# Introduction to Partial Differential Equations

Jason Murphy

May 13, 2017

# Contents

# Chapter 1

# Introduction

These notes are based on an undergraduate introductory PDE course I gave at UC Berkeley in the Fall of 2014. While searching for a textbook, I could not find exactly what I was looking for. Between the textbooks of Strauss and Evans, I could find essentially everything I wanted to cover; however, I found that I wanted something a bit more in-depth than the book of Strauss and a bit more basic than the book of Evans. Inspired largely by these two textbooks, I prepared my own notes, which formed the basis of this document.

The prerequisites for this class consisted of multivariable calculus, linear algebra, and differential equations. In particular, I could not assume familiarity with real analysis. Thus, many issues related to convergence are dealt with rather informally, and many technical details are glossed over.

These notes are a work in progress and will be updated over time.

## 1.1 Derivation of some common PDE

*Example* 1.1.1 (Heat equation). Let $\Omega \subset \mathbb{R}^d$ be an open region of space. Let $u(t,x)$ denote the density of some quantity in $\Omega$ (e.g. 'heat') at time $t$ and position $x$. The goal is to describe how the distribution of heat evolves in time.

Denote by $\mathbf{F}(t,x)$ the 'flux density' of the heat at $(t,x)$. Heat flows from regions of high density to low density, so we approximate $\mathbf{F} = -k\nabla u$ for some $k > 0$.

Now fix a small subregion $V \subset \Omega$. The amount of heat in $V$ is given by

$$\int_V u(t,x)\,dx.$$

We wish to measure

$$\tfrac{d}{dt}\int_V u(t,x)\,dx = \int_V \tfrac{\partial u}{\partial t}(t,x)\,dx.$$

To do so, we need to measure the flow of heat through the boundary $\partial V$.

Let $\mathbf{n}(x)$ be the outward-pointing unit normal vector at $x \in \partial V$. Then

$$\int_V \tfrac{\partial u}{\partial t}\,dx = -\int_{\partial V} \mathbf{F}\cdot\mathbf{n}\,dS = \int_{\partial V} k\nabla u\cdot\mathbf{n}\,dS.$$

Applying the divergence theorem, we find

$$\int_V \tfrac{\partial u}{\partial t}\,dx = \int_V \operatorname{div} k\nabla u\,dx = \int_V k\Delta u\,dx. \quad (*)$$

For any $(t,x)$, we now consider regions $V \subset \Omega$ that 'shrink down' to $(t,x)$. As $(*)$ holds on each $V$, we conclude

$$\tfrac{\partial u}{\partial t}(t,x) = k\Delta u(t,x).$$

This is the *heat equation*.

*Example* 1.1.2 (Laplace equation). We revisit the setting of Example 1, but suppose the system has reached equilibrium. In particular, $u = u(x)$. (The density is constant in time).

Consider a subregion $V \subset \Omega$. As the system is at equilibrium,

$$\int_V u\,dx$$

is constant. Thus the flow through $\partial V$ is zero:

$$\int_{\partial V} k\nabla u\cdot\mathbf{n}\,dS = 0.$$

Applying the divergence theorem, we find

$$\int_V \Delta u\,dx = 0.$$

As $V$ was arbitrary, we deduce

$$\Delta u(x) = 0.$$

This is the *Laplace equation*.

**Definition 1.1.1.** Solutions to Laplace's equation are called **harmonic**.

*Example* 1.1.3 (Wave equation). Let $\Omega \subset \mathbb{R}^d$ be an open region of space inhabited by an elastic object (for example, a string if $d = 1$, a drumhead if $d = 2$). Let $u(t, x)$ denote the displacement of the object in some fixed direction. The goal is to derive an equation for $u$. We will primarily use '$F = ma$'.

Let $V \subset \Omega$ be a small subregion. Force (tension) acts on the object through the boundary $\partial V$. For elastic bodies, a reasonable approximation is $\mathbf{F} = -k\nabla u$ for some $k > 0$.

The net force on the object in the region $V$ is thus

$$-\int_{\partial V} \mathbf{F} \cdot \mathbf{n}\, dS = \int_{\partial V} k\nabla u \cdot \mathbf{n}\, dS.$$

The acceleration of the object within $V$ is given by

$$\tfrac{d^2}{\partial t^2} \int_V u\, dx = \int_V \tfrac{\partial^2 u}{\partial t^2}\, dx.$$

Assuming constant mass density $\rho$, Newton's law ($F = ma$) gives

$$\int_{\partial V} k\nabla u \cdot \mathbf{n}\, dS = \rho \int_V \tfrac{\partial^2 u}{\partial t^2}\, dx.$$

Applying the divergence theorem we find

$$\int_V (\tfrac{\partial^2 u}{\partial t^2} - c\Delta u)\, dx = 0, \quad c = \tfrac{k}{\rho}.$$

As $V$ was arbitrary, we deduce

$$\tfrac{\partial^2 u}{\partial t^2} - c\Delta u = 0.$$

This is the *wave equation*.

# Chapter 2

# Review of Topology and Calculus

## 2.1 Properties of $\mathbb{R}^d$

We work in Euclidean space $\mathbb{R}^d$, which is a vector space with the standard basis $\{e^1, \ldots, e^d\}$, where $e^1 = (1, 0, \ldots, 0)$, etc. For $x \in \mathbb{R}^d$, we write $x = (x_1, \ldots, x_d)$ to denote $x = x_1 e^1 + \cdots + x_d e^d$. We call $x_1, \ldots, x_d$ the **components** of $x$.

The standard inner product on $\mathbb{R}^d$ is given by $x \cdot y = x_1 y_1 + \cdots + x_d y_d$. The norm of a vector is defined by $|x| = \sqrt{x \cdot x}$. The distance between two points is given by $d(x, y) = |x - y|$.

We use the metric on $\mathbb{R}^d$ to define its topology, that is, the notion of open/closed sets and convergence/limits.

For $x \in \mathbb{R}^d$ and $r > 0$, we define the **open ball**

$$B_r(x) = \{y \in \mathbb{R}^d : d(x, y) < r\}.$$

A set $\Omega \subset \mathbb{R}^d$ is **open** if for any $x \in \Omega$, there exists $r > 0$ such that $B_r(x) \subset \Omega$. A set $C \subset \mathbb{R}^d$ is **closed** if its complement $\mathbb{R}^d \backslash C := \{x \in \mathbb{R}^d : x \notin C\}$ is open.

The **closure** of $\Omega$, denoted $\overline{\Omega}$, is the intersection of all closed sets containing $\Omega$

A sequence $\{x_n\}_{n=1}^{\infty} \subset \mathbb{R}^d$ **converges** to $\ell \in \mathbb{R}^d$ if

$$\text{for any } \varepsilon > 0 \text{ there exists } N \text{ such that } |x_n - \ell| < \varepsilon \text{ for all } n \geq N.$$

We call $\ell$ the **limit** of the sequence and write $x_n \to \ell$

A set $C$ is closed if whenever $\{x_n\}_{n=1}^{\infty} \subset C$ and $x_n \to \ell$, we have $\ell \in C$.

A set $K \subset \mathbb{R}^d$ is **compact** if it is closed and bounded. (This is not a definition of compact, but rather a theorem about compact sets in $\mathbb{R}^d$.)

## 2.2 Calculus Review

Fix an open set $\Omega \subset \mathbb{R}^d$ and a function $f : \Omega \to \mathbb{R}$.

A function $f$ is **continuous at** $x \in \Omega$ if whenever $\{x_n\}_{n=1}^{\infty} \subset \Omega$ and $x_n \to x$, we have $f(x_n) \to f(x)$. The function $f$ is **continuous** if it is continuous at each $x \in \Omega$

A function $f$ is **differentiable at** $x \in \Omega$ if there exists $v \in \mathbb{R}^d$ such that for any $\{h_n\}_{n=1}^{\infty} \subset \mathbb{R}^d \backslash \{0\}$ such that $h_n \to 0$, we have

$$\frac{f(x+h_n)-f(x)-v\cdot h_n}{|h_n|} \to 0.$$

We call $v$ the **gradient of** $f$ **at** $x$, and we write $v = \nabla f(x)$.

The components of $\nabla f(x)$ are the **partial derivatives of** $f$, denoted $\frac{\partial f}{\partial x_i}(x)$. That is, $\nabla f(x) = \big(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_d}(x)\big)$.

The function $f$ is **differentiable** if it is differentiable at each $x \in \Omega$.

Note that $\frac{\partial f}{\partial x_i}$ defines a function from $\Omega \to \mathbb{R}$, which may also be differentiable. Thus we have second order derivatives $\frac{\partial^2 f}{\partial x_j \partial x_i}$, as well as higher order derivatives. Clairaut's theorem states that 'mixed' partial derivatives (e.g. $\frac{\partial^2 f}{\partial x_i \partial x_j}$ versus $\frac{\partial^2 f}{\partial x_j \partial x_i}$) are equal if they are continuous.

We say $f$ is **smooth** if it has continuous derivatives of every order — we write $f \in C^{\infty}$.

The **support** of $f$ is defined as follows: $\operatorname{supp}(f) := \overline{\{x : f(x) \neq 0\}}$. If $f$ is smooth and $\operatorname{supp}(f)$ is compact, we write $f \in C_c^{\infty}$.

A **vector field** is a function $\mathbf{F} : \mathbb{R}^d \to \mathbb{R}^d$. We write $\mathbf{F}$ in terms of its component functions: $\mathbf{F} = (F_1, \ldots, F_d)$.

