18.152: Introduction to Partial Differential Equations

Thomas Walpuski
Massachusetts Institute of Technology

2016-04-24
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


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


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:

<table>
<thead>
<tr>
<th>Component</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>homework</td>
<td>40%</td>
</tr>
<tr>
<td>midterm</td>
<td>20%</td>
</tr>
<tr>
<td>final exam</td>
<td>40%</td>
</tr>
</tbody>
</table>

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? 8
  1.1 What is a PDE? ................................. 8
  1.2 Some interesting examples ...................... 9
  1.3 Questions about PDE ............................ 11
  1.4 Linear PDE .................................. 12
  1.5 More stuff you should know .................... 14

2 Ordinary Differential Equations 15
  2.1 Uniqueness of solutions ....................... 16
  2.2 Local existence theorems ...................... 19
  2.3 Dependence of solutions on initial conditions . 20

3 First-order PDE 21
  3.1 Change of coordinates ......................... 21
  3.2 Warm-up: the transport equation .............. 22
  3.3 The method of characteristics ................ 24

4 The inviscid Burgers’ equation 27

5 Introduction to the Heat Equation 32
  5.1 Boundary conditions ............................ 33
  5.2 A toy model .................................. 35
  5.3 Solution of the heat equation with Dirichlet boundary conditions in dimension one .................. 36

6 Uniqueness for the Heat Equation 39
  6.1 Energy method .................................. 40
  6.2 Weak maximum principle ....................... 41

7 The Heat Kernel on \( \mathbb{R}^n \) 43

8 Introduction to the Wave Equation 47
  8.1 A toy model .................................. 47
  8.2 The wave equation on \([0, 1]\) ...................... 48
  8.3 d’Alembert’s formula .......................... 48
9 The Wave Equation in dimension three and energy methods 51
  9.1 Spherical means ................................................. 51
  9.2 Kirchhoff’s formula ............................................. 52
  9.3 Uniqueness of solutions to the Cauchy problem .......... 54
  9.4 Finite propagation speed ..................................... 55

10 The Fourier Transform 57
  10.1 The Fourier Transform on $L^1$ .............................. 57
  10.2 Fourier inversion on Schwartz functions ................. 59

11 The Fourier Transform (continued) 62
  11.1 Plancherel’s theorem .......................................... 62
  11.2 Tempered distributions ...................................... 63
  11.3 A derivation of the Heat Kernel ............................ 65

12 Introduction Curve-Shortening Flow 66

13 Laplace’s and Poisson’s equations 67
  13.1 Dirichlet’s principle .......................................... 67
  13.2 Connection with holomorphic functions ................. 68
  13.3 Uniqueness via the energy method ......................... 69
  13.4 The weak maximum principle ............................... 69
  13.5 $W^{2,2}$–estimates ........................................... 70
  13.6 Dirichlet’s principle (cont.) ................................. 70

14 The Poincaré inequalities 74
  14.1 Dirichlet–Poincaré inequality .............................. 74
  14.2 Neumann–Poincaré inequality .............................. 75
  14.3 The Li–Schoen proof of the Dirichlet–Poincaré inequality . 77

15 Mean-value properties of harmonic functions 79
  15.1 Strong maximum principle .................................. 79
  15.2 $C^k$–estimates ................................................. 80
  15.3 Liouville’s Theorem .......................................... 81
  15.4 Harnack’s Inequality ......................................... 81

16 Weyl’s Lemma 82
  16.1 Proof using Heat Kernel ..................................... 82
  16.2 Proof via Mean-value property ............................ 85
  16.3 Unique continuation and the frequency function ........ 87
17 Green’s functions
  17.1 Derivations of the Green’s function . . . . . . . . . . . . . . . . . 91
  17.1.1 Derivation from the heat kernel . . . . . . . . . . . . . . 91
  17.1.2 Derivation using the Fourier transform . . . . . . . . . . 92
  17.2 Solving the Poisson equation with $C^1$ inhomogeneity . . . . . . 93
  17.3 Representation formula for solutions of the Dirichlet problem . . . 94
  17.4 The Dirichlet problem on a ball . . . . . . . . . . . . . . . . . . 95

18 Perron’s Method
  18.1 Harmonic replacement . . . . . . . . . . . . . . . . . . . . . . 98
  18.2 The maximal subharmonic function . . . . . . . . . . . . . . 98
  18.3 The boundary condition . . . . . . . . . . . . . . . . . . . . . . 99

19 Minimal hypersurfaces and the Bernstein problem

20 $L^2$ regularity theory for second order elliptic operators in divergence
   form
  20.1 Existence of weak solutions . . . . . . . . . . . . . . . . . . . 108
  20.2 $L^2$ regularity for $\Delta$ . . . . . . . . . . . . . . . . . . . 109
  20.3 $L^2$ interior regularity for variable coefficients operators . . . . . 110
  20.4 $L^2$ boundary regularity . . . . . . . . . . . . . . . . . . . . 114

21 Hölder spaces
  21.1 Definition and basic properties of Hölder spaces . . . . . . . . . . 116
  21.2 Integral characterisation of Hölder spaces . . . . . . . . . . . . 116
  21.3 Morrey’s theorem . . . . . . . . . . . . . . . . . . . . . . . . . 120

22 Campanato estimates
  22.1 Estimates for weak solutions . . . . . . . . . . . . . . . . . . . 121
  22.2 Comparison estimates . . . . . . . . . . . . . . . . . . . . . . . 122
  22.3 Operators with variable coefficients . . . . . . . . . . . . . . . 123

23 Higher Regularity
  23.1 $C^{1,\alpha}$ estimate . . . . . . . . . . . . . . . . . . . . . . . . 127
  23.2 Bootstrapping $C^{k,\alpha}$ estimates . . . . . . . . . . . . . . . 129
  23.3 Boundary estimates . . . . . . . . . . . . . . . . . . . . . . . . 129
  23.4 Existence theorems . . . . . . . . . . . . . . . . . . . . . . . . . 130

A Divergence Theorem

B Metric spaces

6
C  Fourier Series on \([0, 1]\)  

D  Dominated Convergence Theorem
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:

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

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 \textit{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:

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

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

where $F$ is some function.

\textbf{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 $v: U \to \mathbb{R}^{n+1}$ 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:

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

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.
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

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

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. \hfill \Box

2.1 Uniqueness of solutions

Example 2.10 (Failure of Uniqueness). Imagine an infinitely high\textsuperscript{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 $\phi_T : \mathbb{R} \rightarrow \mathbb{R}$ 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 $F : U \rightarrow \mathbb{R}^m$ 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 $\text{Lip}(F) := \sup\{ \frac{|F(x) - F(y)|}{|x - y|} : x, y \in U \} \in [0, \infty]$ the Lipschitz constant of $F$. Clearly, $F$ is Lipschitz continuous if and only if $\text{Lip}(F) < \infty$.

\textsuperscript{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 $y, z \in V$

$$|F(s, y) - F(s, z)| \leq L.$$  

Remark 2.15. This is 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 $t \in I$

$$g(t) \leq B e^{A|t-t_0|}.$$

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 \( g(t) \leq G(t) \).

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

$$\dot{x} = F(t, x) \quad \text{and} \quad x(t_0) = x_0$$

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

$$T(\phi)(t) := x_0 + \int_{t_0}^{t} F(s, \phi(s)) \, ds$$

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 $M := \sup\{|F(s, x_0)| : s \in [t_0 - \delta, t_0 + \delta]|$. 

19
Step 2. Set \( I := [t_0 - \delta, t_0 + \delta] \) and 
\[
X := C^0(I; \overline{B}_\varepsilon(x_0)) = \{ \phi: I \to \overline{B}_\varepsilon(x_0) \text{ continuous} \}
\]
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
\[
|T\phi(t) - x_0| \leq \int_{t_0}^t |F(s, \phi(s)) - F(s, x_0)| \, ds + |F(s, x_0)| \, ds 
\]
\[
\leq \delta(L\varepsilon + M) \leq \varepsilon.
\]

Step 3. \( C^0(I; \overline{B}_\varepsilon(x_0)) \) 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
\[
d(T\phi_1, T\phi_2) = \sup_I \left| \int_{t_0}^t F(s, \phi_1(s)) - F(s, \phi_2(s)) \, ds \right|
\]
\[
\leq \sup_I \int_{t_0}^t |F(s, \phi_1(s)) - F(s, \phi_2(s))| \, ds
\]
\[
\leq \delta \int_{t_0}^t |\phi_1(s) - \phi_2(s)| \, ds
\]
\[
\leq \delta Ld(\phi_1, \phi_2) = \gamma d(\phi_1, \phi_2).
\]

This completes the proof. \( \square \)

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 \([-\varepsilon, \varepsilon] \times \overline{V} \subset U\), 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. \qed

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

\[ v(x) \cdot \nabla u(x) = 0 \]  

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

\[ \partial_s x(s) = v(x(s)). \]  

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:
\[ \partial_s u(\phi(s)) = \partial_s \phi(s) \cdot \nabla u(\phi(s)) = v(\phi(s)) \cdot \nabla u(\phi(s)) = 0. \] \qed

22
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

$$u|_{\Gamma} = f_0, \quad (\partial_\nu u)|_{\Gamma} = f_1, \quad \ldots \quad (\partial_{\nu}^k u)|_{\Gamma} = f_k.$$ 

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

By the chain rule
$$\partial_s y(s) = \nabla u(x(s)) \cdot \partial_s x(s).$$

If $x$ where such that
$$\partial_s x(s) = b(x, y(s)),$$
then $y$ solves the ODE
$$\partial_s y(s) = -c(x(s), y(s)) y(s).$$

Definition 3.14. The characteristic equation for (3.2) is the ODE
\begin{equation}
\begin{aligned}
\partial_s x(s) &= b(x(s), y(s)) \\
\partial_s y(s) &= -c(x(s), y(s)) y(s).
\end{aligned}
\tag{3.15}
\end{equation}

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 $f: \Gamma \rightarrow \mathbb{R}$. 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 $x \in U_0$, \[d\Phi(0, x) = \begin{pmatrix} \left. b(x, f(x)) \right|_{0} & 0 \\ \text{id}_{T_x \Gamma} \end{pmatrix}\]

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 \( u : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R} \)

\[
\frac{\partial}{\partial t} u + u \frac{\partial}{\partial x} u = 0. \tag{4.2}
\]

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

\[
\begin{align*}
\frac{\partial}{\partial t} u + (u \cdot \nabla) u + \nu \Delta u &= -\nabla p \\
\text{div } u &= 0.
\end{align*} \tag{4.5}
\]

for a pair of maps \( u : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^n \), the **flow velocity**, and \( p : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R} \), the **pressure**.\(^3\) 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**

\[
\frac{\partial}{\partial t} u + u \frac{\partial}{\partial x} u - \nu \frac{\partial^2}{\partial x^2} u = 0.
\]

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

\[
\frac{\partial}{\partial t} u + \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) = 0;
\]

\(^2\text{Here } u \cdot \nabla = \sum_{i=1}^n u_i \partial_i. \text{ The whole expression } \frac{\partial}{\partial t} + (u \cdot \nabla) \text{ is sometimes called the } \text{material derivative.} \text{ It is the time derivative from the perspective of a particle moving along the flow.}\)

\(^3\text{The role of the pressure is somewhat secondary since it can be recovered from the flow velocity using the equation}\)

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

\text{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 \( t \in [0, T) \) 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
\[ \partial_s t(s) = 1, \]
\[ \partial_s x(s) = y(s), \]
\[ \partial_s y(s) = 0. \]

Moreover, note that
\[ \partial_s^2 x(s) = \partial_s y(s) = 0. \]

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.

---

*The name conservation law comes from the fact that solutions have the property \( \partial_t \int_{\mathbb{R}} u(t, x) \, dx = 0 \) (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): \( u: [0, \infty) \times \mathbb{R} \to \mathbb{R} \) 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 \( u: [0, 1) \times \mathbb{R} \to \mathbb{R} \) defined by
\[
u(t, x) = \frac{x}{t - 1}.
\]

Exercise 4.10. Find an example of Cauchy data for which no solution \( u: [0, \varepsilon) \times \mathbb{R} \to \mathbb{R} \) 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” \( u: [0, 1) \times \mathbb{R} \to \mathbb{R} \) defined by
\[
u(t, x) := \begin{cases} 
1 & x \leq t \\
\frac{1 - x}{t - 1} & x \in [t, 1] \\
0 & x \geq 1
\end{cases}
\]
(4.12)
I wrote “solution” because \( u(t, x) \) is not \( C^1 \).
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

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

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

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

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 $u$ on $[0, \infty) \times \mathbb{R}$ 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 $f(x_0) > f(x_1)$.
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 \( \Phi : [0, \infty) \times \mathbb{R} \rightarrow [0, \infty) \times \mathbb{R} \) 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. \( \qed \)
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 heat equation is the PDE

\[ \partial_t u(t, x) + \Delta u(t, x) = 0 \]

for a function \( u : [0, T) \times U \rightarrow \mathbb{R} \) 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 \( \sigma : [0, T) \times U \rightarrow \mathbb{R} \). The heat equation with source term \( \sigma \) is the PDE

\[ \partial_t u(t, x) + \Delta u(t, x) = \sigma(t, x). \]

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

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, somewhat 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
\[
(5.9) \quad u(0, x) = f(x) \quad \text{for all} \ x \in U.
\]
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**

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

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

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

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

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 \( \partial U \).

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

The Dirichlet, Neumann, Robin, or mixed initial-boundary value problem for the heat equation 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.

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,

\[ (\partial_t + \Delta)u(t, x) = (\partial_t + \Delta) \int_0^t u_\tau(t - s, x) \, ds \]

\[ = u_t(0, x) + \int_0^t (\partial_t + \Delta)u_\tau(t - \tau, x) \, d\tau \]

\[ = \sigma(t, x). \]
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) \[ x(t) = e^{-At}x(0). \]

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

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

then

\[
\int_U (\Delta f)g = \int_U f(\Delta g) \quad \text{and} \quad \int_U (\Delta f)f \geq 0.
\]

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.

\[ \square \]
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

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

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

$$u(t, 0) = u(t, 1) = 0 \quad \text{for all } t > 0,$$

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 $t \in [0, \infty)$,

$$U(t) = \sum_{n=1}^{\infty} e^{-\langle n\pi, \cdot \rangle} \langle f_n, f \rangle f_n \text{ in } L^2.$$
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 \( t \geq T_0 > 0 \)
\[
\sum_{n=1}^{\infty} (\pi n)^{2\ell} e^{-(\pi n)^2 t} |\langle f_n, f \rangle| \leq \left( \sum_{n=1}^{\infty} (\pi n)^{2\ell} e^{-2(\pi n)^2 T_0} \right)^{1/2} \left( \sum_{n=1}^{\infty} |\langle f_n, f \rangle|^2 \right)^{1/2} < \infty.
\]

**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 sumable, 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 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 $v = 0$. 

39
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

$$\dot{E}^2 \leq E \ddot{E}.$$  

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

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

But (6.3) means that

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

Thus $\log E$ is a convex function which tends to $-\infty$ as $t \to T$. This is impossible.
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 \( \partial \rho U_T \), i.e.,
\[
\max_{\bar{U}_T} u = \max_{\partial \rho U_T} u.
\]Note that this gives another proof of Theorem 6.1. The function \( v \) satisfies the hypothesis of Theorem 6.5 and \( v|_{\partial \rho U_T} = 0 \) 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, \( x_0 \in \bar{U} \) 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, \( \Delta u = -\text{tr} \text{Hess}(u) \geq 0 \). 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
\[
\max_{\bar{U}_T} u_\varepsilon = \max_{\partial \rho U_T} u_\varepsilon.
\]Because
\[
u_\varepsilon \leq u \leq u_\varepsilon + \varepsilon T,
\]
we also have
\[
\max_{\partial U_T} u - \varepsilon T \leq \max_{U_T} u - \varepsilon T \leq \max_{U_T} u_{\varepsilon} \leq \max_{\partial U_T} u_{\varepsilon} \leq \max_{\partial U_T} u.
\]

Since this is true for all \( \varepsilon > 0 \), we have
\[
\max_{U_T} u = \max_{\partial U_T} u. \quad \square
\]

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

**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
\[
\begin{align*}
\partial_t u + \Delta u &= \sigma \\
\partial_t v + \Delta v &= \tau,
\end{align*}
\]
then
\[
\max_{U_T} |u - v| \leq \max_{\partial U_T} |u - v| + T \max_{U_T} |\sigma - \tau|.
\]

**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
\[
\max_{U_T} w - tM \leq \max_{\partial U_T} w - tM \leq \max_{\partial U_T} w
\]
and
\[
\max_{U_T} -w - tM \leq \max_{\partial U_T} -w - tM \leq \max_{\partial U_T} -w.
\]
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

\[
\left(\int_{\mathbb{R}} e^{-s^2} \, ds\right)^2 = \int_{\mathbb{R}^2} e^{-|z|^2} \, dz = \int_0^{2\pi} \int_0^\infty r e^{-r^2} \, dr \, d\theta = 2\pi \int_0^\infty r e^{-r^2} \, dr = \pi.
\]

The last step uses

\[d(e^{-r^2}) = -2r e^{-r^2}.
\]
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 \( k \in \mathbb{N}_0 \).

