Partial Differential Equations

MAST90133

Volker Schlue
University of Melbourne

Semester 2, 2021

(version: November 23, 2023)

Contents

Preface

I have written these lecture notes as an introduction to partial differential equations, which I taught in 2021, and 2023, in the Master of Science course at the University of Melbourne.

This class is mostly compiled from selected chapters of the *Princeton Lectures in Analysis*, by Elias Stein and Rami Shakarchi, as well as from selected chapter of Fritz John's book on PDEs. The student will encounter here more Fourier analysis than in other introductions, certainly compared to my first encounter with the subject, but it won't hurt, and it offers an easy connection point to typical third years classes.

The class can be taught over 12 weeks, assignments were written separately and are not included here. — VS, Melbourne, November 2023.

Part I.

Introduction

Lecture 2.

Wave and heat equations: a first look

Further Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 1, Section 1, 2), and (John, *Partial differential equations*, Chapter 2, Section 4)

2.1. One-dimensional wave equation

Many of you will have encountered the one-dimensional wave equation, as it arises for example in the descriptions of a vibrating string. As a graph $y = u(t, x)$ over the x-axis, the height function satisfies

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \tag{2.1}$$

where c is the speed by which a vibration can travel in the string.

It is easy to see that the **traveling waves** $u(t, x) = f(x + ct)$, and $u(t, x) = f(x - ct)$ are solutions for *any* twice differentiable function f. In fact, let us show that we can always find twice differentiable functions f, and g such that

$$u(t, x) = f(x + ct) + g(x - ct)\,. \tag{2.2}$$

To see this we introduce new variables

$$\xi = x + ct \qquad \eta = x - ct \tag{2.3}$$

then the new unknown $v(\xi, \eta) = u((\xi - \eta)/2c, (\xi + \eta)/2)$ satisfies

$$\frac{\partial^2 v}{\partial \xi \partial \eta} = 0\,. \tag{2.4}$$

We can integrate this relation twice to find $v = f(\xi) + g(\eta)$ for some functions f, and g.

Remark 2.1. The formula (2.2) can interpreted as the statement that the general solution to (2.1) is a superposition of a solution $v_+ = f(x + ct)$ solving $\partial_t v - c\partial_x v = 0$ and a solution $v_- = f(x - ct)$ solving $\partial_t v + c\partial_x v = 0$. These are examples for first-order equations that we will study in Lecture 4.

Moreover these functions are uniquely determined from the **initial conditions**

$$u(0,x) = u_0(x) \qquad \partial_t u(0,x) = u_1(x)\,. \tag{2.5}$$

Since $u(0,x) = f(x) + g(x) = u_0(x)$, and $\partial_t u(0,x) = cf'(x) - cg'(x) = u_1(x)$, it follows after differentiating the first relation, and adding it to the second that

$$2cf'(x) = cu_0'(x) + u_1(x) \tag{2.6}$$

and similarly after subtracting,

$$2cg'(x) = cu_0'(x) - u_1(x) \tag{2.7}$$

and hence there are constants C_1, and C_2 so that

$$f(x) = \frac{1}{2}\Big[u_0(x) + \frac{1}{c}\int_0^x u_1(y)\mathrm{d}y\Big] + C_1 \tag{2.8}$$

and

$$g(x) = \frac{1}{2}\Big[u_0(x) - \frac{1}{c}\int_0^x u_1(y)\mathrm{d}y\Big] + C_2 \tag{2.9}$$

Since $f + g = u_0$ it follows that $C_1 + C_2 = 0$, and therefore the solution to (2.1) is given in terms of the initial conditions by the formula:

$$u(t,x) = \frac{1}{2}\Big[u_0(x+ct) + u_0(x-ct)\Big] + \frac{1}{2c}\int_{x-ct}^{x+ct} u_1(y)\mathrm{d}y \tag{2.10}$$

This formula is known as **d'Alembert's formula**.

We see from this formula that $u(t,x)$ is determined uniquely by the values of the initial data on the interval $[x-ct, x+ct]$ whose end points are the points of intersection of the characteristics through (t,x) with the x-axis. This interval is the **domain of dependence** for the solution u at the point (t,x); see Fig. 2.1 and compare to the discussion in Lecture 4. Conversely, the values of the initial data at $(0,x_0)$ can influence the solution u only at point (t,x) that lie in the "forward cone" with vertex at $(0,x_0)$, namely with $x_0 - ct < x < x_0 + ct$; see Figure 2.1.

Another idea to represent the solution — which is familiar from courses in physics, but also reappears in the modern theory of partial differential equations — is the idea that it might be possible to *separate variables* and write the solution in the form $u(t,x) = \varphi(x)\psi(t)$. This leads very quickly to the system

$$\psi''(t) - \lambda\psi(t) = 0 \tag{2.11a}$$

$$\varphi''(x) - \lambda\varphi(x) = 0 \tag{2.11b}$$

where λ is a *constant*. Now taking $\lambda < 0$ and considering the case when the string is attached at $x = 0$ and $x = \pi$, namely imposing the conditions $\varphi(0) = \varphi(\pi) = 0$, one finds that for $\lambda = -m^2$, where m is an integer the solution

$$u_m(t,x) = \Big(a_m\cos(mt) + b_m\sin(mt)\Big)\sin(mx) \tag{2.12}$$

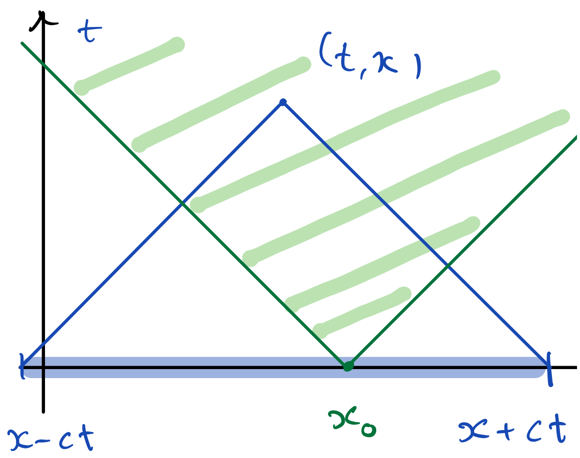

Figure 2.1.: Domain of dependence and domain of influence for the one-dimensional wave equation.

for some constants a_m, b_m. These solutions are called **harmonics** (they correspond to the overtones of say a violin string), and the higher m the higher the frequency of the **standing wave**. In view of the **linearity** of the equation, the question is if *any* solution could written as a **superposition** of these harmonics. If so we would have in particular at $t = 0$ that

$$\sum_{m=1}^{\infty} a_m \sin(mx) = u_0(x)\,. \tag{2.13}$$

The question for which functions u_0 — with $u_0(0) = u_0(\pi) = 0$ — one can find constants a_m such that this identity holds is the starting point of **Fourier Analysis**.

2.2. Heat equation

Another familiar equation is the time-dependent heat equation, arising for example as a model for the temperature distribution $u(t, x, y)$ on a metal plate:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{2.14}$$

In thermal equilibrium when $\partial u/\partial t = 0$, the temperature distributions satisfies the **steady-state heat equation**:

$$\Delta u = 0 \tag{2.15}$$

where $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the **Laplacian** on $\mathbb{R}^2$.

Consider the unit disc in the plane

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\} \tag{2.16}$$

whose boundary is the unit circle C. The problem of finding a solution to (2.15) with prescribed boundary condition $u = f$ on C is called the **Dirichlet problem** (on the unit disc).

Since the boundary condition is most easily expressed in polar coordinates (r, θ), it is convenient to note that in the plane

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial \theta^2}\,. \tag{2.17}$$

Now if we pursue the idea to seperate variables, namely an ansatz of the form $u(t, x, y) = F(r)G(\theta)$, then we find

$$\frac{r^2 F''(r) + rF'(r)}{F(r)} = -\frac{G''(\theta)}{G(\theta)} \tag{2.18}$$

Since each side depends on a different variable, they must both be constant, say equal to λ. We get on one hand $G''(\theta) + \lambda G(\theta) = 0$ which is solved by the 2π-periodic function $G(\theta) = Ae^{im\theta} + Be^{-im\theta}$ provided $\lambda = m^2$ and m is an integer. On the other hand we get

$$r^2 F''(r) + rF'(r) - m^2 F(r) = 0 \tag{2.19}$$

which has the simple solutions $F(r) = r^m$, and $F(r) = r^{-m}$ for $m \neq 0$, and $F(r) = 1$, and $F(r) = \log(r)$ in the case $m = 0$. Since the solutions $r^{-|m|}$ and $\log(r)$ are unbounded we dismiss them and we are left with the special solutions:

$$u_m(r, \theta) = r^{|m|} e^{im\theta} \qquad (m \in \mathbb{Z}) \tag{2.20}$$

In view of the **linearity** of (2.15) one may ask if a general solution can be obtained as the *superposition*:

$$u(r, \theta) = \sum_{m \in \mathbb{Z}} a_m r^{|m|} e^{im\theta} \tag{2.21}$$

If this were true, then we would have in particular

$$u(1, \theta) = \sum_{m=-\infty}^{\infty} a_m e^{im\theta} = f(\theta)\,. \tag{2.22}$$

The question for which functions f on $[0, 2\pi]$ with $f(0) = f(2\pi)$ it is possible to find coefficients a_m so that this holds is answered by Fourier Analysis.

Problems

1. Let u_0 and u_1 in (2.5) have compact support. Show that the solution $u(t, x)$ of (2.1) has compact support in x for each t. Show that the functions f, and g in (2.2) can have compact support only when

$$\int_{-\infty}^{\infty} u_1(x)\mathrm{d}x = 0\,. \tag{2.23}$$

2. Solve

$$u_{tt} - c^2 u_{xx} = x^2 \qquad (t > 0) \tag{2.24}$$
$$u(0,x) = x \qquad u_t(0,x) = 0\,. \tag{2.25}$$

   Hint: First find a special time-independent solution of the PDE.

3. Verify (2.17), and also show that

$$|\frac{\partial u}{\partial x}|^2 + |\frac{\partial u}{\partial y}|^2 = |\frac{\partial u}{\partial r}|^2 + \frac{1}{r^2}|\frac{\partial u}{\partial \theta}|^2 \tag{2.26}$$

4. Show that if $m \in \mathbb{N}$ the only solutions of the differential equation (2.18) which are twice differentiable when $r > 0$ are those indicated above.

   Hint: If F solves (2.18) then write $F(r) = g(r)r^m$, and find the equation satisfied by g, and conclude that $rg' + 2mg$ is constant.

Digression: Convolutions

Further Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 2, Section 1, 3, 4).

In the previous Lecture we have arrived at the question if any "reasonable" function f on $[-\pi, \pi]$ can be expressed as a **Fourier series**

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx} . \tag{2.1}$$

(We also call functions on $[-\pi, \pi]$, and 2π periodic functions on $\mathbb{R}$, *functions on the circle.*)

Since

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{imx} e^{-inx} \mathrm{d}x = \begin{cases} 0 & \text{if } n \neq m \\ 1 & \text{if } n = m \end{cases} \tag{2.2}$$

we can expect that

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \mathrm{d}x \tag{2.3}$$

The number a_n thus defined is called the n^{th} **Fourier coefficient** of f.

When considering the question of *convergence* of the Fourier series we are naturally lead to the partial sums

$$S_N(f)(x) = \sum_{n=-N}^{N} a_n e^{inx} \tag{2.4}$$

which are *trigonometric polynomials* of order N, and the question of convergence can be phrased in terms of convergence of $S_N(f)$ to f as $N \to \infty$.

Since

$$S_N(f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) \Big(\sum_{n=-N}^{N} e^{in(x-y)} \Big) \mathrm{d}y = (f * D_N)(x) \tag{2.5}$$

where we have inserted the defintion (2.3) into the expression (2.4) and interchanged sum and integral, we see that the partial sums S_N can be written as a *convolution* of f with the **Dirichlet kernel**

$$D_N(x) = \sum_{n=-N}^{N} e^{inx} . \tag{2.6}$$

Similarly we will see that the series that we have encountered in the discussion of the steady state heat equation on the disc can be written as a convolution with the **Poisson kernel**

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} . \tag{2.7}$$

2.1. Convolutions

Given two 2π-periodic integrable functions f and g on $\mathbb{R}$ their **convolution** on $[-\pi, \pi]$ is defined by

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)\mathrm{d}y \tag{2.8}$$

Remark 2.1. Loosely speaking, convolutions correspond to "weighted averages". For example if $g = 1$, then $f * g$ is constant and equal to the average of f on $[-\pi, \pi]$.

Proposition 2.1. *Suppose f and g are 2π periodic integrable functions on $\mathbb{R}$. Then*

1. $f * (g + h) = f * g + f * h$
2. $(cf) * g = c(f * g) = f * (cg)$
3. $f * g = g * f$
4. $(f * g) * h = f * (g * h)$
5. *$f * g$ is continuous.*

Remark 2.2. The last statement shows that the convolution $f * g$ is *more regular* than the functions f, and g, which are here merely assumed to be (Riemann) integrable.

A key observation in the context of Fourier analysis is that *the Fourier coefficients of $f * g$ are the product of the Fourier coefficients of f and g,* but we will not immediately need that here.

2.2. Good Kernels

A family of kernels $\{K_n(x)\}_{n=1}^{\infty}$ on the circle is said to be a family of **good kernels** if it satisfies the following properties:

1. For all $n \geq 1$,
$$\frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(x)\mathrm{d}x = 1 \tag{2.9}$$
2. There exists $M > 0$ such that for all $n \geq 1$,
$$\int_{-\pi}^{\pi} |K_n(x)|\mathrm{d}x \leq M \tag{2.10}$$

3. For every $\delta > 0$,

$$\int_{\delta \leq |x| \leq \pi} |K_n(x)| \mathrm{d}x \to 0 \qquad \text{as } n \to \infty \tag{2.11}$$

These kernels are bounded functions which have unit mass, but they concentrate near the origin as $n \to \infty$. They are also referred to as **approximations to the identity** which is a terminology that comes from the following general result:

Theorem 2.2. *Let $\{K_n\}$ be a family of good kernels, and f an integrable function on the circle. Then*

$$\lim_{n \to \infty} (f * K_n)(x) = f(x) \tag{2.12}$$

whenever f is continuous at x. If f is continuous everywhere, then the above limit is uniform.

Remark 2.3. We have thought of convolutions as weighted averages. Here

$$(f * K_n)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-y) K_n(y) \mathrm{d}y \tag{2.13}$$

the weights K_n concentrate its mass at $y = 0$ as n becomes large, until in the limit the full mass is assigned at $y = 0$.

Proof. First note that by the first property of good kernels

$$(f * K_n)(x) - f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(y) \Big[f(x-y) - f(x) \Big] \mathrm{d}y \,. \tag{2.14}$$

Now if $\epsilon > 0$ and f is continuous at x, choose $\delta > 0$ so that $|y| < \delta$ implies $|f(x-y) - f(x)| < \epsilon$. Then

$$\begin{aligned} \Big| (f * K_n)(x) - f(x) \Big| &\leq \frac{1}{2\pi} \int_{|y|<\delta} |K_n(y)| \Big| f(x-y) - f(x) \Big| \mathrm{d}y \\ &\quad + \frac{1}{2\pi} \int_{\delta \leq |y| < 2\pi} |K_n(y)| \Big| f(x-y) - f(x) \Big| \mathrm{d}y \\ &\leq \frac{\epsilon M}{2\pi} + \frac{2B}{2\pi} \int_{\delta \leq |y| < 2\pi} |K_n(y)| \mathrm{d}y \end{aligned} \tag{2.15}$$

where $|f(x)| \leq B$ is a bound for f. Thus by the third property of good kernels we can choose N large enough, so that for all $n \geq N$

$$\Big| (f * K_n)(x) - f(x) \Big| \leq C\epsilon \tag{2.16}$$

for some constant $C > 0$ independent of n.

$\square$

2.3. Convergence of Fourier series

Returning to the question of convergence of $S_N(f)(x) \to f(x)$ one might be tempted to think that the Dirichlet kernel is a good kernel, but unfortunately *this is not the case*. In fact one can show that

$$\int_{-\pi}^{\pi} |D_N(x)| \mathrm{d}x \geq c \log N \qquad \text{as } N \to \infty \tag{2.17}$$

in violation of the second property. Nonetheless one immediately verfies that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(x) \mathrm{d}x = 1 \tag{2.18}$$

and so the first property is actually verfied. The reason that the integral of the absolute value is large, while the mean value of D_N is 1, is that $D_N(x)$ oscillates rapidly as N gets large. All this suggests that the *pointwise convergence* of Fourier series is a difficult question, which we will not pursue further here.

However, we note that

1. If f is continuously differentiable then the Fourier series converges to f uniformly.
2. If f is merely integrable, then $S_N(f) \to f$ in the *mean square sense*, namely

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| S_N(f)(\theta) - f(\theta) \right|^2 \mathrm{d}\theta \to 0 \qquad \text{as } N \to \infty \tag{2.19}$$

Supplementary Problems

1. Show that the Dirichlet kernel can be expressed in closed form as

$$D_N(x) = \frac{\sin((N + 1/2)x)}{\sin(x/2)} \tag{2.20}$$

Hint: Note that $D_N = \sum_{n=0}^{N} \omega^n + \sum_{n=-N}^{-1} \omega^n$ with $\omega = e^{ix}$ and sum up these geometric progressions separately.

Lecture 3.

Dirichlet's problem on the disc

Recommended Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 2, Section 5.4).

3.1. Dirichlet's problem on the unit disc

Recall that in Lecture 2 we argued that the solution to Dirichlet's problem on the unit disc, namely the solution to $\Delta u = 0$ in the unit disc with $u = f$ on the boundary, should be given by

$$u(r, \theta) = \sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{in\theta} \tag{3.1}$$

where a_n are the Fourier coefficients of f. We will now show that this expression can be written as the convolution of the Poisson kernel with f. Indeed, using the definition of the Fourier coefficients,

$$\begin{aligned} u(r, \theta) = & \sum_{n=-\infty}^{\infty} r^{|n|} \Big(\frac{1}{2\pi} \int_{-\infty}^{\infty} f(\varphi) e^{-in\varphi} \mathrm{d}\varphi \Big) e^{in\theta} \\ = & \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\varphi) \Big(\sum_{n=-\infty}^{\infty} r^{|n|} e^{-in(\varphi - \theta)} \Big) \mathrm{d}\varphi \\ = & \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\varphi) P_r(\theta - \varphi) \mathrm{d}\varphi \end{aligned} \tag{3.2}$$

where $P_r(\theta)$ is the **Poisson kernel**

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} \tag{3.3}$$

Remark 3.1. Since f is integrable, $|a_n|$ is uniformly bounded in n, so the series (3.1) converges absolutely and uniformly for each $0 \leq r < 1$, which also justifies the interchange of the sum and integral (3.2).

Lemma 3.1. *If* $0 \le r < 1$*, then*

$$P_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2} \tag{3.4}$$

and $\{P_r\}$ *is a family of* good kernels *on the circle as* $r \to 1$ *from below.*

Proof. We have

$$P_r(\theta) = \sum_{n=0}^{\infty} \omega^n + \sum_{n=1}^{\infty} \overline{\omega}^n \tag{3.5}$$

with $\omega = re^{i\theta}$ and both series converge absolutely. The first sum equals $1/(1-\omega)$, and the second $\overline{\omega}/(1-\overline{\omega})$, so

$$P_r(\theta) = \frac{1 - |\omega|^2}{|1-\omega|^2} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}\,. \tag{3.6}$$

Now if $1/2 \le r \le 1$ and $\delta \le |\theta| \le \pi$, then

$$1 - 2r\cos\theta + r^2 = (1-r)^2 + 2r(1 - \cos\theta) \ge c_\delta > 0 \tag{3.7}$$

where $c_\delta > 0$ is a constant that only depends on δ. Thus in the range $\delta \le |\theta| \le \pi$

$$|P_r(\theta)| \le \frac{1 - r^2}{c_\delta} \tag{3.8}$$

which implies the third property of good kernels:

$$\int_{\delta \le |\theta| \le \pi} |P_r(\theta)| \mathrm{d}\theta \to 0 \qquad \text{as } r \to 1\,. \tag{3.9}$$

Since $P_r(\theta) \ge 0$ we can infer the second from the first property, which we derive using the expression (3.3), and integrating term by term (which is justified by absolute convergence the series):

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(\theta) \mathrm{d}\theta = 1 \tag{3.10}$$

□

It now follows immediately from Theorem 2.2 that the series (3.1) has the correct values on the boundary at points of continuity, because in view of (3.2),

$$u(r, \theta) = (f * P_r)(\theta) \to f(\theta) \qquad \text{as } r \to 1 \tag{3.11}$$

at every point $\theta \in [-\pi, \pi]$ where f is continuous.

Theorem 3.2. *Let* f *be an integrable function on the circle. Then* u *defined on the unit disc by*

$$u(r, \theta) = (f * P_r)(\theta) \tag{3.12}$$

has the following properties:

1. *u is twice continuously differentiable on the unit disc and satisfies*

$$\Delta u = 0\,. \tag{3.13}$$

2. *If f is continuous at θ, then*

$$\lim_{r\to 1} u(r,\theta) = f(\theta)\,. \tag{3.14}$$

   If f is continuous everywhere, then this limit is uniform.

3. *If f is continuous, then the* unique *solution to the Dirichlet problem, which is twice differentiable on the unit disc and satisfies* (3.13)*, and for which the limit* (3.14) *is uniform, is given by* (3.12)*.*

Proof. Given an integrable function f, the Fourier coefficients a_m are defined and uniformly bounded, and in view of the calculation (3.2), the function $u(r,\theta) = (P_r * f)(\theta)$ can be expressed as the series (3.1). Since this series is absolutely convergent it can be differentiated term by term and $\Delta u = 0$ can be verified using the expression (2.17) for the Laplacian in polar coordinates. We have also seen that (3.14) follows from Theorem 2.2.

For the uniqueness statement note that given another solution u_1 with said properties, the difference $v = u - u_1$ solves $\Delta v = 0$ with vanishing boundary data. Then in polar coordinates

$$\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} + \frac{1}{r^2}\frac{\partial^2 v}{\partial \theta^2} = 0\,. \tag{3.15}$$

Moreover since $v(r,\theta)$ is 2π-periodic in θ for each $0 < r < 1$, $v(r,\cdot)$ has a Fourier series with coefficients $a_n(r)$. Multiplying the above equation by $e^{-in\theta}$ and integrating by parts

$$\frac{1}{2\pi}\frac{1}{r^2}\int_{-\pi}^{\pi} \frac{\partial^2 v}{\partial \theta^2} e^{-in\theta}\mathrm{d}\theta = -\frac{n^2}{r^2} a_n(r) \tag{3.16}$$

we obtain that $a_n(r)$ satisfies the differential equation satisfied by $F(r)$ in (2.18),

$$a_n''(r) + \frac{1}{r}a_n'(r) - \frac{n^2}{r^2}a_n(r) = 0\,. \tag{3.17}$$

We have seen (in the Problems for Lecture 2) that for $n > 0$ this has as its only bounded solution $a_n(r) = A_n r^n$. Since the limit of v is uniform,

$$A_n = \lim_{r\to 1} a_n(r) = \frac{1}{2\pi}\lim_{r\to 1}\int_{-\pi}^{\pi} v(r,\theta)e^{-in\theta}\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}(f(\theta)-f(\theta))\cdot e^{-in\theta}\mathrm{d}\theta = 0 \tag{3.18}$$

and we conclude that $a_n(r) = 0$ for $n > 0$. Similiarly for $n \leq 0$.

The uniqueness statement can thus be inferred from the following uniqueness result for Fourier series:[1] *If f is continuous on the circle and the Fourier coefficients $a_n = 0$ for all $n \in \mathbb{Z}$, then $f = 0$.*

□

[1] The proof of this result, which we will not discuss there, also uses a family of good kernels; cf. (Stein and Shakarchi, *Fourier analysis*, Chapter 2, Section 2).

3.2. Heat equation on the circle

We have now solved the steady state heat equation on the unit disc. Recall that this problem arose in the study of the heat equation $\partial_t u = \triangle u$ to which we now return. We could entertain for a moment the $d = 1$ dimensional case with periodic boundary conditions, we would describe the heat flow on the circle:

$$\partial_t u = \frac{\partial^2 u}{\partial \theta^2} \tag{3.19}$$

The unknown $u(\theta, t)$ is a 2π-periodic function in θ, and could describe the evolution of an initial temperature distribution $u(\theta, 0) = f(\theta)$. Multiplying the equation by $e^{-in\theta}$, and integrating by parts as above gives that

$$\partial_t a_n(t) = -n^2 a_n(t) \tag{3.20}$$

which has the unique solution $a_n(t) = a_n(0)e^{-n^2 t}$, where $a_n(0)$ ought to be the Fourier coefficients of f. This yields that

$$u(t, \theta) = \sum_n a_n(t) e^{in\theta} = (f * H_t)(\theta) \tag{3.21}$$

where H_t is the **heat kernel on the circle**

$$H_t(\theta) = \sum_{n=-\infty}^{\infty} e^{-n^2 t} e^{in\theta} \,. \tag{3.22}$$

Remark 3.2. There are several analogies to the Poisson kernel: For example if we set $r = e^{-\tau}$ in the Poisson kernel, for $0 < r < 1$ with some $\tau > 0$, then $P_r(\theta)$ takes the form

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} e^{-|n|\tau} e^{in\theta} \,. \tag{3.23}$$

The heat kernel is also *positive*, and a *good kernel.* (Neither of these properties are obvious from the the above expression.) However, unlike for the Poisson kernel, there does not appear to be an elementary closed expression for the heat kernel.

We will pursue a similar strategy to derive a formula for the time-dependent heat equation on the line, and in higher dimension, but to do that we will first have to recall the basic properties of the Fourier transform.

Problems

1. We have seen that

$$\lim_{r\to 1} (P_r * f)(\theta) = f(\theta) \tag{3.24}$$

at points θ where f is continuous. In this problem we study the behaviour at certain points of discontinuity. An integrable function has a **jump discontinuity** at θ if the two limits

$$\lim_{\substack{h\to 0\\ h>0}} f(\theta+h) = f(\theta^+)\,, \qquad \lim_{\substack{h\to 0\\ h>0}} f(\theta-h) = f(\theta^-)\,, \tag{3.25}$$

exist. Prove that if f has a jump discontinuity at θ, then

$$\lim_{r\to 1}(P_r * f)(\theta) = \frac{f(\theta^+)+f(\theta^-)}{2} \tag{3.26}$$

Hint: Explain why

$$\frac{1}{2\pi}\int_{-\pi}^{0} P_r(\theta)\mathrm{d}\theta = \frac{1}{2\pi}\int_{0}^{\pi} P_r(\theta)\mathrm{d}\theta = \frac{1}{2} \tag{3.27}$$

and then revisit the proof of (3.24).

Part II.

Quasi-linear equations in two variables

Lecture 4.

First order equations: Cauchy problem and characteristics

Further Reading

(John, *Partial differential equations*, Chapter 1).

We begin with the simple equation

$$\partial_x u + c\partial_y u = 0 \tag{4.1}$$

for an unknown function $u(x, y)$, where c is a constant.

We can view solutions of (4.1) as graphs in the plane whose level sets are straight lines:

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t, \xi + ct) = 0\,. \tag{4.2}$$

Hence the value of $u(t, x)$ only depends on ξ, which parametrizes the lines $y = \xi + cx$ on which u is constant. The general form of the solution to (4.1) is thus:

$$u(x, y) = f(\xi) = f(y - cx)\,. \tag{4.3}$$

Since $f(y) = u(0, y)$ the solution is clearly determined everywhere from its values at $x = 0$. For fixed x, the graph of $y \mapsto u(x, y)$ is obtained by translating the graph of f, without changing shape.

The idea to understand solutions of a partial differential equation by solving *systems of ordinary differential equations* is productive in the context of **general first-order equations** for a function $u(x, y, \cdots)$,

$$F(x, y, \ldots, u, u_x, u_y, \ldots) = 0\,. \tag{4.4}$$

We will explore this approach in this lecture is the special case of *quasi-linear first-order equations*, namely equations which are linear in the derivatives.

4.1. Quasi-linear equations

Consider the equation

$$a(x, y, u)\partial_x u + b(x, y, u)\partial_y u = c(x, y, u)\,. \tag{4.5}$$

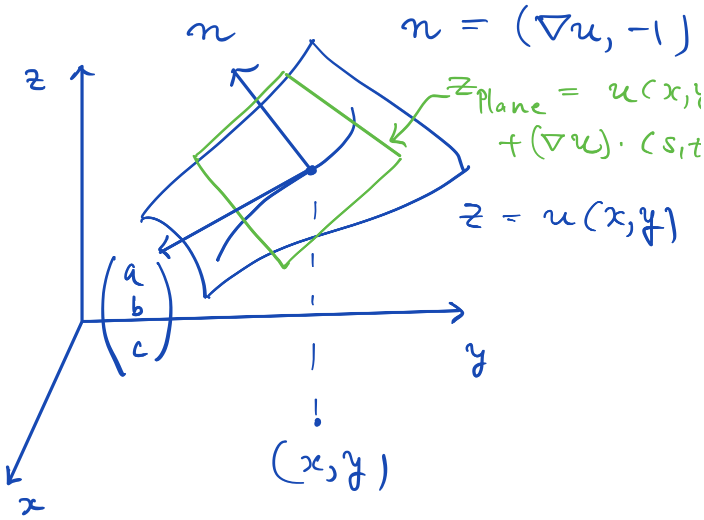

Figure 4.1.: The *characteristic direction* (a, b, c) is contained in the tangent plane to the surface.

The graph of functions $z = u(x, y)$ are surfaces in $\mathbb{R}^3$; solutions of (4.5) are also called **integral surfaces**.

Geometrically, the equation (4.5) states that at every point on the graph of u, the vector $V = (a, b, c)$ is orthogonal to the normal $n = (\partial_x u, \partial_y u, -1)$; cf. Fig 4.1:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \begin{pmatrix} \partial_x u \\ \partial_y u \\ -1 \end{pmatrix} = v \cdot n = 0 \tag{4.6}$$

The curve $\gamma(t) = (x(t), y(t), z(t))$ which is defined as the solution to the ODE

$$\dot{\gamma}(t) = V(\gamma(t)) \,, \tag{4.7}$$

thus has the property that its tangent vector $\dot{\gamma}(t)$ is tangent to the integral surface at every point. In components,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a(x, y, z) \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = b(x, y, z) \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = c(x, y, z) \,. \tag{4.8}$$

Assuming that the functions a, b, c are continuously differentiable we know from the basic theory of ODEs that through each point passes exactly one curve. These curves are also called **characteristics**.

Proposition 4.1. *Let $p = (x_0, y_0, z_0)$ be a point on the integral surface $S = \{(x, y, z) : z = u(x, y)\}$. Let $\gamma : I \mapsto \mathbb{R}^3$ be a characteristic curve through $p = \gamma(0)$. Then*

$$\gamma(t) \in S \quad (t \in I) \,.$$

Proof. Let $\gamma(t) = (x(t), y(t), z(t))$ be the solution to the system of differential equations (4.8) with initial conditions $\gamma(0) = p$. We compute that

$$U(t) = z(t) - u(x(t), y(t)) \tag{4.9}$$

satisfies

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}U(t) =& c(x, y, z) - a(x, y, z)\partial_x u(x, y) - b(x, y, z)\partial_y u(x, y) \\ =& c(x, y, U + u(x, y)) \\ & - a(x, y, U + u(x, y))\partial_x u(x, y) - b(x, y, U + u(x, y))\partial_y u(x, y)\end{aligned} \tag{4.10}$$

This is a ordinary differential equation for U, with initial condition $U(0) = 0$. Moreover $U(t) = 0$ is a solution because u satisfies (4.5). By uniqueness of solutions this is the only solution, and hence $z(t) = u(x(t), y(t))$, which says that $\gamma(t)$ is contained in S. □

4.2. Example of an initial value problem

Consider the quasi-linear equation

$$\partial_t u + u\partial_x u = 0 \,. \tag{4.11}$$

Interpretation. We can view (4.11) as the equation for a velocity field u of a continuum of particles on the line which are not accelerated: If $x(t)$ is a solution to

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = u(t, x(t)) \tag{4.12}$$

where u satisfies the equation (4.11), then

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}(t) = \partial_t u(t, x(t)) + u(t, x(t))\partial_x u(t, x(t)) = 0 \,. \tag{4.13}$$

Here (t, x) are coordinates, and solutions $z = u(t, x)$ are graphs over the tx-plane. We will see that the level sets of this graph are the trajectories of the particles, which are in turn the projections of the characteristics to the tx plane.

The characteristic equations here are

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1 \qquad \frac{\mathrm{d}x}{\mathrm{d}s} = z \qquad \frac{\mathrm{d}z}{\mathrm{d}s} = 0 \tag{4.14}$$

Here we can prescribe **initial data**, at

$$t = 0 : \qquad z = h(x) \,, \tag{4.15}$$

and the function $h(x)$ should uniquely determine the corresponding solution $u(t, x)$, so that $u(0, x) = h(x)$; see Fig. 4.2.

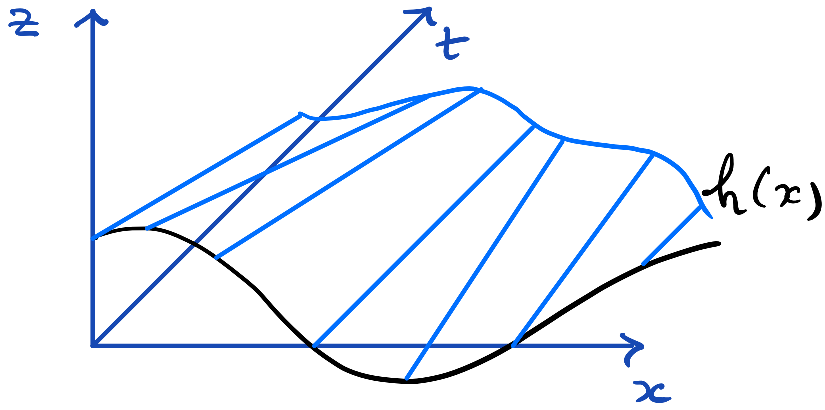

Figure 4.2.: Initial value problem for (4.11).

We can find the characteristic curves emanating from the points $(0, y, h(y))$ by solving the equations with the initial conditions $t(0) = 0$, $x(0) = y$, $z(0) = h(y)$. Hence

$$t = s \qquad z = h(y) \qquad x(s) = y + h(y)s\,. \tag{4.16}$$

So the projection of the characteristic curves to the tx-plane is

$$x(t) = y + h(y)t \tag{4.17}$$

along which u has the constant value $u = h(y)$. They are precisely the unaccelerated particle trajectories defined in (4.12).

We can write the solution u in *implicit* form, by finding the value of y for a given point (t, x):

$$z = u(t, x) = h(y) = h(x - u(t, x)t)\,. \tag{4.18}$$

Note the similarity of the solution to (4.3).

Blow-up. Consider two characteristics

$$\gamma_{y_1}(t) = (t, x_1(t), h(y_1))\,, \quad \gamma_{y_2}(t) = (t, x_2(t), h(y_2))\,, \tag{4.19}$$

through the points $(0, y_1, h(y_1))$, and $(0, y_2, h(y_2)$, respectively. In the projection to the tx-plane, these lines intersect if for some $t_0 > 0$,

$$y_1 + h(y_1)t_0 = y_2 + h(y_2)t_0\,, \tag{4.20}$$

and that time is given by

$$t_0 = \frac{y_1 - y_2}{h(y_2) - h(y_1)} \,. \tag{4.21}$$

If h is decreasing on an interval that contains (y_1, y_2), $y_1 < y_2$, then such a time $t_0 > 0$ always exists, and the solution u cannot be defined at the point (t_0, x_0), $x_0 = x_1(t_0) = x_2(t_0)$, because it would need to take both the values $h(y_1)$ and $h(y_2)$. A *global in time* solution can only be defined in the case that the data h is a non-decreasing function.

We want to understand in which sense the solution becomes singular. Let h have compact support, and consider consider the behaviour of $\partial_x u$ along a characteristic γ_y. First it follows from (4.18) that

$$\partial_x u(t,x) = h'(x - u(t,x)t)(1 - \partial_x u(t,x)t) \tag{4.22}$$

and thus along γ_y:

$$\partial_x u(\gamma_y(t)) = \frac{h'(y)}{1 + h'(y)t} \tag{4.23}$$

Hence for points y with $h'(y) < 0$, we find that

$$\lim_{t \to T} \partial_x u(\gamma_y(t)) = \infty \qquad T(y) = -\frac{1}{h'(y)} \,. \tag{4.24}$$

The smallest time $T = \min_y T(y)$ for which this happens is where $|h'(y)|$ has a maximum. Therefore there cannot exist a continuously differentiable solution beyond the time T.

4.3. Cauchy problem

For the general first order quasi-linear equation (4.5) in two variables there is of course no preferred "time-variable" but the idea of generating solutions using families of characteristic curves persists, and leads to the Cauchy problem.