We use the following notation for a function $u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, with $(t, x) \in \mathbb{R} \times \mathbb{R}^d$:

$$\begin{aligned}
&u_t = \partial_t u = \tfrac{\partial u}{\partial t},\\
&u_{x_i} = \partial_{x_i} u = \partial_i u = \tfrac{\partial u}{\partial x_i},\\
&\tfrac{\partial^2 u}{\partial x_i \partial x_j} = u_{x_i x_j} = \partial_{x_i x_j} u = \partial_{ij} u, \quad \text{etc.},\\
&\nabla u = (u_{x_1}, \ldots, u_{x_d}) = \operatorname{grad} u.
\end{aligned}$$

The **divergence** of $\mathbf{F} : \mathbb{R}^d \to \mathbb{R}^d$ is given by $\operatorname{div} \mathbf{F} = \nabla \cdot \mathbf{F} = \sum_{i=1}^d \frac{\partial F_i}{\partial x_i}$.

The **Laplacian** of $u$ is given by $\Delta u = \operatorname{div} \nabla u = \sum_{i=1}^d u_{x_i x_i}$.

## 2.3 Further Topics

**Theorem 2.3.1** (Gauss–Green formula)**.** *Let $\Omega \subset \mathbb{R}^d$ be an open region with smooth boundary $\partial\Omega$. Let $f : \overline{\Omega} \to \mathbb{R}$ be a smooth function. Let $\boldsymbol{n}(x) = (n_1(x), \ldots, n_d(x))$ denote the outward-pointing unit normal vector at $x \in \partial\Omega$. Then for $i \in \{1, \ldots, d\}$,*

$$\int_\Omega u_{x_i}\, dx = \int_{\partial\Omega} u n_i\, dS.$$

- **Change of variables:**

$$\underbrace{\int_{\mathbb{R}^d} f(x + x_0)\, dx = \int_{\mathbb{R}^d} f(y)\, dy,}_{y=x+x_0, \quad dy=dx} \qquad \underbrace{\int_{\mathbb{R}^d} f(cx)\, dx = |c|^{-d} \int_{\mathbb{R}^d} f(y)\, dy\,.}_{y=cx, \quad dy=|c|^d dx}$$

- **Polar coordinates:**

$$\underbrace{\int_{\mathbb{R}^d} f(x)\, dx = \int_0^\infty \int_{\partial B_1(0)} f(r\theta)\, r^{d-1}\, dS\, dr\,.}_{x=r\theta, \quad dx=r^{d-1}\, dr\, dS(\theta)}$$

- An **operator** is a function whose input/output are functions.

  Important examples include **differential operators**, for example

  - $\partial_{x_i}$ (partial derivative)
  - $\Delta$ (Laplacian)
  - $\partial_t - \Delta$ (heat operator)
  - $\partial_t^2 - \Delta$ (d'Alembertian)

## 2.4 Convolution and Distributions

**Definition 2.4.1.** Let $f, g \in C_c^\infty(\mathbb{R}^d)$. The **convolution** of $f$ and $g$ is the function

$$f * g : \mathbb{R}^d \to \mathbb{R}$$

defined by

$$f * g(x) = \int_{\mathbb{R}^d} f(x - y) g(y)\, dy.$$

Note: this makes sense even if $g$ is merely continuous, and even in much more general settings.

**Properties.**

- $f * g = g * f$
- $\partial_i(f * g) = (\partial_i f) * g = f * (\partial_i g)$ (similarly for differential operators)
- for $f, g \in C_c^\infty$, we have $f * g \in C_c^\infty$

**Approximate identities.**

- Let $\phi \in C_c^\infty(\mathbb{R}^d)$ satisfy $\int \phi \, dx = 1$, with $\text{supp}(\phi) \subset B_1(0)$.
- Define $\{\phi_n\}_{n=1}^\infty$ via $\phi_n(x) = n^d \phi(nx)$. (draw picture)
- Note
$$\int \phi_n(x)\, dx = \underbrace{\int n^d \phi(nx)\, dx}_{y=nx, \quad dy = n^d\, dx} = \int \phi(y)\, dy = 1$$
with $\text{supp}(\phi) \subset B_{1/n}(0)$.
- **Fact:** For $f \in C_c^\infty(\mathbb{R}^d)$, we have $f * \phi_n(x) \to f(x)$ for all $x$

**Question.** Is there a function $\delta_0$ such that $f * \delta_0(x) = f(x)$ ?

**Answer.** No, there is no such *function*. But there is such a "distribution".

**Definition 2.4.2.** A **distribution** is a functional $u : C_c^\infty(\mathbb{R}^d) \to \mathbb{R}$ that is

(i) linear, i.e. $u(\alpha f + \beta g) = \alpha u(f) + \beta u(g)$

(ii) continuous, i.e. if $f_k \to f$ then $u(f_k) \to u(f)$.

We denote the set of distributions by $\mathcal{D}(\mathbb{R}^d)$. Distributions are sometimes called 'generalized functions'. In this context, elements of $C_c^\infty$ are often called 'test functions'. The space $\mathcal{D}$ is the "dual space" of $C_c^\infty$.

*Example* 2.4.1.

- if $u \in C_c^\infty$, then we may think of $u$ as a distribution by defining
$$u(f) = \int u(x) f(x)\, dx. \quad (*)$$

- even for more general functions $u$, we may think of $u$ as a distribution via $(*)$
- **Dirac delta:** $\delta_0(f) = f(0)$

**Remark 2.4.3.** We will define $f * u$ for $f \in C_c^\infty$ and $u \in \mathcal{D}$ below and show $f * \delta_0 = f$.

**Derivatives of distributions.** For $u \in \mathcal{D}$, we would like to define $\partial_i u \in \mathcal{D}$.

First, suppose $u \in C_c^\infty$ and take a 'test function' $f \in C_c^\infty$.

By the integration by parts formula,

$$\underbrace{\int \partial_i u \, f \, dx}_{\partial_i u(f)} = \underbrace{-\int u \, \partial_i f \, dx}_{-u(\partial_i f)}.$$

Thus for $u \in \mathcal{D}$, we can **define** $\partial_i u \in \mathcal{D}$ by

$$\partial_i u(f) := -u(\partial_i f) \quad \text{for } f \in C_c^\infty.$$

This is very powerful: now we can take derivatives of non-differentiable functions!

Similarly, if $\mathcal{L}$ is a differential operator we can make sense of $\mathcal{L}u$.

**Convolution with distributions.** We would like to define $f * u$ for $f \in C_c^\infty$ and $u \in \mathcal{D}$.

First, suppose $f, u \in C_c^\infty$. Then

$$f * u(x) = u * f(x) = \int u(y) f(x-y) \, dy$$

**Notation:** $\tau_x f(y) = f(y-x)$ (translation)
$\tilde{f}(y) = f(-y)$ (reflection)

Then we have $f(x-y) = \widetilde{\tau_x f}(y)$, so that we can write

$$f * u(x) = \underbrace{\int u(y) \widetilde{\tau_x f}(y) \, dy}_{u(\widetilde{\tau_x f})}.$$

Thus for $f \in C_c^\infty$ and $u \in \mathcal{D}$ we **define** $f * u$ to be the function

$$f * u(x) := u(\widetilde{\tau_x f}).$$

**General principle.** Whatever you can define for $C_c^\infty$ you can also define for $\mathcal{D}$, provided the definitions agree when restricted to $C_c^\infty$.

Let's verify that $f * \delta_0 = f$ for all $f \in C_c^\infty$:

$$f * \delta_0(x) = \delta_0(\widetilde{\tau_x f}) = \widetilde{\tau_x f}(0) = f(x).$$

**Definition 2.4.4.** Let $\mathcal{L}$ be a differential operator. A distribution $\Phi \in \mathcal{D}$ is a **fundamental solution** for $\mathcal{L}$ if $\mathcal{L}(\Phi) = \delta_0$.

The key property of the fundamental solution is the following: if $\Phi$ is a fundamental solution for $\mathcal{L}$, then the solution to $\mathcal{L}u = f$ is given by

$$u = f * \Phi.$$

Indeed,

$$\mathcal{L}(u) = \mathcal{L}(f * \Phi) = f * \mathcal{L}(\Phi) = f * \delta_0 = f.$$

Question: why does $f * \phi_n(x) \to f(x)$ ?

# Chapter 3

# The Laplace Equation

## 3.1 The Fundamental Solution

We seek the fundamental solution for the Laplace operator $-\Delta$, that is, a distribution $\Phi$ such that $-\Delta\Phi = \delta_0$. Then we can solve **Poisson's equation** $-\Delta u = f$ on $\mathbb{R}^d$ by defining $u = f * \Phi$.

We focus on $d \geq 3$ (see the exercises for $d = 1, 2$). We begin by trying to find a function $u$ such that $\Delta u = 0$ on $\mathbb{R}^d \backslash \{0\}$.

The key observation is the following: the Laplace equation is invariant under rotations (see the exercises). So we look for a function that is also invariant under rotations:

$$u(x) = v(r), \qquad r = |x| = \sqrt{x_1^2 + \cdots + x_d^2}.$$

We compute:

$$\begin{aligned}
\partial_i u &= v'(r)\partial_i r = v'(r)\frac{x_i}{|x|}, \qquad ' = \frac{d}{dr}, \\
\partial_i \partial_i u &= v''(r)\frac{x_i^2}{|x|^2} + v'(r)\left[\frac{1}{|x|} - \frac{x_i^2}{|x|^3}\right], \\
\Delta u &= \sum_{i=1}^{d} \partial_i \partial_i u = v''(r) + \tfrac{d-1}{r} v'(r)
\end{aligned}$$

Thus

$$\begin{aligned}
&\Delta u = 0 \quad \text{if and only if} \quad v'' + \tfrac{d-1}{r} v' = 0, \\
&\implies \tfrac{v''}{v'} = -\tfrac{d-1}{r} \\
&\implies [\log(v')]' = -\tfrac{d-1}{r}, \\
&\implies \log(v') = -(d-1)\log r + C_1 = \log r^{-(d-1)} + C_1, \\
&\implies v' = C_2 r^{-(d-1)} \\
&\implies v = -C r^{-(d-2)}.
\end{aligned}$$

Therefore

$$u(x) = C|x|^{-(d-2)} \quad \text{solves} \quad -\Delta u = 0 \quad \text{on} \quad \mathbb{R}^d \backslash \{0\} \quad \text{for} \quad d \geq 3.$$

We now show show that we can choose $C$ so that $\Phi(x) = C|x|^{-(d-2)}$ is a fundamental solution, that is, $-\Delta\Phi = \delta_0$. We consider $\Phi(x)$ as a distribution and take a test function $f \in C_c^\infty$. Then

$$\begin{aligned}
-\Delta\Phi(f) = f(0) &\iff -\Phi(\Delta f) = f(0) \\
&\iff -\int \Phi(x)\Delta f(x)\,dx = f(0) \\
&\iff -\int C|x|^{-(d-2)}\Delta f(x)\,dx = f(0) \\
&\iff \int |x|^{-(d-2)}\Delta f(x)\,dx = -\tfrac{1}{C} f(0).
\end{aligned}$$

Therefore we need to compute $\int |x|^{-(d-2)}\Delta f(x)\,dx$ for an arbitrary test function $f$.

