \nabla u(\Psi(x)),
\]
or, equivalently,
\[ \nabla u(\Psi(x)) = (d\Psi(x)^{-1})^t \nabla \tilde{u}(x); \]
hence,
\[ b(\Psi(x), u(\Psi(x))) \cdot \nabla u(\Psi(x)) = \tilde{b}(x, \tilde{u}(x)) \cdot \nabla \tilde{u}(x). \]
Trivially, we also have
\[ c(\Psi(x), y)u(\Psi(x)) = \tilde{c}(x, y)\tilde{u}(x). \]

This completes the proof. \(\square\)

### 3.2 Warm-up: the transport equation

Let us now consider a simple example which will, however, play an important rôle in the theory.

Example 3.6. A transport equation is a PDE of the form
\[ v(x) \cdot \nabla u(x) = 0 \tag{3.7} \]
for a function \( u : U \to \mathbb{R} \) with \( U \subset \mathbb{R}^n \) open and \( v : U \to \mathbb{R}^n \) a vector field on \( U \).

The transport equation (3.7) is dual to ODE corresponding to \( v \) in the sense of the following proposition, and this is what makes transport equations easy—at least, theoretically.

Proposition 3.8. Suppose \( \phi : I \to U \) is a solution of the ODE
\[ \partial_s x(s) = v(x(s)). \tag{3.9} \]
If \( u : U \to \mathbb{R} \) is a solution of (3.7), then it is constant along \( \phi \), i.e.,
\[ \partial_s u(\phi(s)) = 0. \]

Remark 3.10. The reason we parametrise \( \phi \) by \( s \) and not \( t \) is that often in (3.2) one of the components of \( x \) is time and we want to reserve \( t \) for that purpose.

Proof. The proof is a simple computation using the chain-rule:
\[ \partial_s u(\phi(s)) = \partial_s \phi(s) \cdot \nabla u(\phi(s)) = v(\phi(s)) \cdot \nabla u(\phi(s)) = 0. \quad \square \]
Example 3.11. Suppose we want to solve transport equation on $\mathbb{R}^{n+1}$ with

$$v = \frac{\partial}{\partial t} + w$$

for some non-zero constant $w \in \mathbb{R}^n$. (Recall, that we use coordinates $(t, x_1, \ldots, x_n)$ on $\mathbb{R}^{n+1}$.)

The solutions of (3.9) are the straight lines

$$x(s) = (s_0, x_0) + s(1, w);$$

hence, the solutions to (3.7) must be functions of the form

$$u(t, x) := f(x - tw)$$

with $f : \mathbb{R}^n \to \mathbb{R}$ some function.

If we specify $u$ on the hypersurface $\{t = 0\}$, that is, we fix $u(0, \cdot) = f$, then there is a unique solution. Even if $f$ is not differentiable, the function $u(s, x) := f(x - sw)$ still satisfies (3.7), since $u$ still has a derivatives in the direction of the vector $v$ and this is all that is needed to make sense of (3.7).

In the light of the previous example we introduce the following notion which will come up over and over again in this class (it is a generalisation of the notion of boundary or initial condition).

Definition 3.12. Let $U \subset \mathbb{R}^n$ be an open subset and consider a PDE (1.2) for a function $u : U \to \mathbb{R}$. Let $\Gamma$ be a smooth oriented hypersurface with outward-pointing unit normal vector $\nu$, let $k \in \mathbb{N}_0$ and $(f_0, \ldots, f_k) : \Gamma \to \mathbb{R}^k$ be a $k$–tuple of functions. We say that $u : U \to \mathbb{R}$ solves the Cauchy problem for (1.2) with Cauchy hypersurface $\Gamma$ and Cauchy data $(f_0, \ldots, f_k)$ if

- the function $u$ solves the PDE (1.2), and
- we have

$$u|_\Gamma = f_0, \quad (\partial_\nu u)|_\Gamma = f_1, \quad \cdots \quad (\partial^k_\nu u)|_\Gamma = f_k.$$

Exercise 3.13. In the situation of Example 3.11, find the general form of a solution of the inhomogeneous transport equation

$$v \cdot \nabla u(t, x) = g(t, x).$$

Here we take $g : \mathbb{R}^{n+1} \to \mathbb{R}$ to be some continuous function. (Hint: What is the analogue of Proposition 3.8?)
3.3 The method of characteristics

Suppose that \( u : U \to \mathbb{R} \) is a solution of (3.2). Let \( I \to U, s \mapsto x(s) \) be a path in \( U \). Set \( y(s) := u(x(s)) \).

By the chain rule

\[
\frac{\partial}{\partial s} y(s) = \nabla u(x(s)) \cdot \frac{\partial}{\partial s} x(s).
\]

If \( x \) where such that

\[
\frac{\partial}{\partial s} x(s) = b(x, y(s)),
\]

then \( y \) solves the ODE

\[
\frac{\partial}{\partial s} y(s) = -c(x(s), y(s))y(s).
\]

**Definition 3.14.** The characteristic equation for (3.2) is the ODE

\[
\begin{align*}
\frac{\partial}{\partial s} x(s) &= b(x(s), y(s)) \\
\frac{\partial}{\partial s} y(s) &= -c(x(s), y(s))y(s).
\end{align*}
\]

(3.15)

If \( s \mapsto (x(s), y(s)) \) is a solution of (3.15), then we say \( x \) is the projected characteristic of (3.2).

**Remark 3.16.** If the coefficient \( b(x, u(x)) \) in (3.2) does not actually depend on \( u(x) \) (this is the case, e.g., for first order linear PDE), then (3.15) partially decouples and the ODE for \( x \) no longer involves \( y \). One consequence of this is that the projected characteristics for a such first-order PDE never intersect. This simplifies the problem of solving (3.2) quite a bit. In particular, some of the phenomena we are going to encounter for the inviscid Burgers’ equation cannot occur.

We will now consider the Cauchy problem with Cauchy hypersurface \( \Gamma \) and Cauchy data \( f : \Gamma \to \mathbb{R} \). The strategy now is to solve (3.15) with initial conditions of the form \( (x_0, f(x_0)) \) for \( x \in \Gamma \) and piece these together to get a solution of (3.2). We cannot always do this for two basic reasons:

- The projected characteristics may not fill out all of \( U \).
- The projected characteristics may intersect.

One way the projected characteristics could fail fill out all of \( U \) is if the projection of the characteristic through \( (x_0, f(x_0)) \) is tangent to \( \Gamma \) at \( x_0 \). If there are no such “bad” points on \( \Gamma \), then we can at least solve near \( \Gamma \).

**Definition 3.17.** A point \( x_0 \in \Gamma \) is called non-characteristic for (3.2) if

\[
\nu(x) \cdot b(x_0, f(x_0)) \neq 0.
\]
Theorem 3.18 (Local solvability). Suppose $x_0 \in \Gamma$ is non-characteristic. Then there exists a neighbourhood $V$ of $x_0 \in \Gamma$ and a unique solution $u : V \to \mathbb{R}$ of (3.2) with Cauchy data $f$ on $\Gamma \cap V$.

Proof. We can choose coordinates $(s, x_1, \ldots, x_{n-1})$ near $x_0$ such that $\Gamma$ locally is cut out by $s = 0$ and $x_0$ has coordinates $(0, \ldots, 0)$. By abuse of notation we still denote the coefficients of (3.2) by $b$ and $c$. The normal vector-field $\nu$ is now nothing but $\partial_s$ and we still have $\nu \cdot b(x_0, f(x_0)) \neq 0$.

For a sufficiently small neighbourhood $U_0$ of $0 \in \Gamma$ and $I = (-\delta, \delta)$, we can find $(\Phi, \Upsilon) : I \times U_0 \to U \times \mathbb{R}$ such that

$$\Phi(0, x) = x \quad \text{and} \quad \Upsilon(0, x) = f(x)$$

and

$${\partial}_s \Phi(s, x) = b(\Phi(s, x), \Upsilon(s, x)) \quad \text{and} \quad {\partial}_s \Upsilon(s, x) = -c(\Phi(s, x), \Upsilon(s, x)) \Upsilon(s, x).$$

For $x \in U_0$,

$$d\Phi(0, x) = \begin{pmatrix} b(x, f(x)) & 0 \\ \text{id}_{T_x \Gamma} \end{pmatrix}$$

with the top-left entry being $\nu \cdot b((0, x), f(0, x))$. Since $x_0$ is non-characteristic, we can assume that $\nu \cdot b((0, x), f(0, x)) \neq 0$ for all $x \in U_0$. Hence, $\Phi : I \times U_0 \to U$ is a diffeomorphism near $x_0$. Denote its image by $V$.

A function $u : V \to \mathbb{R}$ solves (3.2) with Cauchy data $f$ on $V_0 = V \cap \Gamma$ if and only the function $\tilde{u} : I \times U_0 \to \mathbb{R}$ defined by

$$\tilde{u}(s, x) := u(\Phi(s, x))$$

solves the equation

$${\partial}_s \tilde{u}(s, x) = -c(\Phi(s, x), \tilde{u}(s, x)) \tilde{u}(s, x)$$

and $\tilde{u}(0, x) = f(x)$. For this PDE the asserted statement is clear.

Equivalently, the computation preceding this theorem asserts that if a solution $u$ exists then it must be of the form

$$u(s, x) = \Upsilon(\Phi^{-1}(s, x));$$

moreover, this formula also does define a solution as a consequence of the computation in the proof of Proposition 3.5. \qed
Although, the previous theorem guarantees the existence of unique local solutions, the question of global solvability quite involved. Many issues can already arise in the linear case.

**Example 3.19.** Suppose \( v = r \partial_r = \sum_{i=1}^{n} x_i \partial_i \) and \( \Gamma = \partial B_1(0) \). Then the Cauchy problem of (3.7) with Cauchy data \( f : \Gamma \to \mathbb{R} \) has a solution on all of \( \bar{B}_1(0) \) if and only if \( f \) is constant.