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 \( s = r/2\sqrt{t} \)

\[
II_{\varepsilon,t} \leq C_1 \int_{|x| \geq \varepsilon} \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2 / 4t} \, dx \\
= C_2 t^{-n/2} \int_{\varepsilon}^{\infty} r^{n-1} e^{-r^2 / 4t} \, dr \\
= C_3 \int_{\varepsilon/2\sqrt{t}}^{\infty} s^{n-1} e^{-s^2} \, ds \to 0 \quad \text{as } \varepsilon/\sqrt{t} \to \infty.
\]

It follows that

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

which establishes the assertion. \( \square \)

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 $r > 0$ at $(t, x) \in \mathbb{R}^{n+1}$ is the set

$$B_r^p(t, x) := \{(y, s) \in \mathbb{R}^{n+1} : s \leq t, \Phi(x - y, t - s) \geq r^{-n}\}.$$ 

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

$$(s, y) \in B_1^p(0, 0) \iff (r^2 s, ry) \in B_r^p(0, 0).$$

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 $B_r^p(t, x) \subset U_T$, then

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

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

- The map

$$r \mapsto \frac{1}{4r^n} \int_{B_r^p(t, x)} u(s, y) \frac{|x - y|^2}{|t - s|^2} \, dy \, ds$$

is constant.

- For all $r > 0$,

$$(7.10) \quad \frac{1}{4r^n} \int_{B_r^p(t, x)} \frac{|x - y|^2}{|t - s|^2} \, dy \, ds = 1;$$

hence,

$$u(t, x) = \lim_{r \downarrow 0} \frac{1}{4r^n} \int_{B_r^p(t, x)} u(s, y) \frac{|x - y|^2}{|t - s|^2} \, dy \, ds.$$
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 $\gamma : [0, 1] \rightarrow U_T$ is a piece-wise linear path with

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

and with decreasing $t$–component. Its 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 $\bar{U}_{t_0}$.

$\square$
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 \( u: I \times U \to \mathbb{R} \) 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

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

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.
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

$$u(t, x) := \sum_{n=1}^{\infty} \left( a_n \cos(n\pi t) + \frac{b_n}{n\pi} \sin(n\pi t) \right) f_n$$

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 $f, g \in C^\infty([0, 1])$ are compactly supported in $(0, 1)$, 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 $C^\infty([0, 1])$ is compactly supported in $(0, 1)$, 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

$$v_R(t, x) := h(t - x) \quad \text{or} \quad v_L(t, x) := h(t + x)$$

48
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

\begin{equation}
(8.12) \quad u(t, x) := \frac{1}{2} (f(x + t) + f(x - t)) + \frac{1}{2} \int_{x-t}^{x+t} g(y) \, dy.
\end{equation}

**Proof.** First note that $u$ defined by (8.12) does solve the Cauchy problem. For uniqueness, suppose $u$ is such a solution. 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

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

which integrates to

\[ u_L(x) = \frac{1}{2} f(x) + \frac{1}{2} \int_0^y g(y) \, dy \quad \text{and} \quad u_R(x) = \frac{1}{2} f(-x) + \frac{1}{2} \int_{-x}^0 g(y) \, dy. \]

This gives the asserted formula for $u$. \qed

**Remark 8.13.** Note that if $f \in C^k(\mathbb{R})$ 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

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

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

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

Recall that

\[
\int_U := \frac{1}{\text{vol}(U)} \int_U .
\]

**Proposition 9.1.** Suppose \( u : [0, \infty) \times \mathbb{R}^n \to \mathbb{R} \) is a \( C^2 \) solution of (8.2) with initial conditions \( f \) and \( g \). Define \( U_x : [0, \infty) \times (0, \infty) \to \mathbb{R} \) and \( F_x, G_x : (0, \infty) \to \mathbb{R} \) by

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

\[
F_x(r) := \int_{\partial B_r(x)} f(y) \, dy, \quad \text{and}
\]

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

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

\[
U_x(0, \cdot) = F_x \quad \text{and} \quad \partial_t U_x(0, \cdot) = G_x.
\]

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 $\omega_n$ the volume of the $n$–dimensional unit-sphere. We compute

$$\partial_t U_x(t, r) = \partial_r \left( \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} u(t, y) \, dy \right)$$

$$= \partial_r \frac{1}{\omega_{n-1}} \int_{\partial B_1(0)} u(t, x + rz) \, dz$$

$$= \frac{1}{\omega_{n-1}} \int_{\partial B_1(0)} \langle \nabla u(t, x + rz), z \rangle \, dz$$

$$= \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} \langle \nabla u(t, y), (y - x)/r \rangle \, dz$$

$$= \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} \nabla u(t, y) \, dz$$

$$= \frac{1}{\omega_{n-1} r^{n-1}} \int_{B_r(x)} -\Delta u(t, y) \, dz.$$  

Here we use Theorem A.4 in the last step. From this we obtain

$$\partial_t^2 U_x(t, r) = \frac{1 - n}{\omega_{n-1} r^n} \int_{B_r(x)} -\Delta u(t, y) \, dz.$$  

$$+ \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} -\Delta u(t, y) \, dz$$

$$= -\frac{n - 1}{r} \partial_t U_x(t, r) + \int_{\partial B_r(x)} -\Delta u(t, y) \, dz.$$  

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

### 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 [Eva10, 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
\[
(\partial_t^2 - \partial_r^2)\tilde{U} = 0 \quad \text{on } [0, \infty) \times (0, \infty)
\]
\[
\tilde{U} = \tilde{F}_x \quad \text{and} \quad \partial_t \tilde{U} = \tilde{G}_x \quad \text{on } \{0\} \times (0, \infty)
\]
\[
\tilde{U} = 0 \quad \text{on } [0, \infty) \times \{0\}.
\]

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 \(0 \leq r \leq t\)
\[
\tilde{U}_x(t, r) = \frac{1}{2} \left( \tilde{F}_x(t + r) - \tilde{F}_x(t - r) \right) + \frac{1}{2} \int_{t-r}^{t+r} \tilde{G}_x(y) \, dy.
\]

From this we can recover \(u(t, x)\) as follows:
\[
u(t, x) = \lim_{r \to 0} \frac{\tilde{U}_x(t, r)}{r} = \partial_r \tilde{U}_x(t, 0) = \partial_r \tilde{F}_x(t) + \tilde{G}_x(t).
\]

Finally, we compute
\[
\partial_r \tilde{F}_x(t) = \partial_t \left( t \int_{\partial B_t(x)} f(y) \, dy \right) = \int_{\partial B_t(x)} f(y) + t \nabla_{\nu} f(t) \, dy \quad \text{and}
\]
\[
\tilde{G}_x(t) = \int_{\partial B_t(x)} t g(y) \, dy.
\]

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
\[
u(t, x) := \int_{\partial B_t(x)} f(y) + t \nabla_{\nu} f(y) + t g(y) \, dy.
\]

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:

\[
(9.9) \quad u(t, x) = \frac{1}{2} \int_{B_t(x)} \frac{tf(y) + t \langle \nabla f(y), y - x \rangle + t^2 g(y)}{\sqrt{t^2 - |x - y|^2}} \, dy
\]

9.3 Uniqueness of solutions to the Cauchy problem

Suppose \( U \) is a bounded open subset of \( \mathbb{R}^n \) and \( T > 0 \)

Theorem 9.10. Given \( f, g : \bar{U} \rightarrow \mathbb{R} \), there exists at most one \( u \in C^2([0, T] \times \bar{U}) \) such that

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

with

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

and satisfying a fixed choice of boundary conditions.

Proof. If there are more than one, then the difference of two such solution gives us a \( w : [0, T] \times U \rightarrow \mathbb{R} \) satisfying

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

\[
w(t, x) = 0 \quad \text{for} \ x \in \partial U
\]

\[
w(0, \cdot) = \partial_t w(0, \cdot) = 0.
\]

We introduce the following (somewhat ad-hoc) energy quantity

\[
E(t) := \frac{1}{2} \int_U |\partial_t w|^2 + |\nabla w|^2
\]

Since

\[
\partial_t E(t) = \int_U \langle \partial_t^2 w, \partial_t w \rangle + \langle \nabla w, \nabla \partial_t w \rangle
\]

\[
= \int_U \langle \partial_t^2 w + \Delta w, \partial_t w \rangle = 0,
\]

this energy is unchanging in \( t \). It must therefore vanish since \( E(0) = 0 \). \( E(t) = 0 \), however, implies \( \partial_t w(t) = 0 \) and thus \( w(t) = 0 \) since \( w(0, \cdot) = 0 \). \( \square \)
9.4 Finite propagation speed

The following is a demonstration of the finite propagation speed of waves and also shows uniqueness of the Cauchy problem on \( \mathbb{R}^n \).

**Theorem 9.11.** Let \( u \in C^2([0, \infty) \times \mathbb{R}^n) \) be a solution of (8.2). Fix \( (t_0, x_0) \in (0, \infty) \times \mathbb{R}^n \). Set
\[
C := \{(t, x) : t \in [0, t_0], x \in B_{t_0-t}(x_0)\}.
\]
If \( u(0, x) = \partial_t u(0, x) = 0 \) for all \( x \in B_{t_0}(x_0) \), then \( u(t, x) = 0 \) for all \( (t, x) \in C \).

**Proof.** Again, we introduce an ad-hoc energy quantity
\[
E(t) := \frac{1}{2} \int_{B_{t_0-t}(x)} |\partial_t u|^2 + |\nabla u|^2.
\]
We know that \( E(0) = 0 \). Now we compute how \( E \) changes in \( t \):
\[
\partial_t E(t) = \frac{1}{2} \partial_t \int_{B_{t_0-t}(x)} |\partial_t u|^2 + |\nabla u|^2
\]
\[
= \int_{B_{t_0-t}(x)} \langle \partial_t^2 u, \partial_t u \rangle + \langle \nabla w, \nabla \partial_t w \rangle
\]
\[
- \frac{1}{2} \int_{B_{t_0-t}(x)} |\partial_t w|^2 + |\nabla w|^2
\]
\[
= \int_{B_{t_0-t}(x)} \langle \partial_t^2 u + \Delta w, \partial_t w \rangle
\]
\[
+ \int_{\partial B_{t_0-t}(x)} \langle \partial_t u, w_t \rangle - \frac{1}{2} \left( |\partial_t w|^2 + |\nabla w|^2 \right)
\]
\[
= \int_{\partial B_{t_0-t}(x)} \langle \partial_t u, w_t \rangle - \frac{1}{2} \left( |\partial_t w|^2 + |\nabla w|^2 \right) \leq 0.
\]
Here, the last inequality uses \( ab \leq \frac{1}{2} (a^2 + b^2) \).

Thus \( E \) is non-increasing in \( t \). Since \( E(0) = 0 \) and \( E \geq 0 \), we must have \( E = 0 \). Thus \( \partial_t u = 0 \) (and \( \nabla u = 0 \)) within \( C \); hence, \( u = 0 \) in \( C \) because \( u(0, x) = 0 \) for all \( x \in B_{t_0}(x_0) \). \( \square \)

**Exercise 9.12** (Equipartition of energy). Suppose \( u \in C^2(\mathbb{R} \times [0, \infty)) \) solves the Cauchy problem for the wave equation with compactly supported Cauchy data \((f, g)\). Define the kinetic energy \( T \) and the potential energy \( U \) by
\[
T(t) := \frac{1}{2} \int_{\mathbb{R}} |\partial_t u|^2 \quad \text{and} \quad U(t) := \frac{1}{2} \int_{\mathbb{R}} |\nabla u|^2.
\]
Prove that there is a $T > 0$ such that for all $t \geq T$ the kinetic energy equals the potential energy, i.e.,

$$T(t) = U(t) \quad \text{for } t \geq T!$$

How does $T$ depend on $f$ and $g$? (Hint: Use d’Alembert’s formula!)


**LECTURE 10. The Fourier Transform**

The next two lectures will give an introduction to the Fourier Transform. Throughout these lectures all our functions under consideration are a priori complex valued.

Formally, the Fourier Transform of a function $f: \mathbb{R}^n \to \mathbb{C}$ is the function $\hat{f}: \mathbb{R}^n \to \mathbb{C}$ defined by the formula.

\[
(10.1) \quad \mathcal{F}(f)(y) = \hat{f}(y) := \int_{\mathbb{R}^n} f(x) e^{-2\pi i \langle x, y \rangle} \, dx.
\]

I write “formally”, because actually making sense of (10.1) for suitable large class of functions and making sure that all the theorems one would like to be true really hold is a bit requires some work.

### 10.1 The Fourier Transform on $L^1$

**Proposition 10.2.** If $f \in L^1(\mathbb{R}^n)$, then $\hat{f} \in C^0(\mathbb{R}^n)$ and

\[
\|\hat{f}\|_{L^\infty} \leq \|f\|_{L^1}.
\]

In particular, the Fourier Transform defines a bounded linear map $\mathcal{F}: L^1(\mathbb{R}^n) \to C^0(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$.

**Proof.** To prove the estimate, note that for every $y \in \mathbb{R}^n$ we have

\[
|\hat{f}(y)| \leq \int_{\mathbb{R}^n} |f(x) e^{-2\pi i \langle x, y \rangle}| \, dx \leq \int_{\mathbb{R}^n} |f(x)| \, dx = \|f\|_{L^1};
\]

hence, $\|\hat{f}\|_{L^\infty} \leq \|f\|_{L^1}$.

To prove that $\hat{f}$ is continuous, note that for all $r > 0$ we have

\[
|\hat{f}(y) - \hat{f}(z)| \leq \int_{B_r(0)} |f(x)||e^{-2\pi i \langle x, y \rangle} - e^{-2\pi i \langle x, z \rangle}| + \int_{\mathbb{R}^n \setminus B_r(0)} |f(x)||e^{-2\pi i \langle x, y \rangle} - e^{-2\pi i \langle x, z \rangle}| \leq \|e^{-2\pi i \langle x, y \rangle} - e^{-2\pi i \langle x, z \rangle}\|_{L^\infty(B_r(0))}\|f\|_{L^1} + 2 \int_{\mathbb{R}^n \setminus B_r(0)} |f(x)|.
\]

For fixed $r > 0$, we have

\[
\lim_{z \to y} \|e^{-2\pi i \langle x, y \rangle} - e^{-2\pi i \langle x, z \rangle}\|_{L^\infty(B_r(0))} = 0.
\]

57
Thus for all \( r > 0 \), we have
\[
\lim_{z \to x} |\hat{f}(y) - \hat{f}(z)| \leq 2 \int_{\mathbb{R}^n \setminus B_r(0)} |f(x)|.
\]
Since
\[
\lim_{r \to \infty} \int_{\mathbb{R}^n \setminus B_r(0)} |f(x)| = 0,
\]
it follows that \( \hat{f} \) is continuous.

Proposition 10.2 has a slight improvement called the Riemann-Lebesgue Lemma.

**Lemma 10.3** (Riemann-Lebesgue Lemma). If \( f \in L^1(\mathbb{R}^n) \), then \( \hat{f} \) decays at infinity, that is,
\[
\lim_{r \to \infty} \left\| \hat{f} \right\|_{L^\infty(\mathbb{R}^n \setminus B_r(0))} = 0.
\]

**Exercise 10.4.** Prove the Riemann-Lebesgue Lemma. *Hint:* Use that smooth compactly supported function are dense in \( L^1 \).

Possibly the most important property of the Fourier Transform is that it “diagonalises” differentiation.