First let $\varepsilon > 0$ and write

$$\int |x|^{-(d-2)}\Delta f(x)\,dx = \underbrace{\int_{B_\varepsilon(0)} |x|^{-(d-2)}\Delta f(x)\,dx}_{I} + \underbrace{\int_{\mathbb{R}^d \backslash B_\varepsilon(0)} |x|^{-(d-2)}\Delta f(x)\,dx}_{II}.$$

For $I$, first use polar coordinates:

$$\int_{B_\varepsilon(0)} |x|^{-(d-2)}\,dx = \int_0^\varepsilon \int_{\partial B_1(0)} r^{-d+2} r^{d-1}\,dS\,dr = \int_0^\varepsilon r\,dr = \tfrac{1}{2}\varepsilon^2.$$

So

$$|I| \leq \max_{x\in\mathbb{R}^d} |\Delta f(x)| \cdot \tfrac{1}{2}\varepsilon^2 \leq c\varepsilon^2.$$

For $II$, we integrate by parts: (draw picture)

$$\int_{\mathbb{R}^d\setminus B_\varepsilon(0)} |x|^{-(d-2)}\Delta f(x)\,dx = -\int_{\mathbb{R}^d\setminus B_\varepsilon(0)} \nabla(|x|^{-(d-2)})\cdot\nabla f(x)\,dx \qquad III$$
$$+\int_{\partial B_\varepsilon(0)} |x|^{-(d-2)}\nabla f(x)\cdot\mathbf{n}(x)\,dS \qquad IV$$

We have

$$|IV| \le \varepsilon^{-(d-2)} \max_{x\in\mathbb{R}^d} |\nabla f(x)| \cdot \text{Area}(\partial B_\varepsilon(0)) \le c\varepsilon^{-(d-2)}\varepsilon^{d-1} \le c\varepsilon.$$

For $III$ we integrate by parts again: (draw picture)

$$III = \int_{\mathbb{R}^d\setminus B_\varepsilon(0)} \Delta\big(|x|^{-(d-2)}\big) f(x)\,dx \quad (\equiv 0)$$
$$-\int_{\partial B_\varepsilon(0)} \nabla\big(|x|^{-(d-2)}\big) f(x)\cdot\mathbf{n}(x)\,dS. \qquad V$$

For $x \in \partial B_\varepsilon(0)$ we have

$$\nabla\big(|x|^{-(d-2)}\big) = -(d-2)|x|^{-d}x = -(d-2)\varepsilon^{-d}x,$$
$$\mathbf{n}(x) = -\frac{x}{|x|} = -\varepsilon^{-1}x,$$
$$\nabla\big(|x|^{-(d-2)}\big)\cdot\mathbf{n}(x) = (d-2)\varepsilon^{-d-1}|x|^2 = (d-2)\varepsilon^{-(d-1)}.$$

Let $\alpha_d = \text{Volume}(B_1(0))$, so that $\text{Area}(\partial B_\varepsilon(0)) = d\varepsilon^{d-1}\alpha_d$.

$$V = -(d-2)\varepsilon^{-(d-1)}\int_{\partial B_\varepsilon(0)} f(x)\,dS = -(d-2)d\alpha_d \frac{1}{\text{Area}(\partial B_\varepsilon(0))}\int_{\partial B_\varepsilon(0)} f(x)\,dS.$$

Putting it all together: for any $\varepsilon > 0$, we have

$$\int |x|^{-(d-2)}\Delta f(x)\,dx = -d(d-2)\alpha_d \underbrace{\frac{1}{\text{Area}(\partial B_\varepsilon(0))}\int_{\partial B_\varepsilon(0)} f(x)\,dS}_{\text{average value of } f \text{ on } \partial B_\varepsilon(0)} + \mathcal{O}(\varepsilon)$$

We send $\varepsilon \to 0$ to find

$$\int |x|^{-(d-2)}\Delta f(x)\,dx = -d(d-2)\alpha_d f(0).$$

We conclude that

$$\Phi(x) = \tfrac{1}{d(d-2)\alpha_d}|x|^{-(d-2)}$$

is a fundamental solution for $-\Delta$ for $d \ge 3$.

**Conclusion.** The solution to Poisson's equation $-\Delta u = f$ on $\mathbb{R}^d$ is given by

$$u(x) = f * \Phi(x) = \tfrac{1}{d(d-2)\alpha_d}\int_{\mathbb{R}^d} \tfrac{f(y)}{|x-y|^{(d-2)}}\,dy.$$

## 3.2 Green's Functions

So far, we have solved Poisson's equation $-\Delta u = f$ on all of $\mathbb{R}^d$ by convolution with the fundamental solution $\Phi$. We turn to the **Dirichlet problem**, which is an example of a **boundary value problem**.

Let $\Omega \subset \mathbb{R}^d$ be an open, bounded set with smooth boundary $\partial\Omega$. Given functions $f$ and $g$, we study the problem

$$\begin{cases} -\Delta u = f & \text{on } \Omega \\ u = g & \text{on } \partial\Omega. \end{cases} \qquad (*)$$

We will break $(*)$ into two subproblems:

$$\begin{cases} -\Delta u = f & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \qquad (*_1) \qquad\qquad \begin{cases} -\Delta u = 0 & \text{on } \Omega \\ u = g & \text{on } \partial\Omega \end{cases} \qquad (*_2)$$

Note if $u_1$ solves $(*_1)$ and $u_2$ solves $(*_2)$, then $u = u_1 + u_2$ solves $(*)$.

We first turn to $(*_1)$. For the case of $\Omega = \mathbb{R}^d$, we computed $-\Delta\Phi = \delta_0$, so that

$$u(x) = \int_{\mathbb{R}^d} \Phi(x-y) f(y)\, dy \quad \text{solves} \quad -\Delta u = f.$$

The key fact about $\Phi$ was the following: $u = -\Delta u * \Phi$, that is,

$$-\int_{\mathbb{R}^d} \Phi(x-y)\Delta u(y)\, dy = u(x).$$

We will look for a solution to $(*_1)$ of the form

$$u(x) = \int_\Omega G(x,y) f(y)\, dy.$$

The function $G$ will be called the **Green's function** for $\Omega$.

We need to find $G$ so that

$$(i) \quad -\int_\Omega G(x,y)\Delta u(y)\, dy = u(x) \quad \text{whenever} \quad u = 0 \quad \text{on} \quad \partial\Omega.$$

This will guarantee $-\Delta u = f$ on $\Omega$. To have $u = 0$ on $\partial\Omega$ we also need

$$(ii) \quad G(x,y) = 0 \text{ for } x \in \partial\Omega,\ y \in \Omega.$$

**First try:** can we just take $G(x,y) = \Phi(x-y)$ again, as in the case of $\mathbb{R}^d$?

For $x \in \Omega$, we have

$$
\begin{aligned}
&-\int_\Omega \Phi(x-y)\Delta u(y)\,dy \\
&\quad = \int_\Omega \nabla[\Phi(x-y)]\cdot\nabla u(y)\,dy - \int_{\partial\Omega}\Phi(x-y)\nabla u(y)\cdot\mathbf{n}(y)\,dS \\
&\quad = -\int_\Omega \Delta\Phi(x-y)u(y)\,dy + \underbrace{\int_{\partial\Omega} u(y)\nabla[\Phi(x-y)]\cdot\mathbf{n}(y)\,dS}_{u=0 \text{ on } \partial\Omega} - \int_{\partial\Omega}\Phi(x-y)\nabla u(y)\cdot\mathbf{n}(y)\,dS \\
&\quad = \int_\Omega \delta_0(x-y)u(y)\,dy - \int_{\partial\Omega}\Phi(x-y)\nabla u(y)\cdot\mathbf{n}(y)\,dS. \\
&\quad = u(x) - \int_{\partial\Omega}\Phi(x-y)\nabla u(y)\cdot\mathbf{n}(y)\,dS.
\end{aligned}
$$

We missed...

**Second try:** How about $G(x,y) = \Phi(x-y) - e(x,y)$ for some "corrector" function $e$?

Computing as above (using $u = 0$ on $\partial\Omega$), we find

$$\int_\Omega e(x,y)\Delta u(y) = \int_\Omega \Delta e(x,y)u(y)\,dy + \int_{\partial\Omega} e(x,y)\nabla u(y)\cdot\mathbf{n}(y)\,dS.$$

So

$$
\begin{aligned}
-\int_\Omega G(x,y)\Delta u(y)\,dy = u(x) &\\
&+ \int_\Omega \Delta e(x,y)u(y)\,dy + \int_{\partial\Omega}[e(x,y)-\Phi(x-y)]\nabla u(y)\cdot\mathbf{n}(y)\,dS
\end{aligned}
$$

We should therefore choose $e(x,y)$ such that

$$
\begin{aligned}
&(a) \quad -\Delta_y e(x,y) = 0 \quad \text{for } x,y \in \Omega \\
&(b) \quad e(x,y) = \Phi(x-y) \quad \text{for } x \in \Omega,\ y \in \partial\Omega.
\end{aligned}
$$

Together $(a)$ and $(b)$ guarantee that $G(x,y) = \Phi(x-y) - e(x,y)$ satisfies $(i)$.