Let Γ be a differentiable curve in $\mathbb{R}^3$,

$$\Gamma : s \mapsto (f(s), g(s), h(s)) \,. \tag{4.25}$$

We want to find a solution u so that $h(s) = u(f(s), g(s))$; cf. Fig 4.3. Let $P = (f(s_0), g(s_0), h(s_0))$, and for each s let $(X(s,t), Y(s,t), Z(s,t))$ be the solution to the system of ordinary differential equations

$$\frac{\mathrm{d}X}{\mathrm{d}t} = a(X,Y,Z) \qquad X(s,0) = f(s) \tag{4.26a}$$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = b(X,Y,Z) \qquad Y(s,0) = g(s) \tag{4.26b}$$

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = c(X,Y,Z) \qquad Z(s,0) = h(s) \tag{4.26c}$$

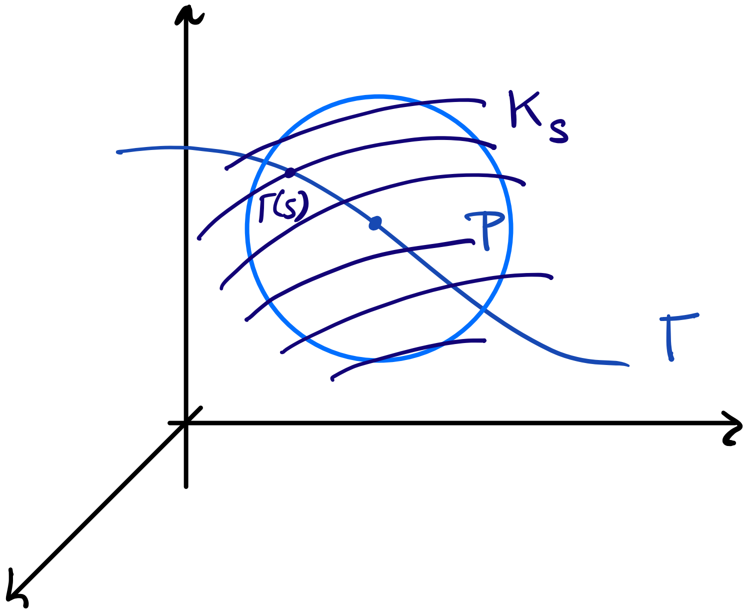

Figure 4.3.: Cauchy problem for (4.5).

Then $K_s : t \mapsto (X(s,t), Y(s,t), Z(s,t))$ is a family of characteristic curves, and for $a, b, c \in \mathrm{C}^1$ the map $(s,t) \mapsto K_s(t)$ is continuously differentiable in a neighbourhood of $(s_0, 0)$, by the local existence theory for ODEs. It remains to show that the surface

$$\Sigma = \{(x,y,z) : x = X(s,t), y = Y(s,t), z = Z(s,t), |s - s_0| + t < \epsilon\} \tag{4.27}$$

can be written as the graph $z = u(x,y)$ of a function, which is then the solution to (4.5) with the prescribed *data* on Γ. This is possible if we can invert the map

$$(s,t) \mapsto (X(s,t), Y(s,t)) \tag{4.28}$$

in a neighborhood of $(s_0, 0)$, for then we can write

$$z = u(x,y) = Z(S(x,y), T(x,y))\,. \tag{4.29}$$

By the implicit function theorem the inverse $(x,y) \mapsto (S(x,y), T(x,y))$ exists if

$$\det \begin{pmatrix} \frac{\partial X}{\partial s} & \frac{\partial Y}{\partial s} \\ \frac{\partial X}{\partial t} & \frac{\partial Y}{\partial t} \end{pmatrix} (s_0, 0) \neq 0 \tag{4.30}$$

which is a condition that reduces to

$$f'(s_0) b(P) - g'(s_0) a(P) \neq 0\,. \tag{4.31}$$

The following observation helps us understand this condition geometrically.

Lemma 4.2. *If* (4.31) *fails, then* $\dot{\Gamma}(s_0)$ *is colinear to* $V(P) = (a(P), b(P), c(P))$.

In other words, the condition (4.30) is satisfied as long as Γ is chosen to be a *non*-characteristic curve.

Proof. If for a solution u *this condition fails* then at P,

$$f'(s_0)b(P) - g'(s_0)a(P) = 0 \,. \tag{4.32}$$

while along Γ,

$$h'(s_0) = f'(s_0)u_x + g'(s_0)u_y \,, \quad c(P) = a(P)u_x + b(P)u_y \,, \tag{4.33}$$

where u_x and u_y are evaluated at $(f(s_0), g(s_0))$,

To show that the vectors $(a, b, c)(P)$ and $(f', g', h')(s_0)$ are colinear, we can verify that the cross product vanishes. The equation (4.32) states that the third component in the cross product vanishes. For the other components, note that (4.33) implies that at P,

$$bh' - cg' = 0 \,, \qquad ah' - cf' = 0 \,. \tag{4.34}$$

□

Problems

1. Solve the following initial value problems.
   a) $u_x + u_y = u^2 \qquad u(x, 0) = h(x)\,.$
   b) $u_y = xuu_x \qquad u(x, 0) = x$
2. Show that the solution u of the quasi-linear PDE

$$u_y + a(u)u_x = 0 \tag{4.35}$$

   with initial condition $u(x, 0) = h(x)$ is given implicitly by

$$u = h(x - a(u)y) \,. \tag{4.36}$$

   Show that the solution becomes singular for some positive y, unless $a(h(s))$ is a nondecreasing function of s.

Lecture 5.

Second order equations: three types of equations

Further Reading

(John, *Partial differential equations*, Chapter 2).

5.1. Characteristics for quasi-linear second order equations

Consider the quasi-linear second-order PDE in two variables,

$$a\,u_{xx} + 2b\,u_{xy} + c\,u_{yy} = d \tag{5.1}$$

where the coefficients a, b, c, d are allowed to depend on x, y, u, u_x, u_y.

As in the case of first-order equations, we would like to understand the Cauchy problem, namely the problem of constructing solutions from a suitable prescription of "data".

Given a curve $\gamma(t) = (f(t), g(t))$ we may first prescribe the values of u along the curve:

$$u(\gamma(t)) = h(t)\,. \tag{5.2}$$

Then also the derivative *along* the curve is fixed:

$$\frac{\mathrm{d}}{\mathrm{d}t} u \circ \gamma(t) = h'(t) = u_x(\gamma(t)) f'(t) + u_y(\gamma(t)) g'(t) \tag{5.3}$$

In particular, the three functions

$$u(\gamma(t)) = h(t) \qquad u_x(\gamma(t)) = \phi(t) \qquad u_y(\gamma(t)) = \psi(t)\,. \tag{5.4}$$

are not independent. Geometrically, it is clear that the *normal derivative* is independent:

$$\frac{\partial u}{\partial n}(\gamma(t)) = \nabla u(\gamma(t)) \cdot n(t) = k(t) \qquad n(t) = (-g'(t), f(t))\,. \tag{5.5}$$

In other words, the functions $h(t)$, and $k(t)$ may be freely prescribed. Then both partial derivatives $u_x(\gamma(t)) = \phi(t)$, and $u_y(\gamma(t)) = \psi(t)$ can be computed along the curve γ.

With the values of u along γ, and the first partial derivatives u_x and u_y prescribed along γ from data, we want to understand how the equation determines the second derivatives.

Since u_x and u_y is now prescribed along γ, we can derive the following conditions:

$$\frac{\mathrm{d}}{\mathrm{d}t}u_x \circ \gamma(t) = u_{xx}(\gamma(t))f'(t) + u_{xy}(\gamma(t))g'(t) = \phi'(t) \tag{5.6}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}u_y \circ \gamma(t) = u_{yx}(\gamma(t))f'(t) + u_{yy}(\gamma(t))g'(t) = \psi'(t)\,. \tag{5.7}$$

Together with the equation (5.1), we then have three equations for the three unknowns u_{xx}, u_{xy}, and u_{yy}, which we can write as the system of equations:

$$\begin{pmatrix} f' & g' & 0 \\ 0 & f' & g' \\ a & 2b & c \end{pmatrix} \begin{pmatrix} u_{xx} \\ u_{xy} \\ u_{yy} \end{pmatrix} = \begin{pmatrix} \phi'(t) \\ \psi'(t) \\ d(\gamma(t), h(t), \phi(t), \psi(t)) \end{pmatrix} \tag{5.8}$$

Therefore all second derivatives of u along γ are uniquely determined by h, ϕ, ψ, unless

$$\Delta = \det \begin{pmatrix} f' & g' & 0 \\ 0 & f' & g' \\ a & 2b & c \end{pmatrix} = 0\,. \tag{5.9}$$

We say the curve γ is **characteristic** if $\Delta = 0$, and **non-characteristic** if $\Delta \neq 0$ along γ. Note that

$$\Delta(t) = a\,(g'(t))^2 - 2b\,f'(t)g'(t) + c\,(f'(t))^2 \tag{5.10}$$

where $a, b, c = a, b, c(\gamma(t), h(t), \phi(t), \psi(t))$, so Δ on γ depends on the equation (5.1) *and the Cauchy data*.

In fact, let us separate the dependence on the choice of the curve γ, from the structure of the equation. We can view (5.10) as the value of a quadratic form, or metric, on directions tangential to the curve γ. For this purpose let the metric g be defined by

$$g = c\ \mathrm{d}x \otimes \mathrm{d}x - b\ \mathrm{d}x \otimes \mathrm{d}y - b\ \mathrm{d}y \otimes \mathrm{d}x + a\ \mathrm{d}y \otimes \mathrm{d}y\,. \tag{5.11}$$

If we rename the coordinates to $x^1 = x$, $x^2 = y$, we can also express this as

$$g = \sum_{i,j=1}^{2} g_{ij}\mathrm{d}x^i\mathrm{d}x^j\,, \qquad (g_{ij}) = \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \tag{5.12}$$

and as a quadratic from

$$g(V, V) = \sum_{i,j=1}^{2} g_{ij}V^iV^j\,. \tag{5.13}$$

Therefore,

$$\Delta(t) = g(\dot\gamma(t), \dot\gamma(t))\,, \qquad \dot\gamma(t) = \begin{pmatrix} f'(t) \\ g'(t) \end{pmatrix}. \tag{5.14}$$

Remark 5.1. *Along a non-characteristic curve the Cauchy data determines uniquely all second order partial derivatives.* In fact, all higher order derivatives are determined in that case; for differentiating (5.1) in x, and applying (5.3) to u_{xx} and u_{xy}, gives a system of three equations which again can be solved for u_{xxx}, u_{xyx}, and u_{yyx} along γ, provided $\Delta \neq 0$ and u and all its derivatives up to second order are known along γ. This suggests that, at least locally, one should be able to obtain a solution to the Cauchy problem by writing down a power series; this is made precise for *analytic data* by the **Cauchy-Kovalewski** theorem.

The determinant $\Delta(t)$ thus may only vanish if the quadratic form associated to g has a non-trivial null space.

Definition 5.1. The *characterisitic set* associated to a solution u of (5.1) at a point $p = (x, y, u, u_x, u_y)$ is given by

$$C_p = \Big\{V \in \mathbb{R}^2 : g_p(V, V) = 0, V \neq 0\Big\} \subset \mathrm{T}_p\mathbb{R}^2 . \tag{5.15}$$

It clearly depends on the coefficient functions in (5.1), and the solution u, as to whether this set is non-empty. More explicitly, for $V = (x, y)$, we have

$$g(V, V) = cx^2 - 2bxy + ay^2 , \tag{5.16}$$

and so the relevant quantity here is the discrimenant of the quadratic:

$$\det(g) = ac - b^2 . \tag{5.17}$$

Definition 5.2. We say (5.1) is **elliptic** if $\det(g) > 0$, **parabolic** if $\det(g) = 0$, and **hyperbolic** if $\det(g) < 0$.

Note that for a quasi-linear equation this notion depends both on the point (x, y) *and* the solution (u, u_x, u_y) at (x, y).

In the *linear* case, when a, b, and c only depend on (x, y), we can view $\Delta = 0$ as an ordinary differential equation for γ.

Exercise 5.1. Show that C_p is empty in the elliptic case, that C_p is a line in the parabolic case, and a cone in the hyperbolic case.

Definition 5.3. A *characteristic curve* associated to a solution u of (5.1) is a curve $\gamma(t) \in \mathbb{R}^2$ whose tangent vector is characteristic at every point on the curve,

$$g_{p(t)}(\dot\gamma(t), \dot\gamma(t)) = 0 , \quad p(t) = (\gamma(t), u(\gamma(t)), u_x(\gamma(t)), u_y(\gamma(t))) . \tag{5.18}$$

Exercise 5.2. Write (5.18) as an ODE for $\gamma(t) = (f(t), g(t))$. Suppose $f'(t) \neq 0$ in a neighborhood of a point $\gamma(t_0)$, and suppose (5.1) is either hypbolic, or parabolic in that neighborhood. Reparametrize γ so that $f'(t) = 1$, and verify that a *characteristic* curve $\gamma(t) = (t, g(t))$ is obtained by solving an ODE for g:

$$y' = \frac{b \pm \sqrt{b^2 - ac}}{a} \tag{5.19}$$

where a, b, c are functions of $(x, y, u(x, y), u_x(x, y), u_y(x, y))$.

5.2. Linear second order equations

Consider the *linear* second-order equation

$$au_{xx} + 2bu_{xy} + cu_{yy} + 2du_x + 2eu_y + fu = 0 \tag{5.20}$$

with coefficients a, b, c, d, e, f depending only on (x, y).

We have seen that associated to (5.20) is the metric g given by (5.12). When (5.12) is elliptic, or hyperbolic, then $\det g \neq 0$, and the components of the *inverse* of g are

$$g^{-1} = (g^{-1})^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}, \qquad \Big((g^{-1})^{ij}\Big) = \frac{1}{\det(g)} \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \tag{5.21}$$

and we can write (5.20) in the form

$$(g^{-1})^{ij} \frac{\partial^2 u}{\partial x^i \partial x^j} = F(u, \nabla u), \tag{5.22}$$

where F is linear in u, and the derivatives $\partial_{x^i} u, i = 1, 2$. Here $(x^1 = x, x^2 = y)$.

We now want to show that under a suitable change of variables

$$(x, y) \mapsto (\xi, \eta) \tag{5.23}$$

the equation can be brought into a certain **normal form**:

There exists a change of variables, so that (5.20) takes the form

Elliptic case:

$$\partial^2_{\xi\xi} u + \partial^2_{\eta\eta} u = G(u, \partial_\xi u, \partial_\eta u), \tag{5.24}$$

in other words, the principal part becomes the Laplacian in $\mathbb{R}^2$,

Hyperbolic case:

$$\partial^2_{\xi\eta} u = G(u, \partial_\xi u, \partial_\eta u), \tag{5.25}$$

in other words, the principal part becomes the d'Alembertian in $\mathbb{R}^{1+1}$,

where G is linear in u and its first derivatives.

In general, for a locally invertible transformation,

$$y^a = f^a(x^1, x^2), \quad a = 1, 2, \tag{5.26}$$

we compute

$$\frac{\partial}{\partial x^i} = \sum_{a=1}^{2} \frac{\partial f^a}{\partial x^i} \frac{\partial}{\partial y^a}, \tag{5.27}$$

and hence in the new coordinates

$$g^{-1} = (g^{-1})^{ij} \frac{\partial f^a}{\partial x^i} \frac{\partial f^b}{\partial x^j} \frac{\partial}{\partial y^a} \otimes \frac{\partial}{\partial y^b}, \tag{5.28}$$

the components are

$$(g^{-1})^{ab} = \sum_{i,j=1}^{2} (g^{-1})^{ij} \frac{\partial f^a}{\partial x^i} \frac{\partial f^b}{\partial x^j} \,. \tag{5.29}$$

In the new coordinates the equation reads

$$(g^{-1})^{ab} \frac{\partial^2 u}{\partial y^a \partial y^b} = G(u, \nabla u) \,, \tag{5.30}$$

for some function G linear in u and ∇u.

Remark 5.2. The differential operator $(g^{-1})^{ij}\partial^2_{ij}$ is *not* coordinate invariant, but it is the principle part of the **Laplace-Beltrami** operator associated to g,

$$\triangle_g u = \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^i} \Big(g^{ij} \sqrt{\det(g)} \frac{\partial}{\partial x^j} u \Big) \tag{5.31}$$

which can be expressed in any coordinates. The equation (5.30) is a statement about the *principal part* of the equation, while all first order terms are absorbed in the right hand side.

Exercise 5.3. Verify, that in different notation, this shows that for a locally invertible coordinate transformation

$$\xi = \phi(x, y) \qquad \eta = \psi(x, y) \,, \tag{5.32}$$

the equation (5.20) in the new coordinates (ξ, η) takes the form

$$A(\xi, \eta)\partial^2_{\xi\xi} u + 2B(\xi, \eta)\partial_{\xi\eta} u + C(\xi, \eta)\partial^2_{\eta\eta} u = F \tag{5.33}$$

where

$$A(\xi, \eta) = a\phi_x^2 + 2b\phi_x\phi_y + c\phi_y^2 \tag{5.34}$$

$$B(\xi, \eta) = a\phi_x\psi_x + b(\phi_x\psi_y + \psi_y\phi_x) + c\phi_y\psi_y \tag{5.35}$$

$$C(\xi, \eta) = a\psi_x^2 + 2b\psi_x\psi_y + c\psi_y^2 \,. \tag{5.36}$$

and F is linear in the first derivatives $\partial_\xi u$, and $\partial_\eta u$.

Elliptic case. In this case want to achieve $B = 0$, which in view of (5.35) amounts to requiring that the transformations ϕ and ψ solve

$$\left\langle \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \phi_x \\ \phi_y \end{pmatrix}, \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} \right\rangle = 0 \tag{5.37}$$

which we can achieve by choosing $\nabla\phi$ to be orthogonal to the image of $\nabla\psi$ with the adjoint matrix, so for example

$$\frac{\partial \phi}{\partial x^i} = \epsilon_{ij} (g^{-1})^{jk} \frac{\partial \psi}{\partial x^k} \tag{5.38}$$

where $\epsilon_{12} = \sqrt{\det g}$ and ϵ_{ij} is antisymmetric. The integrability condition $\partial^2_{xy}\phi = \partial^2_{yx}\phi$ then reads

$$\sum_{i,j=1}^{2} \frac{\partial}{\partial x^i}\Big(g^{ij}\sqrt{\det(g)}\frac{\partial}{\partial x^j}\phi\Big) = 0\,. \tag{5.39}$$

which means that the equation takes normal form provided the transformations satisfy the **Beltrami equation**:

$$\triangle_g \phi = 0\,, \qquad \triangle_g \psi = 0\,. \tag{5.40}$$

Hyperbolic case. Here the aim is to achieve that $A = 0$ and $C = 0$. Since the diagonal components of $(g^{-1})^{aa}$, $a = 1, 2$, are of the form

$$\sum_{i,j=1}^{2} (g^{-1})^{ij}\frac{\partial f}{\partial x^i}\frac{\partial f}{\partial x^j} \qquad \text{for } f = f^a, a = 1, 2\,, \tag{5.41}$$

this can be arranged provided $f^a : a = 1, 2$ solve the **eikonal equation**

$$(g^{-1})^{ij}\frac{\partial \phi}{\partial x^i}\frac{\partial \phi}{\partial x^j} = 0\,. \tag{5.42}$$

Solutions are also called **optical functions**, and have following properties:

We will show next that the level sets of solutions to the eikonal equation are characteristic curves. It is thus natural to call the the variables $\xi = \phi(x, y)$, and $\eta = \psi(x, y)$ for which $A = C = 0$ *characteristic coordinates.*

The vector V with components

$$V^i = \sum_j (g^{-1})^{ij}\frac{\partial \phi}{\partial x^j} \tag{5.43}$$

is *tangential* to the level sets of ϕ, because

$$V \cdot \nabla\phi = \sum_i V^i \partial_i \phi = 0\,. \tag{5.44}$$

Moreover V is *characteristic*:

$$g(V, V) = g_{ij}V^iV^j = g_{ij}g^{im}\frac{\partial \phi}{\partial x^m}g^{jn}\frac{\partial \phi}{\partial x^n} = \frac{\partial \phi}{\partial x^j}g^{jn}\frac{\partial \phi}{\partial x^n} = 0\,, \tag{5.45}$$

(summed over repeated indices).

Therefore solutions to the eikonal equation can be *generated* from characteristic curves, namely by solving the ODEs (5.19).

In summary, we can define the functions (ξ, η) in such a way that the coordinate lines $\xi = \text{const.}$ and $\eta = \text{const.}$ are characteristic curves. Then (5.20) in the coordinates takes the normal form (5.25).

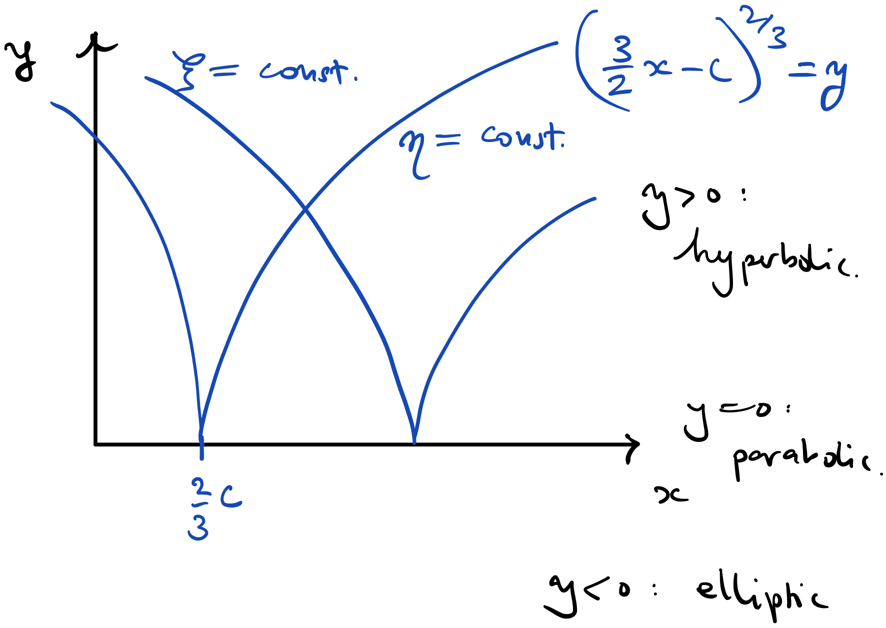

Figure 5.1.: Characteristics of the Tricomi equation.

5.3. Tricomi equation

As an example of a linear equation which changes its type consider the **tricomi equation**

$$u_{yy} - y u_{xx} = 0 \,. \tag{5.46}$$

Here

$$\det(g) = -y \,, \tag{5.47}$$

hence this equation is elliptic for $y < 0$, hyperbolic for $y > 0$, and parabolic for $y = 0$. Moreover, in the upper half-plane $y > 0$ the ODEs (5.19) for the characteristic curves $\gamma(x) = (x, y(x))$ are

$$y' = \pm y^{-1/2} \,. \tag{5.48}$$

We can integrate this equation and find that the two families of characteristic curves are given by

$$y^{3/2} = \pm \frac{3}{2} x + C \,; \tag{5.49}$$

cf. Fig. 5.1. Hence we can choose as coordinate transformations

$$\phi(x, y) = 2y^{3/2} + 3x \qquad \psi(x, y) = 2y^{3/2} - 3x \tag{5.50}$$

and, after computing the first order terms, the equation takes the normal form

$$3(\xi+\eta)\partial^2_{\xi\eta}u + 2(\partial_\xi u - \partial_\eta u) = 0\,. \tag{5.51}$$

Exercise 5.4. Verify that the functions ϕ and ψ satisfy the eikonal equation (5.42).

Problems

1. Revisit the one-dimensional wave equation from Lecture 2 and verify that the coordinates ξ and η introduced there are characteristic in the sense of this Lecture.

2. Find by power series expansion in t the solution to the initial value problem

$$u_{tt} - u_{xx} = u \tag{5.52}$$

$$u(x,0) = e^x \qquad u_t(x,0) = 0\,. \tag{5.53}$$

3. Let u be the solution to the quasi-linear equation

$$a(u_x,u_y)u_{xx} + 2b(u_x,u_y)u_{xy} + c(u_x,u_y)u_{yy} = 0\,. \tag{5.54}$$

   Introduce new independent variables ξ, η and a new unknown function ϕ by

$$\xi = u_x(x,y) \qquad \eta = u_y(x,y) \qquad \phi = xu_x + yu_y - u\,. \tag{5.55}$$

   Prove that ϕ as a function of ξ, and η, satisfies $x = \phi_\xi$, and $y = \phi_\eta$, and the *linear* partial differential equation

$$a(\xi,\eta)\phi_{\eta\eta} - 2b(\xi,\eta)\phi_{\xi\eta} + c(\xi,\eta)\phi_{\xi\xi} = 0\,. \tag{5.56}$$

Part III.

Linear PDE and Fourier Analysis

Review: The Fourier Transform on $\mathbb{R}$

Further Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 5, Section 1).

In this lecture we recall several properties of the **Fourier transform** $\hat{f}$ of a function f on the real line $\mathbb{R}$, defined by

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\xi}\mathrm{d}x \qquad (\xi \in \mathbb{R}) \tag{6.1}$$

In the first place, one may define this for functions of *moderate decrease*, namely continuous functions satisfying the bound

$$|f(x)| \leq \frac{A}{1+x^2} \tag{6.2}$$

for some constant $A > 0$, which ensures that the limit

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \lim_{N\to\infty} \int_{-N}^{N} f(x)\mathrm{d}x \tag{6.3}$$

exists. However, while the Fourier transform of a function of moderate decrease is clearly bounded, it is not clear if $\hat{f}$ itself enjoys any specific decay property. To remedy this, one considers instead the space of **rapidly decreasing** functions.

6.1. Schwartz space

The *Schwartz space* $\mathcal{S}(\mathbb{R})$ on $\mathbb{R}$ consists of the set of all indefinitely differentiable functions f so that f and *all* its derivatives f', f'', ..., are *rapidly decreasing*, in the sense that

$$\sup_{x\in\mathbb{R}} |x|^k |f^{(l)}(x)| < \infty \qquad (k, l \in \mathbb{N})\,. \tag{6.4}$$

The space $\mathcal{S}(\mathbb{R})$ is a vectorspace over $\mathbb{C}$, closed under differentiation and multiplication by polynomials.

Exercise 6.1. Show that if $f \in \mathcal{S}(\mathbb{R})$, then $f' \in \mathcal{S}(\mathbb{R})$, and $xf(x) \in \mathcal{S}(\mathbb{R})$.

Example 6.1. An example of a function in Schwartz space is the *Gaussian* defined by

$$f(x) = e^{-x^2}\,. \tag{6.5}$$

The **Fourier transform** of a function $f \in \mathcal{S}(\mathbb{R})$ is defined by (6.1). We also use the notation $f(x) \to \hat{f}(\xi)$ to mean that the function $f(x)$ is transformed to $\hat{f}(\xi)$ under the Fourier transform.

Proposition 6.1. *If $f \in \mathcal{S}(\mathbb{R})$, then*

1. $f(x+h) \longrightarrow \hat{f}(\xi)e^{2\pi ih\xi}$ *whenever* $h \in \mathbb{R}$
2. $f(x)e^{-2\pi ixh} \longrightarrow \hat{f}(\xi+h)$ *whenever* $h \in \mathbb{R}$
3. $f(\delta x) \longrightarrow \delta^{-1}\hat{f}(\delta^{-1}\xi)$ *whenever* $\delta > 0$.
4. $f'(x) \to 2\pi i\xi \hat{f}(\xi)$
5. $-2\pi ixf(x) \longrightarrow \frac{\mathrm{d}}{\mathrm{d}\xi}\hat{f}(\xi)$

In particular, up to factors of $2\pi i$, the Fourier transform interchanges differentiation and multiplication by x. This is the key property that makes the Fourier transform a central object in the theory of (partial) differential equations.

The main motivation for the Schwartz space is that the Fourier transform is a bijection on the Schwartz space. First observe however, that if $f \in \mathcal{S}(\mathbb{R})$ then $\hat{f}$ is bounded, and moreover $\xi^k \partial_\xi^l \hat{f}(\xi)$ is bounded, because by the above Proposition, this is the Fourier transform of

$$\frac{1}{(2\pi i)^k}\Big(\frac{\mathrm{d}}{\mathrm{d}x}\Big)^k\Big[(-2\pi ix)^l f(x)\Big]\,. \tag{6.6}$$

We have shown:

Theorem 6.2. *If $f \in \mathcal{S}(\mathbb{R})$, then $\hat{f} \in \mathcal{S}(\mathbb{R})$.*

6.2. Fourier Inversion

The next major result is known as the **Fourier inversion** theorem:

Theorem 6.3. *If $f \in \mathcal{S}(\mathbb{R})$, then*

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi ix\xi}\mathrm{d}\xi \tag{6.7}$$

The proof of this theorem uses that the *Gaussian is a good kernel*, in the sense that the family of functions

$$K_\delta(x) = \delta^{-1/2}e^{-\pi x^2/\delta} \tag{6.8}$$

have the properties of good kernels discussed in Lecture 2.2, namely

1. $\int_{-\infty}^{\infty} K_\delta(x)\mathrm{d}x = 1$
2. $\int_{-\infty}^{\infty} |K_\delta(x)|\mathrm{d}x \leq M$

3. For every $\eta > 0$, we have

$$\int_{|x|>\eta} |K_\delta(x)|\mathrm{d}x \to 0 \qquad (\delta \to 0)\,. \tag{6.9}$$

The fundamental property of the Gaussian is that it equals its Fourier transform.

Theorem 6.4. *If $f(x) = e^{-\pi x^2}$, then $\hat{f}(\xi) = f(\xi)$.*

Taking this fact for granted, we infer from the dilation property of Prop 6.1 that

$$\widehat{K_\delta}(\xi) = e^{-\pi\delta\xi^2}\,. \tag{6.10}$$

Note that as $\delta \to 0$, the function K_δ peaks at the origin, while $\widehat{K_\delta}$ flattens out.

Proof that K_δ is a family of good kernels. The first property now follows from

$$\int_{-\infty}^{\infty} K_\delta(x)\mathrm{d}x = \widehat{K_\delta}(0) = 1\,.$$

Since $K_\delta \geq 0$, this also implies the second property with $M = 1$. For the third property, note that by a change of variables

$$\int_{|x|>\eta} K_\delta(x)\mathrm{d}x = \int_{|y|>\eta\delta^{-1/2}} e^{-\pi y^2}\mathrm{d}y \to 0 \qquad (\delta \to 0)\,.$$

□

Similary to the arguments in Lecture 2.2 we can now prove:

Theorem 6.5. *If $f \in \mathcal{S}(\mathbb{R})$, then*

$$(f * K_\delta)(x) \to f(x) \qquad \textit{uniformly in } x \textit{ as } \delta \to 0\,. \tag{6.11}$$

Let us now return to the proof of Theorem 6.3, which relies on one hand on the properties of the Gaussian, and on the other on the following **multiplication formula**:

Theorem 6.6. *If $f, g \in \mathcal{S}(\mathbb{R})$, then*

$$\int_{-\infty}^{\infty} f(x)\hat{g}(x)\mathrm{d}x = \int_{-\infty}^{\infty} \hat{f}(y)g(y)\mathrm{d}y\,. \tag{6.12}$$

This formula follows essentially by integrating the function $f(x)g(y)e^{-2\pi ixy}$ alternatively first in x or y.

Proof of Theorem 6.3. Let $G_\delta(x) = e^{-\pi\delta x^2}$ so that $\widehat{K_\delta} = G_\delta$, and also $\widehat{G_\delta}(\xi) = K_\delta(\xi)$. By the multiplication formula we get

$$\int_{-\infty}^{\infty} f(x)K_\delta(x)\mathrm{d}x = \int_{-\infty}^{\infty} \hat{f}(\xi)G_\delta(\xi)\mathrm{d}\xi\,. \tag{6.13}$$

Since K_δ is a good kernel, the integral on the left hand side converges to $f(0)$ as $\delta \to 0$, and thus

$$f(0) = \int_{-\infty}^{\infty} \hat{f}(\xi)\mathrm{d}\xi\,, \tag{6.14}$$

which proves the inversion formula for $x = 0$. The general formula now follows easily by setting $F(y) = f(x+y)$ so that

$$f(x) = F(0) = \int_{-\infty}^{\infty} \hat{F}(\xi)\mathrm{d}\xi = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi i x\xi}\mathrm{d}\xi\,. \tag{6.15}$$

□

This result is also expressed by the statement that

$$\mathcal{F} \circ \mathcal{F}^* = \mathcal{F}^* \circ \mathcal{F} = I \tag{6.16}$$

where $\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ and $\mathcal{F}^* : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ are the mappings

$$\mathcal{F}(f)(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\xi}\mathrm{d}\xi \tag{6.17}$$

$$\mathcal{F}^*(g)(\xi) = \int_{-\infty}^{\infty} g(x)e^{2\pi i x\xi}\mathrm{d}\xi \tag{6.18}$$

The Fourier transform $\mathcal{F}$ is a *bijective* mapping on the Schwartz space.

6.3. Plancherel's formula

Above we have already used the **convolution** of two functions $f, g \in \mathcal{S}(\mathbb{R})$, defined by

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-t)g(t)\mathrm{d}t\,. \tag{6.19}$$

Note that for fixed x, the function $f(x-t)g(t)$ is of rapid decrease in t, hence this integral converges. We mention a few further results about convolutions of Schwartz functions.

Theorem 6.7. *If $f, g \in \mathcal{S}(\mathbb{R})$, then*

1. $f * g \in \mathcal{S}(\mathbb{R})$
2. $f * g = g * f$
3. $\widehat{f * g}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$

The last property — which shows that the Fourier transform interchanges convolutions with products — is the most important statement, which is proven as follows:

Proof. Consider the function $F(x, y) = f(y)g(x-y)e^{-2\pi i x\xi}$, which is rapidly decreasing in y for fixed x, and also in x for fixed y. Integrating in y first gives, on one hand, $(f * g)(x)e^{-2\pi i x\xi}$, and then in x gives $\widehat{f * g}(\xi)$. On the other hand, integrating in x first yields $f(y)\hat{g}(\xi)e^{-2\pi i y\xi}$, which after integrating in y gives $\hat{f}(\xi)\hat{g}(\xi)$. This proves the identity *3.* □

The Schwartz space can be equipped with a Hermitian inner product

$$(f,g) = \int_{-\infty}^{\infty} f(x)\overline{g(x)}\mathrm{d}x \tag{6.20}$$

whose associated norm is $\|f\| = (f,f)^{1/2}$.

One can use the above properties of convolutions of Schwartz functions to prove that the Fourier transform is a *unitary transformation* on $\mathcal{S}(\mathbb{R})$.

Theorem 6.8 (Plancherel)**.** *If* $f \in \mathcal{S}(\mathbb{R})$ *then* $\|\hat{f}\| = \|f\|$.

Proof. Let

$$h(x) = \int_{-\infty}^{\infty} f(y)\overline{f(y-x)}\mathrm{d}y \tag{6.21}$$

Then $h(0) = \|f\|^2$, and by Fourier inversion

$$h(0) = \int_{-\infty}^{\infty} \hat{h}(\xi)\mathrm{d}\xi \tag{6.22}$$

We can compute the Fourier transform of h because it is a convolution: $\hat{h}(\xi) = \hat{f}(\xi)\overline{\hat{f}(\xi)} = |\hat{f}(\xi)|^2$, because

$$\int_{-\infty}^{\infty} \overline{f(-x)}e^{-2\pi i x\xi}\mathrm{d}x = -\int_{\infty}^{-\infty} \overline{f(x)}e^{2\pi i x\xi}\mathrm{d}x = \overline{\hat{f}(\xi)}\,. \tag{6.23}$$

Thus also $h(0) = \|\hat{f}\|^2$. □

Supplementary Problems

1. Let f be the characteristic function on $[-1,1]$ defined by

$$f(x) = \chi_{[-1,1]}(x) = \begin{cases} 1, & |x| \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{6.24}$$

   Although f is not continuous, the integral defining its Fourier transform still makes sense. Show that

$$\hat{f}(\xi) = \frac{\sin(2\pi\xi)}{\pi\xi}\,. \tag{6.25}$$

2. This problem gives an example of a **bump function**.

   a) Suppose $a < b$, and f is the function $f(x) = 0$ if $x \leq a$ or $x \geq b$, and for $a < x < b$:

$$f(x) = e^{-\frac{1}{x-a}}e^{-\frac{1}{b-x}} \tag{6.26}$$

   Show that f is indefinitely differentiable on $\mathbb{R}$.

b) Give an indefinitely differentiable function F on $\mathbb{R}$ such that $F(x) = 0$ for $x \leq a$, and $F(x) = 1$ if $x \geq b$ which is strictly increasing on $[a, b]$.

   Hint: Consider $F(x) = c \int_{-\infty}^{x} f(t) \mathrm{d}t$ for some appropriate constant c.

c) Let $\delta > 0$ be such that $a + \delta < b - \delta$. Show that there exists an indefinitely differentiable function g which vanishes for $x \leq a$, and $x \geq b$, and $g(x) = 1$ on $[a + \delta, b - \delta]$, and moreover is strictly monotone on $[a, a + \delta]$ and $[b - \delta, b]$.

Lecture 6.

Heat equation

Further Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 5, Section 2).

