18.152: Introduction to Partial Differential Equations

Thomas Walpuski
Massachusetts Institute of Technology

2016-11-10

Contents

1 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 1.1. A *(system of) ordinary differential equations (ODEs) of first order* is an equation of the form

$$\dot{x}(t) = F(t, x(t)) \tag{1.2}$$

with $F\colon U \to \mathbf{R}^n$ and $U \subset \mathbf{R} \times \mathbf{R}^n$ open.

We also call (1.2) a *dynamical system* and call a *time-dependent vector field*.

Remark 1.3. Any system of ODEs of any order can be converted to a system of ODEs of first order by introducing auxiliary variables.

Exercise 1.4. If you unfamiliar with the idea of Remark 1.3, please, read up on this.

Let us introduce some language to talk about solutions of (1.2).

Definition 1.5. A *solution of* (1.2) is a differentiable map $\phi\colon I \to \mathbf{R}^n$ defined on an interval $I \subset \mathbf{R}$ 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 $\phi\colon I \to \mathbf{R}^n$ is an *integral curve* of the time-dependent vector field F.

Definition 1.6. A system of ODE of first order (1.2) together with an equation of the form

$$x(t_0) = x_0 \tag{1.7}$$

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

A solution $\phi\colon I \to \mathbf{R}^n$ to (1.2) is called a *solution to the IVP* (1.2) *and* (1.7) 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 1.8. *Suppose F is continuous. A function $\phi\colon I \to \mathbf{R}^n$ is a solution to the IVP* (1.2) *and* (1.7) *if and only if for all $t \in I$ we have*

$$\phi(t) = \phi(t_0) + \int_{t_0}^{t} F(s, \phi(s))\,\mathrm{d}s. \tag{1.9}$$

Proof. If F is continuous, then any solution ϕ to (1.2) 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. □

1.1 Uniqueness of solutions

Example 1.10 (Failure of Uniqueness)**.** Imagine an infinitely high[1] cylindrical container filled with water and a drain at the bottom. Denote by x the water level (in some suitable units). If the container is empty, then it stays empty and nothing changes; if the water is at level $x \geq 0$, then it drains at rate $2\sqrt{x}$. That is, x is governed by the IVP

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

with

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

For *each* $0 \leq T \leq \infty$, the function $\phi_T \colon \mathbf{R} \to \mathbf{R}$ 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 1.11. A function $F \colon U \to \mathbf{R}^m$ 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

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

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

Remark 1.12. 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 1.13. *Suppose that* $F \colon U \to \mathbf{R}^n$, $(t, x) \mapsto F(t, x)$ *is continuous and for each* $(t, x) \in U$ *there exists a neighbourhood* K *of* $t \in \mathbf{R}$ *and* V *of* $x \in \mathbf{R}^n$ *such that* $K \times V \subset U$ *and there exists a constant* $L > 0$ *such that for all* $s \in K$ *and* $y, z \in V$

$$\frac{|F(s, y) - F(s, z)|}{|y - z|} \leq L. \tag{1.14}$$

[1]This is ridiculous of course, but the model becomes easier to describe assuming this.

Remark 1.15. 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 1.16 (Uniqueness Theorem)**.** *Assume Hypothesis 1.13. If $\phi_1, \phi_2 \colon I \to \mathbf{R}^n$ are solutions to (1.2) and for some $t_0 \in I$ we have*

$$\phi_1(t_0) = \phi_2(t_0),$$

then

$$\phi_1 = \phi_2.$$

Remark 1.17. Note that F in Example 1.10 is not Lipschitz in any neighbourhood of 0.

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

Lemma 1.18 (Grönwall's Lemma)**.** *Let $g \colon I \to \mathbf{R}$ be a continuous function with $g \geq 0$ and $t \in I$. If $A, B \geq 0$ are constants such that for all $t \in I$*

$$g(t) \leq A\left|\int_{t_0}^{t} g(s)\,\mathrm{d}s\right| + B,$$

then for all $t \in I$

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

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

The function defined by

$$G(t) := A\int_{t_0}^{t} g(s)\,\mathrm{d}s + B$$

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 1.16. 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 ϕ_1 and ϕ_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 $t \in \mathbf{R}$ and $V \subset \mathbf{R}^n$ of $x_0 = \phi_1(t_0) = \phi_2(t_0) \in \mathbf{R}^n$ such that $K \times V \subset U$ and (1.14) holds for all $t \in K$ and $y, z \in U$. Using the integral form of the ODE (1.9), for $t \in K \cap I$, we can write

$$|\delta(t)| \leq \left| \int_{t_0}^{t} |F(s, \phi_1(s)) - F(s, \phi_2(s))| \, \mathrm{d}s \right| \leq L \left| \int_{t_0}^{t} |\delta(t)| \, \mathrm{d}s \right|.$$

By Lemma 1.18 with $B = 0$, $\delta = 0$ in $K \cap I$; hence, $J \supset K \cap I$. □

1.2 Local existence theorems

Exercise 1.19. Find a function $F\colon \mathbf{R} \to \mathbf{R}$ such that there is no solution to the IVP

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

Theorem 1.20 (Picard–Lindelöf)**.** *Assume Hypothesis 1.13. For every $(t_0, x_0) \in U$ there is an interval I containing t_0 on which the IVP*

$$\dot{x} = F(t, x) \quad \textit{and} \quad x(t_0) = x_0 \tag{1.21}$$

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

$$\phi(t) = x_0 + \int_{t_0}^{t} F(s, \phi(s)) \, \mathrm{d}s$$

has the shape of a fixed point equation

$$\phi = T\phi$$

where $T\colon X \to X$ is defined by

$$T(\phi)(t) := x_0 + \int_{t_0}^{t} F(s, \phi(s)) \, \mathrm{d}s \tag{1.22}$$

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 1.13, we can fix $\delta, \varepsilon, L > 0$ such that (1.14) holds for all $s \in [t_0 - \delta, t_0 + \delta]$ and $y, z \in \bar{B}_{2\varepsilon}(x_0)$. We can also assume, by making ε and δ smaller, that*

$$\gamma := \delta L$$

and

$$\delta(L\varepsilon + M) \leq \varepsilon$$

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

Step 2. *Set $I := [t_0 - \delta, t_0 + \delta]$ and*

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

The formula (1.22) defines a map $T\colon\ X \to X$.

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

$$\begin{aligned}
|T\phi(t) - x_0| &\leq \left| \int_{t_0}^{t} |F(s, \phi(s)) - F(s, x_0)| + |F(s, x_0)| \,\mathrm{d}s \right| \\
&\leq \delta(L\varepsilon + M) \leq \varepsilon.
\end{aligned}$$

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

This is a question on problem set #1.

Step 4. *$T\colon\ X \to X$ is a contraction.*

It follows from (1.14) that

$$\begin{aligned}
d(T\phi_1, T\phi_2) &= \sup_I \left| \int_{t_0}^{t} F(s, \phi_1(s)) - F(s, \phi_2(s)) \,\mathrm{d}s \right| \\
&\leq \sup_I \left| \int_{t_0}^{t} |F(s, \phi_1(s)) - F(s, \phi_2(s))| \,\mathrm{d}s \right| \\
&\leq \sup_I \int_{t_0}^{t} L|\phi_1(s) - \phi_2(s)| \,\mathrm{d}s \\
&\leq \delta L d(\phi_1, \phi_2) = \gamma d(\phi_1, \phi_2).
\end{aligned}$$

This completes the proof. □

Remark 1.23. The conclusion of Theorem 1.20 holds under weaker assumptions: Peano's theorem asserts that it suffices for F to be continuous.

1.3 Dependence of solutions on initial conditions

Theorem 1.24. *Suppose F is C^1. If $\varepsilon > 0$ and $V \subset \mathbf{R}^n$ is a bounded open subset, such that $[-\varepsilon, \varepsilon] \times \bar{V} \subset U$, then there is a $\delta \in (0, \varepsilon)$ and a C^1 map $\Phi\colon (-\delta, \delta) \times V \to U$ such that*

$$\Phi(0, x_0) = x_0 \tag{1.25}$$

and for each $x_0 \in V$, $t \mapsto \Phi(t, x_0)$ is an integral curve of (1.2).

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 [**Duistermaat2000**]*Lemma B.4.

2 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 [1]*Lectures 1 and 2.

Definition 2.1. A *first order quasi-linear PDE* is a PDE of the form

$$b(x, u(x)) \cdot \nabla u(x) + c(x, u(x))u(x) = 0 \tag{2.2}$$

for a function $u\colon U \to \mathbf{R}$ with $U \subset \mathbf{R}^n$ an open set, $b\colon U \times \mathbf{R}^n \to \mathbf{R}^n$ and $c\colon U \times \mathbf{R}^n \to \mathbf{R}$ smooth.

Remark 2.3. Note that if b, c do not depend on the second variable, then (2.2) is a first order linear PDE.

Remark 2.4. We think of x as a "space variable". Sometimes, however, we work on space-time. In these cases we replace $\mathbf{R}^n$ above by $\mathbf{R}^{n+1}$ and x by (t, x).

2.1 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 find a good coordinate system in which (2.2) becomes very simple.

Proposition 2.5. *Let* $\Psi\colon V \to U$ *be a* C^1*–diffeomorphism and* $u\colon U \to \mathbf{R}$. *Define* $\tilde{u}\colon V \to \mathbf{R}$ *by*

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

The function u *satisfies* (2.2) *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

$$\tilde{b}(x, y) := (\mathrm{d}\Psi(x))^{-1} b(\Psi(x), y)$$
$$\tilde{c}(x, y) := c(\Psi(x), y).$$

Proof. By the chain rule

$$\nabla\tilde{u}(x) = \mathrm{d}\Psi(x)^t \nabla u(\Psi(x)),$$

or, equivalently,

$$\nabla u(\Psi(x)) = \left(\mathrm{d}\Psi(x)^{-1}\right)^t \nabla\tilde{u}(x);$$

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. □

2.2 Warm-up: the transport equation

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

Example 2.6. A *transport equation* is a PDE of the form

$$v(x) \cdot \nabla u(x) = 0 \tag{2.7}$$

for a function $u\colon U \to \mathbf{R}$ with $U \subset \mathbf{R}^n$ open and $v\colon U \to \mathbf{R}^n$ a vector field on U.

The transport equation (2.7) 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 2.8. *Suppose $\phi\colon I \to U$ is a solution of the ODE*

$$\partial_s x(s) = v(x(s)). \tag{2.9}$$

If $u\colon U \to \mathbf{R}$ is a solution of (2.7), then it is constant along ϕ, i.e.,

$$\partial_s u(\phi(s)) = 0.$$

Remark 2.10. The reason we parametrise ϕ by s and not t is that often in (2.2) 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:

$$\partial_s u(\phi(s)) = \partial_s \phi(s) \cdot \nabla u(\phi(s)) = v(\phi(s)) \cdot \nabla u(\phi(s)) = 0.$$ □

Example 2.11. Suppose we want to solve transport equation on $\mathbf{R}^{n+1}$ with

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

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

The solutions of (2.9) are the straight lines

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

hence, the solutions to (2.7) must be functions of the form

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

with $f\colon \mathbf{R}^n \to \mathbf{R}$ 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 (2.7), since u still has a derivatives in the direction of the vector v and this is all that is needed to make sense of (2.7).

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 2.12. Let $U \subset \mathbf{R}^n$ be an open subset and consider a PDE (**??**) for a function $u\colon U \to \mathbf{R}$. Let Γ be a smooth oriented hypersurface with outward-pointing unit normal vector ν, let $k \in \mathbf{N}_0$ and $(f_0, \dots, f_k)\colon \Gamma \to \mathbf{R}^k$ be a k–tuple of functions. We say that $u\colon U \to \mathbf{R}$ solves the *Cauchy problem* for (**??**) with *Cauchy hypersurface* Γ and *Cauchy data* $(f_0, \dots, f_k)$ if

- the function u solves the PDE (**??**), and
- we have

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

Exercise 2.13. In the situation of Example 2.11, find the general form of a solution of the inhomogeneous transport equation

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

Here we take $g\colon \mathbf{R}^{n+1} \to \mathbf{R}$ to be some continuous function. (*Hint*: What is the analogue of Proposition 2.8?)

2.3 The method of characteristics

Suppose that $u\colon U \to \mathbf{R}$ is a solution of (2.2). Let $I \to U, s \mapsto x(s)$ 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 2.14. The *characteristic equation* for (2.2) is the ODE

$$\begin{aligned} \partial_s x(s) &= b(x(s), y(s))) \\ \partial_s y(s) &= -c(x(s), y(s))y(s). \end{aligned} \tag{2.15}$$

If $s \mapsto (x(s), y(s))$ is a solution of (2.15), then we say x is the *projected characteristic* of (2.2).

Remark 2.16. If the coefficient $b(x, u(x))$ in (2.2) does not actually depend on $u(x)$ (this is the case, e.g., for first order linear PDE), then (2.15) 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 (2.2) 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 Γ and Cauchy data $f\colon \Gamma \to \mathbf{R}$. The strategy now is to solve (2.15) with initial conditions of the form $(x_0, f(x_0))$ for $x \in \Gamma$ and piece these together to get a solution of (2.2). 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 Γ at x_0. If there are no such "bad" points on Γ, then we can at least solve near Γ.

Definition 2.17. A point $x_0 \in \Gamma$ is called *non-characteristic* for (2.2) if

$$\nu(x) \cdot b(x_0, f(x_0)) \neq 0.$$

Theorem 2.18 (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 $u\colon V \to \mathbf{R}$ of (2.2) with Cauchy data f on $\Gamma \cap V$.*

Proof. We can choose coordinates $(s, x_1, \dots, x_{n-1})$ near x_0 such that Γ locally is cut out by $s = 0$ and x_0 has coordinates $(0, \dots, 0)$. By abuse of notation we still denote the coefficients of (2.2) by b and c. The normal vector-field ν is now nothing but ∂_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 $(\Phi, \Upsilon)\colon I \times U_0 \to U \times \mathbf{R}$ such that

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

and

$$\begin{aligned} \partial_s \Phi(s, x) &= b(\Phi(s, x), \Upsilon(s, x)) \quad \text{and} \\ \partial_s \Upsilon(s, x) &= -c(\Phi(s, x), \Upsilon(s, x))\Upsilon(s, x). \end{aligned}$$

For $x \in U_0$,

$$\mathrm{d}\Phi(0, x) = \left(b(x, f(x)) \middle| \begin{matrix} 0 \\ \mathrm{id}_{T_x\Gamma} \end{matrix} \right)$$

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, $\Phi\colon I \times U_0 \to U$ is a diffeomorphism near x_0. Denote its image by V.

A function $u\colon V \to \mathbf{R}$ solves (2.2) with Cauchy data f on $V_0 = V \cap \Gamma$ if and only the function $\tilde{u}\colon I \times U_0 \to \mathbf{R}$ 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 2.5. □

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 2.19. Suppose $v = r\partial_r = \sum_{i=1}^n x_i \partial_i$ and $\Gamma = \partial B_1(0)$. Then the Cauchy problem of (2.7) with Cauchy data $f \colon \Gamma \to \mathbf{R}$ has a solution on all of $\bar{B}_1(0)$ if and only if f is constant.

Example 2.20. Suppose $v = \partial_1$ and $\Gamma = \{0\} \times \mathbf{R}^{n-1} \cup \{1\} \times \mathbf{R}^{n-1}$. Then the Cauchy problem of (2.7) with Cauchy data $f \colon \Gamma \to \mathbf{R}$ has a solution on all of $[0, 1] \times \mathbf{R}^{n-1}$ if and only if $f(0, \cdot) = f(1, \cdot)$.

3 The inviscid Burgers' equation

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

Definition 3.1. The *inviscid Burgers' equation* is the following PDE for a function $u\colon [0,\infty)\times\mathbf{R}\to\mathbf{R}$

$$\partial_t u + u\partial_x u = 0. \tag{3.2}$$

Remark 3.3. You might want to think of this is a transport equation where the speed depends on u itself.

Remark 3.4 (Burgers' equation and Navier–Stokes). The incompressible Navier–Stokes equation is the PDE[2]

$$\begin{aligned} \partial_t u + (u\cdot\nabla)u + \nu\Delta u &= -\nabla p \\ \operatorname{div} u &= 0. \end{aligned} \tag{3.5}$$

for a pair of maps $u\colon [0,\infty)\times\mathbf{R}^n\to\mathbf{R}^n$, the *flow velocity*, and $p\colon [0,\infty)\times\mathbf{R}^n\to\mathbf{R}$, the *pressure*.[3] Here ν is a constant. This is a model for the motion of an incompressible viscous fluid.

The second equation in (3.5) is the continuity equation and asserts that no material is created by the flow. If we want to force $\nabla p = 0$, then we can no longer expect this second equation to hold. In dimension one, the equation we arrive at by setting $\nabla p = 0$ and dropping the second equation is *Burger's equation*

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

Setting $\nu = 0$ we obtain (3.2).

Roughly speaking, in (3.5), $(u\cdot\nabla)u$ is a bad term, while the viscosity term $\nu\Delta u$ is good.

Remark 3.6. The inviscid Burgers' equation can be written as

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

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.[4]

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

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

$$\Delta p = -\operatorname{div}\left((u\cdot\nabla)u + \Delta u\right)$$

and the boundary conditions.

[4]The name conservation law comes from the fact that solutions have the property $\partial_t\int_{\mathbf{R}} u(t,x)\,\mathrm{d}x = 0$ (subject to some technical hypotheses of course).

We will be concerned with the Cauchy problem for (3.2) with Cauchy data $f\colon\ \mathbf{R} \to \mathbf{R}$ prescribed along the Cauchy hypersurface $\{0\} \times \mathbf{R}$, that is, we prescribe

$$u(0, x) = f(x).$$

Exercise 3.7 (Conservation of the spatial L^2–norm)**.** Suppose $u\colon\ [0, T) \times \mathbf{R} \to \mathbf{R}$ is a solution of (3.2) and

$$\int_{\mathbf{R}} |u(0, x)|^2 \,\mathrm{d}x = \int_{\mathbf{R}} |f(x)|^2 \,\mathrm{d}x < \infty.$$

Prove that the spatial L^2–norm is conserved, that is, for all for all $t \in [0, T)$ we have

$$\int_{\mathbf{R}} |u(t, x)|^2 \,\mathrm{d}x = \int_{\mathbf{R}} |u(0, x)|^2 \,\mathrm{d}x.$$

Let us now apply the method of characteristics to the inviscid Burgers' equation. The characteristic equation takes the particularly simple form

$$\begin{aligned} \partial_s t(s) &= 1, \\ \partial_s x(s) &= y(s), \\ \partial_s y(s) &= 0. \end{aligned}$$

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 (3.2) in concrete examples.

Example 3.8. 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 (3.2): $u\colon\ [0, \infty) \times \mathbf{R} \to \mathbf{R}$ defined by

$$u(t, x) := \frac{x}{1 + t}.$$

Example 3.9. 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 (3.2) up to $s = 1$ by $u\colon [0,1) \times \mathbf{R} \to \mathbf{R}$ defined by

$$u(t,x) = \frac{x}{t-1}.$$

Exercise 3.10. Find an example of Cauchy data for which no solution $u\colon [0,\varepsilon) \times \mathbf{R} \to \mathbf{R}$ exists for any value of $\varepsilon > 0$.

Example 3.11. 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" $u\colon [0,1) \times \mathbf{R} \to \mathbf{R}$ defined by

$$u(t,x) := \begin{cases} 1 & x \leq t \\ \frac{1-x}{1-t} & x \in [t,1] \\ 0 & x \geq 1 \end{cases} \tag{3.12}$$

I wrote "solution" because $u(t,x)$ is not C^1.

Exercise 3.13. Suppose $\phi \in C_c^\infty((0,1) \times \mathbf{R})$ is a smooth function on $(0,1) \times \mathbf{R}$ with compact support. We call such a ϕ a test function. If u is a C^1–solution of (3.2), then integration by parts shows that

$$\int_{(0,1)\times\mathbf{R}} (\partial_t \phi) u + (\partial_x \phi) \frac{u^2}{2} \, \mathrm{d}t \, \mathrm{d}x = 0. \tag{3.14}$$

Show that this is also true for (3.12).

Remark 3.15. We say that u is a *weak solution* of the inviscid Burgers' equation if (3.14) 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, (3.14) appears quite natural.

Remark 3.16. The solution u can be extended to $t \geq 1$ by

$$u(t, x) = \begin{cases} 1 & x \leq \frac{1+t}{2} \\ 0 & x > \frac{1+t}{2} \end{cases} \tag{3.17}$$

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 3.18 (Criterion for singularity formation)**.** *Let $f \in C^1(\mathbf{R})$. There is a C^1 solution $u\colon [0, \infty) \times \mathbf{R} \to \mathbf{R}$ to the Cauchy problem for (3.2) with Cauchy data f if and only if $f'(x) \geq 0$.*

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

$$x_0 < x_1 \quad \text{and} \quad f(x_0) > f(x_1).$$

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 $s = s_0 > 0$, ξ_1 and ξ_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 $\Phi\colon [0, \infty) \times \mathbf{R} \to [0, \infty) \times \mathbf{R}$ by

$$\Phi(s, x) := (s, x + f(x)s)$$

This is a C^1–map and its derivative is

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

Since

$$\det \mathrm{d}\Phi = 1 + f'(x)t > 0$$

for all $s \geq 0$, Φ is a diffeomorphism (initially locally, but one can see that it is a global diffeomorphism). Thus the method of characteristics provides the desired solution. □

4 Introduction to the Heat Equation

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

Definition 4.1. The *heat equation* is the PDE

$$\partial_t u(t,x) + \Delta u(t,x) = 0 \tag{4.2}$$

for a function $u\colon [0,T)\times U \to \mathbf{R}$ with $U \subset \mathbf{R}^n$ and $T > 0$.

Often one also has to deal with the following inhomogeneous version of (4.2).

Definition 4.3. Let $U \subset \mathbf{R}^n$, $T > 0$ and $\sigma\colon [0,T)\times U \to \mathbf{R}$. The *heat equation with source term σ* is the PDE

$$\partial_t u(t,x) + \Delta u(t,x) = \sigma(t,x). \tag{4.4}$$

for a function $u\colon [0,T)\times U \to \mathbf{R}$.