**Proposition 10.5.** If \( f, \partial_{x_k} f \in L^1(\mathbb{R}^n) \), then
\[
\hat{\partial_{x_k} f}(y) = 2\pi i y_k \hat{f}(y).
\]
Conversely, if \( f, -2\pi i x_k f \in L^1(\mathbb{R}^n) \), then \( \hat{f} \) is differentiable in the direction of \( y_k \) and
\[
\partial_{y_k} \hat{f}(y) = \mathcal{F}(-2\pi i x_k f)(y).
\]

**Proof.** To verify the first assertion we compute
\[
\hat{\partial_{x_k} f}(y) = \int \partial_{x_k} f(x) e^{-2\pi i (x,y)} \, dx
\]
\[
= 2\pi i y_k \int f(x) e^{-2\pi i (x,y)} \, dx.
\]

To prove the second assertion observe that, by the hypothesis,
\[
f(x) \partial_{y_k} e^{-2\pi i (x,y)} = -2\pi i x_k f(x) e^{-2\pi i (x,y)}
\]
is absolutely integrable, that is, in \( L^1 \). By **Proposition D.2**, it follows that \( \hat{f} \) is differentiable in the direction of \( y_k \) and
\[
\partial_{y_k} \hat{f}(y) = \int f(x) \partial_{y_k} e^{-2\pi i (x,y)}
\]
\[
= \int -2\pi i x_k f(x) e^{-2\pi i (x,y)} = \mathcal{F}(-2\pi i x_k f)(y).
\]
Another important property of the Fourier Transform is the way in which it interacts with convolutions. Recall that

\[(f \ast g)(y) := \int f(y - x)g(x) \, dx;\]

and if \(f, g \in L^1(\mathbb{R}^n)\), then \(f \ast g \in L^1(\mathbb{R}^n)\); in fact,

\[\|f \ast g\|_{L^1} \leq \|f\|_{L^1}\|g\|_{L^1}.\]

**Proposition 10.6.** If \(f, g \in L^1(\mathbb{R}^n)\), then

\[\hat{f} \ast \hat{g} = \hat{f} \cdot \hat{g}.\]

**Proof.** We compute

\[
\begin{align*}
\hat{f} \ast \hat{g}(x) &= \int \left( \int f(y - z)g(z) \, dz \right) e^{-2\pi i \langle x, y \rangle} \, dy \\
&= \int \left( \int f(y - z)e^{-2\pi i \langle x, y - z \rangle} \cdot g(z)e^{-2\pi i \langle x, z \rangle} \, dy \, dz \\
&= \hat{f}(x) \cdot \hat{g}(x). \quad \blacksquare
\end{align*}
\]

### 10.2 Fourier inversion on Schwartz functions

Formally, inverse Fourier Transform of a function \(g : \mathbb{R}^n \to \mathbb{C}\) is the function \(\hat{g} : \mathbb{R}^n \to \mathbb{C}\) defined by

\[\mathcal{F}^{-1}(g)(y) = \hat{g}(y) := \int_{\mathbb{R}^n} g(x)e^{2\pi i \langle x, y \rangle} \, dx.\]

There are various heuristic reasons why \(\mathcal{F}^{-1}\) should be the inverse of \(\mathcal{F}\). However, \(\mathcal{F} : L^1 \to L^\infty\) and, a priori, (10.7) does not make any sense if we just know that \(g \in L^\infty(\mathbb{R}^n)\). This can be addressed by working with a suitable function space with respect to which the Fourier Transform behaves in a “more symmetric way”.

**Definition 10.8.** A function \(f \in C^\infty(\mathbb{R}^n)\) is called a **Schwartz function** if for all \(k, \ell \in \mathbb{N}_0\)

\[\sup_{x \in \mathbb{R}^n} |x|^k |\nabla^\ell f| < \infty.\]

The set of all Schwartz functions on \(\mathbb{R}^n\) is denote by \(\mathcal{S}(\mathbb{R}^n)\).

**Proposition 10.9.** The Fourier transform defines a linear map \(\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)\).
Proof. This is a consequence of Proposition 10.5. \hfill \square

**Theorem 10.10** (Fourier inversion). If \( f \in S(\mathbb{R}^n) \), then
\[
\mathcal{F}^{-1} \circ \mathcal{F}(f) = f \quad \text{and} \quad \mathcal{F} \circ \mathcal{F}^{-1}(f) = f.
\]

Let me first tell you how a non-rigorous proof would go. Just by writing out the definitions we have
\[
\mathcal{F}^{-1} \circ \mathcal{F}(f)(x) = \int \int f(z) \cdot e^{-2\pi i (z-x,y)} \, dz \, dy.
\]
Interchanging the integral we get
\[
= \int f(z) \left( \int e^{-2\pi i (z-x,y)} \, dy \right) \, dz.
\]
The integral in parenthesis can be computed to be the distribution \( \delta(y-x) \). Thus the entire expression is \( f * \delta = f \).

There are two (related) difficulties with this argument: first one cannot exchange the order of integration because \( |f(z) \cdot e^{-2\pi i (z,y-x)}| \) is not integrable on \( \mathbb{R}^n \times \mathbb{R}^n \), second saying that integral in parenthesis is \( \delta(y-x) \) does not make sense. In fact, the latter is really just a restatement of what needs to be proved.

We will be able to salvage this argument. The key idea is to “regularise” \( e^{-2\pi i (z,y-x)} \). Before we can get started with this we need to compute the Fourier Transform of a Gaussian.

**Proposition 10.11** (Fourier Transform of a Gaussian). For all \( a > 0 \), we have
\[
\hat{e}^{-a\pi |\cdot|^2} = a^{-\frac{n}{2}} e^{-\frac{\pi}{a} |\cdot|^2}.
\]

**Proof.** For \( \lambda > 0 \) and \( f \in L^1(\mathbb{R}^n) \), we have
\[
\hat{f}(\lambda \cdot) = \lambda^{-n} \hat{f}(\cdot/\lambda).
\]
Thus it suffices verify the assertion for \( a = 1 \). Moreover, by Fubini, it suffices to consider \( n = 1 \).

We compute
\[
\hat{e}^{-\pi (\cdot)^2}(y) = \int_{\mathbb{R}} e^{-\pi x^2 + 2\pi i xy} \, dx = e^{-\pi y^2} \int_{\mathbb{R}} e^{-\pi (x+iy)^2} \, dx.
\]
A simple way to evaluate the last integral is to note that $e^{-\pi z^2}$ has no pole on any strip of the form $\{z \in \mathbb{C} : \text{Im}(z) \in [a, b]\}$ thus

$$\int_{\mathbb{R}+iy} e^{-\pi z^2} \, dz$$

does not depend on $y$, and for $y = 0$ we know the integral to be one. \hfill \square

**Proof of Theorem 10.10.** Given $f \in \mathcal{S}(\mathbb{R}^n)$, we need to show that

$$\mathcal{F}^{-1} \circ \mathcal{F}(f)(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(z) \cdot e^{-2\pi i (z-x,y)} \, dz \, dy = f(x).$$

The function

$$f(z) \cdot e^{-2\pi i (z-x,y) - 4\pi^2 t|y|^2}$$

is absolutely integrable over $\mathbb{R}^n \times \mathbb{R}^n$, thus by Fubini’s theorem

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(z) \cdot e^{-2\pi i (z-x,y) - 4\pi^2 t|y|^2} \, dz \, dy = \int_{\mathbb{R}^n} f(z) \left( \int_{\mathbb{R}^n} e^{-2\pi i (z-x,y) - 4\pi^2 t|y|^2} \, dy \right) \, dz.$$

By Proposition 10.11, the integral in parenthesis is $K_t(z-x)$ with

$$K_t(\cdot) := \frac{1}{(4\pi t)^{n/2}} e^{-|\cdot|^2/4t}.$$

(Recall that this is the heat kernel on $\mathbb{R}^n$.) Thus the above integral is

$$(f \ast K_t)(z).$$

In the lecture on the heat kernel we saw that

$$\lim_{t \to 0} (f \ast K_t)(z) = f(z).$$

We also have

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(z) \cdot e^{-2\pi i (z-x,y) - 4\pi^2 t|y|^2} \, dz \, dy = \int_{\mathbb{R}^n} \hat{f}(y) e^{2\pi i (x,y) - 4\pi^2 t|y|^2} \, dy.$$

Since $\hat{f}$ is integrable, by Theorem D.1, we have

$$\lim_{t \to 0} \int_{\mathbb{R}^n} \hat{f}(y) e^{2\pi i (x,y) - 4\pi^2 t|y|^2} \, dy = \mathcal{F}^{-1} \circ \mathcal{F}(f)(x).$$

\hfill \square
LECTURE 11. The Fourier Transform (continued)

11.1 Plancherel’s theorem

Theorem 11.1 (Plancherel’s theorem). If \( f, g \in S(\mathbb{R}^n) \), then

\[
\langle f, g \rangle_{L^2} = \langle \hat{f}, \hat{g} \rangle_{L^2}
\]

Proof. By Theorem 10.10, we can assume that \( g = \tilde{h} \). The assertion is then equivalent to

\[
\langle f, \tilde{h} \rangle_{L^2} = \langle \hat{f}, \hat{h} \rangle_{L^2}.
\]

Using the fact that \(|f(x)||h(y)|\) is integrable, we compute

\[
\langle f, \tilde{h} \rangle_{L^2} = \int f(x) \tilde{h}(x) \, dx = \int \int f(x) \left( \int \tilde{h}(y) e^{-2\pi i (x,y)} \, dy \right) \, dx = \int \left( \int f(x) e^{-2\pi i (x,y)} \, dx \right) \cdot \tilde{h}(y) \, dy,
\]

\[
= \int \hat{f}(y) \tilde{h}(y) \, dy = \langle \hat{f}, \hat{h} \rangle_{L^2}.
\]

Proposition 11.2. There is a unique bounded linear map \( \mathcal{F} : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n) \) which agrees with the Fourier Transform on \( S(\mathbb{R}^n) \). This map is an isometry.

Proof. The set \( S(\mathbb{R}^n) \) is dense in \( L^2(\mathbb{R}^n) \). Thus for any \( f \in L^2(\mathbb{R}^n) \) we can pick a sequence \((f_i)\) in \( S(\mathbb{R}^n) \) which converges to \( f \) in \( L^2 \). By Theorem 11.1,

\[
\| \mathcal{F}(f_i) - \mathcal{F}(f_j) \|_{L^2} = \| f_i - f_j \|_{L^2}.
\]

Therefore \( \mathcal{F}(f_i) \) is a Cauchy sequence and has a limit

\[
\lim_{i \to \infty} \mathcal{F}(f_i) \in L^2.
\]

If \( (\tilde{f}_i) \) were another sequence in \( S(\mathbb{R}^n) \) converging to \( f \), then

\[
\lim_{i \to \infty} \| \mathcal{F}(f_i) - \mathcal{F}(\tilde{f}_i) \|_{L^2} \leq \lim_{i \to \infty} \| f_i - f \|_{L^2} + \| \tilde{f}_i - f \|_{L^2} = 0.
\]
Thus the limit only depends on $f$. We define

$$\mathcal{F}(f) := \lim_{i \to \infty} \mathcal{F}(f_i) \in L^2.$$ 

The linearity of $\mathcal{F}$ on $\mathcal{S}(\mathbb{R}^n)$ immediately implies the linearity of the extension to $L^2(\mathbb{R}^n)$.

It follows from Theorem 11.1 that $\mathcal{F}$ is an isometry. 

### 11.2 Tempered distributions

Any continuous function $f : \mathbb{R}^n \to \mathbb{C}$ is completely determined by the linear map $T_f : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$

$$T_f(\phi) := \int_{\mathbb{R}^n} f \phi.$$ 

In the proof of Theorem 11.1 we saw that if $f, g \in \mathcal{S}(\mathbb{R}^n)$, then

$$\int_{\mathbb{R}^n} f \hat{g} = \int_{\mathbb{R}^n} \hat{f} g.$$ 

Thus if $f \in \mathcal{S}(\mathbb{R}^n)$, then

$$T_f(\hat{\phi}) = T_f(\hat{\phi}).$$

Note that the right-hand side of this equation makes sense, even if $f$ is not in $\mathcal{S}(\mathbb{R}^n)$. These observations leads us to the definition of tempered distributions and vast extension of the applicability of the Fourier Transform.

Roughly speaking a tempered distribution is an element in “the dual space” of $\mathcal{S}(\mathbb{R}^n)$. However, since $\mathcal{S}(\mathbb{R}^n)$ is an infinite dimensional vector space one has to be a bit careful with what one means by “the dual space”.

**Definition 11.3.** A tempered distribution $T$ is a linear map $T : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ with the following continuity property: If $(\phi_i)$ is a sequence in $\mathcal{S}(\mathbb{R}^n)$ such that for all $k, \ell \in \mathbb{N}_0$

$$\lim_{i \to \infty} \sup_{x \in \mathbb{R}^n} |x|^k |\nabla^\ell \phi_i| = 0,$$

then

$$\lim_{i \to \infty} T(\phi_i) = 0.$$ 

We write $\mathcal{S}^*(\mathbb{R}^n)$ for the space of all tempered distributions on $\mathbb{R}^n$.

**Example 11.4.** Suppose $f : \mathbb{R}^n \to \mathbb{C}$ is slowly growing, that is, there is some $k \in \mathbb{N}_0$ and $c > 0$ such that

$$|f(x)| \leq c(1 + |x|)^k.$$
Then
\[ T_f(\phi) := \int_{\mathbb{R}^n} f \phi \]
defines a tempered distribution.

If no confusion can arise, we will often just write \( f \) for the distribution \( T_f \) defined by \( f \).

**Example 11.5.** The “\( \delta \)–function”
\[ \delta(\phi) := \phi(0) \]
is a tempered distribution.

**Definition 11.6.** If \( T \) is a tempered distribution, we define its Fourier Transform and inverse Fourier Transform by
\[ \mathcal{F}(T) = \hat{T} := T \circ \mathcal{F} \quad \text{and} \quad \mathcal{F}^{-1}(T) = \check{T} := T \circ \mathcal{F}^{-1}. \]

Thus we extend the Fourier Transform and its inverse to linear maps
\[ \mathcal{F} : \mathcal{S}^*(\mathbb{R}^n) \to \mathcal{S}^*(\mathbb{R}^n) \quad \text{and} \quad \mathcal{F}^{-1} : \mathcal{S}^*(\mathbb{R}^n) \to \mathcal{S}^*(\mathbb{R}^n). \]

By Theorem 10.10 we have
\[ \mathcal{F} \circ \mathcal{F}^{-1} = \text{id}_{\mathcal{S}^*(\mathbb{R}^n)} \quad \text{and} \quad \mathcal{F}^{-1} \circ \mathcal{F} = \text{id}_{\mathcal{S}^*(\mathbb{R}^n)}. \]

Note that “trick” allows us to make sense of the Fourier Transform even for functions which are slowly growing and for which (10.1) wouldn’t make any sense at all.

Two key identities for the Fourier Transform are
\[ \mathcal{F}(\delta) = 1 \quad \text{and} \quad \mathcal{F}(1) = \delta. \]
The first is trivial. The second is not at all trivial, however; the bulk of the proof of Theorem 10.10 was proving this identity.

Most things one can do with functions one can also do with tempered distributions:

**Definition 11.7 (Differentiation).** Suppose \( T \) is a tempered distribution on \( \mathbb{R}^n \). We define its derivative by \( x_k \) as
\[ (\partial_{x_k} T)(\phi) := T(-\partial_{x_k} \phi). \]
Definition 11.8 (Multiplication by a function). Suppose $T$ is a tempered distribution on $\mathbb{R}^n$. If $f$ is smooth and all of its derivatives (of arbitrary order) are slowly growing, then we define

$$(fT)(\phi) := T(f\phi).$$

Definition 11.9 (Convolution). Suppose $T$ is a tempered distribution on $\mathbb{R}^n$ and $f \in S(\mathbb{R}^n)$. We define $T \ast f \in C^\infty(\mathbb{R}^n)$ by

$$(T \ast f)(\phi) = T(f(-\cdot) \ast \phi).$$

With these definitions in place one can readily see that all the identities we derived in the previous lecture for the Fourier Transform on $S(\mathbb{R}^n)$ carry over to $S^*(\mathbb{R}^n)$.

11.3 A derivation of the Heat Kernel

Suppose $u: [0,T] \times \mathbb{R}^n \to \mathbb{R}$.

$$\begin{align*}
\partial_t u + \Delta u &= 0 \\
u(0, \cdot) &= f.
\end{align*}$$

Denote by $\hat{u}$ the Fourier Transform of $u$ in the spacial directions. Then we know that

$$\begin{align*}
\partial_t \hat{u} + 4\pi^2 |x|^2 \hat{u} &= 0 \\
\hat{u}(0, \cdot) &= \hat{f}.
\end{align*}$$

Thus

$$\hat{u}(t, \cdot) = e^{-4\pi^2 |x|^2 t} \hat{f}(\cdot).$$

By Proposition 10.11,

$$\mathcal{F}^{-1}(e^{-4\pi^2 |x|^2 t}) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t} =: K_t(x).$$

Thus by Proposition 10.6 and Theorem 10.10

$$u(t, \cdot) = \mathcal{F}^{-1}(\hat{K}_t \cdot \hat{f}) = K_t \ast f.$$

Exercise 11.10. Find the fundamental solution of Schrödinger’s equation

$$- i \partial_t \psi + \Delta \psi = 0.$$
LECTURE 12. Introduction Curve-Shortening Flow

There are no notes for this lecture, and the content of this lectures is not examinable.
Lecture 13. Laplace’s and Poisson’s equations

Fix an open subset $U \subset \mathbb{R}^n$.