**Example 3.20.** Suppose \( v = \partial_1 \) and \( \Gamma = \{0\} \times \mathbb{R}^{n-1} \cup \{1\} \times \mathbb{R}^{n-1} \). Then the Cauchy problem of (3.7) with Cauchy data \( f : \Gamma \to \mathbb{R} \) has a solution on all of \([0, 1] \times \mathbb{R}^{n-1}\) if and only if \( f(0, \cdot) = f(1, \cdot) \).
LECTURE 4. The inviscid Burgers’ equation

We will now see another example where serious issues arise due to non-linearities.

Definition 4.1. The inviscid Burgers’ equation is the following PDE for a function $u: [0, \infty) \times \mathbb{R} \to \mathbb{R}$

(4.2) \[ \partial_t u + u \partial_x u = 0. \]

Remark 4.3. You might want to think of this is a transport equation where the speed depends on $u$ itself.

Remark 4.4 (Burgers’ equation and Navier–Stokes). The incompressible Navier–Stokes equation is the PDE

(4.5) \[ \partial_t u + (u \cdot \nabla) u + \nu \Delta u = -\nabla p \]
\[ \text{div } u = 0. \]

for a pair of maps $u: [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$, the flow velocity, and $p: [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$, the pressure. Here $\nu$ is a constant. This is a model for the motion of an incompressible viscous fluid.

The second equation in (4.5) is the continuity equation and asserts that no material is created by the flow. If we want to force $\nabla p = 0$, then we can no longer expect this second equation to hold. In dimension one, the equation we arrive at by setting $\nabla p = 0$ and dropping the second equation is Burger’s equation

\[ \partial_t u + u \partial_x u - \nu \partial_x^2 u = 0. \]

Setting $\nu = 0$ we obtain (4.2).

Roughly speaking, in (4.5), $(u \cdot \nabla) u$ is a bad term, while the viscosity term $\nu \Delta u$ is good.

Remark 4.6. The inviscid Burgers’ equation can be written as

\[ \partial_t u + \partial_x \left( \frac{u^2}{2} \right) = 0; \]

Here $u \cdot \nabla = \sum_{i=1}^n u_i \partial_i$. The whole expression $\partial_t + (u \cdot \nabla)$ is sometimes called the material derivative. It is the time derivative from the perspective of a particle moving along the flow.

The role of the pressure is somewhat secondary since it can be recovered from the flow velocity using the equation

\[ \Delta p = -\text{div}((u \cdot \nabla) u + \Delta u) \]

and the boundary conditions.
hence, it is a special case of a conservation law, i.e., a PDE of the form
\[ \partial_t u(t, x) + \partial_x F(u(t, x)) = 0 \]
for some function \( F \).

We will be concerned with the Cauchy problem for (4.2) with Cauchy data \( f : \mathbb{R} \to \mathbb{R} \) prescribed along the Cauchy hypersurface \( \{0\} \times \mathbb{R} \), that is, we prescribe
\[ u(0, x) = f(x). \]

Exercise 4.7 (Conservation of the spatial \( L^2 \)-norm). Suppose \( u : [0, T) \times \mathbb{R} \to \mathbb{R} \) is a solution of (4.2) and
\[ \int_{\mathbb{R}} |u(0, x)|^2 \, dx = \int_{\mathbb{R}} |f(x)|^2 \, dx < \infty. \]
Prove that the spatial \( L^2 \)-norm is conserved, that is, for all for all \( t \in [0, T) \) we have
\[ \int_{\mathbb{R}} |u(t, x)|^2 \, dx = \int_{\mathbb{R}} |u(0, x)|^2 \, dx. \]

Let us now apply the method of characteristics to the inviscid Burgers’ equation. The characteristic equation takes the particularly simple form
\[
\begin{align*}
\partial_s t(s) &= 1, \\
\partial_s x(s) &= y(s), \\
\partial_s y(s) &= 0.
\end{align*}
\]
Moreover, note that
\[ \partial^2_s x(s) = \partial_s y(s) = 0. \]
Thus for the projected characteristic \((t(s), x(s))\) emanating from \((0, x_0)\) we have
\[ \partial_s x(s) = f(x_0). \]
We conclude that the projected characteristics are of the form
\[ t(s) = s \quad \text{and} \quad x(s) = x_0 + f(x_0)s. \]

From this we can easily compute the solution to (4.2) in concrete examples.

\[ ^4 \text{The name conservation law comes from the fact that solutions have the property } \partial_t \int_{\mathbb{R}} u(t, x) \, dx = 0 \text{ (subject to some technical hypotheses of course).} \]
Example 4.8. Suppose $f(x) = x$. Then the characteristic curves are
\[ t(s) = s \quad \text{and} \quad x(s) = x_0(1 + s). \]
The projected characteristics never intersect (in positive time); hence, we can construct a global solution to (4.2): $u : [0, \infty \times \mathbb{R} \rightarrow \mathbb{R}$ defined by
\[ u(t, x) := \frac{x}{1 + t}. \]

Example 4.9. Let’s slightly change the previous example. Suppose $f(x) = -x$. Then the characteristic curves are
\[ t(s) = s \quad \text{and} \quad x(s) = x_0(1 - s). \]
All the projected characteristics intersect at $s = 1$. Thus (unless $f$ is constant) we can only solve (4.2) up to $s = 1$ by $u : [0, 1) \times \mathbb{R} \rightarrow \mathbb{R}$ defined by
\[ u(t, x) = \frac{x}{t - 1}. \]

Exercise 4.10. Find an example of Cauchy data for which no solution $u : [0, \varepsilon) \times \mathbb{R} \rightarrow \mathbb{R}$ exists for any value of $\varepsilon > 0$.

Example 4.11. Suppose
\[ f(x) = \begin{cases} 
1 & x \leq 0 \\
1 - x & x \in [0, 1] \\
0 & x \geq 1. 
\end{cases} \]
(If the fact that $f$ is not smooth, just piece-wise linear, bothers you can smooth out the kinks. This doesn’t drastically change what is going to happen, but makes the analysis more cumbersome.) The projected characteristics are
\[ s \mapsto \begin{cases} 
(s, x + s) & x \leq 0 \\
(s, x + (1 - x)s) & x \in [0, 1] \\
(s, x) & x \geq 1. 
\end{cases} \]
For $t \geq 1$, these start to intersect.

So initially we only get the “solution” $u : [0, 1) \times \mathbb{R} \rightarrow \mathbb{R}$ defined by
\[ u(t, x) := \begin{cases} 
1 & x \leq t \\
\frac{1 - x}{1 - t} & x \in [t, 1] \\
0 & x \geq 1. 
\end{cases} \]
(4.12)

I wrote “solution” because $u(t, x)$ is not $C^1$. 27
Exercise 4.13. Suppose $\phi \in C^\infty_c((0, 1) \times \mathbb{R})$ is a smooth function on $(0, 1) \times \mathbb{R}$ with compact support. We call such a $\phi$ a test function. If $u$ is a $C^1$-solution of (4.2), then integration by parts shows that

$$\int_{(0, 1) \times \mathbb{R}} (\partial_t \phi) u + (\partial_x \phi) \frac{u^2}{2} \, dt \, dx = 0.$$  

Show that this is also true for (4.12).

Remark 4.15. We say that $u$ is a weak solution of the inviscid Burgers’ equation if (4.14) holds for all test functions. Weak solutions play a very important role in PDE. One is often told to first find a weak solution (this is surprisingly often possible using functional analysis methods) and then prove using regularity theory that the solution is smooth.

Sometimes one can only prove a certain amount of regularity for solutions, like in the above example.

Another perspective on weak solutions is that in general asking for a smooth solution is too much because the actual values of $u(t, x)$ are not physically accessible (= measurable with experiments), while quantities like

$$\int \phi u, \int (\partial_x \phi) u^2, \ldots$$

are—at least in principle. From this standpoint, (4.14) appears quite natural.

Remark 4.16. The solution $u$ can be extended to $t \geq 1$ by

$$u(t, x) = \begin{cases} 1 & x \leq \frac{1+t}{2} \\ 0 & x > \frac{1+t}{2} \end{cases}$$

as a weak solution.

Whenever singularities in a PDE can occur, it is interesting and important to understand when they will occur. For the inviscid Burger’s equation the answer is rather simple.

Theorem 4.18 (Criterion for singularity formation). Let $f \in C^1(\mathbb{R})$. There is a $C^1$ solution $u: [0, \infty) \times \mathbb{R} \to \mathbb{R}$ to the Cauchy problem for (4.2) with Cauchy data $f$ if and only if $f'(x) \geq 0$.

Proof. Suppose $f'(x) < 0$ for some $x \in \mathbb{R}$. Then we can find points $x_0, x_1 \in \mathbb{R}$ with $x_0 < x_1$ and $f(x_0) > f(x_1)$. 

28
The characteristic curves emanating from these points are

\[ \xi_0(s) = (s, x_0 + f(x_0)s) \quad \text{and} \quad \xi_1(s) = (s, x_1 + f(x_1)s). \]

Since

\[ x_0 + f(x_0)s = x_1 + f(x_1)s \iff x_0 - x_1 = (f(x_1) - f(x_0))s \]

has a solution for some \( s = s_0 > 0 \), \( \xi_1 \) and \( \xi_2 \) intersect which prohibits a \( C^1 \) solution from existing up until time \( s_0 \).

Conversely, let us show that a global solution exists if \( f'(x) \geq 0 \). We can define a map \( \Phi : [0, \infty) \times \mathbb{R} \to [0, \infty) \times \mathbb{R} \) by

\[ \Phi(s, x) := (s, x + f(x)s) \]

This is a \( C^1 \)-map and its derivative is

\[ d\Phi = \begin{pmatrix} 1 & 0 \\ f(x) & 1 + f'(x)s \end{pmatrix}. \]

Since

\[ \det d\Phi = 1 + f'(x)s > 0 \]

for all \( s \geq 0 \), \( \Phi \) is a diffeomorphism (initially locally, but one can see that it is a global diffeomorphism). Thus the method of characteristics provides the desired solution. \( \square \)
Lecture 5. Introduction to the Heat Equation

The next three or four lectures will be concerned with the study of the heat equation.

**Definition 5.1.** The \textit{heat equation} is the PDE

\begin{equation}
\partial_t u(t,x) + \Delta u(t,x) = 0
\end{equation}

for a function \( u : [0, T) \times U \to \mathbb{R} \) with \( U \subset \mathbb{R}^n \) and \( T > 0 \).