Moreover $(b)$ implies

$$(ii') \qquad G(x,y) = 0 \quad \text{for} \quad x \in \Omega,\ y \in \partial\Omega.$$

This is not quite condition $(ii)$. However, we will show the following:

$$\textbf{Claim:} \quad G(x,y) = G(y,x).$$

Then it is clear that $(ii')$ implies $(ii)$.

**Proof of claim.** Fix $x, y \in \Omega$. Define $v(z) = G(x, z)$. By $(ii')$, we know $v = 0$ on $\partial\Omega$. Thus

$$\begin{aligned} G(x,y) &= v(y) \\ &= -\int_\Omega G(y,z)\Delta v(z)\,dz \\ &= -\int_\Omega G(y,z)\Delta G(x,z)\,dz \\ &= \int_\Omega G(y,z)\delta_0(x-z)\,dz \\ &= G(y,x). \end{aligned}$$

**Conclusion (part one).** Suppose $\Omega \subset \mathbb{R}^d$ is an open, bounded set with smooth boundary $\partial\Omega$. Suppose that $e$ is a "corrector" function satisfying

$$\begin{cases} -\Delta_y e(x,y) = 0 & \text{for } x, y \in \Omega \\ e(x,y) = \Phi(x-y) & \text{for } x \in \Omega,\ y \in \partial\Omega. \end{cases} \qquad (**)$$

Then

$$G(x,y) = \Phi(x-y) - e(x,y)$$

is the **Green's function** for $\Omega$.

The function

$$u(x) = \int_\Omega G(x,y) f(y)\,dy$$

solves $(*_1)$, that is:

$$\begin{cases} -\Delta u = f & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

**Remark 3.2.1.** The question of the existence of the Green's function is still open, since we still do not know how to solve $(**)$.

How about $(*_2)$? We can actually the solution to $(*_1)$. Suppose $v$ solves

$$\begin{cases} -\Delta v = -\Delta g & \text{on } \Omega \\ v = 0 & \text{on } \partial\Omega. \end{cases}$$

Then

$$u = g - v \quad \text{solves} \quad (*_2),$$

since we have $-\Delta u = 0$ on $\Omega$ and $u = g$ on $\partial\Omega$.

From above, we know

$$\begin{aligned} v(x) &= -\int_\Omega G(x,y)\Delta g(y)\,dy \\ &= \int_\Omega \nabla G(x,y)\cdot\nabla g(y)\,dy - \underbrace{\int_{\partial\Omega} G(x,y)\nabla g(y)\cdot\mathbf{n}(y)\,dS}_{G(x,y)=0 \text{ for } x\in\Omega,\ y\in\partial\Omega} \\ &= -\int_\Omega \Delta G(x,y)\,g(y)\,dy + \int_{\partial\Omega} g(y)\nabla G(x,y)\cdot\mathbf{n}(y)\,dS \\ &= \int_\Omega \delta_0(x-y)\,g(y)\,dy + \int_{\partial\Omega} g(y)\nabla G(x,y)\cdot\mathbf{n}(y)\,dS \\ &= g(x) + \int_{\partial\Omega} g(y)\nabla G(x,y)\cdot\mathbf{n}(y)\,dS. \end{aligned}$$

Hence

$$u(x) = g(x) - v(x) = -\int_{\partial\Omega} g(y)\nabla G(x,y)\cdot\mathbf{n}(y)\,dS \quad \text{solves} \quad (*_2).$$

**Conclusion (part two).** If $G$ is the **Green's function** for $\Omega$, then the function

$$u(x) = \int_\Omega G(x,y)f(y)\,dy - \int_{\partial\Omega} g(y)\nabla_y G(x,y)\cdot\mathbf{n}(y)\,dS$$

solves

$$\begin{cases} -\Delta u = f & \text{on } \Omega \\ u = g & \text{on } \partial\Omega. \end{cases}$$

We will next study a few cases when the Green's function can be explicitly computed.

Recall that for $\Omega\subset\mathbb{R}^d$, the **Green's function** for $\Omega$ is the function

$$G(x,y) = \Phi(x-y) - e(x,y),$$

where $\Phi$ is the fundamental solution for $-\Delta$ and $e$ is a "**corrector**" function satisfying

$$\begin{cases} -\Delta_y e(x,y) = 0 & \text{for } x,y\in\Omega \\ e(x,y) = \Phi(x-y) & \text{for } x\in\Omega,\ y\in\partial\Omega. \end{cases} \qquad (*)$$

We compute the Green's function for two specific cases.

*Example* 3.2.1 (Half-space). Let

$$\Omega = \mathbb{R}^d_+ = \{x \in \mathbb{R}^d : x_d \geq 0\}.$$

Then $\partial\Omega = \{x \in \mathbb{R}^d : x_d = 0\}$.

For a point $y \in \Omega$, define the **reflection** of $y$ by

$$\tilde{y} = (y_1, \ldots, y_{d-1}, -y_d).$$

We let

$$e(x, y) = \Phi(x - \tilde{y}) = \Phi(\tilde{x} - y) \quad \longleftarrow \Phi = \Phi(|x|) \text{ (draw picture).}$$

Then for $x, y \in \Omega$, we have $\tilde{x} \notin \Omega$, and hence

$$-\Delta_y e(x, y) = -\Delta_y \Phi(\tilde{x} - y) = \delta_0(\tilde{x} - y) = 0.$$

Moreover for $y \in \partial\Omega$ we have $y = \tilde{y}$, and hence

$$e(x, y) = \Phi(x - \tilde{y}) = \Phi(x - y) \quad \text{for} \quad x \in \Omega,\ y \in \partial\Omega.$$

Thus the Green's function for the half-space is given by

$$G(x, y) = \Phi(x - y) - \Phi(\tilde{x} - y), \quad \tilde{x} = (x_1, \ldots, x_{d-1}, -x_d). \quad \square$$

*Example* 3.2.2 (Unit Ball). Let $\Omega = B_1(0)$, so $\partial\Omega = \partial B_1(0)$.

For a point $x \in \Omega \backslash \{0\}$, define the **inversion** of $x$ by

$$\tilde{x} = \frac{x}{|x|^2}.$$

As a first try, suppose we let $e(x, y) = \Phi(\tilde{x} - y)$ again, since $-\Delta\Phi(\tilde{x} - y) = 0$. However, in this case we will not have $\Phi(\tilde{x} - y) = \Phi(x - y)$ for $x \in \Omega$, $y \in \partial\Omega$, since $|\tilde{x} - y| \neq |x - y|$. Indeed, let $x \in \Omega$ and $y \in \partial\Omega$. Then $|y| = 1$, so

$$\begin{aligned}
|\tilde{x} - y|^2 &= |\tilde{x}|^2 - 2\tilde{x} \cdot y + |y|^2 \\
&= \tfrac{1}{|x|^2} - 2\tfrac{x}{|x|^2} \cdot y + 1 \\
&= \tfrac{1}{|x|^2}\left[1 - 2x \cdot y + |x|^2\right] \\
&= \tfrac{1}{|x|^2}\left[|y|^2 - 2x \cdot y + |x|^2\right] \\
&= \tfrac{1}{|x|^2}|x - y|^2.
\end{aligned}$$

So we find $|x-y| = |x|\,|\tilde{x}-y|$ for $x \in \Omega$ and $y \in \partial\Omega$.

For a second try, let us take

$$e(x,y) = \Phi(|x|\,|\tilde{x}-y|).$$

Then $e(x,y) = \Phi(x-y)$ for $x \in \Omega$, $y \in \partial\Omega$.

But do we have $-\Delta_y e(x,y) = 0$ ? In fact we do! Recall $\Phi(z) = c|z|^{-(d-2)}$ (in dimensions $d \geq 3$). Then

$$\Phi(|x|\,|\tilde{x}-y|) = |x|^{-(d-2)}\Phi(\tilde{x}-y),$$

so

$$-\Delta_y e(x,y) = -|x|^{-(d-2)}\Delta\Phi(\tilde{x}-y) = 0$$

as above.

Thus the Green's function for the unit ball is given by:

$$G(x,y) = \Phi(x-y) - \Phi\big(|x|(\tilde{x}-y)\big), \quad \tilde{x} = \tfrac{x}{|x|^2}.$$

So far, we have attempted to find solutions to the Laplace and Poisson equation. We next discuss some properties of solutions.