6.1. Time-dependent heat equation on the real line

Consider the **heat equation** on the line,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{6.1}$$

with initial condition

$$u(x,0) = f(x)\,. \tag{6.2}$$

Let us first take the Fourier transform, formally in (6.1), then we see that the Fourier transform in x of the solution must satisfy

$$\frac{\partial \hat{u}}{\partial t}(\xi, t) = -4\pi^2\xi^2\hat{u}(\xi,t)\,, \tag{6.3}$$

which is a first-order differential equation in t, with the solution

$$\hat{u}(\xi,t) = A(\xi)e^{-4\pi^2\xi^2 t}\,. \tag{6.4}$$

We can accomodate the initial condition by setting $A(\xi) = \hat{f}(\xi)$.

We recognise this solution precisely as the Fourier transform of the kernels K_δ, $\widehat{K_\delta}(\xi) = e^{-\pi\delta\xi^2}$, with $\delta = 4\pi t$. Therefore we define the **heat kernel** of the line by

$$\mathcal{H}_t(x) = K_{4\pi t}(x) = \frac{1}{(4\pi t)^{1/2}}e^{-x^2/4t}\,, \tag{6.5}$$

and we can see, cf. **Review** of the Fourier transform above, that $\mathcal{H}_t * f$ should be the desired solution of the heat equation.

Theorem 6.1. *Given $f \in \mathcal{S}(\mathbb{R})$, let*

$$u(x,t) = (f * \mathcal{H}_t)(x) \qquad (t > 0) \tag{6.6}$$

where $\mathcal{H}_t$ is the heat kernel. Then

1. *The function $u(x,t)$ is twice continuously differentiable (for $t>0$) and solves* (6.1).

2. $u(x,t) \to f(x)$ *uniformly in x as $t \to 0$.*

3. *We have as $t \to 0$,*

$$\int_{-\infty}^{\infty} |u(x,t) - f(x)|^2 \mathrm{d}x \to 0\,. \tag{6.7}$$

Proof. We can express $u(x,t)$ using the Fourier inversion formula as

$$u(x,t) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{-4\pi^2\xi^2 t} e^{2\pi i \xi x} \mathrm{d}\xi\,. \tag{6.8}$$

This shows in particular that $u(x,t)$ is indefinitely differentiable in x, and t, for $t>0$. Moreover uniform convergence to $f(x)$, as $t \to 0$, follows because $\mathcal{H}_t$ is a good kernel. The last property follows from Plancharel's theorem

$$\int_{-\infty}^{\infty} |u(x,t) - f(x)|^2 \mathrm{d}x = \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 |e^{-4\pi^2\xi^2 t} - 1|^2 \mathrm{d}x \to 0 \qquad (t \to 0) \tag{6.9}$$

which can be readily seen by splitting up the integral over $\mathbb{R}$ in $[-N,N]$ and its complement.

□

This theorem guarantees the existence of a solution to (6.1) for any initial condition $f \in \mathcal{S}(\mathbb{R})$ in (6.2). The solution is also *unique* as can be seen most directly using the **energy**-type quantity $E(t)$ at time t of the solution $u(x,t)$:

$$E(t) = \int_{\mathbb{R}} |u(x,t)|^2 \mathrm{d}x\,. \tag{6.10}$$

If we assume that $u(x,0)=0$, then $E(0)=0$, and we compute

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \int_{\mathbb{R}} (2u\partial_t u)(x,t)\mathrm{d}x = 2\int_{\mathbb{R}} (u\partial_x^2 u)(x,t)\mathrm{d}x = -2\int_{\mathbb{R}} (\partial_x u)^2(x,t)\mathrm{d}x \leq 0 \tag{6.11}$$

and hence $E(t)=0$, and thus $u(x,t)=0$ for all $t>0$.

Exercise 6.1. What is needed here to justify this computation?

6.2. The steady-state heat equation in the upper half plane

A problem that can be solved in a similar fashion is the steady state heat equation in the upper half plane,

$$\triangle u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \qquad (y>0) \tag{6.12}$$

with the boundary condition

$$u(x,0) = f(x)\,. \tag{6.13}$$

Taking the Fourier transform, first formally, in the x variable, we obtain

$$-4\pi^2\xi^2\hat{u} + \frac{\partial^2 \hat{u}}{\partial y^2} = 0 \tag{6.14}$$

which again is an ordinary second order differential equation in y with constant coefficients, with the general solution

$$\hat{u}(\xi, y) = A(\xi)e^{-2\pi|\xi|y} + B(\xi)e^{2\pi|\xi|y} \tag{6.15}$$

We dismiss the exponentially growing solution, and accomodate the boundary condition by choosing $A(\xi) = \hat{f}(\xi)$, which leaves us with

$$\hat{u}(\xi, y) = \hat{f}(\xi)e^{-2\pi|\xi|y} \tag{6.16}$$

Thus if we can find a function $\mathcal{P}_y(x)$ such that $\widehat{\mathcal{P}_y} = e^{-2\pi|\xi|y}$, then the solution to (6.12) with the boundary condition (6.13) can be expressed as

$$u(x, y) = (\mathcal{P}_y * f)(x)\,. \tag{6.17}$$

It turns out this function is the **Poisson kernel** for the upper half plane,

$$\mathcal{P}_y(x) = \frac{1}{\pi}\frac{y}{x^2 + y^2} \qquad (x \in \mathbb{R}, y > 0) \tag{6.18}$$

In fact, we can compute

$$\mathcal{P}_y(x) = \int_{-\infty}^{\infty} e^{-2\pi|\xi|y}e^{2\pi i\xi x}\mathrm{d}\xi \tag{6.19}$$

by evaluating the integrals on $(-\infty, 0)$ and $(0, \infty)$ separately. We find

$$\int_0^\infty e^{-2\pi\xi y}e^{2\pi i\xi x}\mathrm{d}\xi = \int_0^\infty e^{2\pi i(iy+x)\xi}\mathrm{d}\xi = \frac{e^{2\pi i(x+iy)\xi}}{2\pi i(x+iy)}\Big|_0^\infty = -\frac{1}{2\pi i(x+iy)} \tag{6.20}$$

because $y > 0$, and similarly

$$\int_{-\infty}^0 e^{2\pi\xi y}e^{2\pi i\xi x}\mathrm{d}\xi = \frac{1}{2\pi i(x - iy)}\,. \tag{6.21}$$

Therefore

$$\mathcal{P}_y(x) = -\frac{1}{2\pi i(x+iy)} + \frac{1}{2\pi i(x-iy)} = \frac{1}{\pi}\frac{y}{x^2+y^2}\,. \tag{6.22}$$

Lemma 6.2. *The Poisson kernel is a good kernel on* $\mathbb{R}$ *as* $y \to 0$.

Proof. In view of the Fourier inversion formula we have

$$\widehat{P_y}(\xi) = \int_{-\infty}^{\infty} \mathcal{P}_y(x)e^{-2\pi i x\xi}\mathrm{d}x = e^{-2\pi|\xi|y} \tag{6.23}$$

which shows in particular that

$$\int_{-\infty}^{\infty} \mathcal{P}_y(x)\mathrm{d}x = 1 \tag{6.24}$$

and the same is true for the integral of $|\mathcal{P}_y(x)|$, because $\mathcal{P}_y(x) \geq 0$. It remains to show that for every $\eta > 0$,

$$\int_{|x|>\eta} |\mathcal{P}_y(x)|\mathrm{d}x \to 0 \qquad (y \to 0)\,. \tag{6.25}$$

Now by a change of variables, to $s = x/y$,

$$\int_{\eta}^{\infty} \frac{y}{x^2+y^2}\mathrm{d}x = \int_{\eta/y}^{\infty} \frac{1}{s^2+1}\mathrm{d}s \to 0 \tag{6.26}$$

as $y \to 0$. □

Similarly to before this allows us to prove:

Theorem 6.3. *Given $f \in \mathcal{S}(\mathbb{R})$, let $u(x,y) = (\mathcal{P}_y * f)(x)$. Then*

1. *u is twice differentiable in the upper half plane, and $\triangle u = 0$.*
2. *$u(x,y) \to f(x)$ uniformly as $y \to 0$.*
3. *$\int_{-\infty}^{\infty} |u(x,y) - f(x)|^2\mathrm{d}x \to 0$ as $y \to 0$.*
4. *If $u(x,0) = f(x)$, then u is continuous on the closure of the upper half plane, and vanishes at infinity in the sense that $u(x,y) \to 0$ as $|x| + y \to \infty$.*

Instead of working out the details of this proof, let us think about the *uniqueness* statement, which relies on the **mean value property** of harmonic functions.

Proposition 6.4. *Suppose Ω is an open set in $\mathbb{R}^2$, and let $u \in \mathrm{C}^2$ be a solution to $\triangle u = 0$ in Ω. If the closure of the disc of radius $R > 0$ centered at (x,y) is contained in Ω, then*

$$u(x,y) = \frac{1}{2\pi}\int_0^{2\pi} u(x + r\cos\theta, y + r\sin\theta)\mathrm{d}\theta \tag{6.27}$$

for all $0 \leq r \leq R$.

Proof. Let $U(r,\theta) = u(x + r\cos\theta, y + r\sin\theta)$. Expressing the Laplacian in polar coordinates, the equation implies

$$r\frac{\partial}{\partial r}\Big(r\frac{\partial U}{\partial r}\Big) + \frac{\partial^2 U}{\partial \theta^2} = 0 \tag{6.28}$$

Since

$$\int_0^{2\pi} \partial_\theta^2 U(r,\theta)\mathrm{d}\theta = \partial_\theta U(r,\theta)|_0^{2\pi} = 0 \tag{6.29}$$

we see that

$$F(r) = \frac{1}{2\pi}\int_0^{2\pi} U(r,\theta)\mathrm{d}\theta \tag{6.30}$$

satisfies $r\partial_r(r\partial_r F) = 0$, and consequently $r\partial_r F(r)$ must be constant, so evaluating at $r = 0$, we see that $\partial_r F = 0$, which means that F is constant, which yields: $u(x, y) = F(0) = F(r)$. ☐

We can now prove uniqueness:

Theorem 6.5. *Suppose u is continuous on the closure of the upper half plane, and satisfies $\triangle u = 0$ for $y > 0$. Assume moreover that $u(x, 0) = 0$, and $u(x, y)$ vanishes at infinity. Then $u = 0$.*

Exercise 6.2. Show that the assumption that u vanishes at infinity is necessary.

Proof. Suppose there is a point $u(x_0, y_0)$ in the upper half plane where $u(x_0, y_0) > 0$. Since u vanishes at infinity, we can choose $R > 0$ so that outside the half disc

$$D_R^+ = \{(x, y) : x^2 + y^2 \leq R^2, y \geq 0\} \tag{6.31}$$

we have $u(x, y) \leq u(x_0, y_0)/2$. Since u is continuous it attains a maximum on D_R^+, say at a point $(x_1, y_1) \in D_R^+$:

$$u(x_1, y_1) = M \tag{6.32}$$

Then $u(x, y) \leq M$ on the disc, and in particular $u(x_0, x_0) \leq M$, and so also outside the disc $u(x, y) \leq M/2$. In other words, u is bounded by M on the whole closed upper half plane.

Now by the mean value property,

$$u(x_1, y_1) = \frac{1}{2\pi} \int_0^{2\pi} u(x_1 + r\cos\theta, y_1 + r\sin\theta) \mathrm{d}\theta \tag{6.33}$$

whenever the circle of integration lies in the upper half plane, so for all $r < y_1$.

Exercise 6.3. Show by using continuity of u, that this implies that on the whole circle:

$$u(x_1 + r\cos\theta, y_1 + r\sin\theta) = M \tag{6.34}$$

In particular, by taking $r \to y_1$ and using continuity of u up to the boundary $y = 0$, we conclude that $u(x_1, 0) = M$, in contradiction to the assumption of vanishing boundary values.

☐

Problems

1. Suppose that u is the solution to the heat equation given by $u = f * \mathcal{H}_t$ where $f \in \mathcal{S}(\mathbb{R})$. If we also set $u(x, 0) = f(x)$, prove that u is continuous on the closure of the upper half-plane, and vanishes at infinity, in the sense that

$$u(x, t) \to 0 \qquad \text{as } |x| + t \to \infty\,. \tag{6.35}$$

   Hint: To prove that u vanishes at infinity, show that

a) $|u(x,t)| \le C/\sqrt{t}$

b) $|u(x,t)| \le C/(1+|x|^2) + Ct^{-1/2}e^{-cx^2/t}$.

Use *a)* when $|x| \le t$, and *b)* otherwise.

2. Show that the function

$$u(x,t) = \frac{x}{t}\mathcal{H}_t(x) \tag{6.36}$$

satisfies the heat equation for $t > 0$ and $\lim_{t\to 0} u(x,t) = 0$ for every x, but u is *not* continuous at the origin.

Hint: Approach the origin with (x,t) on the parabola $x^2/4t = c$ where c is a constant.

3. Prove the following uniqueness theorem for harmonic functions in the strip

$$Z = \{(x,y) : 0 < y < 1, -\infty < x < \infty\}. \tag{6.37}$$

Suppose u satisfies $\triangle u = 0$ in Z and is continuous on its closure. If $u(x,0) = u(x,1) = 0$ for all x, and u vanishes at infinity, then $u = 0$ in Z.

Review: The Fourier transform on $\mathbb{R}^d$

Recommended Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 6, Section 2, and 4).

We briefly review the Fourier transform in $\mathbb{R}^d$, and will see that with the appropriate notation, the key theorems, such as the Fourier inversion formula, remain unchanged.

We will use **multi-index notation** for both monomials and differential operators: For $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, and $\alpha = (\alpha_1, \ldots, \alpha_d)$ and d-tuple of non-negative integers, we write

$$x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d} \tag{7.1}$$

$$(\frac{\partial}{\partial x})^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}} \tag{7.2}$$

where $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d$.

The **Schwartz space** $\mathcal{S}(\mathbb{R}^d)$ consists of all indefinitely differentiable functions f on $\mathbb{R}^d$ such that

$$\sup_{x\in\mathbb{R}^d} \left| x^\alpha \partial_x^\beta f(x) \right| < \infty \tag{7.3}$$

for every multi-index α, and β. In other words, f and all its derivatives are required to be *rapidly decreasing.*

In complete analogy to (6.1) we define the **Fourier transform** $\hat{f}$ of a function f on the $\mathbb{R}^d$ by

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x\cdot\xi} \mathrm{d}x \qquad (\xi \in \mathbb{R}^d) \tag{7.4}$$

Similarly to Prop. 6.1 we note that:

Proposition 7.1. *If $f \in \mathcal{S}(\mathbb{R}^d)$, then*

1. $f(x+h) \longrightarrow \hat{f}(\xi) e^{2\pi i h\cdot\xi}$ *whenever* $h \in \mathbb{R}^d$
2. $f(x) e^{-2\pi i x\cdot h} \longrightarrow \hat{f}(\xi + h)$ *whenever* $h \in \mathbb{R}^d$
3. $f(\delta x) \longrightarrow \delta^{-d} \hat{f}(\delta^{-1}\xi)$ *whenever* $\delta > 0$.
4. $\left(\frac{\partial}{\partial x}\right)^\alpha f(x) \to (2\pi i \xi)^\alpha \hat{f}(\xi)$
5. $(-2\pi i x)^\alpha f(x) \longrightarrow \left(\frac{\partial}{\partial \xi}\right)^\alpha \hat{f}(\xi)$

6. *$f(Rx) \longrightarrow \hat{f}(R\xi)$ whenever R is a rotation.*

What is new is really just the last part. Recall that a **rotation** is a linear transformation $R : \mathbb{R}^d \to \mathbb{R}^d$ which preserves the standard inner product,

$$Rx \cdot Ry = x \cdot y = x_1 y_1 + \ldots + x_d y_d \tag{7.5}$$

Equivalently, this condition can be replaced by $R^t = R^{-1}$, which implies that $\det(R) = \pm 1$. If $\det(R) = 1$ we say R is a proper rotation.

The statement for the Fourier transform of $f(Rx)$ now follows from a change of variables to $y = Rx$,

$$\int_{\mathbb{R}^d} f(Rx) e^{-2\pi i x\cdot\xi} \mathrm{d}x = \int_{\mathbb{R}^d} f(y) e^{-2\pi i R^{-1} y\cdot\xi} |\det(R)|^{-1} \mathrm{d}y = \hat{f}(R\xi)\,. \tag{7.6}$$

This property also implies that

Corollary 7.2. *The Fourier transform of a radial function is radial.*

A **radial** function on $\mathbb{R}^d$ is a function that only depends on $|x|$. Equivalently a function is radial if and only if $f(Rx) = f(x)$ for all rotations r. This obviously implies that $\hat{f}(R\xi) = \hat{f}(\xi)$.

This shows that if $f(x) = f_0(|x|)$ for some function f_0, then $\hat{f}(\xi) = F_0(|\xi|)$ for some F_0. But to obtain a formula for F_0 in terms of f_0 in the case $d = 3$, it will be important to compute "the Fourier transform of the surface element of $\mathbb{S}^2$," which will also play a role in the derivation of the representation formula for solutions to the wave equation in dimension $d = 3 + 1$.

Lemma 7.3.

$$\frac{1}{4\pi} \int_{\mathbb{S}^2} e^{-2\pi i \xi\cdot\gamma} \mathrm{d}\sigma(\gamma) = \frac{\sin(2\pi|\xi|)}{2\pi|\xi|} \tag{7.7}$$

Proof. The left hand side is a radial function, so if $|\xi| = \rho$ we can prove the formula with $\xi = (0, 0, \rho)$. We express the integral in polar coordinates:

$$\frac{1}{4\pi} \int_{\mathbb{S}^2} e^{-2\pi i \xi\cdot\gamma} \mathrm{d}\sigma(\gamma) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} e^{-2\pi i \rho \cos\theta} \sin(\theta) \mathrm{d}\theta \mathrm{d}\varphi\,, \tag{7.8}$$

and the stated formula follows with a change of variables $u = -\cos(\theta)$. □

So let $f(x) = f_0(|x|)$ be a radial function on $\mathbb{R}^d$, and $\hat{f}(\xi) = F_0(|\xi|)$ for some F_0 as above. Then in dimension $d = 3$, by Lemma 7.3 with $\rho = |\xi|$,

$$\begin{aligned} F_0(\rho) = \hat{f}(\xi) &= \int_{\mathbb{R}^3} f(x) e^{-2\pi i x\cdot\xi} \mathrm{d}x \\ &= \int_0^\infty f_0(r) \int_{\mathbb{S}^2} e^{-2\pi i r\gamma\cdot\xi} \mathrm{d}\sigma(\gamma) r^2 \mathrm{d}r \\ &= 2\int_0^\infty f_0(r) \frac{\sin(2\pi\rho r)}{\rho r} r^2 \mathrm{d}r \\ &= \frac{2}{\rho} \int_0^\infty \sin(2\pi\rho r) f_0(r) r \mathrm{d}r\,. \end{aligned} \tag{7.9}$$

Now let us first return to the discussion of the Fourier transform more generally. Proposition 7.1 shows in particular that the *Fourier transform maps* $\mathcal{S}(\mathbb{R}^d)$ *to itself.* The main result is the following **Fourier inversion formula**, and **Plancherel**'s theorem for functions on $\mathbb{R}^d$.

Theorem 7.4. *Suppose* $f \in \mathcal{S}(\mathbb{R}^d)$*. Then*

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi i x\cdot\xi} \mathrm{d}\xi \tag{7.10}$$

$$\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 \mathrm{d}\xi = \int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}x\,. \tag{7.11}$$

The proof is again based on the fact that the Gaussian $e^{-\pi|x|^2}$ is its own Fourier transform, and that $K_\delta(x) = \delta^{-d/2} e^{-\pi|x|^2/\delta}$ is a family of good kernels. We omit the details.

Problems

1. (Euler's geometric description of rotations in $\mathbb{R}^3$)

   Show that given a proper rotation R, there exists a unit vector γ so that

   a) R fixes γ, that is, $R(\gamma) = \gamma$.

   b) If $\mathcal{P}$ denotes the plane perpendicular to γ, and passing through the origin, then $R : \mathcal{P} \to \mathcal{P}$ and the restriction of R to $\mathcal{P}$ is a rotation in $\mathbb{R}^2$.

   Hint: To prove that there exists $\gamma \in \mathbb{S}^2$ with $R(\gamma) = \gamma$ (geometrically γ is the direction of the axis of rotation), show that $p(t) = \det(R - tI)$ is a polynomial of degree 3. Then use that $p(0) > 0$ to see that there is a $\lambda > 0$ with $p(\lambda) = 0$. This means that the kernel of $R - \lambda I$ is non-trivial.

Lecture 7.

The wave equation on $\mathbb{R}^d \times \mathbb{R}$

Recommended Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 6, Section 3).

Further reading also (John, *Partial differential equations*, Chapter 5, Section 1)

In this lecture we study the initial value problem for the wave equation on $\mathbb{R}^d \times \mathbb{R}$,

$$-\frac{\partial^2 u}{\partial t^2} + \triangle u = 0 \tag{7.1}$$

where u satisfies the initial conditions

$$u(x,0) = f(x) \qquad \frac{\partial u}{\partial t}(x,0) = g(x)\,. \tag{7.2}$$

Here we restrict ourselves to functions $f, g \in \mathcal{S}(\mathbb{R}^d)$. While the wave equation allows much rougher solutions, we can already see many properties of the general behaviour for initial data in Schwartz space. In particular, we can see using the Fourier transform that there *exist* solutions to the **Cauchy problem**.

7.1. Fourier representation formula

Suppose first that a solution to (7.1) exists with $u(\cdot,t) \in \mathcal{S}(\mathbb{R}^d)$. Then we can take the Fourier transform in $\mathbb{R}^d$, and get

$$\frac{\partial \hat{u}}{\partial t^2} + 4\pi^2|\xi|^2\hat{u} = 0 \tag{7.3}$$

For each fixed $\xi \in \mathbb{R}^d$ this is a differential equation with the solution

$$\hat{u}(\xi,t) = A(\xi)\cos(2\pi|\xi|t) + B(\xi)\sin(2\pi|\xi|t) \tag{7.4}$$

where we can choose $A(\xi)$ and $B(\xi)$ to accomodate the initial conditions:

$$\hat{u}(\xi,0) = \hat{f}(\xi) \qquad \partial_t\hat{u}(\xi,0) = \hat{g}(\xi) \tag{7.5}$$

so therefore we want to set

$$A(\xi) = \hat{f}(\xi) \qquad 2\pi|\xi|B(\xi) = \hat{g}(\xi)\,. \tag{7.6}$$

Theorem 7.1. *Suppose $f, g \in \mathcal{S}(\mathbb{R}^d)$. Then a solution to the Cauchy problem is given by*

$$u(x,t) = \int_{\mathbb{R}^d} \Big[\hat{f}(\xi)\cos(2\pi|\xi|t) + \hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\Big] e^{2\pi i\xi\cdot x}\mathrm{d}\xi \tag{7.7}$$

Exercise 7.1. Verify that $u(x,t)$ defined in this way actually is a solution to (7.1) that satisfies the initial conditions (7.2).

This theorem gives *a* solution to the Cauchy problem, but to say that it is *the* solution, we need to prove that for given initial data there is only one solution to the initial value problem. This argument is based on the **energy** of a solution defined by

$$E(t) = \int_{\mathbb{R}^d} (\partial_t u(t,x))^2 + |\nabla u(t,x)|^2 \mathrm{d}x\,. \tag{7.8}$$

We have, assuming for example that $\partial_t u(t,\cdot)$, $\partial_j u(t,\cdot) \in \mathcal{S}(\mathbb{R}^d)$, for $t \geq 0$,

$$E'(t) = 2\int_{\mathbb{R}^d} \partial_t u(t,x)\Big(\partial_t^2 u(t,x) - \triangle u(t,x)\Big)\mathrm{d}x = 0 \tag{7.9}$$

and so $E(t) = 0$ $(t \geq 0)$, whenever the initial data is trivial with $E(0) = 0$.

Exercise 7.2. In fact a stronger uniqueness statement can be proven demonstrating the *finite speed of propagation* in this problem. See Problems below.

Now recall from the very first lecture (Lecture 2) that for $d = 1$ d'Alembert's formula holds:

$$u(t,x) = \frac{f(x+t) + f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(y)\mathrm{d}y \tag{7.10}$$

We can verify that this formula really does follow from (7.7) in the case $d = 1$: Since $\cos(2\pi|\xi|t) = \cos(2\pi\xi t) = (e^{2\pi i\xi t} + e^{-2\pi i\xi t})/2$,

$$\int_{\mathbb{R}} \hat{f}(\xi)\cos(2\pi|\xi|t)e^{2\pi i\xi x}\mathrm{d}\xi = \frac{1}{2}\int_{\mathbb{R}} \hat{f}(\xi)(e^{2\pi i\xi(t+x)} + e^{2\pi i(x-t)})\mathrm{d}\xi = \frac{1}{2}(f(t+x) + f(x-t)) \tag{7.11}$$

and similarly

$$\begin{aligned} \int_{\mathbb{R}} \hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}e^{2\pi i\xi x}\mathrm{d}\xi &= \int_{\mathbb{R}} \hat{g}(\xi)\frac{e^{2\pi i\xi(t+x)} - e^{-2\pi i\xi(t-x)}}{4\pi i\xi}\mathrm{d}\xi = \\ &= \frac{1}{2}\int_{\mathbb{R}} \hat{g}(\xi)\int_{x-t}^{x+t} e^{2\pi i\xi y}\mathrm{d}y\mathrm{d}\xi = \frac{1}{2}\int_{x-t}^{x+t} g(y)\mathrm{d}y\,. \end{aligned} \tag{7.12}$$

Note that the first term in d'Alembert's formula is an average of the values of f on the boundary of the interval $[x-t, x+t]$, and the second term is up to a factor of t the mean value of the function g over the interval $[x-t, x+t]$. We will see that also in the case $d = 3$ the values of the solution can be expressed as averages over the initial data, on domains that are (causally) related in the same way.

7.2. Spherical means and Huygens principle

Given a function f on $\mathbb{R}^3$, we denote the **spherical mean** over the sphere of radius t centered at x by

$$M_t(f)(x) = \frac{1}{4\pi}\int_{\mathbb{S}^2} f(x - t\gamma)\mathrm{d}\sigma(\gamma)\,, \tag{7.13}$$

where $\mathrm{d}\sigma$ is the area element on $\mathbb{S}^2$.

One can show that if f is a Schwartz function, then so is the spherical mean M_t. Moreover $M_t(f)$ is then indefinitely differentiable in t, and each derivative is a Schwartz function. We will now show that the solution found in Theorem 7.1 may be expressed in terms of spherical means of the initial data.

Theorem 7.2. *In dimension $d = 3$, the solution to the Cauchy problem (7.1, 7.2) is given by*

$$u(x,t) = \frac{\partial}{\partial t}(tM_t(f)(x)) + tM_t(g)(x)\,. \tag{7.14}$$

The proof relies on the observation that

$$\widehat{M_t(f)}(\xi) = \hat{f}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|t} \tag{7.15}$$

We can make sense of this formula as follows: $M_t(f)$ can be viewed as the convolution of f with surface element $\mathrm{d}\sigma$, and thus can expect the Fourier transform to be the product of the corresponding Fourier transforms. Indeed we have already computed the Fourier transform of the surface measure in Lemma 7.3, and so to verify the formula let us write

$$\begin{aligned}\widehat{M_t(f)}(\xi) =& \frac{1}{4\pi}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2} f(x - t\gamma)\mathrm{d}\sigma(\gamma)e^{-2\pi i x\cdot\xi}\mathrm{d}x \\ =& \frac{1}{4\pi}\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} f(y)e^{-2\pi i y\cdot\xi}e^{-2\pi i t\gamma\cdot\xi}\mathrm{d}y\mathrm{d}\sigma(\gamma) = \frac{1}{4\pi}\int_{\mathbb{S}^2}\hat{f}(\xi)e^{-2\pi i t\gamma\cdot\xi}\mathrm{d}\sigma(\gamma)\,.\end{aligned} \tag{7.16}$$

Proof. First let us set $f = 0$. Then according to Theorem 7.1,

$$u(x,t) = t\int_{\mathbb{R}^3}\hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|t}e^{2\pi i\xi\cdot x}\mathrm{d}\xi = t\int_{\mathbb{R}^3}\widehat{M_t(g)}(\xi)e^{2\pi i\xi\cdot x}\mathrm{d}\xi = tM_t(g)(x)\,. \tag{7.17}$$

Second if $g = 0$ then by Theorem 7.1,

$$\begin{aligned}u(x,t) =& \int_{\mathbb{R}^d}\hat{f}(\xi)\cos(2\pi|\xi|t)e^{2\pi i\xi\cdot x}\mathrm{d}\xi \\ =& \frac{\partial}{\partial t}\Big(t\int_{\mathbb{R}^d}\hat{f}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|t}e^{2\pi i\xi\cdot x}\mathrm{d}\xi\Big) = \frac{\partial}{\partial t}\Big(tM_t(f)\Big)\,.\end{aligned} \tag{7.18}$$

In view of the linearity of the problem the general statement follows. □

This formula tells us that solutions to the wave equation depend on the initial data in the following way: To compute the value $u(x,t)$, draw the **backward light cone** with vertex at (x,t):

$$C_{(x,t)} = \Big\{(y,s) \in \mathbb{R}^3 \times \mathbb{R} : |s-t| = |y-x|, s \leq t\Big\} \tag{7.19}$$

Then only the values of (f,g) on (or strictly speaking in an immediate neighborhood of) the intersection of $C_{(x,t)}$ with the hyperplane $t=0$ in $\mathbb{R}^{3+1}$ enter the formula for $u(x,t)$. This has various consequences which one way or another are referred to as **Huygens principle**.

Exercise 7.3. Suppose f, and g are compactly supported. In fact, let us assume that f, and g, vanish outside a ball $B_r(x_0)$. Sketch the support of the solution u to the Cauchy problem.

7.3. Method of descent

Consider the wave equation on $\mathbb{R}^2 \times \mathbb{R}$,

$$-\partial_t^2 u + \triangle u = 0\,, \qquad \triangle = \partial_{x_1}^2 + \partial_{x_2}^2\,, \tag{7.20}$$

with initial data $f,g \in \mathcal{S}(\mathbb{R}^2)$. We could try to solve this equation by first extending f,g trivially to functions on $\mathbb{R}^3$,

$$\tilde{f}(x_1,x_2,x_3) = f(x_1,x_2)\,, \qquad \tilde{g}(x_1,x_2,x_3) = g(x_1,x_2)\,, \tag{7.21}$$

however that does not work because $\tilde{f}$, and $\tilde{g}$ are not Schwartz functions. But we could introduce a **cutoff**, such as a smooth function of compact support $\eta_T \in \mathcal{S}(\mathbb{R})$, with the property that $\eta_T(x) = 1$ whenever $|x| \leq T$. Then

$$\tilde{f}^\flat(x_1,x_2,x_3) = f(x_1,x_2)\eta_T(x_3)\,, \qquad \tilde{g}^\flat(x_1,x_2,x_3) = g(x_1,x_2)\eta_T(x_3)\,, \tag{7.22}$$

are Schwartz functions, $\tilde{f}^\flat, \tilde{g}^\flat \in \mathcal{S}(\mathbb{R}^3)$, and we can solve the Cauchy problem with this data on $\mathbb{R}^3 \times \mathbb{R}$. The question is whether the solution $\tilde{u}^\flat$ is also independent of x_3. To answer this question, note that also $\partial_{x_3}\tilde{u}^\flat$ solves the wave equation

$$-\partial_t^2(\partial_{x_3}\tilde{u}^\flat) + \triangle(\partial_{x_3}\tilde{u}^\flat) = 0 \tag{7.23}$$

and has trivial initial data for $|x_3| \leq T$. Therefore by the uniqueness theorem below (Thm. 7.4) we have that

$$\partial_{x_3}\tilde{u}^\flat(x,t) = 0 \qquad |x_3| \leq T - t\,. \tag{7.24}$$

In particular $u(x_1,x_2,t) = \tilde{u}^\flat(x_1,x_2,0,t)$ is a solution to the $d=2$-dimensional wave equation for $t \leq T$. Since T is arbitrary, we obtain the statement that solutions to the 2-dimensional wave equation can be obtained by restricting solutions to the 3-dimensional wave equation with trivially extended data. It remains to compute what form they take exactly.

Theorem 7.3. *A solution of the Cauchy problem for the wave equation in two dimensions with initial data $f, g \in \mathcal{S}(\mathbb{R}^2)$ is given by*

$$u(x,t) = \frac{\partial}{\partial t}(t\widetilde{M}_t(f)(x)) + t\widetilde{M}_t(g)(x)\,, \tag{7.25}$$

where

$$\widetilde{M}_t(f)(x) = \frac{1}{2\pi}\int_{|y|\leq 1} f(x-ty)\frac{\mathrm{d}y}{\sqrt{1-|y|^2}}\,. \tag{7.26}$$

To see this we just need to compute the restriction $M_t(F)(x_1,x_2,0)$ in the case that $F(x_1,x_2,x_3)$ is independent of x_3:

$$\begin{aligned} M_t(F)(x_1,x_2,0) =& \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} F(x_1 - t\sin\theta\cos\varphi, x_2 - t\sin\theta\sin\varphi)\sin\theta\mathrm{d}\theta\mathrm{d}\varphi \\ =& \frac{1}{2\pi}\int_0^{2\pi}\int_0^{1} f(x_1 - tr\cos\varphi, x_2 - tr\sin\varphi)\frac{r\mathrm{d}r\mathrm{d}\varphi}{\sqrt{1-r^2}} = \widetilde{M}_t(f)(x_1,x_2) \end{aligned} \tag{7.27}$$

where $f(x_1,x_2) = F(x_1,x_2,0)$.

The most important difference to the three dimensional case is that $u(x,t)$ depends on the values of f and g on the *whole disc* $B_t(x)$, *not* just its boundary, the circle of radius t centered at x.

7.4. Further comments

There are several qualitative features of solutions to the wave equation that we can read off from the formula given in Theorem 7.2:

$$\begin{aligned} u(x,t) =& \frac{\partial}{\partial t}(tM_t(f)(x)) + tM_t(g)(x) \\ =& \frac{1}{4\pi}\int_{\mathbb{S}^2} f(x-t\gamma) - t\nabla f(x-t\gamma)\cdot\gamma + tg(x-t\gamma)\mathrm{d}\sigma(\gamma) \end{aligned} \tag{7.28}$$

- The value of $u(x,t)$ depends on the values of f, its derivative ∇f, and g on the sphere $S(x,t)$ centered at x and of radius t.
- The initial data at $t=0$ near a point y can only influence u at the time t in points (x,t) near the cone $|x-y|=t$.
- Let the support of f and g be contained in $\Omega \subset \mathbb{R}^3$. For $u(x,t)\neq 0$, the point x has to lie on a sphere of radius t centered at some $y \in \Omega$. Hence the union of $S(y,t)$ over $y \in \Omega$ contains the support of u at time t.
- The support of u at time t is always contained in the "envelope" of the spheres $S(y,t)$ with centers $y \in \partial\Omega$. However, it can be smaller: Take for example $\Omega = B(0,\rho)$ a ball of radius ρ. Then the sphere $S(x,t)$ for $t>\rho$ will only have a point in common

with Ω when x lies in the annulus bounded by the spheres $S(0, t-\rho)$ and $S(0, t+\rho)$. In particular, for any fixed x and all sufficiently large t, namely $t > |x| + \rho$ we have $u(x,t) = 0$. This is also referred to as **strong Huygens principle**.

- While the support of the solution with initial data of compact support expands, the solutions *decays in time*. This is qualitatively the meaning of **dispersion**.

  Assume f, and g are supported in $B(0,\rho)$. We see that contributions to the integral (7.28), which we now write as

$$u(x,t) = \frac{1}{4\pi t^2} \int_{S(x,t)} f(y) + \nabla f(y) \cdot (y - x) \cdot + t g(y) \mathrm{d}\sigma(y) \tag{7.29}$$

  arise only from that portion of $S(x,t)$ that intersect $B(0,\rho)$. The area of this intersection cannot be larger than the area of $\partial B(0,\rho)$, so

$$|S(x,r) \cap B(0,\rho)| \leq 4\pi\rho^2 \,. \tag{7.30}$$

  Therefore

$$|u(x,t)| \leq \frac{\rho^2}{t} \times \Big(\sup |f| + \sup |\nabla f| + \sup |g| \Big) \,. \tag{7.31}$$

- The formula (7.28) also indicates that *the solution u can be less regular than the initial data.* There is a possible loss of one order of differentiability: Namely $u(\cdot, 0) \in \mathrm{C}^k$, $\partial_t u(\cdot, 0) \in \mathrm{C}^{k+1}$ initially, only guarantees that $u(\cdot, t) \in \mathrm{C}^{k-1}$, $\partial_t u(\cdot, t) \in \mathrm{C}^k$ at a later time. This can be understood in terms of a **focussing effect**, that is only present for $d > 1$. For example the second derivative could become infinite at some time $t > 0$, even though they are bounded for $t = 0$. Note that while this behaviour may occur *pointwise*, the L^2-based *energy norm* of (7.8) remains bounded, in fact conserved as we have seen in (7.9).