Often one also has to deal with the following inhomogeneous version of (5.2).

**Definition 5.3.** Let \( U \subset \mathbb{R}^n, T > 0 \) and \( \sigma : [0, T) \times U \to \mathbb{R} \). The \textit{heat equation with source term} \( \sigma \) is the PDE

\begin{equation}
\partial_t u(t,x) + \Delta u(t,x) = \sigma(t,x).
\end{equation}

for a function \( u : [0, T) \times U \to \mathbb{R} \).

**Remark 5.5.** Of course, \( \Delta \) denotes the Laplace operator (or the Laplacian, for short). It acts on the spacial variables only:

\[ \Delta u(t,x) = -\sum_{i=1}^{n} \partial^2_{x_i} u(t,x). \]

I prefer to put the minus sign in \( \Delta \), because it makes the operator \( \Delta \) positive. Also, various natural Laplace type operators in geometry come with a minus sign. \( \Delta \) as defined here is often called the geometer’s Laplacian. (The other sign convention is popular among analysts; hence, the so defined Laplacian is often called the analysts’ Laplacian.)

**Remark 5.6.** The heat equation models the time-evolution of temperature distributions. Since this is not a physics class, I will refrain from discussing how exactly one arrives at the heat equation, but here is a sketch: Consider a small space region \( U \) and denote by \( E_U(t) \) the amount of thermal energy contained in \( U \). If \( u(t,x) \) is the temperature distribution, then in suitable units

\[ E_U(t) = \int_U u(t,x). \]

On the one hand we have, some what trivially,

\[ \partial_t E_U = \int_U \partial_t u(t,x); \]
on the other hand if $q$ is the heat flux vector, then by the divergence theorem

$$\partial_t E_U = -\int_{\partial U} \langle q, \nu \rangle = -\int_U \text{div} \, q.$$ 

Now by *Fourier’s law of heat conduction*:

$$q(t, x) = -\nabla u(t, x);$$

hence,

$$\int_U \partial_t u(t, x) = \partial_t E_U = -\int_U \Delta u.$$ 

Since this is supposed to hold for every space region $U$, $u$ better be a solution of (5.2).

Observe that there are two key assumptions here: (1) there is such a thing as the heat flux vector and (2) the heat flux vector is given by Fourier’s law. One can summarise this colloquially as: $u$ diffuses.

The heat equation is a constant coefficient linear equation and consequently it is invariant under spacetime-translations. It has a further interesting symmetry.

**Proposition 5.7** (Parabolic rescaling). If $u$ solves (5.2), then for each $\lambda > 0$ so does

$$u_\lambda(t, x) := u(\lambda^2 t, \lambda x).$$

**Exercise 5.8.** Prove Proposition 5.7!

### 5.1 Boundary conditions

We will usually be interested in solving the heat equation (5.2) subject to an *initial condition*

(5.9) \hspace{1cm} u(0, x) = f(x) \quad \text{for all } x \in U.

If $U$ is a bounded domain in $\mathbb{R}^n$ with boundary $\partial U$, one typically imposes one of the following boundary conditions.

**Definition 5.10.** Given $g: \partial U \to \mathbb{R}$, we consider the following type of boundary conditions:

- *Dirichlet boundary conditions*

(5.11) \hspace{1cm} u(t, x) = g(x) \quad \text{for all } (t, x) \in (0, T] \times \partial U,
• Neumann boundary conditions

\[ \partial_\nu u(t, x) = g(x) \quad \text{for all } (t, x) \in (0, T] \times \partial U, \tag{5.12} \]

• and, given also \( \alpha > 0 \), Robin boundary conditions

\[ \partial_\nu u(t, x) + \alpha u(t, x) = g(x) \quad \text{for all } (t, x) \in (0, T] \times \partial U. \tag{5.13} \]

If \( \partial U \) is disconnected, we can impose mixed boundary conditions, that is, we impose one of the above boundary conditions on each component of \( \partial U \).

\[ \Delta \text{ Remark 5.14. Mostly we will consider homogeneous boundary conditions, that is, } f = 0. \text{ (The general case can easily be reduced to this case.) If I write some thing like “we impose Dirichlet boundary conditions” and make no mention of } f \text{ at all, then } f = 0. \]

The Dirichlet, Neumann, Robin, or mixed initial-boundary value problem for the heat equation is the is the condition for \( u \) to solve (5.2), (5.9) and either (5.11), (5.12), (5.13) or a mixed boundary condition. It turns out that each of these problems is a “good” (that is well-posed) problem; in particular, they have (essentially) unique solutions.

\[ \text{Remark 5.15 (Duhamel’s principle). The problem of solving (5.4) can be reduced to the homogeneous equation via Duhamel’s principle.} \]

Suppose we want to solve

\[ \partial_t u + \Delta u = \sigma \]

with Dirichlet boundary data \( g \) and initial data \( f \). At the expense of changing \( \sigma \), we can assume that \( g = 0 \). Moreover, by subtracting the solution to the Dirichlet initial-boundary value problem with initial data \( f \), we can also assume that \( f = 0 \).

Let \( u_\tau \) denote the solution to (5.2) with Dirichlet boundary conditions and initial condition \( \sigma(\tau, \cdot) \), and define

\[ u(t, x) = \int_0^t u_\tau(t - \tau, x) \, d\tau. \]

This function clearly satisfies the initial and boundary conditions; moreover,

\[ \begin{align*}
(\partial_t + \Delta) u(t, x) &= (\partial_t + \Delta) \int_0^t u_\tau(t - s, x) \, ds \\
&= u_t(0, x) + \int_0^t (\partial_t + \Delta) u_\tau(t - \tau, x) \, d\tau \\
&= \sigma(t, x).
\end{align*} \]
5.2 A toy model

Initially very formally, although a lot of this can be made rigorous, we can think of $u$ as a map $U : [0, T] \rightarrow C^2(U)$, i.e., as a path in the infinite dimensional vector space $C^2(U, \mathbb{R})$ and think of (5.2) as an ODE. As a toy model replace $C^2(U)$ by a finite dimensional vector space $\mathbb{R}^n$ and $\Delta$ by a matrix $A \in \mathbb{R}^{n \times n}$. The ODE

$$\partial_t x(t) + Ax(t) = 0$$

can be solved very easily. One extremely concise version of writing the solution is as

$$(5.16) \quad x(t) = e^{-At}x(0).$$

If we knew more about $A$, then we could also say more about the trajectories defined by (5.16).

**Proposition 5.17.** If $f, g \in C^2(\bar{U})$ satisfy

- $f|_{\partial U} = g|_{\partial U} = 0,$
- $\partial_{\nu} f|_{\partial U} = \partial_{\nu} g|_{\partial U} = 0 \text{ or}$
- $\partial_{\nu} f|_{\partial U} + \alpha f|_{\partial U} = \partial_{\nu} g|_{\partial U} + \alpha g|_{\partial U} = 0 \text{ with } \alpha > 0,$

then

$$\int_{\bar{U}} (\Delta f) g = \int_{\bar{U}} f(\Delta g) \quad \text{and} \quad \int_{\bar{U}} (\Delta f) f \geq 0.$$  

**Proof.** The first assertion is a consequence of Theorem A.4 provided that

$$\int_{\partial U} f(\partial_{\nu} g) - (\partial_{\nu} f) g = 0.$$

Given either of the first two boundary conditions this clearly holds. For the last boundary condition the integrand is equal to $\alpha (fg - fg) = 0$.

The second assertion too is a consequence of Theorem A.4 provided that

$$\int_{\partial U} -(\partial_{\nu} f) f \geq 0.$$  

Again, given either of the first two boundary conditions this clearly holds. For the last boundary condition the integrand is equal to $\alpha f^2 \geq 0$. This is where $\alpha > 0$ comes in.
This justifies specialising to the case where \( A \) (our model for \( \Delta \)) is a symmetric matrix with non-negative spectrum. In this case we can understand (5.16) more concretely. From linear algebra we know that there is an orthonormal basis \((e_i)\) of \(\mathbb{R}^n\) consisting of eigenvectors. Denote the corresponding eigenvalues by \(\lambda_i \geq 0\). We can write (5.16) as

\[
\sum_{i=1}^{n} e^{-\lambda_i t} \langle e_i, x(0) \rangle \cdot e_i,
\]

In particular, every solution decays (fast) as \( t \to \infty \), except for the potential zero eigenvalues which stay put. For the heat equation this corresponds to everything decaying to zero temperature (for Dirichlet and Robin boundary conditions) or a finite temperature (for Neumann boundary conditions).

To make at least some of the preceding discussion rigorous one needs the spectral theory of unbounded operators on Banach spaces, which is a somewhat advanced topic in functional analysis. On the interval \([0,1]\), the spectral theory of the Laplace operator is very simple and is completely captured by the theory of Fourier series. In the next section we will give a very satisfactory answer the heat equation on \([0,1]\) under Dirichlet boundary conditions.

### 5.3 Solution of the heat equation with Dirichlet boundary conditions in dimension one

**Theorem 5.19.** Let \( f \in L^2([0,1]) \). There is a unique function \( u : (0, \infty) \times [0,1] \to \mathbb{R} \) which is smooth, solves (5.2), satisfies the Dirichlet boundary conditions

\[
u(t, 0) = u(t, 1) = 0 \text{ for all } t > 0,
\]

and the initial condition \( f \) in the sense that

\[
\lim_{t \to 0} u(t, \cdot) \to f \text{ in } L^2([0,1]).
\]

**Proof.** We proceed in four steps.

**Step 1.** Suppose there is such a function \( u \). Define a continuous map \( U : [0, \infty) \to L^2([0,1]) \) by

\[
U(t) := \begin{cases} 
 u(t, \cdot) & t > 0, \\
 f & t = 0.
 \end{cases}
\]

Then, for each \( t \in [0, \infty) \),

\[
U(t) = \sum_{n=1}^{\infty} e^{-(n\pi)^2 t} \langle f_n, f \rangle f_n \text{ in } L^2.
\]
For \( n \in \mathbb{N} \), define continuous functions \( a_n : [0, \infty) \to \mathbb{R} \) by
\[
a_n(t) = \langle U(t), f_n \rangle.
\]
Since \( u \) solves (5.2), for \( t > 0 \)
\[
\partial_t a_n + (n\pi)^2 a_n = 0.
\]
Hence, for \( t > 0 \)
\[
a_n(t) = c_n e^{-(n\pi)^2 t}
\]
for some constant \( c_n \). Since \( a_n \) is continuous, \( c_n = a_n(0) = \langle f_n, f \rangle \). This proves the asserted identity.