Remark 4.5. Of course, Δ denotes the Laplace operator (or the Laplacian, for short). It acts on the spacial variables only:

$$\Delta u(t,x) = -\sum_{i=1}^n \partial_{x_i}^2 u(t,x).$$

I prefer to put the minus sign in Δ, because it makes the operator Δ positive. Also, various natural Laplace type operators in geometry come with a minus sign. Δ 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 4.6. The heat equation models the time-evolution of temperature distributions. Since this is not a physics class, I will refrain from discussing how exactly one arrives at the heat equation, but here is a sketch: Consider a small space region U and denote by $E_U(t)$ the amount of thermal energy contained in U. If $u(t,x)$ is the temperature distribution, then in suitable units

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

On the one hand we have, some what trivially,

$$\partial_t E_U = \int_U \partial_t u(t,x);$$

on the other hand if q is the heat flux vector, then by the divergence theorem

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

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 (4.2).

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 4.7 (Parabolic rescaling)**.** *If u solves (4.2), then for each $\lambda > 0$ so does*

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

Exercise 4.8. Prove Proposition 4.7!

4.1 Boundary conditions

We will usually be interested in solving the heat equation (4.2) subject to an *initial condition*

$$u(0,x) = f(x) \quad \text{for all } x \in U. \tag{4.9}$$

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

Definition 4.10. Given $g\colon \partial U \to \mathbf{R}$, we consider the following type of boundary conditions:

- *Dirichlet boundary conditions*

$$u(t,x) = g(x) \quad \text{for all } (t,x) \in (0,T] \times \partial U, \tag{4.11}$$

- *Neumann boundary conditions*

$$\partial_\nu u(t,x) = g(x) \quad \text{for all } (t,x) \in (0,T] \times \partial U, \tag{4.12}$$

- and, given also $\alpha > 0$, *Robin boundary conditions*

$$\partial_\nu u(t,x) + \alpha u(t,x) = g(x) \quad \text{for all } (t,x) \in (0,T] \times \partial U. \tag{4.13}$$

If ∂U is disconnected, we can impose *mixed boundary conditions*, that is, we impose one of the above boundary conditions on each component of ∂U.

⚠ *Remark* 4.14. Mostly we will consider homogeneous boundary conditions, that is, $f = 0$. (The general case can easily be reduced to this case.) If I write some thing 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 is the condition for u to solve (4.2), (4.9) and either (4.11), (4.12), (4.13) 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 4.15 (Duhamel's principle). The problem of solving (4.4) 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 σ, 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_τ denote the solution to (4.2) with Dirichlet boundary conditions and initial condition $\sigma(\tau, \cdot)$, and define

$$u(t, x) = \int_0^t u_\tau(t - \tau, x)\, \mathrm{d}\tau.$$

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

$$\begin{aligned}(\partial_t + \Delta)u(t, x) &= (\partial_t + \Delta) \int_0^t u_\tau(t - s, x)\, \mathrm{d}s \\ &= u_t(0, x) + \int_0^t (\partial_t + \Delta)u_\tau(t - \tau, x)\, \mathrm{d}\tau \\ &= \sigma(t, x).\end{aligned}$$

4.2 A toy model

Initially very formally, although a lot of this can be made rigorous, we can think of u as a map $U\colon [0, T] \to C^2(U)$, i.e., as a path in the infinite dimensional vector space $C^2(U, \mathbf{R})$ and think of (4.2) as an ODE. As a toy model replace $C^2(U)$ by a finite dimensional vector space $\mathbf{R}^n$ and Δ by a matrix $A \in \mathbf{R}^{n\times n}$. The ODE

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

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

$$x(t) = e^{-At} x(0). \tag{4.16}$$

If we knew more about A, then we could also say more about the trajectories defined by (4.16).

Proposition 4.17. *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

$$\int_U (\Delta f) g = \int_U f(\Delta g) \quad \textit{and} \quad \int_U (\Delta f) f \geq 0.$$

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. □

This justifies specialising to the case where A (our model for Δ) is a symmetric matrix with non-negative spectrum. In this case we can understand (4.16) more concretely. From linear algebra we know that there is a orthonormal basis (e_i) of $\mathbf{R}^n$ consisting of eigenvectors. Denote the corresponding eigenvalues by $\lambda_i \geq 0$. We can this write (4.16) as

$$x(t) = \sum_{i=1}^{n} e^{-\lambda_i t} \langle e_i, x(0) \rangle \cdot e_i. \tag{4.18}$$

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.

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

Theorem 4.19. *Let $f \in L^2([0,1])$. There is a unique function $u\colon (0,\infty) \times [0,1] \to \mathbf{R}$ which is smooth, solves (4.2), 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 $U\colon [0,\infty) \to L^2([0,1])$ by*

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

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

$$U(t) = \sum_{n=1}^{\infty} e^{-(n\pi)^2 t} \langle f_n, f\rangle f_n \quad \text{in } L^2.$$

For $n \in \mathbf{N}$, define continuous functions $a_n\colon [0,\infty) \to \mathbf{R}$ by

$$a_n(t) = \langle U(t), f_n\rangle .$$

Since u solves (4.2), 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 $k \in \mathbf{N}_0$ and $T_0 > 0$, the series*

$$u(t,x) := \sum_{n=1}^{\infty} e^{-(n\pi)^2 t} \langle f_n, f\rangle f_n(x) \tag{4.20}$$

converges in $C^k([T_0,\infty)\times[0,1])$. In particular, (4.20) defines a smooth function $u\colon (0,\infty)\times[0,1] \to \mathbf{R}$.

This follows from the same kind of argument as Proposition C.10 and the observation that for $\ell \geq 0$ and $t \geq T_0 > 0$

$$\sum_{n=1}^{\infty} (\pi n)^{\ell} e^{-(\pi n)^2 t} |\langle f_n, f \rangle| \leq \left(\sum_{n=1}^{\infty} (\pi n)^{2\ell} e^{-2(\pi n)^2 T_0} \right)^{1/2} \left(\sum_{n=1}^{\infty} |\langle f_n, f \rangle|^2 \right)^{1/2} < \infty.$$

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 (4.2) and satisfies the Dirichlet boundary conditions. The assertion of about the initial value is obvious. □

Remark 4.21. 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 4.22. 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 $u\colon [0, \infty) \times [0, 1] \to \mathbf{R}$ satisfying the initial value

$$u(t, \cdot) = f.$$

Exercise 4.23. Prove the analogue of Theorem 4.19 for Neumann boundary conditions. (*Hint:* The key is to find a suitable analogue of Theorem C.8.)

Exercise 4.24. Suppose $\sigma \in C^{\infty}([0, \infty) \times [0, 1])$. Give a formula for the solution of (4.4) with homogeneous Dirichlet boundary conditions and initial condition $f = 0$. *Hint:* Use Duhamel's principle.

Exercise 4.25. 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 $u\colon (0, \infty) \times [0, 1] \to \mathbf{R}$ which solves (4.2), 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 4.26. Is there a $f \in L^2([0, 1])$ such that if $u\colon (0, \infty) \times [0, 1]$ is as in Theorem 4.19, then the Fourier series of $u(1, x)$ is

$$u(1, x) = \sum_{n=1}^{\infty} \frac{f_n}{n^2}?$$

5 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 $\mathbf{R}^n$ 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 $f\colon \bar{U}_T \to \mathbf{R}$ which are k–times continuously differentiable in the t–direction and ℓ–times continuously differentiable in the x–direction.

Theorem 5.1 (Uniqueness)**.** *Given functions $\sigma\colon U_T \to \mathbf{R}$, $u_0\colon \bar{U} \to \mathbf{R}$ 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 5.2 (Backwards uniqueness)**.** *Given functions $\sigma\colon U_T \to \mathbf{R}$, $u_T\colon \bar{U} \to \mathbf{R}$ 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 $v = 0$.

5.1 Energy method

Proof of Theorem 5.1. 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

$$\begin{aligned}\dot{E}(t) &= \partial_t \int_U |v(t,x)|^2 \, \mathrm{d}x \\ &= 2 \int_U \langle \partial_t v(t,x), v(t,x) \rangle \, \mathrm{d}x \\ &= -2 \int_U \langle \Delta v(t,x), v(t,x) \rangle \, \mathrm{d}x \leq 0\end{aligned}$$

by Proposition 4.17. Since $E(t) \geq 0$, it follows that $E(t) = 0$ for all t. Thus $v = 0$. □

Proof of Theorem 5.2. 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 4.17,

$$\begin{aligned}\ddot{E}(t) &= -4 \int_U \langle \Delta v(t,x), \partial_t v(t,x) \rangle \, \mathrm{d}x \\ &= 4 \int_U |\Delta v(t,x)|^2 \, \mathrm{d}x \geq 0.\end{aligned}$$

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

$$\dot{E}^2 \leq E\ddot{E}. \tag{5.3}$$

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

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

But (5.3) means that

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

Thus $\log E$ is a convex function which tends to $-\infty$ as $t \to T$. This is impossible. □

Exercise 5.4. If $f \colon [0, T) \to \mathbf{R}$ is a C^2–function with $f'' \geq 0$, then

$$f(t) \geq f(0) + f'(0)t.$$

5.2 Weak maximum principle

Theorem 5.5. *Let $u \in C^{1,2}(\bar{U}_T)$ be a subsolution of the heat equation, i.e.,*

$$\partial_t u + \Delta u \leq 0.$$

Then u attains its maximum in $\bar{U}_T$ on $\partial_p U_T$, i.e.,

$$\max_{\bar{U}_T} u = \max_{\partial_p U_T} u.$$

Note that this gives another proof of Theorem 5.1. The function v satisfies the hypothesis of Theorem 5.5 and $v|_{\partial_p U_T} = 0$ thus $v \leq 0$, and the same holds for $-v$; hence, $v = 0$.

Proof of Theorem 5.5.

Step 1. *The assertion holds if $\partial_t u + \Delta u < 0$.*

If the maximum is attained at a point (t_0, x_0) in U_T, then $\partial_t u(t_0, x_0) \geq 0$—in fact, the case > 0 is possible only if $t_0 = T$. It follows that

$$\Delta u(t_0, x_0) < 0.$$

However, $x_0 \in U$ is also a local maximum for $u(t_0, \cdot)$; hence, the Hessian $\mathrm{Hess}(u) = (\partial_i \partial_j u)$ is negative semi-definite; thus, $\Delta u = -\operatorname{tr} \mathrm{Hess}(u) \geq 0$. This is a contradiction.

Step 2. *We prove the theorem.*

For each $\varepsilon > 0$, define a new function $u_\varepsilon \in C^{1,2}(\bar{U}_T)$ by

$$u_\varepsilon(t, x) := u(t, x) - \varepsilon t.$$

Since

$$(\partial_t + \Delta) u_\varepsilon < 0,$$

the previous step shows that

$$\max_{\bar{U}_T} u_\varepsilon = \max_{\partial_p U_T} u_\varepsilon.$$

Because

$$u_\varepsilon \leq u \leq u_\varepsilon + \varepsilon T,$$

we also have

$$\max_{\partial_p U_T} u - \varepsilon T \leq \max_{\bar{U}_T} u - \varepsilon T \leq \max_{\bar{U}_T} u_\varepsilon \leq \max_{\partial_p U_T} u_\varepsilon \leq \max_{\partial_p U_T} u.$$

Since this is true for all $\varepsilon > 0$, we have

$$\max_{\bar{U}_T} u = \max_{\partial_p U_T} u. \qquad \square$$

Proposition 5.6 (Comparison)**.** *If $u, v \in C^{1,2}(\bar{U}_T)$ satisfy*

$$\begin{aligned} \partial_t u + \Delta u &= \sigma \quad \textit{and} \\ \partial_t v + \Delta v &= \tau, \end{aligned}$$

and

$$u \le v \textit{ on } \partial U_T \quad \textit{and} \quad \sigma \le \tau,$$

then

$$u \le v \textit{ in } \bar{U}_T.$$

Proof. This follows from a direct application of Theorem 5.5 to $u - v$. □

Proposition 5.7 (Stability estimate)**.** *If $u, v \in C^{1,2}(\bar{U}_T)$ satisfy*

$$\begin{aligned} \partial_t u + \Delta u &= \sigma \quad \textit{and} \\ \partial_t v + \Delta v &= \tau, \end{aligned}$$

then

$$\max_{\bar{U}_T} |u - v| \le \max_{\partial_p U_T} |u - v| + T \max_{\bar{U}_T} |\sigma - \tau|.$$

Proof. Set $w := u - v$ and $M := \max_{\bar{U}_T} |\sigma - \tau|$. Both $w - tM$ and $-w - tM$ are subsolutions of the heat equation. Apply Theorem 5.5 to both to obtain

$$\max_{\bar{U}_T} w - tM \le \max_{\partial_p U_T} w - tM \le \max_{\partial_p U_T} w$$

and

$$\max_{\bar{U}_T} -w - tM \le \max_{\partial_p U_T} -w - tM \le \max_{\partial_p U_T} -w.$$

These two inequalities are equivalent to the assertion. □

6 The Heat Kernel on $\mathbf{R}^n$

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

Definition 6.1. The *heat kernel* on $\mathbf{R}^n$ is the function $\Phi\colon (0,\infty)\times\mathbf{R}^n \to [0,\infty)$ defined by

$$\Phi(t,x) := \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}. \tag{6.2}$$

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

Theorem 6.3. *Suppose $f \in C^0(\mathbf{R}^n)$ is bounded. Define $u\colon (0,\infty)\times\mathbf{R}^n \to \mathbf{R}$ by*

$$u(t,x) := \int_{\mathbf{R}^n} \Phi(t,x-y) f(y)\,\mathrm{d}y.$$

Then the function u is smooth, satisfies

$$\partial_t u + \Delta u = 0,$$

and, for each $x \in \mathbf{R}^n$,

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

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

Proposition 6.4. *Φ solves the heat equation, i.e.,*

$$\partial_t \Phi + \Delta\Phi = 0.$$

Exercise 6.5. Prove Proposition 6.4.

Proposition 6.6. *For each $t > 0$,*

$$\int_{\mathbf{R}^n} \Phi(t,x)\,\mathrm{d}x = 1.$$

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

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

To compute the integral in the product note that

$$\left(\int_{\mathbf{R}} e^{-s^2}\,\mathrm{d}s\right)^2 = \int_{\mathbf{R}^2} e^{-|z|^2}\,\mathrm{d}z = \int_0^{2\pi}\int_0^\infty re^{-r^2}\,\mathrm{d}r = 2\pi\int_0^\infty re^{-r^2}\,\mathrm{d}r = \pi.$$

The last step uses

$$\mathrm{d}(e^{-r^2}) = -2re^{-r^2}.$$

□

Proof of Theorem 6.3. It is clear that u is smooth and solves the heat equation (by differentiating under the integral, which is justified because, for each $T_0 > 0$, Φ is bounded in $C^k([T_0, \infty) \times \mathbf{R}^n)$ for each $k \in \mathbf{N}_0$).

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

$$\begin{aligned}
&\left| \int_{\mathbf{R}^n} \Phi(t, x-y) f(y) \, \mathrm{d}y - f(x_0) \right| \\
&\quad = \left| \int_{\mathbf{R}^n} \Phi(t, x-y)(f(y) - f(x_0)) \, \mathrm{d}y \right| \\
&\quad \le \int_{\mathbf{R}^n} \Phi(t, x-y) |f(y) - f(x_0)| \, \mathrm{d}y \\
&\quad \le \int_{B_\varepsilon(x_0)} \Phi(t, x-y) |f(y) - f(x_0)| \, \mathrm{d}y \\
&\qquad + \int_{\mathbf{R}^n \setminus B_\varepsilon(x_0)} \Phi(t, x-y) |f(y) - f(x_0)| \, \mathrm{d}y \\
&\quad =: \mathrm{I}_{\varepsilon,t} + \mathrm{II}_{\varepsilon,t}
\end{aligned}$$

for any $\varepsilon > 0$.

We have

$$\mathrm{I}_{\varepsilon,t} \le \sup_{B_\varepsilon(x_0)} |f(y) - f(x_0)| \to 0 \text{ as } \varepsilon \downarrow 0.$$

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

$$\begin{aligned}
\mathrm{II}_{\varepsilon,t} &\le C_1 \int_{|x| \ge \varepsilon} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \, \mathrm{d}x \\
&= C_2 t^{-n/2} \int_\varepsilon^\infty r^{n-1} e^{-\frac{r^2}{4t}} \, \mathrm{d}r \\
&= C_3 \int_{\varepsilon/2\sqrt{t}}^\infty s^{n-1} e^{-s^2} \, \mathrm{d}s \to 0 \quad \text{as } \varepsilon/\sqrt{t} \to \infty.
\end{aligned}$$

It follows that

$$\lim_{\varepsilon \to 0} \lim_{t \to 0} \mathrm{I}_{\varepsilon,t} + \mathrm{II}_{\varepsilon,t} = 0,$$

which establishes the assertion. □

By Duhamel's principle we also get a formula for the solution of the inhomogeneous heat equation.

Theorem 6.7. *Suppose $\sigma \in C^\infty([0, \infty) \times \mathbf{R}^n)$. Then $u\colon (0, \infty) \times \mathbf{R}^n \to \mathbf{R}$ defined by*

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

is smooth, solves

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

and, for all $x \in \mathbf{R}^n$,

$$\lim_{t \downarrow 0} u(t,x) = 0.$$

The proof of this theorem is similar to (but more involved than) the proof of Theorem 6.3.

Definition 6.8. The *heat ball* of radius a $r > 0$ at $(t,x) \in \mathbf{R}^{n+1}$ is the set

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

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 6.9 (Mean-value property)**.** *Suppose* $U \subset \mathbf{R}^n$ *is open and* $T > 0$*. If* $u \in C^\infty(U_T)$ *is a solution of* (4.2) *and* $B_r^p(t,x) \subset U_T$*, then*

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

Proof sketch. For the detailed proof, see [**Evans2010**]*pp. 53–54. The argument presented there proceeds by proving the following:

- The map
$$r \mapsto \frac{1}{4r^n} \int_{B_r^p(t,x)} u(s,y) \frac{|x-y|^2}{|t-s|^2} \,\mathrm{d}y\mathrm{d}s$$
is constant.

- For all $r > 0$,
$$\frac{1}{4r^n} \int_{B_r^p(t,x)} \frac{|x-y|^2}{|t-s|^2} \,\mathrm{d}y\mathrm{d}s = 1; \tag{6.10}$$
hence,
$$u(t,x) = \lim_{r \downarrow 0} \frac{1}{4r^n} \int_{B_r^p(t,x)} u(s,y) \frac{|x-y|^2}{|t-s|^2} \,\mathrm{d}y\mathrm{d}s.$$

□

Using the mean-value property we can prove:

Theorem 6.11 (Strong maximum principle)**.** *Suppose* U *is a bounded open and* connected *subset of* $\mathbf{R}^n$, $T > 0$ *and* $u \in C^\infty(\bar{U}_T)$ *satisfies* (4.2)*. 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 $B_r^p(t_0, x_0) \subset U_t$, then by (6.10)

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

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 $B_r^p(t_0, x_0)$.

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

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

and with decreasing t–component. Its not hard to see that γ can be covered finitely many heat balls "centered" along the path. Thus u is constant along γ.

Since γ was arbitrary (and by continuity of u), u is constant on $\bar{U}_{t_0}$. □

7 Introduction to the Wave Equation

Definition 7.1. The *wave equation* is the PDE

$$\partial_t^2 u + \Delta u = 0 \tag{7.2}$$

for a function $u\colon I \times U \to R$ with $U \subset \mathbf{R}^n$ open and $I \subset \mathbf{R}$ an interval.

A solution u of (7.2) 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 2.12), that is for fixed $f, g\colon U \to \mathbf{R}$ 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 4.1.

7.1 A toy model

A toy model for the heat equation is the ODE

$$\partial_t^2 x(t) + Ax(t)$$

for a function $x\colon [0, T] \to \mathbf{R}^n$ and symmetric non-negative definite matrix A. In a 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 (4.2). 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.

7.2 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

$$u(t, x) := \sum_{n=1}^\infty \left(a_n \cos(n\pi t) + \frac{b_n}{n\pi} \sin(n\pi t) \right) f_n \tag{7.3}$$

defines a solution of (7.2) 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 7.4. *If $f, g \in C^\infty([0, 1])$ are compactly supported in $(0, 1)$, then (7.3) defines a smooth function $u: \mathbf{R} \times [0, 1] \to \mathbf{R}$ solving (7.2) with initial condition (f, g) and satisfying Dirichlet boundary conditions.*

Exercise 7.5. 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 ℓ and $\|f\|_{C^k}$) such that the Fourier coefficients a_k of f satisfy

$$|a_k| \le c/k^\ell.$$

Exercise 7.6. Use the previous exercise to prove Theorem 7.4.

Remark 7.7. If you're stuck on these exercises, please, consult [1]*Lecture 5.

7.3 d'Alembert's formula

Let us now consider the case of an infinitely long vibrating string, i.e., $U = \mathbf{R}$. 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 $\mathbf{R}$,

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

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

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

i.e., the solutions of the transport equations

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

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

Proposition 7.8. *Every solution u of the wave equation on $\mathbf{R}$ 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 7.9 (Finite propagation speed). Let us first note that Proposition 7.8 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 7.10. Also note that the amplitude of the waves does not decrease in space. A one-dimensional world would be very noisy!

Proof of Proposition 7.8. 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

$$v(q,s) = v(q,0) + \int_0^s \partial_s v(a,0)\,\mathrm{d}a.$$

This translates back to the asserted statement about u. □

The general solution to the wave equation on $\mathbf{R}$ is given by

Theorem 7.11. *Suppose $f \in C^2(\mathbf{R})$ and $g \in C^1(\mathbf{R})$. Then there is a unique solution $u \in C^2(\mathbf{R} \times \mathbf{R})$ to the wave equation with Cauchy data (f, g). This solution can be written as*

$$u(t,x) := \frac{1}{2}\left(f(x+t) + f(x-t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g(y)\,\mathrm{d}y. \tag{7.12}$$

Proof. First note that u defined by (7.12) 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

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

This gives the asserted formula for u. □

Remark 7.13. Note that if $f \in C^k(\mathbf{R})$ and $g \in C^{k-1}(\mathbf{R})$, then $u \in C^k(\mathbf{R})$. However, it typically won't be any smoother.