## 3.3 Properties of Solutions

**Theorem 3.3.1** (Mean Value Property)**.** *Let $u$ be a smooth solution to $-\Delta u = 0$ on an open set $\Omega \subset \mathbb{R}^d$. If $\overline{B_R(x)} \subset \Omega$, then*

$$u(x) = \frac{1}{Area(\partial B_R(x))}\int_{\partial B_R(x)} u(y)\,dS \qquad (\textit{average of } u \textit{ over } \partial B_R(x))$$

*Proof.* Let

$$A_R := d\alpha_d R^{d-1} = \text{Area}(\partial B_R(x)) \qquad \longleftarrow \text{recall } \alpha_d = \text{Volume}(B_1(0))$$

Define $f : (0,R] \to \mathbb{R}$ via

$$f(r) := \frac{1}{A_r}\int_{\partial B_r(x)} u(y)\,dS = \underbrace{\frac{r^{d-1}}{A_r}\int_{\partial B_1(0)} u(x+rz)\,dS}_{z=\frac{y-x}{r},\ dS(z)=r^{-(d-1)}\,dS(y)} = \frac{1}{A_1}\int_{\partial B_1(0)} u(x+rz)\,dS.$$

By the continuity of $u$, we have

$$u(x) = \lim_{r\to 0} f(r). \qquad (1)$$

On the other hand, we can compute

$$\begin{aligned} f'(r) &= \frac{1}{A_1}\int_{\partial B_1(0)} \nabla u(x+rz)\cdot z\, dS(z) \qquad (y = x+rz) \\ &= \tfrac{1}{r^{d-1}A_1}\int_{\partial B_r(x)} \nabla u(y)\cdot \tfrac{y-x}{r}\, dS(y) \\ &= \tfrac{1}{A_r}\int_{\partial B_r(x)} \nabla u(y)\cdot \mathbf{n}(y)\, dS(y) \\ &= \tfrac{1}{A_r}\int_{B_r(x)} \Delta u(y)\, dy \qquad \text{(Gauss–Green formula)} \\ &= 0. \end{aligned}$$

Thus $f$ is constant. In particular $f(r) = f(R)$ for all $r \in (0, R]$. Therefore, from (1) we deduce

$$u(x) = f(R) = \frac{1}{A_R}\int_{\partial B_R(x)} u(y)\, dS.$$

□

In fact, we will prove that we also have

$$u(x) = \frac{1}{\text{Volume}(B_R(x))}\int_{B_R(x)} u(y)\, dy.$$

**Corollary 3.3.2.** *We also have that*

$$u(x) = \frac{1}{\mathit{Volume}(B_R(x))}\int_{B_R(x)} u(y)\, dy = \frac{1}{\alpha_d R^d}\int_{B_R(x)} u(y)\, dy.$$

*Proof.* Indeed, using polar coordinates:

$$\begin{aligned} \int_{B_R(x)} u(y)\, dy &= \int_0^R \int_{\partial B_1(x)} u(r\theta)\, dS\, r^{d-1}\, dr \\ &= \int_0^R \int_{\partial B_r(x)} u(z)\, dS\, dr \qquad \big(z = r\theta, \quad dS(z) = r^{d-1}\, dS(\theta)\big) \\ &= \int_0^R d\alpha_d r^{d-1} u(x)\, dr \qquad \text{(mean value property)} \\ &= \alpha_d R^d u(x). \end{aligned}$$

□

Thus if $u$ is harmonic, then $u$ satisfies the mean value property. The converse is true as well! (This can be shown by using similar computations.)

Next we will prove the maximum principle for harmonic functions. First, we need a definition.

**Definition 3.3.3.** A set $\Omega \subset \mathbb{R}^d$ is **connected** if $\Omega$ **cannot** be written as a disjoint union of non-empty open sets.

**Remark 3.3.4.** $\Omega$ connected $\iff$ only subsets of $\Omega$ that are both open and closed (in $\Omega$) are $\emptyset$ and $\Omega$.

(Indeed, if $\Omega = A \cup B$ where $A, B$ are disjoint non-empty open sets then $A$ and $B$ are both open and closed.)

**Theorem 3.3.5** (Maximum principle)**.** *Suppose $u$ is smooth and $\Omega \subset \mathbb{R}^d$ is open, bounded, and connected. Suppose*

$$-\Delta u = 0 \quad \text{in } \Omega \quad \text{and let} \quad M := \max_{x \in \overline{\Omega}} u(x).$$

*If there exists $x_0 \in \Omega$ such that $u(x_0) = M$, then $u \equiv M$ in $\Omega$.*

That is, non-constant harmonic functions attain their maximum value on the boundary.

*Proof.* We let

$$A = \{x \in \Omega : u(x) = M\}.$$

First notice that $A \neq \emptyset$, since $x_0 \in A$ by assumption.

Second, we claim $A$ is closed. Indeed, suppose $\{x_n\} \subset A$ and $x_n \to x \in \Omega$. Then since $u$ is continuous, we have $u(x_n) \to u(x)$. But since $x_n \in A$, we have $u(x_n) \equiv M$. Thus $u(x) = M$. That is, $x \in A$.

Third, we claim $A$ is open. Indeed, suppose $x \in A$. Then since $\Omega$ is open, we may find $R > 0$ so that $B_R(x) \subset \Omega$. Then using the mean value property, we have

$$M = u(x) = \frac{1}{\mathrm{Vol}(B_R(x))} \int_{B_R(x)} u(y)\, dy \leq \frac{1}{\mathrm{Vol}(B_R(x))} \int_{B_R(x)} M\, dy \leq M.$$

This implies that $u(y) = M$ for $y \in B_R(x)$. That is $B_R(x) \subset A$. This shows that $A$ is open.

As $\Omega$ is connected and $A$ is non-empty, open, and closed, we conclude that $A = \Omega$. That is, $u(x) \equiv M$ on $\Omega$. □

The maximum principle has an important corollary:

**Corollary 3.3.6** (Uniqueness for Poisson's equation)**.** *Let $\Omega \subset \mathbb{R}^d$ be open, bounded, and connected. Suppose $u_1$ and $u_2$ are both smooth solutions to*

$$\begin{cases} -\Delta u = f & \text{on } \Omega \\ u = g & \text{on } \partial\Omega \end{cases}$$

*Then $u_1 \equiv u_2$ on $\Omega$.*

*Proof.* Let $u = u_1 - u_2$. Then $u$ solves $-\Delta u = 0$ on $\Omega$, with $u = 0$ on $\partial\Omega$. By the maximum principle, we get $u \leq 0$ on $\Omega$. That is, $u_1 \leq u_2$ on $\Omega$. Next let $\tilde{u} = u_2 - u_1$. Then $\tilde{u}$ solves $-\Delta\tilde{u} = 0$ on $\Omega$ with $\tilde{u} = 0$ on $\partial\Omega$. By the maximum principle, we get $\tilde{u} \leq 0$ on $\Omega$. That is, $u_2 \leq u_1$ on $\Omega$. We conclude $u_1 \equiv u_2$ on $\Omega$. □

# Chapter 4

# The Heat Equation

## 4.1 The Fundamental Solution

We turn to the heat equation $\partial_t u - \Delta u = 0$. ($\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$)

We study the **initial-value problem** (also called the **Cauchy problem**):

$$\begin{cases} \partial_t u - \Delta u = 0 & \text{for } (t,x) \in (0,\infty) \times \mathbb{R}^d \\ u(0,x) = f(x) & \text{for } x \in \mathbb{R}^d \end{cases} \qquad (IVP)$$

We first find the fundamental solution. We look for a function so that $\partial_t u - \Delta u = 0$ away from $(t,x) = (0,0)$.

Observation 1: if $u$ solves $\partial_t u - \Delta u = 0$, then so does $u(\lambda^2 t, \lambda x)$ for any $\lambda > 0$.

We look for a solution that is invariant under this rescaling, for example

$$u(t,x) = v(\tfrac{x}{\sqrt{t}}) \quad \text{for some function} \quad v. \qquad (*)$$

However, this is not a good guess. To see why not, we need the following:

Observation 2: if $u$ solves $\partial_t u - \Delta u = 0$, then $\int u(t,x)\,dx$ is constant in $t$.

(Physically, this is clear. For a proof, see HW.)

If we make the guess of $(*)$, however, we find

$$\begin{aligned} \int u(t,x)\,dx &= \int v(\tfrac{x}{\sqrt{t}})\,dx \\ &= t^{d/2} \int v(y)\,dy \qquad (y = \tfrac{x}{\sqrt{t}}, \quad dy = t^{-d/2}dx) \\ &\neq \text{constant in } t \end{aligned}$$

To fix this, we look for a solution of the form

$$u(t,x) = t^{-d/2} v(\tfrac{x}{\sqrt{t}}).$$

Then we have scaling invariance **and** $\int u\,dx$ constant in $t$.
As before, we also assume $v(x) = v(|x|)$. Then

$$u(t,x) = t^{-d/2} v\big(\tfrac{|x|}{\sqrt{t}}\big).$$

Let's write $y = \frac{|x|}{\sqrt{t}}$, so that $u(t,x) = t^{-d/2} v(y)$.
Then

$$\begin{aligned} \partial_t u &= t^{-\frac{d}{2}} v'(y) \cdot -\tfrac{1}{2} t^{-\frac{3}{2}} |x| - \tfrac{d}{2} t^{-\frac{d}{2}-1} v \\ &= -\tfrac{1}{2} t^{-\frac{d}{2}-1} y v' - \tfrac{d}{2} t^{-\frac{d}{2}-1} v. \end{aligned}$$

$$\begin{aligned} u_{x_i} &= t^{-\frac{d}{2}} v'(y) \cdot t^{-\frac{1}{2}} \tfrac{x_i}{|x|} \\ &= t^{-\frac{d}{2}-\frac{1}{2}} v'(y) \tfrac{x_i}{|x|} \end{aligned}$$

$$u_{x_i x_i} = t^{-\frac{d}{2}-\frac{1}{2}} v'(y) \big[\tfrac{1}{|x|} - \tfrac{x_i^2}{|x|^3}\big] + t^{-\frac{d}{2}-1} v'' \tfrac{x_i^2}{|x|^2}$$

$$\begin{aligned} \Delta u &= t^{-\frac{d}{2}-\frac{1}{2}} v'(y) \tfrac{d-1}{|x|} + t^{-\frac{d}{2}-1} v'' \\ &= t^{-\frac{d}{2}-1} v'(y) \tfrac{d-1}{y} + t^{-\frac{d}{2}-1} v'' \end{aligned}$$

$$\partial_t u - \Delta u = -t^{-\frac{d}{2}-1} \big[\tfrac{1}{2} y v' + \tfrac{d}{2} v + v'' + \tfrac{d-1}{y} v'\big].$$

So we want $\frac{1}{2} y v' + \frac{d}{2} v + v'' + \frac{d-1}{y} v' = 0$.
Or, multiplying by $y^{d-1}$:

$$\tfrac{1}{2} y^d v' + \tfrac{d}{2} v y^{d-1} + v'' y^{d-1} + (d-1) y^{d-2} v' = 0$$

$$\tfrac{1}{2} (v y^d)' + (v' y^{d-1})' = 0$$

$$\tfrac{1}{2} v y^d + v' y^{d-1} = 0$$

$$v' = -\tfrac{1}{2} y v$$

$$v = Ce^{-y^2/4}$$
$$u(t,x) = Ct^{-d/2}e^{-|x|^2/4t}.$$

The **fundamental solution** will be $\Phi(t,x) = Ct^{-d/2}e^{-|x|^2/4t}$ for some $C > 0$. (See below)

Consider again the problem (IVP).