Definition 13.1. Laplace’s equation is the PDE

\begin{equation}
\Delta u = 0
\end{equation}

for a function $u: U \to \mathbb{R}$. A function $u: U \to \mathbb{R}$ is called harmonic if it satisfies (13.2).

Definition 13.3. Given $f: U \to \mathbb{R}$, Poisson’s equation is the PDE

\begin{equation}
\Delta u = g
\end{equation}

for a function $u: U \to \mathbb{R}$.

Remark 13.5. We can think of any solution to (13.2) and (13.4) as a steady state solution of an (inhomogeneous) heat or wave equation.

Note that since (13.2) and (13.4) do not depend on time, do not have to/cannot specify initial conditions. You might think that it is boring to study equations without time-dependence. However, many situations are inherently time-independent and (versions of) (13.2) and (13.4) can be used to great profit. Even if you really only care about time-dependent PDE, studying the Laplace operator $\Delta$ in depth will be important: note, for example, how the solution to the heat equation is basically equivalent to understanding all possible solutions $(u, \lambda)$ of the eigenvalue problem

\[ \Delta u = \lambda^2 u. \]

13.1 Dirichlet’s principle

Suppose $U$ is a bounded domain in $\mathbb{R}^n$ with smooth boundary and $f: \partial U \to \mathbb{R}$ smooth. Consider the space

\[ C^\infty_f(\bar{U}) := \{ u : u|_{\partial U} = f \}. \]

We can think of $u \in C^\infty_f(U)$ as describing the dislocation of a membrane over $U$ which is supported by a boundary frame with dislocation given by $f$. A model for the energy of such a membrane is

\[ E(u) := \frac{1}{2} \int_U |\nabla u|^2. \]
Physics tells us that the physically realised membrane $u$ should minimise $E(u)$. Suppose $u$ minimises $E(u)$. Then for each variation $\phi \in C^\infty_0(\bar{U})$ we have

$$0 = \partial_t E(u + t\phi)|_{t=0} = \int_U \langle \nabla u, \nabla \phi \rangle = \int_U \langle \Delta u, \phi \rangle.$$ 

Thus we must have $\Delta u = 0$, for otherwise we could find a $\phi$ for which the right-hand side of the above does not vanish.

This shows that if there is a smooth minimiser $u \in C^\infty(\bar{U})$ of $E$, then it must be harmonic. More generally, solutions of

$$\Delta u = g$$

are characterized by minimizing the functional

$$E(u) := \int_U \frac{1}{2} |\nabla u|^2 - u \cdot g.$$ 

We say that harmonic functions can be characterised by a variational principle. Finding such a variational principle for a PDE is extremely useful because finding minima of a functional is often easier than solving a PDE. We will pick up this thread later in this lecture.

### 13.2 Connection with holomorphic functions

**Definition 13.6.** Let $U \subset \mathbb{C}$ be a domain in the complex plane. A complex-valued function $f : U \to \mathbb{C}$ is called holomorphic at $z_0 \in U$ if the limit

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists. It is called holomorphic if it is holomorphic at points in $U$.

**Exercise 13.7.** A complex-valued function $f : U \to \mathbb{C}$ is holomorphic if and only if it satisfies the Cauchy–Riemann equation

$$\frac{\partial f}{\partial \bar{z}} = 0$$

with

$$\frac{\partial f}{\partial \bar{z}} := \partial_x f + i \partial_y f = (\partial_x u - \partial_y v) + i(\partial_y u + \partial_x v).$$

Here $u := \text{Re } f$ and $v := \text{Im } f$, and $x := \text{Re } z$ and $y = \text{Re } z$.

**Remark 13.9.** The Cauchy–Riemann equation is a system of first order PDE on the vector-valued function $(u, v)$. 

68
I hope you recall from your class on complex analysis that the theory of holomorphic functions is incredibly rich. In particular, there are a great number of holomorphic functions and they enjoy spectacular regularity properties.

**Proposition 13.10.** If \( f: U \rightarrow \mathbb{R} \) is holomorphic, then both \( u := \text{Re } f \) and \( v := \text{Im } f \) are harmonic.

**Proof.** Holomorphic functions are \( C^2 \)—in fact, analytic. By (13.8),
\[
\partial_x u = \partial_y v \quad \text{and} \quad \partial_y u = -\partial_x v.
\]
Thus
\[
\Delta u = -\partial_x^2 u - \partial_y^2 u = -\partial_x \partial_y v + \partial_y \partial_x v = 0.
\]
(In the last step we used Schwarz’s theorem that for a \( C^2 \) function partial derivatives commute.)

**Remark 13.11.** The converse of this is also true: every harmonic function \( u: U \rightarrow \mathbb{R} \) on a domain \( U \subset \mathbb{C} \) is the real part of a holomorphic function provided \( U \) is simply-connected. A nice explanation of this fact that uses only complex analysis can be found in this answer of Christian Blatter’s to a question on MSE.

### 13.3 Uniqueness via the energy method

**Theorem 13.12.** Suppose \( U \) is bounded with \( C^1 \) boundary. There is at most one function \( u: C^2(\overline{U}) \cap C^1(U) \rightarrow \mathbb{R} \) satisfying Poisson’s equation (13.4) satisfying Dirichlet, Robin or mixed boundary conditions (cf. Section 5.1). For Neumann boundary conditions any two solutions differ by a locally constant function.

**Proof.** The difference \( w \) satisfies Laplace’s equation with homogeneous boundary conditions and integration by parts shows that
\[
0 = \int_U (\Delta w, w) \geq \int_U |\nabla w|^2.
\]
It follows that \( \nabla w = 0 \).

### 13.4 The weak maximum principle

**Theorem 13.13.** Suppose \( U \) is bounded and \( f \in C^2(\overline{U}) \) satisfies
\[
\Delta u \leq 0,
\]
then
\[
\max_{\overline{U}} u = \max_{\partial U} u.
\]

Exercise 13.15. Suppose \( f \in C^2(\mathbb{R}^n) \) is harmonic and “decays at infinity” in the sense that
\[
\lim_{r \to \infty} \sup_{\mathbb{R}^n \setminus B_r(0)} |f| = 0.
\]
Show that \( f = 0 \).

13.5 \( W^{2,2} \)-estimates

Theorem 13.16. For \( f \in C_0^\infty(\mathbb{R}^n) \), we have
\[
\|\nabla^2 f\|_{L^2} = \|\Delta f\|_{L^2}.
\]

Proof. This is a simple computation
\[
\|\nabla^2 f\|_{L^2}^2 = \sum_{i,j=1}^n \int_{\mathbb{R}^n} (\partial_i \partial_j f)(\partial_i \partial_j f)
\]
\[
= \sum_{i,j=1}^n \int_{\mathbb{R}^n} (\partial_i \partial_j f)(\partial_j \partial_i f)
\]
\[
= \sum_{i,j=1}^n \int_{\mathbb{R}^n} (\partial_j \partial_i \partial_j f)(\partial_i f)
\]
\[
= \sum_{i,j=1}^n \int_{\mathbb{R}^n} (-\partial_j \partial_i \partial_j f)(\partial_i f)
\]
\[
= \sum_{i,j=1}^n \int_{\mathbb{R}^n} (\partial_i f)(-\partial_j \partial_i f)
\]
\[
= \|\Delta f\|_{L^2}.
\]

Remark 13.17. Although this result has a very simple proof, it is surprising that the size of the Hessian of \( f \) can be controlled just in terms of the Laplacian (at least when the size is measured in \( L^2 \)).

13.6 Dirichlet’s principle (cont.)

Can we construct such a minima for \( E \)? Let us try the direct method \( E \) is bounded below. Thus \( E_0 := \inf \{ E(u) : u \in C_0^\infty(U) \} \geq 0 \) (in particular, it is finite) and we can find a sequence \( (u_i) \) such that
\[
E(u_i) \to E_0.
\]
If there were a \( u_\infty \) such that (some subsequence) of \((u_i)\) converges to \( u_\infty \) in \( C^\infty \), then \( E(u_\infty) = E_0 \). The problem is that \textit{a priori} such a \( u_\infty \) does not need to exist, and even if it existed the minimising sequence will typically not converge in \( C^\infty \).

Nevertheless, one can make this idea rigorous. First of all note that it suffices to consider the case of homogeneous boundary conditions, that is, \( f = 0 \). We introduce the following space of functions with zero boundary values.

**Definition 13.18.** If \( u \in C^\infty(U) \), set
\[
\| u \|_{W^{1,2}}^2 := \int_U |\nabla u|^2 + |u|.
\]
Denote by \( W^{1,2}(U) \) and \( W^{1,2}_0(U) \) the completion of \( C^\infty(U) \) and \( C^\infty_0(U) \) with respect to \( \| \cdot \|_{W^{1,2}} \) respectively.

**Remark 13.19.** In the next lecture we will proof the Dirichlet–Poincaré inequality, which says that if \( U \) is bounded and \( u \in C^\infty_0(U) \), then
\[
\int_U |u|^2 \leq c \int_U |\nabla u|^2
\]
with a constant \( c > 0 \) depending only on \( U \). As a consequence of this we might equally well use the norm
\[
\| u \| := \int_U |\nabla u|^2
\]
on \( W^{1,2}_0(U) \).

Now, consider the functional
\[
E(u) := \int_U \frac{1}{2} |\nabla u|^2 - u g.
\]

**Proposition 13.20.** A function \( u \in W^{1,2}_0(U) \) satisfies
\[
E(u) = \inf \{ E(v) : v \in W^{1,2}_0(U) \}
\]
if and only if
\[
\Delta u = g
\]
in the weak sense, that is, for all \( \phi \in W^{1,2}_0(U) \)
\[
\int_U (\Delta u) \phi = \int_U g \phi.
\]
Proof. $E : W^{1,2}_0(U) \to \mathbb{R}$ is a smooth map and its derivative at $u$ is

$$dE_u[\dot{u}] = \int_U (\Delta u - g) \dot{u}.$$ 

If $u$ minimizes $E_\rho$, this vanishes for all $\dot{f}u \in W^{1,2}_0(U)$; hence $u$ satisfies $\Delta u = g$ in the weak sense.

Conversely, if $u$ satisfies $\Delta u = g$ in the weak sense and $v \in W^{1,2}_0(U)$, then

$$\int_U |\nabla u|^2 - ug = \int_U (\Delta u)u - ug$$

$$= \int_U (\Delta u)v - vg$$

$$= \int_U \langle \nabla u, \nabla v \rangle - vg$$

$$\leq \int_U \frac{1}{2} |\nabla u|^2 + \frac{1}{2} |\nabla v|^2 - vg.$$ 

Thus $E(u) \leq E(v)$.

Let $u_i$ be a minimising sequence, i.e.,

$$\inf \{ E(v) : v \in W^{1,2}_0(U) \} = \lim_{i \to \infty} E(u_i).$$

By the Dirichlet–Poincaré inequality, for some $c > 0$ and all $\varepsilon > 0$, we have

$$\|u_i\|_{W^{1,2}} \leq cE(u_i) + c \int_U |u||g| \leq cE(u_i) + c\varepsilon \|u\|_{L^2}^2 + c\varepsilon^{-1}\|g\|.$$ 

Taking $\varepsilon = 1/2c$, we get

$$\|u_i\|_{W^{1,2}}^2 \leq 2cE(u_i) + 4c^2\|g\|.$$ 

In particular, $\|u_i\|_{W^{1,2}}$ is bounded independent of $i$.

Then $f_i$ converges weakly in $H_0^1$ to a limit $f$. That is for all $g \in H_0^1$,

$$\langle f_i, g \rangle \to \langle f, g \rangle.$$ 

Proposition 13.21 (sequential weak-* compactness). Let $H$ be a separable Hilbert space. Let $(x_i) \in H^{2N}$ be a sequence with $\lim \inf \|x_i\|_H < \infty$. Then after passing to a subsequence, $(x_i)$ has a weak-limit, i.e., there exists a $x \in H$ such that for all $y \in H$

$$\langle x, y \rangle = \lim_i \langle x_i, y \rangle.$$ 

72
Proof. This is (a consequence of) the sequential Banach–Alaoglu Theorem which states that the closed unit ball of the dual space of a separable normed vector space is sequentially compact in the weak-* topology and the fact that Hilbert spaces are isomorphic to their dual spaces. (See Theorem 3.2.1 in Bühler–Salamon’s Lecture Notes on Functional Analysis for a proof.)

Remark 13.22. There is also a non-sequential Banach–Alaoglu Theorem which states that the unit ball of a dual space of a separable normed vector space is compact in the weak-* topology. This follows easily from Tychonoff’s theorem stating that arbitrary(!) products of compact spaces are compact. However, the proof of Tychonoff’s theorem uses the Axiom of Choice whose use one might want to avoid to construct something as concrete as a solution to a PDE.

This allows us to take some limit \( u \in W^{1,2}_0(U) \) of the minimizing sequence \( u_i \). The following guarantees that \( u \) actually minimizes \( E \).

**Proposition 13.23** (weak-* lower semi-continuity of \( E \)). Suppose \( u_i \in (W^{1,2}_0(U))^N \) weakly converges to \( u \in W^{1,2}_0(U) \). Then

\[
E(u) \leq \liminf_{i \to \infty} E(u_i).
\]

Proof. By weak convergence

\[
\int_U |\nabla u|^2 - gu = \lim_i \langle \nabla u_i, \nabla u \rangle - gu_i \\
\leq \liminf_i \int \frac{1}{2} |\nabla u|^2 + \frac{1}{2} |\nabla u_i|^2 - gu_i.
\]

This implies the assertion.

It follows that \( u \) really does minimize \( E \) and thus satisfies

\[
\Delta u = g
\]

in the weak sense. In one of the following lectures we when \( u \) satisfies the Poisson equation in the strong sense.

Remark 13.24. Note that the above argument works with very little regularity on \( g \). In fact, it suffices that \( g \) define an element of the dual space of \( W^{1,2}_0(U) \).
LECTURE 14. The Poincaré inequalities

Let $U$ be a bounded open subset of $\mathbb{R}^n$.

14.1 Dirichlet–Poincaré inequality

**Theorem 14.1** (Dirichlet–Poincaré inequality). There is a constant $c > 0$ such that for all $u \in W^{1,2}_0(U)$ we have

$$\int_U |u|^2 \leq c \int_U |
abla u|^2. \tag{14.2}$$

**Proof.** By density it suffices to prove this for $u \in C_0^\infty(U)$. Moreover, it suffices to prove this for $U = Q := (0, 1)^n$, since after rescaling we can assume that $U \subset Q$ and $C_0^\infty(U) \subset C_0^\infty(Q)$.

Since

$$u(x_1, x_2, \ldots, x_n) = \int_0^{x_1} \partial_1 u(s, x_2, \ldots, x_n) \, ds,$$

we have

$$|u(x_1, x_2, \ldots, x_n)| \leq \int_0^1 |\partial_1 u(s, x_2, \ldots, x_n)| \, ds.$$

By Cauchy–Schwarz,

$$\int_0^1 |\partial_1 u(s, x_2, \ldots, x_n)| \, ds \leq \left( \int_0^1 |\partial_1 u(s, x_2, \ldots, x_n)|^2 \, ds \right)^{1/2}.$$

Thus

$$|u(x_1, \ldots, x_n)|^2 \leq \int_0^1 |\partial_1 u(s, x_2, \ldots, x_n)|^2 \, ds$$

Integrating over $x_1$ yields

$$\int_0^1 |u(x_1, \ldots, x_n)|^2 \, dx_1 \leq \int_0^1 |\partial_1 u(x_1, x_2, \ldots, x_n)|^2 \, dx_1,$$

and a further integration over $x_2, \ldots, x_n$ yields the Dirichlet–Poincaré inequality. \qed

**Exercise 14.3.** Show that there is a constant $c > 0$ (depending only on $n$) such that for all $r > 0$ and all $u \in W^{1,2}_0(B_r(0))$ we have

$$\int_{B_r(0)} |u|^2 \leq cr^2 \int_{B_r(0)} |
abla u|^2.$$
**Exercise 14.4.** Prove the $L^p$ Dirichlet–Poincaré inequality: For each $p \in [1, \infty)$ there exists a constant $c > 0$ such that for all $u \in C_0^\infty(U)$ we have

$$\int_U |u|^p \leq c \int_U |\nabla u|^p.$$  

*Hint:* Adapt the proof of Theorem 14.1 by using Hölder’s inequality

$$\int_U fg \leq \left( \int_U |f|^p \right)^{1/p} \left( \int_U |g|^q \right)^{1/q}$$

for satisfying $1 = \frac{1}{p} + \frac{1}{q}$.