**Step 2.** There is at most one function \( u \) with the asserted properties.

The assignment \( u \mapsto U \) is injective and by the previous step \( U \) is uniquely determined by \( f \).

**Step 3.** For each \( k \in \mathbb{N}_0 \) and \( T_0 > 0 \), the series
\[
(5.20) \quad u(t, x) := \sum_{n=1}^{\infty} e^{-(n\pi)^2 t} \langle f_n, f \rangle f_n(x)
\]
converges in \( C^k([T_0, \infty) \times [0, 1]) \). In particular, (5.20) defines a smooth function \( u : (0, \infty) \times [0, 1] \to \mathbb{R} \).

This follows from the same kind of argument as Proposition C.10 and the observation that for \( \ell \geq 0 \) and \( t \geq T_0 > 0 \)
\[
\sum_{n=1}^{\infty} n^\ell e^{-n^2 t} |\langle f_n, f \rangle| \leq \left( \sum_{n=1}^{\infty} n^{2\ell} e^{-2n^2 T_0} \right)^{1/2} \left( \sum_{n=1}^{\infty} |\langle f_n, f \rangle|^2 \right)^{1/2} < \infty.
\]

**Step 4.** The function \( u \) has the asserted properties.

Each of the functions \((t, x) \mapsto e^{-n^2 t} \langle f_n, f \rangle f_n(x)\) solves (5.2) and satisfies the Dirichlet boundary conditions. The assertion of about the initial value is obvious.

\( \square \)

**Remark 5.21.** The formula for \( u \) can be written more verbosely as
\[
u(t, x) = 2 \sum_{n=1}^{\infty} e^{-n^2 t} \left( \int_0^1 \sin(n\pi x) f(x) \, dx \right) \sin(n\pi x).
\]
Exercise 5.22. Prove that if the sequence \( \langle f_n, f \rangle \) is summable, i.e.,
\[
\sum_{n=1}^{\infty} |\langle f_n, f \rangle| < \infty,
\]
then \( u \) extends to a continuous function \( u : [0, \infty) \times [0, 1] \to \mathbb{R} \) satisfying the initial value
\[
u(t, \cdot) = f.
\]

Exercise 5.23. Prove the analogue of Theorem 5.19 for Neumann boundary conditions. (Hint: The key is to find a suitable analogue of Theorem C.8.)

Exercise 5.24. Suppose \( \sigma \in C^\infty([0, \infty) \times [0, 1]) \). Give a formula for the solution of (5.4) with homogeneous Dirichlet boundary conditions and initial condition \( f = 0 \). Hint: Use Duhamel’s principle.

Exercise 5.25. Define \( f \in L^2([0, 1]) \) by
\[
f(x) = \begin{cases} x & x \leq 1/2 \\ 1/2 - x & x \geq 1/2. \end{cases}
\]
Find a smooth function \( u : (0, \infty) \times [0, 1] \to \mathbb{R} \) which solves (5.2), satisfies the Dirichlet boundary conditions
\[
u(t, 0) = u(t, 1) = 0 \quad \text{for all } t > 0,
\]
and with initial condition \( f \) in the sense that
\[
\lim_{t \to 0} u(t, \cdot) = f \text{ in } L^2([0, 1]).
\]

Exercise 5.26. Is there a \( f \in L^2([0, 1]) \) such that if \( u : (0, \infty) \times [0, 1] \) is as in Theorem 5.19, then the Fourier series of \( u(1, x) \) is
\[
u(1, x) = \sum_{n=1}^{\infty} \frac{f_n}{n^2}?
\]
LECTURE 6. Uniqueness for the Heat Equation

In this lecture we will address the uniqueness question for the heat equation in two separate ways: via the energy method and via the weak maximum principle.

To properly state the main theorems, we need to make some definitions. Let $U$ be a bounded open subset of $\mathbb{R}^n$ and $T > 0$. Set

$$U_T := (0, T] \times U, \quad \bar{U}_T := [0, T] \times \bar{U}, \quad \text{and} \quad \partial_p U_T := \bar{U}_T \setminus U_T.$$ 

Moreover, we write $C^{k,\ell}(\bar{U}_T)$ for the space of continuous functions $f: \bar{U}_T \to \mathbb{R}$ which are $k$–times continuously differentiable in the $t$–direction and $\ell$–times continuously differentiable in the $x$–direction.

**Theorem 6.1** (Uniqueness). Given functions $\sigma: U_T \to \mathbb{R}$, $u_0: \bar{U} \to \mathbb{R}$ and a choice of either Dirichlet, Neumann, Robin or mixed boundary conditions, there exists at most one $u \in C^{1,2}(\bar{U}_T)$ such that

$$\partial_t u + \Delta u = \sigma \text{ in } U_T$$

with

$$u(0, \cdot) = u_0$$

and satisfying the chosen boundary conditions.

**Theorem 6.2** (Backwards uniqueness). Given functions $\sigma: U_T \to \mathbb{R}$, $u_T: \bar{U} \to \mathbb{R}$ and a choice of either Dirichlet, Neumann, Robin or mixed boundary conditions, there exists at most one $u \in C^\infty(\bar{U}_T)$ such that

$$\partial_t u + \Delta u = \sigma \text{ in } U_T$$

with

$$u(T, \cdot) = u_T$$

and satisfying the chosen boundary conditions.

In both of these cases, if there were two solutions $u_1$ and $u_2$ we could subtract them to obtain a solution $v := u_1 - u_2$ of

$$\partial_t v + \Delta v = 0 \text{ in } U_T$$

satisfying either Dirichlet, Neumann, Robin or mixed boundary conditions. What we need to prove is that if either $v(0, \cdot) = 0$ or $v(T, \cdot) = 0$, then $v = 0$. 

37
6.1 Energy method

Proof of Theorem 6.1. Suppose \( v \) is as above with \( v(0, \cdot) = 0 \). Define an energy function by

\[
E(t) := \int_U |v(t, x)|^2 \, dx.
\]

We know that \( E(0) = 0 \).

We compute

\[
\dot{E}(t) = \partial_t \int_U |v(t, x)|^2 \, dx
= 2 \int_U \langle \partial_t v(t, x), v(t, x) \rangle \, dx
= -2 \int_U \langle \Delta v(t, x), v(t, x) \rangle \, dx \leq 0
\]

by Proposition 5.17. Since \( E(t) \geq 0 \), it follows that \( E(t) = 0 \) for all \( t \). Thus \( v = 0 \).

Proof of Theorem 6.2. Suppose \( v \) is as above with \( v(T, \cdot) = 0 \). We define the energy as before. We now know that \( E(T) = 0 \). Since \( \dot{E} \leq 0 \), we know that \( E(t) \geq 0 \) for \( t \in [0, T) \). There is no loss in assuming that, in fact, \( E(t) > 0 \) for \( t \in [0, T) \).

We compute, using Proposition 5.17,

\[
\ddot{E}(t) = -4 \int_U \langle \Delta v(t, x), \partial_t v(t, x) \rangle \, dx
= 4 \int_U |\Delta v(t, x)|^2 \, dx \geq 0.
\]

So we know that \( E \) is a decreasing convex function, but this is not good enough. However, we can do better: by using Cauchy–Schwarz

(6.3) \[ \dot{E}^2 \leq E \ddot{E}. \]

The function \( \log E : [0, T) \to \mathbb{R} \) satisfies

\[
\lim_{t \to T} \log E(t) = -\infty.
\]

But (6.3) means that

\[
\partial_t^2 \log E = \frac{\dot{E}^2}{E^2} - \frac{E \ddot{E}}{E^2} = \frac{E \dddot{E} - \dot{E}^2}{E^2} \geq 0.
\]

Thus \( \log E \) is a convex function which tends to \( -\infty \) as \( t \to T \). This is impossible. \( \Box \)
Exercise 6.4. If \( f : [0, T) \to \mathbb{R} \) is a \( C^2 \)-function with \( f'' \geq 0 \), then
\[
 f(t) \geq f(0) + f'(0)t.
\]

6.2 Weak maximum principle

Theorem 6.5. Let \( u \in C^{1,2}(\bar{U}_T) \) be a subsolution of the heat equation, i.e.,
\[
 \partial_t u + \Delta u \leq 0.
\]
Then \( u \) attains its maximum in \( \bar{U}_T \) on \( \partial p U_T \), i.e.,
\[
 \max_{\bar{U}_T} u = \max_{\partial p U_T} u.
\]

Note that this gives another proof of Theorem 6.1. The function \( v \) satisfies the hypothesis of Theorem 6.5 and \( v|_{\partial p U_T} = 0 \) thus \( v \leq 0 \), and the same holds for \( -v \); hence, \( v = 0 \).

Proof of Theorem 6.5.

Step 1. The assertion holds if \( \partial_t u + \Delta u < 0 \).

If the maximum is attained at a point \((t_0, x_0)\) in \( U_T \), then \( \partial_t u(t_0, x_0) \geq 0 \)—in fact, the case \( > 0 \) is possible only if \( t_0 = T \). It follows that
\[
 \Delta u(t_0, x_0) < 0.
\]
However, \( x_0 \in U \) is also a local maximum for \( u(t_0, \cdot) \); hence, the Hessian \( \text{Hess}(u) = (\partial_i \partial_j u) \) is negative semi-definite; thus, \( \Delta u = -\text{tr Hess}(u) \geq 0 \). This is a contradiction.

Step 2. We prove the theorem.