Define

$$u(t,x) = [f * \Phi(t,\cdot)](x) = \int_{\mathbb{R}^d} f(y)\Phi(t,x-y)\,dx. \qquad (**)$$

Note $\partial_t u = f * \partial_t \Phi$ and $\Delta u = f * \Delta\Phi$, so

$$\partial_t u - \Delta u = f * [\partial_t \Phi - \Delta\Phi] = 0 \quad \text{for} \quad (t,x) \in (0,\infty) \times \mathbb{R}^d.$$

So $u$ solves the PDE. What about the initial condition?

Note $\Phi(t,x)$ form approximate identities as $t \to 0$, provided

$$\int \Phi(t,x)\,dx = Ct^{-d/2} \int e^{-|x|^2/4t}\,dx = 1.$$

In particular, we will have

$$\lim_{t\to 0} u(t,x) = \lim_{t\to 0}[f * \Phi(t,\cdot)](x) = f(x),$$

so that $(**)$ defines a solution to (IVP).

**Claim:** $(4\pi t)^{-d/2} \int e^{-|x|^2/4t}\,dx = 1$ (see homework)

Therefore

$$\Phi(t,x) = (4\pi t)^{-d/2}e^{-|x|^2/2t}$$

is the **fundamental solution** for the heat equation.

The function

$$u(t,x) = (4\pi t)^{-d/2} \int_{\mathbb{R}^d} e^{-|x-y|^2/4t} f(y)\,dy$$

solves

$$\begin{cases} \partial_t u - \Delta u = 0 & \text{for } (t,x) \in (0,\infty) \times \mathbb{R}^d \\ u(0,x) = f(x) & \text{for } x \in \mathbb{R}^d \end{cases} \qquad (IVP).$$

**Remark 4.1.1.** (i) Infinite speed of propagation.

(ii) Inhomogeneous problem $u_t - \Delta u = F$. Solved by 'Duhamel's principle' (more later)

## 4.2 The Mean Value Property

**Definition 4.2.1.**

(i) (Parabolic cylinder) For $\Omega \subset \mathbb{R}^d$ and $T > 0$ we define $\Omega_T \subset \mathbb{R} \times \mathbb{R}^d$ by $\Omega_T = \Omega \times (0, T]$.

(ii) (Parabolic boundary) We define $\Gamma_T = \overline{\Omega_T} \backslash \Omega_T$.

(iii) (Heat ball) For $t \in \mathbb{R}$, $x \in \mathbb{R}^d$, and $r > 0$ define

$$E_r(t, x) = \{(s, y) \in \mathbb{R} \times \mathbb{R}^d : \ s \leq t \text{ and } \Phi(t - s, x - y) \geq \tfrac{1}{r^d}\}.$$

To understand the heat ball definition, consider $E_1(0, 0)$. For $s \leq 0$,

$$(4\pi|s|)^{-d/2} e^{-|y|^2/4|s|} \geq 1 \quad \iff \quad |y|^2 \leq -2d|s| \ \log(4\pi|s|).$$

Thus for $-\frac{1}{4\pi} \leq s \leq 0$ we get $y$ in a ball of radius $\sqrt{-2d|s| \ \log(4\pi|s|)}$.

For $E_r(0, 0)$: for $-\frac{r^2}{4\pi} \leq s \leq 0$ we get $y$ in a ball of radius $\sqrt{-2d|s| \ \log(\frac{4\pi|s|}{r^2})}$.

**Theorem 4.2.2** (Mean value property.)**.** *Let $u$ solve $\partial_t - \Delta u = 0$ on $\Omega_T$. Suppose $E_R(t, x) \subset \Omega_T$. Then*

$$u(t, x) = \frac{1}{4R^d} \iint_{E_R(t,x)} u(s, y) \frac{|x - y|^2}{(t - s)^2} \, dy \, ds.$$

*Proof.* Without loss of generality, suppose $(t, x) = (0, 0)$. (see homework)

Define $f : (0, R] \to \mathbb{R}$ by

$$f(r) = \frac{1}{4r^d} \iint_{E_r(0,0)} u(s, y) \frac{|y|^2}{s^2} \, dy \, ds.$$

We will show that

(i) $\lim_{r \to 0} f(r) = u(0, 0)$

(ii) $f'(r) \equiv 0$

With (i) and (ii) we can conclude $u(0, 0) = f(R) = \iint_{E_R(0,0)} u(s, y) \frac{|y|^2}{s^2} \, dy \, ds$, as needed.

For (i), we need the following

**Fact.** $\displaystyle \frac{1}{4r^d} \iint_{E_r(0,0)} \frac{|y|^2}{s^2} \, dy \, ds = 1$ for all $r > 0$. (long computation)

Thus

$$\begin{aligned}|f(r)-u(0,0)| &\le \frac{1}{4r^d}\iint_{E_r(0,0)} \big|u(s,y)-u(0,0)\big|\frac{|y|^2}{s^2}\,dy\,ds.\\ &\le \max_{(s,y)\in E_r(0,0)} |u(s,y)-u(0,0)|\\ &\to 0 \qquad \text{as } r\to 0,\end{aligned}$$

since $u$ is continuous and $|(s,y)|\to 0$ for all $(s,y)\in E_r(0,0)$ as $r\to 0$.

For (ii), we change variables via

$$y = rz \quad \text{and} \quad s = r^2\sigma.$$

Then

$$dy\,ds = r^{d+2}\,dz\,d\sigma \quad \text{and} \quad \frac{|y|^2}{s^2} = r^{-2}\frac{|z|^2}{\sigma^2}.$$

Moreover

$$\begin{aligned}&(4\pi|s|)^{-d/2}e^{-|y|^2/4|s|} \ge r^{-d}\\ \iff &(4\pi r^2|\sigma|)^{-d/2}e^{-|z|^2/4|\sigma|} \ge r^{-d}\\ \iff &(4\pi|\sigma|)^{-d/2}e^{-|z|^2/4|\sigma|} \ge 1\end{aligned}$$

so $(s,y)\in E_r(0,0) \iff (\sigma,z)\in E_1(0,0)$. Thus

$$f(r) = \frac{1}{4}\iint_{E_1(0,0)} u(r^2\sigma, rz)\frac{|z|^2}{\sigma^2}\,dz\,d\sigma.$$

We compute and undo the change of variables:

$$\begin{aligned}f'(r) &= \frac{1}{4}\iint_{E_1(0,0)}\Big[2r\sigma[\partial_\sigma u](r^2\sigma,rz) + [\nabla_z u](r^2\sigma,rz)\cdot z\Big]\frac{|z|^2}{\sigma^2}\,dz\,d\sigma\\ &= \frac{1}{4r^{d+1}}\iint_{E_r(0,0)} \underbrace{2\partial_s u(s,y)\frac{|y|^2}{s}}_{I} + \underbrace{\nabla u(s,y)\cdot y\frac{|y|^2}{s^2}}_{II}\,dy\,ds.\end{aligned}$$

We now introduce an auxiliary function $\psi = \psi(r,s,y)$ such that $\psi = 0$ on $\partial E_r(0,0)$, that is:

$$(4\pi|s|)^{-d/2}e^{-|y|^2/4|s|}r^d = 1 \implies \psi = 0$$

We take $\psi(r,s,y) = -\frac{d}{2}\log(-4\pi s) + \frac{|y|^2}{4s} + d\log r.$

Note $\nabla_y \psi = \frac{y}{2s}, \quad \partial_s \psi = -\frac{d}{2s} - \frac{|y|^2}{4s^2}$.
We now rewrite things and integrate by parts:

$$\begin{aligned}
I &= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} \partial_s u \; y \cdot \nabla_y \psi \, ds \, dy \\
&= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} -d \, \partial_s u \, \psi - y \cdot \nabla \partial_s u \, \psi \, ds \, dy \qquad \text{(parts in } y\text{)} \\
&= -\frac{1}{r^{d+1}} \iint_{E_r(0,0)} -d \, \partial_s u \, \psi + y \cdot \nabla u \, \partial_s \psi \, ds \, dy \qquad \text{(parts in } s\text{)} \\
&= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} -d \, \partial_s u \, \psi - y \cdot \nabla u \left( \frac{d}{2s} + \frac{|y|^2}{4s^2} \right) ds \, dy \\
&= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} -d \, \partial_s u \, \psi - \frac{d}{2s} y \cdot \nabla u \, ds \, dy \qquad - II.
\end{aligned}$$

So, using $u_t = \Delta u$,

$$\begin{aligned}
f'(r) &= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} -d \, \partial_s u \, \psi - \frac{d}{2s} y \cdot \nabla u \, ds \, dy \\
&= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} -d \, \Delta u \, \psi - \frac{d}{2s} y \cdot \nabla u \, ds \, dy \\
&= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} d \nabla u \cdot \nabla \psi - \frac{d}{2s} y \cdot \nabla u \, ds \, dy \\
&= \frac{1}{r^{d+1}} \iint_{E_r(0,0)} \frac{d}{2s} y \cdot \nabla u - \frac{d}{2s} y \cdot \nabla u \, ds \, dy \\
&= 0.
\end{aligned}$$

□

## 4.3 The Maximum Principle and Uniqueness

Recall for $\Omega \subset \mathbb{R}^d$ and $T > 0$ we define $\Omega_T = \Omega \times (0, T]$ and $\Gamma_T = \overline{\Omega_T} \backslash \Omega_T$.