Problems

1. For the formulation of the following uniqueness property of solutions to the wave equation the following terminology is relevant: Let $B(x_0, r_0)$ be the closed ball of radius r_0 centered at x_0 in $\mathbb{R}^d$. Then the **domain of dependence** of the ball $B(x_0, r_0)$ (viewed as a subset of $\mathbb{R}^d \times \{0\}$) is the following subset of $\mathbb{R}^d \times \mathbb{R}$:

$$\mathcal{D}(B(x_0, r_0)) = \Big\{ (t,x) : 0 \leq t \leq r_0, |x - x_0| \leq r_0 - t \Big\} \,. \tag{7.32}$$

   Sketch the this domain.
   Theorem 7.4 (Finite speed of propagation)**.** *Suppose that $u(t,x)$ is C^2 function on $\mathbb{R}^d \times \mathbb{R}$ that solves the wave equation (7.1), and suppose that for some $x_0 \in \mathbb{R}^d$, and $r_0 > 0$,*

$$u(0,x) = \partial_t u(0,x) = 0 \qquad : \textit{for all} \quad x \in B(x_0, r_0) \,. \tag{7.33}$$

   Then $u(t,x) = 0$ for all $(t,x) \in \mathcal{D}(B(x_0, r_0))$.

Prove this theorem in three steps:

a) Let

$$B_t(x_0, r_0) = \{x : |x - x_0| \leq r_0 - t\} \tag{7.34}$$

be the intersection of $\mathcal{D}(B(x_0, r_0))$ with $\{t\} \times \mathbb{R}^d$, and consider the energy

$$E(t) = \int_{B_t(x_0,r_0)} e(t,x)\mathrm{d}x \tag{7.35}$$

where e denotes the energy density

$$e(t,x) = \frac{1}{2}\Big((\partial_t u)^2 + |\nabla u|^2\Big)(t,x) \tag{7.36}$$

and $\nabla u = (\partial_{x^1} u, \dots, \partial_{x^d} u)$ is the gradient on $\mathbb{R}^d$.

Show that

$$E'(t) = \int_{B_t(x_0,r_0)} \partial_t e \mathrm{d}x - \int_{\partial B_t(x_0,r_0)} e \mathrm{d}\sigma \tag{7.37}$$

where $\mathrm{d}\sigma$ denotes the surface element on the sphere $\partial B_t(x_0, r_0)$ of radius r_0 centered at x_0 in $\{t\} \times \mathbb{R}^d$.

b) Compute $\partial_t e$, and show that for any solution to the wave equation,

$$\int_{B_t(x_0,r_0)} \partial_t e \mathrm{d}x = \int_{\partial B_t(x_0,r_0)} \partial_t u \; \nabla u \cdot n \mathrm{d}\sigma \tag{7.38}$$

where n is the unit normal to $\partial B_t(x_0, r_0)$.

c) Use the Cauchy Schwarz inequality to conclude that $E'(t) \leq 0$.

Hint: Use the divergence theorem in a) and b).

Decay in time using Fourier representation

Further Reading

(Luk, *Introduction to nonlinear wave equations*, Section 2)

We have seen in (7.31) that solutions to the linear wave equation, arising from compactly supported initial data, decay in time. In this note, we shall try to explain this decay rate in $d = 3$ dimensions, using the Fourier representation formula, for initial data which is rapidly decaying.

According to Theorem 7.1 the terms that are contributing to $u(x,t)$ are integrals of the form

$$\int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi i(\xi\cdot x \pm |\xi| t)} \mathrm{d}\xi\,, \qquad \int_{\mathbb{R}^d} \hat{g}(\xi) \frac{e^{2\pi i(\xi\cdot x \pm |\xi| t)}}{2\pi i |\xi|} \mathrm{d}\xi\,. \tag{7.1}$$

We may introduce spherical coordinates,

$$\xi = (\rho \sin\theta \cos\varphi, \rho \sin\theta \sin\varphi, \rho\cos\theta) \tag{7.2}$$

and let us consider for example the latter interal

$$I = \int_{\mathbb{R}^d} \hat{g}(\xi) \frac{e^{2\pi i(\xi\cdot x + |\xi| t)}}{2\pi i |\xi|} \mathrm{d}\xi = \int_0^\infty \int_0^\pi \int_0^{2\pi} \hat{g}(\xi) e^{2\pi i \xi\cdot x} \frac{e^{2\pi i \rho t}}{2\pi i \rho} \sin(\theta) \rho^2 \mathrm{d}\varphi \mathrm{d}\theta \mathrm{d}\rho \tag{7.3}$$

and we can integrate by parts in ρ, using that

$$\frac{1}{2\pi i t} \frac{\partial}{\partial \rho} e^{2\pi i \rho t} = e^{2\pi i \rho t}\,. \tag{7.4}$$

This yields

$$I = -\frac{1}{(2\pi i)^2 t} \int_0^\infty \int_0^\pi \int_0^{2\pi} e^{2\pi i \rho t} \frac{\partial}{\partial \rho}\Big(\hat{g}(\xi) e^{2\pi i \xi\cdot x} \rho \sin(\theta)\Big) \mathrm{d}\varphi \mathrm{d}\theta \mathrm{d}\rho \tag{7.5}$$

without any boundary terms at $\rho = \infty$, nor at $\rho = 0$. Moreover

$$\frac{\partial}{\partial \rho}\Big(\hat{g}(\xi) e^{2\pi i \xi\cdot x} \rho \sin(\theta)\Big) = \Big(\nabla \hat{g} \cdot \xi + \hat{g}(\xi) 2\pi i \xi \cdot x + \hat{g}(\xi)\Big) e^{2\pi i \xi \cdot x} \sin(\theta) \tag{7.6}$$

Note that the second term depends on x, so it is not possible to obtain a bound uniform in x in this way. Furthermore, one cannot integrate by parts indefintely in this way, already the next integration by parts produces a nonvanishing boundary term at the origin:

$$\begin{aligned} I =& -\frac{1}{(2\pi i)^2 t}\int_0^\infty\int_0^\pi\int_0^{2\pi} e^{2\pi i\rho t}\Big(\nabla\hat{g}\cdot\xi+\hat{g}(\xi)2\pi i\xi\cdot x+\hat{g}(\xi)\Big)e^{2\pi i\xi\cdot x}\sin(\theta)\mathrm{d}\varphi\mathrm{d}\theta\mathrm{d}\rho \\ =&\frac{1}{(2\pi i)^3t^2}\int e^{2\pi i\rho t}\frac{\partial}{\partial\rho}\Big[\Big(\nabla\hat{g}\cdot\xi+\hat{g}(\xi)2\pi i\xi\cdot x+\hat{g}(\xi)\Big)e^{2\pi i\xi\cdot x}\sin(\theta)\Big]\mathrm{d}\varphi\mathrm{d}\theta\mathrm{d}\rho \\ &+\frac{4\pi}{(2\pi i)^3t^2}\hat{g}(0) \end{aligned} \tag{7.7}$$

This type of argument leads to the following estimate:

Proposition 7.1. *Let u be the solution to the Cauchy problem in $d=3$ dimensions, with initial data $f,g\in\mathcal{S}(\mathbb{R}^3)$. Then for any $R>0$, there exists a constant $C>0$ such that*

$$|u(x,t)|\leq\frac{C}{t^2}\qquad(|x|\leq R)\,. \tag{7.8}$$

A slightly different way to carry out these integrations by parts, which yields a uniform bound in x, is the following: First by a rotation of the coordinate axes let us assume that $x=(0,0,r)$, and let us introduce cylindrical coordinates such that

$$\xi=(\rho\cos(\phi),\rho\sin(\phi),\zeta)\qquad|\xi|^2=\rho^2+\zeta^2 \tag{7.9}$$

In these coordinates we have

$$\frac{\partial}{\partial\rho}e^{2\pi i|\xi|t}=e^{2\pi i|\xi|t}\cdot 2\pi it\frac{\rho}{|\xi|} \tag{7.10}$$

and crucially

$$\frac{\partial}{\partial\rho}e^{2\pi i\xi\cdot x}=\frac{\partial}{\partial\rho}e^{2\pi i\zeta r}=0\,, \tag{7.11}$$

which now leads to

$$\begin{aligned} I =& \int_{\mathbb{R}^d}\hat{g}(\xi)\frac{e^{2\pi i(\xi\cdot x+|\xi|t)}}{2\pi i|\xi|}\mathrm{d}\xi \\ =&\frac{1}{(2\pi i)^2t}\int_0^\infty\int_0^{2\pi}\int_{-\infty}^\infty\hat{g}(\xi)e^{2\pi i\zeta r}\frac{1}{\rho}\frac{\partial}{\partial\rho}\Big(e^{2\pi i|\xi|t}\Big)\rho\mathrm{d}\zeta\mathrm{d}\phi\mathrm{d}\rho \\ =&-\frac{2\pi}{(2\pi i)^2t}\int_{-\infty}^\infty\hat{g}(0,0,\zeta)e^{2\pi i(\zeta r+|\zeta|t)}\mathrm{d}\zeta \\ &-\frac{1}{(2\pi i)^2t}\int_0^\infty\int_0^{2\pi}\int_{-\infty}^\infty\nabla\hat{g}(\xi)\cdot(\cos\phi,\sin\phi,0)e^{2\pi i(\zeta r+|\xi|t)}\mathrm{d}\zeta\mathrm{d}\phi\mathrm{d}\rho\,. \end{aligned} \tag{7.12}$$

In conclusion:

Proposition 7.2. *Let u be the solution to the Cauchy problem in $d=3$ dimensions, with initial data $f,g\in\mathcal{S}(\mathbb{R}^3)$. Then*

$$|u(x,t)|\leq\frac{C}{t} \tag{7.13}$$

where C only depends on the initial data.

Duhamel's principle

Further Reading

(John, *Partial differential equations*, Chapter 5.1 (c)).

In this note we want to show that using Theorem 7.2 we can also solve the inhomogeneous problem,

$$\Box u = F \tag{7.1}$$

for any given "source" function $F(t,x)$, with prescribed data

$$u(0,x) = f(x)\,, \qquad \partial_t u(0,x) = g(x)\,. \tag{7.2}$$

In view of the linearity of the problem it sufficies to solve (7.1) with trivial initial data

$$u(0,x) = 0 \qquad \partial_t u(0,x) = 0\,, \tag{7.3}$$

and we will assume for simplicity that $F \in \mathcal{S}(\mathbb{R}^{3+1})$.

The idea of our approach is similar to the "method of variation of constants" for ODEs, in the sense that solutions to the inhomogeneous equation can be written as a superposition of solutions to the homogeneous equation, with variable coefficients, to be determined.

Here we make the following ansatz for solutions to the *inhomogeneous* problem (7.1):

$$u(t,x) = \int_0^t U_s(t,x)\mathrm{d}s \tag{7.4}$$

where U_s, for each s, is a solution to the *homogeneous problem*:

$$\Box U_s = 0 \qquad t \geq s \tag{7.5}$$

with initial conditions to be determined.

We compute

$$\partial_t u(t,x) = U_t(t,x) + \int_0^t \partial_t U_s(t,x)\mathrm{d}s \tag{7.6}$$

and will choose

$$U_s(s,x) = 0\,. \tag{7.7}$$

Then we compute further that

$$\begin{aligned}\partial_t^2 u(t,x) =& \partial_t U_t(t,x) + \int_0^t \partial_t^2 U_s(t,x)\mathrm{d}s \\ =& -F(t,x) + \int_0^t \triangle U_s(t,x)\mathrm{d}s = -F(t,x) + \triangle u(t,x)\end{aligned} \tag{7.8}$$

and can thus arrage for u as defined in (7.4) to be a solution to (7.1), provided we impose

$$\partial_t U_s(s,x) = -F(s,x)\,. \tag{7.9}$$

In other words, the family of functions U_s are chosen to be the solutions of

$$\Box U_s = 0 \qquad (t \geq s) \qquad u(s,x) = 0 \qquad \partial_t u(s,x) = -F(s,x)\,. \tag{7.10}$$

Setting

$$V_s(t,x) = U_s(t+s,x)\,, \tag{7.11}$$

the functions V_s solve

$$\Box V_s = 0 \quad (t \geq 0) \qquad V_s(0,x) = 0 \qquad \partial_t V_s(0,x) = -F(s,x)\,. \tag{7.12}$$

According to Theorem 7.2, as expressed for instance in (7.28), we have

$$V_s(t,x) = -\frac{1}{4\pi}\int_{\mathbb{S}^2} tF(s,x-t\gamma)\mathrm{d}\sigma(\gamma) = -\frac{1}{4\pi t}\int_{|y-x|=t} F(s,y)\mathrm{d}S(y) \tag{7.13}$$

and therefore

$$\begin{aligned} u(t,x) &= \int_0^t U_s(t,x)\mathrm{d}s = \int_0^t V_s(t-s,x)\mathrm{d}s \\ &= -\frac{1}{4\pi}\int_0^t \frac{1}{t-s}\int_{|y-x|=t-s} F(s,y)\mathrm{d}S(y)\mathrm{d}s\,. \end{aligned} \tag{7.14}$$

Remark 7.1. Thus the solution u at a point (t,x) depends only on the values of F on the truncated backwards lightcone

$$C_0^-(t,x) = \Big\{(s,y) : |y-x| = t-s, 0 \leq s \leq t\Big\}\,. \tag{7.15}$$

This means that the solution can be expressed as

$$u(t,x) = -\frac{1}{4\pi}\int_{C_0^-(t,x)} F\,\mathrm{d}\mu_C \tag{7.16}$$

where $\mathrm{d}\mu_C$ is a measure supported only on the lightcone. In fact,

$$\begin{aligned} \frac{1}{4\pi}\int_{C_0^-(t,x)} F\,\mathrm{d}\mu_C &= \frac{1}{4\pi}\int_0^r \frac{1}{r}\int_{|y-x|=r} F(t-r,y)\mathrm{d}S(y)\mathrm{d}r = \\ = \frac{1}{4\pi}\int_0^r \frac{1}{r}\int_{\mathbb{S}^2} F(t-r,x+r\gamma)r^2\mathrm{d}\sigma(\gamma)\mathrm{d}r &= \frac{1}{4\pi}\int_{\substack{y\in\mathbb{R}^3\\|y|\leq t}} F(t-|y-x|,y)\frac{\mathrm{d}y}{|y-x|} \end{aligned} \tag{7.17}$$

If we remove the truncation, and integrate on the whole backward lightcone

$$C^-(t,x) = \Big\{(s,y) : |y-x| = s-t, s \leq t\Big\} \tag{7.18}$$

then we obtain the "retarded solution",

$$u(t,x) = -\int_{C^-(t,x)} F\mathrm{d}\mu_C = -\frac{1}{4\pi}\int_{\mathbb{R}^3} F(t-|y-x|,y)\frac{\mathrm{d}y}{|y-x|} \tag{7.19}$$

namely the solution to (7.1) with trivial data at $t \to -\infty$.

Lecture 8.

The Radon transform and its applications

Further Reading

(Stein and Shakarchi, *Fourier analysis*, Chapter 6, Section 5).

A problem that arises in *medical imaging* is the following. Consider an object which occupies a certain domain $\mathcal{O} \subset \mathbb{R}^3$ in space. We are interested in the **density** of some of its constituents, and we have a means of "scanning" the object which provides some information of the density. For example, the *X-ray* is a beam that is sent through the object, and we can measure the brightness of the beam before and after it passes through the object. Conceptually each beam is associated with a line $L \subset \mathbb{R}^3$, and the measurement is related to the quantity:

$$X(\rho)(L) = \int_L \rho \tag{8.1}$$

The problem is if it is possible to *reconstruct* the function $\rho : \mathcal{O} \to [0, \infty)$ from knowledge of the function $X(\rho)$. Now $X(\rho)$ can be thought of as a function of four variables (it is a function on the space of lines in $\mathbb{R}^3$), and ρ is a function of three variables. This suggests that it should be possible to determine ρ from $X(\rho)$, and in fact we will show this ρ can be determined with *less* information: Instead of associating to each line L a number $\int_L \rho$, consider the quantity

$$\mathcal{R}(\rho)(\mathcal{P}) = \int_{\mathcal{P}} \rho \tag{8.2}$$

which denotes the integral of the function ρ over a given *plane* $\mathcal{P} \subset \mathbb{R}^3$. More precisely, let us parametrize the planes in $\mathbb{R}^3$ by a unit vector $\gamma \in \mathbb{S}^2$ (the normal to the plane), and a number $t \in \mathbb{R}$ (its distance to the origin):

$$\mathcal{P}_{t,\gamma} = \{x \in \mathbb{R}^3 : x \cdot \gamma = t\} \tag{8.3}$$

Given a function $\rho \in \mathcal{S}(\mathbb{R}^3)$, we define

$$\int_{\mathcal{P}_{\gamma,t}} \rho = \int_{\mathbb{R}^2} f(t\gamma + u_1 e_2 + u_2 e_2) \mathrm{d}u_1 \mathrm{d}u_2 \tag{8.4}$$

where e_1, e_2 are chosen such that (e_1, e_2, γ) is an orthonormal basis for $\mathbb{R}^3$.

Exercise 8.1. Verify that this is well-defined, namely the integral is independent of the the choice of basis. Moreover, convince yourself that for any $\gamma \in \mathbb{S}^2$,

$$\int_{\mathbb{R}^3} \rho = \int_{-\infty}^{\infty} \Big(\int_{\mathcal{P}_{t,\gamma}} \rho \Big) \mathrm{d}t \tag{8.5}$$

The function $\mathcal{R}(\rho)(t,\gamma) = \mathcal{R}(\rho)(\mathcal{P}_{t,\gamma})$ is called the **Radon transform** of ρ . Note that knowledge of the **X-ray transform** $X(\rho)$ determines the Radon transform, because integrals over planes can be expressed in terms of integrals over lines.

Note that with this parametrization $\mathcal{R}(\rho)$ can be thought of as a function on the cylinder $\mathbb{R} \times \mathbb{S}^2$. The relevant class of functions will be those that are rapidly decreasing in t, uniformly in γ. In other words, let us define $\mathcal{S}(\mathbb{R} \times \mathbb{S}^2)$ to be the space of continuous functions $F(t,\gamma)$ which are indefinitely differentiable in t, and satisfy

$$\sup |t|^k \Big| \frac{\mathrm{d}^l F}{\mathrm{d}t^l} \Big| < \infty \,, \qquad \text{for all } k, l \geq 0 \,. \tag{8.6}$$

Instead of immediately trying to solve the *reconstruction problem*, we could first address the *uniqueness problem*: Suppose we know $\mathcal{R}(f) = \mathcal{R}(g)$. Is it then true that $f = g$?

Lemma 8.1. *If $f \in \mathcal{S}(\mathbb{R}^3)$, then $\mathcal{R}(f)(t,\gamma) \in \mathcal{S}(\mathbb{R})$ for each fixed γ. Moreover the Fourier transform of $\mathcal{R}(f)(\cdot,\gamma)$, for fixed γ, is*

$$\widehat{\mathcal{R}(f)}(s,\gamma) = \hat{f}(s\gamma) \,. \tag{8.7}$$

Proof. By definition

$$\begin{aligned} \widehat{\mathcal{R}(f)}(s,\gamma) &= \int_{-\infty}^{\infty} \Big(\int_{\mathcal{P}_{t,\gamma}} f \Big) e^{-2\pi i s t} \mathrm{d}t \\ &= \int_{-\infty}^{\infty} \int_{\mathbb{R}^2} f(t\gamma + u_1 e_1 + u_2 e_2) e^{-2\pi i s t} \mathrm{d}u_1 \mathrm{d}u_2 \mathrm{d}t \end{aligned} \tag{8.8}$$

and we can write $st = s\gamma \cdot (t\gamma + u)$, because $u = u_1 e_1 + u_2 e_2$ is orthogonal to γ. Therefore this integral equals

$$\int_{-\infty}^{\infty} \int_{\mathbb{R}^2} f(t\gamma + u) e^{-2\pi i s\gamma\cdot(t\gamma+u)} \mathrm{d}u \mathrm{d}t = \hat{f}(s\gamma) \tag{8.9}$$

after a rotation of the basis (γ, e_1, e_2) to the standard basis in $\mathbb{R}^3$. □

As a corollary we can answer the uniqueness question in the affirmative: If $\mathcal{R}(f) = 0$ then $\hat{f} = 0$ which implies that $f = 0$ by the Fourier inversion theorem, for $f \in \mathcal{S}(\mathbb{R}^3)$.

Let us now to turn to the *reconstruction problem*, namely the question if f can be recovered from its Radon transform.

The Radon transform $\mathcal{R}$ sends functions on $\mathbb{R}^3$ to functions on the set of planes in $\mathbb{R}^3$. We define the **dual Radon transform** $\mathcal{R}^*$, which sends functions on the set of planes to functions on $\mathbb{R}^3$, by

$$\mathcal{R}^*(F)(x) = \int_{\mathbb{S}^2} F(x\cdot\gamma, \gamma) \mathrm{d}\sigma(\gamma) \,. \tag{8.10}$$

Here F is viewed as a function on $\mathbb{R} \times \mathbb{S}^2$ via the identification $F(t,\gamma) = F(\mathcal{P}_{t,\gamma})$. Note that given x and γ, the plane $\mathcal{P}_{x\cdot\gamma,\gamma}$ contains x, so the above is an integral over all planes that pass through $x \in \mathbb{R}^3$.

Remark 8.1. The transformation $\mathcal{R}^*$ is *dual* to $\mathcal{R}$ in the sense that

$$(\mathcal{R}(f), F) = (f, \mathcal{R}^*(F)) \tag{8.11}$$

where the Hermitian inner product on the right is on $\mathcal{S}(\mathbb{R}^3)$, and on $\mathcal{S}(\mathbb{R} \times \mathbb{S}^2)$ on the left hand side.

Theorem 8.2. *If $f \in \mathcal{S}(\mathbb{R}^3)$, then*

$$\triangle(\mathcal{R}^*\mathcal{R}(f)) = -8\pi^2 f\,. \tag{8.12}$$

In particular, the theorem provides a *formula* for f in terms of its Radon transform.

Proof. In view of the previous Lemma,

$$\mathcal{R}(f)(t,\gamma) = \int_{-\infty}^{\infty} \hat{f}(s\gamma)e^{2\pi i s t}\mathrm{d}s \tag{8.13}$$

and hence

$$\mathcal{R}^*(\mathcal{R}(f))(x) = \int_{\mathbb{S}^2}\int_{-\infty}^{\infty} \hat{f}(s\gamma)e^{2\pi i s x\cdot\gamma}\mathrm{d}s\mathrm{d}\sigma(\gamma)\,. \tag{8.14}$$

We can now compute that

$$\begin{aligned}\triangle(\mathcal{R}^*\mathcal{R}(f))(x) &= \int_{\mathbb{S}^2}\int_{-\infty}^{\infty} \hat{f}(s\gamma)(2\pi i s)^2 e^{2\pi i s x\cdot\gamma}\mathrm{d}s\mathrm{d}\sigma(\gamma)\\ &= -8\pi^2\int_0^{\infty}\int_{\mathbb{S}^2} \hat{f}(s\gamma)e^{2\pi i x\cdot s\gamma}s^2\mathrm{d}\sigma(\gamma)\mathrm{d}s = -8\pi^2 f(x)\,.\end{aligned} \tag{8.15}$$

□

Problems

1. Establish the identity (8.11). In other words prove that

$$\int_{\mathbb{R}}\int_{\mathbb{S}^2} \mathcal{R}(f)(t,\gamma)\overline{F(t,\gamma)}\mathrm{d}\sigma(\gamma)\mathrm{d}t = \int_{\mathbb{R}^3} f(x)\overline{\mathcal{R}^*(F)(x)}\mathrm{d}x \tag{8.16}$$

for all $f \in \mathcal{S}(\mathbb{R}^3)$, and $F \in \mathcal{S}(\mathbb{R} \times \mathbb{S}^2)$, and

$$\mathcal{R}(f) = \int_{\mathcal{P}_{t,\gamma}} f\,, \tag{8.17}$$

and

$$\mathcal{R}^*(F)(x) = \int_{\mathbb{S}^2} F(x\cdot\gamma,\gamma)\mathrm{d}\sigma(\gamma)\,. \tag{8.18}$$

Hint: Consider the integral

$$\int\int\int f(t\gamma + u_1e_1 + u_2e_2)\overline{F(t,\gamma)}\mathrm{d}t\mathrm{d}\sigma(\gamma)\mathrm{d}u_1\mathrm{d}u_2\,. \tag{8.19}$$

Integrating in u first gives the left hand side, while integrating in t and u, and setting $x = t\gamma + u_1e_2 + u_2e_2$, gives the right hand side.

Supplementary Problem

The following problems relate to the **X-ray transform**, namely Radon transform in dimesion $d = 2$, where the solution to the reconstruction problem is *more complicated* than in $d = 3$ dimensions.

1. Recall that the X-ray transform of a function ρ is given by
$$X(\rho)(L) = \int_L \rho \tag{8.20}$$
where L is a line in $\mathbb{R}^2$.
We can parametrize this as follows: For each (t, θ) with $t \in \mathbb{R}$ and $|\theta| \leq \pi$, let $L_{t,\theta}$ denote the line in the xy-plane given by
$$x \cos\theta + y \sin\theta = t\,. \tag{8.21}$$
This is the line perpendicular to the direction $(\cos\theta, \sin\theta)$ at "distance" $t \in \mathbb{R}$ from the origin. Show that for $f \in \mathcal{S}(\mathbb{R}^2)$ the X-ray transform of f is then parametrized by
$$X(f)(t, \theta) = \int_{L_{t,\theta}} f = \int_{-\infty}^{\infty} f(t\cos\theta + u\sin\theta, t\sin\theta - u\cos\theta)\mathrm{d}u\,. \tag{8.22}$$

2. Show that if $f \in \mathcal{S}(\mathbb{R}^2)$ and $X(f) = 0$, then $f = 0$, by taking the Fourier transform in one variable.

3. For $F \in \mathcal{S}(\mathbb{R} \times \mathbb{S}^1)$, define the **dual X-ray transform** $X^*(F)$ by integrating Fover all lines that pass through the point (x, y), namely all lines $L_{t,\theta}$ with $x\cos\theta + y\sin\theta = t$:
$$X^*(F)(x, y) = \int F(x\cos\theta + y\sin\theta, \theta)\mathrm{d}\theta\,. \tag{8.23}$$
Check that X^* is in fact dual to X, in the sense that, if $f \in \mathcal{S}(\mathbb{R}^2)$ and $F \in \mathcal{S}(\mathbb{R}\times\mathbb{S}^1)$, then
$$\int\int X(f)(t,\theta)\overline{F(t,\theta)}\mathrm{d}t\mathrm{d}\theta = \int\int f(x,y)\overline{X^*(F)(x,y)}\mathrm{d}x\mathrm{d}y\,. \tag{8.24}$$

4. For every real number $a > 0$, define the operator $(-\triangle)^a$ by the formula
$$(-\triangle)^a f(x) = \int_{\mathbb{R}^d} (2\pi|\xi|)^{2a} \hat{f}(\xi) e^{2\pi i\xi\cdot x}\mathrm{d}\xi \tag{8.25}$$
whenever $f \in \mathcal{S}(\mathbb{R}^d)$.
   a) Check that when a is a natural number then $(-\triangle)^a = (-\triangle) \circ \cdots \circ (-\triangle)$ (a-times) agrees with the usual definition.
   b) Verify that $(-\triangle)^a f$ is indefinitely differentiable.

5. Show that reconstruction formula for the X-ray transform reads
$$(-\triangle)^{1/2} X^*(X(f)) = 4\pi f\,. \tag{8.26}$$

Part IV.

Hilbert spaces and Weak Solutions

Review: The Hilbert space $\mathrm{L}^2(\mathbb{R}^d)$

Recommended Reading

(Stein and Shakarchi, *Real analysis*, Chapter 4, Sections 1-3).

We will first discuss the prime example of a Hilbert space, namely the space of **square integrable functions**, and then discuss Hilbert spaces more generally.

9.1. The space $\mathrm{L}^2(\mathbb{R}^d)$

The space $\mathrm{L}^2(\mathbb{R}^d)$ is the collection of all square integrable functions on $\mathbb{R}^d$, and consists of all complex-valued *measurable* functions f on $\mathbb{R}^d$ that satisfy

$$\int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}x < \infty \tag{9.1}$$

The space of square integrable functions is naturally equipped with the inner product

$$(f, g) = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\mathrm{d}x \tag{9.2}$$

and the corresponding **norm** on $\mathrm{L}^2(\mathbb{R}^d)$ is given by

$$\|f\|_2 = \|f\|_{\mathrm{L}^2(\mathbb{R}^d)} = (f, f)^{1/2} = \Big(\int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}x\Big)^{1/2}. \tag{9.3}$$

The integral that appears here is with respect to the *Lebesgue measure*, and consequently the condition $\|f\| = 0$ only implies $f(x) = 0$ *almost everywhere*. Therefore functions that agree almost everywhere should be identified, and $\mathrm{L}^2(\mathbb{R}^d)$ should be defined as the space of equivalence classes under this identification, but in practice this distinction is conveniently forgotten, and we think of elements in $\mathrm{L}^2(\mathbb{R}^d)$ as *functions*.

In the above definition of the inner product we need to know that $f\overline{g}$ is integrable on $\mathbb{R}^d$ whenever f and g are square-integrable. This and other basic properties are gathered in the following:

Proposition 9.1. *The space* $\mathrm{L}^2(\mathbb{R}^d)$ *has the following properties:*

1. $\mathrm{L}^2(\mathbb{R}^d)$ *is a vectorspace.*

2. *The Cauchy-Schwarz inequality holds:* $|(f, g)| \leq \|f\|\|g\|$

3. *If* $g \in \mathrm{L}^2(\mathbb{R}^d)$ *is fixed, then the map* $f \mapsto (f, g)$ *is linear in* f*, and also* $(f, g) = \overline{(g, f)}$

4. *The triangle inequality holds* $\|f + g\| \leq \|f\| + \|g\|$*.*

We turn our attention to the notion of a limit in the space $\mathrm{L}^2(\mathbb{R}^d)$. The norm induces a **metric** d on $\mathrm{L}^2(\mathbb{R}^d)$:

$$d(f, g) = \|f - g\|_2 \tag{9.4}$$

A sequence of functions f_n is a **Cauchy sequence** if $d(f_n, f_m)$ as $n, m \to \infty$. Moreover a sequence f_n is said to **converge** to $f \in \mathrm{L}^2(\mathbb{R}^d)$ if $d(f_n, f) \to 0$ as $n \to \infty$.

The central property that motivates Lebesgue's theory of integration is that these spaces — namely the space of *integrable functions* $\mathrm{L}^1(\mathbb{R}^d)$ and here $\mathrm{L}^2(\mathbb{R}^d)$, and more generally $\mathrm{L}^p(\mathbb{R}^d)$ — are *complete*, meaning that every Cauchy sequence in $\mathrm{L}^2(\mathbb{R}^d)$ converges to a function in $\mathrm{L}^2(\mathbb{R}^d)$. This is *not* true in the context of Riemann integrable functions.

Theorem 9.2. *The space* $\mathrm{L}^2(\mathbb{R}^d)$ *is complete.*

The proof uses the main covergence theorems of Lebesgue's integration theory: *monotone convergence*, and *dominated covergence*.

Another property worth pointing out is that:

Theorem 9.3. *The space* $\mathrm{L}^2(\mathbb{R}^d)$ *is* ***separable****.*

This means that there exists a *countable collection* of functions $\{f_k\}$ in $\mathrm{L}^2(\mathbb{R}^d)$ such that their linear combinations are **dense** in $\mathrm{L}^2(\mathbb{R}^d)$.

9.2. Hilbert spaces

More generally, the properties we have seen above for $\mathrm{L}^2(\mathbb{R}^d)$ are shared by all **Hilbert spaces** $\mathcal{H}$:

1. $\mathcal{H}$ is a vector space of $\mathbb{C}$ or $\mathbb{R}$

2. $\mathcal{H}$ is equipped with an inner product $(\cdot, \cdot)$ so that
   a) $f \to (f, g)$ is linear on $\mathcal{H}$ for every fixed $g \in \mathcal{H}$
   b) $(f, g) = \overline{(g, f)}$
   c) $(f, f) \geq 0$ for all $f \in \mathcal{H}$

   The induced norm is $\|f\| = (f, f)^{1/2}$

3. $\|f\| = 0$ if and only if $f = 0$

4. The Cauchy-Schwarz and triangle inequalities hold

$$|(f, g)| \leq \|f\|\|g\| \qquad \|f + g\| \leq \|f\| + \|g\| \tag{9.5}$$

   for all $f, g \in \mathcal{H}$.

5. $\mathcal{H}$ is complete with respect to $d(f,g) = \|f - g\|$

6. $\mathcal{H}$ is separable.

Remark 9.1. In the context of Hilbert spaces we write $\lim_{n\to\infty} f_n = f$ to mean that $\lim_{n\to\infty} \|f_n - f\| = 0$.

Apart from the square-integrable functions on $\mathbb{R}^d$, another important example are the square-summable sequences.

Example 9.1. The space

$$l^2(\mathbb{Z}) = \Big\{a = (\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots) : a_i \in \mathbb{C}, \sum_{n=-\infty}^{\infty} |a_n|^2 < \infty\Big\} \tag{9.6}$$

is a Hilbert space with the inner product in $l^2(\mathbb{Z})$ defined by

$$(a,b) = \sum_{k=-\infty}^{\infty} a_k \overline{b_k}\,. \tag{9.7}$$

9.2.1. Orthogonality

The inner product allows us to define a notion of orthogonality: Two elements $f, g \in \mathcal{H}$ are **perpendicular** if

$$(f,g) = 0\,, \tag{9.8}$$

and we often write $f \perp g$.

Proposition 9.4. *If $f \perp g$, then $\|f+g\|^2 = \|f\|^2 + \|g\|^2$.*

A countable collection $\{e_1, e_2, \ldots\}$ of vectors in $\mathcal{H}$ is called **orthonormal** if

$$(e_k, e_l) = \delta_{kl}\,. \tag{9.9}$$

Moreover, an orthonormal subset $\{e_k\}_{k=1}^{\infty}$ of $\mathcal{H}$ is called an **orthonormal basis** for $\mathcal{H}$ if finite linear combinations of these elements are *dense* in $\mathcal{H}$, or equivalently:

Theorem 9.5. *The following properties of an orthonormal set $\{e_k\}$ are equivalent.*

1. *Finite linear combinations of elements in $\{e_k\}$ are dense in $\mathcal{H}$.*

2. *If $f \in \mathcal{H}$ and $(f, e_j) = 0$ for all j, then $f = 0$.*

3. *If $f \in \mathcal{H}$, and $S_N(f) = \sum_{k=1}^{N} a_k e_k$ where $a_k = (f, e_k)$, then $S_N(f) \to f$ as $N \to \infty$.*

4. *If $a_k = (f, e_k)$ then $\|f\|^2 = \sum_{k=1}^{\infty} |a_k|^2$. **(Parseval's identity)***

Any Hilbert space has an orthonormal basis, which can be constructed by Gram-Schmidt starting from a countable collection which is dense and given because $\mathcal{H}$ is separable.

Example 9.2. It is useful to keep in mind the example of **Fourier series**, where

$$\mathcal{H} = \mathrm{L}^2([-\pi,\pi]) \tag{9.10}$$

with inner product

$$(f,g) = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\mathrm{d}x\,. \tag{9.11}$$

As we have seen the functions

$$e_n(x) = e^{inx} \qquad (n\in\mathbb{Z}) \tag{9.12}$$

are orthonormal and the coefficients

$$a_k = (f,e_k) = \int_{-\pi}^{\pi} f(x)e^{-inx}\mathrm{d}x \tag{9.13}$$

are precisely the Fourier coefficients. Hence *2.* of Theorem 9.5 above looks like *uniqueness* of Fourier series: If $a_n = 0$ for all n, then $f = 0$. We have already referred to this statement in Lecture 3 for *continuous* functions, but it is needed here more generally for functions $f \in \mathrm{L}^2([-\pi,\pi]) \subset \mathrm{L}^1([-\pi,\pi])$, namely *integrable functions.*

Theorem 9.6. *Suppose $f \in \mathrm{L}^1([-\pi,\pi])$. If $a_n = 0$ for all n then $f(x) = 0$ almost everywhere.*

In view of Theorem 9.5 this implies that the Fourier series converges in L^2,

$$\|f - S_N(f)\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - S_N(x)|^2\mathrm{d}x \to 0 \qquad (N\to\infty) \tag{9.14}$$

where $S_N(f) = \sum_{|n|\leq N} a_n e^{inx}$, and Parseval's identity holds.

9.2.2. Unitary mappings

Definition 9.1. A linear mapping $U : \mathcal{H} \to \mathcal{H}'$ between two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ is **unitary** if

1. U is a bijection
2. $\|Uf\|_{\mathcal{H}'} = \|f\|_{\mathcal{H}}$ for all $f \in \mathcal{H}$.

Moreover two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ are **unitarily equivalent** if there exists a unitary mapping $U : \mathcal{H} \to \mathcal{H}'$.