For each \( \varepsilon > 0 \), define a new function \( u_\varepsilon \in C^{1,2}(\bar{U}_T) \) by
\[
 u_\varepsilon(t, x) := u(t, x) - \varepsilon t.
\]
Since
\[
 (\partial_t + \Delta) u_\varepsilon < 0,
\]
the previous step shows that
\[
 \max_{\bar{U}_T} u_\varepsilon = \max_{\partial p U_T} u_\varepsilon.
\]
Because
\[
 u_\varepsilon \leq u \leq u_\varepsilon + \varepsilon T,
\]
39
we also have
\[
\max_{\partial_p U_T} u - \varepsilon T \leq \max_{U_T} u - \varepsilon T \leq \max_{U_T} u_{\varepsilon} \leq \max_{\partial_p U_T} u_{\varepsilon} \leq \max_{\partial_p U_T} u.
\]
Since this is true for all \( \varepsilon > 0 \), we have
\[
\max_{U_T} u = \max_{\partial_p U_T} u.
\]

**Proposition 6.6 (Comparison).** If \( u, v \in C^{1,2}(\bar{U}_T) \) satisfy
\[
\partial_t u + \Delta u = \sigma \quad \text{and} \quad \partial_t v + \Delta v = \tau,
\]
and
\[
u \leq v \quad \text{on} \quad \partial U_T \quad \text{and} \quad \sigma \leq \tau,
\]
then
\[
u \leq v \quad \text{in} \quad \bar{U}_T.
\]

**Proof.** This follows from a direct application of Theorem 6.5 to \( u - v \). \( \square \)

**Proposition 6.7 (Stability estimate).** If \( u, v \in C^{1,2}(\bar{U}_T) \) satisfy
\[
\partial_t u + \Delta u = \sigma \quad \text{and} \quad \partial_t v + \Delta v = \tau,
\]
then
\[
\max_{U_T} |u - v| \leq \max_{\partial_p U_T} |u - v| + T \max_{U_T} |\sigma - \tau|.
\]

**Proof.** Set \( w := u - v \) and \( M := \max_{U_T} |\sigma - \tau| \). Both \( w - tM \) and \( -w - tM \) are subsolutions of the heat equation. Apply Theorem 6.5 to both to obtain
\[
\max_{U_T} w - tM \leq \max_{\partial_p U_T} w - tM \leq \max_{\partial_p U_T} w
\]
and
\[
\max_{U_T} -w - tM \leq \max_{\partial_p U_T} -w - tM \leq \max_{\partial_p U_T} -w.
\]
These two inequalities are equivalent to the assertion. \( \square \)
LECTURE 7. The Heat Kernel on $\mathbb{R}^n$

In today’s lecture we study the heat kernel on $\mathbb{R}^n$, which is in some sense the universal solution of the heat equation and thus also often called the fundamental solution.

**Definition 7.1.** The heat kernel on $\mathbb{R}^n$ is the function $\Phi : (0, \infty) \times \mathbb{R}^n \to [0, \infty)$ defined by

\[
\Phi(t, x) := \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}.
\]

The importance of the heat kernel stems from the following fact.

**Theorem 7.3.** Suppose $f \in C^0(\mathbb{R}^n)$ is bounded. Define $u : (0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ by

\[
u(t, x) := \int_{\mathbb{R}^n} \Phi(t, x - y) f(y) \, dy.
\]

Then the function $u$ is smooth, satisfies $\partial_t u + \Delta u = 0$, and, for each $x \in \mathbb{R}^n$,

\[
\lim_{t \downarrow 0} u(t, x) = f(x).
\]

Before we can proof this, we need to verify the following two propositions.

**Proposition 7.4.** $\Phi$ solves the heat equation, i.e.,

\[
\partial_t \Phi + \Delta \Phi = 0.
\]

**Exercise 7.5.** Prove Proposition 7.4.

**Proposition 7.6.** For each $t > 0$,

\[
\int_{\mathbb{R}^n} \Phi(t, x) \, dx = 1.
\]

**Proof.** In the coordinates $\xi = x/2\sqrt{t}$ the integral becomes

\[
\frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} e^{-|\xi|^2} \, d\xi = \prod_{i=1}^{n} \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-\xi_i^2} \, d\xi_i.
\]

To compute the integral in the product note that

\[
\left( \int_{\mathbb{R}} e^{-s^2} \, ds \right)^2 = \int_{\mathbb{R}^2} e^{-|z|^2} \, dz = \int_{0}^{2\pi} \int_{0}^{\infty} re^{-r^2} \, dr \, d\theta = 2\pi \int_{0}^{\infty} re^{-r^2} \, dr = \pi.
\]

The last step uses

\[
d(e^{-r^2}) = -2re^{-r^2}.
\]
Proof of Theorem 7.3. It is clear that $u$ is smooth and solves the heat equation (by differentiating under the integral, which is justified because, for each $T_0 > 0$, $\Phi$ is bounded in $C^k([T_0, \infty) \times \mathbb{R}^n)$ for each $k \in \mathbb{N}$).

So we only need to prove the last assertion. For $x_0 \in \mathbb{R}^n$, by the previous proposition,

$$\left| \int_{\mathbb{R}^n} \Phi(t, x - y)f(y) \, dy - f(x_0) \right|$$

$$= \left| \int_{\mathbb{R}^n} \Phi(t, x - y)(f(y) - f(x_0)) \, dy \right|$$

$$\leq \int_{\mathbb{R}^n} \Phi(t, x - y)|f(y) - f(x_0)| \, dy$$

$$\leq \int_{B_\varepsilon(x_0)} \Phi(t, x - y)|f(y) - f(x_0)| \, dy$$

$$+ \int_{\mathbb{R}^n \setminus B_\varepsilon(x_0)} \Phi(t, x - y)|f(y) - f(x_0)| \, dy$$

$$=: I_{\varepsilon,t} + II_{\varepsilon,t}$$

for any $\varepsilon > 0$.

We have

$$I_{\varepsilon,t} \leq \sup_{B_\varepsilon(x_0)} |f(y) - f(x_0)| \to 0 \text{ as } \varepsilon \downarrow 0.$$  

Since $f$ is bounded, for $s = r/2\sqrt{t}$

$$II_{\varepsilon,t} \leq C_1 \int_{|x| \geq \varepsilon} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \, dx$$

$$= C_2 t^{-n/2} \int_{\varepsilon}^{\infty} r^{n-1} e^{-\frac{r^2}{4t}} \, dr$$

$$= C_3 \int_{\varepsilon/2\sqrt{t}}^{\infty} s^{n-1} e^{-s^2} \, ds \to 0 \text{ as } \varepsilon/\sqrt{t} \to \infty.$$  

It follows that

$$\lim_{\varepsilon \to 0} \lim_{t \to 0} I_{\varepsilon,t} + II_{\varepsilon,t} = 0,$$

which establishes the assertion. \qed

By Duhamel’s principle we also get a formula for the solution of the inhomogeneous heat equation.
**Theorem 7.7.** Suppose $\sigma \in C^\infty((0, \infty) \times \mathbb{R}^n)$. Then $u : (0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$u(t, x) := \int_0^t \frac{1}{(4\pi(t-s))^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4(t-s)}} \sigma(y, s) \, dy \, ds$$

is smooth, solves

$$(\partial_t + \Delta)u = \sigma$$

and, for all $x \in \mathbb{R}^n$, $\lim_{t \downarrow 0} u(t, x) = 0$.

The proof of this theorem is similar to (but more involved than) the proof of Theorem 7.3.

**Definition 7.8.** The heat ball of radius $a > 0$ at $(t, x) \in \mathbb{R}^{n+1}$ is the set

$$B^p_a(t, x) := \{(y, s) : s \leq t, \Phi(x - y, t - s) \geq a^{-n}\}.$$

Note that this ball is compatible with parabolic rescaling based at $(t, x)$:

$$(s, y) \in B^p_1(0, 0) \iff (r^2 s, r y) \in B^p_r(0, 0).$$

**Theorem 7.9 (Mean-value property).** Suppose $U \subset \mathbb{R}^n$ is open and $T > 0$. If $u \in C^\infty(U_T)$ is a solution of (5.2) and $B^p_1(t, x) \subset U_T$, then

$$u(t, x) = \frac{1}{4\pi^n} \int_{B^p_1(t, x)} u(s, y) \frac{|x - y|^2}{|t - s|^2} \, dy \, ds.$$

**Proof sketch.** For the detailed proof, see [Eva10, pp. 53–54]. The argument presented there proceeds by proving the following:

- The map
  $$r \mapsto \frac{1}{4\pi^n} \int_{B^p_1(t, x)} u(s, y) \frac{|x - y|^2}{|t - s|^2} \, dy \, ds$$
  is constant.

- For all $r > 0$,
  $$\frac{1}{4\pi^n} \int_{B^p_1(t, x)} \frac{|x - y|^2}{|t - s|^2} \, dy \, ds = 1;$$
  hence,
  $$u(t, x) = \lim_{r \downarrow 0} \frac{1}{4\pi^n} \int_{B^p_1(t, x)} u(s, y) \frac{|x - y|^2}{|t - s|^2} \, dy \, ds.$$
Using the mean-value property we can prove:

**Theorem 7.11** (Strong maximum principle). Suppose $U$ is a bounded open and connected subset of $\mathbb{R}^n$, $T > 0$ and $u \in C^\infty(\bar{U}_T)$ satisfies (5.2). If there exists a $(t_0, x_0) \in U_T$ such that

$$u(t_0, x_0) = \max_{\bar{U}_T} u,$$

then $u$ is constant on $\bar{U}_{t_0}$.

**Proof.** Suppose $u$ achieves its maximum at $(t_0, x_0) \in U_T$. If $r > 0$ is small enough so that $B^p_r(t_0, x_0) \subset U_t$, then by (7.10)

$$0 = \frac{1}{4r^n} \int_{B^p_r(t_0, x_0)} (u(t_0, x_0) - u(s, y)) \frac{|x - y|^2}{|t - s|^2}.$$

But for all $(s, y) \in U_T$, we have $u(t_0, x_0) - u(s, y) \geq 0$. Therefore $u$ must be constant on $B^p_r(t_0, x_0)$.

Suppose $\gamma : [0, 1] \to U_T$ is a piece-wise linear path with

$$\gamma(0) = (t_0, x_0)$$

and with decreasing $t$–component. It is not hard to see that $\gamma$ can be covered finitely many heat balls “centered” along the path. Thus $u$ is constant along $\gamma$.