**Theorem 4.3.1** (Maximum principle)**.** *Suppose $\Omega \subset \mathbb{R}^d$ is open, bounded, and connected. Suppose $u_t - \Delta u = 0$ on $\Omega_T$ for some $T > 0$. Let $M := \max_{(t,x) \in \overline{\Omega_T}} u(t,x)$. If there exists $(t_0, x_0) \in \Omega_T$ such that $u(t_0, x_0) = M$, then $u(t,x) \equiv M$ on $\overline{\Omega_{t_0}}$.*

In particular, solutions attain their maximum on $\Gamma_T$.

*Proof.* Notice that for $r > 0$ small enough, we have $E_r(t_0, x_0) \subset \Omega_T$. By the mean value formula,

$$M = u(t_0, x_0) = \frac{1}{4r^n} \iint_{E_r(t_0,x_0)} u(s, y) \frac{|x_0 - y|^2}{(t_0 - s)^2} \, dy \, ds \leq M,$$

since

$$\frac{1}{4r^n} \iint_{E_r(t_0,x_0)} \frac{|x_0 - y|^2}{(t_0 - s)^2} \, dy \, ds = 1 \qquad \text{(long computation)}.$$

Thus $u = M$ for $(t, x) \in E_r(t_0, x_0)$.

Now suppose $(s_0, y_0) \in \Omega_T$ is a point with $s_0 < t_0$ that can be connected to to $(t_0, x_0)$ by a line segment $L \subset \Omega_T$. We will show $u = M$ on $L$.

Let

$$r_0 = \min\{s \geq s_0 : \ u(t, x) = M \text{ for } (t, x) \in L, \ s \leq t \leq t_0\}.$$

We need to show that $r_0 = s_0$. Suppose toward a contradiction that $r_0 > s_0$. Then we can find $z_0 \in \Omega$ so that $(r_0, z_0) \in \Omega_T \cap L$ and $u(r_0, z_0) = M$.

But then as above we have $u = M$ on $E_r(r_0, z_0)$ for some small $r > 0$.

But $E_r(r_0, z_0)$ contains points $(s, y) \in L$ with $s < r_0$, which is a contradiction.

Thus $u \equiv M$ on $L$.

Finally fix any $(t, x) \in \Omega_T$ with $t < t_0$. We want to show $u(t, x) = M$.

We can connect $(t_0, x_0)$ to $(t, x)$ by a piecewise linear path in $\Omega_T$.

By the arguments above, $u = M$ on each segment, so we finally get that $u(t, x) = M$. □

**Corollary 4.3.2** (Uniqueness)**.** *Let $\Omega \subset \mathbb{R}^d$ be open, bounded, and connected, and let $T > 0$. Suppose $u_1$ and $u_2$ are smooth solutions to*

$$\begin{cases} u_t - \Delta u = f & \text{on } \Omega_T \\ u = g & \text{on } \Gamma_T. \end{cases}$$

*Then $u_1 \equiv u_2$ on $\Omega_T$.*

*Proof.* Define $v = u_1 - u_2$. Then $v$ solves

$$\begin{cases} v_t - \Delta v = 0 & \text{on } \Omega_T \\ v = 0 & \text{on } \Gamma_T. \end{cases}$$

Thus the maximum principle implies that $v \leq 0$ on $\Omega_T$, so $u_1 \leq u_2$ on $\Omega_T$.

Similarly, with $w = u_2 - u_1$ we deduce $u_2 \leq u_1$ on $\Omega_T$. Thus $u_1 = u_2$ on $\Omega_T$. □

There is also maximum principle/uniqueness for the case $\Omega = \mathbb{R}^d$, but they require an extra condition.

**Proposition 4.3.3** (Maximum principle, $\mathbb{R}^d$ version)**.** *Suppose $u$ solves*

$$\begin{cases} u_t - \Delta u = 0 & (t,x) \in (0,T] \times \mathbb{R}^d \\ u(0,x) = f(x) & x \in \mathbb{R}^d \end{cases}$$

*If $u$ satisfies*

$$u(t,x) \leq Ae^{a|x|^2} \quad \textit{for } (t,x) \in [0,T] \times \mathbb{R}^d,$$

*then*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} u(t,x) = \sup_{x\in\mathbb{R}^d} g(x).$$

**Remark 4.3.4.** Here sup means...

**Proposition 4.3.5** (Uniqueness, $\mathbb{R}^d$ version)**.** *Suppose $u_1, u_2$ solve*

$$\begin{cases} u_t - \Delta u = 0 & (t,x) \in (0,T] \times \mathbb{R}^d \\ u(0,x) = f(x) & x \in \mathbb{R}^d \end{cases}$$

*Suppose that $u_1, u_2$ satisfy*

$$u_1(t,x), u_2(t,x) \leq Ae^{a|x|^2} \quad \textit{for } (t,x) \in [0,T] \times \mathbb{R}^d.$$

*Then $u_1 \equiv u_2$.*

To see why this extra condition is necessary, consider the following example.

*Example* 4.3.1. Consider the heat equation in dimension $d = 1$:

$$\begin{cases} u_t - u_{xx} = 0 & (t,x) \in (0,\infty) \times \mathbb{R} \\ u(0,x) = 0 & x \in \mathbb{R} \end{cases} \qquad (*)$$

Clearly $u(x) \equiv 0$ is a solution to $(*)$ However, we can define a second solution $v$ as follows. Let

$$g(t) = \begin{cases} e^{-1/t^2} & t > 0 \\ 0 & t = 0 \end{cases}$$

Then

$$v(t,x) = g(t) + g'(t)\tfrac{x^2}{2!} + g''(t)\tfrac{x^4}{4!} + g'''(t)\tfrac{x^6}{6!} + \dots$$

is also a solution to $(*)$. (Non-trivial issue: convergence of this series.)

Indeed,

$$v_{xx} = 0 + g'(t) + g''(t)\tfrac{x^2}{2!} + g'''(t)\tfrac{x^4}{4!} + \dots,$$

$$v_t = g'(t) + g''(t)\tfrac{x^2}{2!} + g'''(t)\tfrac{x^4}{4!} + \dots,$$

and

$$v(0,x) = g(0) + g'(0)\tfrac{x^2}{2!} + g''(0)\tfrac{x^4}{4!} + \dots = 0.$$

(Note that $v$ has rapid growth as $|x| \to \infty$.)

# Chapter 5

# The Wave Equation

## 5.1 The Fundamental Solution

We turn to the **wave equation**: $u_{tt} - \Delta u = 0 \qquad (t,x) \in \mathbb{R} \times \mathbb{R}^d$.

Here $u = u(t,x)$, $\quad u_{tt} = \frac{\partial^2 u}{\partial t^2}$, $\quad \Delta u = \sum_{j=1}^{d} \frac{\partial^2 u}{\partial x_j^2}$.

We write $\Box u = u_{tt} - \Delta u$. The operator $\Box$ is called the **d'Alembertian**.

We study the following initial value problem.

$$(IVP) \quad \begin{cases} \Box u = 0 & (t,x) \in (0,\infty) \times \mathbb{R}^d \\ u(0,x) = f(x) & x \in \mathbb{R}^d \\ u_t(0,x) = g(x) & x \in \mathbb{R}^d. \end{cases}$$

(This time we need to impose two initial conditions since the equation is second order in $t$.)

As usual, we want to compute the fundamental solution $\Phi$. We will solve

$$\begin{cases} \Box \Phi = 0 \\ \Phi(0) = 0, \quad \Phi_t(0) = \delta_0. \end{cases} \qquad (*)$$

Then the solution to (IVP) will be given by

$$u(t,x) = [f * \partial_t \Phi(t,\cdot)](x) + [g * \Phi(t,\cdot)](x).$$

We can write this more succintly as

$$u(t) = f * \partial_t \Phi(t) + g * \Phi(t).$$

Indeed, first we have

$$\Box u = f * \partial_t(\Box \Phi) + g * \Box \Phi = 0,$$

so that $u$ solves the PDE.

For the initial conditions: First, $u(0) = f * \delta_0 + f * 0 = f$. Next, since

$$u_t = f * \Phi_{tt} + g * \Phi_t = f * \Delta\Phi + g * \Phi_t = \Delta(f * \Phi) + g * \Phi_t,$$

we have $u_t(0) = \Delta(f * 0) + g * \delta_0 = g$.

### 5.1.1 The Case $d = 1$

Then $(*)$ reads

$$\begin{cases} \Phi_{tt} - \Phi_{xx} = 0, \\ \Phi(0) = 0, \quad \Phi_t(0) = \delta_0. \end{cases}$$

Inspired by the factorization $\partial_{tt} - \partial_{xx} = (\partial_t - \partial_x)(\partial_t + \partial_x)$, we introduce the variables

$$\xi = x + t, \quad \eta = x - t, \qquad\qquad t = \tfrac{1}{2}(\xi - \eta), \quad x = \tfrac{1}{2}(\xi + \eta).$$

We write

$$\Phi(t, x) = \Phi(\tfrac{\xi-\eta}{2}, \tfrac{\xi+\eta}{2}) =: v(\xi, \eta).$$

We have

$$\begin{aligned} v_\xi &= \tfrac{1}{2}\Phi_t + \tfrac{1}{2}\Phi_x \\ v_{\xi\eta} &= -\tfrac{1}{4}\Phi_{tt} + \tfrac{1}{4}\Phi_{tx} - \tfrac{1}{4}\Phi_{tx} + \tfrac{1}{4}\Phi_{xx} \\ &= -\tfrac{1}{4}(\Phi_{tt} - \Phi_{xx}). \end{aligned}$$

Thus we want $v_{\xi\eta} = 0$. The general solution to this ODE is given by

$$v(\xi, \eta) = a(\xi) + b(\eta) \implies \Phi(t, x) = a(x + t) + b(x - t)$$

for some $a, b$.