The reason this notion is worth recalling is that *all infinite dimensional Hilbert spaces are unitarily equivalent to* $l^2(\mathbb{Z})$, and every finite-dimensional Hilbert space over $\mathbb{C}$, or over $\mathbb{R}$, is equivalent to $\mathbb{C}^d$ (or $\mathbb{R}^d$) for some $d \in \mathbb{N}$.

Proposition 9.7. *Any two infinite-dimensional Hilbert spaces are unitarily equivalent.*

Proof. If $\{e_i\}$ is an orthonormal basis for $\mathcal{H}$ and $\{e_i'\}$ is an orthonormal basis for $\mathcal{H}'$, then define

$$Uf = \sum_i a_i e_i' \qquad f = \sum_i a_i e_i \,. \tag{9.15}$$

This mapping is linear and invertible, and by Parseval's identity

$$\|Uf\| = \sum_i |a_i|^2 = \|f\| \,. \tag{9.16}$$

□

Example 9.3. Continuing with the example of Fourier series, Example 9.2 above, we can view

$$U : f \mapsto \{a_n\} \tag{9.17}$$

as a map from $\mathcal{H} = L^2([-\pi,\pi])$ to $l^2(\mathbb{Z})$. Indeed, we have seen that for any square integrable function $f \in \mathcal{H}$, the sequence $\{a_n\}$ of Fourier coefficients is square integrable, by Parseval's identity, and the norm is presevered:

$$\|Uf\|_{l^2(\mathbb{Z})} = \sum_k |a_k|^2 = \|f\|_2 \tag{9.18}$$

This mapping is also linear, and one-to-one by Theorem 9.6. Therefore it remains to show that U is onto, which then implies that U is a unitary correspondence. Given $\{a_n\} \in l^2(\mathbb{Z})$, we have that $S_N = \sum_{|n|\leq N} a_n e_n$ is a Cauchy sequence, because

$$\|S_N - S_M\|_2 = \sum_{M\leq|n|\leq N} |a_n|^2 \to 0 \qquad (M,N\to\infty) \tag{9.19}$$

Since $\mathcal{H}$ is complete there exists $f \in \mathcal{H}$ such that $\|f - S_N\| \to 0$ as $N \to \infty$. Moreover the Fourier coefficients of f are $(f, e_k) = \lim_{N\to\infty}(S_N, e_k) = a_k$, so $Uf = \{a_n\}$.

This example is in itself a *motivation* for the space $\mathrm{L}^2([-\pi,\pi])$: While for any *Riemann*-integrable function f on the circle, the Fourier coefficients are in $l^2(\mathbb{Z})$, and Parseval's identity holds:

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = \frac{1}{2\pi}\int_{\pi}^{\pi} |f(\theta)|^2 \mathrm{d}\theta \,, \tag{9.20}$$

it is *not* true that for any sequence in $\{a_n\} \in l^2(\mathbb{Z})$ we can find a Riemann integrable function whose Fourier coefficients are a_n. The reason is that the space of Riemann integrable functions is *not* complete, and the feat of Lebesgue measure theory is to provide such a completion $\mathrm{L}^2([-\pi,\pi])$. With respect to these complete spaces the relationship between a function on the circle and its Fourier coefficients is a unitary equivalence.

9.2.3. Supplement: Fourier series

A few comments about the proof of Theorem 9.6.

Recall from (3.3) the Poisson kernel $P_r(y) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{iny}$. We have already shown in Lecture 3 that for Riemann integrable functions f,

$$\sum_{n=-\infty}^{\infty} a_n r^n e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-y) P_r(y) \mathrm{d}y \,. \tag{9.21}$$

More generally, for functions in $\mathrm{L}^1([-\pi,\pi])$ this can be established with the help of the dominated convergence theorem, after evaluating the right hand side term by term. We have also seen in Lecture 3 that $P_r(\theta)$ is a family of good kernels, so we would like to conclude with Theorem 2.2 that $(f * P_r)(x) \to f(x)$ as $r \to 1$, which holds at every point of continuity of f, but which we cannot assume here. Instead we apply Theorem 9.8 which gives that $(f * P_r)(x) \to f(x)$ *almost everywhere*. In conclusion,

$$\sum_{n=-\infty}^{\infty} a_n r^n e^{inx} \to f(x) \qquad \text{for almost every } x \text{ as} \qquad r \to 1\,. \tag{9.22}$$

In particular, if $a_n = 0$ then $f(x) = 0$ almost everywhere.

Problems

1. a) Show that neither the inclusion $\mathrm{L}^2(\mathbb{R}^d) \subset \mathrm{L}^1(\mathbb{R}^d)$ nor the inclusion $\mathrm{L}^1(\mathbb{R}^d) \subset \mathrm{L}^2(\mathbb{R}^d)$ is valid.

   b) Note, however, that if f is supported on a set E of finite measure and if $f \in \mathrm{L}^2(\mathbb{R}^d)$, then by the Cauchy-Schwarz inequality applied to $f\chi_E$, $f \in \mathrm{L}^1(\mathbb{R}^d)$ with

   $$\|f\|_{L^1(\mathbb{R}^d)} \le |E|^{1/2} \|f\|_{L^2(\mathbb{R}^d)} \tag{9.23}$$

   c) If f is bounded, say $|f(x)| \le M$, and $f \in \mathrm{L}^1(\mathbb{R}^d)$, then $f \in \mathrm{L}^2(\mathbb{R}^d)$ with

   $$\|f\|_{L^2(\mathbb{R}^d)} \le M^{1/2} \|f\|_{L^1(\mathbb{R}^d)}^{1/2} \,. \tag{9.24}$$

2. Let $\eta(t)$ be a fixed strictly positive continuous function on $[a,b]$. Define $\mathcal{H}_\eta$ to be the space of all measurable functions f on $[a,b]$ such that

   $$\int_a^b |f(t)|^2 \eta(t) \mathrm{d}t < \infty \tag{9.25}$$

   and the inner product on $\mathcal{H}_\eta$ by

   $$(f,g)_\eta = \int_a^b f(t)\overline{g(t)}\eta(t)\mathrm{d}t \,. \tag{9.26}$$

   Show that $\mathcal{H}_\eta$ is a Hilbert space and that the mapping

   $$U : f \mapsto \eta^{1/2} f \tag{9.27}$$

   gives a unitary correspondence between $\mathcal{H}_\eta$ and $\mathrm{L}^2([a,b])$.

Supplement: Approximation to the identity

In the Digresssion following Lecture 2 we have introduced the notion of **good kernels** K_δ and shown that whenever f is bounded, then $(f * K_\delta)(x) \to f(x)$ as $\delta \to 0$ at every point of continuity of f. To obtain a similar conclusion for functions $f \in \mathrm{L}^1(\mathbb{R}^d)$, namely that this is true *almost everywhere*, we need to strengthen our assumptions on the kernels K_δ, and the resulting *narrower* class of kernels will be referred to as **approximations to the identity**.

A family of integrable functions K_δ on $\mathbb{R}^d$ is an *approximation to the identity* if

1. $\int_{\mathbb{R}^d} K_\delta(x)\mathrm{d}x = 1$ for all $\delta > 0$
2. $|K_\delta(x)| \leq A\delta^{-d}$ for all $\delta > 0$.
3. $|K_\delta(x)| \leq A\delta/|x|^{d+1}$ for all $\delta > 0$ and $x \in \mathbb{R}^d$.

Remark 9.2. An approximation to the identity is always a family of good kernels. In other words, these conditions are stronger than those for good kernels.

Theorem 9.8. *If $\{K_\delta\}$ is an approximation to the identity and $f \in \mathrm{L}^1(\mathbb{R}^d)$, then*

$$(f * K_\delta)(x) \to f(x) \qquad (\delta \to 0) \tag{9.28}$$

for every x in the Lebesgue set of f. In particular, the limit holds for almost every *x.*

Review: Linear subspaces and maps

Recommended Reading

(Stein and Shakarchi, *Real analysis*, Chapter 4, Sections 4).

9.1. Closed subspaces and orthogonal projections

In an infinite dimensional Hilbert space $\mathcal{H}$, such as $\mathrm{L}^2(\mathbb{R}^d)$, not all *linear subspaces* are closed. For example the space of Riemann integrable functions $\mathcal{R} \subset \mathrm{L}^2(\mathbb{R}^d)$, or Schwartz space $\mathcal{S}(\mathbb{R}^d) \subset \mathrm{L}^2(\mathbb{R}^d)$ are *not* closed. A **linear subspace** $\mathcal{S} \subset \mathcal{H}$ is **closed** if, whenever $f_n \subset \mathcal{S}$ converges to some $f \in \mathcal{H}$, then f belongs to $\mathcal{S}$.

Lemma 9.1. *Suppose $\mathcal{S}$ is a closed subspace of $\mathcal{H}$ and $f \in \mathcal{H}$. Then*

1. *There exists a (unique) element $g_0 \in \mathcal{S}$ which is closest to f, in the sense that*

$$\|f - g_0\| = \inf_{g \in \mathcal{S}} \|f - g\| \,. \tag{9.1}$$

2. *The element $f - g_0$ is perpendicular to $\mathcal{S}$,*

$$(f - g_0, g) = 0 \qquad \textit{for all } g \in \mathcal{S} \,. \tag{9.2}$$

Proof of 1. If $f \notin \mathcal{S}$ then $d = \inf_{g \in \mathcal{S}} \|f - g\| > 0$ because $\mathcal{S}$ is closed. Consider a sequence $\{g_n\}$ in $\mathcal{S}$ such that $\|f - g_n\| \to d$. This tells us in particular that the sequence $\{g_n\}$ is *bounded*, and in *finite dimensions* we could infer that there exists a convergent subsequence g_{n_k}, which would give us the element $g_0 = \lim_{k\to\infty} g_{n_k}$. However, it is precisely this type of *compactness that fails* in general.

Instead we can show that g_n is a Cauchy sequence using the **parallelogram law**, which states that

$$\|x + y\|^2 + \|x - y\|^2 = 2\Big[\|x\|^2 + \|y\|^2\Big] \qquad (x, y \in \mathcal{H}) \,. \tag{9.3}$$

Applied to $x = f - g_n$ and $y = f - g_m$ we get

$$\|2f - (g_n + g_m)\|^2 + \|g_n - g_m\|^2 = 2\Big[\|f - g_n\|^2 + \|f - g_m\|^2\Big] \tag{9.4}$$

and since $g_n + g_m \in \mathcal{S}$, $\|2f - (g_n + g_m)\|^2 \geq 4d^2$, so it follows

$$\|g_n - g_m\|^2 \leq 2\Big[\|f - g_n\|^2 + \|f - g_m\|^2\Big] - 4d^2 \,, \tag{9.5}$$

and hence $\|g_n - g_m\|^2 \to 0$ as $m, n \to \infty$, because $\|f - g_n\| \to d$ as $n \to \infty$. Since $\mathcal{H}$ is complete $g_n \to g$ for some $g \in \mathcal{H}$, and since $\mathcal{S}$ is closed, $g \in \mathcal{S}$. $\square$

Further to the concept of orthogonality we define the **orthogonal complement** $\mathcal{S}^\perp$ of any subspace $\mathcal{S}$ by:

$$\mathcal{S}^\perp = \{f \in \mathcal{H} : (f, g) = 0\,,\ g \in \mathcal{S}\} \tag{9.6}$$

Note that $\mathcal{S}^\perp$ is always a *closed* subspace. Indeed, in general if $f_n \to f$, then $(f_n, g) \to (f, g)$ because by the Cauchy-Schwarz inequality:

$$|(f_n - f, g)| \leq \|f_n - f\|\|g\| \to 0 \qquad (n \to \infty)\,. \tag{9.7}$$

Hence if $f_n \in \mathcal{S}^\perp$, so $(f_n, g) = 0$ for all $g \in \mathcal{S}$, then $(f, g) = 0$ for all $g \in \mathcal{S}$, so $f \in \mathcal{S}^\perp$.

Proposition 9.2. *If $\mathcal{S}$ is a closed subspace of a Hilbert space $\mathcal{H}$, then*

$$\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp\,. \tag{9.8}$$

The proposition says that every element $f \in \mathcal{H}$ can be written *uniquely* as $f = g + h$, where $g \in \mathcal{S}$ and $h \in \mathcal{S}^\perp$; we say $\mathcal{H}$ is the **direct sum** of $\mathcal{S}$ and $\mathcal{S}^\perp$.

Proof. For any $f \in \mathcal{H}$, choose $g_0 \in \mathcal{S}$ as in the Lemma, then

$$f = f - g_0 + g_0\,, \tag{9.9}$$

and $(f - g_0, g_0) = 0$. To prove that the decomposition is unique not that if $f = g + h = \tilde{g} + \tilde{h}$, then $g - \tilde{g} = \tilde{h} - h$, and since $\mathcal{S} \cap \mathcal{S}^\perp = \{0\}$, we get $g = \tilde{g}$, and $\tilde{h} = h$. $\square$

This decomposition also entails a natural projection onto $\mathcal{S}$ defined by

$$P_{\mathcal{S}}(f) = g\,, \qquad \text{where } f = g + h\,, \quad g \in \mathcal{S}, h \in \mathcal{S}^\perp\,. \tag{9.10}$$

This map is called the **orthogonal projection** on $\mathcal{S}$ and satisfies the following properties:

1. $P_{\mathcal{S}}$ is linear
2. $P_{\mathcal{S}}(f) = f$ whenever $f \in \mathcal{S}$ and $P_{\mathcal{S}}(f) = 0$ whenever $f \in \mathcal{S}^\perp$
3. $\|P_{\mathcal{S}}(f)\| \leq \|f\|$ for all $f \in \mathcal{H}$.

Example 9.1. Given a function $f \in \mathrm{L}^2([-\pi, \pi])$ the partial sum

$$S_N(f)(x) = \sum_{n=-N}^{N} a_n e^{inx} \qquad a_n = (f, e_n) \qquad e_n(x) = e^{inx} \tag{9.11}$$

is a projection to the closed subspace spanned by $\{e_{-N}, \ldots, e_N\}$. We have seen that this projection can also be expressed as

$$S_N(f)(x) = \frac{1}{2} \int_{-\pi}^{\pi} D_N(x - y) f(y) \mathrm{d}y \tag{9.12}$$

where D_N is the Dirichlet kernel.

9.2. Linear transformations

We have already encountered two classes of linear maps: unitary mappings and orthogonal projections. There are many other important classes of linear transformations, such as "compact operators", and "closed operators", and "linear functionals" which will play an important role in Lecture 9.

A mapping $T : \mathcal{H}_1 \to \mathcal{H}_2$ from one Hilbert space to another is a **linear transformation** (or **linear operator**) if $T(af + bg) = aT(f) + bT(g)$ for all scalars a, b and $f, g \in \mathcal{H}_1$. A linear operator is **bounded** if there exists a $M > 0$ so that

$$\|T(f)\|_{\mathcal{H}_2} \leq M\|f\|_{\mathcal{H}_1}\,, \qquad f \in \mathcal{H}_1\,. \tag{9.13}$$

The **norm** of T is the smallest such number M, or more precisely the greatest lower bound of the set of numbers M for which this inequality holds:

$$\|T\| = \inf\{M\} \tag{9.14}$$

Example 9.2. The **identity** $I(f) = f$ is a unitary operator, and an orthogonal projection, with $\|I\| = 1$.

Lemma 9.3.

$$\|T\| = \sup\{|(Tf, g)| : \|f\| \leq 1\,, \|g\| \leq 1\} \tag{9.15}$$

Proof. Let A be the number defined by the right hand side.

Let $M \geq \|T\|$, then by the Cauchy-Schwarz inequality, whenever $\|f\| \leq 1$, $\|g\| \leq 1$,

$$|(Tf, g)| \leq \|Tf\|\|g\| \leq M\,, \tag{9.16}$$

we see that M is an upper bound, so $A \leq M$, and thus $A \leq \|T\|$.

Now let $M \geq A$. If we can prove that then (9.13) holds, this shows that $\|T\| \leq A$. We can assume that $Tf \neq 0$ for otherwise the inequality is trivial. Set

$$f' = \frac{f}{\|f\|} \qquad g' = \frac{Tf}{\|Tf\|} \tag{9.17}$$

then by assumption

$$(Tf', g') \leq A \leq M \tag{9.18}$$

but $(Tf', g') = \|Tf\|/\|f\|$, which shows that $\|Tf\| \leq M\|f\|$. □

A linear transformation is **continuous** if $T(f_n) \to T(f)$ whenever $f_n \to f$. It is well known that for *linear* operators on a Hilbert space continuity and boundedness are equivalent.

Proposition 9.4. *A linear operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ is bounded if and only if it is continuous.*

Problems

1. (Revisit this problem after Lecture 9.) Let $P = P_{\mathcal{S}}$ be the orthogonal projection onto a closed subspace $\mathcal{S}$ in a Hilbert space $\mathcal{H}$.
   a) Show that $P^2 = P$ and $P^* = P$
   b) Conversely, if P is a any bounded operator satisfying $P^2 = P$ and $P^* = P$, prove that $P = P_{\mathcal{S}}$ is the orthogonal projection associated to the some closed subspace $\mathcal{S}$ of $\mathcal{H}$.
   c) Using the projection $P_{\mathcal{S}}$, prove that if $\mathcal{S}$ is a closed subspace of a Hilbert space, then $\mathcal{S}$ is also a Hilbert space.
2. Suppose P_1 and P_2 are a pair of orthogonal projections on S_1 and S_2, respectively. Then P_1P_2 is an orthogonal projection if and only if P_1 and P_2 commute, that is $P_1P_2 = P_2P_1$. In this case, P_1P_2 projects onto $\mathcal{S}_1 \cap \mathcal{S}_2$.

Lecture 9.

Riesz representation theorem and adjoints

Further Reading

(Stein and Shakarchi, *Real analysis*, Chapter 4, Sections 5.1, 5.2, and 2.3, and Chapter 5, Section 1).

9.1. Linear functionals

A **linear functional** l is a linear map from a Hilbert space $\mathcal{H}$ to the underlying field of scalars,

$$l : \mathcal{H} \to \mathbb{C}\,, \tag{9.1}$$

where the $\mathbb{C}$ is equipped with the norm $|\cdot|$. A linear functional l is **continuous** if $\lim_{n\to\infty} l(f_n) = l(f)$, whenever $f_n \to f$ in $\mathcal{H}$.

Example 9.1. For each fixed $g \in \mathcal{H}$, a linear functional is defined by

$$l_g(f) = (f, g)\,. \tag{9.2}$$

This functional is clearly bounded, hence continuous, by the Cauchy Schwarz inequality. In fact,

$$\|l_g\| = \|g\|\,. \tag{9.3}$$

The remarkable fact is that these are *all* the linear continuous functionals on a Hilbert space.

Theorem 9.1 (Riesz representation theorem)**.** *Let l be a continuous linear functional on a Hilbert space $\mathcal{H}$. Then there exists $g \in \mathcal{H}$ such that*

$$l(f) = (f, g) \qquad f \in \mathcal{H}\,. \tag{9.4}$$

Moreover $\|l\| = \|g\|$.

Proof. The idea is to consider the *null space* of l,

$$\mathcal{S} = \Big\{ f \in \mathcal{H} : l(f) = 0 \Big\} \tag{9.5}$$

If $\mathcal{S} = \mathcal{H}$ then $l = 0$ and we can choose $g = 0$. Suppose then that $\mathcal{S} \subsetneq \mathcal{H}$. The important point is that $\mathcal{S}$ is a *closed subspace*. Indeed, if $f_n \to f \in \mathcal{H}$, where $f_n \in \mathcal{S}$, then

$$l(f) = \lim_{n\to\infty} l(f_n) = 0\,, \tag{9.6}$$

because l is continuous, so $f \in \mathcal{S}$.

Now choose any $f_0 \in \mathcal{H} \setminus \mathcal{S}$, then by Prop. 9.2, we can find some $g_0 \in \mathcal{S}$, and $0 \neq h_0 \in \mathcal{S}^\perp$,

$$f_0 = g_0 + h_0\,. \tag{9.7}$$

Since $h_0 \neq 0$ we can set

$$h = \frac{h_0}{\|h_0\|}\,. \tag{9.8}$$

Now given any $f \in \mathcal{H}$, we can find a linear combination u of f and h in $\mathcal{S}$:

$$u = l(h)f - l(f)h \qquad l(u) = 0 \tag{9.9}$$

This means that $(u, h) = 0$, and since $\|h\| = 1$, we infer

$$0 = (u, h) = (l(h)f, h) - l(f) \tag{9.10}$$

or alternatively,

$$l(f) = (f, g) \qquad g = \overline{l(h)}h\,. \tag{9.11}$$

Therefore $l = l_g$, and as we have seen above $\|l_g\| = \|g\|$.

□

9.2. Adjoints

As a first application of the Riesz representation theorem we will infer the existence of an "adjoint" transformation.

Theorem 9.2. *Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. There exists a unique bounded linear operator T^* on $\mathcal{H}$, the so-called **adjoint** of T, so that:*

1. $(Tf, g) = (f, T^*g)$
2. $\|T\| = \|T^*\|$
3. $(T^*)^* = T$.

Proof. We can view, for each fixed $g \in \mathcal{H}$,

$$l(f) = (Tf, g) \tag{9.12}$$

as a linear functional, which is bounded because T is bounded. Consequently, by the Riesz representation theorem, there exists $h \in \mathcal{H}$ so that

$$l(f) = (f, h)\,. \tag{9.13}$$

We can then define

$$T^*g = h\,. \tag{9.14}$$

It is clear that T^* is linear, and by definition $(Tf, g) = (f, h) = (f, T^*g)$. In fact, for any $f, g \in \mathcal{H}$,

$$(T^*f, g) = \overline{(g, T^*f)} = \overline{(Tg, f)} = (f, Tg) \tag{9.15}$$

which shows that $(T^*)^* = T$. Moreover by Lemma 9.3,

$$\begin{aligned} \|T^*\| &= \sup\{|(T^*f, g)| : \|f\| \le 1, \|g\| \le 1\} \\ &= \sup\{|(f, Tg)| : \|f\| \le 1, \|g\| \le 1\} = \|T\|\,. \end{aligned} \tag{9.16}$$

□

Remark 9.1. An operator is called **symmetric**[1] if $T^* = T$. In this case, one can show that

$$\|T\| = \sup\{|(Tf, f)| : \|f\| = 1\}\,. \tag{9.17}$$

Remark 9.2. If T and S are bounded linear transformations of $\mathcal{H}$ to itself, then so is TS. Moreover

$$(TS)^* = S^*T^*\,, \tag{9.18}$$

because $(TSf, g) = (Sf, T^*g) = (f, S^*T^*g)$.

Remark 9.3. There is an associated **bilinear from** to each bounded linear operator T,

$$B(f, g) = (Tf, g)\,. \tag{9.19}$$

More precisely, B is linear in f, but *conjugate linear* in g. Also by Cauchy-Schwarz,

$$|B(f, g)| \le \|T\|\|f\|\|g\|\,. \tag{9.20}$$

Conversely, if B is a bilinear form, which satisfies

$$|B(f, g)| \le M\|f\|\|g\| \tag{9.21}$$

for some $M > 0$, then by same argument gave the existence of the adjoint, we can show that there exists a unique bounded linear transformation T so that

$$B(f, g) = (Tf, g)\,. \tag{9.22}$$

9.3. Extensions and completions

In applications of the Riesz representation theorem the linear functional in question, l, is often *at first* not defined on the whole Hilbert space $\mathcal{H}$, but merely a dense subspace $\mathcal{H}_0$. For example, when we discuss constant coefficient partial differential equations, we will deal with functionals first defined on $\mathcal{H}_0 = C_0^\infty$, namely the space of smooth functions of compact support, as a subspace of $\mathcal{H} = \mathrm{L}^2$. We thus need to discuss the notion of an *extension* of a linear functional.

[1] For *bounded* operators the terms *symmetric*, *self-adjoint*, and *essentially self-adjoint* can all be used synomymously, but for unbounded operators their definitions diverge.

9.3.1. Pre-Hilbert spaces

A **pre-Hilbert space** is a space $\mathcal{H}_0$ which has all the properties of a Hilbert space *except* completeness. A **completion** of $\mathcal{H}_0$ is a Hilbert space $\mathcal{H}$ such that

1. $\mathcal{H}_0 \subset \mathcal{H}$
2. $(f, g)_0 = (f, g)$ whenever $f, g \in \mathcal{H}_0$
3. $\mathcal{H}_0$ is dense in $\mathcal{H}$.

In most cases of interest to us the completions will be known to us, and the spaces $\mathcal{H}_0$ will be explicit subspaces of a given Hilbert space $\mathcal{H}$. However, it is worth pointing out that a completion *always exists*, and is unique up to isomorphisms.

Proposition 9.3. *Given any pre-Hilbert space $\mathcal{H}_0$ there exists a completion $\mathcal{H}$ of $\mathcal{H}_0$.*

The proof proceeds by the construction of $\mathcal{H}$ as the collection of *Cauchy sequences* $\{f_n\}$ with $f_n \in \mathcal{H}_0$. One defines an equivalence relation in this collection by saying that $\{f_n\}$ is equivalent to $\{f'_n\}$ if $f_n - f'_n$ converges to 0 as $n \to \infty$. $\mathcal{H}$ is then taken to be the space of equivalence classes. Note that $\mathcal{H}$ contains $\mathcal{H}_0$ in the form of the elements $\{f_n\}$, with $f_n = f \in \mathcal{H}_0$. An inner product on $\mathcal{H}$ is defined by

$$(f, g) = \lim_{n\to\infty} (f_n, g_n) \tag{9.23}$$

where the sequences $\{f_n\}$, and $\{g_n\}$ represent f, and g respectively. One can show that $\mathcal{H}$ with this inner product is indeed *complete*.

9.3.2. Extensions

Suppose $\mathcal{H}$ is the completion of a pre-Hilbert space $\mathcal{H}_0$, and suppose l_0 is a linear functional on $\mathcal{H}_0$ which is bounded, namely

$$|l_0(f)| \le M\|f\| \qquad (f \in \mathcal{H}_0) \tag{9.24}$$

We would like to define the *extension* l of l_0 to $\mathcal{H}$. For any $f \in \mathcal{H}$ we can choose $f_n \in \mathcal{H}_0$ such that $f_n \to f$ in $\mathcal{H}$. Therefore

$$|l(f_n) - l(f_m)| \le M\|f_n - f_m\| \to 0 \qquad (m, n \to \infty) \tag{9.25}$$

and thus $\{l(f_n)\}$ is a Cauchy sequence, and we may define

$$l(f) = \lim_{n\to\infty} l(f_n) \tag{9.26}$$

This is well-defined, namely independent of the choice of the sequence $\{f_n\}$: If g_n is another sequence in $\mathcal{H}_0$ that converges to f in $\mathcal{H}$, then $|l(f_n) - l(g_n)| \le M\|f_n - g_n\| \le M\|f_n - f\| + M\|f - g_n\| \to 0$, and we conclude that $\{l(g_n)\}$ converges to the same limit as $\{l(f_n)\}$. The extension l is a bounded linear functional on $\mathcal{H}$, with $|l(f)| \le M\|f\|$.

This is a special case of the following

Lemma 9.4 (Extension principle)**.** *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ denote Hilbert spaces with norms $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively. Suppose $\mathcal{S}$ is a dense subspace of $\mathcal{H}_1$ and $T_0 : \mathcal{S} \to \mathcal{H}_2$ is a linear transformation that satisfies*

$$\|T_0(f)\|_2 \leq M\|f\|_1 \qquad (f \in \mathcal{S})\,. \tag{9.27}$$

Then T_0 extends to a unique linear transformation $T : \mathcal{H}_1 \to \mathcal{H}_2$ that satisfies

$$\|T(f)\|_2 \leq M\|f\|_1 \qquad (f \in \mathcal{H}_1)\,. \tag{9.28}$$

9.3.3. Fourier transform on L^2

An example of this extension procedure is the definition of the Fourier transform on $\mathrm{L}^2(\mathbb{R}^d)$. We have previously defined the Fourier transform on Schwartz space, so let us denote for this purpose by $\mathcal{F}_0$ the map

$$\mathcal{F}_0(f) = \hat{f} \qquad (f \in \mathcal{S}(\mathbb{R}^d)) \tag{9.29}$$

This is a linear map from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$, which is *bounded* in view of Plancherel's identity:

$$\int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 \mathrm{d}\xi = \int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}x\,. \tag{9.30}$$

The idea is, as suggested above, to define the Fourier transform $\mathcal{F}$ as the extension of $\mathcal{F}_0$ to $\mathrm{L}^2(\mathbb{R}^d)$: If $\{f_n\}$ is a sequence in Schwartz space that converges to f in $\mathrm{L}^2(\mathbb{R}^d)$, then $\{\mathcal{F}_0(f_n)\}$ will converge to an element in $\mathrm{L}^2(\mathbb{R}^d)$ which we will *define* as the Fourier transform of f. For this procedure to work we need that every L^2 function can indeed be approximated by functions in Schwartz space.

Lemma 9.5. *The space $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathrm{L}^2(\mathbb{R}^d)$. In other words, given any $f \in \mathrm{L}^2(\mathbb{R}^d)$, there is a sequence $\{f_n\} \subset \mathcal{S}(\mathbb{R}^d)$ such that $\|f - f_n\|_{\mathrm{L}^2(\mathbb{R}^d)} \to 0$ as $n \to \infty$.*

Thus we can apply Lemma 9.4 to the case $\mathcal{H}_1 = \mathcal{H}_2 = \mathrm{L}^2(\mathbb{R}^d)$, $\mathcal{S} = \mathcal{S}(\mathbb{R}^d)$, and $T_0 = \mathcal{F}_0$. This yields a bounded linear map $\mathcal{F}$, which is the the extension of $\mathcal{F}_0$ in the sense that for any $f \in \mathrm{L}^2(\mathbb{R}^d)$,

$$\mathcal{F} = \lim_{n\to\infty} \mathcal{F}_0(f_n) \tag{9.31}$$

where $f_n \in \mathcal{S}(\mathbb{R}^d)$ is an approximating sequence $f_n \to f$ in $\mathrm{L}^2(\mathbb{R}^d)$.

Theorem 9.6. *The Fourier transform $\mathcal{F}_0$, initially defined on $\mathcal{S}(\mathbb{R}^d)$, has a (unique) extension $\mathcal{F}$ to a unitary mapping of $\mathrm{L}^2(\mathbb{R}^d)$ to itself. In particular, for all $f \in \mathrm{L}^2(\mathbb{R}^d)$,*

$$\|\mathcal{F}(f)\|_{\mathrm{L}^2(\mathbb{R}^d)} = \|f\|_{\mathrm{L}^2(\mathbb{R}^d)}\,. \tag{9.32}$$

Problems

1. If T is a bounded linear operator on a Hilbert space, prove that

$$\|TT^*\| = \|T^*T\| = \|T\|^2 = \|T^*\|^2\,. \tag{9.33}$$

2. Prove the statement about bilinear forms made in Remark 9.3.

Lecture 10.

Constant coefficient partial differential equations

Further Reading

(Stein and Shakarchi, *Real analysis*, Chapter 5, Sections 3).

In this lecture we turn to general linear partial differential equations in $\mathbb{R}^d$ of the form

$$L(u) = f \tag{10.1}$$

where L is a differential operator on with *constant coefficients* $a_\alpha \in \mathbb{C}$,

$$L = \sum_{|\alpha| \leq n} a_\alpha \Big(\frac{\partial}{\partial x}\Big)^\alpha \tag{10.2}$$

The wave equation, the heat equation, and Laplace's equation are all examples of this class of equations, and as in these special cases we may try to understand the solution in the case that u and f are in Schwartz space, which then leads to the equation

$$P(\xi)\hat{u}(\xi) = \hat{f}, \tag{10.3}$$

where $P(\xi)$ is the **characteristic polynomial** of L defined by

$$P(\xi) = \sum_{|\alpha| \leq n} a_\alpha (2\pi i \xi)^\alpha \,. \tag{10.4}$$

Thus a solution u, if it exists, and is in $\mathcal{S}(\mathbb{R}^d)$, would be determined by

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{P(\xi)} \,. \tag{10.5}$$

This is not always possible, and too restrictive in general, but nonetheless the characterisitic polynomial is an important concept related to (10.1). A larger class of solutions, understood in a wider sense, are the following "weak solutions."

10.1. Weak solutions

Let $\Omega \subset \mathbb{R}^d$ be an open set, then we denote by $C_0^\infty(\Omega)$ the *smooth functions of compact support* in Ω.

Lemma 10.1. *The space* $C_0^\infty(\Omega)$ *is dense in* $L^2(\Omega)$.

Proof. For $f \in L^2(\mathbb{R}^d)$, set

$$g_M(x) = \begin{cases} f(x) & \text{if } |x| \leq M \text{ and } d(x, \Omega^c) \geq 1/M \text{ and } |f(x)| \leq M \\ 0 & \text{otherwise}, \end{cases} \tag{10.6}$$

then $g_M(x) \to f(x)$ as $M \to \infty$ almost everywhere in Ω, and $|f(x) - g_M(x)|^2 \leq 4|f(x)|^2$ so by dominated convergence, $\|g_M - f\| \to 0$. Now for given $\epsilon > 0$, let $g = g_M$, with M so that

$$\|f - g_M\|_{L^2(\Omega)} < \epsilon/2 . \tag{10.7}$$

It remains to approximate g by a smooth function of compact support, which can be achieved by a process called **regularization** (or **mollification**), which yields a smooth function of compact support $g * K_\delta$, where K_δ is an approximation to the identity, such that $(g * K_\delta)(x) \to g(x)$ almost everywhere. In fact, by choosing $\delta < 1/M$, we can ensure that the support of $g_M * K_\delta$ is contained in Ω. Therefore again by the dominated convergence theorem $\|g * K_\delta - g\| \to 0$, and so we can find $\delta < 1/M$ such that

$$\|g - g * K_\delta\| < \epsilon/2 . \tag{10.8}$$

□

Besides $C_0^\infty(\Omega)$ we will also use the spaces $C^n(\Omega)$ consisting of all function f on Ω with continuous partial derivatives up to order n, and the space $C^n(\bar{\Omega})$ consisiting of those functions on $\bar{\Omega}$ that can be extended to functions in $\mathbb{R}^d$ that belong to $C^n(\mathbb{R}^d)$. We note, for each natural number n,

$$C_0^\infty(\Omega) \subset C^n(\bar{\Omega}) \subset C^n(\Omega) . \tag{10.9}$$

Now for $\varphi, \psi \in C_0^\infty(\Omega)$, we have by integration by parts,

$$\int_{-\infty}^{\infty} (\partial_{x_j}\varphi)(x)\overline{\psi}(x)\mathrm{d}x_j = -\int_{-\infty}^{\infty} \varphi(x)\overline{\partial_{x_j}\psi}(x)\mathrm{d}x_j , \tag{10.10}$$

because both functions have compact support. More generally,

$$(L\varphi, \psi) = \int_{\mathbb{R}^d} (L\varphi)(x)\overline{\psi(x)}\mathrm{d}x = (\varphi, L^*\psi) \tag{10.11}$$

where L^* is the (formal) **adjoint operator** of L defined by

$$L^* = \sum_{|\alpha| \leq n} (-1)^{|\alpha|}\overline{a_\alpha}\frac{\partial^{|\alpha|}}{\partial x^\alpha} . \tag{10.12}$$

Note that the identity (10.11) continues to hold for $\varphi \in \mathrm{C}^n(\Omega)$, because the boundary terms still vanish with $\psi \in \mathrm{C}_0^\infty(\Omega)$. In particular if (10.1) holds in the **strong sense**, namely $u \in \mathrm{C}^n(\Omega)$, and

$$Lu = f \tag{10.13}$$

then we have

$$(f, \psi) = (u, L^*\psi) \qquad \psi \in \mathrm{C}_0^\infty(\Omega)\,. \tag{10.14}$$

Definition 10.1. For any $f \in \mathrm{L}^2(\Omega)$, a function $u \in \mathrm{L}^2(\Omega)$ is a **weak solution** to the equation $Lu = f$ in Ω, if for all $\psi \in \mathrm{C}_0^\infty(\Omega)$,

$$(u, L^*\psi) = (f, \psi)\,. \tag{10.15}$$

As we have seen any strong solution is a weak solution of $Lu = f$. However, not all weak solutions are strong solutions, in the ordinary sense.