Exercise 7.14. 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

$$u(t,x) = \begin{cases} \frac{1}{2}\left(f(x+t) + f(x-t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g(y)\,\mathrm{d}y & 0 \le t \le x, \\ \frac{1}{2}\left(f(x+t) - f(t-x)\right) + \frac{1}{2}\int_{-x+t}^{x+t} g(y)\,\mathrm{d}y & 0 \le x \le t. \end{cases}$$

State and prove corresponding analogue of Theorem 7.11. (*Hint:* Use the reflection principle!)

8 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.

8.1 Spherical means

Recall that

$$\fint_U := \frac{1}{\mathrm{vol}(U)} \int_U .$$

Proposition 8.1. *Suppose* $u\colon [0,\infty) \times \mathbf{R}^n \to \mathbf{R}$ *is a* C^2 *solution of* (7.2) *with initial conditions* f *and* g*. Define* $U_x\colon [0,\infty) \times (0,\infty) \to \mathbf{R}$ *and* $F_x, G_x\colon (0,\infty) \to \mathbf{R}$ *by*

$$\begin{aligned} U_x(t,r) &:= \fint_{\partial B_r(x)} u(t,y)\,\mathrm{d}y, \\ F_x(r) &:= \fint_{\partial B_r(x)} f(y)\,\mathrm{d}y, \quad \textit{and} \\ G_x(r) &:= \fint_{\partial B_r(x)} g(y)\,\mathrm{d}y. \end{aligned}$$

Then $U_x\colon [0,\infty) \times (0,\infty)$ *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

$$U_x(0,\cdot) = F_x \quad \textit{and} \quad \partial_t U_x(0,\cdot) = G_x.$$

Moreover,

$$\lim_{r\to 0} U_x(t,r) = u(t,x) \quad \textit{and} \quad \lim_{r\to 0} \partial_t U_x(t,r) = \partial_t u(t,x).$$

Proof. First, note that

$$\partial_t^2 U_x(t,r) = \fint_{\partial B_r(x)} \partial_t^2 u(t,y)\,\mathrm{d}y$$

by differentiating under the integral.

Denote by ω_n the volume of the n–dimensional unit-sphere. We compute

$$\begin{aligned}
\partial_r U_x(t,r) &= \partial_r \left(\frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} u(t,y)\, \mathrm{d}y \right) \\
&= \partial_r \frac{1}{\omega_{n-1}} \int_{\partial B_1(0)} u(t, x + rz)\, \mathrm{d}z \\
&= \frac{1}{\omega_{n-1}} \int_{\partial B_1(0)} \langle \nabla u(t, x + rz), z \rangle \, \mathrm{d}z \\
&= \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} \langle \nabla u(t,y), (y-x)/r \rangle \, \mathrm{d}z \\
&= \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} \nabla_\nu u(t,y)\, \mathrm{d}z \\
&= \frac{1}{\omega_{n-1} r^{n-1}} \int_{B_r(x)} -\Delta u(t,y)\, \mathrm{d}z.
\end{aligned}$$

Here we use Theorem A.4 in the last step.

From this we obtain

$$\begin{aligned}
\partial_r^2 U_x(t,r) = &\frac{1-n}{\omega_{n-1} r^n} \int_{B_r(x)} -\Delta u(t,y)\, \mathrm{d}z. \\
&+ \frac{1}{\omega_{n-1} r^{n-1}} \int_{\partial B_r(x)} -\Delta u(t,y)\, \mathrm{d}z \\
= &-\frac{n-1}{r} \partial_r U_x(t,r) + \fint_{\partial B_r(x)} -\Delta u(t,y)\, \mathrm{d}z.
\end{aligned}$$

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

8.2 Kirchhoff's formula

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

$$\tilde{U}_x = rU_x, \quad \tilde{F}_x = rF_x, \quad \text{and} \quad \tilde{G}_x = rG_x.$$

Remark 8.2. 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 [**Evans2010**]*Section 2.4.d. From the resulting solution formulae for n even can be obtain by dimensional reduction.

From Proposition 8.1 we derive that

$$\begin{aligned}
(\partial_t^2 - \partial_r^2)\tilde{U}_x &= 0 && \text{on } [0,\infty) \times (0,\infty) \\
\tilde{U}_x = \tilde{F}_x \quad \text{and} \quad \partial_t \tilde{U}_x &= \tilde{G}_x && \text{on } \{0\} \times (0,\infty) \\
\tilde{U} &= 0 && \text{on } [0,\infty) \times \{0\}.
\end{aligned}$$

Exercise 8.3. Check this!

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

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

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

$$\begin{aligned} u(t, x) &= \lim_{r \downarrow 0} \frac{\tilde{U}_x(t, r)}{r} \\ &= \partial_r \tilde{U}_x(t, 0) \\ &= \partial_r \tilde{F}_x(t) + \tilde{G}_x(t). \end{aligned}$$

Finally, we compute

$$\begin{aligned} \partial_r \tilde{F}_x(t) &= \partial_t \left(t \fint_{\partial B_t(x)} f(y)\,\mathrm{d}y\right) \\ &= \fint_{\partial B_t(x)} f(y) + t\nabla_\nu f(t)\,\mathrm{d}y \quad \text{and} \\ \tilde{G}_x(t) &= \fint_{\partial B_t(x)} t g(y)\,\mathrm{d}y. \end{aligned}$$

Theorem 8.4. *Suppose $f \in C^3$ and $g \in C^2$. The Cauchy problem for the wave equation on $\mathbf{R}^3$ has a unique solution given by* Kirchhoff's formula

$$u(t, x) := \fint_{\partial B_t(x)} f(y) + t\nabla_\nu f(y) + t g(y)\,\mathrm{d}y. \tag{8.5}$$

Proof. The previous computations show that if there is a solution it must satisfy (8.5). A computation shows that u defined by (8.5) solves the wave equation. □

Remark 8.6 (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 $n \ge 3$ 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 8.7. Note that we *lose* differentiabilty compared to the initial data! (This phenomenon gets worse and worse as n grows.)

Exercise 8.8 (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:

$$u(t, x) = \frac{1}{2}\fint_{B_t(x)} \frac{t f(y) + t\langle \nabla f(y), y - x\rangle + t^2 g(y)}{\sqrt{t^2 - |x - y|^2}}\,\mathrm{d}y \tag{8.9}$$

8.3 Uniqueness of solutions to the Cauchy problem

Suppose U is a bounded open subset of $\mathbf{R}^n$ and $T > 0$

Theorem 8.10. *Given $f, g\colon \bar{U} \to \mathbf{R}$, there exists at most one $u \in C^2([0,T] \times \bar{U})$ such that*

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

with

$$u(0,\cdot) = f \quad \textit{and} \quad \partial_t u(0,\cdot) = g$$

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 $w\colon [0,T]\times U \to \mathbf{R}$ satisfying

$$\begin{aligned} \partial_t^2 w + \Delta w &= 0 \\ w(t,x) &= 0 \quad \text{for } x \in \partial U \\ w(0,\cdot) = \partial_t w(0,\cdot) &= 0. \end{aligned}$$

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

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

Since

$$\begin{aligned} \partial_t E(t) &= \int_U \left\langle \partial_t^2 w, \partial_t w \right\rangle + \langle \nabla w, \nabla \partial_t w \rangle \\ &= \int_U \left\langle \partial_t^2 w + \Delta w, \partial_t w \right\rangle = 0, \end{aligned}$$

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 $w(0,\cdot) = 0$. □

8.4 Finite propagation speed

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

Theorem 8.11. *Let $u \in C^2([0,\infty) \times \mathbf{R}^n$ be a solution of (7.2). Fix $(t_0, x_0) \in (0,\infty) \times \mathbf{R}^n$. 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

$$E(t) := \frac{1}{2}\int_{B_{t_0-t}(x)} |\partial_t w|^2 + |\nabla w|^2.$$

We know that $E(0) = 0$. Now we compute how E changes in t:

$$\begin{aligned}
\partial_t E(t) &= \partial_t \frac{1}{2}\int_{B_{t_0-t}(x)} |\partial_t u|^2 + |\nabla u|^2 \\
&= \int_{B_{t_0-t}(x)} \left\langle \partial_t^2 u, \partial_t u\right\rangle + \langle \nabla u, \nabla \partial_t u\rangle \\
&\quad - \frac{1}{2}\int_{\partial B_{t_0-t}(x)} |\partial_t u|^2 + |\nabla u|^2 \\
&= \int_{B_{t_0-t}(x)} \left\langle \partial_t^2 u + \Delta u, \partial_t u\right\rangle \\
&\quad + \int_{\partial B_{t_0-t}(x)} \langle \partial_\nu u, \partial_t u\rangle - \frac{1}{2}\left(|\partial_t u|^2 + |\nabla u|^2\right) \\
&= \int_{\partial B_{t_0-t}(x)} \langle \partial_\nu u, \partial_t u\rangle - \frac{1}{2}\left(|\partial_t u|^2 + |\nabla u|^2\right) \leq 0.
\end{aligned}$$

Here, the last inequality uses $ab \leq \frac{1}{2}\left(a^2 + b^2\right)$.

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 $x \in B_{t_0}(x_0)$. □

Exercise 8.12 (Equipartition of energy)**.** Suppose $u \in C^2(\mathbf{R} \times [0, \infty))$ solves the Cauchy problem for the have equation with *compactly supported* Cauchy data (f, g). Define the kinetic energy T and the potential energy U by

$$T(t) := \frac{1}{2}\int_{\mathbf{R}} |\partial_t u|^2 \quad \text{and} \quad U(t) := \frac{1}{2}\int_{\mathbf{R}} |\nabla u|^2.$$

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!

9 The Fourier Transform

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

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

$$\mathcal{F}(f)(y) = \hat{f}(y) := \int_{\mathbf{R}^n} f(x)e^{-2\pi i\langle x,y\rangle}\,\mathrm{d}x. \tag{9.1}$$

I write "formally", because actually making sense of (9.1) 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.

9.1 The Fourier Transform on L^1

Proposition 9.2. *If $f \in L^1(\mathbf{R}^n)$, then $\hat{f} \in C^0(\mathbf{R}^n)$ and*

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

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

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

$$|\hat{f}(y)| \leq \int_{\mathbf{R}^n} |f(x)e^{-2\pi i\langle x,y\rangle}|\,\mathrm{d}x \leq \int_{\mathbf{R}^n} |f(x)|\,\mathrm{d}x = |f|_{L^1};$$

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

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

$$\begin{aligned}
|\hat{f}(y) - \hat{f}(z)| &\leq \int_{B_r(0)} |f(x)||e^{-2\pi i\langle x,y\rangle} - e^{-2\pi i\langle x,z\rangle}| \\
&\quad + \int_{\mathbf{R}^n\setminus B_r(0)} |f(x)||e^{-2\pi i\langle x,y\rangle} - e^{-2\pi i\langle x,z\rangle}| \\
&\leq |e^{-2\pi i\langle x,y\rangle} - e^{-2\pi i\langle x,z\rangle}|_{L^\infty(B_r(0))}\|f\|_{L^1} + 2\int_{\mathbf{R}^n\setminus B_r(0)} |f(x)|.
\end{aligned}$$

For fixed $r > 0$, we have

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

Thus for all $r > 0$, we have

$$\lim_{z\to x} |\hat{f}(y) - \hat{f}(z)| \leq 2\int_{\mathbf{R}^n\setminus B_r(0)} |f(x)|.$$

Since

$$\lim_{r\to\infty}\int_{\mathbf{R}^n\setminus B_r(0)}|f(x)|=0,$$

it follows that $\hat{f}$ is continuous. □

Proposition 9.2 has a slight improvement called the Riemann-Lebesgue Lemma.

Lemma 9.3 (Riemann-Lebesgue Lemma)**.** *If $f\in L^1(\mathbf{R}^n)$, then $\hat{f}$ decays at infinity, that is,*

$$\lim_{r\to\infty}\|\hat{f}\|_{L^\infty(\mathbf{R}^n\setminus B_r(0))}=0.$$

Exercise 9.4. 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 9.5. *If $f,\partial_{x_k}f\in L^1(\mathbf{R}^n)$, then*

$$\widehat{\partial_{x_k}f}(y)=2\pi i y_k\hat{f}(y).$$

Conversely, if $f,-2\pi i x_k f\in L^1(\mathbf{R}^n)$, 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

$$\begin{aligned}\widehat{\partial_{x_k}f}(y)&=\int\partial_{x_k}f(x)e^{-2\pi i\langle x,y\rangle}\mathrm{d}x\\&=2\pi i y_k\int f(x)e^{-2\pi i\langle x,y\rangle}\mathrm{d}x.\end{aligned}$$

To prove the second assertion observe that, by the hypothesis,

$$f(x)\partial_{y_k}e^{-2\pi i\langle x,y\rangle}=-2\pi i x_k f(x)e^{-2\pi i\langle x,y\rangle}$$

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

$$\begin{aligned}\partial_{y_k}\hat{f}(y)&=\int f(x)\partial_{y_k}e^{-2\pi i\langle x,y\rangle}\\&=\int -2\pi i x_k f(x)e^{-2\pi i\langle x,y\rangle}=\mathcal{F}(-2\pi i x_k f)(y).\end{aligned}$$

□

Another important property of the Fourier Transform is the way in which it interacts with convolutions. Recall that

$$(f * g)(y) := \int f(y-x)g(x)\,\mathrm{d}x;$$

and if $f, g \in L^1(\mathbf{R}^n)$, then $f * g \in L^1(\mathbf{R}^n)$; in fact,

$$\|f * g\|_{L^1} \leq \|f\|_{L^1}\|g\|_{L^1}.$$

Proposition 9.6. *If $f, g \in L^1(\mathbf{R}^n)$, then*

$$\widehat{f * g} = \hat{f} \cdot \hat{g}.$$

Proof. We compute

$$\begin{aligned}
\widehat{f * g}(x) &= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} f(y-z)g(z)\,\mathrm{d}z \right) e^{-2\pi i \langle x,y \rangle}\,\mathrm{d}y \\
&= \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} f(y-z)e^{-2\pi i \langle x,y-z \rangle} \cdot g(z)e^{-2\pi i \langle x,z \rangle}\,\mathrm{d}y\mathrm{d}z \\
&= \hat{f}(x) \cdot \hat{g}(x).
\end{aligned}$$

□

9.2 Fourier inversion on Schwartz functions

Formally, inverse Fourier Transform of a function $g\colon \mathbf{R}^n \to \mathbf{C}$ is the function $\check{g}\colon \mathbf{R}^n \to \mathbf{C}$ defined by

$$\mathcal{F}^{-1}(g)(y) = \check{g}(y) := \int_{\mathbf{R}^n} g(x)e^{2\pi i \langle x,y \rangle}\,\mathrm{d}x. \tag{9.7}$$

There are various heuristic reasons why $\mathcal{F}^{-1}$ should be the inverse of $\mathcal{F}$. However, $\mathcal{F}\colon L^1 \to L^\infty$ and, a priori, (9.7) does not make any sense if we just know that $g \in L^\infty(\mathbf{R}^n)$. 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 9.8. A function $f \in C^\infty(\mathbf{R}^n)$ is called a *Schwartz function* if for all $k, \ell \in \mathbf{N}_0$

$$\sup_{x \in \mathbf{R}^n} |x|^k |\nabla^\ell f| < \infty.$$

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

Proposition 9.9. *The Fourier transform defines a linear map $\mathcal{F}\colon \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n)$.*

Proof. This is a consequence of Proposition 9.5. □

Theorem 9.10 (Fourier inversion)**.** *If $f \in \mathcal{S}(\mathbf{R}^n)$, 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

$$\mathcal{F}^{-1} \circ \mathcal{F}(f)(x) = \int \int f(z) \cdot e^{-2\pi i \langle z-x, y\rangle} \, dz \, dy.$$

Interchanging the integral we get

$$= \int f(z) \left(\int e^{-2\pi i \langle z-x, y\rangle} \, dy \right) dz.$$

The integral in parenthesis can be computed to be the distribution $\delta(y-z)$. 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 $|f(z) \cdot e^{-2\pi i \langle z, y-x\rangle}|$ is not integrable on $\mathbf{R}^n \times \mathbf{R}^n$, 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" $e^{-2\pi i \langle z, y-x\rangle}$. Before we can get started with this we need to compute the Fourier Transform of a Gaussian.

Proposition 9.11 (Fourier Transform of a Gaussian)**.** *For all $a > 0$, we have*

$$\widehat{e^{-a\pi|\cdot|^2}} = a^{-\frac{n}{2}} e^{-\frac{\pi}{a}|\cdot|^2}.$$

Proof. For $\lambda > 0$ and $f \in L^1(\mathbf{R}^n)$, we have

$$\widehat{f(\lambda\, \cdot)} = \lambda^{-n} \hat{f}(\cdot/\lambda).$$

Thus it suffices verify the assertion for $a = 1$. Moreover, by Fubini, it suffices to consider $n = 1$.

We compute

$$\begin{aligned} \widehat{e^{-\pi(\cdot)^2}}(y) &= \int_{\mathbf{R}} e^{-\pi x^2 + 2\pi i x y} \, dx \\ &= e^{-\pi y^2} \int_{\mathbf{R}} e^{-\pi(x+iy)^2} \, dx \end{aligned}$$

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 $\{z \in \mathbf{C} : \mathrm{Im}(z) \in [a, b]\}$ thus

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

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

Proof of Theorem 9.10. Given $f \in \mathcal{S}(\mathbf{R}^n)$, we need to show that

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

The function

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

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

$$\begin{aligned}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} & f(z) \cdot e^{-2\pi i\langle z-x,y\rangle - 4\pi^2 t|y|^2}\,\mathrm{d}z\,\mathrm{d}y \\ &= \int_{\mathbf{R}^n} f(z)\left(\int_{\mathbf{R}^n} e^{-2\pi i\langle z-x,y\rangle - 4\pi^2 t|y|^2}\,\mathrm{d}y\right)\mathrm{d}z.\end{aligned}$$

By Proposition 9.11, 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 $\mathbf{R}^n$.) Thus the above integral is

$$(f * K_t)(z).$$

In the lecture on the heat kernel we saw that

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

We also have

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} f(z) \cdot e^{-2\pi i\langle z-x,y\rangle - 4\pi^2 t|y|^2}\,\mathrm{d}z\,\mathrm{d}y = \int_{\mathbf{R}^n} \hat{f}(y) e^{2\pi i\langle x,y\rangle - 4\pi^2 t|y|^2}\mathrm{d}y.$$

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

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

10 The Fourier Transform (continued)

10.1 Plancherel's theorem

Theorem 10.1 (Plancherel's theorem)**.** *If $f, g \in \mathcal{S}(\mathbf{R}^n)$, then*

$$\langle f, g\rangle_{L^2} = \left\langle \hat{f}, \hat{g}\right\rangle_{L^2}$$

Proof. By Theorem 9.10, we can assume that $g = \check{h}$. The assertion is then equivalent to

$$\left\langle f, \check{h}\right\rangle_{L^2} = \left\langle \hat{f}, h\right\rangle_{L^2}.$$

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

$$\begin{aligned}
\left\langle f, \check{h}\right\rangle_{L^2} &= \int f(x)\bar{\check{h}}(x)\,\mathrm{d}x \\
&= \int f(x)\left(\int \bar{h}(y)e^{-2\pi i\langle x,y\rangle}\,\mathrm{d}y\right)\mathrm{d}x \\
&= \int\int f(x)e^{-2\pi i\langle x,y\rangle}\cdot\bar{h}(y)\,\mathrm{d}y\mathrm{d}x \\
&= \int\left(\int f(x)e^{-2\pi i\langle x,y\rangle}\,\mathrm{d}x\right)\cdot\bar{h}(y)\,\mathrm{d}y \\
&= \int \hat{f}(y)\bar{h}(y)\,\mathrm{d}y = \left\langle \hat{f}, h\right\rangle_{L^2}.
\end{aligned}$$ □

Proposition 10.2. *There is a unique bounded linear map $\mathcal{F}: L^2(\mathbf{R}^n) \to L^2(\mathbf{R}^n)$ which agrees with the Fourier Transform on $\mathcal{S}(\mathbf{R}^n)$. This map is an isometry.*

Proof. The set $\mathcal{S}(\mathbf{R}^n)$ is dense in $L^2(\mathbf{R}^n)$. Thus for any $f \in L^2(\mathbf{R}^n)$ we can pick a sequence (f_i) in $\mathcal{S}(\mathbf{R}^n)$ which converges to f in L^2. By Theorem 10.1,

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

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 $\mathcal{S}(\mathbf{R}^n)$ converging to f, then

$$\lim_{i\to\infty}\|\mathcal{F}(f_i) - \mathcal{F}(\tilde{f}_i)\|_{L^2} \le \lim_{i\to\infty}\|f_i - f\|_{L^2} + \|\tilde{f}_i - f\|_{L^2} = 0.$$

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 $\mathcal{S}(\mathbf{R}^n)$ immediately implies the linearity of the extension to $L^2(\mathbf{R}^n)$. It follows from Theorem 10.1 that $\mathcal{F}$ is an isometry. □

10.2 Tempered distributions

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

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

In the proof of Theorem 10.1 we saw that if $f, g \in \mathcal{S}(\mathbf{R}^n)$, then

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

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

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

Note that the right-hand side of this equation makes sense, even if f is not in $\mathcal{S}(\mathbf{R}^n)$. 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 $\mathcal{S}(\mathbf{R}^n)$. However, since $\mathcal{S}(\mathbf{R}^n)$ is an infinite dimensional vector space one has to be a bit careful with what one means by "the dual space".

Definition 10.3. A *tempered distribution* T is a linear map $T\colon\ \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$ with the following continuity property: If (ϕ_i) is a sequence in $\mathcal{S}(\mathbf{R}^n)$ such that for all $k, \ell \in \mathbf{N}_0$

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

then

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

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

Example 10.4. Suppose $f\colon\ \mathbf{R}^n \to \mathbf{C}$ is *slowly growing*, that is, there is some $k \in \mathbf{N}_0$ and $c > 0$ such that

$$|f(x)| \le c(1 + |x|)^k.$$

Then

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

defines a tempered distribution.