*Remark 14.5.* The best constant for the Dirichlet–Poincaré inequality on $U$ is the number

$$c_D(U) := \sup_{u \in C_0^\infty(U)} \frac{\int_U |u|^2}{\int_U |\nabla u|^2}.$$  

Since for $u \in C_0^\infty(U)$ we have

$$\int_U |\nabla u|^2 = \int_U \langle \Delta u, u \rangle,$$

$$\frac{1}{c_D(U)} = \lambda_1^D(U) := \inf_{u \in C_0^\infty(U)} \frac{\int_U \langle \Delta u, u \rangle}{\int_U |u|^2}.$$  

It turns out that $\lambda_1^D$ is the smallest eigenvalue of $\Delta$ with Dirichlet boundary condition. This is why Theorem 14.1 is called the Dirichlet–Poincaré inequality.

**Exercise 14.6.** Find an upper bound for $c(U)$ in terms of the geometry of $U$.

If one does not impose that $u$ vanishes on the boundary of $U$, then (14.2) cannot hold. Because locally constant functions violate this bound. However this is the only problem.

### 14.2 Neumann–Poincaré inequality

**Theorem 14.7** (Neumann–Poincaré inequality). Suppose $U$ is connected. There is a constant $c > 0$ such that for all $u \in W^{1,2}(U)$ we have

$$\int_U |u - \bar{u}|^2 \leq c \int_U |\nabla u|^2.$$  

Here

$$\bar{u} := \int_U u.$$
Proof for $U = B_r(0)$. It suffices to consider $B_1 := B_1(0)$. We compute

$$\int_{B_1} |u - \bar{u}|^2 = \int_{B_1} \left( \int_{B_1} |u(x) - u(y)| \, dy \right)^2 \, dx \leq \text{vol}(B_1)^{-1} \int_{B_1} \int_{B_1} |u(x) - u(y)|^2 \, dy \, dx.$$ 

Since the straight-line connection $x, y \in B_1$ is contained in $B_1$, we have

$$|u(x) - u(y)| \leq \int_0^1 |\partial_t u(x + t(y - x))| \, dt \leq |x - y| \int_0^1 |(\nabla u)(x + t(y - x))| \, dt \leq 2 \left( \int_0^1 |(\nabla u)(x + t(y - x))|^2 \, dt \right)^{1/2}.$$ 

Combining both observations using the symmetry between $x$ and $y$, and substituting $z := ty + (1 - t)x$ we have

$$\int_{B_1} |u - \bar{u}|^2 \leq \frac{4}{\text{vol}(B_1)} \int_{B_1} \int_{B_1} \int_0^1 |(\nabla u)(x + t(y - x))|^2 \, dt \, dy \, dx = \frac{8}{\text{vol}(B_1)} \int_{B_1} \int_{B_1} \int_0^{1/2} |(\nabla u)(x + t(y - x))|^2 \, dt \, dy \, dx = \frac{8}{\text{vol}(B_1)} \int_0^{1/2} \int_{B_1} \frac{1}{1 - t} \int_{B_1 - t(y)} |(\nabla u)(z)|^2 \, dz \, dy \, dt \leq 8 \int_{B_1} |(\nabla u)(z)|^2.$$

Proof sketch for general $U$. The typical proof of this result for general $U$ is more involved and proceeds by a contradiction argument. I don’t expect you to know this proof, but for completeness let me give a outline of the argument.

Suppose there is no such constant. Then there exists a sequence $u_k \in W^{1,2}(U)$ such that

$$\int_U |u_k - \bar{u}_k|^2 \geq k \int_U |\nabla u_k|^2.$$ 

There is no loss in assuming that $\bar{u}_k = 0$ and $\int_U |u_k|^2 = 1$. Then $|\nabla u_k|_{L^2} \to 0$. Thus one the one hand $|u_k|_{W^{1,2}}$ is uniformly bounded and thus a subsequence has
a weak-* $W^{1,2}$ limit $u$, which one can show satisfies $\nabla u = 0$ and thus is constant. On the other hand Rellich’s Theorem asserts that a further subsequence has a $L^2$ limit, which must also be $u$, and thus

$$\bar{u} = 0 \quad \text{and} \quad \|u\|_{L^2} = 1.$$ 

But this contradicts $u$ being constant. □

**Remark 14.9.** The best constant for the Neumann–Poincaré inequality on $U$ is the number

$$c_N(U) := \sup_{u \in C^\infty(U)} \frac{\int_U |u - \bar{u}|^2}{\int_U |\nabla u|^2}.$$ 

By Theorem A.4, for $u \in C^\infty(U)$ we have

$$\int_U |\nabla u|^2 = \int_U \langle \Delta u, u \rangle + \int_{\partial U} (\partial_\nu u) u$$

Thus if $\partial_\nu u = 0$ and $\Delta u = \lambda u$, then

$$\int_U |\nabla u|^2 = \lambda \int_U |u|^2.$$ 

Constants satisfy this identity with $\lambda = 0$. That is the smallest eigenvalue of the Laplacian with Neumann boundary conditions is zero. To find the next smallest eigenvalue $\lambda_1^N(U)$ we work $L^2$–orthogonal to the constants, i.e., we make the assumption that $\bar{u} = 0$:

$$\lambda_1^N(U) := \inf \left\{ \frac{\int_U |\nabla u|^2}{\int_U |u|^2} : u \in C^\infty(U), \partial_\nu u = 0, \bar{u} = 0 \right\}.$$ 

Then we have

$$\frac{1}{c_N(U)} \leq \lambda_1^N(U).$$

### 14.3 The Li–Schoen proof of the Dirichlet–Poincaré inequality

The above proof of Theorem 14.1 is the “standard proof”. Another proof of the $L^1$ Dirichlet–Poincaré inequality was discovered by Peter Li and Richard Schoen [LS84].

It suffices to consider the case $U = B_1$. 

---

5 I do not expect you to be able to reconstruct this proof.
Proposition 14.10. Fix \( y \in \partial B_2 \), and consider the function \( \rho : B_1 \to (0, \infty) \) defined by
\[
\rho(x) := |x - y|.
\]
Then
\[
ne^{-3n} \leq -\Delta e^{-n\rho} \quad \text{and} \quad |\nabla e^{-n\rho}| \leq ne^{-n\rho}.
\]

Proof. By a direct computation we have
\[
\Delta \rho = -(n - 1)\rho^{-1}
\]
and for \( c > 0 \)
\[
\Delta e^{-c\rho} = ce^{-c\rho}(-\Delta \rho - c|\nabla \rho|^2)
= ce^{-c\rho}(-\Delta \rho - c)
= ce^{-c\rho}((n - 1)\rho^{-1} - c)
\leq ce^{-c\rho}((n - 1) - c).
\]
Thus
\[
\Delta e^{-n\rho} \leq -ne^{-n\rho} \leq -ne^{-3n}.
\]
It is clear that \(|\nabla e^{-n\rho}| \leq ne^{-n\rho}|.

Using Theorem A.4 it follows that
\[
ne^{-3n} \int_{B_1} |u| \leq \int_{B_1} -\Delta e^{-n\rho}|u| \leq \int_{B_1} |\nabla e^{-n\rho}||\nabla u| \leq n \int_{B_1} |\nabla u|.
\]
This proves the \( L^1 \) Dirichlet–Poincaré inequality.

Remark 14.11. Note that the \( L^1 \) Dirichlet–Poincaré inequality implies all \( L^p \) Dirichlet–Poincaré inequalities:
\[
\int_U |u|^p \leq c \int_U |\nabla |u|^p| \leq cp \int_U |u|^{p-1}|\nabla u| \leq cp \left( \int_U |u|^p \right)^{\frac{p-1}{p}} \left( \int_U |\nabla u|^p \right)^{1/p}.
\]
(Note that \( \frac{1}{p} + \frac{p-1}{p} = 1 \).
Lecture 15. Mean-value properties of harmonic functions

Let $U$ be a bounded open subset of $\mathbb{R}^n$.

**Theorem 15.1** (Mean-value property). Let $u : U \to \mathbb{R}$ be a harmonic function. If $B_r(x) \subset U$, then

$$u(x) = \int_{\partial B_r(x)} u \quad \text{and} \quad u(x) = \int_{B_r(x)} u.$$

**Proof.** Note that the second identity follows from the first by integration.

To prove the first identity, we define $f : (0, r] \rightarrow \mathbb{R}$ by

$$f(s) := \int_{\partial B_s(x)} u.$$

By continuity of $u$, we have

$$\lim_{s \downarrow 0} f(s) = u(x).$$

The following computation shows that $f(s)$ is independent of $s$, thus proving the identity:

$$f'(s) = \int_{\partial B_s(x)} \partial_\nu u = -\frac{n}{r} \int_{B_s(x)} \Delta u = 0.$$

Here we use Theorem A.4 and that $\text{vol}(\partial B_r) = n\text{vol}(B_r)$.

**Remark 15.2.** Compare this with Theorem 7.9.

**Remark 15.3.** In the next lecture we will see that the converse is also true, see Theorem 16.12.

As we will see in the rest of this lecture, this simple observation has remarkable consequences.

### 15.1 Strong maximum principle

**Theorem 15.4** (Strong maximum principle). Suppose $U$ is connected. Suppose $u \in C^2(\bar{U})$ is harmonic. If there is a $x \in U$ such that

$$u(x) = \max_{\bar{U}} u,$$

then $u$ is constant.
Proof. Suppose
\[ u(x) = M := \max_U u. \]

Then for each \( y \in B_r(x) \subset U \), we must have \( u(y) = M \) because otherwise \( u(y) < M \) and thus \( f_{B_r(x)} u < M \). It follows that \( u \) is locally constant and thus constant, since \( U \) is connected. \( \square \)

**Exercise 15.5.** Suppose \( u, v : U \to \mathbb{R} \) are both harmonic and set
\[ f := u|_{\partial U} \quad \text{and} \quad g := v|_{\partial U}. \]

Prove that
\[ f \geq g, \quad \text{but} \quad f \neq g \implies u > v \quad \text{in} \ U. \]

Also, show that
\[ \|u - v\|_{L^\infty(U)} \leq \|f - g\|_{L^\infty(\partial U)}. \]

### 15.2 \( C^k \)–estimates

**Theorem 15.6.** For each \( k \in \mathbb{N}_0 \), there is a constant \( c_k \) such that for all \( B_r(x) \subset U \) we have
\[ |\nabla^k u(x)| \leq \frac{c_k}{r^{n+k}} \|u\|_{L^1(B_r(x))} \]

**Proof.** The proof is by induction. For \( k = 0 \), the assertion is clear from Theorem 15.1.

To show that the assertion holds for \( k + 1 \). Note that \( \partial_{i_{k+1}} \partial_{i_k} \cdots \partial_{i_1} u \) is harmonic. Thus by Theorem 15.1, using Theorem A.4, and the induction hypothesis:

\[
|\partial_{i_{k+1}} \partial_{i_k} \cdots \partial_{i_1} u(x)| = \left| \int_{B_{r/2}(x)} \partial_{i_{k+1}} \partial_{i_k} \cdots \partial_{i_1} u \right|
\leq \frac{2n}{r} \left| \int_{\partial B_{r/2}(x)} (\partial_{i_k} \cdots \partial_{i_1} u) \nu_{i_{k+1}} \right|
\leq \frac{2n}{r} \|\partial_{i_k} \cdots \partial_{i_1} u\|_{L^\infty(B_{r/2}(x))}
\leq \frac{2n+k+1}{r^{n+k+1}} \|u\|_{L^1(B_r(x))}
\leq \frac{c_{k+1}}{r^{n+k+1}} \|u\|_{L^1(B_r(x))}.
\]

Here we used that \( \text{vol}(\partial B_r) = n \text{vol}(B_r) \). \( \square \)
15.3 Liouville’s Theorem

**Theorem 15.7** (Liouville). Let \( u : \mathbb{R}^n \to \mathbb{R} \) be a harmonic function. If there exists a constant \( M > 0 \) such that \( |u| \leq M \) on all of \( \mathbb{R}^n \), then \( u \) is constant.

**Proof.** By Theorem 15.6

\[
|\nabla u|(x) \leq \frac{c}{r} \int_{B_r(x)} |u| \leq \frac{cM}{r}.
\]

Since \( r > 0 \) is arbitrary, it follows that \( \nabla u = 0 \).

**Exercise 15.8.** Let \( u : \mathbb{R}^n \to \mathbb{R} \) be a harmonic function. Suppose there are constants \( A, B > 0 \) and \( \alpha \in (0, 1) \) such that

\[
|u|(x) \leq A + B|x|^\alpha.
\]

Prove that \( u \) is constant.

**Remark 15.9.** Note that there are plenty of harmonic functions with that satisfy the above hypothesis with \( \alpha = 1 \)!

15.4 Harnack’s Inequality

**Theorem 15.10** (Harnack’s inequality). Suppose \( V \subset \mathbb{R}^n \) is an open subset such that \( \overline{V} \subset U \). Then there exists a constant \( c > 0 \) (depending on \( V \)) such that all whenever \( u : U \to \mathbb{R} \) is harmonic and \( u \geq 0 \) then

\[
\sup_V u \leq c \inf_V u.
\]

**Proof.** Set \( r := d(V, \partial U) \). If \( y \in V \cap B_{r/2}(x) \), then \( B_{r/2}(x) \subset B_r(y) \) and thus

\[
u(x) = \int_{B_{r/2}(x)} u \leq 2^n \int_{B_r(y)} u = 2^n u(y).
\]

(Note that the inequality uses that \( u \geq 0 \).

\( V \) can be covered by a finite number of balls of radius \( r \) with center in \( V \). If there \( N \) balls, then the above argument gives

\[
u(x) \leq 2^{nN} u(y)
\]

for every pair \( x, y \in V \).
Lecture 16. Weyl’s Lemma

Definition 16.1. A function \( u : U \to \mathbb{R} \) is called weakly harmonic if \( u \in L^1_{\text{loc}} \), i.e., it is locally integrable\(^6\) and for each \( \phi \in C_0^\infty(U) \) we have
\[
\int_U u \Delta \phi = 0.
\]

Exercise 16.2. Prove that if \( u \) is slowly growing (see Section 11.2) and weakly harmonic on \( \mathbb{R}^n \), and if
\[
K_t(x) := \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t}
\]
denotes the heat kernel, then
\[
\int_{\mathbb{R}^n} (\Delta_x K_t(x - y)) u(y) \, dy = 0.
\]
(\( \Delta_x \) means that we take the Laplacian of \( K_t(x - y) \) as a function of \( x \).)

Theorem 16.3 (Weyl’s Lemma). Every weakly harmonic function is smooth and harmonic.

This is a truly remarkable statement. We start with a function \( u \) that could a priori be extremely rough, not even smooth enough to write down \( \Delta u = 0 \); and then the fact that it is weakly harmonic is enough to deduce its smoothness. Although the proof we will give relies very heavily on properties of the Laplace’s equation, this type of statement hold for the vast class of so-called elliptic equations. This is often referred to as elliptic regularity.

Remark 16.4. Theorem 16.3 extends to the space of Schwartz distributions \( \mathcal{D}(U) \), the dual space of \( C_0^\infty(U) \).

Remark 16.5. Note that \( \text{Re}(\frac{1}{z}) : \mathbb{C}^* \to \mathbb{R} \) is not in \( L^1_{\text{loc}}(\mathbb{R}^2) \).

16.1 Proof using Heat Kernel

Proof of Theorem 16.3 using the heat kernel. We begin with the case \( U = \mathbb{R}^n \) and \( u \) slowly growing; not because the general case reduces to it, but because it makes the proof idea crystal clear.

Let \( K_t(x) := \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t} \) denote the heat kernel on \( \mathbb{R}^n \). Define
\[
u_t(x) := (K_t \ast u)(x) = \int_{\mathbb{R}^n} K_t(x - y) u(y) \, dy.
\]

\(^6\)In case you don’t know what this means assume \( u \) to be continuous.
From Section 7 we know that for each \( t > 0 \), \( u_t \) is smooth, and for each \( x \in \mathbb{R}^n \),

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

Using the fact that \( K_t \) solves the heat equation, i.e.,

\[
\partial_t K_t = -\Delta K_t
\]

and Exercise 16.2, we compute that for each \( t > 0 \)

\[
\partial_t u_t(x) = \int_{\mathbb{R}^n} \partial_t K_t(x-y)u(y) \, dy
\]

\[
= -\int_{\mathbb{R}^n} (\Delta K_t(x-y))u(y) \, dy
\]

\[
= 0.
\]

But this implies that \( u_t(x) \) does not depend on \( t \) at all! Hence, \( u = u_t \) and therefore smooth.

**Remark 16.6.** Before pressing on, let me briefly point out what the idea is here: Suppose we are given a weakly harmonic function \( u \), then we consider the heat flow \( u_t \) starting at \( u \). The function \( (t,x) \mapsto u(x) \) already does solve the heat equation in the weak sense; hence, by uniqueness (in a stronger form than we proved in this class) it agrees with the smooth solution that is known to exist.