Example 10.1. Consider the $1+1$-dimensional wave equation. Here

$$L(u) = -\partial_t^2 u + \partial_x^2 u \tag{10.16}$$

where $(t, x) \in \mathbb{R} \times \mathbb{R}$. Consider initial data corresponding to a "plucked string", so

$$u(0, x) = f(x) \qquad \partial_t u(0, x) = 0\,, \tag{10.17}$$

where f is a piecewise linear function, say on the interval $[0, \pi]$,

$$f(x) = \begin{cases} \frac{h}{p} x & 0 \leq x \leq p \\ h - \frac{h}{\pi - p}(x - p) & p \leq x \leq \pi \end{cases} \tag{10.18}$$

where $h > 0$, and $0 < p < \pi$. If we extend f as an odd function to $[-\pi, \pi]$, and then to $\mathbb{R}$ as a 2π-periodic function, then we expect that by d'Alembert's formula

$$u(x, t) = \frac{f(x+t) + f(x-t)}{2}\,. \tag{10.19}$$

Note that by construction $u(0, t) = u(\pi, t) = 0$, but u is *not* twice continuously differentiable, hence not a strong solution. Nevertheless it *is* a weak solution:

We need to verify that

$$(u, L^*\psi) = 0 \qquad \psi \in \mathrm{C}_0^\infty(\mathbb{R}^2)\,. \tag{10.20}$$

Let us define

$$u_n(x, t) = \frac{f_n(x+t) + f_n(x-t)}{2} \tag{10.21}$$

where f_n approximates f in the sense that f_n is a sequence of smooth functions on $\mathbb{R}$ such that $f_n \to f$ uniformly on every closed interval. Then $L(u_n) = 0$ and hence $(u_n, L^*\psi) = 0$ for every $\psi \in \mathrm{C}_0^\infty(\mathbb{R}^2)$, which implies by uniform convergence that

$$0 = (u_n, L^*\psi) = \int_{\mathbb{R}^2} u_n(t, x)\overline{(L^*\psi)(t, x)}\mathrm{d}t\mathrm{d}x \longrightarrow (u, L^*\psi)\,. \tag{10.22}$$

Example 10.2. Another instructive example is the operator $L = \frac{\mathrm{d}}{\mathrm{d}x}$ on $\mathbb{R}$. If $\Omega = (0, 1)$, then $u \in \mathrm{L}^2(\Omega)$ is a weak solution to $Lu = f$, for some $f \in \Omega$, if and only if there is an *absolutely continuous* function F on $[0, 1]$ such that $F(x) = u(x)$, and $F'(x) = f(x)$ almost everywhere. Cf. Problems below.

10.2. Existence of weak solutions

Having discussed the notion of a weak solution, we will now prove that they always exist for constant coefficient partial differential equations.

Theorem 10.2. *Suppose Ω is a bounded open subset of $\mathbb{R}^d$. Given a linear partial differential operator L with constant coefficients, there exists a bounded linear operator K on $\mathrm{L}^2(\Omega)$ such that*

$$L(Kf) = f \qquad \text{in the weak sense} \tag{10.23}$$

whenever $f \in \mathrm{L}^2(\Omega)$. In other words, $u = K(f)$ is a weak solution to $L(u) = f$.

The map K is sometimes called the solution map. It cannot exist unless L is surjective. Now for *bounded operators* T on a Hilbert space $\mathcal{H}$,

$$\ker(T^*) = \mathrm{range}(T)^\perp \,, \tag{10.24}$$

because if $g \in \mathrm{range}(T)^\perp$ is equivalent to $(g, Tu) = 0$ for all $u \in \mathcal{H}$, which is equivalent to $(T^*g, u) = 0$ for all $u \in \mathcal{H}$, which holds if and only $T^*g = 0$. So a bounded operator is surjective if its adjoint has trivial kernel, namely if the adjoint is injective. This argument does *not* directly apply to the differential operator L, but we can nonetheless prove the following central estimate, which is a quatitative version of the required injectivity property of L^*:

Proposition 10.3. *There exists a constant C such that for all $\psi \in \mathrm{C}_0^\infty(\Omega)$,*

$$\|\psi\|_{\mathrm{L}^2(\Omega)} \leq C \|L^*\psi\|_{\mathrm{L}^2(\Omega)} \,. \tag{10.25}$$

Now consider the pre-Hilbert space $\mathcal{H}_0 = \mathrm{C}_0^\infty(\Omega)$ equipped with the inner product

$$(\varphi, \psi)_0 = (L^*\varphi, L^*\psi) \,, \qquad \|\varphi\|_0 = \|L^*\varphi\|_{\mathrm{L}^2(\Omega)} \,. \tag{10.26}$$

In particular L^* is then *bounded* as a map from $\mathcal{H}_0$ to $\mathrm{L}^2(\Omega)$. We have seen in Proposition 9.3, that there exists a completion $\mathcal{H}$ of $\mathcal{H}_0$, and Proposition 10.3 tells us that any Cauchy sequence in $\mathcal{H}_0$ is also a Cauchy sequence in $\mathrm{L}^2(\Omega)$, which is complete, thus we can identify $\mathcal{H}$ with a subspace of $\mathrm{L}^2(\Omega)$,

$$\mathcal{H} \subset \mathrm{L}^2(\Omega) \,. \tag{10.27}$$

Moreover, in view of Lemma 9.4 L^* extends to a bounded linear transformation on $\mathcal{H}$, which we again denote by L^*. To avoid confusion, let us denote the inner product in $\mathcal{H}$ by $\langle \cdot, \cdot \rangle$.

The aim is to show that for any $f \in \mathrm{L}^2(\Omega)$, there exists $u \in \mathrm{L}^2(\Omega)$ such that for all $\psi \in \mathrm{C}_0^\infty(\Omega)$,

$$(f, \psi) = (u, L^*\psi) \tag{10.28}$$

and u depends linearly on f. So let us define for fixed $f \in \mathrm{L}^2(\Omega)$, a linear functional l_0 on $\mathcal{H}_0$,

$$l_0(\psi) = (\psi, f) \tag{10.29}$$

which is continuous, because by the Cauchy-Schwartz inequality, and Proposition 10.3,

$$|l_0(\psi)| \leq \|\psi\|_2\|f\|_2 \leq C\|\psi\|_0\|f\|_2 \leq M\|\psi\|_0 \tag{10.30}$$

with $M = C\|f\|_2$. Therefore l_0 extends to a linear bounded functional l on $\mathcal{H}$, and

$$|l(\psi)| \leq M\|\psi\|\,, \qquad \psi \in \mathcal{H}\,. \tag{10.31}$$

The **Riesz representation theorem** now gives the existence of an element $U \in \mathcal{H}$, such that

$$l(\psi) = \langle \psi, U\rangle \tag{10.32}$$

so in particular for all $\psi \in \mathrm{C}_0^\infty(\Omega)$,

$$(\psi, f) = l_0(\psi) = \langle \psi, U\rangle = (L^*\psi, L^*U) = (L^*\psi, u) \tag{10.33}$$

where $u = L^*U \in \mathrm{L}^2(\Omega)$. The assignment $K : f \mapsto u$ is linear, so it remains to show that $K : \mathrm{L}^2(\Omega) \to \mathrm{L}^2(\Omega)$ is bounded. We have

$$\|Kf\|_{\mathrm{L}^2(\Omega)} = \|u\|_{\mathrm{L}^2(\Omega)} = \|L^*U\|_{\mathrm{L}^2(\Omega)} = \|U\|_{\mathcal{H}} = \|l\| \leq C\|f\|_2\,. \tag{10.34}$$

10.2.1. Comments on the proof of Proposition 10.3

The proof of the key estimate, in this generality, relies on a fair amount of complex analysis.

First note that in view of Plancherel's identity the claim is that for any smooth function $\psi \in \mathrm{C}_0^\infty(\Omega)$,

$$\|\psi\|_{\mathrm{L}^2(\Omega)} = \|\hat{\psi}\|_{\mathrm{L}^2(\mathbb{R}^d)} \leq C\|L^*\psi\|_{\mathrm{L}^2(\Omega)} \tag{10.35}$$

and

$$\|L^*\psi\|_{\mathrm{L}^2(\Omega)} = \|\widehat{L^*\psi}\|_{\mathrm{L}^2(\mathbb{R}^d)} \qquad \widehat{L^*\psi}(\xi) = Q(\xi)\hat{\psi}(\xi) \tag{10.36}$$

where Q is the characteristic polynomial of L^*.

The approach is to reduce the estimate to an inequality for holomorphic functions and polynomials. For simplicity consider the $d = 1$-dimensional case. Suppose $f \in \mathrm{L}^2(\mathbb{R})$ is a function supported on the interval $[-M, M]$. Then

$$\hat{f}(\xi) = \int_{-M}^{M} f(x)e^{-2\pi i x\xi}\mathrm{d}x\,. \tag{10.37}$$

We can then extend $\hat{f}$ to a holomorphic function in the complex plane $\mathbb{C}$,

$$\hat{f}(\xi + i\eta) = \int_{-M}^{M} f(x)e^{2\pi x\eta}e^{-2\pi i x\xi}\mathrm{d}x\,, \tag{10.38}$$

and by Plancherel's identity, viewing the left hand side as the Fourier transform of the function $f(x)e^{2\pi x\eta}$ (for $|x| \leq M$) for fixed $\eta > 0$,

$$\int_{-\infty}^{\infty} |\hat{f}(\xi + i\eta)|^2\mathrm{d}\xi \leq e^{4\pi M\eta}\int_{-\infty}^{\infty} |f(x)|^2\mathrm{d}x\,. \tag{10.39}$$

Given $\Omega \subset \mathbb{R}$, and $\psi \in C_0^\infty(\Omega)$, we may apply this consideration to the function

$$f(x) = (L^*\psi)(x) \tag{10.40}$$

with $M >$ chosen such that $\Omega \subset [-M, M]$. Since $\hat{f}(\xi) = Q(\xi)\hat{\psi}(\xi)$ this yields

$$\int_{-\infty}^{\infty} |Q(\xi + i\eta)\hat{\psi}(\xi + i\eta)|^2 \mathrm{d}\xi \leq e^{4\pi M\eta} \int_{-\infty}^{\infty} |L^*\psi|^2 \mathrm{d}x\,. \tag{10.41}$$

and by substituting $\xi = \xi' + \cos(\theta)$, and setting $\eta = \sin(\theta)$, we obtain in particular

$$\int_{-\infty}^{\infty} |Q(\xi + \cos\theta + i\sin\theta)\hat{\psi}(\xi + \cos\theta + i\sin\theta)|^2 \mathrm{d}\xi \leq e^{4\pi M} \int_{-\infty}^{\infty} |L^*\psi|^2 \mathrm{d}x\,. \tag{10.42}$$

The point is that the left hand side, after integration in θ, actually bounds $|\hat{\psi}(\xi)|^2$, namely

$$|\hat{\psi}(\xi)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |Q(\xi + \cos\theta + i\sin\theta)\hat{\psi}(\xi + \cos\theta + i\sin\theta)|^2 \mathrm{d}\theta\,. \tag{10.43}$$

This peculiar inequality is a consequence of the following Lemma.

Lemma 10.4. *Suppose $P(z) = z^m + \ldots + a_1 z + a_0$ is a polynomial of degree m. If F is a holomorphic function on $\mathbb{C}$, then*

$$|F(0)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |P(e^{i\theta})F(e^{i\theta})|^2 \mathrm{d}\theta\,. \tag{10.44}$$

An application of the Lemma to $F(z) = \hat{\psi}(\xi + z)$, and $P(z) = Q(\xi + z)$, gives the inequality (10.43), provided we choose without loss of generality the leading order coefficient in Q to be 1. Finally, integrating in ξ, gives the desired estimate:

$$\begin{aligned}\int_{-\infty}^{\infty} |\hat{\psi}(\xi)|^2 &\leq \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} |Q(\xi + \cos\theta + i\sin\theta)\hat{\psi}(\xi + \cos\theta + i\sin\theta)|^2 \mathrm{d}\xi \mathrm{d}\theta \\ &\leq e^{4\pi M} \int_{-\infty}^{\infty} |L^*\psi|^2 \mathrm{d}x\,. \end{aligned} \tag{10.45}$$

It remains to understand the statement of the Lemma which in the case $P = 1$ reads

$$|F(0)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |F(e^{i\theta})|^2 \mathrm{d}\theta \tag{10.46}$$

This in turn is an immediate consequence of the mean value property of holomorphic functions, cf. Lecture 6, Proposition 6.4,

$$F(\zeta) = \frac{1}{2\pi} \int_0^{2\pi} F(\zeta + re^{i\theta}) \mathrm{d}\theta\,. \tag{10.47}$$

The general case actually follows from (10.46) with a suitable factorization of P.

Supplement: Mollification

Consider a function φ on $\mathbb{R}^d$ with the following properties:

1. φ is smooth
2. the support of φ is contained in the unit ball
3. $\varphi \geq 0$
4. $\int_{\mathbb{R}^d} \varphi(x)\mathrm{d}x = 1$

For instance,

$$\varphi(x) = \begin{cases} ce^{-\frac{1}{1-|x|^2}} & |x| \leq 1 \\ 0 & |x| \geq 1 \end{cases} \tag{10.48}$$

is a function with these properties, provided we choose c such that its integral is 1.

Given such a function φ we obtain an approximation to the identity defined by

$$K_\delta(x) = \frac{1}{\delta^d}\varphi(x/\delta)\,. \tag{10.49}$$

Given a bounded function g, supported on a bounded set, the convolution $g * K_\delta$ is also bounded, and supported on a bounded set,

$$(g * K_\delta)(x) = \int_{\mathbb{R}^d} g(y)K_\delta(x-y)\mathrm{d}y\,. \tag{10.50}$$

Indeed if g is supported in Ω, then $g * K_\delta$ is supported in $\Omega_1 = \{x \in \mathbb{R}^d : d(x,\Omega) < 1\}$, provided $\delta \leq 1$. Moreover, if $|g(x)| \leq M$, then

$$|(g * K_\delta)(x)| \leq \sup|g| \int_{\mathbb{R}^d} K_\delta(y)\mathrm{d}y \leq M \tag{10.51}$$

independently of $\delta > 1$. Finally, $g * K_\delta$ is smooth, because we can differentiate under the integral, and all derivatives are also supported in Ω_1.

Supplementary Problems

1. Suppose F and G are two integrable functions on a bounded interval $[a,b]$. Show that $F' = G$ in the weak sense if and only if F is absolutely continuous and $F'(x) = G(x)$ for almost every x.

   Hint: By definition

$$\int_a^b G(x)\psi(x)\mathrm{d}x = -\int_a^b F(x)\psi'(x)\mathrm{d}x \tag{10.52}$$

for any smooth function ψ of compact support in $[a, b]$. Now choose ψ_n to be an approximation to the piece-wise linear function $\psi^{(h)}_{\alpha,\beta}$, where

$$\psi^{(h)}_{\alpha,\beta}(x) = \begin{cases} 1 & \alpha \le x \le \beta \\ 0 & 0 \le x \le \alpha - h, \text{ or } \beta + h \le x \le 1 \\ \frac{1}{h}(x - \alpha + h) & \alpha - h \le x \le \alpha \\ 1 - \frac{1}{h}(x - \beta) & \beta \le x \le \beta + h \end{cases} \tag{10.53}$$

and evaluate the integrals.

2. Let $\mathcal{H}$ denote a Hilbert space $\mathcal{H}$, and $\mathcal{L}(\mathcal{H})$ the vector space of all bounded linear operators on $\mathcal{H}$. Given $\mathcal{L}(\mathcal{H})$, we define the operator norm by

$$\|T\| = \inf\{B : \|Tv\| \le B\|v\|\,, \text{ for all } v \in \mathcal{H}\}\,. \tag{10.54}$$

   a) Show that $\|T_1 + T_2\| \le \|T_1\| + \|T_2\|$ whenever $T_1, T_2 \in \mathcal{L}(\mathcal{H})$.
   b) Prove that $d(T_1, T_2) = \|T_1 - T_2\|$ defines a metric in $\mathcal{L}(\mathcal{H})$.
   c) Show that $\mathcal{L}(\mathcal{H})$ is complete in the metric d.

3. There are several senses in which a sequence of bounded operators $\{T_n\}$ can converge to a bounded operator T in a Hilbert space $\mathcal{H}$. First there is **norm convergence**, that is $\|T_n - T\| \to 0$. Next, there is **strong convergence**, that requires that

$$T_n f \to T f \qquad (n \to \infty) \tag{10.55}$$

for every $f \in \mathcal{H}$. Finally, there is **weak convergence** that requires that

$$(T_n f, g) \to (T f, g) \qquad (n \to \infty) \tag{10.56}$$

for every pair of vectors $f, g \in \mathcal{H}$.

   a) Show by examples that weak convergence does not imply strong convergence, nor does strong convergence imply norm convergence.
   b) Show that for any bounded operator T there is a sequence $\{T_n\}$ of bounded operators of finite rank so that $T_n \to T$ strongly as $n \to \infty$.

Lecture 11.

Dirichlet's problem

Recommended Reading

(Stein and Shakarchi, *Real analysis*, Chapter 5, Sections 4).

Let us return to the boundary value problem for Laplace's equation in two dimensions:

Dirichlet problem: Given a bounded open set Ω in $\mathbb{R}^2$, and a continuous function f on the boundary $\partial\Omega$, find a function u such that

$$\triangle u = 0 \qquad : \text{in } \Omega \tag{11.1a}$$

$$u = f \qquad : \text{on } \partial\Omega\,. \tag{11.1b}$$

We have solved this problem in special cases:

In Lecture 3 we have seen that in the case of the unit disc, $\Omega = D = \{x : |x| < 1\}$, in $\mathbb{R}^2$, the solution is given by

$$u(r\cos\theta, r\sin\theta) = \frac{1}{2}\int_{-\pi}^{\pi} f(\varphi)P_r(\theta-\varphi)\mathrm{d}\varphi \tag{11.2}$$

where P_r is the Poisson kernel for the unit disc.

We have also obtained, in Lecture 6, the solution to the problem in the case of the upper half plane, $\Omega = \{(x,y) : y > 0\}$, which is given similarly by

$$u(x,y) = \int_{-\infty}^{\infty} \mathcal{P}_y(x-t)f(t)\mathrm{d}t \tag{11.3}$$

where $\mathcal{P}_y$ is the Poisson kernel for the upper half-plane.

In general, however, namely when Ω or f do not have special symmetries, *there are no explicit solutions*, and other methods are need to prove the existence and uniqueness of solutions.

Another approach is given by the idea that the solution to (11.1) should be given by a function which, compared to any other function with the prescribed boundary values, *minimizes the energy*:

$$\mathcal{D}(u) = \int_\Omega |\nabla u|^2 \tag{11.4}$$

Moreover, given the positivity of the **Dirichlet energy** $\mathcal{D}(u) \geq 0$, a minimizer should always exist.

Lemma 11.1. *Suppose there exists a function $u \in \mathrm{C}^2(\overline{\Omega})$ that minimizes $\mathcal{D}(u)$ among all $U \in \mathrm{C}^2(\overline{\Omega})$ with $U|_{\partial\Omega} = f$. Then u is harmonic, $\triangle u = 0$ in Ω.*

Proof. For functions F and G in $\mathrm{C}^2(\overline{\Omega})$ define the inner product

$$\langle F, G\rangle = \int_\Omega \nabla F \cdot \overline{\nabla G}\,. \tag{11.5}$$

If $v \in \mathrm{C}^2(\overline{\Omega})$ is a function with trivial boundary data $v|_{\partial\Omega} = 0$, then $u + \epsilon v$ has the prescribed boundary values f, so by assumption

$$\mathcal{D}(u + \epsilon v) \geq \mathcal{D}(u)\,. \tag{11.6}$$

However,

$$\mathcal{D}(u + \epsilon v) = \langle u + \epsilon v, u + \epsilon v\rangle = \mathcal{D}(u) + \epsilon\langle u, v\rangle + \epsilon\langle v, u\rangle + \epsilon^2\mathcal{D}(v) \tag{11.7}$$

which shows that

$$\epsilon\langle u, v\rangle + \epsilon\langle v, u\rangle + \epsilon^2\mathcal{D}(v) \geq 0\,. \tag{11.8}$$

Since ϵ can be both positive or negative, this can happen only if $\mathrm{Re}\langle u, v\rangle = 0$. Similarly, by the same argument for the function $u + i\epsilon v$, we infer that $\mathrm{Im}\langle u, v\rangle = 0$.

Therefore, by integration by parts for all $v \in \mathrm{C}^2(\overline{\Omega})$ with $v|_{\partial\Omega} = 0$,

$$0 = \langle u, v\rangle = -\int_\Omega (\triangle u)\,\overline{v} \tag{11.9}$$

which implies that $\triangle u = 0$ in Ω. □

This Lemma is a strong indication that the problem can be solved using **Dirichlet's principle**, or more generally by an "**action principle**" — namely the idea that the solution to the boundary value problem can be found as the function that minimizes the Dirichlet energy, or more generally an action — however, there are also indications that it cannot:

Example 11.1. Consider the simpler one-dimensional problem of minimizing the integral

$$D(\varphi) = \int_{-1}^{1} |x\varphi'(x)|^2 \mathrm{d}x \tag{11.10}$$

among all C^1 functions on $[-1, 1]$ that satisfy $\varphi(-1) = -1$, and $\varphi(1) = 1$.

Claim: The minimum value of the integral is zero, however there is no C^1 function with the prescribed boundary values that achieves this minimum.

Let ψ be any smooth function on the real line with $\psi(x) = -1$ for $x \leq -1$, and $\psi(x) = 1$ for $x \geq 1$, and $\psi'(x) \geq 0$ for $-1 \leq x \leq 1$. Then set, for each $0 < \epsilon < 1$,

$$\varphi_\epsilon(x) = \begin{cases} 1 & x \geq \epsilon \\ \psi(x/\epsilon) & -\epsilon < x < \epsilon \\ -1 & x \leq -\epsilon \end{cases} \tag{11.11}$$

Then φ_ϵ has the desired boundary values, and

$$D(\varphi_\epsilon) = \int_{-\epsilon}^{\epsilon} |x\psi'(x/\epsilon)/\epsilon|^2 \mathrm{d}x = \epsilon \int_{-1}^{1} |y\psi'(y)|^2 \mathrm{d}y \to 0 \qquad (\epsilon \to 0)\,. \tag{11.12}$$

However, if $D(\varphi) = 0$ for some function $\varphi \in \mathrm{C}^1([-1,1])$ then $\varphi'(x) = 0$ which means that φ is constant, and cannot satisfy the boundary conditions.

This example shows that positivity of the integral alone does not imply the existence of a minimizing function. Another concern is that the integral may not be finite for a solution.

Example 11.2. There are functions f on the circle so that the solution u to the Dirichlet problem on the unit disc given by (11.2) has *infinite* Dirichlet energy $\mathcal{D}(u)$. Examples include functions f which are continuous but not differentiable.

Despite these difficulties this approach does indeed lead to success, if applied appropriately. As already suggested in the proof of Lemma 11.1, we can view the space $\mathrm{C}^1(\overline{\Omega})$ of "competing functions" (for the minimum) as a pre-Hilbert space $\mathcal{H}_0$ endowed with the inner product (11.5). The solution will then be found in the *completion* $\mathcal{H}$ of $\mathcal{H}_0$, and this will require some analysis of this problem in $\mathrm{L}^2(\Omega)$.

More precisely, consider $\mathcal{H}_0 = \mathrm{C}^1(\overline{\Omega})$ endowed with the inner product (11.5), then the corresponding norm is

$$\|u\| = \|\nabla u\|_{\mathrm{L}^2(\Omega)}\,. \tag{11.13}$$

Note that $\|u\| = 0$ implies that u is a constant, thus $\mathcal{H}_0$ should really be defined as equivalence classes of continuously differentiable functions on $\overline{\Omega}$ which differ by constants.

Now let $\mathcal{H}$ be the completion of $\mathcal{H}_0$, and let S_0 be the linear subspace of $\mathrm{C}^1(\overline{\Omega})$ consisting of functions that vanish on the boundary of Ω. Note that distinct elements in S_0 remain distinct in $\mathcal{H}_0$ under the above equivalence relation, so S_0 can be identified with a subspace of $\mathcal{H}_0$. Moreover, let S be the closure of S_0 in $\mathcal{H}$, and let P_S be the orthogonal projection of $\mathcal{H}$ onto S.

In order to solve Dirichlet's problem in these spaces, let us first make the additional assumption that f is the restriction to $\partial\Omega$ of some function $F \in \mathrm{C}^1(\overline{\Omega})$,

$$f = F|_{\partial\Omega}\,. \tag{11.14}$$

Then we seek the solution u as the limit, in a suitable sense, of a sequence $\{u_n\}$ of functions $u_n \in \mathrm{C}^1(\overline{\Omega})$ with $u_n|_{\partial\Omega} = F|_{\partial\Omega}$, with the property that $\|u_n\|$ converges to the minimum value of the Dirichlet energy. This means that $v_n = F - u_n \in S_0$, and $u_n = F - v_n$ is a sequence that minimizes the distance from F to S_0 in $\mathcal{H}$.

We now apply the main Lemma about closed subspaces and orthogonal projections from Lecture 9.2.3. Since $v_n \in S_0 \subset S$, and S is closed, we know from (the proof of) Lemma 9.1 that the limit $v_n \to v \in S$ exists, and hence also $u = F - v = \lim_{n\to\infty} u_n$ exists, and satisfies

$$\|u\| = \|F - v\| = \inf_n \|F - v_n\|\,. \tag{11.15}$$

In fact,

$$u = F - v \in S^{\perp}\,, \qquad v = P_S(F)\,, \qquad u = F - P_S(F)\,. \tag{11.16}$$

We could now proceed as for Proposition 10.3 to show that for any function $v \in S_0$,

$$\int_\Omega |v|^2 \le C \int_\Omega |\nabla v|^2\,. \tag{11.17}$$

However, we will give a simpler proof of this special case in Lecture 13. At any rate, we can apply this inequality to $v_n - v_m$ which yields that

$$\|v_n - v_m\|_{\mathrm{L}^2(\Omega)} \le C\|v_n - v_m\| \to 0 \qquad (n, m \to \infty)\,, \tag{11.18}$$

so that v_n is a Cauchy sequence in $\mathrm{L}^2(\Omega)$, and v_n also converges to v in $\mathrm{L}^2(\Omega)$, hence also u_n to u in $\mathrm{L}^2(\Omega)$. In particular,

$$S \subset \mathcal{H} \subset \mathrm{L}^2(\Omega)\,. \tag{11.19}$$

Finally, let us show that u is (what we will define to be) **weakly harmonic**: In view of (11.16) we have, on the one hand, that for any $\psi \in \mathrm{C}_0^\infty(\Omega) \subset S_0$,

$$\langle u, \psi\rangle = 0\,. \tag{11.20}$$

On the other hand, by integration by parts,

$$\langle u_n, \psi\rangle = \int_\Omega \nabla u_n \cdot \overline{\nabla\psi} = -\int_\Omega u_n \overline{\triangle\psi} = -(u_n, \triangle\psi)_{\mathrm{L}^2(\Omega)} \tag{11.21}$$

which shows that

$$(u, \triangle\psi)_2 = \lim_{n\to\infty}(u_n, \triangle\psi)_2 = -\lim_{n\to\infty}\langle u_n, \psi\rangle = -\langle u, \psi\rangle = 0\,. \tag{11.22}$$

For the resolution of Dirichlet's problem with this approach we thus still need to answer several questions:

1. What can we say about weakly harmonic functions? Are weakly harmonic functions harmonic in the classical sense?
2. The purported solution u to the problem is here constructed as the limit of a sequence $\{u_n\}$ of continuous functions on $\overline{\Omega}$ and $u_n|_{\partial\Omega} = f$ for each n. But is u itself continuous in $\overline{\Omega}$, and does it satisfy $u|_{\partial\Omega} = f$?
3. We restricted our argument to the case that the boundary data f is induced by a function $F \in \mathrm{C}^1(\overline{\Omega})$. How can we remove this restriction?

Lecture 12.

Harmonic functions

Recommended Reading

(Stein and Shakarchi, *Real analysis*, Chapter 5, Sections 4.1).

Let $\Omega \subset \mathbb{R}^d$ be an open set in $\mathbb{R}^d$. A function u is called **harmonic** in Ω if $u \in \mathrm{C}^2(\Omega)$ solves

$$\triangle u = 0\,. \tag{12.1}$$

A weak solution to this equation is called **weakly harmonic**, namely $u \in \mathrm{L}^2(\Omega)$ is weakly harmonic in Ω if for every $\psi \in \mathrm{C}_0^\infty(\Omega)$ it holds

$$(u, \triangle\psi) = 0\,. \tag{12.2}$$

The remarkable fact is that:

Theorem 12.1. *Any weakly harmonic function u in Ω can be redefined on a set of measure zero resulting in a function which is harmonic in Ω.*

This is closely related to the **mean-value property** of harmonic functions:

Theorem 12.2. *If u is harmonic in Ω, then u satisfies the* ***mean-value property****, namely for every ball $B(x_0)$ centered at x_0 such that $\overline{B(x_0)} \subset \Omega$, it holds*

$$u(x_0) = \frac{1}{|B(x_0)|}\int_{B(x_0)} u(x)\mathrm{d}x\,. \tag{12.3}$$

Conversely, *a continuous function in Ω satisfying the mean-value property is harmonic.*

We have already seen in Lecture 6 that harmonic functions, in two dimensions, satisfy the mean value property. The proof that this holds in any dimension $d \geq 2$, relies on Green's formula, and a choice of test functions which relates to the fundamental solutions of the Laplacian; cf. Part **??**. In this lecture we will focus on the converse statement.

A consequence is the simplest version of the **maximum principle**.

Corollary 12.3. *Suppose Ω is a bounded open set. If u is continuous on $\overline{\Omega}$ and harmonic in Ω, then*

$$\max_{x\in\overline{\Omega}} |u(x)| = \max_{\partial\Omega} |u(x)|\,, \tag{12.4}$$

namely the maximum is attained on the boundary $\partial\Omega = \overline{\Omega} \setminus \Omega$.

Proof. Since both $\overline{\Omega}$, and $\partial\Omega$ is compact, both maxima are attained. Suppose then that the maximum is attained at an interior point $x_0 \in \Omega$. By the mean-value property

$$|u(x_0)| \leq \frac{1}{|B(x_0,r)|}\int_{B(x_0,r)} |u(x)|\mathrm{d}x \tag{12.5}$$

for every ball $\overline{B(x_0,r)} \subset \Omega$. We know that $|u(x)| \leq |u(x_0)|$ throughout $B(x_0,r)$, but if $|u(x')| < |u(x_0)|$ at some point $x' \in B(x_0,r)$, then by continuity,

$$\frac{1}{|B(x_0,r)|}\int_{B(x_0,r)} |u(x)|\mathrm{d}x < |u(x_0)|\,, \tag{12.6}$$

which contradicts the above, so $|u(x)| = |u(x_0)|$ for all $x \in B(x_0,r)$. This is true for all $B(x_0,r) \subset \Omega$, and choosing r_0 to be the least upper bound of the set of radii for which this inclusion holds, we get by continuity $|u(\tilde{x})| = |u(x_0)|$, where $\tilde{x} \in \overline{B(x_0,r_0)} \cap \partial\Omega$.
□

Let us first try to understand the converse direction in Theorem 12.2. For this purpose, let $\varphi(y) = \Phi(|y|)$ be a continuous radial function on $\mathbb{R}^d$, that vanishes outside the unit ball, $\Phi(|y|) = 0$ for $|y| > 1$, and satisfies the condition that $\int \varphi(y)\mathrm{d}y = 1$.

Lemma 12.4. *Suppose u is continuous and satisfies the mean value property* (12.3) *in Ω, then*

$$u(x_0) = \int_{\mathbb{R}^d} u(x_0 - ry)\varphi(y)\mathrm{d}y = \int_{\mathbb{R}^d} u(x_0 - y)\varphi_r(y)\mathrm{d}y = (u * \varphi_r)(x_0) \tag{12.7}$$

whenever $\overline{B(x_0,r)} \subset \Omega$, where

$$\varphi_r(y) = r^{-d}\varphi(y/r)\,. \tag{12.8}$$

In particular, if $\varphi \geq 0$ is *smooth*, then φ_r is an approximation to the identity as described in Lecture 10, and $u * \varphi_r$ is a **smooth regularisation** (or mollification) of u; the Lemma states that whenever $x \in \Omega$, and $r < d(x,\Omega)$,

$$u(x) = (u * \varphi_r)(x)\,. \tag{12.9}$$

In other words, a continuous function in Ω which satisfies the mean value property equals its own regularization, in particular such a function is smooth.

Proof of Lemma 12.4. The idea is to approximate in a suitable sense

$$\int_{\mathbb{R}^d} u(x_0 - ry)\varphi(y)\mathrm{d}y \sim \sum_{j=1}^{N} \Phi(j/N)\int_{B(j)\setminus B(j-1)} u(x_0 - ry)\mathrm{d}y \tag{12.10}$$

where $N \in \mathbb{N}$ is a suitably large integer, and $B(j) = \{y : |y| < j/N\}$. Then by the mean value property,

$$\int_{B(j)\setminus B(j-1)} u(x_0 - ry)\mathrm{d}y = u(x_0)(|B(j)| - |B(j-1)|) \tag{12.11}$$

and hence

$$\sum_{j=1}^{N}\Phi(j/N)\int_{B(j)\setminus B(j-1)}u(x_0-ry)\mathrm{d}y\sim u(x_0)\int_{\mathbb{R}^d}\varphi(y)\mathrm{d}y=u(x_0)\,. \tag{12.12}$$

One can make this argument precise, by proving that for any bounded function ψ on $\mathbb{R}^d$,

$$\int_{\mathbb{R}^d}\varphi(y)\psi(y)\mathrm{d}y=\lim_{N\to\infty}\sum_{j=1}^{N}\Phi(j/N)\int_{B(j)-B(j-1)}\psi(y)\mathrm{d}y\,. \tag{12.13}$$

□

Proof of the converse in Theorem 12.2. In view of (12.9) it remains to show that a smooth function which satisfies the mean value property is harmonic. First note that by Taylor's theorem, for every $x_0\in\Omega$,

$$u(x_0+x)-u(x_0)=\nabla u(x_0)\cdot x+\frac{1}{2}(x\cdot\nabla)^2u(x_0)+\epsilon(x) \tag{12.14}$$

where $\epsilon(x)=\mathcal{O}(|x|^3)$. If we integrate this identity in x over the ball of radius r, then the left hand side vanishes by the mean value property. For the right hand side we note that by symmetry,

$$\int_{|x|\le r}x_j\mathrm{d}x=0\qquad\int_{|x|\le r}x_ix_j\mathrm{d}x=0\qquad(i\neq j) \tag{12.15}$$

$$\int_{|x|\le r}x_j^2\mathrm{d}x=r^2\int_{|x|\le r}(x_j/r)^2\mathrm{d}x=r^{2+d}\int_{|x|\le 1}x_j^2\mathrm{d}x=cr^{2+d}\,, \tag{12.16}$$

for some constant $c>0$ independent of j. Therefore

$$0=\frac{1}{2}\triangle u(x_0)r^{2+d}+\int_{|x|\le r}\epsilon(x)\mathrm{d}x\,,\qquad\mathcal{O}\Big(\int_{|x|\le r}\epsilon(x)\mathrm{d}x\Big)=\mathcal{O}(r^{d+3}) \tag{12.17}$$

and thus, after dividing by r^d, and taking the limit $r\to 0$, $\triangle u(x_0)=0$. □

Proof of Theorem 12.1. Let us assume that u is weakly harmonic in Ω. For each $\epsilon>0$, define the open set

$$\Omega_\epsilon=\{x\in\Omega:d(x,\partial\Omega)>\epsilon\}\,. \tag{12.18}$$

Then the regularisation $u_r=u*\varphi_r$ is defined in Ω_ϵ for $r<\epsilon$, and is a smooth function there. Now for any $\psi\in\mathrm{C}_0^\infty(\Omega_\epsilon)$ we have

$$\begin{aligned}(u_r,\triangle\psi)&=\int_\Omega u_r(x)\triangle\psi(x)\mathrm{d}x\\&=\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}u(x-ry)\varphi(y)\triangle\psi(x)\mathrm{d}y\mathrm{d}x\\&=\int_{\mathbb{R}^d}(u,\triangle\psi_{ry})_2\varphi(y)\mathrm{d}y=0\end{aligned} \tag{12.19}$$

because u is weakly harmonic, $(u, \psi_{ry}) = 0$, where $\psi_{ry}(x) = \psi(x + ry) \in C_0^\infty(\Omega)$, for $r < \epsilon$, and $|y| < 1$. This shows that u_r is weakly harmonic in Ω_ϵ, and since it is also smooth, it is harmonic in Ω_ϵ.

Next we will prove that whenever $x \in \Omega_\epsilon$, and $r_1 + r_2 < \epsilon$,

$$u_{r_1} = u_{r_2} \,. \tag{12.20}$$

We have just shown that u_{r_1} is harmonic, so it satisfies the mean value property, and thus by Lemma 12.4 equals its own regularisation:

$$u_{r_1} * \varphi_{r_2} = u_{r_1} \tag{12.21}$$

Since convolutions are commutative, we get that

$$u_{r_1} = (u * \varphi_{r_1}) * \varphi_{r_2} = (u * \varphi_{r_2}) * \varphi_{r_1} = u_{r_2} * \varphi_{r_1} = u_{r_2} \,. \tag{12.22}$$

Fixing r_2, we can take the limit $r_1 \to 0$, and recall that by the properties of an approximation to the identity,

$$\lim_{r_1 \to 0} u_{r_1}(x) = \lim_{r_1 \to 0} (u * \varphi_{r_1})(x) = u(x) \qquad \text{almost everywhere}\,, \tag{12.23}$$

which shows that $u(x) = u_{r_2}(x)$ for almost every $x \in \Omega_\epsilon$. Thus u can indeed be corrected on a set of measure zero, by setting it equal to u_{r_2}, for the result to be a harmonic function. Moreover, $\epsilon > 0$ is arbitrary.