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

Example 10.5. The "δ–function"

$$\delta(\phi) := \phi(0)$$

is a tempered distribution.

Definition 10.6. 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

$$\mathcal{F}\colon\ \mathcal{S}^*(\mathbf{R}^n) \to \mathcal{S}^*(\mathbf{R}^n) \quad \text{and} \quad \mathcal{F}^{-1}\colon\ \mathcal{S}^*(\mathbf{R}^n) \to \mathcal{S}^*(\mathbf{R}^n).$$

By Theorem 9.10 we have

$$\mathcal{F} \circ \mathcal{F}^{-1} = \mathrm{id}_{\mathcal{S}^*(\mathbf{R}^n)} \quad \text{and} \quad \mathcal{F}^{-1} \circ \mathcal{F} = \mathrm{id}_{\mathcal{S}^*(\mathbf{R}^n)}.$$

Note that "trick" allows us to make sense of the Fourier Transform even for functions which are slowly growing and for which (9.1) 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 9.10 was proving this identity.

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

Definition 10.7 (Differentiation)**.** Suppose T is a tempered distribution on $\mathbf{R}^n$. We define its derivative by x_k as

$$(\partial_{x_k} T)(\phi) := T(-\partial_{x_k}\phi).$$

Definition 10.8 (Multiplication by a function)**.** Suppose T is a tempered distribution on $\mathbf{R}^n$. If f is smooth and f all of its derivatives (of arbitrary order) are slowly growing, then we define

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

Definition 10.9 (Convolution)**.** Suppose T is a tempered distribution on $\mathbf{R}^n$ and $f \in \mathcal{S}(\mathbf{R}^n)$. We define $T * f \in C^\infty(\mathbf{R}^n)$ by

$$(T * f)(\phi) = T(f(-\cdot) * \phi)$$

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

10.3 A derivation of the Heat Kernel

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

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

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

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

Thus

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

By Proposition 9.11,

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

Thus by Proposition 9.6 and Theorem 9.10

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

Exercise 10.10. Find the fundamental solution of Schrödinger's equation

$$-i\partial_t \psi + \Delta \psi = 0. \tag{10.11}$$

11 Introduction Curve-Shortening Flow

There are no notes for this lecture, and the content of this lectures is not examinable.

12 Laplace's and Poisson's equations

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

Definition 12.1. *Laplace's equation* is the PDE

$$\Delta u = 0 \tag{12.2}$$

for a function $u\colon U \to \mathbf{R}$. A function $u\colon U \to \mathbf{R}$ is called *harmonic* if it satisfies (12.2).

Definition 12.3. Given $f\colon U \to \mathbf{R}$, *Poisson's equation* is the PDE

$$\Delta u = g \tag{12.4}$$

for a function $u\colon U \to \mathbf{R}$.

Remark 12.5. We can think of any solution to (12.2) and (12.4) as a *steady state* solution of an (inhomogeneous) heat or wave equation.

Note that since (12.2) and (12.4) 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) (12.2) and (12.4) can be used to great profit. Even if you really only care about time-dependent PDE, studying the Laplace operator Δ in depth will be important: note, for example, how the solution to the heat equation is basically equivalent to understanding all possible solutions (u, λ) of the eigenvalue problem

$$\Delta u = \lambda^2 u.$$

12.1 Dirichlet's principle

Suppose U is a bounded domain in $\mathbf{R}^n$ with smooth boundary and $f\colon \partial U \to \mathbf{R}$ 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 ϕ for which the right-hand side of the above does not vanish.

This shows that *if* there is a smooth minimiser $u \in C_f^\infty(\bar{U})$ 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.

12.2 Connection with holomorphic functions

Definition 12.6. Let $U \subset \mathbf{C}$ be a domain in the complex plane. A complex-valued function $f\colon U \to \mathbf{C}$ 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 12.7. A complex-valued function $f\colon U \to \mathbf{C}$ is holomorphic if and only if it satisfies the *Cauchy–Riemann equation*

$$\frac{\partial f}{\partial \bar{z}} = 0 \tag{12.8}$$

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 $u :=$ Re f and $v :=$ Im f, and $x :=$ Re z and $y =$ Re z.

Remark 12.9. The Cauchy–Riemann equation is a *system* of first order PDE on the vector-valued function (u, v).

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 12.10. *If $f\colon U \to \mathbf{R}$ is holomorphic, then both $u :=$ Re f and $v :=$ Im f are harmonic.*

Proof. Holomorphic functions are C^2—in fact, analytic. By (12.8),

$$\partial_x u = \partial_y v \quad \text{and} \quad \partial_y u = -\partial_x v.$$

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 12.11. The converse of this is also true: every harmonic function $u\colon U \to \mathbf{R}$ on a domain $U \subset \mathbf{C}$ 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.

12.3 Uniqueness via the energy method

Theorem 12.12. *Suppose U is bounded with C^1 boundary. There is at most one function $u\colon C^2(U) \cap C^1(\bar{U}) \to \mathbf{R}$ satisfying Poisson's equation (12.4) satisfying Dirichlet, Robin or mixed boundary conditions (cf. Section 4.1). 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

$$0 = \int_U \langle \Delta w, w \rangle \geq \int_U |\nabla w|^2.$$

It follows that $\nabla w = 0$. □

12.4 The weak maximum principle

Theorem 12.13. *Suppose U is bounded and $f \in C^2(\bar{U})$ satisfies*

$$\Delta u \leq 0,$$

then

$$\max_{\bar{U}} u = \max_{\partial U} u.$$

Exercise 12.14. Prove Theorem 12.13.

Exercise 12.15. Suppose $f \in C^2(\mathbf{R}^n)$ is harmonic and "decays at infinity" in the sense that

$$\lim_{r\to\infty} \sup_{\mathbf{R}^n \setminus B_r(0)} |f| = 0.$$

Show that $f = 0$.

12.5 $W^{2,2}$–estimates

Theorem 12.16. *For $f \in C_c^\infty(\mathbf{R}^n)$, we have*

$$\|\nabla^2 f\|_{L^2} = \|\Delta f\|_{L^2}.$$

Proof. This is a simple computation

$$\begin{aligned}
\|\nabla^2 f\|_{L^2}^2 &= \sum_{i,j=1}^n \int_{\mathbf{R}^n} (\partial_i\partial_j f)(\partial_i\partial_j f) \\
&= \sum_{i,j=1}^n \int_{\mathbf{R}^n} (\partial_i\partial_j f)(\partial_j\partial_i f) \\
&= \sum_{i,j=1}^n \int_{\mathbf{R}^n} (-\partial_j\partial_i\partial_j f)(\partial_i f) \\
&= \sum_{i,j=1}^n \int_{\mathbf{R}^n} (-\partial_i\partial_j\partial_j f)(\partial_i f) \\
&= \sum_{i,j=1}^n \int_{\mathbf{R}^n} (-\partial_j\partial_j f)(-\partial_i\partial_i f) \\
&= \|\Delta f\|_{L^2}.
\end{aligned}$$

□

Remark 12.17. 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).

12.6 Dirichlet's principle (cont.)

Can we construct such a minima for E? Let us try the *direct method* E is bounded below. Thus $E_0 := \inf\{E(u) : u \in C_f^\infty(U)\} \geq 0$ (in particular, it is finite) and we can find a sequence (u_i) such that

$$E(u_i) \to E_0.$$

If there where a u_∞ such that (some subsequence) of (u_i) converges to u_∞ in C^∞, then $E(u_\infty) = E_0$. The problem is that *a priori* such a u_∞ does not need to exist, and even if it existed the minimising sequence will typically not converge in C^∞.

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 12.18. 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 $W_0^{1,2}(U)$ the completion of $C^\infty(U)$ and $C_0^\infty(U)$ with respect to $\|\cdot\|_{W^{1,2}}$ respectively.

Remark 12.19. In the next lecture we will proof the Dirichlet–Poincaré inequality, which says that if U is bounded and $u \in C_0^\infty(U)$, 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 $W_0^{1,2}(U)$.

Now, consider the functional

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

Proposition 12.20. *A function $u \in W_0^{1,2}(U)$ satisfies*

$$E(u) = \inf\{E(v) : v \in W_0^{1,2}(U)\}$$

if and only if

$$\Delta u = g$$

in the weak sense, that is, for all $\phi \in W_0^{1,2}(U)$

$$\int_U (\Delta u)\phi = \int_U g\phi.$$

Proof. $E\colon W_0^{1,2}(U) \to \mathbf{R}$ is a smooth map and its derivative at u is

$$\mathrm{d}E_u[\hat{u}] = \int_U (\Delta u - g)\hat{u}.$$

If u minimizes E_ρ, this vanishes for all $\hat{f}u \in W_0^{1,2}(U)$; hence u satisfies $\Delta u = g$ in the weak sense.

Conversely, if u satisfies $\Delta u = g$ in the weak sense and $v \in W_0^{1,2}(U)$, then

$$\begin{aligned}\int_U |\nabla u|^2 - ug &= \int_U (\Delta u)u - ug \\ &= \int_U (\Delta u)v - vg \\ &= \int_U \langle \nabla u, \nabla v\rangle - vg \\ &\leq \int_U \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|\nabla v|^2 - vg.\end{aligned}$$

Thus $E(u) \leq E(v)$. □

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

$$\inf\{E(v) : v \in W_0^{1,2}(U)\} = \lim_{i\to\infty} E(u_i).$$

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

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

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

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

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

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

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 12.22. 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 $u \in W_0^{1,2}(U)$ of the minimizing sequence u_i. The following guarantees that u actually minimizes E.

Proposition 12.23 (weak-* lower semi-continuity of E)**.** *Suppose $u_i \in (W_0^{1,2}(U))^{\mathbf{N}}$ weakly converges to $u \in W_0^{1,2}(U)$. Then*

$$E(u) \le \liminf_{i\to\infty} E(u_i).$$

Proof. By weak convergence

$$\begin{aligned}\int_U |\nabla u|^2 - gu &= \lim_i \int \langle \nabla u_i, \nabla u\rangle - gu_i \\ &\le \liminf_i \int \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|\nabla u_i|^2 - gu_i.\end{aligned}$$

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 12.24. 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 $W_0^{1,2}(U)$.

13 The Poincaré inequalities

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

13.1 Dirichlet–Poincaré inequality

Theorem 13.1 (Dirichlet–Poincaré inequality)**.** *There is a constant $c > 0$ such that for all $u \in W_0^{1,2}(U)$ we have*

$$\int_U |u|^2 \le c \int_U |\nabla u|^2. \tag{13.2}$$

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

$$u(x_1, x_2, \dots x_n) = \int_0^{x_1} \partial_1 u(s, x_2, \dots, x_n)\,\mathrm{d}s,$$

we have

$$|u(x_1, x_2, \dots x_n)| \le \int_0^1 |\partial_1 u(s, x_2, \dots, x_n)|\,\mathrm{d}s.$$

By Cauchy–Schwarz,

$$\int_0^1 |\partial_1 u(s, x_2, \dots, x_n)|\,\mathrm{d}s \le \left(\int_0^1 |\partial_1 u(s, x_2, \dots, x_n)|^2\,\mathrm{d}s\right)^{1/2}.$$

Thus

$$|u(x_1, \dots, x_n)|^2 \le \int_0^1 |\partial_1 u(s, x_2, \dots, x_n)|^2\,\mathrm{d}s$$

Integrating over x_1 yields

$$\int_0^1 |u(x_1, \dots, x_n)|^2\,\mathrm{d}x_1 \le \int_0^1 |\partial_1 u(x_1, x_2, \dots, x_n)|^2\,\mathrm{d}x_1,$$

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

Exercise 13.3. Show that there is a constant $c > 0$ (depending only on n) such that for all $r > 0$ and all $u \in W_0^{1,2}(B_r(0))$ we have

$$\int_{B_r(0)} |u|^2 \le cr^2 \int_{B_r(0)} |\nabla u|^2.$$

Exercise 13.4. 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

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

Hint: Adapt the proof of Theorem 13.1 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 13.5. 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 λ_1^D is the smallest eigenvalue of Δ with Dirichlet boundary condition. This is why Theorem 13.1 is called the Dirichlet–Poincaré inequality.

Exercise 13.6. 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 (13.2) cannot hold. Because locally constant functions violate this bound. However this is the only problem.

13.2 Neumann–Poincaré inequality

Theorem 13.7 (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

$$\bar{u} := \fint_U u.$$

Proof for $U = B_r(0)$. It suffices to consider $B_1 := B_1(0)$. We compute

$$\begin{aligned}
\int_{B_1} |u - \bar{u}|^2 &= \int_{B_1} \left| \fint_{B_1} u(x) - u(y)\, \mathrm{d}y \right|^2 \mathrm{d}x \\
&= \mathrm{vol}(B_1)^{-2} \int_{B_1} \left| \int_{B_1} u(x) - u(y)\, \mathrm{d}y \right|^2 \mathrm{d}x \\
&\leq \mathrm{vol}(B_1)^{-1} \int_{B_1} \int_{B_1} |u(x) - u(y)|^2\, \mathrm{d}y\, \mathrm{d}x.
\end{aligned}$$

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

$$
\begin{aligned}
|u(x) - u(y)| &\le \int_0^1 |\partial_t u(x + t(y - x))| \, \mathrm{d}t \\
&\le |x - y| \int_0^1 |(\nabla u)(x + t(y - x))| \, \mathrm{d}t \\
&\le 2 \left(\int_0^1 |(\nabla u)(x + t(y - x))|^2 \, \mathrm{d}t \right)^{1/2}.
\end{aligned}
$$

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

$$
\begin{aligned}
\int_{B_1} |u - \bar{u}|^2 &\le \frac{4}{\mathrm{vol}(B_1)} \int_{B_1} \int_{B_1} \int_0^1 |(\nabla u)(x + t(y - x))|^2 \, \mathrm{d}t \, \mathrm{d}y \, \mathrm{d}x \\
&= \frac{8}{\mathrm{vol}(B_1)} \int_{B_1} \int_{B_1} \int_0^{1/2} |(\nabla u)(x + t(y - x))|^2 \, \mathrm{d}t \, \mathrm{d}y \, \mathrm{d}x \\
&= \frac{8}{\mathrm{vol}(B_1)} \int_0^{1/2} \int_{B_1} \frac{1}{1 - t} \int_{B_{1-t}(ty)} |(\nabla u)(z)|^2 \, \mathrm{d}z \, \mathrm{d}y \, \mathrm{d}t \\
&\le 8 \int_{B_1} |(\nabla u)(z)|^2.
\end{aligned}
$$

□

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

$$
\int_U |u_k - \bar{u}_k|^2 \ge k \int_U |\nabla u_k|^2. \tag{13.8}
$$

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 13.9. 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

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

13.3 The Li–Schoen proof of the Dirichlet–Poincaré inequality

The above proof of Theorem 13.1 is the "standard proof". Another proof of the L^1 Dirichlet–Poincaré inequality was discovered by Peter Li and Richard Schoen [4].[5]

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

Proposition 13.10. *Fix $y \in \partial B_2$, and consider the function $\rho \colon B_1 \to (0, \infty)$ 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 $c > 0$

$$\begin{aligned} \Delta e^{-c\rho} &= ce^{-c\rho}(-\Delta\rho - c|\nabla\rho|^2) \\ &= ce^{-c\rho}(-\Delta\rho - c) \\ &= ce^{-c\rho}((n-1)\rho^{-1} - c) \\ &\leq ce^{-c\rho}((n-1) - c). \end{aligned}$$

[5] I do *not* expect you to be able to reconstruct this proof.

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 13.11. 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 $\frac{1}{p} + \frac{p-1}{p} = 1$.)

14 Mean-value properties of harmonic functions

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

Theorem 14.1 (Mean-value property)**.** *Let $u\colon U \to \mathbf{R}$ be a harmonic function. If $B_r(x) \subset U$, then*

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

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

To prove the first identity, we define $f\colon (0, r] \to \mathbf{R}$ by

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

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:

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

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

Remark 14.2. Compare this with Theorem 6.9.

Remark 14.3. In the next lecture we will see that the converse is also true, see Theorem 15.12.

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

14.1 Strong maximum principle

Theorem 14.4 (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

$$u(x) = M := \max_{\bar{U}} u.$$

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

Exercise 14.5. Suppose $u, v\colon U \to \mathbf{R}$ are both harmonic and set

$$f := u|_{\partial U} \quad \text{and} \quad g := v|_{\partial U}.$$

Prove that

$$f \geq g, \text{but } f \neq g \implies u > v \quad \text{in } U.$$

Also, show that

$$\|u - v\|_{L^\infty(U)} \leq \|f - g\|_{L^\infty(\partial U)}.$$

14.2 C^k–estimates

Theorem 14.6. *For each $k \in \mathbf{N}_0$, 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 14.1.

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 14.1, using Theorem A.4, and the induction hypothesis:

$$\begin{aligned}
|\partial_{i_{k+1}}\partial_{i_k}\cdots\partial_{i_1}u|(x) &= \left|\fint_{B_{r/2}(x)} \partial_{i_{k+1}}\partial_{i_k}\cdots\partial_{i_1}u\right| \\
&\leq \frac{2n}{r}\left|\fint_{\partial B_{r/2}(x)} (\partial_{i_k}\cdots\partial_{i_1}u)\nu_{i_{k+1}}\right| \\
&\leq \frac{2n}{r}\|\partial_{i_k}\cdots\partial_{i_1}u\|_{L^\infty(B_{r/2}(x))} \\
&\leq \frac{2^{n+k+1}nc_k}{r^{n+k+1}}\|u\|_{L^1(B_r(x))} \\
&= \frac{c_{k+1}}{r^{n+k+1}}\|u\|_{L^1(B_r(x))}.
\end{aligned}$$

Here we used that $\operatorname{vol}(\partial B_r) = n\operatorname{vol}(B_r)$. □

14.3 Liouville's Theorem

Theorem 14.7 (Liouville)**.** *Let $u\colon \mathbf{R}^n \to \mathbf{R}$ be a harmonic function. If there exists a constant $M > 0$ such that $|u| \leq M$ on all of $\mathbf{R}^n$, then u is constant.*

Proof. By Theorem 14.6

$$|\nabla u|(x) \leq \frac{c}{r}\fint_{B_r(x)} |u| \leq \frac{cM}{r}.$$

Since $r > 0$ is arbitrary, it follows that $\nabla u = 0$. □

Exercise 14.8. Let $u\colon \mathbf{R}^n \to \mathbf{R}$ 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 14.9. Note that there are plenty of harmonic functions with that satisfy the above hypothesis with $\alpha = 1$!

14.4 Harnack's Inequality

Theorem 14.10 (Harnack's inequality)**.** *Suppose $V \subset \mathbf{R}^n$ is an open subset such that $\bar{V} \subset U$. Then there exists a constant $c > 0$ (depending on V) such that all whenever $u\colon U \to \mathbf{R}$ 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

$$u(x) = \fint_{B_{r/2}(x)} u \leq 2^n \fint_{B_r(y)} u = 2^n u(y).$$

(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 $x, y \in V$. ◻

15 Weyl's Lemma

Definition 15.1. A function $u\colon U \to \mathbf{R}$ is called *weakly harmonic* if $u \in L^1_{\mathrm{loc}}$, 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 15.2. Prove that if u is slowly growing (see Section 10.2) and weakly harmonic on $\mathbf{R}^n$, and if

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

denotes the heat kernel, then

$$\int_{\mathbf{R}^n} (\Delta_x K_t(x-y))u(y)\,\mathrm{d}y = 0.$$

(Δ_x means that we take the Laplacian of $K_t(x-y)$ as a function of x.)

Theorem 15.3 (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 15.4. Theorem 15.3 extends to the space of Schwartz distributions $\mathcal{D}(U)$, the dual space of $C_0^\infty(U)$.

Remark 15.5. Note that $\mathrm{Re}(\frac{1}{z})\colon \mathbf{C}^* \to \mathbf{R}$ is not in $L^1_{\mathrm{loc}}(\mathbf{R}^2)$.

15.1 Proof using Heat Kernel

Proof of Theorem 15.3 using the heat kernel. We begin with the case $U = \mathbf{R}^n$ 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 $\mathbf{R}^n$. Define

$$u_t(x) := (K_t * u)(x) = \int_{\mathbf{R}^n} K_t(x-y)u(y)\,\mathrm{d}y.$$

From Section 6 we know that for each $t > 0$, u_t is smooth, and for each $x \in \mathbf{R}^n$,

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

[6] In case you don't know what this means assume u to be continuous.

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

$$\partial_t K_t = -\Delta K_t$$

and Exercise 15.2, we compute that for each $t > 0$

$$\begin{aligned}\partial_t u_t(x) &= \int_{\mathbf{R}^n} \partial_t K_t(x-y)u(y)\,\mathrm{d}y \\ &= -\int_{\mathbf{R}^n} (\Delta K_t(x-y))u(y)\,\mathrm{d}y \\ &= 0.\end{aligned}$$

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

Remark 15.6. 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 $\mathbf{R}^n$. This, however, is easily remedied. Fix $r > 0$ such that $B_{4r}(x_0) \subset U$ and fix a smooth function $\chi\colon \mathbf{R}^n \to [0,1]$ supported in $B_{4r}(x_0)$ and equal to one on $B_{2r}(x_0)$. Define $v \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ by

$$v := \chi \cdot u.$$

Set

$$v_t(x) := (K_t * v)(x).$$

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 15.7. *For each $k \geq 0$, there is a constant $c_k > 0$ (independent of t) such that*

$$\|\partial_t v_t\|_{C^k(\bar{B}_1(x_0))} \leq c_k.$$

As a consequence of this, the sequence $(v_{1/i})$ is Cauchy in each $C^k(\bar{B}_1(x_0))$; hence, they converge to a smooth limit function which must agree with $u|_{\bar{B}_1(x_0)}$ because this is the point-wise limit.

The proof of Proposition 15.7 is very simple in principle, but writing out all the details will be a bit hairy.