Since $\gamma$ was arbitrary (and by continuity of $u$), $u$ is constant on $\bar{U}_{t_0}$. □
LECTURE 8. Introduction to the Wave Equation

Definition 8.1. The wave equation is the PDE

\[ \partial_t^2 u + \Delta u = 0 \]  

for a function \( u: I \times U \to R \) with \( U \subset \mathbb{R}^n \) open and \( I \subset \mathbb{R} \) an interval.

A solution \( u \) of (8.2) is a model for a dislocation of a membrane (or string) over a domain \( U \).

When studying the wave equation one typically imposes initial data/Cauchy data on the Cauchy hypersurface \( \Gamma = \{0\} \times U \) (cf. Definition 3.12), that is for fixed \( f, g: U \to \mathbb{R} \) we require that

\[ u(0, x) = f(x) \quad \text{and} \quad \partial_t u(0, x) = g(x). \]

If \( U \) is bounded, one typically further imposes boundary conditions as in Section 5.1.

8.1 A toy model

A toy model for the heat equation is the ODE

\[ \partial_t^2 x(t) + Ax(t) \]

for a function \( x: [0, T] \to \mathbb{R}^n \) and symmetric non-negative definite matrix \( A \). In an orthonormal basis \( (e_i) \) consisting of eigenvectors with eigenvalues \( \lambda_i^2 \geq 0 \) we can write the solution of the ODE as

\[ x(t) = \sum_{i=1}^{n} \cos(\lambda_i t) \langle e_i, x(0) \rangle \cdot e_i + \sum_{i=1}^{n} \sin(\lambda_i t) \langle e_i, x'(0) \rangle \cdot e_i. \]

This looks somewhat similar, to what we saw in (5.2). However there are key differences: We need to specify \( x(0) \) and \( x'(0) \) to determine \( x(t) \) uniquely. Moreover and most importantly, the time-dependent coefficients of the \( e_i \) do not decay as \( t \to \infty \) and they do not blow up as \( t \to -\infty \). This foreshadows two key properties of the wave equation: solutions do not become smoother as \( t \) increases (or “roughness of the initial conditions propagates in time”) and the wave equation behaves well backwards in time.

45
8.2  The wave equation on \([0, 1]\)

The preceding discussion suggests that for \(f, g \in L^2([0, 1])\) with

\[
    f = \sum_{n=1}^{\infty} a_n f_n \quad \text{and} \quad g = \sum_{n=1}^{\infty} b_n f_n
\]

the formal expression

\[
    u(t, x) := \sum_{n=1}^{\infty} (a_n \cos(n\pi t) + b_n \sin(n\pi t)) f_n
\]

defines a solution of (8.2) with Cauchy data \((f, g)\) and Dirichlet boundary conditions.

The crucial difficulty in making sense of this is that the coefficients of \(f_n\) do not decay fast in \(t\); hence, \(u(t, x)\) need not be in \(C^2(I \times [0, 1])\) for any interval \(0 \in I\). Nevertheless, we have the following.

**Theorem 8.4.** If \(f\) and \(g\) are smooth, then (8.3) defines a smooth function \(u : \mathbb{R} \times [0, 1] \to \mathbb{R}\) solving (8.2) with initial condition \((f, g)\) and satisfying Dirichlet boundary conditions.

**Exercise 8.5.** Prove that if \(f\) is \(C^\infty([0, 1])\), then for each \(\ell > 0\) there is a constant \(c > 0\) (depending on \(\ell\) and \(\|f\|_{C^k}\)) such that the Fourier coefficients \(a_k\) of \(f\) satisfy

\[
    |a_k| \leq c/k^\ell.
\]

**Exercise 8.6.** Use the previous exercise to prove Theorem 8.4.

**Remark 8.7.** If you’re stuck on these exercises, please, consult [Arn04, Lecture 5].

8.3  d’Alembert’s formula

Let us now consider the case of an infinitely long vibrating string, i.e., \(U = \mathbb{R}\). There is a beautiful solution formula for the Cauchy problem in this case, called d’Alembert’s formula. To find this first notice that on \(\mathbb{R}\),

\[
\partial_t^2 + \Delta = \partial_t^2 - \partial_x^2 = (\partial_t + \partial_x)(\partial_t - \partial_x).
\]

One consequence of this is that any(!) function of the form

\[
    v_R(t, x) := h(t - x) \quad \text{or} \quad v_L(t, x) := h(t + x)
\]
i.e., the solutions of the transport equations

\[(\partial_t + \partial_x)v = 0 \quad \text{and} \quad (\partial_t - \partial_x)v = 0\]

also solve the wave equation (provided \(h\) is smooth enough). These are called right- and left-travelling waves respectively.

**Proposition 8.8.** Every solution \(u\) of the wave equation on \(\mathbb{R}\) can be written uniquely as the sum of a right- and a left-travelling wave

\[u(t, x) = u_R(t - x) + u_L(t + x)\]

with \(u_R(0) = u_L(0)\).

**Remark 8.9 (Finite propagation speed).** Let us first note that Proposition 8.8 implies that waves propagate at finite speed (in fact, speed one): Suppose \(u\) is a solution of the wave equation, then \(u(t, x)\) depends only on the initial data on \([x - t, x + t]\).

**Remark 8.10.** Also note that the amplitude of the waves does not decrease in space. A one-dimensional world would be very noisy!

**Proof of Proposition 8.8.** It is convenient to define new coordinates

\[q := t - x \quad \text{and} \quad s := t + x.\]

These are called characteristic (or null) coordinates for reasons that will soon become apparent. The derivatives in the old and new coordinates are related by

\[\partial_q = \frac{1}{2}(-\partial_t + \partial_x) \quad \text{and} \quad \partial_s = \frac{1}{2}(\partial_t + \partial_x).\]

With respect to these coordinates the wave equation is equivalent to

\[\partial_q \partial_s v = 0.\]

for \(u(t, x) = v(t - x, t + x)\). Hence,

\[\partial_s v(q, s) = \partial_s v(q, 0)\]

which integrates to

\[v(q, s) = v(q, 0) + \int_0^s \partial_s v(a, 0) \, da.\]

This translates back to the asserted statement about \(u\).  \(\square\)
The general solution to the wave equation on $\mathbb{R}$ is given by

**Theorem 8.11.** Suppose $f \in C^2(\mathbb{R})$ and $g \in C^1(\mathbb{R})$. Then there is a unique solution $u \in C^2(\mathbb{R} \times \mathbb{R})$ to the wave equation with Cauchy data $(f, g)$. This solution can be written as

$$u(t, x) := \frac{1}{2}(f(x + t) + f(x - t)) + \frac{1}{2} \int_{x-t}^{x+t} g(y) \, dy.$$  

**(8.12)**

**Proof.** First note that $u$ defined by (8.12) does solve the Cauchy problem. We can write $u(t, x) = u_R(t - x) + u_L(t + x)$ with $u_R(0) = u_L(0) = \frac{1}{2} f(0)$.

The initial conditions amount to the equations

$$f(x) = u_R(-x) + u_L(x) \quad \text{and} \quad g(x) = u'_R(-x) + u'_L(x).$$

Differentiating the first equation yields

$$f'(x) = -u'_R(-x) + u'_L(x).$$

This leads to

$$u'_L(x) = \frac{1}{2}(f'(x) + g(x)) \quad \text{and} \quad u'_R(x) = \frac{1}{2}(-f'(-x) + g(-x)),$n

which integrates to

$$u_L(x) = \frac{1}{2} f(x) + \frac{1}{2} \int_0^y g(y) \, dy \quad \text{and} \quad u_R(x) = \frac{1}{2} f(-x) + \frac{1}{2} \int_0^{-x} g(y) \, dy.$$n

This gives the asserted formula for $u$. 

**Remark 8.13.** Note that $f \in C^k(\mathbb{R})$ and $g \in C^{k-1}(\mathbb{R})$, then $u \in C^k(\mathbb{R})$. However, it typically won’t be any smoother.

**Exercise 8.14.** There is a variant of d’Alembert’s formula for the case $U = [0, \infty)$ and with Dirichlet boundary conditions at $x = 0$: In this case

$$u(t, x) = \begin{cases}  
\frac{1}{2}(f(x + t) + f(x - t)) + \frac{1}{2} \int_{x-t}^{x+t} g(y) \, dy & 0 \leq t \leq x, \\
\frac{1}{2}(f(x + t) - f(t - x)) + \frac{1}{2} \int_{x-t}^{x+t} g(y) \, dy & 0 \leq x \leq t.
\end{cases}$$

State and prove corresponding analogue of **Theorem 8.11.** (*Hint:* Use the reflection principle!)}
LECTURE 9. The Wave Equation in dimension three

This is a draft!

There are explicit solution formulae for the wave equation in each dimension \( n \). One way to derive them is by the method of spherical means. The general case is quite tedious, so we restrict to \( n = 3 \) which yields Kirchhoff’s formula. While the derivation is a bit hairy, the final formula is quite beautiful, so please bear with me.

9.1 Spherical means

Proposition 9.1. Suppose \( u : [0, \infty) \times \mathbb{R}^n \) is a \( C^2 \) solution to Equation 8.2 with initial conditions \( f \) and \( g \). Define \( U, F, G : [0, \infty) \times (0, \infty) \to \mathbb{R} \) by

\[
U_x(t, r) := \int_{\partial B_r(x)} u(t, y) \, dy,
\]

\[
F_x(t, r) := \int_{\partial B_r(x)} F(t, y) \, dy,
\]

\[
G_x(t, r) := \int_{\partial B_r(x)} G(t, y) \, dy.
\]

Then \( U_x : [0, \infty) \times (0, \infty) \) solves the PDE

\[
\partial_t^2 U_x - \partial_r^2 U_x - \frac{n-1}{r} \partial_r U_x = 0
\]

with initial conditions

\[
U_x(t, \cdot) = F_x \quad \text{and} \quad \partial_t U_x(t, \cdot) = G_x.
\]

Moreover,

\[
\lim_{r \to 0} U_x(t, r) = u(t, x) \quad \text{and} \quad \lim_{r \to 0} \partial_t U_x(t, r) = \partial_t u(t, x).
\]

Proof. First, note that

\[
\partial_t^2 U_x(t, r) = \int_{\partial B_r(x)} \partial_t^2 u(t, y) \, dy
\]

by differentiating under the integral.
Denote by $\omega_{n-1}$ the volume of the $(n-1)$-dimensional unit-sphere. We compute
\[
\partial_r U_x(t, r) = \partial_r \left( \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} u(t, y) \, dy \right)
\]
\[
= \partial_r \int_{\partial B_1(0)} u(t, x + rz) \, dz
\]
\[
= \int_{\partial B_1(0)} \langle \nabla u(t, x + rz), z \rangle \, dz
\]
\[
= \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} \langle \nabla u(t, y), (y - x)/r \rangle \, dz
\]
\[
= \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} \nabla u(t, y) \, dz
\]
\[
= \frac{1}{\omega_{n-1} r^{n-1}} \int_{B_r(x)} -\Delta u(t, y) \, dz.
\]

Here we use integration by parts in the last step.