□

The fact that a harmonic function equals its regularisation also shows:

Corollary 12.5. *Every harmonic function is indefinitely differentiable.*

Another consequence of the argument above is:

Corollary 12.6. *Suppose $\{u_n\}$ is a sequence of harmonic functions in Ω that converge uniformly on compact subsets of Ω to a function u as $n \to \infty$. Then u is harmonic.*

Proof. Since u_n is harmonic, it satisfies the mean value property

$$u_n(x_0) = \frac{1}{|B(x_0, r)|} \int_{B(x_0,r)} u_n(x) \mathrm{d}x \tag{12.24}$$

for every $x_0 \in \Omega$, and $\overline{B(x_0, r)} \subset \Omega$. Thus by uniform convergence it follows that also u has this property, and hence u is harmonic. □

Problems

1. Suppose u is harmonic on the punctured unit disc

$$\Omega_\circ = \{x \in \mathbb{R}^2 : 0 < |x| < 1\}\,. \tag{12.25}$$

   a) Show that if u is also continuous at the origin, then u is harmonic throughout the unit disc $\Omega = \{|x| < 1\}$.

      Hint: Show that u is weakly harmonic.

   b) Show that the Dirichlet problem for the punctured unit disc $\Omega_\circ$ is in general not solvable.

Lecture 13.

Dirichlet's problem in two dimensions: boundary regularity

Recommended Reading

(Stein and Shakarchi, *Real analysis*, Chapter 5, Section 4).

In this lecture we return to the boundary value problem in two dimensions:

Dirichlet problem: Let Ω be an open bounded set in $\mathbb{R}^2$. Given a continuous function f on the boundary $\partial\Omega$, find a function u that is continuous on $\overline{\Omega}$, harmonic in Ω, and such that $u|_{\partial\Omega} = f$.

As a consequence of the maximum principle the solution to this problem is *unique*:

Proof. Suppose u_1, and u_2 are solutions to the Dirichlet problem, then also $u_1 - u_2$ is harmonic in Ω, so by Corollary 12.3,

$$\max_{x\in\overline{\Omega}} |u_1(x) - u_2(x)| = 0\,, \tag{13.1}$$

because $u_1(x) = u_2(x) = f(x)$ on the boundary $x \in \partial\Omega$.

□

For *existence*, we have so far obtained a *weak solution* in Lecture 11.

Recall that for this purpose we have considered the completion $\mathcal{H}$ of the pre-Hilbert space $\mathcal{H}_0 = \mathrm{C}^1(\overline{\Omega})$, endowed with the inner product

$$\langle u, v\rangle = \int_\Omega \nabla u \cdot \overline{\nabla v} \mathrm{d}x\,. \tag{13.2}$$

In general, a completion $\mathcal{H}$ can be identified with the space of Cauchy sequences in $\mathcal{H}_0$. However, the following Lemma showed that given a Cauchy sequence $u_n \in \mathcal{H}_0$ with respect to the norm induced by (13.2), then

$$\|u_n - u_m\|_{\mathrm{L}^2(\Omega)} \leq c_\Omega \|u_n - u_m\|_{\mathcal{H}} \to 0 \tag{13.3}$$

hence $\{u_n\}$ is also a Cauchy sequence in $\mathrm{L}^2(\Omega)$, which is complete, hence $u = \lim_{n\to\infty} u_n \in \mathrm{L}^2(\Omega)$. In other words,

$$\mathcal{H} \subset \mathrm{L}^2(\Omega)\,. \tag{13.4}$$

Lemma 13.1. *Let Ω be an open bounded set in $\mathbb{R}^d$. Then for some constant c_Ω which only depends on Ω, we have for all $v \in \mathrm{C}^1(\overline{\Omega})$ with $v|_{\partial\Omega} = 0$,*

$$\int_\Omega |v(x)|^2 \mathrm{d}x \leq c_\Omega \int_\Omega |\nabla v(x)|^2 \mathrm{d}x \,. \tag{13.5}$$

Proof. Consider first the one-dimensional case $v \in \mathrm{C}^1(\overline{\Omega})$ where $\Omega = (a, b)$ is an interval in $\mathbb{R}$. If $f(a) = 0$, then $f(x) = \int_a^x f'(t)\mathrm{d}t$, and by Cauchy-Schwarz

$$|f(x)|^2 \leq |I| \int_I |f'(x)|^2 \mathrm{d}x \tag{13.6}$$

and the inequality follows after another integration with $c_\Omega = |I|^2$. The general inequality can be deduced from this special case by considering, for $(x_1, x') \in \Omega$, the slices $J(x') = \{x_1 : (x_1, x') \in \Omega\}$. The subset $J(x')$ of the real line can be written as a disjoint union of open intervals I_j, and so

$$\int_{I_j} |v(x_1, x')|^2 \mathrm{d}x_1 \leq |I_j|^2 \int_{I_j} |\nabla v(x_1, x')|^2 \mathrm{d}x_1 \tag{13.7}$$

which after summation in j, and integration in x' gives the inequality with $c_\Omega \leq d(\Omega)^2$, where $d(\Omega)$ is the diameter of Ω. □

Finally, we have made the additional assumption[1] that f is the restriction of some function $F \in \mathrm{C}^1(\overline{\Omega})$,

$$f = F|_{\partial\Omega} \,. \tag{13.8}$$

Then we have considered the sequence

$$u_n = F - v_n \,, \qquad v_n \in \mathrm{C}^1(\overline{\Omega}) \,, \qquad v_n|_{\partial\Omega} = 0 \,, \tag{13.9}$$

which *minimizes the Dirichlet energy*, and proven that v_n, and hence u_n, converges in $\mathcal{H}$, and in $\mathrm{L}^2(\Omega)$, to some $v \in \mathrm{L}^2(\Omega)$, and $u \in \mathrm{L}^2(\Omega)$, respectively. Moreover, we have shown that u is *weakly harmonic*.

It now follows from the results of Lecture 12, Theorem 12.1, that after possibly redefining the function $u \in \mathrm{L}^2(\Omega)$ on a set of measure zero, this function is *harmonic*, in particular $u \in \mathrm{C}^2(\Omega)$, and

$$\triangle u = 0 \,. \tag{13.10}$$

It remains to show that u is continuous *up to the boundary* on $\overline{\Omega}$, and that u satisfies the boundary condition $u = F$ on $\partial\Omega$.

It turns out that this can only be proven under some conditions on the nature of the boundary $\partial\Omega$.

[1] We will discuss the in supplement below how this assumption can be removed.

13.1. Boundary regularity and main theorem in two dimensions

Let T_0 be a triangle in the plane with two equal sides of length l which make an angle α at their common vertex. The length $l > 0$, and the angle $\alpha > 0$ are arbitrary small but fixed in the following discussion. We will denote by T triangles which are congruent to T_0, namely obtained from T_0 by translation and rotation.

Definition 13.1. A domain Ω satisfies the *outside-triangle condition*, if for every point $x \in \partial\Omega$ we can find a triangle T congruent to T_0 with vertex x, such that the interior of T lies outside of Ω, namely $T \cap \overline{\Omega} = \{x\}$ and $T \cap \Omega = \emptyset$.

Remark 13.1. Any domain Ω whose boundary is a polygonal curve satisfies the outside-triangle condition. Moreover, if the boundary $\partial\Omega$ is made up of *Lipschitz curves,* so in particular if the boundary is a C^1 curve, then the outside triangle condition is also satisfied.

For these domains the Dirichlet problem is always solvable.

Theorem 13.2. *Let $\Omega \subset \mathbb{R}^2$ be an open bounded domain which satisfies the outside-triangle condition for some $l, \alpha > 0$, and let f be any continuous function on $\partial\Omega$. Then there exists a unique solution to the Dirichlet problem u which is continuous on $\overline{\Omega}$ and satisfies $u|_{\partial\Omega} = f$.*

The theorem relies on a refinement of Lemma 13.1.

Proposition 13.3. *For any bounded open set $\Omega \subset \mathbb{R}^2$ that satisfies the outside-triangle condition there are constants $c_1 < 1$, and $c_2 > 1$ such that the following holds. For any $z \in \Omega$, and any $v \in \mathrm{C}^1(\overline{\Omega})$, with $v|_{\partial\Omega} = 0$, we have*

$$\int_{B(z,c_1\delta(z))} |v(x)|^2 \mathrm{d}x \leq C\delta^2(z) \int_{B(z,c_2\delta(z))\cap\Omega} |\nabla v(x)|^2 \mathrm{d}x \,, \tag{13.11}$$

where $\delta(z) = dist(z, \partial\Omega)$ and C is a constant that only depends on the diameter of Ω, and the parameters of the triangle T_0.

Let us now explain how this estimate can be used to prove Theorem 13.2. Let $u_n \in \mathrm{C}^1(\overline{\Omega})$ be as in (13.9). Then for each $v_n = F - u_n$ the bound (13.11) holds, and since $v_n \to v$ in $\mathcal{H}$ *and* $\mathrm{L}^2(\Omega)$, we have that (13.11) also holds for the limit

$$\int_{B(z,c_1\delta(z))} |(u-F)(x)|^2 \mathrm{d}x \leq C\delta^2 \int_{B(z,c_2\delta(z))\cap\Omega} |\nabla(u-F)(x)|^2 \mathrm{d}x \,. \tag{13.12}$$

We want to show that for any $y \in \partial\Omega$,

$$\lim_{z \to y, z \in \Omega} u(z) = F(y) \,. \tag{13.13}$$

In order to use the estimate, let us not consider the values $u(z)$ directly, but rather their averages over discs centered at z of radius $c_1\delta(z)$. For any function $f \in \mathrm{C}^1(\overline{\Omega})$, let us denote by

$$\bar{f}(z) = \frac{1}{4\pi(c_1\delta(z))^2} \int_{B(z,c_1\delta(z))} f(x) \mathrm{d}x \tag{13.14}$$

Then, first by Cauchy-Schwarz,

$$|\overline{u-F}(z)| \leq \frac{1}{\sqrt{4\pi}c_1\delta(z)}\Big(\int_{B(z,c_1\delta(z))}|(u-F)(x)|^2\mathrm{d}x\Big)^{1/2} \tag{13.15}$$

and then using (13.12), gives

$$|(\bar{u}-\bar{F})(z)|^2 \leq \frac{C}{4\pi c_1^2}\int_{B(z,c_2\delta(z))\cap\Omega}|\nabla(u-F)(x)|^2\mathrm{d}x\,. \tag{13.16}$$

Now since u is harmonic in Ω, it satisfies the mean value property, and so

$$\bar{u}(z) = u(z)\,. \tag{13.17}$$

Moreover note that $\delta(z) \leq |z-y|$, so $\delta(z) \to 0$ as $z \to y$. Therefore

$$|\bar{F}(z)-F(y)| \leq \frac{1}{4\pi(c_1\delta(z))^2}\int_{B(z,c_1\delta(z))}|F(x)-F(y)|\mathrm{d}x \leq \max_{x\in B(z,c_1\delta(z))}|F(x)-F(y)| \to 0 \tag{13.18}$$

as $z \to y$, because F is continuous up to the boundary. Furthermore, $u - F \in \mathcal{H}$, so in particular $\nabla(u-F)$ is square integrable on Ω, and hence

$$\int_{B(z,c_2\delta(z))\cap\Omega}|\nabla(u-F)(x)|^2\mathrm{d}x \to 0 \tag{13.19}$$

because $|B(z,c_1\delta(z))| \to 0$ as as $z \to y$. In summary,

$$|u(z)-F(y)| \leq |\bar{u}(z)-\bar{F}(z)| + |\bar{F}(z)-F(y)| \to 0\,, \qquad \text{as } z \to y \text{ in } \Omega\,. \tag{13.20}$$

13.2. Proof of Proposition 13.3

Let $z \in \Omega \subset \mathbb{R}^2$, and $\delta = \mathrm{dist}(z,\partial\Omega)$. By assumption Ω satisfies the outside-triangle condition for some $l > 0$, and $\alpha > 0$. We can assume that $\delta < l/2$, for otherwise the estimate already follows as in Lemma 13.1.

Choose $y \in \partial\Omega$ so that $\delta = |z-y|$. Let T be the triangle with vertex at y, and let β be the smaller of the angles that the line from y to z makes with the sides of length l of the triangle T; then $\beta \leq \pi - \alpha/2$. Now choose coordinates (x_1, x_2) in the plane so that y is at the origin, and both said lines make an angle γ with the x_2-axis; then $\gamma \geq \alpha/4$. In these coordinates

$$z = (-\delta\sin\gamma, \delta\cos\gamma)\,. \tag{13.21}$$

Let us choose $c_1 < \sin\gamma$, and consider the disc $B(z, c_1\delta)$. We construct a rectangle R as in Figure 13.1, so that in particular

$$B(z,c_1\delta) \subset R = [-\delta\sin\gamma - c_1\delta, -\delta\sin\gamma + c_1\delta] \times [-L, \delta\cos\gamma + c_1\delta]\,. \tag{13.22}$$

Here the rectangle R intersects the x_1-axis at the points $P_1 = (-a, 0)$ where $a = \delta\sin\gamma + c_1\delta$, and L is chosen so that the point

$$P_2 = (-a, -L) \in T\,. \tag{13.23}$$

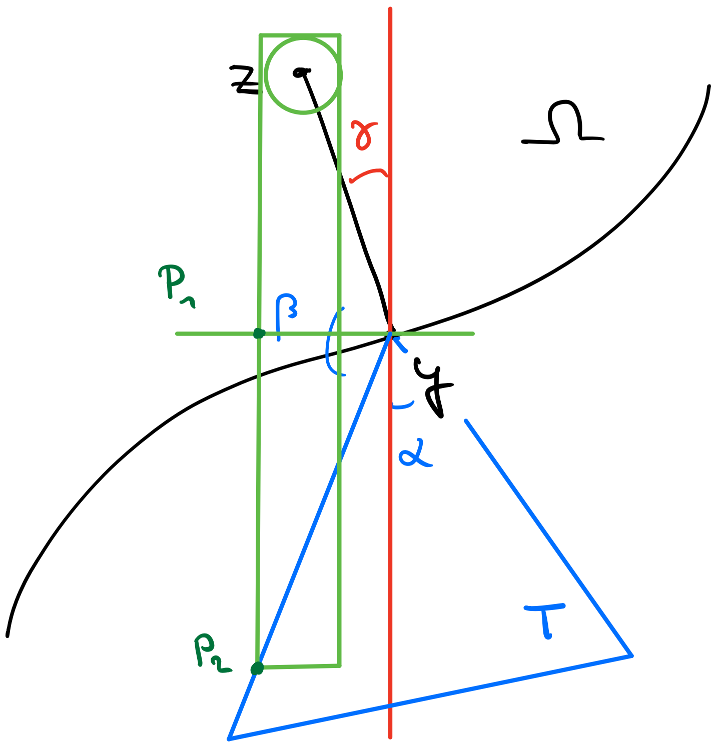

Figure 13.1.: Construction of the rectangle R using the outside triangle condition.

Since $\tan\gamma = a/L$, we find $L = a/\tan\gamma < 2\delta\cos\gamma$, because $c_1 < \sin\gamma$. In particular, the distance of P_2 from y is $L/\cos\gamma < 2\delta < l$. Moreover, the width of the rectangle is $\leq 2\delta$, and its height is $\leq 4\delta$.

Suppose $(x_1, x_2) \in R \cap \Omega$. We know that $(x_1, -L) \in T$, and T does not intersect with Ω, so there is point $(x_1, x_2') \in \partial\Omega$, and we can define the interval $I(x_1) = (x_2', \delta\cos\gamma + c_1\delta]$, which has length $|I(x_1)| \leq 4\delta$, and the property that

$$R \cap \Omega = \bigcup_{x \in [-a, -a+2c_1\delta]} \{x\} \times I(x)\,. \tag{13.24}$$

Integrating in $x_2 \in I(x)$, we have as in (13.6) that

$$\int_{I(x_1)} |v(x_2)|^2 \mathrm{d}x_2 \leq |I(x_1)|^2 \int_{I(x_1)} |\partial_{x_2} v(x_1, x_2)|^2 \mathrm{d}x_2 \tag{13.25}$$

and after integrating also in x_1 we get

$$\int_{R\cap\Omega} |v|^2 \leq (4\delta)^2 \int_{R\cap\Omega} |\nabla v|^2 \tag{13.26}$$

In view of (13.22), namely $B(z, c_1\delta) \subset R$, and Figure 13.1, namely $R \subset B(z, c_2\delta)$, this implies the statement of the Proposition, provided we choose $c_2 > 3$.

Supplement: Extension principle

The assumption that $f = F|_{\partial\Omega}$ for some $F \in \mathrm{C}^1(\overline{\Omega})$ can be removed with the help of an extension principle.

Lemma 13.4. *Let f be a continuous function on a compact subset Γ of $\mathbb{R}^d$. Then there exists a function G on $\mathbb{R}^d$ that is continuous so that $G|_{\partial\Gamma} = f$.*

The function G can be regularized, cf. Lecture 12, by defining a sequence of smooth functions $F_n = G * \varphi_{1/n}$, which has the property that $F_n \to f$ uniformly on Γ. Now solve the Dirichlet problem with boundary values $F_n|_{\partial\Omega}$ for each n, which yields a sequence of solutions $U_n \in \mathrm{C}^2(\Omega)$, continuous up to the boundary,

$$\triangle U_n = 0 \qquad U_n = F_n|_{\partial\Omega}\,. \tag{13.27}$$

We can now apply the maximum principle to see that U_n converges uniformly to a function u that is continuous on $\overline{\Omega}$,

$$\max_{x\in\overline{\Omega}} |U_n(x) - U_m(x)| = \max_{\partial\Omega} |F_n(x) - F_m(x)| \to 0 \qquad (n, m \to \infty) \tag{13.28}$$

and $u = \lim_{n\to\infty} U_n$ has the property that $u|_{\partial\Omega} = f$. Moreover by virtue of Corollary 12.6, the function u is harmonic.

Lecture 14.

Sobolev spaces in $\mathbb{R}^d$

Further Reading

(Evans, *Partial differential equations*, Chapter 5)

(Folland, *Introduction to Partial Differential Equations*, Chapter 6)

Given a function $f \in \mathrm{L}^2(\mathbb{R}^d)$, we say f is *weakly differentiable* up to order k, if for every multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, with $|\alpha| \le k$, there exists $g_\alpha \in \mathrm{L}^2(\mathbb{R}^d)$ so that

$$\Big(\frac{\partial}{\partial x}\Big)^\alpha f = g_\alpha \tag{14.1}$$

holds in the weak sense, namely

$$(-1)^\alpha \int_{\mathbb{R}^d} f\, \partial_x^\alpha \varphi \, \mathrm{d}x = \int_{\mathbb{R}^d} g_\alpha \, \varphi \, \mathrm{d}x\,, \qquad \varphi \in \mathrm{C}_0^\infty(\mathbb{R}^d)\,. \tag{14.2}$$

If g_α exists, we usually denote by $D^\alpha f = g_\alpha$ the **weak derivative** of f.

Remark 14.1. If a weak derivative exists, it is unique. Indeed, if g_α and g'_α are weak derivatives corresponding to the same multi-index α, then we have

$$\int_{\mathbb{R}^d} (g_\alpha - g'_\alpha)\varphi = 0 \tag{14.3}$$

which shows that $g_\alpha = g'_\alpha$ almost everywhere.

Remark 14.2. A *classical derivative* is also a weak derivative in the above sense. Indeed if a function f is k-times continuously differentiable, then (14.2) holds by integration by parts, with $D^\alpha f = \partial_x^\alpha f$.

In this lecture we will introduce the **Sobolev space** $\mathrm{H}^k(\mathbb{R}^d)$ of functions $u \in \mathrm{L}^2(\mathbb{R}^d)$ whose weak derivatives $D^\alpha u$ exist and are in $\mathrm{L}^2(\mathbb{R}^d)$, up to order $|\alpha| \le k$. This is a Hilbert space with the inner product

$$\langle u, v\rangle = \sum_{|\alpha|\le k} (D^\alpha u, D^\alpha v)_{\mathrm{L}^2(\mathbb{R}^d)}\,. \tag{14.4}$$

The corresponding norm is

$$\|u\|_k = \Big(\sum_{|\alpha|\le k} \|D^\alpha u\|^2_{\mathrm{L}^2(\mathbb{R}^d)}\Big)^{1/2}\,. \tag{14.5}$$

Remark 14.3. In a similar fashion we can introduce the spaces $\mathrm{H}^k(\Omega)$ on a domain $\Omega \subset \mathbb{R}^d$ consisting of functions u whose weak derivatives $D^\alpha u$ exist up to order k and are in $\mathrm{L}^2(\Omega)$. Here in the definition of weak derivative in (14.2) we replace $\mathrm{C}_0^\infty(\mathbb{R}^d)$ by $\mathrm{C}_0^\infty(\Omega)$.

The relevance of Sobolev spaces to PDE theory stems on one hand from the fact that they are Hilbert spaces (and hence the theory of linear operators can be applied), and on the other hand from the crucial fact that they embed into the space of classically continously differentiable functions, namely for any $m > d/2 + k$ we have

$$\mathrm{H}^m \subset \mathrm{C}^k\,. \tag{14.6}$$

This is the simplest version of the **Sobolev embedding theorem**. It says that if $f \in \mathrm{H}^m(\mathbb{R}^d)$, $m > d/2$, then f can be corrected on a set of measure zero so that f becomes continuous, and in fact $f \in \mathrm{C}^k(\mathbb{R}^d)$ for $k < m - d/2$.

In this sense Sobolev spaces allow us to measure the differentiability properties of functions in $\mathbb{R}^d$ purely in terms of L^2-norms.

Remark 14.4. The analogous statement of the Sobolev embedding theorem on *bounded* domains Ω requires slighly more care and we will return to the precise statement below.

For the proof of the Sobolev embedding theorem it is useful to characterize Sobolev spaces using the Fourier transform.

Exercise 14.1. Suppose $f \in \mathrm{L}^2(\mathbb{R}^d)$, and α a multi-index. Prove that the weak derivative $D^\alpha f$ exists if and only if $(2\pi i\xi)^\alpha \hat{f}(\xi) \in \mathrm{L}^2(\mathbb{R}^d)$. In other words, prove that there exists $g_\alpha \in \mathrm{L}^2(\mathbb{R}^d)$ so that (14.1) holds weakly, *if and only if*

$$(2\pi i\xi)^\alpha \hat{f}(\xi) = \widehat{g_\alpha}(\xi) \in \mathrm{L}^2(\mathbb{R}^d)\,. \tag{14.7}$$

Lemma 14.1. *For $m \in \mathbb{N}$,*

$$H^m(\mathbb{R}^d) = \left\{ f \in \mathrm{L}^2(\mathbb{R}^d) : (1 + |\xi|^2)^{m/2} \hat{f} \in \mathrm{L}^2(\mathbb{R}^d) \right\}. \tag{14.8}$$

In particular, the norms $\|f\|_m$ and $\|\hat{f}(\xi)(1 + |\xi|^2)^{m/2}\|_{\mathrm{L}^2}$ are equivalent.

Proof. If $f \in \mathrm{L}^2$ with $(1 + |\xi|^2)^{m/2}\hat{f}(\xi) \in \mathrm{L}^2(\mathbb{R}^d)$, then

$$(2\pi i\xi)^\alpha \hat{u}(\xi) \in \mathrm{L}^2(\mathbb{R}^d) \qquad |\alpha| \leq m\,, \tag{14.9}$$

because $|\xi^\alpha| \leq |\xi|^m \leq (1+|\xi|^2)^{m/2}$. Thus in view of Exercise 14.1 there exists $g_\alpha \in \mathrm{L}^2(\mathbb{R}^d)$, so that $\partial_x^\alpha f = g_\alpha$ in the weak sense, and $\|g_\alpha\| = \|\widehat{g_\alpha}\| = \|(2\pi i\xi)^\alpha \hat{f}\|$. In particular,

$$\sum_{|\alpha|\leq m} \|D^\alpha f\|^2 \leq C\|(1 + |\xi|^2)^{m/2}\hat{f}(\xi)\|^2\,.$$

Conversely, again by Exercise 14.1, for $f \in \mathrm{H}^m(\mathbb{R}^d)$, we know that $(2\pi i\xi)^\alpha \hat{f}(\xi) \in \mathrm{L}^2$ for all $|\alpha| \leq m$. Now we can find $C > 0$ so that

$$(1 + |\xi|^2)^m \leq C \sum_{|\alpha|\leq m} |\xi^\alpha|^2\,; \tag{14.10}$$

indeed the right hand side includes the term $|\alpha| = 0$, and so $\sum_{|\alpha|\le m} |\xi^\alpha|^2 \ge 1$. Moreover both $(1+|\xi|^2)^m$ and $\sum_{|\alpha|=m} |\xi^\alpha|^2$ are homogeneous of degree $2m$, and so their quotient is homogeneous of degree 0, and hence bounded for large ξ. We can take C to be the supremum of $(1+|\xi|^2)^m / \sum_{|\alpha|=m} |\xi^\alpha|^2$ on the unit sphere. Therefore

$$\|\hat{f}(\xi)(1+|\xi|^2)^{m/2}\|^2 \le C \sum_{|\alpha|\le k} \|D^\alpha u\|^2 .$$

□

The following result now relates the existence of sufficiently many weak derivatives to classical pointwise derivatives.

Theorem 14.2 (Sobolev embedding)**.** *If $m > k + d/2$, then*

$$\mathrm{H}^m(\mathbb{R}^d) \subset \mathrm{C}^k(\mathbb{R}^d) . \tag{14.11}$$

More precisely, if $m > k + d/2$, every element of H^m agrees with a C^k function almost everywhere, and

$$\sup_{|\alpha|\le k} \sup_{x\in\mathbb{R}^d} |\partial_x^\alpha f(x)| \le C_{s,k} \|f\|_m . \tag{14.12}$$

Proof. The main observation here is that

$$\int_{\mathbb{R}^d} (1+|\xi|^2)^{k-m} \mathrm{d}\xi < \infty \tag{14.13}$$

precisely when $m > k + d/2$. Indeed, evaluating this integral in polar coordinates, we see that

$$\int_{\mathbb{R}^d} (1+|\xi|^2)^{k-m} \mathrm{d}\xi = A_d \int_0^\infty (1+r^2)^{k-m} r^{d-1} \mathrm{d}r \tag{14.14}$$

which is finite, provided $2k - 2m + d - 1 < -1$, i.e. $m > k + d/2$.

Suppose now that $m > k + d/2$, $|\alpha| \le k$, and $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then by Fourier inversion, and the Cauchy-Schwarz inequality,

$$\begin{aligned}\sup_x |\partial_x^\alpha \varphi| \le& \int_{\mathbb{R}^d} |(2\pi i\xi)^\alpha \widehat{\varphi}(\xi)| \mathrm{d}\xi \\ \le& \Big(\int_{\mathbb{R}^d} (1+|\xi|^2)^{k-m} \mathrm{d}\xi\Big)^{1/2} \Big(\int_{\mathbb{R}^d} (1+|\xi|^2)^m |\widehat{\varphi}(\xi)|^2 \mathrm{d}\xi\Big)^{1/2} \\ \le& C\|\varphi\|_m .\end{aligned} \tag{14.15}$$

So given $f \in \mathrm{H}^m(\mathbb{R}^d)$, choose a sequence $\varphi_j \in \mathcal{S}(\mathbb{R}^d)$ so that $\|\varphi_j - u\|_m \to 0$. Then by the above inequality, applied to $\varphi = \varphi_i - \varphi_j$, we get that

$$\sup_x |\partial_x^\alpha \varphi_i(x) - \partial_x^\alpha \varphi_j(x)| \le C\|\varphi_i - \varphi_j\|_m \to 0 \tag{14.16}$$

which shows that $\{\partial_x^\alpha \varphi_i\}$ converges uniformly for all $|\alpha| \le k$. Hence the limit $f \in \mathrm{C}^k$, and $\partial_x^\alpha \varphi_i \to \partial_x^\alpha f$ uniformly for each $|\alpha| \le k$. □

Exercise 14.2. The proof uses that $\mathcal{S}(\mathbb{R}^d)$ is dense in $\mathrm{H}^m(\mathbb{R}^d)$. Prove this statement.

Corollary 14.3. *If $u \in \mathrm{H}^m$ for all $m \in \mathbb{N}$, then $u \in \mathrm{C}^\infty$.*

Application to the wave equation. Finally let us look at the wave equation for an application of Sobolev spaces. Consider the Cauchy problem on $\mathbb{R}^{d+1}$,

$$\begin{cases} -\partial_t^2 u + \triangle u = 0 & t > 0 \\ u = f\,, \partial_t u = g\,, & t = 0\,. \end{cases} \tag{14.17}$$

We have seen that the energy is conserved:

$$E[u](t) = E[u](0)\,, \qquad E[u](t) = \frac{1}{2}\int_{\mathbb{R}^d} (\partial_t u)^2(t,x) + |\nabla u(t,x)|^2 \mathrm{d}x \tag{14.18}$$

Moreover by commuting the equation with the derivatives ∂_x^α, we see that also the *higher order energies* are conserved:

$$E_k(t) = \sum_{|\alpha|\leq k} E[\partial_x^\alpha u](t) = E_k(0) \tag{14.19}$$

This is the reason that Sobolev regularity is propagated for solutions to the wave equation.

Theorem 14.4. *Suppose $f \in \mathrm{H}^k(\mathbb{R}^d)$, and $g \in \mathrm{H}^{k-1}(\mathbb{R}^d)$ for some $k \in \mathbb{N}$, then*

$$u(t,\cdot) \in \mathrm{H}^k(\mathbb{R}^d) \quad \textit{for all } t > 0\,.$$

Proof. This statement follows directly from the Fourier representation obtained in Theorem 7.1: Since

$$\hat{u}(\xi,t) = \hat{f}(\xi)\cos(2\pi|\xi|t) + \hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|} \tag{14.20}$$

it follows from Plancherel's theorem that if $f, g \in \mathrm{L}^2(\mathbb{R}^d)$, then $u(t,\cdot) \in \mathrm{L}^2(\mathbb{R}^d)$. Moreover

$$\widehat{\partial_x^\alpha u}(\xi,t) = (2\pi i\xi)^\alpha \hat{f}(\xi)\cos(2\pi|\xi|t) + (2\pi i\xi)^\alpha \hat{g}(\xi)\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|} \tag{14.21}$$

which shows that under the stated assumptions, $\partial_x^\alpha u(t,x) \in \mathrm{L}^2(\mathbb{R}^d)$, $|\alpha| \leq k$.

□

Remark 14.5. Consider the Cauchy problem with initial data $f \in C^k$, and of compact support, and $g = 0$ for simplicity. Then, also $f \in \mathrm{H}^k(\mathbb{R}^3)$, and we can then infer from the above theorem and the Sobolev embedding that

$$u(t,\cdot) \in \mathrm{C}^{k-[d/2]}\,, \tag{14.22}$$

where $[d/2]$ is the integer part $d/2$. This fits well with our observation in Lecture 7, where we have seen in the case $d = 3$ that even though $f \in \mathrm{C}^k$, the solution may only be $u(t,\cdot) \in \mathrm{C}^{k-1}$. Similarly in higher dimensions, we can see from the spherical means formula that for data in C^k, the solution may be no better than $u(t,\cdot) \in \mathrm{C}^{k-[d/2]}$, which is the maximum discrepancy between weak derivatives in L^2 and continuous derivatives allowed by the Sobolev embedding theorem.

Problems

1. Let $u \in \mathrm{L}^2(\mathbb{R}^d)$ be weakly differentiable, and let $u_r = u * \varphi_r$, where $\varphi_r = r^{-d}\varphi(x/r)$, and φ as in the Supplement to Lecture 10. Show that mollification commutes with the operation of taking a weak derivative:
$$D^\alpha u_r(x) = (D^\alpha u)_r(x)\,. \tag{14.23}$$

2. The conclusion of the Sobolev embedding theorem fails when $s = d/2$. Consider the case $d = 2$, and let
$$f(x) = \big(\log(1/|x|)\big)^\alpha \eta(x) \tag{14.24}$$
where η is a smooth cutoff function with $\eta = 1$ for x near the origin, but $\eta(x) = 0$ for $|x| \geq 1/2$. Let $0 < \alpha < 1/2$.
   a) Verify that $\frac{\partial f}{\partial x_1}$ and $\frac{\partial f}{\partial x_2}$ are in $\mathrm{L}^2(\mathbb{R}^d)$ in the weak sense.
   b) Show that f cannot be corrected on a set of measure zero such that the resulting function is continuous at the origin.

3. Consider the linear partial differential operator
$$L = \sum_{|\alpha|\leq n} a_\alpha \Big(\frac{\partial}{\partial x}\Big)^\alpha\,. \tag{14.25}$$
We say L is **elliptic** if
$$|P(\xi)| \geq c|\xi|^n \tag{14.26}$$
for some $c > 0$, and all $\xi \in \mathbb{R}^d$ sufficiently large, where
$$P(\xi) = \sum_{|\alpha|\leq n} a_\alpha (2\pi i \xi)^\alpha \tag{14.27}$$
is the characteristic polynomial associated to L.
   a) Check that L is elliptic if and only if
   $$\sum_{|\alpha|=n} a_\alpha (2\pi\xi)^\alpha \tag{14.28}$$
   vanishes only when $\xi = 0$.
   b) If L is elliptic, prove that for some $C > 0$ the inequality
   $$\Big\|\big(\frac{\partial}{\partial x}\big)^\alpha \varphi\Big\| \leq C\Big(\|L\varphi\|_{\mathbb{R}^d} + \|\varphi\|_{\mathrm{L}^2(\mathbb{R}^d)}\Big) \tag{14.29}$$
   holds for all $\varphi \in \mathrm{C}_0^\infty(\mathbb{R}^d)$ and $|\alpha| \leq n$.
   c) Conversely, prove that if (14.29) holds, then L is elliptic.

Additional: Sobolev spaces on bounded domains

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. We can define as before $\mathrm{H}^m(\Omega)$ as the space of all functions $u \in \mathrm{L}^2(\Omega)$ whose weak derivatives $D^\alpha u$ exist in $\mathrm{L}^2(\Omega)$, $|\alpha| \leq m$.

An important subspace is the closed subspace $\mathrm{H}_0^m(\mathbb{R}^d)$ obtained as the closure of $\mathrm{C}_0^\infty(\Omega)$ functions, taken in the norm of $\mathrm{H}^m(\Omega)$. While for functions $u \in \mathrm{H}^m(\Omega)$ we have

$$\int_\Omega \varphi D^\alpha u = \int_\Omega (-1)^{|\alpha|} u \partial_x^\alpha \varphi \tag{14.1}$$

for all $\varphi \in \mathrm{C}_0^\infty(\Omega)$, this identity holds for functions $u \in \mathrm{H}_0^m(\Omega)$ even if $\varphi \in \mathrm{C}^\infty(\Omega)$ is not compactly supported in Ω. Moreover since (14.1) can be obtained for $u, \varphi \in \mathrm{C}^\infty(\mathbb{R}^d)$ by successive integration by parts, under the assumption that $D^\alpha u$ vanish on $\partial\Omega$ up to order $|\alpha| \leq m-1$, this suggests that the weak derivatives $D^\alpha u$ of a function $u \in \mathrm{H}_0^m(\Omega)$ should vanish on the boundary up to order $m-1$ in some sense.

Lemma 14.1. *Let Ω be a bounded domain, with C^1 boundary $\partial\Omega$, and $U_\sigma = \{x \in \Omega : d(x, \partial\Omega) < \sigma\}$. Then for all $u \in \mathrm{H}_0^m(\Omega)$,*

$$\|u\|_{\mathrm{H}^{m-1}(U_\sigma)} \leq C\sqrt{\sigma}\|u\|_{\mathrm{H}^m(U_{K\sigma})} \qquad (\sigma < \sigma_0) \tag{14.2}$$

where $K > 1$, and $C, \sigma_0 > 0$ are positive constants that depend only on Ω.

Remark 14.1. We have already seen a version of this Lemma in Proposition 13.3 for the case $m = 1$. Similary to how it has been applied in Lecture 13, we can see here that it implies vanishing of weak derivatives up to order $m-1$ on the boundary, in the sense that

$$\lim_{\sigma \to 0} \frac{1}{\sigma} \int_{U_\sigma} |D^\alpha u|^2 = 0 \qquad (|\alpha| \leq m-1)\,. \tag{14.3}$$

Instead of pursuing the proof of Lemma 14.1, let us make the observation that the norm on $\mathrm{H}_0^m(\Omega)$ is equivalent to the Sobolev norm at "top order", namely

$$\|u\|^2_{\mathrm{H}_0^m(\Omega)} = \int_\Omega \sum_{|\alpha|=m} |D^\alpha u|^2\,. \tag{14.4}$$

To see this, note that for any $\varphi \in \mathrm{C}_0^\infty(\Omega)$, and $y \in \Omega$ fixed

$$\begin{aligned}\int_\Omega |\varphi|^2 &= \frac{1}{d}\int_\Omega \nabla_x \cdot ((x-y)|\varphi|^2(x)) - 2\varphi(x-y)\cdot\nabla\varphi(x)\mathrm{d}x \\ &\leq \frac{2\max_{x,y\in\Omega}|y-x|}{d}\|\varphi\|_2\|\nabla\varphi\|_2\end{aligned} \tag{14.5}$$

which proves the **Poincaré inequality**

$$\|\varphi\|_{\mathrm{L}^2(\Omega)} \leq C\|\nabla\varphi\|_{\mathrm{L}^2(\Omega)} . \tag{14.6}$$

Given that smooth functions of compact support are dense in H_0^1, this also holds for $\varphi \in \mathrm{H}_0^1(\Omega)$. Applied to each weak derivative, we then obtain

$$\|u\|_{\mathrm{H}^m(\Omega)}^2 \leq C\|u\|_{\mathrm{H}_0^m(\Omega)}^2 \tag{14.7}$$

Since the converse is obvious, we have shown that these two norms are equivalent on H_0^1.