Let's compute

$$\begin{aligned}\partial_t v_t(x) &= \int_{\mathbf{R}^n} \partial_t K_t(x-y)\chi(y)u(y)\,\mathrm{d}y \\ &= -\int_{\mathbf{R}^n} (\Delta_x K_t(x-y))\chi(y)u(y)\,\mathrm{d}y \\ &= -\int_{\mathbf{R}^n} (\Delta_y K_t(x-y))\chi(y)u(y)\,\mathrm{d}y \\ &= -\int_{\mathbf{R}^n} [\Delta_y(K_t(x-y)\chi(y))]u(y)\,\mathrm{d}y \\ &\quad -\int_{\mathbf{R}^n} L_t(x,y)u(y)\,\mathrm{d}y \\ &= -\int_{\mathbf{R}^n} L_t(x,y)u(y)\,\mathrm{d}y.\end{aligned}$$

with

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

The last step uses that u is weakly harmonic. Now the $\nabla\chi$ and $\Delta\chi$ are supported in the annulus $A := B_{4r}(x_0) \setminus \bar{B}_{2r}(x_0)$ and we only care about $x \in \bar{B}_1(x_0)$. 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 15.8 (Elliptic regularity for Δ)**.** *Suppose $f \in C^\infty(U)$ and $u \in L^1_{\mathrm{loc}}(U)$ is such that for each $\phi \in C_0^\infty(U)$ we have*

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

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

Exercise 15.9. Prove Theorem 15.8!

Remark 15.10. The above proof can be generalised in various ways, see Daniel Stroock's "Weyl's Lemma, one of many".

15.2 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 15.7, but this sort of argument really should not scare you.)

Alternative proof of Theorem 15.3. 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 $B_{4r}(x_0) \subset U$, we need to show that u is smooth on $B_r(x_0)$. We introduce a family of *mollifiers* η_ε: fix $\eta \in C_0^\infty([0,1);[0,\infty))$ with $\int_{[0,1)} s^{n-1}\eta(s)\,\mathrm{d}s = 1$ 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 $u_\varepsilon \colon B_{2r}(x_0) \to \mathbf{R}$ defined by*

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

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

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

Moreover,

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

Exercise 15.11. Prove this!

Step 2. *For each $x \in B_r(x_0)$ and $0 < s < r$,*

$$u(x) = ⨍_{\partial B_s(x)} u(y)\,\mathrm{d}y.$$

(A simplified version of) the argument from Exercise 15.2 shows that

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

Hence, u_ε satisfies

$$u_\varepsilon(x) = ⨍_{B_s(x)} u_\varepsilon(y)\,\mathrm{d}y$$

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

$$u(x) = ⨍_{B_r(x)} u(y)\,\mathrm{d}y.$$

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

$$⨍_{\partial B_s(x)} u(y)\,\mathrm{d}y = ⨍_{B_s(x)} u(y)\,\mathrm{d}y.$$

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

$$\begin{aligned} u_\varepsilon(x) &= \int_{B_\varepsilon(x)} \eta_\varepsilon(x-y)u(y)\,\mathrm{d}y \\ &= \int_0^\varepsilon \frac{s^{n-1}}{\varepsilon^n}\eta(s/\varepsilon) \fint_{\partial B_s(x)} u(y)\,\mathrm{d}y\mathrm{d}s \\ &= u(x)\cdot \int_0^\varepsilon (s/\varepsilon)^{n-1}\eta(s/\varepsilon)\,\mathrm{d}(s/\varepsilon) \\ &= u(x). \end{aligned}$$

□

Theorem 15.12. *If u satisfies the mean-value property, then it is smooth and harmonic.*

Theorem 15.13. *If $u\colon U \to \mathbf{R}$ is harmonic, then it is analytic.*

Proof. See [**Evans2010**]*Section 2.2 Theorem 10. The idea is to show that the Taylor series converges by estimating the remained using Theorem 14.6. (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.) □

Corollary 15.14 (Unique continuation)**.** *Suppose $u\colon U \to \mathbf{R}$ is a harmonic function. If u vanishes on a non-empty open subset of U, then u vanishes identically.*

Exercise 15.15. Consider the annulus

$$A := \{(x,y) \in \mathbf{R}^2 \cong \mathbf{C} : \sqrt{|x|^2+|y|^2} \in [1,2]\} = \bar{B}_2(0) \setminus B_1(0).$$

The function $f\colon A \to \mathbf{R}$ defined by

$$f(x,y) = \frac{x}{x^2+y^2} = \mathrm{Re}(1/z).$$

is harmonic.

Is there a harmonic function $\bar{f}\colon \bar{B}_2(0) \to \mathbf{R}$ such that

$$\bar{f}|_A = f?$$

15.3 Unique continuation and the frequency function

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

Definition 15.16. Let $x \in U$ and $R := d(x, \partial U)$. The *frequency function* $\mathrm{N}_x\colon I \to [0,\infty)$ is defined as

$$\mathrm{N}_x(r) := \frac{r\int_{B_r(x)} |\nabla u|^2}{\int_{\partial B_r(x)} |u|^2} = \frac{rH_x(r)}{h_x(r)}$$

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 15.17. *If u is harmonic, then $\mathrm{N}' \geq 0$ on I.*

Proposition 15.18. *If u is harmonic, then we have*

$$H(r) = \int_{\partial B_r(x)} (\partial_\nu u) u.$$

Proof. This follows from Theorem A.4. □

Proposition 15.19. *If u is harmonic, then*

$$H'(r) = \frac{n-2}{r} H(r) + 2 \int_{\partial B_r(x)} |\partial_r u|^2 \quad \textit{and}$$
$$h'(r) = \frac{n-1}{r} h(r) + 2 \int_{\partial B_r(x)} (\partial_\nu u) u.$$

Proof. We compute

$$\begin{aligned} h'(r) &= \frac{n-1}{r} \int_{\partial B_r(x)} |u|^2 + \int_{\partial B_r(x)} \partial_r |u|^2 \\ &= \frac{n-1}{r} h(r) + 2 \int_{\partial B_r(x)} (\partial_\nu u) u. \end{aligned}$$

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,

$$\begin{aligned} &\int_{\partial B_r(x)} \langle x_i |\nabla u|^2, x_i/r \rangle \\ &= \int_{B_r(x)} \partial_i (x_i |\nabla u|^2) \\ &= \int_{B_r(x)} |\nabla u|^2 + 2 \sum_{j=1}^n x_i (\partial_j u)(\partial_{ij} u) \\ &= \int_{B_r(x)} |\nabla u|^2 - 2 \sum_{j=1}^n \partial_j (x_i \partial_j u)(\partial_i u) + 2 \int_{\partial B_r(x)} \sum_{j=1}^n x_i (\partial_j u)(\partial_i u) \nu_j \\ &= \int_{B_r(x)} |\nabla u|^2 - 2 |\partial_i u|^2 + 2r \int_{\partial B_r(x)} \left(\frac{x_i}{r} \partial_i u \right) (\partial_r u). \end{aligned}$$

Summing this up yields

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

Proof of Proposition 15.17. Using Proposition 15.19, we compute

$$\begin{aligned}
\mathrm{N}'(r) &= \frac{H(r)}{h(r)} + \frac{rH'(r)}{h(r)} - \frac{rH(r)h'(r)}{h(r)^2} \\
&= \frac{H(r)}{h(r)} + \frac{(n-2)H(r)}{h(r)} + \frac{2r\int_{\partial B_r(x)} |\partial_r u|^2}{h(r)} \\
&\quad - \frac{rH(r)}{h(r)^2}\left(\frac{n-1}{r}h(r) + 2\int_{\partial B_r(x)} (\partial_\nu u)u\right) \\
&= \frac{2r}{h(r)^2}\left(\int_{\partial B_r(x)} |u|^2 \cdot \int_{\partial B_r(x)} |\partial_r u|^2 - \left|\int_{\partial B_r(x)} (\partial_\nu u)u\right|^2\right) \\
&\geq 0
\end{aligned}$$

by Cauchy–Schwarz. $\square$

Proposition 15.20. *If u is harmonic and non-vanishing on $B_R(x)$, then $I = (0,R)$ and for $0 < s < r < R$*

$$h(r) = \left(\frac{r}{s}\right)^{n-1} \exp\left(2\int_s^r \mathrm{N}(t)/t\,\mathrm{d}t\right) h(s);$$

in particular,

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

Proof. On I, the formula for h' from Proposition 15.19 can be written in the form.

$$h'(r) = (n-1+2\mathrm{N}(r))h(r)/r;$$

hence,

$$(\log h(r))' = (n-1+2\mathrm{N}(r))/r.$$

This integrates to the stated identity if $[s,r] \in I$. The second assertion follows from Proposition 15.17.

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

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

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 15.14. 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. $\square$

16 Green's functions

Combining the Dirichlet principle and Weyl's Lemma in the form of Theorem 15.8 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 $u \in W_0^{1,2}(U)$. 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 $\mathbf{R}^n$. We want G to be such that for suitable $g\colon\ \mathbf{R}^n \to \mathbf{R}$ the function defined by

$$u = G * g,$$

i.e.,

$$u(x) = \int_{\mathbf{R}^n} G(x-y)g(y)\,\mathrm{d}y$$

satisfies

$$\Delta u = g.$$

In the language of distributions this means that

$$\Delta G = \delta.$$

Definition 16.1. The *Green's function* on $\mathbf{R}^n$ is the function $G\colon\ \mathbf{R}^n \setminus \{0\} \to \mathbf{R}$ defined by

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

16.1 Derivations of the Green's function

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

16.1.1 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

$$\int_0^\infty \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \, \mathrm{d}t = G(x).$$

To see this, change variables $s := \frac{4t}{|x|^2}$ to get

$$\begin{aligned}\int_0^\infty \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} \, \mathrm{d}t &= \frac{1}{4\pi^{n/2}} \int_0^\infty \frac{e^{-\frac{1}{s}}}{s^{n/2}} \, \mathrm{d}s \cdot |x|^{2-n} \\ &= \frac{\Gamma\left(\frac{n}{2} - 1\right)}{4\pi^{n/2}} \cdot |x|^{2-n}.\end{aligned}$$

But

$$\mathrm{vol}(B_1^n) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} = \frac{4\pi^{n/2}}{n(n-2)\Gamma\left(\frac{n}{2} - 1\right)}.$$

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

Theorem 16.2. *If $g \in C_0^\infty(\mathbf{R}^n)$, then $u := G * g \in C^\infty(\mathbf{R}^n)$ and*

$$\Delta u = g.$$

16.1.2 Derivation using the Fourier transform

Using Proposition 9.5, taking the Fourier transform of $\Delta G = \delta$, gives

$$4\pi^2 |y|^2 \hat{G}(y) = 1.$$

Thus

$$G = \mathscr{F}^{-1}\left(\frac{1}{4\pi^2 |\cdot|^2}\right) = \mathscr{F}\left(\frac{1}{4\pi^2 |\cdot|^2}\right)$$

Actually computing the right-hand side is not that trivial, because $|\cdot|^{-2} \notin L^1(\mathbf{R}^n)$ and we need to view it as a tempered distribution: For $n \geq 3$, the integrability fails at infinity, i.e., $|\cdot|^{-2} \notin L^1(\mathbf{R}^n \setminus B_1)$; but we still have $|\cdot|^{-2} \in \mathcal{S}^*(\mathbf{R}^n)$. 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 $A \in \mathrm{SO}(n)$, then

$$\mathscr{F}(f \circ A) = \mathscr{F}(f) \circ A.$$

Therefore, because $|\cdot|^{-2}$ is $\mathrm{SO}(n)$–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.

16.2 Solving the Poisson equation with C^1 inhomogeneity

Theorem 16.3. *If $g \in C_0^1(\mathbf{R}^n)$, then $u := G * g \in C^2(\mathbf{R}^n)$ and*

$$\Delta u = g.$$

Remark 16.4. 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 $g \in C_0^0(\mathbf{R}^n)$ only; however, one can get away with $g \in C_0^{0,\alpha}(\mathbf{R}^n)$ for some $\alpha > 0$, that is, $g \in C_0^0(\mathbf{R}^n)$ 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 16.3.

Exercise 16.5. Prove that $\Delta G = 0$ on $\mathbf{R}^n \setminus \{0\}$. (*Hint:* It helps to know the formula for Δ 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

$$\Delta u(x) = -\sum_{i=1}^n (\partial_i G * \partial_i g)(x) = -\sum_{i=1}^n \int_{\mathbf{R}^n} \partial_i G(x-y) \partial_i g(y)\, \mathrm{d}y.$$

To see that this vanishes, we proceed as follows. For any $\varepsilon > 0$, note that

$$\begin{aligned}\left|\sum_{i=1}^n \int_{B_\varepsilon(x)} \partial_i G(x-y) \partial_i g(y)\, \mathrm{d}y\right| &\leq \|\nabla g\|_{L^\infty} \int_{B_\varepsilon(0)} |\nabla G| \\ &= \varepsilon \|\nabla g\|_{L^\infty} \int_{B_1(0)} |\nabla G|,\end{aligned}$$

by the homogeneity of degree $(1-n)$ of $|\nabla G|$. This goes to zero as $\varepsilon \downarrow 0$. Moreover, by Theorem A.4

$$\begin{aligned}\int_{\mathbf{R}^n \setminus B_\varepsilon(x)} \langle \nabla_y G(x-y), \nabla_y g(y) \rangle\, \mathrm{d}y = &\int_{\mathbf{R}^n \setminus B_\varepsilon(x)} (\Delta_y G(x-y)) g(y)\, \mathrm{d}y \\ &+ \int_{B_\varepsilon(x)} \partial_\nu G(x-y) g(y)\, \mathrm{d}y.\end{aligned}$$

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

$$\begin{aligned}\int_{\partial B_\varepsilon(x)} \partial_\nu G(x-y) g(y)\, \mathrm{d}y &= \int_{\partial B_\varepsilon(0)} -\partial_r G(y) g(x-y)\, \mathrm{d}y \\ &= \int_{\partial B_\varepsilon(0)} \frac{1}{n \mathrm{vol}(B_1^n)} |y|^{1-n} g(x-y)\, \mathrm{d}y \\ &= \fint_{\partial B_\varepsilon(0)} g(x-y)\, \mathrm{d}y \to g(x)\end{aligned}$$

as $\varepsilon \to 0$. Here we used that $n\mathrm{vol}(B_1^n) = \mathrm{vol}(\partial B_1^n)$. This completes the proof. □

16.3 Representation formula for solutions of the Dirichlet problem

Suppose now that $U \subset \mathbf{R}^n$ is a bounded open subset with smooth boundary. Suppose for each $x \in U$, ϕ_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* ϕ_x is difficult in general; however, for simple regions U formulas for ϕ_x are known. Set

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

Proposition 16.6 (Representation formula)**.** *Given any $u \in C^2(\bar{U})$, for any $x \in U$ we have*

$$u(x) = -\int_{\partial U} u(y)\partial_\nu G_U(x, y)\,\mathrm{d}y + \int_U \Delta u(y)G_U(y, x).$$

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 ∂U (this is why the correction term ϕ_x is needed!), and

$$\int_{\partial B_\varepsilon(x)} G_U(x, y)\partial_\nu u(y)\,\mathrm{d}y \to 0$$

as $\varepsilon \to 0$. The argument from Theorem 16.3 shows that

$$u(x) = \lim_{\varepsilon\to 0}\int_{\partial B_\varepsilon} \partial_\nu G_U(x, y)u(y).$$

Here ν is the unit normal pointing out of U_ε. □

Remark 16.7. Since ϕ_x is often hard to find, it is useful to remark that we also have

$$u(x) = \int_{\partial U} G(x - y)\partial_\nu u(y) - u(y)\partial_\nu G(x - y)\,\mathrm{d}y + \int_U \Delta u(y)G(y, x).$$

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 16.6 allows us to recover u from $g := \Delta u$ and $f := u|_{\partial U}$. However, we do not make the assertion that Proposition 16.6 for prescribed values of g and f does the define a solution.

16.4 The Dirichlet problem on a ball

Exercise 16.8. The formula
$$\phi_x(y) := G(|x|y - x/|x|)$$
defines a correction term for $B := B_1(0) \subset \mathbf{R}^n$.

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

Remark 16.9. 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 16.10. We define the *Poisson kernel* (for $r > 0$) to be
$$P_r(x,y) := \frac{r^2 - |x|^2}{r\mathrm{vol}(\partial B_1)||x-y|^n}.$$

Exercise 16.11. 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 16.12. Prove that there is a constant $c > 0$ such that for all $u \in C^\infty(\bar{B}_1(0))$ we have
$$\|u\|_{L^\infty(B_{1/2}(0))} \le c\left(\|\Delta u\|_{L^\infty(B_1(0))} + \|u\|_{L^1(\partial B_1(0))}\right)$$
(*Hint:* Use Proposition 16.6.)

Theorem 16.13. *Given $f \in C^0(\partial B_r)$, set*
$$u(x) := \begin{cases} \int_{\partial B_r} g(y)P_r(x,y)\,\mathrm{d}y & x \in B \\ g(x) & x \in \partial B. \end{cases}$$
Then $u \in C^\infty(B_r) \cap C^0(\bar{B}_r)$ 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
$$\Delta_x \partial_{y_j} G_B(x,y) = \Delta_x \partial_{x_j} G_B(x,y) = \partial_{x_j}\Delta_x G_B(x,y) = 0.$$
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
$$g(x) = \lim_{z\to x}\int_{\partial B} g(y)P_1(z,y)\,\mathrm{d}y.$$

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

17 Perron's Method

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

$$\begin{aligned} \Delta u &= 0 \quad \text{in } U \\ u|_{\partial U} &= f. \end{aligned}$$

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 17.1. A upper semi-continuous function $v\colon U \to \mathbf{R}$ is called *subharmonic* if one of the following equivalent conditions are satisfied:

- If $V \subset U$ is open and $u\colon V \to \mathbf{R}$ is harmonic and $v \leq u$ on ∂V, then $v \leq u$ on V.
- For all $\bar{B}_r(x) \subset U$, we have
$$v(x) \leq \fint_{\partial B_x} v.$$
- For all $\bar{B}_r(x) \subset U$, we have
$$v(x) \leq \fint_{B_x} v.$$

Exercise 17.2. 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 17.3. Show that if $\{v_1, \dots v_k\}$ are subharmonic, then so is $\max\{v_1, \dots v_k\}$.

17.1 Harmonic replacement

Definition 17.4. Suppose $v \in C^0(\bar{U})$ is subharmonic and $\bar{B}_r(x) \subset U$. Then the *harmonic replacement* $\tilde{v}\colon \bar{U} \to \mathbf{R}$ is defined by

$$\tilde{v}(x) := \begin{cases} v(x) & x \notin B_r(x) \\ \int_{\partial B_r(x)} P_r(x, y) v(y)\, \mathrm{d}y & x \in B_r(x). \end{cases}$$

Here P_r is the Poisson kernel.

Proposition 17.5. *If $v \in C^0(\bar{U})$ 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 $u\colon V \to \mathbf{R}$ 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. □

17.2 The maximal subharmonic function

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

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

Proposition 17.6. *The function $u\colon U \to \mathbf{R}$ defined by*

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

is harmonic.

Remark 17.7. Note that for any $v \in S_f$, $\sup_U v \leq \sup_{\partial U} f$. Thus u is well-defined.

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

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

Let $(v_i) \in S_f^{\mathbf{N}}$ such that $u(x) = \lim_{i\to\infty} v_i(x)$. By Exercise 17.3 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 14.10, 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 $(w_i) \in S_f^{\mathbf{N}}$ 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

$$\tilde{v}(x) = \fint_{B_r(x)} \tilde{v} < \fint_{B_r(x)} \tilde{w} = \tilde{w}(x).$$

□

17.3 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 ∂U potentially being very badly behaved.

Definition 17.8. Let $x \in \partial U$. A function $\beta \in C^0(\bar{U})$ is called a *barrier at x* if $-\beta|_U$ 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 17.9. If β 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 17.10. $B := B_1(0) \subset \mathbf{R}^n$ is regular. To see this we construct a barrier at $x_0 := (1, 0, \ldots, 0)$. Let G denote the Green's function on $\mathbf{R}^n$ 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 $\bar{B}$ closest to z_0 (that is: for all $x \in \bar{B}$ we have $|x - z_0| \geq |x_0 - z_0|$ with equality if and only if $x = x_0$), $\beta(y) > 0$ for all $y \in \bar{B} \setminus \{0\}$.

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 $z_0 \notin \bar{U}$.

Exercise 17.11. Show that $0 \in \partial\{x \in \mathbf{R}^n : 0 < |x| < 1\}$ $(n \geq 2)$ is not regular.

Exercise 17.12. Show that $Q = (0, 1)^n$ is regular.

Proposition 17.13. *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 β a barrier at x. For every $\varepsilon > 0$, and $c > 0$ sufficiently large

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

defines a function $v \in S_f$ and similarly

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

defines a function in S_f^*.

Thus

$$v_\varepsilon \leq u \leq u^* \leq v_\varepsilon^*.$$

Since $\lim_{U \ni z \to x} v_\varepsilon(z) = f(x) - \varepsilon$ and $\lim_{U \ni z \to x} v_\varepsilon^*(z) = f(x) + \varepsilon$ and ε is arbitrary, the assertion follows. □

Theorem 17.14. *There following are equivalent:*

1. *∂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 β to be the solution of the Dirichlet problem with boundary value $f(y) := -|x - y|$. □

18 Minimal hypersurfaces and the Bernstein problem

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

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

The graph of u is the set

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

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

$$\mathscr{A}(u) := \mathrm{vol}(\Gamma(u)) = \int_U \sqrt{1 + |\nabla u|^2}.$$

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

The question of whether or not a solution minimiser of the functional $\mathscr{A}$ 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 $\mathscr{A}$ is given by

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

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

$$H := -\,\mathrm{div}\left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}}\right) = 0. \tag{18.1}$$

Definition 18.2. An *graphical minimal hypersurface* in $\mathbf{R}^{n+1}$ is a graph

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

of a C^2 map $u\colon U \to \mathbf{R}$ defined on a domain $U \subset \mathbf{R}^n$ which satisfies (18.1).

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

Remark 18.3. H is called the mean curvature of $\Gamma(u)$.

Remark 18.4. Note that a minimal hypersurface need not minimise $\mathscr{A}$; in fact, if $U = \mathbf{R}^n$, then $\mathscr{A}$ is usually not defined.