From this we obtain
\[
\partial_r^2 U_x(t, r) = \frac{1}{\omega_{n-1} r^n} \int_{B_r(x)} -\Delta u(t, y) \, dz.
\]
\[
+ \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} -\Delta u(t, y) \, dz
\]
\[
= -\frac{n-1}{r} \partial_r U_x(t, r) + \int_{\partial B_r(x)} -\Delta u(t, y) \, dz.
\]

These formula show that $U_x(t, r)$ satisfies the stated PDE. The rest of the proposition is clear.

9.2 Kirchhoff’s formula

Now suppose $n = 3$. Define $U_x, F_x, G_x: [0, \infty) \times [0, \infty) \to \mathbb{R}$ by

\[
\tilde{U}_x = r U_x, \quad \tilde{F}_x = r F_x, \quad \text{and} \quad \tilde{G}_x = r G_x.
\]

Remark 9.2. The transformation of $U_x, F_x, G_x$ you have to do in dimension $n \geq 3$ is more involved. The transformation for $n$ odd can be found in [Evans Section 2.4.d]. From the resulting solution formulae for $n$ even can be obtain by dimensional reduction.
From Proposition 9.1 we derive that

\[(\partial_r^2 - \partial_t^2)\tilde{U}_x = 0\]  
\[\tilde{U}_x = \tilde{F}_x\quad \text{and}\quad \partial_t \tilde{U}_x = \tilde{G}_x\]  
\[\tilde{U} = 0\]  
on \([0, \infty) \times (0, \infty)\)

on \([0, \infty) \times (0, \infty)\)

Exercise 9.3. Check this!

This is a heat equation on \([0, \infty)\) with Dirichlet boundary conditions at \(x = 0\) and we know that for \(0 \leq r \leq t\)

\[\tilde{U}_x(t, r) = \frac{1}{2} \left( \tilde{F}_x(r + t) - \tilde{F}_x(t - r) \right) + \frac{1}{2} \int_{t-r}^{t+r} \tilde{G}_x(y) \, dy.\]

From this we can recover \(u(t, x)\) as follows:

\[u(t, x) = \lim_{r \downarrow 0} \frac{\tilde{U}_x(t, r)}{r}\]
\[= \partial_r \tilde{U}_x(t, 0)\]
\[= \partial_r \tilde{F}_x(t) + \tilde{G}_x(t).\]

Finally, we compute

\[\partial_r \tilde{F}_x(t) = \partial_t \left( t \int_{\partial B_t(x)} f(y) \, dy \right)\]
\[= \partial_t \left( \frac{1}{4\pi t} \int_{\partial B_t(x)} f(y) \, dy \right)\]
\[= \int_{\partial B_t(x)} (f(y) + \langle \nabla f(t), y - x \rangle) \, dy\]
and

\[\tilde{G}_x(t) = \int_{\partial B_t(x)} tg(y) \, dy.\]

**Theorem 9.4.** Suppose \(f \in C^3\) and \(g \in C^2\). The Cauchy problem for the wave equation on \(\mathbb{R}^3\) has a unique solution given by Kirchhoff’s formula

\[u(t, x) := \int_{\partial B_t(x)} tf(y) + t\nabla_y f(y) + tg(y) \, dy.\]

**Proof.** The previous computations show that if there is a solution it must satisfy (9.5). A computation shows that \(u\) defined by (9.5) solves the wave equation.
Remark 9.6 (Huygens’ principle and finite propagation speed). Note that \( u(t, x) \) depends only on the initial data on the sphere \( \partial B_t(x) \). This is true in all odd dimensions \( n \geq 3 \) and is called Huygens’ principle. The weaker statement that \( u(t, x) \) only depends on the initial data on ball \( B_t(x) \) holds true in all dimensions. It corresponds to the fact that waves propagate at finite speed.

Remark 9.7. Note that we lose differentiability compared to the initial data! (This phenomenon gets worse and worse as \( n \) grows.)

Exercise 9.8 (Dimensional reduction). Note that if \( u(t, x_1, x_2) \) solves the 2-dimensional wave equation then the function \( v(t, x_1, x_2, x_3) := u(t, x_1, x_2) \) solves the 3-dimensional wave equation. Use this to derive Poisson’s formula for the solution to the Cauchy problem in dimension two:

\[
(9.9) \quad u(t, x) = \frac{1}{2} \int_{B_t(x)} \frac{tf(y) + t^2 \nabla f(y) + t^2 h(y)}{\sqrt{t^2 - |x - y|^2}} \, dy
\]
A. Divergence Theorem

The Divergence Theorem is a higher dimensional analogue of the Fundamental Theorem of Calculus. It and its various ramifications will be used repeatedly in this class.

**Definition A.1.** Let $U$ be a open subset. The divergence of a differentiable vector-field $v : U \to \mathbb{R}^n$ is the function $\text{div} \, v : U \to \mathbb{R}$ defined by

$$\text{div} \, v := \sum_{i=1}^{n} \partial_i v_i.$$  

Here $v_i$ are the components of $v$.

**Theorem A.2** (Divergence Theorem). Let $U$ be a open subset with $C^1$ boundary and let $\nu : \partial U \to \mathbb{R}^n$ be the outward-pointing normal vector-field to $\partial U$. If $v : \bar{U} \to \mathbb{R}^n$ is a continuously differentiable vector field, then

$$\int_U \text{div} \, v = \int_{\partial U} \langle v, \nu \rangle.$$  

The expression $\langle v, \nu \rangle$ denotes the inner product of the vector-fields $v$ and $\nu$.

Applying Theorem A.2 to the vector-field $v := fg \cdot e_i$, tells us how to integrate by parts in $\mathbb{R}^n$.

**Theorem A.3** (Integration by parts). Let $U$ be a open subset with $C^1$ boundary. If $f, g : \bar{U} \to \mathbb{R}$ are $C^1$ functions, then

$$\int_U (\partial_i f)g + \int_U f(\partial_i g) = \int_{\partial U} fg \nu_i$$  

for each $i = 1, \ldots, n$. Here $\nu_i$ denotes the $i$–th component of the outward-pointing normal vector field $\nu$.

We will frequently use the following identities, which follow directly from the preceeding theorem.

**Theorem A.4** (Green’s identities). Let $U$ be a open subset with $C^1$ boundary. If $f, g : \bar{U} \to \mathbb{R}$ are $C^2$ functions, then

$$\int_U (\Delta f)g = \int_U \langle \nabla f, \nabla g \rangle - \int_{\partial U} (\partial_{\nu} f)g$$

and

$$\int_U (\Delta f)g - f(\Delta g) = \int_{\partial U} f(\partial_{\nu} g) - (\partial_{\nu} f)g.$$  

Here $\partial_{\nu} f = \langle \nabla f, \nu \rangle$ is the derivative of $f$ in the direction of $\nu$; similarly for $\partial_{\nu} g$.  

53
B. Metric spaces

Metric spaces are an abstraction of the notion of a space for which you can say how far any two points are apart from each other. Some very basic results in the theory of metric spaces can already be exploited to enormous profit for PDE, and this is why we introduce this seemingly unrelated abstract concept here. (It is also very interesting to study metric spaces just by themselves, and there is a surprising amount of theory. In this appendix, however, we are not even scraping the surface.)

Definition B.1. A metric space is a pair \((X, d)\) consisting of a set \(X\) and a function \(d: X \times X \to [0, \infty)\), often called the distance, such that for any \(x, y, z \in X\) we have

- \(d(x, y) = 0\) if and only if \(x = y\),
- \(d(x, y) = d(y, x)\), and
- \(d(x, z) \leq d(x, y) + d(y, z)\).

Remark B.2. All of these axioms except the second one are very natural when talking about distances. The second one, sometimes called symmetry, asserts that it takes as long to get from \(x\) to \(y\) as it takes the other way around from \(y\) to \(x\), and we all know cases where this is not a reasonable assumption in real life. Nevertheless, symmetry is one of the axioms for a metric space.

Remark B.3. One often commits abuse of notation and talks about the metric space \(X\), sweeping \(d\) under the rug. This is usually acceptable, when \(d\) is clear from the context, but if you stumble across a “metric space \(X\)” and have no idea what on earth \(d\) is supposed to be, then you have every right to and should complain!

Let me give a list of important examples of metric spaces appearing throughout this class.

Example B.4. The space of real numbers \(\mathbb{R}\) together with the function \(d: \mathbb{R} \times \mathbb{R} \to [0, \infty)\) defined by

\[
d(x, y) = |x - y|
\]

constitute a metric space.

Exercise B.5. If \((X_1, d_1)\) and \((X_2, d_2)\) are metric spaces, then so is \(X := X_1 \times X_2\) together with

\[
d((x_1, x_2), (y_1, y_2)) := d(x_1, y_1) + d(x_2, y_2)\]

In particular, \(\mathbb{R}^n\) is a metric space.
Example B.6. Given $U \subset \mathbb{R}^n$, we set

$$C^k(U, \mathbb{R}^m) := \{ f : U \to \mathbb{R}^m \text{ \(k\) times continuous differentiable} \}$$

and define $d : C^k(U, \mathbb{R}^m) \times C^k(U, \mathbb{R}^m) \to [0, \infty)$ by

$$d(f, g) := \sum_{i=0}^{k} \sup_{x \in U} |\nabla^i f(x) - \nabla^i g(x)|.$$

The pair $(C^k(U, \mathbb{R}^m), d)$ is a metric space. If $m = 1$, then we usually just write $C^k(U)$.