Let us now prove the general case. We will show that \( u \) is smooth in a neighbourhood of each point \( x_0 \in U \). The problem with carrying over the above proof is that we cannot smooth \( u \) by convolution with the heat kernel, since \( u \) is not defined on all of \( \mathbb{R}^n \). This, however, is easily remedied. Fix \( r > 0 \) such that \( B_{4r}(x_0) \subset U \) and fix a smooth function \( \chi : \mathbb{R}^n \to [0,1] \) supported in \( B_{4r}(x_0) \) and equal to one on \( B_{2r}(x_0) \). Define \( v \in L^1_{\text{loc}}(\mathbb{R}^n) \) by

\[
v := \chi \cdot u.
\]

Set

\[
v_t(x) := (K_t \ast v)(x).
\]

These functions are still smooth for each \( t > 0 \) and also for each \( x \in B_1(x_0) \)

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

The functions \( v_t \) now really do depend on \( t \); however, in a controlled way:
Proposition 16.7. For each $k \geq 0$, there is a constant $c_k > 0$ (independent of $t$) such that

$$\|\partial_t v_t\|_{C^k(B_1(x_0))} \leq c_k.$$ 

As a consequence of this, the sequence $(v_{1/i})$ is Cauchy in each $C^k(B_1(x_0))$; hence, they converge to a smooth limit function which must agree with $u|_{B_1(x_0)}$ because this is the point-wise limit.

The proof of Proposition 16.7 is very simple in principle, but writing out all the details will be a bit hairy.

Let's compute

$$\partial_t v_t(x) = \int_{\mathbb{R}^n} \partial_t K_t(x - y) \chi(y) u(y) \, dy$$

$$= -\int_{\mathbb{R}^n} (\Delta_y K_t(x - y)) \chi(y) u(y) \, dy$$

$$= -\int_{\mathbb{R}^n} (\Delta_y K_t(x - y)) \chi(y) u(y) \, dy$$

$$= -\int_{\mathbb{R}^n} [\Delta_y (K_t(x - y) \chi(y))] u(y) \, dy$$

$$- \int_{\mathbb{R}^n} L_t(x, y) u(y) \, dy$$

$$= -\int_{\mathbb{R}^n} L_t(x, y) u(y) \, dy.$$

with

$$L_t(x, y) := -2 \langle \nabla_y K_t(x - y), \nabla \chi(y) \rangle + K_t(x - t) \Delta \chi(y).$$

The last step uses that $u$ is weakly harmonic. Now the $\nabla \chi$ and $\Delta \chi$ are supported in the annulus $A := B_{4r}(x_0) \setminus B_{2r}(x_0)$ and we only care about $x \in \overline{B}_1(x_0)$. Since the function $K_t(x - \cdot)$ is smooth on $A$ and has uniformly bounded $C^k$-norms, the same holds for $L_t$ and this proves the proposition and thus the theorem.

Theorem 16.8 (Elliptic regularity for $\Delta$). Suppose $f \in C^\infty(U)$ and $u \in L^1_{\text{loc}}(U)$ is such that for each $\phi \in C^0_0(U)$ we have

$$\int_U u \Delta \phi = \int_U f \cdot \phi,$$

then $u$ is smooth and $\Delta u = f$.

Exercise 16.9. Prove Theorem 16.8!

Remark 16.10. The above proof can be generalised in various ways, see Daniel Stroock’s “Weyl’s Lemma, one of many”.

84
16.2 Proof via Mean-value property

Most PDE books give the following proof using the mean-value property of harmonic functions. (You may like usual proof better because it avoids the apparently technical Proposition 16.7, but this sort of argument really should not scare you.)

Alternative proof of Theorem 16.3. We present the argument in three steps. The first two steps together proof that that weakly harmonic functions satisfy the mean-value property; the last step proves that functions satisfying the mean-value property are smooth and thus harmonic.

Given $B_{4r}(x_0) \subseteq U$, we need to show that $u$ is smooth on $B_r(x_0)$. We introduce a family of mollifiers $\eta_\varepsilon$: fix $\eta \in C_0^\infty([0, 1); [0, \infty))$ with $\int_{[0,1]} s^{n-1} \eta(s) \, ds = 1$ and set

$$\eta_\varepsilon(x) := \frac{1}{\varepsilon^n \omega_{n-1}} \eta \left( \frac{|x|}{\varepsilon} \right)$$

**Step 1.** For each $0 < \varepsilon \leq r$, the function $u_\varepsilon : B_{2\varepsilon}(x_0) \rightarrow \mathbb{R}$ defined by

$$u_\varepsilon(x) = \int_{B_\varepsilon(x)} \eta_\varepsilon(x - y) u(y) \, dy$$

is smooth. For each $x \in B_{2\varepsilon}(x_0)$

$$\lim_{\varepsilon \downarrow 0} u_\varepsilon(x) = u(x).$$

Moreover,

$$\lim_{\varepsilon \downarrow 0} \int_{B_{2\varepsilon}(x_0)} |u_\varepsilon - u| = 0.$$

**Exercise 16.11.** Prove this!

**Step 2.** For each $x \in B_{\varepsilon}(x_0)$ and $0 < s < r$,

$$u(x) = \int_{\partial B_s(x)} u(y) \, dy.$$

(A simplified version of) the argument from Exercise 16.2 shows that

$$\Delta u_\varepsilon(x) = 0.$$

Hence, $u\varepsilon$ satisfies

$$u_\varepsilon(x) = \int_{B_s(x)} u_\varepsilon(y) \, dy$$

85
for each $x \in B_r(x_0)$ and $0 < s < r$. The previous step implies that the same holds for $u$, i.e.,

$$u(x) = \int_{B_r(x)} u(y) \, dy.$$  

This is not quite what we are looking for; however, we can differentiate this expression by $s$ to get

$$\int_{\partial B_s(x)} u(y) \, dy = \int_{B_s(x)} u(y) \, dy.$$  

This yields the desired assertion.

**Step 3.** Finally, we come full circle and prove that

$$u = u_\varepsilon$$

in $B_r(x_0)$; in particular, $u$ is smooth in $B_r(x_0)$.

We compute

$$u_\varepsilon(x) = \int_{B_r(x)} \eta_\varepsilon(x - y) u(y) \, dy$$

$$= \int_0^\varepsilon \frac{s^{n-1}}{s^n - \eta(s/\varepsilon)} \int_{\partial B_s(x)} u(y) \, dy \, ds$$

$$= u(x) \cdot \int_0^\varepsilon (s/\varepsilon)^{n-1} \eta(s/\varepsilon) \, d(s/\varepsilon)$$

$$= u(x).$$  

\[\square\]

**Theorem 16.12.** If $u$ satisfies the mean-value property, then it is smooth and harmonic.

**Theorem 16.13.** If $u: U \to \mathbb{R}$ is harmonic, then it is analytic.

**Proof.** See [Eva10, Section 2.2 Theorem 10]. The idea is to show that the Taylor series converges by estimating the remained using Theorem 15.6. (The derivation I presented does not yield estimates on $c_k$ which are good enough, but they can be improved by adjusting the radius $r/2$ of the smaller ball to a more clever choice depending on $k$.)  

\[\square\]

**Corollary 16.14** (Unique continuation). Suppose $u: U \to \mathbb{R}$ is a harmonic function. If $u$ vanishes on a non-empty open subset of $U$, then $u$ vanishes identically.
Exercise 16.15. Consider the annulus
\[ A := \{(x, y) \in \mathbb{R}^2 : \sqrt{|x|^2 + |y|^2} \in [1, 2]\} = \bar{B}_2(0) \setminus B_1(0). \]
The function \( f: A \to \mathbb{R} \) defined by
\[ f(x, y) = \frac{x}{x^2 + y^2} = \text{Re}(1/z). \]
is harmonic.
Is there a harmonic function \( \bar{f}: \bar{B}_2(0) \to \mathbb{R} \) such that
\[ \bar{f}|_A = f? \]

16.3 Unique continuation and the frequency function

There is another proof of unique continuation using the frequency function, which is also an important concept.

Definition 16.16. Let \( x \in U \) and \( R := d(x, \partial U) \). The frequency function \( N_x: I \to [0, \infty) \) is defined as
\[ N_x(r) := \frac{r \int_{B_r(x)} |\nabla u|^2}{\int_{\partial B_r(x)} |u|^2} = \frac{rH_x(r)}{h_x(r)} \]
with
\[ H_x(r) := \int_{B_r(x)} |\nabla u|^2 \quad \text{and} \quad h_x(r) := \int_{\partial B_r(x)} |u|^2. \]
Here \( I := \{r \in (0, R) : h(r) \neq 0\} \).

We now drop \( x \) from the notation.

Proposition 16.17. If \( u \) is harmonic, then \( N' \geq 0 \) on \( I \).

Proposition 16.18. If \( u \) is harmonic, then we have
\[ H(r) = \int_{\partial B_r(x)} (\partial_r u)u. \]

Proof. This follows from Theorem A.4. \qed

Proposition 16.19. If \( u \) is harmonic, then
\[ H'(r) = \frac{n-2}{r} H(r) + 2 \int_{\partial B_r(x)} |\partial_r u|^2 \quad \text{and} \]
\[ h'(r) = \frac{n-1}{r} h(r) + 2 \int_{\partial B_r(x)} (\partial_r u)u. \]
Proof. We compute

\[ h'(r) = \frac{n-1}{r} \int_{\partial B_r(x)} |u|^2 + \int_{\partial B_r(x)} \partial_r |u|^2 \]

\[ = \frac{n-1}{r} h(r) + 2 \int_{\partial B_r(x)} (\partial_r u) u. \]

The computation for \( H' \) is a bit more involved. First note that

\[ H'(r) = \int_{\partial B_r(x)} |\nabla u|^2 = \sum_{i=1}^n \int_{\partial B_r(x)} \langle x_i |\nabla u|^2, x_i/r \rangle. \]

Then compute, using Theorem A.3,

\[
\int_{\partial B_r(x)} \langle x_i |\nabla u|^2, x_i/r \rangle \\
= \int_{B_r(x)} \partial_i (x_i |\nabla u|^2) \\
= \int_{B_r(x)} |\nabla u|^2 + 2 \sum_{j=1}^n x_i (\partial_j u) (\partial_i u) \\
= \int_{B_r(x)} |\nabla u|^2 - 2 \sum_{j=1}^n \partial_j (x_i \partial_j u) (\partial_i u) + 2 \int_{\partial B_r(x)} \sum_{j=1}^n x_i (\partial_j u) (\partial_i u) \nu_j \\
= \int_{B_r(x)} |\nabla u|^2 - 2 |\partial_i u|^2 + 2 r \int_{\partial B_r(x)} \left( \frac{x_i}{r} \partial_i u \right) (\partial_r u). \\
\]

Summing this up yields

\[ H(r) = (n-2) \int_{B_r(x)} |\nabla u|^2 + 2 r \int_{\partial B_r(x)} |\partial_r u|^2. \]

\[ \square \]
Proof of Proposition 16.17. Using Proposition 16.19, we compute

\[ N'(r) = \frac{H(r)}{h(r)} + \frac{r H'(r)}{h(r)} - \frac{r H(r) h'(r)}{h(r)^2} \]

\[ = \frac{H(r)}{h(r)} + \frac{(n - 2)H(r)}{h(r)} + \frac{2r \int_{\partial B_r(x)} |\partial_r u|^2}{h(r)} \]

\[ - \frac{r H(r)}{h(r)^2} \left( \frac{n - 1}{r} h(r) + 2 \int_{\partial B_r(x)} (\partial_{\nu} u) \right) \]

\[ = \frac{2r}{h(r)^2} \left( \int_{\partial B_r(x)} |u|^2 \cdot \int_{\partial B_r(x)} |\partial_r u|^2 - \left| \int_{\partial B_r(x)} (\partial_{\nu} u) \right|^2 \right) \]

\[ \geq 0 \]

by Cauchy–Schwarz. \( \square \)

**Proposition 16.20.** If \( u \) is harmonic and non-vanishing on \( B_R(x) \), then \( I = (0, R) \) and for \( 0 < s < r < R \)

\[ h(r) = \left( \frac{r}{s} \right)^{n-1} \exp \left( 2 \int_{s}^{r} \frac{N(t)}{t} \, dt \right) h(s); \]

in particular,

\[ \left( \frac{r}{s} \right)^{n-1+2N(s)} h(s) \leq h(r) \leq \left( \frac{r}{s} \right)^{n-1+2N(r)} h(s). \]

**Proof.** On \( I \), the formula for \( h' \) from Proposition 16.19 can be written in the form.

\[ h'(r) = (n - 1 + 2N(r)) h(r)/r; \]

hence,

\[ (\log h(r))' = (n - 1 + 2N(r))/r. \]

This integrates to the stated identity if \([s, r] \in I\). The second assertion follows from Proposition 16.17.

To see that \( I = (0, R) \), first note that since \( u \) does not vanish on \( B_R(x) \) we have \( I \neq \emptyset \). We also know trivially that \( I \) is open. Suppose \((r_1, r_2) \subset I\) is a maximal interval contained in \( I \). If \( s \in (r_1, r_2) \), it follows that

\[ h(r_1) \geq (r_1/s)^{n-1+2N(s)} h(s) \quad \text{and} \quad h(r_2) \geq (r_2/s)^{n-1+2N(s)} h(s). \]

Thus \( r_1 = 0 \), because if \( r_1 > 1 \) it follows that \( h(r_1) > 0 \) contradicting the maximality; similarly \( r_2 = R \). \( \square \)
Alternative proof of Corollary 16.14. If $u$ does not vanish identically on $U$, then there exists a $B_r(x) \subset U$ such that $u$ vanishes on $B_{r/2}(x)$ but $h_{x}(r) > 0$. This contradicts the previous proposition. \qed
LECTURE 17.  Green’s functions

Combining the Dirichlet principle and Weyl’s Lemma in the form of Theorem 16.8 allows us to find a smooth solution $u \in C^\infty(U)$ of the Poisson equation

$$\Delta u = g$$

for $g \in C^\infty(U) \cap C^0(\bar{U})$ satisfying the homogeneous Dirichlet boundary conditions in the sense that $u \in W^{1,2}_0(U)$. The construction of this solution however is somewhat abstract. This is “good enough” for most applications, but it is nice to have a more concrete ways to represent/find solutions. This is where Green’s functions, the fundamental solutions of the Laplace equation, come into play.

Let us begin with the Green’s function $G$ on $\mathbb{R}^n$. We want $G$ to be such that for suitable $g: \mathbb{R}^n \to \mathbb{R}$ the function defined by

$$u = G * g,$$

i.e.,

$$u(x) = \int_{\mathbb{R}^n} G(x - y)g(y) \, dy$$

satisfies

$$\Delta u = g.$$

In the language of distributions this means that

$$\Delta G = \delta.$$

**Definition 17.1.** The Green’s function on $\mathbb{R}^n$ is the function $G: \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ defined by

$$G(x) := \begin{cases} -\frac{1}{2\pi} \log |x| & n = 2, \\ \frac{1}{n(n-2)\text{vol}(B_1)} |x|^{2-n} & n > 2. \end{cases}$$

**17.1  Derivations of the Green’s function**

Let me briefly present two ways to arrive at the above formula (for $n \geq 3$).

**17.1.1  Derivation from the heat kernel**

We know that the heat kernel $K_t$ satisfies

$$(\partial_t + \Delta)K_t = 0$$
for $t > 0$, $\lim_{t \to \infty} K_t = 0$ and $\lim_{t \to 0} K_t = \delta$ (as a distribution). Thus (somewhat formally)
\[ \Delta \int_0^\infty K_t = -\int_0^\infty \partial_t K_t = \delta. \]

For $n > 2$ and $x \neq 0$ we have
\[ \int_0^\infty \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \, dt = G(x). \]

To see this, change variables $s := \frac{4t}{|x|^2}$ to get
\[ \int_0^\infty \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \, dt = \frac{1}{4\pi^{n/2}} \int_0^\infty \frac{e^{-\frac{1}{4s}}}{s^{n/2}} \, ds \cdot |x|^{-n} = \frac{\Gamma \left( \frac{n}{2} - 1 \right)}{4\pi^{n/2}} \cdot |x|^{-n}. \]

But
\[ \text{vol}(B^n_1) = \frac{\pi^{n/2}}{\Gamma \left( \frac{n}{2} + 1 \right)} = \frac{4\pi^{n/2}}{n(n-2)\Gamma \left( \frac{n}{2} - 1 \right)}. \]

Arguing with just a bit more care (that is writing everything integrated against test functions $\phi \in C_0^\infty(\mathbb{R}^n)$), the above can be made into a rigorous proof of the following theorem.