Remark 18.5. More generally, a minimal hypersurface in $\mathbf{R}^{n+1}$ is subset M that locally in suitable coordinates can be written as the graph of a function satisfying (18.1). (This allows, in particular, for the hyperplane over which M is a graph to rotate.)

Example 18.6. If $u\colon U \to \mathbf{R}$ is affine, i.e, it is of the form

$$u(x) = u_0 + \langle x, v\rangle$$

for some $u_0 \in \mathbf{R}$ and $v \in \mathbf{R}^n$, then $\nabla u = 0$ and (18.1) holds trivially.

We compute

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

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 18.7. The graph of $u\colon (-r, r) \to \mathbf{R}$ defined by

$$
u(x) = \sqrt{r^2 - x^2}
$$

is the upper half circle $\mathbf{R}^2$ of radius r.

Let us compute H. Since

$$
u' = -\frac{x}{\sqrt{r^2-x^2}},
$$

we have

$$
\begin{aligned}
\left(1+(u')^2\right)^{3/2} &= \left(1+\frac{x^2}{r^2-x^2}\right)^{3/2} = \left(\frac{r^2}{r^2-x^2}\right)^{3/2} \quad \text{and}\\
u'' &= -\frac{1}{(r^2-x^2)^{1/2}} - \frac{x^2}{(r^2-x^2)^{3/2}} = -\frac{r^2}{(r^2-x^2)^{3/2}}.
\end{aligned}
$$

Thus

$$
H = 1/r.
$$

Example 18.8. The catenoids are a family of a minimal hypersurfaces in $\mathbf{R}^3$, which are not graphical. Topologically each catenoid is a cylinder and it can be parametrised for some $r > 0$ by the map $\mathbf{R}\times S^1 \to \mathbf{R}^3$

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

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

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

Exercise 18.9. Prove that $\Gamma(u_\pm)$ are minimal graphs.

Remark 18.10. Although, the study of hypersurfaces in $\mathbf{R}^3$ 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 18.11 (Bernstein's Theorem)**.** *A entire graphical minimal hypersurface in* $\mathbf{R}^3$ *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 18.12. The Bernstein problem is the question whether this is true in dimension $n > 3$ as well. Simons [6] proved that this holds for $n \leq 7$. The problem was settled by Bombieri–De Giorgi–Gusti [2] who proved that for $n \geq 8$ it does not hold.

Proposition 18.13 (Baby Bernstein)**.** *A entire graphical minimal hypersurface in* $\mathbf{R}^2$ *must be affine; hence, a straight line.*

Proof. Equation (18.1) for a function $u\colon \mathbf{R} \to \mathbf{R}$ can be written as

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

Thus, $u'' = 0$, so $u(x) = u_0 + cx$. □

Proof of Theorem 18.11. What we need to show is that any function $u\colon \mathbf{R}^2 \to \mathbf{R}$ 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 ψ_i are obviously bounded.

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

Theorem 18.14. *Let* $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} : \mathbf{R}^2 \to \mathbf{R}^{2\times 2}$ *be such that* $A(x, y)$ *is positive definite for each* $(x, y) \in \mathbf{R}^2$. *If* $u \in C^2(\mathbf{R})$ *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 18.15 (E. Hopf [3], Mickle [5])**.** *If*

$$-\det \operatorname{Hess}(u) = (\partial_x\partial_y u)^2 - (\partial_x^2 u)(\partial_y^2 u) \geq 0 \quad \textit{but} \quad \det \operatorname{Hess}(u) \neq 0,$$

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 18.14 as follows.

Under the hypothesis of Theorem 18.14 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. □

19 L^2 regularity theory for second order elliptic operators in divergence form

In the next couple of lectures I will teach you about one way to deal with linear second order uniformly elliptic operators in divergence form.

Definition 19.1. A second order differential operator L is said to be in *divergence form* if it can be written as

$$Lu = -\sum_{i,j=1}^{n} \partial_i(a_{ij}\partial_j u) + \sum_{k=1}^{n} b_k \partial_k u + cu \tag{19.2}$$

where $A = (a_{ij})\colon U \to \mathbf{R}^{n\times n}$, $b = (b_k)\colon U \to \mathbf{R}^n$ and $c\colon U \to \mathbf{R}$, and U is an open subset of $\mathbf{R}^n$.

We say that L is *uniformly elliptic* if there are constant $\Lambda \geq \lambda > 0$ such that for all $x \in U$, A is symmetric and $\operatorname{spec}(A) \in [\lambda, \Lambda]$.

Hypothesis 19.3. *For the rest of these lecture notes we suppose L is uniformly elliptic.*

Of course, if a_{ij} is differentiable, then

$$Lu = -\sum_{i,j=1}^{n} a_{ij}\partial_i\partial_j u + \sum_{k=1}^{n}\left(b_k - \sum_{\ell=1}^{n}\partial_\ell a_{\ell k}\right)\partial_k u + cu.$$

So one might wonder why bother writing L in this form. The reason is that this is well adopted to the notion of weak solutions.

Definition 19.4. Assume that $A \in L^\infty(U, \mathbf{R}^{n\times n})$, $b \in L^\infty(U, \mathbf{R}^n)$ and $c \in L^\infty(U)$. Let $f \in L^2(U)$ and $g = (g_i) \in L^2(U, \mathbf{R}^n)$. We say that $u \in W^{1,2}(U)$ is a *weak solution* of

$$Lu = f - \sum_{i=1}^{n} \partial_i g_i \tag{19.5}$$

if for each $\phi \in W_0^{1,2}(U)$

$$\int_U \sum_{i,j=1}^{n} a_{ij}(\partial_i u)(\partial_j \phi) + \sum_{k=1}^{n}(b_k\partial_k u)\cdot\phi + cu\cdot\phi = \int_U f\phi + \sum_{i=1}^{n} g_i\partial_i\phi.$$

Observe how the above makes sense even if a_{ij} is not differentiable.

You might ask: why do we introduce those strange g_i terms? A good answer is that in the context of Definition 19.4 L is a linear operator $W^{1,2} \to W^{-1,2} := (W_0^{1,2})^*$ and the general element of $(W_0^{1,2})^*$ is of the same form as the right-hand side of (19.5).

19.1 Existence of weak solutions

Proposition 19.6. *Suppose L is a second order elliptic operator in divergence form with $A \in L^\infty(U, \mathbf{R}^{n\times n})$, $b = 0$, and $c \in L^\infty(U, [0,\infty))$. For every $f \in L^2(U)$ and $g = (g_i) \in L^2(U, \mathbf{R}^n)$, there exists a unique weak solution $u \in W_0^{1,2}(U)$ of*

$$Lu = f - \sum_{i=1}^n \partial_i g_i.$$

Remark 19.7. It is crucial that $u \in W_0^{1,2}(U)$ (and not just $W^{1,2}(U)$, for otherwise uniqueness fails). The observant reader will also note that, in fact, one does not need $c \geq 0$. It suffices to have $c \geq -\varepsilon$ for some small $\varepsilon > 0$ depending on U and λ.

Proof. This can be proved along the lines of Section 12.6, but it is useful to recast this argument in this slightly more abstract setting.

We define an inner product on $W_0^{1,2}(U)$ by

$$\langle u, v\rangle_L := \int_U \sum_{i=1}^n a_{ij}(\partial_i u)(\partial_j v) + cuv.$$

By the Dirichlet–Poincaré inequality for some $\delta > 0$

$$\begin{aligned}\delta\|u\|_{W^{1,2}}^2 &\leq \lambda\|\nabla u\|_{L^2}^2 \\ &\leq \|u\|_L^2 \\ &\leq \Lambda\|\nabla u\|_{L^2}^2 + \sup_U c \cdot \|u\|_{L^2}^2 \leq \delta^{-1}\|\nabla u\|_{W^{1,2}}^2.\end{aligned}$$

This means that the normed vector space $(W_0^{1,2}(U), \|\cdot\|_L)$ is equivalent to $(W_0^{1,2}(U), \|\cdot\|_{W^{1,2}})$. Therefore $(W_0^{1,2}(U), \langle\cdot,\cdot\rangle_L)$ is a Hilbert space.

Now observe that $\alpha\colon W_0^{1,2}(U) \to \mathbf{R}$ defined by

$$\alpha(v) = \int_U fv + \sum_{i=1}^n g_i \partial_i v$$

is a bounded linear functional with respect to $\|\cdot\|_{W^{1,2}}$ and, hence, with respect to $\|\cdot\|_L$. Now by Riesz' representation theorem, there exists a unique $u \in W_0^{1,2}(U)$ such that

$$\langle u, \cdot\rangle_L = \alpha(\cdot).$$

This is exactly the assertion that there is a unique weak solution $u \in W_0^{1,2}(U)$. □

If $b \neq 0$ or c can become negative, then one cannot hope for a result as strong as Proposition 19.6. Neither existence nor uniqueness might hold. There is however, the following

Exercise 19.8 (Lax–Milgram theorem)**.** Suppose H is a Hilbert space and $B\colon\ H \times H \to \mathbf{R}$ is a bounded bilinear form (but possibly non-symmetric) satisfying

$$\lambda\|x\|\|y\| \le |B(x,y)| \le \Lambda\|x\|\|y\|.$$

Show that for any $x \in H$, there exists a unique $\tilde{x} \in H$ such that

$$B(\tilde{x},\cdot) = \langle x,\cdot\rangle.$$

Now suppose u is a weak solution of (19.5). Here are some natural question to ask: How regular is u? Is it a strong solution, i.e., does it literally satisfy (19.5)? Does it attain its boundary values?

The answer to this is roughly: if A, b and c are sufficiently smooth, then u has two more derivatives than f and one more than g; and if ∂U is sufficiently regular then u attains it boundary values.

All of these above questions are qualitative. Closely related is the quantitative question: How large is u in terms if f and g (and $u|_{\partial U}$)? I turns out that one mostly answers these sorts of questions at the same time.

For answering this sort of question one has to make precise what exactly one means be regularity. I turns out that the notions of just C^k are not suitable. The problem with that is that continuity is purely qualitative property. However, in these questions one typically needs something more quantitative. A common choice is to work with in non-linear problems are Hölder spaces. We will come to those soon. For now we will stick with $W^{k,2}$ spaces. Recall that

$$\|u\|_{W^{k,2}} := \|u\|_{L^2} + \|\nabla u\|_{L^2} + \cdots + \|\nabla^k u\|_{L^2}.$$

19.2 L^2 regularity for Δ

From Theorem 12.16, we know that if $u \in C_0^\infty(\mathbf{R}^n)$, then

$$\|u\|_{W^{k+2,2}} \le c\|\Delta u\|_{W^{k,2}} + \|u\|_{W^{1,2}}.$$

In fact this continues to hold true for all $u \in W^{k+2,2}$. This sort of statement is called an *a priori estimate*: provided u is in $W^{k+2,2}$, we can predict an upper bound for $\|u\|_{W^{k+2,2}}$.

For general PDE one has to strictly make a distinction between a priori estimates and questions of regularity. But for constant coefficient operators a priori estimates typically imply regularity. Here is what this means. Suppose $u \in W^{1,2}$ and we know that $\Delta u = f \in L^2$, in the weak sense. Can we conclude that $u \in W^{2,2}$?

Note that constant coefficient operators commute with convolution: That is if $\rho \in C_0^\infty(\mathbf{R}^n)$, then $\Delta u = f$ (in the weak sense) implies

$$\Delta(u * \rho) = \rho * (\Delta u) = \rho * f.$$

Here $\Delta(u * \rho) = \rho * f$ even holds in the strong sense as an identity in $C_0^\infty(\mathbf{R}^n)$ If η_ε is a suitable family of mollifiers and $u_\varepsilon := \eta_\varepsilon * u$ and $f_\varepsilon := \eta_\varepsilon * f$, then

$$u_\varepsilon \to u \in W^{1,2} \quad \text{and} \quad f_\varepsilon \to f \in L^2.$$

But by the above estimate we also know that

$$\|u_\varepsilon - u_\delta\|_{W^{2,2}} \leq \|f_\varepsilon - f_\delta\|_{L^2} + \|u_\varepsilon - u_\delta\|_{W^{1,2}}.$$

Thus u_ε is Cauchy in $W^{2,2}$ has a limit in $W^{2,2}$ must agree with u. It follows that $u \in W^{2,2}$.

A similar argument can be used to show the following

Proposition 19.9. *If $u \in W^{k+1,2}$ and $\Delta u \in W^{k,2}$, then*

$$u \in W^{k+2,2}$$

and

$$\|u\|_{W^{k+2,2}} \leq \|\Delta u\|_{W^{k,2}} + \|u\|_{W^{1,2}}.$$

20 L^2 regularity theory for second order elliptic operators in divergence form (continued)

20.1 L^2 interior regularity for variable coefficients operators

Theorem 20.1. *Let $k \in \mathbf{N}_0$. Suppose $a_{ij} \in W^{k+1,\infty}(U)$ and $b_k, c \in W^{k,\infty}(U)$, and $f \in W^{k,2}(U)$. If $u \in W^{1,2}(U)$ is a weak solution of*

$$Lu = f,$$

then

$$u \in W^{k+2,2}_{\mathrm{loc}}(U)$$

and for each $V \subset\subset U$ there is a constant $c > 0$ (independent of u) such that

$$\|u\|_{W^{k+2,2}(V)} \leq c \left(\|f\|_{W^{k,2}(U)} + \|u\|_{L^2(U)} \right).$$

This cannot be proved in quite the same way as the analogous result for Δ. The standard method to prove it, using difference quotients, goes back to Nirenberg.

For $h \in \mathbf{R} \setminus \{0\}$ and $k \in \{1, \ldots, n\}$ define the difference quotient

$$\partial_k^h u := \frac{u(\cdot + he_k) - u(\cdot)}{h}.$$

These different quotients have an important integration by parts property

$$\int (\partial_k^h u) v = -\int u(\partial_k^{-h} v).$$

There also is a chain rule

$$\partial_k^h (uv) = (\partial_k^h u)v + u^{h,k} \partial_k^h v$$

with $u^{h,k} := u(\cdot + he_k)$.

For us the significance of the difference quotients comes from the following.

Proposition 20.2. *Suppose U is a subset of $\mathbf{R}^n$ and V is an open subset of $\mathbf{R}^n$ such that for all sufficiently small $h \in \mathbf{R}$ we have $V + he_k \in U$. If $u \in L^2(U)$ and $\partial_k u \in L^2(U)$, then there is a constant $c > 0$ such that for all $h \neq 0$ sufficiently small*

$$\|\partial_k^h u\|_{L^2(V)} \leq c\|\partial_k u\|_{L^2(U)}.$$

Conversely, if $u \in L^2(U)$ and

$$c := \limsup_{0 \neq h \to 0} \|\partial_k^h u\|_{L^2(V)} < \infty,$$

then $\partial_k u \in L^2(V)$ and

$$\|\partial_k u\|_{L^2}(V) \leq c$$

Proof. Let me only say something about the second half. We need to explain why the distribution $\partial_k u$ is represented by an L^2 function. Given any $\phi \in C_0^\infty$ and h sufficiently small

$$(\partial_k u)[\phi] = -\int u\partial_k\phi = -\lim_{h\to 0}\int u\partial_k^{-h}\phi = \lim_{h\to 0}\int(\partial_k^h u)\phi.$$

Thus $\partial_k u$ is the weak limit of $(\partial_k^h u)$ (in the space of distributions). But $(\partial_k^h u)$ is uniformly bounded in L^2. Thus its weak limit is also in L^2. □

To not further clutter the exposition we prove the above theorem only in a special case.

Proof of Theorem 20.1 for $k = 0$ and assuming $b = c = 0$. Suppose $u \in W^{1,2}(U)$ is a weak solution of

$$-\sum_{i,j=1}^{n}\partial_i(a_{ij}\partial_j u) = f$$

with $a_{ij} \in C^1$ and $f \in L^2$. We need to prove that if $V \subset U$, then $u|_V \in W^{2,2}$ and satisfies the asserted estimate.

Choose a cut off function χ compactly supported in U and equal to one on V. For $i \in \{1, \dots, n\}$ and $h \in \mathbf{R} \setminus \{0\}$, define a test function by

$$v := -\partial_k^{-h}(\chi^2\partial_k^h u) \in W_0^{1,2}(U).$$

(This is supposed to be an approximation of $\partial_k^2 u$, but we have to use difference quotients and the cut off to ensure that $v \in W_0^{1,2}(U)$. Why we write χ^2 instead of just χ will be come apparent soon.)

Now using this test function in the definition of weak solution we get

$$\int_U \sum_{i,j=1}^{n} -a_{ij}(\partial_i u)\cdot\partial_j\partial_k^{-h}(\chi^2\partial_k^h u) = \int_U -f\partial_k^{-h}(\chi^2\partial_k^h u). \tag{20.3}$$

By discrete integration by parts the left-hand side can be rewritten as follows:

$$
\begin{aligned}
&\int_U \sum_{i,j=1}^{n} -a_{ij}(\partial_i u)\cdot \partial_j \partial_k^{-h}(\chi^2 \partial_k^h u)\\
&= \int_U \sum_{i,j=1}^{n} (\partial_k^h[a_{ij}(\partial_i u)])\cdot \partial_j(\chi^2\partial_k^h u)\\
&= \int_U \sum_{i,j=1}^{n} a_{ij}^{h,k}(\partial_k^h \partial_i u)\cdot \partial_j(\chi^2 \partial_k^h u) + (\partial_k^h a_{ij})(\partial_i u)\cdot \partial_j(\chi^2\partial_k^h u)\\
&= \int_U \sum_{i,j=1}^{n} \chi^2 a_{ij}^{h,k}(\partial_k^h\partial_i u)\cdot(\partial_j \partial_k^h u)\\
&\quad + \int_U \sum_{i,j=1}^{n} 2\chi\partial_j\chi\cdot a_{ij}^{h,k}(\partial_k^h\partial_i u)\cdot(\partial_k^h u)\\
&\qquad + \chi^2(\partial_k^h a_{ij})(\partial_i u)\cdot\partial_j(\partial_k^h u) + 2\chi\partial_j\chi\cdot(\partial_k^h a_{ij})(\partial_i u)\cdot \partial_k^h u\\
&=: \mathrm{I} + \mathrm{II}.
\end{aligned}
$$

We think of I as the good term and of II as the bad term or error term: By uniform ellipticity we have

$$\int_U \chi^2 |\partial_k^h \nabla u|^2 \leq \lambda^{-1}\mathrm{I}.$$

The left-hand side is exactly the term we want to control.

To bound the error term, observe that using $a_{ij}\in W^{1,\infty}$ for some constant $c>0$

$$
\begin{aligned}
|\mathrm{II}| &\leq c\int_U \chi|\partial_k^h\nabla u|(|\partial_k^h u| + |\nabla u|) + \chi|\nabla u||\partial_k^h u|\\
&\leq \int_U \varepsilon\chi^2|\partial_k^h\nabla u|^2 + c\varepsilon^{-1}(|\partial_k^h u|^2 + |\nabla u|^2),
\end{aligned}
$$

for all $\varepsilon>0$. (In the above it is important to still have a χ term. This is why we worked with χ^2).

Now it remains to bound the right hand side of (20.3):

$$
\begin{aligned}
F := \int_U -f\partial_k^{-h}(\chi^2\partial_k^h u) &\leq \int_U \varepsilon^{-1}|f|^2 + \varepsilon|\partial_k^{-h}(\chi^2\partial_k^h u)|^2\\
&\leq \int_U \varepsilon^{-1}|f|^2 + \varepsilon|\partial_k(\chi^2\partial_k^h u)|^2\\
&\leq \int_U \varepsilon^{-1}|f|^2 + 2\varepsilon|2\chi(\partial_k\chi)\partial_k^h u)|^2 + 2\varepsilon\chi^2|\partial_k^h\nabla u|^2\\
&\leq \int_U \varepsilon^{-1}|f|^2 + 2c\varepsilon|\nabla u|^2 + 2\varepsilon\chi^2|\partial_k^h\nabla u|^2
\end{aligned}
$$

Putting all of this together we get

$$\begin{aligned}\lambda \int_U \chi^2 |\partial_k^h \nabla u|^2 &\leq \mathrm{I} \\ &\leq F + |\mathrm{II}| \\ &\leq \int_U 3\varepsilon \chi^2 |\partial_k^h \nabla u|^2 + c\varepsilon^{-1})|\nabla u|^2 \varepsilon^{-1} |f|^2.\end{aligned}$$

If ε is chosen sufficiently small, we get

$$\begin{aligned}\int_v |\partial_k^h \nabla u|^2 &\leq \int_U \chi^2 |\partial_k^h \nabla u|^2 \\ &\leq c \int_U |f|^2 + |\nabla u|^2.\end{aligned}$$

Since the right hand side does not depend on h, it follows that $\nabla u \in W^{1,2}(V)$ and

$$\|u\|_{W^{2,2}(V)} \leq c \left(\|f\|_{L^2(U)} + \|u\|_{W^{1,2}(\mathrm{supp}(\chi))} \right).$$

This estimate is not quite what we claimed. What is missing is to replace $\|u\|_{W^{1,2}(U)}$ by $\|u\|_{L^2(U)}$. This can be achieved by using the test function $\tilde{\chi}^2 u$ where $\tilde{\chi}$ is compactly supported in U and equal to one on $\mathrm{supp}(\chi)$. □

If one is willing to put in a bit more work one can easily deal with the cases where $b, c \in L^\infty$.

The idea behind proof of Theorem 20.1 for $k > 0$ is called *bootstrapping*. Suppose $u \in W^{1,2}$ is a weak solution of $Lu = f \in W^{1,2}$. Then by the $k = 0$ case we know that $u \in W^{2,2}(V)$ for any $V \subset\subset U$ and a computation shows that $\tilde{u} := \partial_\ell u$ is a weak solution of

$$L\tilde{u} = \tilde{f}$$

with

$$\tilde{f} := \partial_\ell f + \sum_{i,j=1}^n \partial_i(\partial_\ell a_{ij} \cdot \partial_j u) - \sum_{k=1}^n \partial_\ell b_k \cdot \partial_k u - \partial_\ell c \cdot u.$$

Thus one derive the $k = 1$ from the $k = 1$ case, and so on.