The Hilbert spaces H_0^m are also the relevant function spaces for the Sobolev embedding theorem on bounded domains:

Theorem 14.2 (Sobolev embedding). *For $m > d/2 + k$,*

$$\mathrm{H}_0^m(\Omega) \subset \mathrm{C}_0^k(\Omega) . \tag{14.8}$$

Here C_0^k denotes the space of k-times continuously differentiable functions f on $\overline{\Omega}$, with $D^\alpha f = 0$ on $\partial\Omega$ for all $|\alpha| \leq k$.

The proof is identical to the proof given in the unbounded case, once it is understood that for a function $f \in \mathrm{L}^2(\Omega)$ we extend f to $\mathbb{R}^d$ by setting $f = 0$ on $\mathbb{R}^d \setminus \Omega$, and give a characterisation as in Lemma 14.1:

Lemma 14.3. *For any bounded domain Ω with C^1 boundary $\partial\Omega$,*

$$H_0^m(\Omega) = \left\{ f \in \mathrm{L}^2(\Omega) : (1+|\xi|^2)^{m/2}\hat{f}(\xi) \in \mathrm{L}^2(\mathbb{R}^d) \right\} . \tag{14.9}$$

An important theorem for bounded domains is the following:

Theorem 14.4 (Rellich's compactness theorem). *Suppose Ω is bounded, and $m \geq 1$. Then the inclusion*

$$\mathrm{H}_0^m(\Omega) \subset \mathrm{H}_0^{m-1}(\Omega) \tag{14.10}$$

is compact. In other words, any bounded sequence $\{u_j\}$ in $\mathrm{H}_0^m(\Omega)$ has a subsequence $\{u_{j_k}\}$ which converges in $\mathrm{H}_0^{m-1}(\Omega)$.

Proof. Evidently it is enough to prove the theorem in the case $m = 1$.

Let $\{u_j\}$ be a bounded sequence in $\mathrm{H}_0^1(\Omega)$, and note that by definition we can choose $v_j \in \mathrm{C}_0^\infty(\Omega)$ such that

$$\|u_j - v_j\|_{\mathrm{H}_0^1(\Omega)} \leq 1/j . \tag{14.11}$$

We can think of $v_j \in \mathrm{C}_0^\infty(\mathbb{R}^d)$ supported in Ω, and in view of (14.7), $\|v_j\|_1$ is bounded independently of j. Denoting by $(v_j) = v_j * \varphi_r$ a smooth regularisation, we then obtain

$$|(v_j)_r(x)| \leq \left| \int_{\mathbb{R}^d} \varphi_r(x-y)v_j(y)\mathrm{d}y \right| \leq C(r) \tag{14.12}$$

where $C(r)$ depends on r, because $|\varphi_r| \leq Cr^{-d}$. Moreover,

$$|D(v_j)_r(x)| \leq \Big|\int_{\mathbb{R}^d} (D\varphi_r)(x-y)v_j(y)\mathrm{d}y\Big| \leq C(r)\,. \tag{14.13}$$

Therefore, for fixed r, the sequence $\{(v_j)_r\}_{j\in\mathbb{N}}$ is uniformly bounded, and *equicontinuous*:

$$|(v_j)_r(x) - (v_j)_r(y)| \leq C(r)|x-y|\,. \tag{14.14}$$

Therefore by the **Arzela-Ascoli theorem**, there is a subsequence $\{(v_{j_k})_r\}_{k\in\mathbb{N}}$ which is *uniformly convergent* on $\mathbb{R}^d$: There is a continuous function w, so that for any $\epsilon > 0$, there is some $K \in \mathbb{N}$ such that for all $x \in \mathbb{R}^d$,

$$k \geq K \Longrightarrow |(v_{j_k})_r(x) - w(x)| < \epsilon/6\,. \tag{14.15}$$

We also have

$$\begin{aligned}
|(v_j)_r(x) - v_j(x)| &= \Big|\int_{\mathbb{R}^d} \varphi_r(x-y)(v_j(y) - v_j(x))\mathrm{d}y\Big| \\
&= \Big|\int_{B_1(0)} \varphi(y)(v_j(x-ry) - v_j(x))\mathrm{d}y\Big| \\
&\leq \Big(\int_{B_1(0)} |(Dv_j)(x-ty)|^2 r^2 \mathrm{d}y\Big)^{1/2}
\end{aligned} \tag{14.16}$$

where we have again used that by the mean value theorem $|v_j(x-ry) - v_j(x)| \leq |Dv_j(x-ty)|r$ for some $0 < t < r$ that depends on y, and so after integrating in x,

$$\|(v_j)_r - v_j\|^2_{\mathrm{L}^2} \leq \int_{B_1(0)} \int_{\mathbb{R}^d} |(Dv_j)(x-ty)|^2 r^2 \mathrm{d}x\mathrm{d}y \leq \|Dv_j\|^2_{\mathrm{L}^2} r^2 \tag{14.17}$$

or,

$$\|(v_j)_r - v_j\|_{\mathrm{L}^2(\Omega)} \leq r\|v_j\|_{\mathrm{H}^1_0(\Omega)} \leq C_1 r\,. \tag{14.18}$$

Therefore, for any $\epsilon > 0$,

$$\begin{aligned}
\|u_{j_k} - u_{j_l}\|_{\mathrm{L}^2(\Omega)} &\leq \|u_{j_k} - v_{j_k}\|_{\mathrm{L}^2(\Omega)} + \|v_{j_k} - v_{j_l}\|_{\mathrm{L}^2(\Omega)} + \|v_{j_l} - u_{j_l}\|_{\mathrm{L}^2(\Omega)} \\
&\leq 1/j_k + \|v_{j_k} - (v_{j_k})_r\|_{\mathrm{L}^2(\Omega)} + \|(v_{j_k})_r - (v_{j_l})_r\|_{\mathrm{L}^2(\Omega)} + \|(v_{j_l})_r - v_{j_l}\|_{\mathrm{L}^2(\Omega)} + 1/j_l \\
&\leq 1/j_k + \epsilon/3 + \epsilon/6 + \epsilon/3 + 1/j_l < \epsilon
\end{aligned} \tag{14.19}$$

provided $r < \epsilon/6C_1$, and k, l sufficiently large. □

Lecture 15.

Local regularity for elliptic equations

Consider the equation

$$\sum_{i,j=1}^{d} D^j(a_{ij} D^i u) = \sum_{j=1}^{d} D^j f_j \tag{15.1}$$

where f_j are given locally square integrable functions, $f_j \in \mathrm{L}^2_{\mathrm{loc}}(\Omega)$, and a_{ij} are given locally bounded functions, $a_{ij} \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)$.

We say $u \in \mathrm{H}^1_{\mathrm{loc}}(\Omega)$ is a **weak solution** of (15.1) if

$$\int_\Omega \sum_{i,j=1}^{n} a_{ij} D^i u D^j \zeta = \int_\Omega \sum_{j=1} f_j D^j \zeta \tag{15.2}$$

for every smooth function $\zeta \in \mathrm{C}^\infty_c(\Omega)$.

Note that if $u \in \mathrm{C}^2(\Omega)$ is twice differentiable and satisfies (15.2) with $a_{ij}, f_j \in \mathrm{C}^1(\Omega)$, then it follows by integration by parts that u is a **strong solution** (or **classical** solution), namely a function that satisfies (15.1) pointwise.

In this lecture we will assume that (15.1) is (strongly) **elliptic**, in the sense that

$$\sum_{i,j=1}^{d} a_{ij}(x)\xi^i\xi^j \geq \mu|\xi|^2 \qquad (x \in \Omega, \xi \in \mathbb{R}^d) \tag{E}$$

for some fixed $\mu > 0$ independently of $x \in \Omega$, and $\xi \in \mathbb{R}^d$.

We will also impose the following explicit **boundedness** assumption:

$$|a_{ij}(x)| \leq M \qquad (x \in \Omega) \tag{B}$$

The aim of this lecture is to prove a *local regularity* result for weak solutions of (15.1) and the first step in that direction is:

Lemma 15.1. *Suppose* (E) *and* (B) *hold, and let $B_R = B_R(x_0)$ be a ball centered at x_0 so that $\overline{B_R} \subset \Omega$. If $u \in \mathrm{H}^1_{loc}(\Omega)$ is a weak solution of* (15.1)*, then for each $\theta \in (0,1)$ we have*

$$\|u\|_{1,B_{\theta R}(x_0)} \leq C\Big(\|u\|_{\mathrm{L}^2(B_R(x_0))} + \sum_{j=1}^{n} \|f_j\|_{\mathrm{L}^2(B_R(x_0))}\Big) \tag{15.3}$$

where C depends only on $R, M/\mu, d$ and θ.

The idea is to use the ellipticity conditon (E) to estimate Du locally in L^2, and then use the weak form of the equation (15.2). For this purpose let us replace ζ in (15.2) by φu, where $\varphi \in \mathrm{C}_c^\infty(\Omega)$:

$$\int_\Omega \sum_{i,j=1}^d a_{ij} D^i u(\varphi D^j u + u D^j \varphi) = \int_\Omega \sum_{j=1}^d f_j(\varphi D^j u + u D^j \varphi) \tag{15.4}$$

Exercise 15.1. This step needs some justification: Show that (15.2) holds with ζ replaced by φh, for each $\varphi \in \mathrm{C}_0^\infty(\Omega)$ and each $h \in \mathrm{H}^1_{\mathrm{loc}}(\Omega)$, by first writing out (15.2) with $\zeta = \varphi h_\sigma$, where h_σ is a mollification of h, and using that $\varphi h_\sigma \to \varphi h$ in $\mathrm{H}^1(\Omega)$.

Therefore, by (E) and (B) we obtain that

$$\mu \int_\Omega |Du|^2 \varphi \le \int_\Omega \sum_{i,j=1}^d a_{ij} D^i u D^j u \, \varphi$$

$$\le M \int_\Omega |Du||u||D\varphi| + \int_\Omega |f|\big(|\varphi||Du| + |D\varphi||u|\big) \tag{15.5}$$

Let us now choose φ to be a cutoff function so that $\varphi = 1$ on the smaller ball $B_{\theta R}$ and $\varphi = 0$ in the complement of B_R. In fact given $\psi \in \mathrm{C}_c^\infty(B_1(0))$ with $\psi = 1$ in $B_\theta(0)$, $\psi \ge 0$ everywhere and $|\partial_x \psi| \le C(\theta)$, we can set $\varphi(x) = \psi((x - x_0)/R)$, the $\varphi \in \mathrm{C}_0^\infty(\Omega)$ with the desired properties and $|\partial\varphi| \le C(\theta)/R$. Replacing φ by φ^2 in (15.5) we then obtain

$$\int_{B_R(x_0)} |Du|^2 \varphi^2 \le \frac{CM}{\mu} \int_\Omega \Big(|u| + |Du|\Big)\Big(|u| + |f|\Big)\varphi \tag{15.6}$$

with some constant C that only depends on θ and R. Finally with an application of the Cauchy inequality $ab \le \epsilon a^2/2 + b^2/(4\epsilon)$ we can absorb the Du term in the left hand side and conclude that

$$\int_{B_R(x_0)} |Du|^2 \varphi^2 \le C \int_{B_R(x_0)} \Big(|u|^2 + |f|^2\Big) \tag{15.7}$$

with some constant C that only depends on θ, R and M/μ. Since $\varphi = 1$ on $B_{\theta R}(x_0)$ this proves the Lemma.

The point is that (15.3) can be used to show that solutions to (15.1) "gain regularity", provided the coefficients a_{ij} and the functions f_j are sufficiently regular. More precisely, let us now assume that for some $k \in \mathbb{N}$, $D^\alpha a_{ij}$ exist for $|\alpha| \le k$ and are in $\mathrm{L}^\infty(\Omega)$ with

$$|D^\alpha a_{ij}(x)| \le M \quad \text{a.e.} x \in \Omega \qquad |\alpha| \le k \tag{B_k}$$

and $f_j \in \mathrm{H}^k(\Omega)$.

Given a classical solution u to (15.1) we see upon differentiating the equation that $v = v_l = \partial_l u$ satisfies

$$\sum_{i,j=1}^d D^j(a_{ij} D^i v) = \sum_{j=1}^d D^j F_j \,, \qquad F_j = F_j^l = \partial_l f_j - \sum_{i=1}^d (\partial_l a_{ij}) D^i u \tag{15.8}$$

We can then apply Lemma 15.1 to obtain

$$\begin{aligned}\|u\|^2_{2,B_{\theta^2 R}(x_0)} &= \sum_{l=1}^d \|v_l\|^2_{1,B_{\theta^2 R}(x_0)} + \|u\|^2_{0,B_{\theta^2 R}(x_0)} \\ &\leq C\sum_{l=1}^d\Big(\|v_l\|^2_{\mathrm{L}^2(B_{\theta R}(x_0))} + \sum_{j=1}^n \|F_j^l\|^2_{\mathrm{L}^2(B_{\theta R}(x_0))}\Big) + \|u\|^2_{0,B_{\theta^2 R}(x_0)} \\ &\leq C(1+M^2)\Big(\|u\|^2_{1,B_{\theta R}(x_0)} + \|f\|^2_{1,B_{\theta R}(x_0)}\Big) \\ &\leq C^2(1+M^2)\Big(\|u\|^2_{\mathrm{L}^2(B_R(x_0))} + \|f\|^2_{\mathrm{H}^1(B_R(x_0))}\Big)\,. \qquad (15.9)\end{aligned}$$

This is however not the proof of the following theorem, because our starting point here is not a strong solution to (15.1) but merely a weak solution $u \in \mathrm{H}^1_{\mathrm{loc}}(\Omega)$ satisfying (15.2).

Theorem 15.2. *If $u \in \mathrm{H}^1_{loc}(\Omega)$ is a weak solution of* (15.1)*, and if* (E)*,* (B) *and* (B_k) *hold for some $k \in \mathbb{N}$, then $u \in \mathrm{H}^{1+k}$ and*

$$\|u\|_{1+k,B_{\theta R}(x_0)} \leq C\Big(\|u\|_{0,B_R(x_0)} + \|f\|_{k,B_R(x_0)}\Big) \qquad (15.10)$$

for any ball $\overline{B_R}(x_0) \subset \Omega$ and any $\theta \in (0,1)$, where C is a constant that only depends on d, k, θ, R, and M, μ.

The idea of differentiating (15.1) is replaced by working instead with difference quotients: For any function u defined on Ω, we define

$$\Delta_h^{(j)} u(x) = \frac{u(x+he_j) - u(x)}{h} \qquad e_j = (0,\ldots,\overset{j^{\text{th}}}{1},\ldots,0) \qquad (15.11)$$

which is defined on the smaller domain

$$\Omega_{|h|} = \{x \in \Omega : d(x,\partial\Omega) > |h|\} \qquad (15.12)$$

From the definition it is clear that

$$\Delta_h^{(j)}(f+g) = \Delta_h^{(j)} f + \Delta_h^{(j)} g \qquad (15.13)$$

$$\Delta_h^{(j)}(fg)(x) = g(x)\Delta_h^{(j)} f(x) + f(x+he_j)\Delta_h^{(j)} g \qquad (15.14)$$

$$D^\alpha \Delta_h^{(j)} f = \Delta_h^{(j)} D^\alpha f\,, \qquad (15.15)$$

Exercise 15.2. Prove the "integration by parts" formula

$$\int_\Omega f\Delta_h^{(j)} g = -\int_\Omega g\Delta_{-h}^{(j)} f \qquad (15.16)$$

for any $f, g \in \mathrm{L}^1(\Omega)$ so that fg is compactly supported in $\Omega_{|h|}$.

We will prove Theorem 15.2 only in the case $k = 1$. The general statement can be obtained by induction on k.

Proof of Theorem 15.2 for $k = 1$. Take $h \neq 0$ so that $\overline{B_R}(x_0) \subset \Omega_{|h|}$, and write down (15.2) with ζ replaced by $\Delta_{-h}^{(j)}\zeta \in \mathrm{C}_c^\infty(\Omega)$ where $\zeta \in \mathrm{C}_c^\infty(\Omega_{|h|})$. Then using the properties of the difference quotient above, we obtain

$$\int_\Omega \sum_{i,j=1}^d a_{ij}(x+he_j)(D^i\Delta_h^{(l)}u)(x)D^j\zeta(x)\mathrm{d}x = \int_\Omega \sum_{j=1} F_j^l D^j\zeta \tag{15.17}$$

where

$$F_j^l = \Delta_h^{(l)} f_j - \sum_{i=1}^d (\Delta_h^{(l)} a_{ij}) D^i u\,. \tag{15.18}$$

Now we can apply apply Lemma 15.1 to $u_h = \Delta_h^{(l)} u$ to get

$$\|u_h\|_{1,B_{\theta R}(x_0)} \leq C\Big(\|u_h\|_{0,B_R(x_0)} + \sum_{j=1}^d \|F_j^l\|_{0,B_R(x_0)}\Big)\,. \tag{15.19}$$

We now need to use that

$$\|\Delta_h^{(l)} f\|_{\mathrm{L}^2(B_R(x_0)} \leq \|f\|_{1,B_{R+|h|}(x_0)} \tag{15.20}$$

Exercise 15.3. More generally, if $f \in \mathrm{H}^k(\Omega)$, then

$$\|\Delta_h^{(l)} f\|_{k-1,\Omega_{|h|}} \leq \|f\|_{k,\Omega}\,. \tag{15.21}$$

We will not prove these facts here in detail.

Similarly, we will now use that by that by (B_k), with $k = 1$,

$$|\Delta_h^{(l)} a_{ij}| \leq M\,. \tag{15.22}$$

Then we obtain that

$$\|u_h\|_{1,B_{\theta R}(x_0)} \leq C(1+M)\Big(\|u\|_{1,B_{R+|h|}(x_0)} + \sum_{j=1}^d \|f_j\|_{1,B_{R+|h|}(x_0)}\Big)\,. \tag{15.23}$$

for some constant C independent of h.

To conlude the proof we would like to take the limit $h \to 0$.

Lemma 15.3. *Let $v \in \mathrm{L}^2_{loc}(\Omega)$ and suppose*

$$\limsup_{h\to 0} \|\Delta_h^{(j)} v\|_{\mathrm{L}^2(K)} \leq C < \infty\,, \tag{15.24}$$

for each compact $K \subset \Omega$. Then v has a weak derivative $D^j v \in \mathrm{L}^2(\Omega)$, and $\Delta_h^{(j)} v \to D^j v$ weakly.

Proof. Let $K \subset \Omega$ be a compact set, and h so that $K \subset \Omega_{|h|} \subset \Omega$. Then by assumption $v_n = \Delta^{(j)}_{1/n} v$ is a bounded sequence in a Hilbert space $L^2(K)$, $n > |h|^{-1}$, there exists a weakly convergent subsequence $v_{n_k} \to g_j$ so that in particular

$$\int_K v_{n_k}\varphi \to \int_K g_j \varphi \qquad (k \to \infty) \tag{15.25}$$

for every $\varphi \in C^\infty_c(\Omega)$. The function $g_j \in L^2(K)$ is the weak derivative of v, because in view of (15.16),

$$\int_K v_{n_k}\varphi = \int_K v \Delta^{(j)}_{-1/n_k}\varphi \to -\int_K v\partial^j_x \varphi \qquad (k \to \infty)\,. \tag{15.26}$$

□

We know from (15.23) that $\|D^\alpha u_h\|_{0,B_{\theta R}(x_0)} = \|\Delta^{(j)}_h D^\alpha u\|_{L^2(B_{\theta R}(x_0))}$ is bounded independently of h, for $|\alpha| \leq 1$, and thus the Lemma shows that $D^\beta u \in L^2(\Omega)$ exists for $|\beta| \leq 2$. Moreover $D^i\Delta^{(l)}_h u \to D^i D^l u$ converges weakly in $L^2(\Omega)$ as $h \to 0$, and similarly $\Delta^{(l)}_h f_j \to D^l f_j$ and $\Delta^{(l)}_h a_{ij} \to D^l a_{ij}$ converge weakly in $L^2(\Omega)$, and thus we can pass to the limit $h \to 0$ in (15.17) and obtain that $D^l v$ precisely satisfies (15.8) weakly, namely

$$\int_\Omega \sum_{i,j=1}^d a_{ij}(x)(D^i D^l u)(x) D^j\zeta(x)\mathrm{d}x = \int_\Omega \sum_{j=1}^d F_j(x) D^j \zeta(x)\mathrm{d}x\,, \tag{15.27}$$

$$F_j = F^l_j = D^l f_j - \sum_{i=1}^d (D^l a_{ij}) D^i u \tag{15.28}$$

As in (15.9) the statement then follows from Lemma 15.1:

$$\|u\|_{2,B_{\theta R}(x_0)} = \|Du\|_{1,B_{\theta R}(x_0)} \leq C(1+M)\Big(\|Du\|_{0,B_R(x_0)} + \|Df\|_{0,B_R(x_0)}\Big) \tag{15.29}$$

which can be applied again to $\|Du\|_{0,B_R(x_0)} = \|u\|_{1,B_R(x_0)}$ after replacing θ by θ^2 above. □

It is now also clear how to iterate the proof: Instead of (15.2) we can use (15.27) as a starting point for the procedure of introducing difference quotients, until after k iterations we obtain the control of $\|u\|_{1+k,B_{\theta R}(x_0)}$ stated, keeping in mind that $\theta \in (0,1)$ is arbitrary.

Together with the Sobolev embedding theorem this now implies a strong regularity result for elliptic equations:

Corollary 15.4. *Suppose $u \in H^1_{loc}(\Omega)$ is a weak solution of* (15.1), *and $l \in \mathbb{N}$, and* (E), (B) *and* (B_k) *all hold for $k > d/2 + l - 1$. Then $u \in C^l(\Omega)$ is l-times continuously differentiable on Ω, and all derivatives up to order l are bounded pointwise. In particular, if a_{ij} and f_j are smooth functions on Ω, then $u \in C^\infty(\Omega)$.*

Problems

1. Suppose $F(x,p)$ is a smooth function of the variables $x \in \Omega$ and $p \in \mathbb{R}^d$. Consider the action functional
$$\mathcal{A}[u] = \int_\Omega F(x, Du(x))\mathrm{d}x\,. \tag{15.30}$$
Derive the Euler-Lagrange equations from a variation through solutions u_t and formulate an ellipticity condition in terms of F.

2. Consider the equation
$$\Delta u = \sum_{j=1}^{d} b_j D^j u + cu + f \tag{15.31}$$
where $b_j, c, f \in \mathrm{C}^\infty(\Omega)$. Formulate the notion of a weak solution $u \in \mathrm{L}^2_{\mathrm{loc}}(\Omega)$ and prove that $u \in \mathrm{C}^\infty(\Omega)$. Note that this would follow from Corollary 15.4 if it was already known that $u \in \mathrm{H}^1_{\mathrm{loc}}(\Omega)$.

3. Suppose that in place of the ellipticity condition (E) we have that the condition that there are constants c_1, c_2 such that
$$\int_\Omega |D\varphi|^2 \le c_1 \int_\Omega \sum_{i,j=1} a_{ij} D^i\varphi D^j\varphi + c_2\|\varphi\|_{\mathrm{L}^2(\Omega)} \tag{C}$$
for every $\varphi \in \mathrm{C}_0^\infty(\Omega)$. Show that the proofs of Lemma 15.1 and Theorem 15.2 can be modified in such a way that it suffices to know (C) in place of (E).

4. Show that (C) is equivalent to (E).

Lecture 16.

Existence of solutions to Dirichlet's problem for elliptic operators in Sobolev spaces

In this lecture we will study the solvability of the Dirichlet problem in its weak formulation for operators of the form

$$Lu = -\sum_{i,j=1}^{d} D^j(a_{ij} D^i u) \tag{16.1}$$

which satisfy the ellipticity condition (E) and the boundedness assumption (B).

Let Ω be a bounded domain in $\mathbb{R}^d$. The Dirichlet problem in this setting is the problem of determining u in Ω such that

$$Lu = f \text{ on } \Omega\,, \qquad u = 0 \text{ on } \partial\Omega\,, \tag{D}$$

where $f = \sum_{i=1}^{d} D^i f_i$ is given. The aim is to develop an existence theory for the following weak formulation: We say that u is a weak solution of the Dirichlet problem (D) provided

$$u \in \mathrm{H}^1_0(\Omega)\,, \qquad \int_\Omega \sum_{i,j=1}^{d} a_{ij} D^i u D^j \varphi = -\int_\Omega \sum_{i=1}^{d} f_i D^i \varphi \qquad \text{for all } \varphi \in \mathrm{H}^1_0(\Omega) \tag{P}$$

Remark 16.1. If $u \in \mathrm{C}^0(\overline{\Omega}) \cap \mathrm{H}^1(\Omega)$ then (P) is equivalent to finding u which satisfies the equation $Lu = f$ in the weak sense of Lecture 15 with $u = 0$ on $\partial\Omega$ in the classical sense.

A family of problems related to (D) is the following more general problem where $\lambda \in \mathbb{R}$ is a free parameter:

$$Lu = \lambda u + f \text{ on } \Omega\,, \qquad u = 0 \text{ on } \partial\Omega\,. \tag{D_λ}$$

Set

$$A_\lambda(u,\varphi) = \int_\Omega \sum_{i,j=1}^{d} a_{ij} D^i u D^j \varphi - \lambda u \varphi \,\mathrm{d}x \tag{16.2}$$

$$F(\varphi) = -\int_\Omega \sum_{i=1}^{d} f_i D^i \varphi \,\mathrm{d}x \tag{16.3}$$

$$\tag{16.4}$$

then the weak formulation of the modified problem (D_λ) is to find $u \in \mathrm{H}^1_0(\Omega)$ so that

$$A_\lambda(u,\varphi) = F(\varphi) \text{ for all } \varphi \in \mathrm{H}^1_0(\Omega)\,. \tag{P_λ}$$

Lemma 16.1 (Lax-Milgram)**.** *Let $\mathcal{H}$ be a real Hilbert space, and let A be a* bounded *and* strictly coercive *bilinear form, namely $A : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$, and $A(u, v)$ is linear in both u, and v with the properties that for some fixed positive constants $\lambda, \Lambda > 0$,*

$$|A(u, v)| \leq \Lambda \|u\| \|v\| \qquad A(u, u) \geq \lambda \|u\|^2 . \tag{16.5}$$

Then there exists an isomorphism T of $\mathcal{H}$ onto $\mathcal{H}$ so that

$$A(u, v) = (Tu, v)\,, \quad \lambda\|u\| \leq \|Tu\| \leq \Lambda\|u\|\,, \qquad u, v \in \mathcal{H}\,. \tag{16.6}$$

Proof. For fixed $u \in \mathcal{H}$, $l(v) = A(u, v)$ is a bounded linear functional, hence by the Riesz representation theorem 9.1 there exists $g \in \mathcal{H}$, so that $l(v) = (v, g)$. Clearly g depends linearly on u, and we have $g = Tu$ for some linear operator T, and $A(u, v) = (v, Tu) = (Tu, v)$ with $\lambda\|u\|^2 \leq |A(u, u)| = |(Tu, u)| \leq \|Tu\|\|u\|$ which shows that

$$\lambda\|u\| \leq \|Tu\| \tag{16.7}$$

In particular, T is injective. Since $\|Tu\|^2 = A(u, Tu) \leq \Lambda\|u\|\|Tu\|$ we see that T is bounded. It remains to show that T is also surjective.

Exercise 16.1. Show that as a consequence of (16.7) the bounded operator T has *closed range*, namely the linear subspace $\mathcal{S} = \{Tu : u \in \mathcal{H}\}$ is closed.

If the range $\mathcal{S}$ were not all of $\mathcal{H}$, then $\mathcal{S}^\perp$ is not empty, cf. Prop. 9.2, and for $v \neq 0$ orthogonal to $\mathcal{S}$, we would have

$$A(v, v) = (Tv, v) = 0 \tag{16.8}$$

which contradicts the assumption. □

The motivation for introducing the problems (P_λ) is that *for some* λ the bilinear form A_λ can be seen to satisfy the assumptions of the Lax-Milgram Lemma: Let $\mathcal{H} = \mathrm{H}^1_0(\Omega)$, and A_λ be the bilinear form on $\mathcal{H}$ defined by (16.2). Then by (E),

$$A_\lambda(u, u) \geq \int_\Omega \mu|Du|^2 - \lambda|u|^2 \geq \mu\|u\|^2_{1,\Omega} \tag{16.9}$$

provided $-\mu - \lambda \geq 0$. Moreover by (B),

$$|A_\lambda(u, v)| \leq \int_\Omega M|Du||Dv| + \mu|u||v| \leq (M + \mu)\|u\|_{1,\Omega}\|v\|_{1,\Omega}\,. \tag{16.10}$$

In summary, if $\lambda_0 < 0$ with $|\lambda_0| \geq \mu$, then A_{λ_0} satisfies the assumptions of Lemma 16.1 with $\mathcal{H} = \mathrm{H}^1_0(\Omega)$. Therefore there exists and isomorphism T from $\mathcal{H}$ onto $\mathcal{H}$, so that

$$A_{\lambda_0}(u, v) = (Tu, v)\,. \tag{16.11}$$

The problem we want to solve is (P_λ). Let us first assume that $f_i \in \mathrm{H}^1(\Omega)$, then $F(v) = (f, v)_{\mathrm{L}^2(\Omega)}$ is a bounded linear functional on $\mathcal{H}$,

$$|F(v)| \leq \|f\|_{\mathrm{L}^2(\Omega)}\|v\|_{\mathrm{L}^2(\Omega)} \leq \|f\|_{\mathrm{L}^2(\Omega)}\|v\|_{1,\Omega}\,. \tag{16.12}$$

Hence once more by the Riesz representation theorem 9.1 there exists $w \in \mathcal{H}$ so that $F(v) = (v, w)_{\mathcal{H}}$. Again w depends linearly on f, and we can write $w = S_0(f)$ for some bounded, linear transformation $S_0 : \mathrm{L}^2(\Omega) \to \mathcal{H}$. Moreover S_0 is injective.

Defining $S = T^{-1} \circ S_0$ we have

$$A_{\lambda_0}(S(f), v) = (TS(f), v) = (S_0(f), v) = (w, v) = (f, v)_{\mathrm{L}^2(\Omega)}\,, \tag{16.13}$$

and thus proven the following existence result.

Lemma 16.2. *Suppose* (E) *and* (B) *hold and* $\lambda_0 \in \mathbb{R}$ *is chosen as above. Then there exists a bounded linear injective operator* $S : \mathrm{L}^2(\Omega) \to \mathcal{H}$ *which is a* solution operator *for the problem* (P_{λ_0}) *in the sense that*

$$A_{\lambda_0}(S(f), \varphi) = (f, \varphi)_{\mathrm{L}^2(\Omega)} \qquad \text{for all } \varphi \in \mathcal{H}, f \in \mathrm{L}^2(\Omega)\,. \tag{16.14}$$

It remains to solve the original problem (P_λ). First note that by (16.14),

$$A_\lambda(u, \varphi) = A_{\lambda_0}(u, \varphi) - (\lambda - \lambda_0)(u, \varphi)_{\mathrm{L}^2(\Omega)} = A_{\lambda_0}\Big(u - (\lambda - \lambda_0)S(u), \varphi\Big) \tag{16.15}$$

and thus $u \in \mathcal{H}$ so that $A_\lambda(u, \varphi) = (f, \varphi)$ for all $\varphi \in \mathcal{H}$ if and only if

$$u - (\lambda - \lambda_0)S(u) = S(f)\,. \tag{16.16}$$

Now if $u = S(w)$ for some $w \in \mathrm{L}^2(\Omega)$, then by the injectivity of S we have

$$w - (\lambda - \lambda_0)\iota \circ S(w) = f\,. \tag{16.17}$$

where $\iota : \mathcal{H} \to \mathrm{L}^2(\Omega)$ is the inclusion map. Conversely, if w solves (16.17) then $u = S(w)$ satisfies (16.16). In conlusion, $u \in \mathcal{H}$ solves (P_λ) if and only if

$$\Big(I - (\lambda - \lambda_0)\iota \circ S\Big)w = f\,, \qquad u = S(w)\,. \tag{16.18}$$

Lemma 16.3. *The map* $\iota \circ S : \mathrm{L}^2(\Omega) \to \mathrm{L}^2(\Omega)$ *is compact.*

Proof. The inclusion map $\iota : \mathcal{H} \to \mathrm{L}^2(\Omega)$ is compact by Rellich's theorem, and S is a bounded map, hence also $\iota \circ S$ is compact. □

The compactness property has significant bearing on the solvability of (16.18), for which we have to understand the null space and range of $I - (\lambda - \lambda_0)\iota \circ S$.

Lemma 16.4. *Suppose* T *is a compact operator on a Hilbert space* $\mathcal{H}$*, and* $\lambda \neq 0$*. Then the dimension of the kernel of* $T - \lambda I$ *is finite. Moreover, the eigenvalues of* T*, namely the set of* $\lambda \in \mathbb{C}$ *for which* $\ker(T - \lambda I) \neq 0$*, form at most a denumerable set* $\lambda_1, \ldots, \lambda_k, \ldots$*, with* $\lambda_k \to 0$ *as* $k \to \infty$*.*

For the operators (16.1) considered in this lecture, the transformation $\iota \circ S$ is in fact symmetric:

$$(\iota \circ S)^* = \iota \circ S\,. \tag{16.19}$$

Exercise 16.2. Show that more generally

$$(\iota \circ S)^* = \iota \circ S_1 \,, \tag{16.20}$$

where S_1 is the solution operator $S_1 : \mathrm{L}^2(\Omega) \to \mathcal{H}$ for adjoint problem

$$A_{\lambda_0}(\zeta, S_1(f)) = (f, \zeta) \qquad \zeta \in \mathcal{H} \,. \tag{16.21}$$

As a result the eigenvalues of $\iota \circ S$ are *real*. We also know that (16.18) is solvable for $\lambda < \lambda_0$. Hence we infer from Lemma 16.4 that there exists a discrete set $\Lambda \subset (\lambda_0, \infty)$ so that for $\lambda \notin \Lambda$, $I - (\lambda - \lambda_0)\iota \circ S$ is injective. Moreover, for those $\lambda \notin \Lambda$ its range is all of $\mathrm{L}^2(\Omega)$:

Lemma 16.5 (Fredholm alternative)**.** *Suppose T is a compact operator on a Hilbert space, and $\lambda \neq 0$. Then $\lambda I - T$ is injective if and only if $\lambda I - T$ is surjective.*

There are thus the following possibilities:

(i) For $\lambda \notin \Lambda$, $I - (\lambda - \lambda_0)\iota \circ S$ is an isomorphism of $\mathrm{L}^2(\Omega)$ onto itself. In this case the problem (P_λ) has a unique solution $u \in \mathrm{H}^1_0(\Omega)$ for any $f \in \mathrm{L}^2(\Omega)$.

(ii) For $\lambda \in \Lambda$, the null space of $I - (\lambda - \lambda_0)\iota \circ S$ is *finite dimensional*. In other words, the problem (P_λ) with $f = 0$ has a set of solutions u_i which span a finite dimensional subspace N_λ of $\mathrm{H}^1_0(\Omega)$.

Bibliography

Evans, Lawrence C. *Partial differential equations.* Second. Vol. 19. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2010, pp. xxii+749. ISBN: 978-0-8218-4974-3. DOI: 10.1090/gsm/019. URL: https://doi.org/10.1090/gsm/019.

Folland, Gerald B. "Fundamental solutions for the wave operator". In: *Expositiones Mathematicae* 15.1 (1997), pp. 25–52.

— *Introduction to Partial Differential Equations.* Ed. by Elias M. Stein Wu-chung Hsiang John Milnor. Mathematical Notes. Princeton University Press, 1976.

John, Fritz. *Partial differential equations.* Third edition. Applied Mathematical Sciences 1. Springer, 1978.

Luk, Jonathan. *Introduction to nonlinear wave equations.* https://web.stanford.edu/ jluk/NWnotes.pdf. available online. 2016.

Stein, Elias M. and Rami Shakarchi. *Fourier analysis.* Vol. 1. Princeton Lectures in Analysis. An introduction. Princeton University Press, Princeton, NJ, 2003, pp. xvi+311. ISBN: 0-691-11384-X.

— *Functional analysis.* Vol. 4. Princeton Lectures in Analysis. Introduction to further topics in analysis. Princeton University Press, Princeton, NJ, 2011, pp. xviii+423. ISBN: 978-0-691-11387-6.

— *Real analysis.* Vol. 3. Princeton Lectures in Analysis. Measure theory, integration, and Hilbert spaces. Princeton University Press, Princeton, NJ, 2005, pp. xx+402. ISBN: 0-691-11386-6.