**Theorem 17.2.** If $g \in C_0^\infty(\mathbb{R}^n)$, then $u := G * g \in C^\infty(\mathbb{R}^n)$ and
\[ \Delta u = g. \]

**17.1.2 Derivation using the Fourier transform**

Using Proposition 10.5, taking the Fourier transform of $\Delta G = \delta$, gives
\[ 4\pi^2 |y|^2 \hat{G}(y) = 1. \]

Thus
\[ G = \mathcal{F}^{-1} \left( \frac{1}{4\pi^2 |\cdot|^2} \right) = \mathcal{F} \left( \frac{1}{4\pi^2 |\cdot|^2} \right) \]

Actually computing the right-hand side is not that trivial, because $|\cdot|^{-2} \notin L^1(\mathbb{R}^n)$ and we need to view it as a tempered distribution: For $n \geq 3$, the integrability fails at infinity, i.e., $|\cdot|^{-2} \notin L^1(\mathbb{R}^n \setminus B_1)$; but we still have $|\cdot|^{-2} \in S^*(\mathbb{R}^n)$. For $n = 2$, however, the problem is more severe since $|\cdot|^{-2} \notin L^1(B_1^2)$ and thus it does not even define a tempered distribution.
We can learn a lot about $G$ without actually doing any computation at all. First, if $A \in \text{SO}(n)$, then
$$ \mathcal{F}(f \circ A) = \mathcal{F}(f) \circ A. $$
Therefore, because $| \cdot |^{-2}$ is $\text{SO}(n)$–invariant, so is $G$; i.e., $G(x) = g(|x|)$ for some $g$. On the other hand if $f$ is homogeneous of degree $k$, then $\hat{f}$ is homogeneous of degree $(-k - n)$. It follows that
$$ G(x) = c|x|^{2-n} $$
for some constant $c$.

### 17.2 Solving the Poisson equation with $C^1$ inhomogeneity

**Theorem 17.3.** If $g \in C^1_0(\mathbb{R}^n)$, then $u := G \ast g \in C^2(\mathbb{R}^n)$ and
$$ \Delta u = g. $$

**Remark 17.4.** Note that this is actually a bit stronger than saying $\Delta G = \delta$ (because $g$ need not be smooth). Also note that our proof does not work with $g \in C^0(\mathbb{R}^n)$ only; however, one can get away with $g \in C^{0,\alpha}_0(\mathbb{R}^n)$ for some $\alpha > 0$, that is, $g \in C^0(\mathbb{R}^n)$ and
$$ \sup_{x \neq y} \frac{|g(x) - g(y)|}{|x - y|^\alpha} < \infty. $$
This observation lead to an important subject called “Schauder theory”.

**Proof of Theorem 17.3.**

**Exercise 17.5.** Prove that $\Delta G = 0$ on $\mathbb{R}^n \setminus \{0\}$. (Hint: It helps to know the formula for $\Delta$ in polar coordinates.)

Note that $\partial_i \partial_j G \sim |x|^{-n}$ near zero, thus it is not integrable and we cannot use just use the exercise and Proposition D.2. However, we can write
$$ \Delta u(x) = -\sum_{i=1}^n (\partial_i G \ast \partial_i g)(x) = -\sum_{i=1}^n \int_{\mathbb{R}^n} \partial_i G(x - y) \partial_i g(y) \, dy. $$
To see that this vanishes, we proceed as follows. For any $\varepsilon > 0$, note that
$$ \left| \sum_{i=1}^n \int_{B_\varepsilon(x)} \partial_i G(x - y) \partial_i g(y) \, dy \right| \leq \|\nabla g\|_{L^\infty} \int_{B_\varepsilon(0)} |\nabla G| $$
$$ = \varepsilon \|\nabla g\|_{L^\infty} \int_{B_\varepsilon(0)} |\nabla G|, $$
by the homogeneity of degree \((1 - n)\) of \(|\nabla G|\). This goes to zero as \(\varepsilon \downarrow 0\). Moreover, by Theorem A.4

\[
\int_{\mathbb{R}^n \setminus B_\varepsilon(x)} \langle \nabla_y G(x-y), \nabla_y g(y) \rangle \, dy = \int_{\mathbb{R}^n \setminus B_\varepsilon(x)} (\Delta_y G(x-y))g(y) \, dy \\
+ \int_{B_\varepsilon(x)} \partial_\nu G(x-y)g(y) \, dy.
\]

Here the first term on the right-hand side vanishes and \(\nu = -\frac{x-y}{|x-y|}\). We compute further

\[
\int_{\partial B_\varepsilon(x)} \partial_\nu G(x-y)g(y) \, dy = \int_{\partial B_\varepsilon(0)} -\partial_\nu G(y)g(x-y) \, dy \\
= \int_{\partial B_\varepsilon(0)} \frac{1}{n\text{vol}(B^n_1)} |y|^{1-n} g(x-y) \, dy \\
= \int_{\partial B_\varepsilon(0)} g(x-y) \, dy \to g(x)
\]
as \(\varepsilon \to 0\). Here we used that \(n\text{vol}(B^n_1) = \text{vol}(\partial B^n)\). This completes the proof. \(\square\)

### 17.3 Representation formula for solutions of the Dirichlet problem

Suppose now that \(U \subset \mathbb{R}^n\) is a bounded open subset with smooth boundary. Suppose for each \(x \in U\), \(\phi_x\) solves the Dirichlet problem

\[
\Delta \phi_x = 0 \quad \text{in } U \quad \text{and} \\
\phi_x(y) = G(x-y) \quad \text{for all } y \in \partial U.
\]

Finding such correction term \(\phi_x\) is difficult in general; however, for simple regions \(U\) formulas for \(\phi_x\) are known. Set

\[G_U(x,y) := G(x-y) - \phi(x,y).\]

**Proposition 17.6** (Representation formula). Given any \(u \in C^2(\bar{U})\), for any \(x \in U\) we have

\[
u(x) = -\int_{\partial U} u(y)\partial_\nu G_U(x,y) \, dy + \int_U \Delta u(y)G_U(y,x).
\]

**Proof.** Apply Theorem A.4 to \(U_\varepsilon := U \setminus B_\varepsilon(x)\) to derive

\[
\int_{\partial U_\varepsilon} u(y)\partial_\nu G_U(x,y) - G_U(x,y)\partial_\nu u(y) = \int_{U_\varepsilon} \Delta u(y)G_U(x,y).
\]
Now $G_U = 0$ on $\partial U$ (this is why the correction term $\phi_x$ is needed!), and
\[
\int_{\partial B_\varepsilon(x)} G_U(x, y) \partial_\nu u(y) \, dy \to 0
\]
as $\varepsilon \to 0$. The argument from Theorem 17.3 shows that
\[
u(x) = \lim_{\varepsilon \to 0} \int_{\partial B_\varepsilon} \partial_\nu G_U(x, y) u(y).
\]
Here $\nu$ is the unit normal pointing out of $U_\varepsilon$.

Remark 17.7. Since $\phi_x$ is often hard to find, it is useful to remark that we also have
\[
u(x) = \int_{\partial U} G(x - y) \partial_\nu u(y) - u(y) \partial_\nu G(x - y) \, dy + \int_U \Delta u(y) g(y, x).
\]

There is a subtlety here: If we know that there is a solution $u \in C^2(\bar{U})$ of a certain Dirichlet problem on $U$, then Proposition 17.6 allows us to recover $u$ from $g := \Delta u$ and $f := u|_{\partial U}$. However, we do not make the assertion that Proposition 17.6 for prescribed values of $g$ and $f$ does the define a solution.

17.4 The Dirichlet problem on a ball

Exercise 17.8. The formula
\[
\phi_x(y) := G(|x|y - x/|x|)
\]
defines a correction term for $B := B_1(0) \subset \mathbb{R}^n$.

Set
\[
G_B(x, y) := G(x - y) - \phi_x(y).
\]

Remark 17.9. Since
\[
||x|y - x/|x|| = |x|^2|y|^2 - 2\langle x, y \rangle + 1,
\]
$G_B(x, y) = G_B(y, x)$. (This is true for all $G_U$, in fact.)

Definition 17.10. We define the Poisson kernel (for $r > 0$) to be
\[
P_r(x, y) := \frac{r^2 - |x|^2}{r \text{vol}(\partial B_1)||x - y||^n}.
\]

Exercise 17.11. For $x \in B_1(x)$ and $y \in \partial B_1(x)$, we have
\[
P_1(x, y) = -\partial_\nu G_B(x, y).
\]
Exercise 17.12. Prove that there is a constant $c > 0$ such that for all $u \in C^\infty(B_1)$ we have
\[
\|u\|_{L^\infty(B_{1/2})} \leq c \left( \|\Delta u\|_{L^\infty(B_1)} + \|u\|_{L^1(\partial B_1)} \right)
\]
(Hint: Use Proposition 17.6.)

Theorem 17.13. Given $f \in C^0(\partial B_r)$, set
\[
u(x) := \begin{cases} \int_{\partial B_r} g(y) P_1(x, y) \, dy & x \in B \\ g(x) & x \in \partial B. \end{cases}
\]
Then $u \in C^\infty(B_r) \cap C^0(\overline{B_r})$ and $\Delta u = 0$ in $B$. That is, $u$ solves the Dirichlet problem for the Laplace equation on $B_r$ with boundary value $g$.

Proof. It suffices to prove this for $r = 1$ (because everything can be pulled back to $B_1$).

Note that $\Delta_x G_B(x, y) = 0$ for all fixed $y \in \partial B$, and similarly
\[
\Delta_x \partial_y G_B(x, y) = \Delta_x \partial_y G_B(x, y) = \partial_y \Delta_x G_B(x, y) = 0.
\]
Thus $\Delta_x P_1(x, y) = 0$. From this it follows that $\Delta u = 0$ using Proposition D.2.

Thus we have to show that
\[
g(x) = \lim_{z \to x} \int_{\partial B} g(y) P_1(z, y) \, dy.
\]
It is a fact that for all $z \in B_1$
\[
\int_{\partial B} P_1(z, y) \, dy = 1.
\]
Thus we have to show that
\[
\lim_{z \to x} \int_{\partial B} (g(y) - g(x)) P_1(z, y) \, dy = 0.
\]
This can be done in the “usual way”: First, we have
\[
\left| \int_{\partial B \setminus B_\varepsilon(x)} (g(y) - g(x)) P_1(z, y) \, dy \right| \leq \sup_{B_\varepsilon(x)} |g(y) - g(x)|,
\]
which goes to zero as $\varepsilon \to 0$ by continuity of $g$. Second, for fixed $\varepsilon > 0$
\[
\left| \int_{\partial B \setminus B_\varepsilon(x)} (g(y) - g(x)) P_1(z, y) \, dy \right| \leq 2\|g\|_{L^\infty} \int_{\partial B \setminus B_\varepsilon(x)} \frac{1 - |z|^2}{\text{vol}(\partial B_1)||z - y||^n} \, dy
\]
goes to zero as $z \to x \in \partial B_1$. 

96
LECTURE 18. Perron’s Method

Perron’s method is a further technique to solve the Dirichlet problem for the Laplace equation:

\[ \Delta u = 0 \quad \text{in } U \]
\[ u|_{\partial U} = f. \]

It is based on the observation that be the maximum principle if \( \Delta v \leq 0 \) and \( v|_{\partial U} = f \), then \( v \leq u \); hence:

- \( u(x) \) is the maximum value of a subharmonic function with this boundary value; and
- we can make any subharmonic function \( v \) larger by replacing \( v|_{B_r(x)} \) with a harmonic function with the same boundary values.

One difficulty is that harmonic replacement does not necessarily produce \( C^2 \) functions, and thus the result might not satisfy \( \Delta v \leq 0 \) in the strong sense. To deal with this we make the following definition.

**Definition 18.1.** A upper semi-continuous function \( v: U \to \mathbb{R} \) is called subharmonic if one of the following equivalent conditions are satisfied:

- If \( V \subset U \) is open and \( u: V \to \mathbb{R} \) is harmonic and \( v \leq u \) on \( \partial V \), then \( v \leq u \) on \( V \).
- For all \( \bar{B}_r(x) \subset U \), we have
  \[ v(x) \leq \int_{\partial B_x} v. \]
- For all \( \bar{B}_r(x) \subset U \), we have
  \[ v(x) \leq \int_{B_x} v. \]

**Exercise 18.2.** Show that these conditions are equivalent. Also, show that if \( v \in C^2(U) \) and \( \Delta v \leq 0 \), then \( v \) is subharmonic.

**Exercise 18.3.** Show that if \( \{v_1, \ldots, v_k\} \) are subharmonic, then so is \( \max\{v_1, \ldots, v_k\} \).
18.1 Harmonic replacement

Definition 18.4. Suppose \( v \in C^0(\overline{U}) \) is subharmonic and \( \overline{B}_r(x) \subset U \). Then the harmonic replacement \( \tilde{v} : \overline{U} \to \mathbb{R} \) is defined by

\[
\tilde{v}(x) := \begin{cases} 
  v(x) & x \notin B_r(x) \\
  \int_{\partial B_r(x)} P_r(x, y)v(y) \, dy & x \in B_r(x).
\end{cases}
\]

Here \( P_r \) is the Poisson kernel.

Proposition 18.5. If \( v \in C^0(\overline{U}) \) is subharmonic and \( \tilde{v} \) is a harmonic replacement, then \( \tilde{v} \) is subharmonic and \( \tilde{v} \geq v \).

Proof. By definition of subharmonic, \( \tilde{v} \geq v \). To see that \( \tilde{v} \) is also subharmonic, suppose \( u : V \to \mathbb{R} \) is a harmonic function with \( v|_{\partial V} \leq \tilde{v}|_{\partial V} \leq u|_{\partial V} \). Since \( v \) is subharmonic, \( v|_V \leq u \). Thus \( \tilde{v}|_{\partial(V \cap B_r(x))} \leq u|_{\partial(V \cap B_r(x))} \). Therefore by the maximum principle \( \tilde{v}|_{V \cap B_r(x)} \leq u|_{V \cap B_r(x)} \). This completes the proof.

18.2 The maximal subharmonic function

Fix \( f \in C^0(\partial U) \). Define

\[
S_f := \{ v \in C^0(\overline{U}) : v \text{ subharmonic and } v|_{\partial U} \leq f \}.
\]

Proposition 18.6. The function \( u : U \to \mathbb{R} \) defined by

\[
u(x) := \sup_{S_f} v(x)\]

is harmonic.

Remark 18.7. Note that for any \( v \in S_f \), \( \sup_U v \leq \sup_{\partial U} f \). Thus \( u \) is well-defined.

Proof. Consider \( \overline{B}_{2r}(x) \subset U \). We will show that \( u \) is harmonic in \( B_r(x) \). Since \( \overline{B}_r(x) \) is arbitrary this will prove the assertion.

Step 1. Construction of a good approximating sequence \( (\tilde{v}_i) \).

Let \( (v_i) \in S_f^N \) such that \( u(x) = \lim_{i \to \infty} v_i(x) \). By Exercise 18.3 we can assume that \( v_i \) is monotone increasing. Denote by \( \tilde{v}_i \) the harmonic replacement of \( v_i \) with respect to \( B_{2r}(x) \). By the maximum principle, \( \tilde{v}_i \) is still monotone increasing. Thus

\[
\lim_{i \to \infty} \tilde{v}_i(x) = u(x).
\]

Step 2. The limit \( \tilde{v} := \lim_{i \to \infty} \tilde{v}_i \) defines a harmonic function on \( B_r(x) \).
By the Harnack inequality Theorem 15.10, applied to \( \pm (\tilde{v}_i - \tilde{v}_j) \), for all \( B_r(x) \)

\[
|\tilde{v}_i - \tilde{v}_j|_{C^0(B_r(x))} \leq c|\tilde{v}_i - \tilde{v}_i|(x).
\]

Thus \( (\tilde{v}_i) \) is a Cauchy sequence in \( C^0(B_r(x)) \). Hence, \( \tilde{v} \) is continuous. Since the \( \tilde{v}_i \) satisfy the mean-value property, so does \( \tilde{v} \). Therefore \( \tilde{v} \) is harmonic.

**Step 3.** \( u = \tilde{v} \) on \( B_r(x) \)

We certainly have \( u \geq \tilde{v} \). Suppose \( u(y) > \tilde{v}(y) \) for some in \( B_r(y) \). Let \( (w_i) \in S^N_f \) be such that \( u(y) = \lim_{i \to \infty} w_i(y) \). Replacing \( w_i \) with \( \max\{w_i, \tilde{v}\} \) we can assume that \( w_i \geq \tilde{v} \). We can use the argument from the previous step to find a harmonic function \( \tilde{w} \in C^0(B_r(x)) \) with

\[
\tilde{v}(y) < \tilde{w}(y) = u(y) \quad \text{and} \quad \tilde{v} \leq \tilde{w} \leq u.
\]

Since \( \tilde{v}(x) = u(x) \), we have

\[
\tilde{v}(x) = \tilde{w}(x) = u(x).
\]

This however contradicts the mean-value property: \( \tilde{v}(y) < \tilde{w}(y) \) implies that

\[
\tilde{v}(x) = \int_{B_r(x)} \tilde{v} < \int_{B_r(x)} \tilde{w} = \tilde{w}(x).
\]

\( \square \)

### 18.3 The boundary condition

Now one question remains: Does \( u \) satisfy the boundary condition? That is, do we have

\[
\lim_{U \ni z \to x} u(z) = f(x)?
\]

Unfortunately, this is not always true. This is related to the boundary \( \partial U \) potentially being very badly behaved.