An important notion that metric spaces inherit from $\mathbb{R}$ is that of a limit.

Definition B.7. Let $(X, d)$ be a metric space. We call $x \in X$ the limit of the sequence $(x_n) \in X^\mathbb{N}$ if

$$\lim_{n \to \infty} d(x, x_n) = 0.$$ 

In this case we write

$$x = \lim_{n \to \infty} x_n.$$ 

Remark B.8. Note that the limit is unique, by the first axiom for metric spaces.

Definition B.9. Let $(X, d)$ be a metric space. A sequence $(x_n) \in X^\mathbb{N}$ is called a Cauchy sequence if for each $\varepsilon > 0$ there exist a $m > 0$ such that for all $k, \ell \geq m$ we have

$$d(x_k, x_\ell) < \varepsilon.$$ 

This means that the elements of $(x_n)$ get increasingly closer to each other as $n$ goes to infinity. Intuitively, one might think that every Cauchy sequence $(x_n)$ needs to have a limit because the points get so close together that they really want to converge. Certainly, this is true in $\mathbb{R}$, but general metric spaces can be horrible, so we make a definition.

Definition B.10. A metric space $(X, d)$ is called complete if every Cauchy sequence has a limit.

Example B.11. $\mathbb{Q}$ with the distance inherited from $\mathbb{R}$ is incomplete.

Exercise B.12. Prove that if $\bar{U} \subset \mathbb{R}^n$ is compact, then $C^0(\bar{U}, \mathbb{R}^m)$ is complete. You can proceed along the following lines: Suppose $(f_n)$ is a Cauchy sequence.

1. Use the limits $\lim_{n \to \infty} f_n(x)$ to construct a map $f : \bar{U} \to \mathbb{R}^m$. 

55
2. Use the Cauchy property of \((f_n)\) to show that \(\lim_{n \to \infty} d(f, f_n) = 0\).

3. Prove that \(f\) is continuous.

**Exercise B.13.** Let \((X, d)\) be a metric space. Set

\[
\bar{X} := \{(x_n) \subset X^N \text{ Cauchy sequence} \} / \sim
\]

with \((x_n) \sim (y_n)\) if and only if

\[
\lim_{n \to \infty} d(x_n, y_n) = 0,
\]

and define \(\bar{d}: \bar{X} \times \bar{X} \to [0, \infty)\) by

\[
\bar{d}((x_n), (y_n)) := \inf_{n \in \mathbb{N}} d(x_n, y_n).
\]

Show that \((\bar{X}, \bar{d})\) is a complete metric space.

**Definition B.14.** Given a metric space \((X, d)\), the metric space \((\bar{X}, \bar{d})\) is called the *completion* of \((X, d)\).

**Example B.15.** The completion of \(\mathbb{Q}\) is \(\mathbb{R}\).

**Example B.16.** Suppose \(U\) is bounded. Consider the \(L^2\)-distance on \(C^0(\bar{U})\) defined by

\[
d_{L^2}(f, g) := \left(\int_U |f - g|^2\right)^{1/2}
\]

\((C^0(\bar{U}), d_{L^2})\) is *not complete*. It is easy to find a discontinuous, but integrable function \(f_\infty\) and a sequence of continuous functions converging to \(f_\infty\) with respect to \(d_{L^2}\). (You may think this is really bad, but in someway this is what makes Fourier series so powerful.)

The space \(L^2(\bar{U})\) is the completion of \((C^0(\bar{U}), d_{L^2})\). If you take a class on measure theory (and if you care about PDE, this is something you should do), then you will learn how to think about the elements of \(L^2(\bar{U})\). Roughly speaking they are functions on \(U\), but you can only define their values at “almost all” points of \(U\).

**Theorem B.17** (Banach’s Fixed Point Theorem). *Let \((X, d)\) be a complete metric space and \(T: X \to X\) a contraction, i.e., for some \(\gamma < 1\) and each \(x, y \in X\),

\[
d(Tx, Ty) < \gamma d(x, y).
\]

Then \(T\) has a unique fixed point \(x_0 \in X\).*
This theorem underlies many PDE applications. It might seem very abstract at first, but (unlike topological fixed point theorems) it is in fact constructive.

**Exercise B.18.** Prove Theorem B.17 along the following lines:

1. Use the contraction property of $T$ to show that there is at most one fixed point.

2. Pick any $x \in X$. Consider the sequence $(x_n) := (T^n x)$. Prove that $(x_n)$ is Cauchy and use completeness of $X$ to extract the fixed point $x_0$.

(What can you say about the rate of convergence in terms of $\gamma$?)
C. Fourier Series on $[0, 1]$

**Definition C.1.** A Hilbert space is a vector space $H$ together with an inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ such that the metric defined by

$$d(x, y) := \|x - y\| := \sqrt{\langle x - y, x - y \rangle}$$

makes $(H, d)$ into a complete metric space.

**Example C.2.** The archetypal example of a Hilbert space is

$$\ell^2 := \left\{ (a_i) \in \mathbb{R}^\infty : \sum_{i=1}^{\infty} |a_i|^2 < \infty \right\}$$

with inner product

$$\langle a_i, b_i \rangle_{\ell^2} := \sum_{i=1}^{\infty} a_i b_i.$$

**Example C.3.** The inner product

$$\langle f, g \rangle_{L^2} := \int_0^1 f(x) g(x) \, dx$$

defined on $C^0([0, 1])$ does not make $C^0([0, 1])$ into a Hilbert space. The completion of $C^0([0, 1])$ with respect to the metric induced by $\langle \cdot, \cdot \rangle_{L^2}$, which we denote by $L^2([0, 1])$, however, is a Hilbert space.

**Remark C.4.** One has to be a bit careful working with $L^2$–spaces, since elements are not functions but only equivalence classes of functions. Functions in the same equivalence class only differ on a set of measure zero, so this problem is “mostly harmless”, but not entirely harmless. Note, however, that the canonical maps $C^0([0, 1]) \to L^2([0, 1])$ is injective and thus we can make statements like “a certain $f \in L^2([0, 1])$ is continuous”. What this means of course is that $f$ can be represented by a continuous function.

**Definition C.5.** An (countable) orthonormal basis of a Hilbert space $(H, \langle \cdot, \cdot \rangle)$ is a sequence $(e_i)_{i \in \mathbb{N}} \in H^\mathbb{N}$ such that

$$\langle e_i, e_j \rangle = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

and, for every $x \in H$, $\langle x, e_i \rangle = 0$ for all $i \in \mathbb{N}$ if and only if $x = 0$. 
Remark C.6. If \( H \) is not finite dimensional, then an orthonormal basis of the Hilbert space \( H \) is not a basis of the vector space \( H \), since not every element can be written as a finite linear combination.

**Proposition C.7.** If \((e_i)\) is a orthonormal basis of a Hilbert space \((H, \langle \cdot , \cdot \rangle)\), then, for every \(x, y \in H\), setting

\[
a_i := \langle x, f_i \rangle \quad \text{and} \quad b_i := \langle y, f_i \rangle
\]

we have

\[
x = \sum_{i=1}^{\infty} a_i e_i
\]

and

\[
\langle x, y \rangle = \sum_{i=1}^{\infty} a_i b_i.
\]

The basic result of Fourier analysis on \([0, 1]\) can be summarised as follows.

**Theorem C.8.** The sequence \((f_n(x) := \sqrt{2} \sin(n \pi x))\) is an orthonormal basis of the Hilbert space \(L^2([0, 1])\).

**Definition C.9.** If \(f \in L^2([0, 1])\), then the **Fourier coefficients** of \(f\) are the sequence \((a_n) \in \mathbb{R}^N\) defined by

\[
a_n := \langle f, f_n \rangle
\]

and the **Fourier series** is the expression

\[
\sum_{n=1}^{\infty} a_n f_n.
\]

What makes this particular orthonormal basis so useful for us is the simple fact that

\[
\Delta f_n = -\partial_x^2 f_n = (n \pi)^2 f_n.
\]

This means that the Laplace operator \(\Delta\) becomes **diagonal** in the orthonormal basis \((f_n)\). Note, however, that the eigenvalues do go to infinity; hence, \(\Delta\) is an **unbounded** linear operator.

There is a tight connection between the regularity of \(f\) and the rate of decay of the Fourier coefficients.
Proposition C.10. Fix $k \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$. If the Fourier coefficients of $f \in L^2([0, 1])$ satisfy

$$(a_n) \in \ell^1_k := \left\{ (b_n) \in \mathbb{R}^N : \sum_{n=1}^{\infty} n^k |b_n| < \infty \right\},$$

then $f \in C^k([0, 1])$ and

$$\lim_{N \to \infty} \left\| f - \sum_{n=1}^{N} a_n f_n \right\|_{C^k} = 0.$$ 

Here

$$\|f\|_{C^k} := \sum_{i=0}^{k} \sup_{x \in [0, 1]} |\nabla^i f(x)|.$$ 

Exercise C.11. Prove this proposition. Here are some hints if you get stuck:

- What is $\|f_n\|_{C^k}$?
- Show that $\sum_{n=1}^{N} a_n f_n$ is a Cauchy sequence in $C^k([0, 1])$.
- Use that if $g = \lim_{N \to \infty} g_N$ in $C^k([0, 1])$, then the same holds true in $L^2([0, 1])$.

Exercise C.12. Define $f \in L^2([0, 1])$ by

$$f(x) = \begin{cases} 1 & x \leq 1/2 \\ -1 & x \geq 1/2. \end{cases}$$

Show that the Fourier coefficients of $f$ are

$$a_{4k+2} = \frac{4\sqrt{2}}{(4k + 2)\pi}$$

and $a_n = 0$ if $n \neq 2 \mod 4$. Show that

$$\lim_{N \to \infty} \sum_{n=1}^{4N+2} a_n f_n \left( \frac{1}{2} - \frac{1}{4N + 2} \right) = 2 \int_{0}^{1} \frac{\sin(\pi x)}{\pi x}.$$

Remark. The right-hand side is approximately 1.18. Thus the partial sums of the Fourier expansion overshoot by about 9% times the height of the discontinuity at 1/2. This is called the Gibbs phenomenon.
References