20.2 L^2 boundary regularity

If $u \in W_0^{1,2}(U)$ is a weak solution of

$$Lu = f,$$

and ∂U is sufficiently smooth, one can expect regularity of u up to the boundary. More precisely, we have

Theorem 20.4. *Let $k \in \mathbf{N}_0$. Suppose that ∂U is C^{k+2}, $a_{ij} \in W^{k+1,\infty}(\bar{U})$, $b_k, c \in W^{k,\infty}(\bar{U})$. If $u \in W_0^{1,2}(U)$ is a weak solution of $Lu = f$ with $f \in W^{k,2}(U)$, then*

$$u \in W^{k+2,2}(U)$$

and for a constant $c > 0$ independent of u

$$\|u\|_{W^{k+2,2}(U)} \leq c\left(\|f\|_{W^{k,2}(U)} + \|u\|_{L^2(U)}\right).$$

Remark 20.5. It is crucial that $u \in W_0^{1,2}(U)$.

Proof sketch. The proof of this theorem is a bit involved. We will only consider the case $k = 0$ and $b = c = 0$. We can cover $\bar{U}$ by finitely many balls $B_r(x)$ and we can assume that these balls are either completely contained in U or $x \in \partial U$. For balls contained in U Theorem 20.1 applies. So we can focus on balls centered at the boundary. By applying a C^{k+2} diffeomorphism we reduce to the case $U = [0, 2) \times B_2^{n-1}(0)$ and $B_r(x) = B_1(0)$. Formally L transforms to a new differential operator and one can (and in fact has to) check that the transformed u still is a weak solution of a $Lu = f$ with the transformed f. You can check this as an exercise.

So now let us assume that $U = [0, 2) \times B_2^{n-1}(0)$ and $B_r(x) = B_1(0)$ (and we still have $b = c = 0$). Let χ be a cut off function compactly supported in U and equal to one on $B_1(0)$. (Note in particular that χ need not vanish for $x_1 = 0$.) For $k \in \{2, \ldots, n\}$ the argument from Theorem 20.1 goes through with

$$v := -\partial_k^{-h}(\chi^2 \partial_k^h u) \in W_0^{1,2}(U)$$

and thus we can show that $\partial_k \nabla u \in L^2(B_1(0))$ with the desired estimates.

The crucial question is: how do we show that $\partial_1 \partial_1 u \in L^2$ (and derive estimates)? Note that we can write $Lu = f$ as

$$-\int_U a_{11}(\partial_1 u)(\partial_1 \phi) = \int_U -\sum_{\substack{i,j=1,\ldots n\\(i,j)\neq(1,1)}} a_{ij}(\partial_i u)(\partial_n \phi) + f\phi.$$

This means that (as a distribution)

$$\partial_1(a_{11}(\partial_1 u)) = \sum_{\substack{i,j=1,\ldots n\\(i,j)\neq(1,1)}} \partial_n(a_{ij}(\partial_i u)) - f =: \rho$$

By the above the $\rho \in L^2$. By uniform ellipticity $a_{11} \geq \lambda > 0$. It follows that

$$\partial_1 \partial_1 u = a_{11}^{-1}(\partial_1(a_{11}(\partial_1 u)) - (\partial_1 a_{11})\partial_1 u \in L^2.$$

This completes our sketch proof. □

21 Hölder spaces

21.1 Definition and basic properties of Hölder spaces

Definition 21.1. For $\alpha > 0$, $k \in \mathbf{N}_0$ and $u \colon U \to \mathbf{R}$ define

$$[u]_{C^{0,\alpha}(\bar{U})} := \sup_{x \neq y \in \bar{U}} \frac{|u(x) - u(y)|}{|x - y|^\alpha}$$

and

$$\|u\|_{C^{k,\alpha}(\bar{U})} := \sum_{\ell=1}^{k} \|\nabla^\ell u\|_{L^\infty(\bar{U})} + [\nabla^\ell u]_{C^{0,\alpha}(\bar{U})}.$$

Set

$$C^{k,\alpha}(\bar{U}) := \{u \in C^k(\bar{U}) : \|u\|_{C^{k,\alpha}(\bar{U})} < \infty\}.$$

Exercise 21.2. $(C^{k,\alpha}(\bar{U}), \|\cdot\|_{C^{k,\alpha}(\bar{U})})$ is a Banach space.

The spaces $C^{k,\alpha}(\bar{U})$ are called Hölder spaces.

Exercise 21.3. If $u \in C^{0,\alpha}$ with $\alpha > 1$, then u is locally constant. (This is why one usually requires $\alpha \in (0, 1]$.)

Exercise 21.4. If $u \in C^{0,\alpha}(\bar{U})$ and $v \in C^0(\bar{U})$, then

$$\|u \cdot v\|_{C^{0,\alpha}(\bar{U})} \leq \|u\|_{C^{0,\alpha}(\bar{U})} \cdot \|v\|_{L^\infty(\bar{U})}.$$

Exercise 21.5. If $f \in C^\infty(\mathbf{R})$ and $u \in C^{0,\alpha}(\bar{U})$, then $f \circ u \in C^{0,\alpha}(\bar{U})$.

The above observations show that Hölder spaces work well with non-linearities.

Remark 21.6. $C^\infty(\bar{U})$ is not dense in $C^{0,\alpha}(\bar{U})$.

21.2 Integral characterisation of Hölder spaces

The main point for the approach to Schauder theory we are going to take is the fact that Hölder continuity can be characterized in terms of integral estimates.

Hypothesis 21.7. *Throughout we make the assumption that U is bounded and there is a constant $c > 0$ such that for all $0 < r < \min\{1, \operatorname{diam}(U)\}$ and all $x \in U$ we have*

$$\operatorname{vol}(U \cap B_r(x)) \geq c r^n.$$

This is sometimes summarized by saying "U is of type A*". This is quite a mild condition. It holds for example if ∂U is C^1.*

We make the following definition.

Definition 21.8. For $1 \leq p < \infty$ and $\lambda > 0$, the *Campanato space* $(\mathcal{L}^{p,\lambda}(U), \|\cdot\|_{\mathcal{L}^{p,\lambda}(U)})$ is the normed vector space defined by

$$\mathcal{L}^{p,\lambda}(U) := \left\{u \in L^p(U) : [u]_{\mathcal{L}^{p,\lambda}(U)} < \infty\right\}$$

and

$$\|u\|_{\mathcal{L}^{p,\lambda}(U)} := \|u\|_{L^p(U)} + [u]_{\mathcal{L}^{p,n}(U)}.$$

Here the *Campanato semi-norm* is defined by

$$[u]_{\mathcal{L}^{p,\lambda}(U)} := \sup_{x \in U, r>0} \left(r^{-\lambda} \int_{B_r(x)\cap U} |u - \bar{u}_{x,r}|^p\right)^{1/p}$$

with

$$\bar{u}_{x,r} := \fint_{B_r(x)\cap U} u.$$

One can show that these spaces are all Banach spaces and it is very interesting and important to study them in their own right. It is a simple but important observation that if $\lambda > n$, then

$$\begin{aligned}\fint_{B_r(x)\cap U} |u - u_{x,r}| &\leq \left(\fint_{B_r(x)\cap U} |u - u_{x,r}|^p\right)^{1/p} \\ &\leq cr^\alpha \left(r^\lambda \int_{B_r(x)\cap U} |u - u_{x,r}|^p\right)^{1/p} \\ &\leq cr^\alpha [u]_{\mathcal{L}^{p,\lambda}(U)}\end{aligned}$$

for $\alpha := \frac{\lambda - n}{p}$. This lies at the heart the following theorem.

Theorem 21.9. *If* $\alpha := \frac{\lambda-n}{p} > 0$*, then* $\mathcal{L}^{p,\lambda}(U) \cong C^{0,\alpha}(\bar{U})$*.*

We prove this in two separate propositions.

Proposition 21.10. *If* $\alpha := \frac{\lambda-n}{p} > 0$*, then* $\mathcal{L}^{p,\lambda}(U) \subset C^{0,\alpha}(\bar{U})$ *and there is a constant* $c > 0$ *such that for all* $u \in \mathcal{L}^{p,\lambda}(U)$

$$\|u\|_{C^{0,\alpha}(\bar{U})} \leq c\|u\|_{\mathcal{L}^{p,\lambda}(U)}.$$

Proof. Suppose $u \in \mathcal{L}^{p,\lambda}(U)$. Note that a priori u is only in L^p and thus defined only almost everywhere (or by a Cauchy sequence up to equivalence).

Step 1. *There is a continuous function* $\bar{u} \in C^0(\bar{U})$ *such that*

$$\bar{u}(x) := \lim_{i\to\infty} u_{x,r_i}$$

for $r_i := 2^{-i} \min\{1, \operatorname{diam}(U)\}$*, and* $\bar{u}$ *is a representative of* u *(i.e., it agrees with* u *almost everywhere). Moreover, for some constant* $c > 0$ *not depending on* u*, we have*

$$|\bar{u}(x) - u_{x,r}| \leq cr^\alpha [u]_{\mathcal{L}^{p,\lambda}(U)} \tag{21.11}$$

Set $\bar{u}_i(x) := u_{x,r_i}$. By basic properties of the integral, $\bar{u}_i \in C^0(\bar{U})$. Now, we compute

$$\begin{aligned}
|u_{x,r_{i+1}} - u_{x,r_i}| &\leq \fint_{B_{r_{i+1}}(x)\cap U} |u - u_{x,r_i}| \\
&\leq c^{-1} r_{i+1}^{-n} \int_{B_{r_i}(x)\cap U} |u - u_{x,r_i}| \\
&\leq 2\mathrm{vol}(B_1)c^{-1} \fint_{B_{r_i}(x)\cap U} |u - u_{x,r_i}| \\
&\leq 2\mathrm{vol}(B_1)c^{-1} \left(\fint_{B_{r_i}(x)\cap U} |u - u_{x,r_i}|^p \right)^{1/p} \\
&\leq 2\mathrm{vol}(B_1)c^{-(1+1/p)} \cdot r_i^{\frac{\lambda-n}{p}} \left(r_i^{-\lambda} \int_{B_{r_i}(x)\cap U} |u - u_{x,r_i}|^p \right)^{1/p} \\
&\leq 2\mathrm{vol}(B_1)c^{-(1+1/p)} \cdot r_i^{\alpha} [u]_{\mathcal{L}^{p,\lambda}(U)}.
\end{aligned}$$

The sequence r_i^{λ} is Cauchy; hence, so is u_{x,r_i}. This means that $\bar{u}$ is well-defined. It also follows that (21.11) holds. Furthermore the right-hand side above does not depend on x, so this in fact shows that $\bar{u}_i$ is a Cauchy sequence in $C^0(\bar{U})$. Thus $\bar{u} \in C^0(\bar{U})$.

The fact that $\bar{u}$ represents u follows from the Lebesgue differentiation theorem.

In what follows we simply write u instead of $\bar{u}$.

Step 2. *We have*

$$\|u\|_{L^\infty(\bar{U})} \leq c\|u\|_{\mathcal{L}^{p,\lambda}(U)}$$

for a constant $c > 0$ independent of u.

We compute, using (21.11),

$$\begin{aligned}
\|u\|_{L^\infty(\bar{U})} &\leq \sup_{x\in U} |u_{x,r_0}| + \sup_{x\in U} |u(x) - u_{x,r_0}| \\
&\leq c\left(\|u\|_{L^p(U)} + [u]_{\mathcal{L}^{p,\lambda}(U)} \right) \\
&\leq c\|u\|_{\mathcal{L}^{p,\lambda}(U)}.
\end{aligned}$$

Step 3. *We have*

$$[u]_{C^{0,\alpha}(\bar{U})} \leq c[u]_{\mathcal{L}^{p,\lambda}(U)}$$

for a constant $c > 0$ independent of u.

Suppose $x, y \in \bar{U}$. Set $r := |x - y|$. Then, using (21.11)

$$\begin{aligned}
|u(x) - u(y)| &\leq |u(x) - u_{x,2r}| + |u_{x,2r} - u_{y,r}| + |u_{y,r} - u(y)| \\
&\leq cr^{\alpha}[u]_{\mathcal{L}^{p,\lambda}(U)} + |u_{x,2r} - u_{y,r}|.
\end{aligned}$$

Finally,

$$\begin{aligned}|u_{x,2r} - u_{y,r}| &\le \fint_{B_r(y)\cap U} |u_{x,2r} - u| \\ &\le c \fint_{B_{2r}(x)\cap U} |u_{x,2r} - u| \\ &\le c r^\alpha [u]_{\mathcal{L}^{p,\lambda}(U)}.\end{aligned}$$

□

Proposition 21.12. *If $\alpha := \frac{\lambda - n}{p} > 0$, there is a constant $c > 0$ such that*

$$[u]_{\mathcal{L}^{p,\lambda}(U)} \le c[u]_{C^{0,\alpha}(\bar{U})}.$$

In particular, $C^{0,\alpha}(\bar{U}) \subset \mathcal{L}^{p,\lambda}(U)$.

Proof. We compute

$$\begin{aligned}\int_{B_r(x)\cap U} |u - u_{x,r}|^p &\le \int_{B_r(x)\cap U} \fint_{B_r(x)\cap U} |u(z) - u(y)|^p \, r dy \, \mathrm{d}z \\ &\le c r^{n+\alpha p} [u]^p_{C^{0,\alpha}(\bar{U})}.\end{aligned}$$

□

In the next lecture, we will also need the following.

Proposition 21.13. *Suppose $u \in L^p(U)$ and $v \in C^{0,\alpha}(U)$, then $uv \in \mathcal{L}^{p,p\alpha}$ and*

$$[uv]_{\mathcal{L}^{p,p\alpha}} \le c\|u\|_{L^p} \cdot [v]_{C^{0,\alpha}}.$$

Proof. This follows from

$$\int_{B_r(x)\cap U} |u(v - v(x))|^p \le (r^\alpha [v]_{C^{0,\alpha}})^p \int_{B_r(x)\cap U} |u|^p.$$

and

$$\begin{aligned}&\int_{B_r(x)\cap U} |(uv)_{x,r} - u \cdot v(x)|^p \\ &\quad = \int_{B_r(x)\cap U} \left| \fint_{B_r(x)\cap U} [(uv)(z) - u(y) \cdot v(x)] \, \mathrm{d}z \right|^p \mathrm{d}y \\ &\quad \le (r^\alpha [v]_{C^{0,\alpha}})^p \int_{B_r(x)\cap U} \left| \fint_{B_r(x)\cap U} |u(z) - u(y)| \, \mathrm{d}z \right|^p \mathrm{d}y.\end{aligned}$$

□

21.3 Morrey's theorem

The above combines well, with the Neumann–Poincaré inequality which we restate here for general p and in its scale invariant form.

Theorem 21.14 (Neumann–Poincaré inequality)**.** *There is a constant such that for all* $u \in C^\infty(\mathbf{R}^n)$ *and all* $x \in \mathbf{R}^n$

$$\int_{B_r(x)} |u - u_{x,r}|^p \le cr^p \int_{B_r(x)} |\nabla u|^p.$$

Theorem 21.15 (Morrey's inequality)**.** *If* $\alpha = 1 - \frac{n}{p} > 0$*, then is a constant* $c > 0$ *such that*

$$[u]_{C^{0,\alpha}(\bar{U})} \le c\|\nabla u\|_{L^p(U)}$$

In particular, $W^{1,p}(U) \hookrightarrow C^{0,\alpha}(\bar{U})$.

Proof. Combine Theorem 21.9 and Theorem 21.14. □

We also need the following closely related result (whose proof is similar to our proof of the Dirichlet–Poincaré inequality).

Theorem 21.16 (L^2 Sobolev inequality)**.** *There is a constant* $c > 0$ *such that for all* $u \in W_0^{1,2}(B_r)$ *we have*

$$\|u\|_{L^{\frac{2n}{n-2}}} \le c\|\nabla u\|_{L^2}.$$

22 Campanato estimates

22.1 Estimates for weak solutions

Proposition 22.1. *Suppose $A = (a_{ij})$ is a (constant) symmetric positive definite matrix with* $\operatorname{spec} A \subset [\lambda, \Lambda]$ *for $0 < \lambda \le \Lambda$. Then there is a constant $c = c(\Lambda/\lambda, n) > 0$ such that the following holds: For all $u \in W^{1,2}(B_r)$ satisfying*

$$\int_{B_r} a_{ij}(\partial_i u)(\partial_j \phi) = 0$$

for all $\phi \in W_0^{1,2}(B_r)$, and for all $s < r$ we have

$$\int_{B_s} |u|^2 \le c\left(\frac{s}{r}\right)^n \int_{B_r} |u|^2$$

and

$$\int_{B_s} |u - u_s|^2 \le c\left(\frac{s}{r}\right)^{n+2} \int_{B_r} |u - u_r|^2.$$

Moreover, $u \in C^\infty(B_r)$.

Note that A is diagonalised by a matrix in $\mathrm{SO}(n)$ and by a further coordinate stretch we can transform A into the identity matrix, that is, there is a $T \in \mathrm{GL}^+(\mathbf{R}^n)$ such that $T \circ A \circ T^* = \mathbf{1}$. Consequently $u \circ T$ is weakly harmonic on $T(B_r)$; hence, it is smooth and thus so is u. This observation allows us to reduce to the case $A = \mathbf{1}$.

Proof in the case $A = \mathbf{1}$. It suffices to consider $r = 1$ (by rescaling) and we can also restrict to $s \in (0, \frac{1}{4}]$ (since for $s > \frac{1}{4}$ the estimate is trivial).

We know that u is given by the representation formula (see Remark 16.7) for $x \in B_{1/4}$ and $t \in (0, 3/4)$ we have

$$u(x) = \int_{\partial B_t(x)} G(x-y)\partial_\nu u(y) - u(y)\partial_\nu G(x,y)\,\mathrm{d}y$$

and thus by integrating t over $[1/8, 1/4]$ and taking the supremum over x we get

$$\|u\|_{L^\infty(B_{1/4})} \le c\|u\|_{W^{1,2}(B_{1/2})}.$$

Now we want to estimate $\|u\|_{W^{1,2}(B_{1/2})}$. To do this pick a $\phi \in C_0^\infty(B_1)$ which is equal to one on $B_{1/2}$. Then by u being harmonic we have

$$\begin{aligned}\int_{B_1} |\phi\nabla u|^2 &\le -2\int_{B_1} \langle \phi\nabla u, u\nabla\phi\rangle \\ &\le \int_{B_1} \frac{1}{2}|\phi\nabla u|^2 + 2|u\nabla\phi|^2.\end{aligned}$$

Cancelling the $\frac{1}{2}|\phi\nabla u|^2$ we get

$$\int_{B_{1/2}} |\nabla u|^2 \leq c \int_{B_1} |u|^2. \tag{22.2}$$

Thus

$$\|u\|_{W^{1,2}(B_{1/2})} \leq c\|u\|_{L^2(B_1)}.$$

It follows that for $s \leq 1/4$

$$\int_{B_s} |u|^2 \leq cs^n \|u\|^2_{L^\infty(B_{1/4})} \leq cs^n \int_{B_1} |u|^2.$$

This proves the first estimate.

To prove the second estimate note that since u is harmonic, so are the components of ∇u. Moreover, by the Neumann–Poincaré inequality and Equation 22.2

$$\begin{aligned}\int_{B_s} |u - u_s|^2 &\leq cs^2 \int_{B_s} |\nabla u|^2 \\ &\leq cs^{n+2} \int_{B_{1/2}} |\nabla u|^2 \\ &\leq cs^{n+2} \int_{B_1} |u - u_1|^2.\end{aligned}$$

This completes the proof. □

22.2 Comparison estimates

Next we show that if u is close to a weak solution v then ∇u almost satisfies the estimate of Proposition 22.1.

Proposition 22.3. *Let $u \in W^{1,2}(B_r)$ and let $v \in W^{1,2}(B_r)$ be as in Proposition 22.1. Then for any $0 < s \leq r$ we have*

$$\int_{B_s} |\nabla u|^2 \leq c\left(\left(\frac{s}{r}\right)^n \int_{B_r} |\nabla u|^2 + \int_{B_r} |\nabla u - \nabla v|^2\right)$$

and

$$\int_{B_s} |\nabla u - (\nabla u)_s|^2 \leq c\left(\left(\frac{s}{r}\right)^{n+2} \int_{B_r} |\nabla u - (\nabla u)_r|^2 + \int_{B_r} |\nabla u - \nabla v|^2\right)$$

where $c = c(\Lambda/\lambda, n) > 0$.

Proof. Set $w := u - v$. Then by Proposition 22.1 for ∇v, we have

$$\begin{aligned}\int_{B_s} |\nabla u|^2 &\le 2 \int_{B_s} |\nabla v|^2 + |\nabla w|^2 \\ &\le c \left(\frac{s}{r}\right)^n \int_{B_r} |\nabla v|^2 + 2 \int_{B_r} |\nabla w|^2 \\ &\le c \left(\frac{s}{r}\right)^n \int_{B_r} |\nabla u|^2 + c \left(1 + \left(\frac{s}{r}\right)^n\right) \int_{B_r} |\nabla w|^2.\end{aligned}$$

This gives the first estimate.

For the second we start with the following exercise.

Exercise 22.4. If U is an bounded open subset of $\mathbf{R}^n$ and $f \in L^2(U)$, then

$$\int_U |f - c|^2$$

is minimal for $c = \fint_U f$.