**Definition 18.8.** Let \( x \in \partial U \). A function \( \beta \in C^0(\bar{U}) \) is called a barrier at \( x \) if \( -\beta \) is subharmonic and

\[
\beta(x) = 0 \quad \text{and} \quad \beta(y) > 0 \text{ for } y \neq x.
\]

A point \( x \in \partial U \) is called regular if there exists a barrier at \( x \). \( U \) is called regular each point \( x \in \partial U \) is regular.
Remark 18.9. If $\beta$ is a local barrier at $x$ in the sense that it defines a barrier for $U \cap B_{2r}(x)$ for some $r > 0$, then

$$
\tilde{\beta}(x) \begin{cases} 
\beta_0 & x \in U \setminus B_r(x) \\
\min \beta, \beta_0 & x \in U \cap B_r(x)
\end{cases}
$$

Here $\beta_0 := \inf_{U \cap (B_{2r}(x) \setminus B_r(x))} \beta$.

Example 18.10. $B := B_1(0) \subset \mathbb{R}^n$ is regular. To see this we construct a barrier at $x_0 := (1, 0, \ldots, 0)$. Let $G$ denote the Green’s function on $\mathbb{R}^n$ and denote $z_0 := (2, 0, \ldots, 0)$. We will assume that $n \geq 2$, so that the Green’s function is decreasing in the distance from 0. Set

$$
\beta(x) := G(x_0 - z_0) - G(x - z_0).
$$

Then $\Delta \beta = 0$ on $B$ and thus $-\beta$ is subharmonic. Also $\beta(x_0) = 0$. Moreover, since $x_0$ is the point in $B$ closest to $z_0$ (that is: for all $x \in B$ we have $|x - z_0| \geq |x_0 - z_0|$ with equality if and only if $x = x_0$), $\beta(y) > 0$ for all $y \in B \setminus \{0\}$.

In a similar way we can construct a barrier at any other boundary point. Note that the crucial point is that each boundary point $x_0 \in \partial U$ is the unique minimizer of $|\cdot - z_0|$ for some $z_0 \not\in \bar{U}$.

Exercise 18.11. Show that $0 \in \partial \{x \in \mathbb{R}^n : 0 < |x| < 1\}$ ($n \geq 2$) is not regular.

Exercise 18.12. Show that $Q = (0, 1)^n$ is regular.

Proposition 18.13. If $x \in \partial U$ is regular, then

$$
\lim_{U \ni z \to x} u(z) = f(x).
$$

Proof. Denote by $u^* \in C^\infty(U)$ the harmonic function defined by

$$
u^*(x) = \inf v^* \in S^*_f v^*(x).
$$

Here $S^*_f = -S_{-f}$. We have $u \leq u^*$.

Denote by $\beta$ a barrier at $x$. For every $\varepsilon > 0$, and $c > 0$ sufficiently large

$$
u_\varepsilon(y) = f(x) - \varepsilon - c\beta(x)
$$

defines a function $v \in S_f$ and similarly

$$
u^*_\varepsilon(y) = f(x) + \varepsilon + c\beta(x)
$$

defines a function in $S^*_f$.

Thus

$$
u_\varepsilon \leq u \leq u^* \leq \nu^*_\varepsilon.
$$

Since $\lim_{U \ni z \to x} \nu_\varepsilon(z) = f(x) - \varepsilon$ and $\lim_{U \ni z \to x} \nu^*_\varepsilon(z) = f(x) + \varepsilon$ and $\varepsilon$ is arbitrary, the assertion follows. \qed
Theorem 18.14. There following are equivalent:

1. $\partial U$ is regular.

2. For every $f \in C^0(\partial U)$, there exists a $u \in C^\infty(U) \cap C^0(\bar{U})$ such that $\Delta u = 0$ on $U$ and $u|_{\partial U} = f$.

Proof. We just proved (1) implies (2). For the converse: given $x \in \partial U$, take $\beta$ to be the solution of the Dirichlet problem with boundary value $f(y) := -|x - y|$. $\square$
Lecture 19. Minimal hypersurfaces and the Bernstein problem

Let $U$ be open subset of $\mathbb{R}^n$, which we assume to be bounded for now. Fix a smooth function $f : \partial U \to \mathbb{R}$. Set

$$C^\infty_f(\bar{U}) := \{u \in C^\infty(\bar{U}) : u|_{\partial U} = f\}.$$  

The graph of $u$ is the set

$$\Gamma(u) := \{(x, f(x)) : x \in U\} \subset \mathbb{R}^{n+1}.$$  

Define $\mathcal{A} : C^\infty_f(\bar{U}) \to [0, \infty)$ by

$$\mathcal{A}(u) := \text{vol}(\Gamma(u)) = \int_\bar{U} \sqrt{1 + |\nabla u|^2}.$$  

This is often called the area functional (even though this term is really only appropriate if $n = 2$).

The question of whether or not a solution minimiser of the functional $\mathcal{A}$ exists for prescribed $f$ is known as Plateau’s problem. This problem has inspired an enormous amount of work and has been the origin of entire subfields of mathematics, like geometric measure theory.

The first variation of $\mathcal{A}$ is given by

$$\delta \mathcal{A}(u) \phi = \int_\bar{U} \left\langle \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}}, \nabla \phi \right\rangle.$$  

This vanishes for all $\phi \in C^\infty_0(U)$ if and only if

$$(19.1) \quad H := -\text{div} \left( \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = 0.$$  

Definition 19.2. An graphical minimal hypersurface in $\mathbb{R}^{n+1}$ is a graph

$$\Gamma(u) := \{(x, f(x)) : x \in U\}$$  

of a $C^2$ map $u : U \to \mathbb{R}$ defined on a domain $U \subset \mathbb{R}^n$ which satisfies (19.1).

We call $\Gamma(u)$ entire if $U = \mathbb{R}^n$.

Remark 19.3. $H$ is called the mean curvature of $\Gamma(u)$.  

102
Remark 19.4. Note that a minimal hypersurface need not minimise $\mathcal{A}$; in fact, if $U = \mathbb{R}^n$, then $\mathcal{A}$ is usually not defined.

Remark 19.5. More generally, a minimal hypersurface in $\mathbb{R}^{n+1}$ is subset $M$ that locally in suitable coordinates can be written as the graph of a function satisfying (19.1). (This allows, in particular, for the hyperplane over which $M$ is a graph to rotate.)

Example 19.6. If $u : U \to \mathbb{R}$ is affine, i.e., it is of the form
$$u(x) = u_0 + \langle x, v \rangle$$
for some $u_0 \in \mathbb{R}$ and $v \in \mathbb{R}^n$, then $\nabla u = 0$ and (19.1) holds trivially.

We compute
$$- \text{div} \left( \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = - \sum_{i=1}^n \partial_i \left( \frac{\partial_i u}{\sqrt{1 + |\nabla u|^2}} \right)$$
$$= \frac{\Delta u}{\sqrt{1 + |\nabla u|^2}} + \sum_{i,j=1}^n \frac{(\partial_i u)(\partial_j u)\partial_i \partial_j u}{(1 + |\nabla u|^2)^{3/2}}$$
$$= - \sum_{i,j=1}^n \left( \frac{\delta_{ij}}{\sqrt{1 + |\nabla u|^2}} - \frac{(\partial_i u)(\partial_j u)}{(1 + |\nabla u|^2)^{3/2}} \right) \cdot \partial_i \partial_j u.$$

If $n = 1$, then this expression simplifies to
$$- \frac{u''}{\sqrt{1 + (u')^2}} + \frac{u''(u')^2}{(1 + (u')^2)^{3/2}} = - \frac{u''}{(1 + (u')^2)^{3/2}}.$$

Example 19.7. The graph of $u : (-r, r) \to \mathbb{R}$ defined by
$$u(x) = \sqrt{r^2 - x^2}$$
is the upper half circle $\mathbb{R}^2$ of radius $r$.

Let us compute $H$. Since
$$u' = -\frac{x}{\sqrt{r^2 - x^2}},$$
we have
$$(1 + (u')^2)^{3/2} = \left(1 + \frac{x^2}{r^2 - x^2}\right)^{3/2} = \left(\frac{r^2}{r^2 - x^2}\right)^{3/2}$$
and
$$u'' = -\frac{1}{(r^2 - x^2)^{1/2}} - \frac{x^2}{(r^2 - x^2)^{3/2}} = -\frac{r^2}{(r^2 - x^2)^{3/2}}.$$
Thus

\[ H = 1/r. \]

**Example 19.8.** The catenoids are a family of a minimal hypersurfaces in \( \mathbb{R}^3 \), which are not graphical. Topologically each catenoid is a cylinder and it can be parametrised for some \( r > 0 \) by the map \( \mathbb{R} \times S^1 \to \mathbb{R}^3 \)

\[
(z, \alpha) \mapsto \begin{pmatrix}
\cosh(z/r) \cdot r \cos(\alpha) \\
\cosh(z/r) \cdot r \sin(\alpha) \\
z
\end{pmatrix}.
\]

We can write (almost all) of the catenoid as the graphs over \( \mathbb{R}^2 \setminus \bar{B}_r(0) \) of the maps

\[
u_{r, \pm}(x, y) := \pm r \cdot \text{arcosh} \left( \frac{\sqrt{x^2 + y^2}}{r} \right).
\]

**Exercise 19.9.** Prove that \( \Gamma(u_{\pm}) \) are minimal graphs.

**Remark** 19.10. Although, the study of hypersurfaces in \( \mathbb{R}^3 \) is an almost ancient subject it is still very active and new results are still being obtained.

In this lecture I want to talk about the following quite surprising result.

**Theorem 19.11** (Bernstein’s Theorem). A entire graphical minimal hypersurface in \( \mathbb{R}^3 \) must be affine; hence, a hyperplane.

One can think of this as a non-linear Liouville type theorem *without* any growth assumptions at infinity.

**Remark** 19.12. The Bernstein problem is the question whether this is true in dimension \( n > 3 \) as well. Simons [Sim68] proved that this holds for \( n \leq 7 \). The problem was settled by Bombieri–De Giorgi–Gusti [BDGG69] who proved that for \( n \geq 8 \) it does not hold.

**Proposition 19.13** (Baby Bernstein). A entire graphical minimal hypersurface in \( \mathbb{R}^2 \) must be affine; hence, a straight line.

**Proof.** Equation (19.1) for a function \( u: \mathbb{R} \to \mathbb{R} \) can be written as

\[
0 = \frac{u''}{(1 + (u')^2)^{3/2}}.
\]

Thus, \( u'' = 0 \), so \( u(x) = u_0 + cx \).
Proof of Theorem 19.11. What we need to show is that any function \( u: \mathbb{R}^2 \to \mathbb{R} \) satisfying
\[
\partial_x^2 u + \partial_y^2 u - 2(\partial_x u)(\partial_y u)\partial_x \partial_y u = 0.
\]
is affine, i.e., \( \nabla u = 0 \). We transform the problem by introducing
\[
\psi_i := \arctan(\partial_i u).
\]
After a somewhat tedious computation one arrives at
\[
a \partial_x^2 \psi_i + 2b \partial_x \partial_y \psi_i + c \partial_y^2 \psi_i = 0.
\]
where
\[
\begin{pmatrix}
a & b \\
b & c
\end{pmatrix} = \begin{pmatrix}
1 + (\partial_y u)^2 & -(\partial_x u)(\partial_y u) \\
-(\partial_x u)(\partial_y u) & 1 + (\partial_x u)^2
\end{pmatrix} =: A.
\]
Note that \( A \) is positive definite and also that \( \psi_i \) are obviously bounded.

This trick reduces Bernstein’s theorem to proving the following Liouville type theorem.

**Theorem 19.14.** Let \( A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^{2 \times 2} \) be such that \( A(x, y) \) is positive definite for each \((x, y) \in \mathbb{R}^2\). If \( u \in C^2(\mathbb{R}) \) is a solution of
\[
a \partial_x^2 u + 2b \partial_x \partial_y u + c \partial_y^2 u = 0,
\]
and
\[
u(x, y) = O(1)
\]
then \( u \) is constant.

Bernstein’s original proof relies on the following theorem, which he stated but gave an incorrect proof of.

**Theorem 19.15** (E. Hopf [Hop50], Mickle [Mic50]). If
\[
- \det \text{Hess}(u) = (\partial_x \partial_y u)^2 - (\partial_x^2 u)(\partial_y^2 u) \geq 0 \quad \text{but} \quad \det \text{Hess}(u) \neq 0,
\]
then \( u \) cannot be bounded (in fact, it cannot be \( o(r) \)).

The proof of this theorem is too involved to present here, but you should feel encouraged to consult the E. Hopf’s and Mickle’s article which give correct proofs of this result. This theorem can be used to prove Theorem 19.14 as follows.

Under the hypothesis of Theorem 19.14 we will prove that
\[
(\partial_x \partial_y u)^2 - (\partial_x^2 u)(\partial_y^2 u) \geq 0.
\]
and equality holds if and only if
\[ \partial_x^2 u = \partial_y^2 u = \partial_x \partial_y u = 0. \]

Using the equation we can write
\[
a \left( (\partial_x \partial_y u)^2 - (\partial_x^2 u)(\partial_y^2 u) \right) = a(\partial_x \partial_y u)^2 + 2b(\partial_x \partial_y u)(\partial_y^2 u) + c(\partial_y^2 u)^2 \geq 0,
\]
since \( A \) is positive definite. This implies the first assertion, since \( a > 0 \) because \( A \) is positive definite.

If the left-hand side vanishes, then again since \( A \) is positive-definite
\[ \partial_x \partial_y u = \partial_y^2 u = 0. \]

To show that \( \partial_x^2 = 0 \) as well note that we also have
\[
c \left( (\partial_x \partial_y u)^2 - (\partial_x^2 u)(\partial_y^2 u) \right) = a(\partial_x^2 u)^2 + 2b(\partial_x^2 u)(\partial_x \partial_y u) + c(\partial_x \partial_y u)^2 \geq 0.
\]
This completes the proof. \( \square \)
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 an 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

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

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

$$\hat{U} (\partial_i f)g + \hat{U} f (\partial_i g) = \hat{\partial U} f (\partial_i \nu_i)$$  

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 preceding 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

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

and

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

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

107
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

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

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 $f : \bar{U} \to \mathbb{R}^m$. 

109
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

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

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

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

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 $x = 0$. 

112
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
\[(a_n) \in \ell^1_k := \{(b_n) \in \mathbb{R}^N : \sum_{n=1}^{\infty} n^k|b_n| < \infty\},\]
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 $n \not\equiv 2 \mod 4$. 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.
D. Dominated Convergence Theorem

Let $U$ be an open subset of $\mathbb{R}^n$.

**Theorem D.1** (Dominated Convergence Theorem). Let $f_n: U \to \mathbb{R}$ be a sequence of integrable functions and $f: U \to \mathbb{R}$ such that for all $x \in U$ we have

$$f(x) = \lim_{n \to \infty} f_n(x)$$

If there exists an integrable function $g: U \to \mathbb{R}$ such that for all $x \in U$

$$|f_n|(x) \leq g(x),$$

then

$$\lim_{n \to \infty} \int_U f_n(x) \, dx = \int_U f(x) \, dx.$$  

**Proposition D.2.** Let $I$ be an interval in $\mathbb{R}$. Let $f: I \times U \to \mathbb{R}$ be such that $f(t, \cdot)$ is integrable for each $t \in \mathbb{R}$ and differentiable in the direction of $t$. If there exists an integrable function $g: U \to \mathbb{R}$ such that for all $t \in I$ and $x \in U$

$$|\partial_t f|(t, x) \leq g(x),$$

then the function $F(t): I \to \mathbb{R}$ defined by

$$F(t) := \int_U f(t, x) \, dx$$

is differentiable and

$$\partial_t F(t) = \int_U \partial_t f(t, x) \, dx.$$  

**Exercise D.3.** Prove Proposition D.2 using Theorem D.1. (Hint: Use the Intermediate Value Theorem: For each $x \in U$, and $t_1, t_2 \in I$ with $t_1 \leq t_2$, there exist a $t \in [t_1, t_2]$ such that $f(t_1, x) - f(t_2, x) = \partial_t f(t, x)$.)
References


116