Now we compute

$$\begin{aligned}\int_{B_s} |\nabla u - (\nabla u)_s|^2 &\le \int_{B_s} |\nabla u - (\nabla v)_s|^2 \\ &\le 2 \int_{B_s} |\nabla v - (\nabla v)_s|^2 + |\nabla w|^2 \\ &\le c \left(\frac{s}{r}\right)^{n+2} \int_{B_r} |\nabla v - (\nabla v)_s|^2 + \int_{B_r} |\nabla w|^2 \\ &\le c \left(\frac{s}{r}\right)^{n+2} \int_{B_r} |\nabla u - (\nabla u)_s|^2 + c \left(1 + \left(\frac{s}{r}\right)^{n+2}\right) \int_{B_r} |\nabla w|^2.\end{aligned}$$

□

22.3 Operators with variable coefficients

We now lift the restriction that a_{ij} be constant and instead require that $a_{ij} \in L^\infty(U)$ and that there are constant $\Lambda, \lambda > 0$ such that for all $x \in U$

$$\operatorname{spec} A(x) \in [\lambda, \Lambda],$$

that is the operator

$$L = -\sum_{ij} \partial_i (a_{ij} \partial_j)$$

is uniformly elliptic. Again we say that u is a weak solution of

$$Lu = f - \sum_i \partial_i h_i$$

for some $f, h_i \in L^2(U)$ and for all $\phi \in W_0^{1,2}(U)$ we have

$$\sum_{ij} \int_U a_{ij}(\partial_i u)(\partial_i \phi) = \int_U f\phi + \sum_i \int_U h\partial_i \phi. \tag{22.5}$$

Proposition 22.6. *Assume that $a_{ij} \in C^0(\bar{U})$. Suppose $u \in W^{1,2}(U)$ is a weak solution of $Lu = f$ and $f \in L^p(U)$, $h \in \mathcal{L}^{2,n-2+2\alpha}$ with $p \in (n/2, n)$ (that is $\alpha := 2 - n/p \in (0,1)$). Then $u \in C^{0,\alpha}_{\text{loc}}(U)$. Moreover, if $V \subset\subset U$ there exists a constant $c = (a_{ij}, V, n) > 0$ such that*

$$\|u\|_{C^{0,\alpha}(V)} \le c\left(\|f\|_{L^p(U)} + \|u\|_{W^{1,2}(U)} + [h]_{\mathcal{L}^{2,n-2+2\alpha}}\right).$$

Proof. If $f = 0$ and a_{ij} were constant, then the assertion would follow (with arbitrary $\alpha \in (0,1)$) from Proposition 22.1 and Theorem 21.9. To be able to talk about the variations in a_{ij} we introduce

$$\omega(r) := \sup\{|a_{ij}(x) - a_{ij}(y)| : x, y \in U \text{ and } |x-y| \le r\}.$$

Now suppose $\bar{B}_s(x) \subset \bar{B}_r(x) \subset U$. (We can assume that $r > 0$ is very small.) Our goal is to show that

$$\int_{B_s(x)} |\nabla u|^2 \le c s^{n-2+2\alpha}\left(\int_U |\nabla u|^2 + \|f\|^2_{L^p}\right).$$

If we can achieve this, then the assertion follows from the Neumann–Poincaré inequality and Theorem 21.9.

The idea for achieving this is to compare u to a solution of a constant coefficient equation with the same boundary values.

Proposition 22.7. *Denote by $v \in W^{1,2}(B_r(x))$ the unique weak solution to*

$$-\sum_{ij} \partial_i(a_{ij}(x)\partial_j v) = 0$$

satisfying $u - v \in W_0^{1,2}(B_r(x))$. Then

$$\int_{B_r(x)} |\nabla(u-v)|^2 \le c\left(\omega(r)^2 \int_{B_r(x)} |\nabla u|^2 + r^\lambda [h]_{\mathcal{L}^{2,\lambda}} + r^{n+2-2n/p}\|f\|^2_{L^p}\right).$$

Proof. First we write (22.5) as

$$\sum_{ij} \int_U a_{ij}(x)(\partial_i u)(\partial_j \phi) = -\sum_{ij}\int_U (a_{ij}(x) - a_{ij})(\partial_i u)(\partial_j \phi)$$
$$- \sum_i \int_U h_i \partial_i \phi + \int_{B_1} f\phi.$$

Denote by $v \in W^{1,2}(B_r(x))$ the unique weak solution to

$$-\sum_{ij} \partial_i(a_{ij}(x)\partial_j v) = 0$$

satisfying $w := u - v \in W_0^{1,2}(B_r(x))$.

Now w satisfies

$$\begin{aligned}\sum_{ij}\int_{B_r(x)} a_{ij}(x)(\partial_i w)(\partial_i \phi) &= -\sum_{ij}\int_{B_r(x)} (a_{ij}(x) - a_{ij})(\partial_i u)(\partial_i \phi) \\ &\quad - \sum_i \int_{B_r(x)} h_i \partial_i \phi + \int_{B_1} f\phi.\end{aligned}$$

for all $\phi \in W_0^{1,2}(B_r(x))$. In particular for $\phi = w$ and using uniform ellipticity.

$$\lambda \int_{B_r(x)} |\nabla w|^2 \leq \sum_{ij}\int_{B_r(x)} (|a_{ij}(x) - a_{ij}||\partial_i u| + |h|)|\partial_i w| + \int_{B_r(x)} |f||v|$$

Note that

$$\sum_{ij}\int_{B_r(x)} (|a_{ij}(x) - a_{ij}||\partial_i u| + |h|)|\partial_i w| \leq \int_{B_r(x)} \frac{\lambda}{2}|\nabla w|^2 + \lambda^{-1}\omega(r)^2|\nabla u|^2 + r^\lambda [h]^2_{\mathcal{L}^{2,\lambda}},$$

and by the Hölder inequality and the L^2 Sobolev inequality (Theorem 21.16)

$$\begin{aligned}\int_{B_r(x)} |f||w| &\leq \left(\int_{B_r(x)} |f|^{\frac{2n}{n+2}}\right)^{\frac{n+2}{2n}} \left(\int_{B_r(x)} |w|^{\frac{2n}{n-2}}\right)^{\frac{n-2}{2n}} \\ &\leq \left(\int_{B_r(x)} |f|^{\frac{2n}{n+2}}\right)^{\frac{n+2}{2n}} \left(\int_{B_r(x)} |\nabla w|^2\right)^{1/2} \\ &\leq \varepsilon^{-1}\left(\int_{B_r(x)} |f|^{\frac{2n}{n+2}}\right)^{\frac{n+2}{n}} + \varepsilon \int_{B_r(x)} |\nabla w|^2.\end{aligned}$$

Moreover, by the Hölder inequality

$$\left(\int_{B_r(x)} |f|^{\frac{2n}{n+2}}\right)^{\frac{n+2}{n}} \leq r^{n+2-2n/p}\left(\int_{B_r(x)} |f|^p\right)^{2/p}$$

Putting all of this together we get for some constant $c > 0$,

$$\int_{B_r(x)} |\nabla w|^2 \leq c\left(\omega(r)^2 \int_{B_r(x)} |\nabla u|^2 + r^\lambda [h]_{\mathcal{L}^{2,\lambda}} + r^{n+2-2n/p}\|f\|^2_{L^p}\right).$$

□

In our situation the above gives

$$\int_{B_r(x)} |\nabla w|^2 \le c\left(\omega(r)^2 \int_{B_r(x)} |\nabla u|^2 + r^{n-2+2\alpha}[h]_{\mathcal{L}^{2,n-2+2\alpha}} + r^{n-2+2\alpha}\|f\|^2_{L^p}\right).$$

Thus by Proposition 22.3 we have for $B_s(x) \subset B_r(x)$

$$\int_{B_s(x)} |\nabla u|^2 \le c\left(\left(\frac{s}{r}\right)^n + \omega(r)^2\right) \int_{B_r(s)} |\nabla u|^2 + cr^{n-2+2\alpha}([h]_{\mathcal{L}^{2,n-2+2\alpha}} + \|f\|^2_{L^p})$$

This seems to be not good enough because the first term on the right-hand side still has the error term $\omega(r)^2$ (although we can assume this to be very small) and in the second term we want to have $s^{n-2+2\alpha}$ instead of $r^{n-2-2\alpha}$. However, Lemma 22.8 shows that the above estimates does in fact imply the desired estimate. □

Lemma 22.8. *Let $\phi\colon [0, R] \to [0, \infty)$ be a non-decreasing function. Suppose there are constant $A, B, \alpha, \beta > 0$ and $\alpha > \beta$ such that for all $0 \le s \le r \le R$ we have*

$$\phi(s) \le A((s/r)^\alpha + \varepsilon) + Br^\beta.$$

Then there is a $\varepsilon_0 > 0$ and $c > 0$ such that if $\varepsilon \le \varepsilon_0$ then

$$\phi(s) \le c(s/r)^\beta(\phi(r) + B).$$

Proof. Pick $\tau \ll 1$ and $\varepsilon_0 \ll 1$ such that

$$A\tau^\alpha = \frac{1}{4}\tau^\beta \quad \text{and} \quad \varepsilon\tau^{-\alpha} \le 1.$$

Then for $s = \tau r$, we get

$$\phi(\tau r) \le \left(\frac{1}{2}\phi(r) + \frac{B}{\tau^\beta} r\right)\tau^\beta.$$

Thus

$$\begin{aligned}
\phi(\tau^{k+1} r) &\le \left(\frac{1}{2}\phi(\tau^k r) + \frac{B}{\tau^\beta}\tau^k r\right)\tau^\beta \\
&\le \left(\phi(r) + \frac{B}{\tau^\beta} r \sum_{i=0}^k 2^{-i}\right)\tau^{(k+1)\beta} \\
&\le \left(\phi(r) + \frac{2Br}{\tau^\beta}\right)\tau^{(k+1)\beta}.
\end{aligned}$$

Now given s pick k such that $\tau^{k+2} r \le s \le \tau^{k+1} r$ to get

$$\phi(s) \le \left(\phi(r)/\tau^\beta + \frac{2BR}{\tau^\beta}\right)(s/r)^\beta.$$

□

23 Higher Regularity

We continue with the setting of the last two lectures.

23.1 $C^{1,\alpha}$ estimate

Proposition 23.1. *Suppose $u \in W^{1,2}(U)$ is a weak solution of $Lu = f - \partial_i h_i$, $p > n$ and $\alpha := 1 - \frac{n}{p}$, $a_{ij} \in C^{0,\alpha}(\bar{U})$, $f \in L^p(U)$, $h \in \mathcal{L}^{2,n+2\alpha}$. Then $\nabla u \in C^{0,\alpha}_{\text{loc}}(U)$ and for each $V \subset\subset U$, there is a constant $c = c(a_{ij}, V, n) > 0$ such that*

$$\|\nabla u\|_{C^{0,\alpha}(\bar{V})} \le c \left(\|f\|_{L^p(U)} + [h]_{\mathcal{L}^{2,n+2\alpha}(U)} + \|u\|_{W^{1,2}(U)}. \right)$$

Proof. We can assume that $V = B_{1/2}$ and $U = B_1$. We will show that for $x \in B_{1/2}$ and $r > 0$ sufficiently small we have

$$\int_{B_r(x)} |\nabla u - (\nabla u)_{x,r}|^2 \le c r^{n+2\alpha} \left(\|f\|^2_{L^p(U)} + [h]_{\mathcal{L}^{2,n+2\alpha}(U)} + \|u\|^2_{W^{1,2}(U)} \right).$$

The assertion then follows from Theorem 21.9.

Suppose $x \in B_{3/4}$ and $r > 0$ is small. (Note that $B_{3/4}$ is larger than V. This is on purpose.) We can write the equation in the form

$$\sum_{ij} \int_U a_{ij}(x)(\partial_i u)(\partial_i \phi) = -\sum_{ij} \int_U (a_{ij}(x) - a_{ij})(\partial_i u)(\partial_i \phi) + \int_{B_1} f\phi.$$

Denote by $v \in W^{1,2}(B_r(x))$ the unique weak solution to

$$-\sum_{ij} \partial_i (a_{ij}(x) \partial_j v) = 0$$

satisfying $w := u - v \in W^{1,2}_0(B_r(x))$, as in Proposition 22.7. Then we know that

$$\int_{B_r(x)} |\nabla(u - v)|^2 \le c \left(r^{2\alpha} [a_{ij}]_{C^{0,\alpha}(\bar{B}_r(x))} \int_{B_r(x)} |u|^2 + r^{n+2\alpha} (\|f\|^2_{L^p} + [h]^2_{\mathcal{L}^{2,n+2\alpha}(U)}) \right).$$

Thus by Proposition 22.3 we have for $B_s(x) \subset B_r(x)$

$$\int_{B_s(x)} |\nabla u|^2 \le c \left(\left(\frac{s}{r} \right)^n + r^{2\alpha} [a_{ij}]_{C^{0,\alpha}} \right) \int_{B_r(s)} |\nabla u|^2 + c r^{n+2\alpha} (\|f\|^2_{L^p} + [h]^2_{\mathcal{L}^{2,n+2\alpha}(U)}) \tag{23.2}$$

and

$$\begin{aligned} \int_{B_s(x)} |\nabla u - (\nabla u)_{x,s}|^2 \le{}& c \left(\frac{s}{r} \right)^{n+2} \int_{B_r(s)} |\nabla u - (\nabla u)_{x,r}|^2 \\ &+ c r^{2\alpha} [a_{ij}]_{C^{0,\alpha}} \int_{B_r(s)} |\nabla u|^2 \\ &+ c r^{n+2\alpha} (\|f\|^2_{L^p} + +[h]^2_{\mathcal{L}^{2,n+2\alpha}(U)}) \end{aligned} \tag{23.3}$$

If a_{ij} is constant, then $[a_{ij}]_{C^{0,\alpha}} = 0$ and the above and Lemma 22.8 imply the assertion. If a_{ij} is not constant, we have to work a bit harder. What we need to do is to control

$$\int_{B_r(s)} |\nabla u|^2.$$

We need to show that this terms is bounded by a constant times r^n.

Using Lemma 22.8 and (23.2) we can show that for all $\delta > 0$

$$\int_{B_s(x)} |\nabla u|^2 \leq cs^{n-2\delta} \left(\|\nabla u\|_{L^2}^2 + \|f\|_{L^p}^2 + +[h]^2_{\mathcal{L}^{2,n+2\alpha}(U)} \right).$$

This means that if $x \in V = B_{1/2} \subset B_{3/4}$, then

$$\int_{B_r(x)} |\nabla u|^2 \leq cr^{n-2\delta} \left(\|\nabla u\|_{L^2}^2 + \|f\|_{L^p}^2 + [h]^2_{\mathcal{L}^{2,n+2\alpha}(U)} \right).$$

Plugging this into (23.3) we get

$$\begin{aligned}
\int_{B_s(x)} |\nabla u - (\nabla u)_{x,s}|^2 \leq c \left(\frac{s}{r}\right)^{n+2} & \int_{B_r(s)} |\nabla u - (\nabla u)_{x,r}|^2 \\
&+ cr^{n+2\alpha-2\delta} \int_{B_1} |\nabla u|^2 \\
&+ cr^{n+2\alpha} (\|f\|_{L^p}^2 + [h]^2_{\mathcal{L}^{2,n+2\alpha}(U)})
\end{aligned}$$

and using Lemma 22.8, we arrive at

$$\int_{B_s(x)} |\nabla u - (\nabla u)_{x,s}|^2 \leq cs^{n+2\alpha-2\delta} \left(\|\nabla u\|_{L^2}^2 + \|f\|_{L^p}^2 \right).$$

This is not quite good enough. However, using Theorem 21.9 we derive that

$$\|\nabla u\|^2_{L^\infty(B_{3/4})} \leq c \left(\|\nabla u\|_{L^2}^2 + \|f\|_{L^p}^2 + [h]^2_{\mathcal{L}^{2,n+2\alpha}(U)} \right);$$

hence,

$$\int_{B_r(s)} |\nabla u|^2 \leq cr^n.$$

This completes the proof. □

23.2 Bootstrapping $C^{k,\alpha}$ estimates

Now differentiating $Lu = f$ in the direction of x_k we see that $v := \partial_k u$ is a weak solution of

$$Lv = \partial_k f + \sum_i \partial_i h_i$$

with $h = (\partial_k a_{ij} \partial_j u) \in C^{0,\alpha} \cong \mathcal{L}^{2,n+2\alpha}$. (This particular point is we why introduced the seemingly odd h_i terms to begin with.) Applying Proposition 22.6 and Proposition 23.1 again, we learn that $\partial_k u \in C^{1,\alpha}$ (with concrete estimates) and by the regularity theory of Δ, in fact, $\partial_k u \in C^2$.

One can keep running this argument to show that in fact $u \in C^\infty$ and get concrete $C^{k,\alpha}$ estimates.

Proposition 23.4. *Suppose $u \in W^{1,2}(U)$ is a weak solution of $Lu = f$ and $a_{ij}, f \in C^{k,\alpha}(U)$. Then $u \in C^{k+2,\alpha}_{\text{loc}}(U)$ and for each $V \subset\subset U$, there is a constant $c = c(a_{ij}, V, n) > 0$ such that*

$$\|u\|_{C^{k+2,\alpha}(\bar{V})} \leq c \left(\|f\|_{C^{k,\alpha}(U)} + \|u\|_{W^{1,2}(U)}. \right)$$

23.3 Boundary estimates

One drawback of Proposition 23.5 is that we only have interior estimates. With some more work one can show the following.

Proposition 23.5. *Suppose $u, \bar{u} \in C^{k+2,\alpha}(\bar{U})$ with $u|_{\partial U} = \bar{u}|_{\partial U}$, $a_{ij} \in C^{k,\alpha}(\bar{U})$ and ∂U is smooth. Then*

$$\|u\|_{C^{k+2,\alpha}(\bar{U})} \leq c \left(\|Lu\|_{C^{k,\alpha}(\bar{U})} + \|\bar{u}\|_{C^{k+2,\alpha}} + \|u\|_{L^2(U)} \right).$$

Exercise 23.6. If $a_{ij}, b_i, c \in C^\infty(\bar{U})$, and a_{ij} is uniformly elliptic, then the estimate in Proposition 23.5 also holds for the differential operator

$$Lu := -\sum_{ij} \partial_i (a_{ij} \partial_j u) + \sum_k b_k \partial_k u + cu,$$

of course with a different constant.

A Divergence Theorem

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

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

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

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 $\nu\colon \partial U \to \mathbf{R}^n$ be the outward-pointing normal vector-field to ∂U. If $v\colon \bar{U} \to \mathbf{R}^n$ is a continuously differentiable vector field, then*

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

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

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

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

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

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

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

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

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

and

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

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

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 $d\colon X \times X \to [0, \infty)$, 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 $\mathbf{R}$ together with the function $d\colon \mathbf{R} \times \mathbf{R} \to [0, \infty)$ 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

$$d\left((x_1, x_2), (y_1, y_2)\right) := d(x_1, y_1) + d(x_2, y_2).$$

In particular, $\mathbf{R}^n$ is a metric space.

Example B.6. Given $U \subset \mathbf{R}^n$, we set

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

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

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

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

An important notion that metric spaces inherit from $\mathbf{R}$ 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 $(x_n) \in X^{\mathbf{N}}$ 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 $(x_n) \in X^{\mathbf{N}}$ 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 $\mathbf{R}$, 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. $\mathbf{Q}$ with the distance inherited from $\mathbf{R}$ is incomplete.

Exercise B.12. Prove that if $\bar{U} \subset \mathbf{R}^n$ is compact, then $C^0(\bar{U}, \mathbf{R}^m)$ 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 $f\colon \bar{U} \to \mathbf{R}^m$.
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

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

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

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

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

$$\bar{d}\,((x_n), (y_n)) := \inf_{n\in\mathbf{N}} d(x_n, y_n).$$

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 $\mathbf{Q}$ is $\mathbf{R}$.

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_∞ and a sequence of continuous functions converging to f_∞ 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 $T\colon X \to X$ 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 γ?)

C Fourier Series on $[0, 1]$

Definition C.1. A Hilbert space is a vector space H together with an inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbf{R}$ 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

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

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

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

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 $(e_i)_{i \in \mathbf{N}} \in H^{\mathbf{N}}$ such that

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

and, for every $x \in H$, $\langle x, e_i \rangle = 0$ for all $i \in \mathbf{N}$ if and only if $x = 0$.

Remark C.6. If H is not finite dimensional, then an orthonormal basis of the Hilbert space H is not a basis of the vector space H, since not every element can be written as a *finite* linear combination.

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

$$a_i := \langle x, f_i\rangle \quad \textit{and} \quad b_i := \langle y, f_i\rangle$$

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 $(a_n) \in \mathbf{R}^{\mathbf{N}}$ 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 Δ becomes *diagonal* in the orthonormal basis (f_n). Note, however, that the eigenvalues do go to infinity; hence, Δ 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* $k \in \mathbf{N}_0 = \{0, 1, 2, \ldots\}$*. If the Fourier coefficients of* $f \in L^2([0,1])$ *satisfy*

$$(a_n) \in \ell_k^1 := \{(b_n) \in \mathbf{R}^{\mathbf{N}} : \sum_{n=1}^{\infty} n^k |b_n| < \infty\},$$

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

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

Show that the Fourier coefficients of f are

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

and $a_n = 0$ if $n \neq 2 \bmod 4$. 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 $\mathbf{R}^n$.

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

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

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

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

then

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

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

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

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

$$F(t) := \int_U f(t,x)\,\mathrm{d}x$$

is differentiable and

$$\partial_t F(t) = \int_U \partial_t f(t,x)\,\mathrm{d}x.$$

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

[1] V. I. Arnold. *Lectures on partial differential equations*. Universitext. Translated from the second Russian edition by Roger Cooke. Springer-Verlag, Berlin; Publishing House PHASIS, Moscow, 2004, pp. x+157. URL: http://dx.doi.org/10.1007/978-3-662-05441-3 (cit. on pp. 11, 36).

[2] E. Bombieri, E. De Giorgi, and E. Giusti. "Minimal cones and the Bernstein problem". In: *Invent. Math.* 7 (1969), pp. 243–268 (cit. on p. 88).

[3] E. Hopf. "On S. Bernstein's theorem on surfaces $z(x, y)$ of nonpositive curvature". In: *Proc. Amer. Math. Soc.* 1 (1950), pp. 80–85 (cit. on p. 89).

[4] P. Li and R. Schoen. "L^p and mean value properties of subharmonic functions on Riemannian manifolds". In: *Acta Math.* 153.3-4 (1984), pp. 279–301. URL: http://dx.doi.org/10.1007/BF02392380 (cit. on p. 64).

[5] E. J. Mickle. "A remark on a theorem of Serge Bernstein". In: *Proc. Amer. Math. Soc.* 1 (1950), pp. 86–89 (cit. on p. 89).

[6] J. Simons. "Minimal varieties in riemannian manifolds". In: *Ann. of Math. (2)* 88 (1968), pp. 62–105 (cit. on p. 88).