Imposing the initial conditions we get

$$a(x) + b(x) = 0, \qquad a'(x) - b'(x) = \delta_0.$$

This implies

$$a'(x) = -b'(x) = \tfrac{1}{2}\delta_0.$$

So we should take

$$a(x) = \tfrac{1}{2}H(x) + C, \quad b(x) = -[\tfrac{1}{2}H(x) + C]$$

for some $C$, where

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases} \qquad \text{“Heaviside function”.} \qquad \text{(See homework)}$$

Thus

$$\Phi(t,x) = \tfrac{1}{2}[H(x+t) - H(x-t)] = \begin{cases} \frac{1}{2}\text{sign}(t) & -|t| \le x \le |t| \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\text{sign}(t) = \begin{cases} 1 & t > 0 \\ -1 & t < 0. \end{cases}$$

So the solution to (IVP) in dimension $d = 1$ for $t \ge 0$ is given by

$$\begin{aligned} u(t,x) &= f * [\tfrac{1}{2}\delta_0(x+t) + \tfrac{1}{2}\delta_0(x-t)] + g * \Phi(t) \\ &= \tfrac{1}{2}f(x+t) + \tfrac{1}{2}f(x-t) + \tfrac{1}{2}\int_{-t}^{t} g(x-y)\,dy \\ &= \tfrac{1}{2}f(x+t) + \tfrac{1}{2}f(x-t) + \tfrac{1}{2}\int_{x-t}^{x+t} g(y)\,dy. \end{aligned}$$

(This is called **d'Alembert's formula**.)

**Remark 5.1.1.** Note that the value $u(t,x)$ depends on $f$ at $x \pm t$ and $g$ on $y \in (x-t, x+t)$. (An example of finite speed of propagation/Huygens principle.)

### 5.1.2 The Case $d = 3$

We want to solve (IVP). Suppose we have a smooth solution $u(t,x)$.

For $r > 0$ define the **spherical means**

$$\begin{aligned} \tilde{u}(t,x,r) &= \frac{1}{4\pi r^2}\int_{\partial B_r(x)} u(t,y)\,dS(y) \quad \left(z = \frac{y-x}{r},\ dS(z) = r^{-2}\,dS(y)\right) \\ &= \frac{1}{4\pi}\int_{\partial B_1(0)} u(t,x+rz)\,dS(z). \end{aligned}$$

By the continuity of $u$, we have

$$\lim_{r\to 0} \tilde{u}(t,x,r) = u(t,x).$$

We also define

$$\tilde{f}(x,r) = \frac{1}{4\pi r^2}\int_{\partial B_r(x)} f(y)\,dS, \qquad \tilde{g}(x,r) = \frac{1}{4\pi r^2}\int_{\partial B_r(x)} g(y)\,dS.$$

We compute

$$\begin{aligned}\tilde{u}_r &= \frac{1}{4\pi}\int_{\partial B_1(0)} \nabla u(t,x+rz)\cdot z\, dS \\ &= \frac{1}{4\pi r^2}\int_{\partial B_r(x)} \nabla u(t,y)\cdot \tfrac{y-x}{r}\, dS \\ &= \frac{1}{4\pi r^2}\int_{B_r(x)} \Delta u(t,y)\, dy \\ &= \frac{1}{4\pi r^2}\int_{B_r(x)} u_{tt}(t,y)\, dy.\end{aligned}$$

So

$$r^2\tilde{u}_r = \frac{1}{4\pi}\int_{B_r(x)} u_{tt}\, dy.$$

Thus

$$\partial_r(r^2\tilde{u}_r) = \partial_r\left(\frac{1}{4\pi}\int_{B_r(x)} u_{tt}\, dS\right) = \frac{1}{4\pi}\int_{\partial B_r(x)} u_{tt}\, dS = r^2\tfrac{1}{4\pi r^2}\int_{\partial B_r(x)} u_{tt}\, dS = r^2\tilde{u}_{tt}.$$

That is,

$$2r\tilde{u}_r + r^2\tilde{u}_{rr} = r^2\tilde{u}_{tt} \qquad \rightsquigarrow \qquad r\tilde{u}_{tt} - r\tilde{u}_{rr} - 2\tilde{u}_r = 0.$$

Now let $v(t,x,r) = r\tilde{u}(t,x,r)$. Then

$$v_{tt} = r\tilde{u}_{rr}, \qquad v_r = r\tilde{u}_r + \tilde{u}, \qquad v_{rr} = \tilde{u}_{rr} + 2\tilde{u}_r.$$

Then

$$v_{tt} - v_{rr} = r\tilde{u}_{tt} - r\tilde{u}_{rr} - 2\tilde{u}_r = 0.$$

So for fixed $x$, $v$ solves the one dimensional wave equation for $t \in \mathbb{R}$, $r > 0$!

We conclude that $v(t,x,r) := r\tilde{u}(t,x,r)$ solves

$$\begin{cases} v_{tt} - v_{rr} = 0 & \text{for } t\in\mathbb{R},\ r>0, \\ v(0,x,r) = F(x,r) & r>0, \\ v_t(0,x,r) = G(x,r) & r>0, \end{cases}$$

where $F = r\tilde{f}$ and $G = r\tilde{g}$.

Applying the result from your homework, we find that for $0 \le r \le t$:

$$v(t,x,r) = \tfrac{1}{2}[F(x,t+r) - F(x,t-r)] + \tfrac{1}{2}\int_{t-r}^{t+r} G(x,y)\, dy,$$

that is,

$$\tilde{u}(t,x,r) = \tfrac{1}{2r}[F(x,t+r) - F(x,t-r)] + \tfrac{1}{2r}\int_{t-r}^{t+r} G(x,y)\,dy.$$

Then

$$u(t,x) = \lim_{r\to 0} \tilde{u}(t,x,r) = \partial_t F(x,t) + G(x,t). \qquad (*)$$

Note

$$G(x,t) = \frac{t}{4\pi t^2}\int_{\partial B_t(x)} g(y)\,dS(y).$$

$$F(x,t) = \frac{t}{4\pi t^2}\int_{\partial B_t(x)} f(y)\,dS(y) = \frac{t}{4\pi}\int_{\partial B_1(0)} f(x+tz)\,dS(z),$$

so

$$\begin{aligned}
\partial_t F &= \frac{1}{4\pi}\int_{\partial B_1(0)} f(x+tz)\,dS(z) + \frac{t}{4\pi}\int_{\partial B_1(0)} \nabla f(x+tz)\cdot z\,dS(z) \\
&= \frac{1}{4\pi t^2}\int_{\partial B_t(x)} f(y)\,dS(y) + \frac{t}{4\pi t^2}\int_{\partial B_t(x)} \nabla f(y)\,dS(y)\cdot\frac{y-x}{t}\,dS(y) \\
&= \frac{1}{4\pi t^2}\left[\int_{\partial B_t(x)} f(y) + \nabla f(y)\cdot(y-x)\,dS(y)\right].
\end{aligned}$$

Continuing from $(*)$, we find that the solution to (IVP) in dimension $d = 3$ is given by

$$u(t,x) = \frac{1}{4\pi t^2}\int_{\partial B_t(x)} f(y) + \nabla f(y)\cdot(y-x) + tg(y)\,dS(y).$$

(This is called **Kirchhoff's formula**.)

**Remark 5.1.2.** (Strong Huygens principle). The value of $u$ at a point $(t,x)$ depends only the initial data on the sphere $\partial B_t(x)$. (This is true in all odd dimensions.)

### 5.1.3 The Case $d = 2$

We use "Hadamard's method of descent".

Suppose $u(t,x)$ is a solution to (IVP) with $d = 2$. Then let us define $\bar{u} : (0,\infty)\times\mathbb{R}^3 \to \mathbb{R}$ by

$$\bar{u}(t,x_1,x_2,x_3) = u(t,x_1,x_2).$$

Then $\bar{u}$ solves

$$\Box \bar{u} = 0, \quad \bar{u}(0) = \bar{f}, \quad \bar{u}_t(0) = \bar{g},$$

where

$$\bar{f}(x_1, x_2, x_3) = f(x_1, x_2), \quad \bar{g}(x_1, x_2, x_3) = g(x_1, x_2).$$

Let's write $x = (x_1, x_2) \in \mathbb{R}^2$ and $\bar{x} = (x_1, x_2, 0) \in \mathbb{R}^3$.

Using $(*)$ above we write

$$u(t,x) = \bar{u}(t,\bar{x}) = \frac{\partial}{\partial t}\left(\frac{t}{4\pi t^2}\int_{\partial \bar{B}_t(\bar{x})} \bar{f}\, dS(\bar{y})\right) + \frac{t}{4\pi t^2}\int_{\partial \bar{B}_t(\bar{x})} \bar{g}\, dS(\bar{y}). \quad (**)$$

Here $\bar{B}_t(\bar{x})$ is the $3d$-ball around $\bar{x}$. We parametrize $\partial \bar{B}_t(\bar{x})$ by defining

$$h_\pm : B_t(x) \to \partial \bar{B}_t(\bar{x}), \quad h_\pm(y) = (y_1, y_2, \pm z(y)),$$

where

$$z(y) = (t^2 - |x-y|^2)^{1/2}.$$

Indeed

$$|\bar{x} - h_\pm(y)|^2 = |x-y|^2 + t^2 - |x-y|^2 = t^2.$$

Note that

$$\nabla z = -\frac{y-x}{(t^2-|x-y|^2)^{1/2}} \implies \sqrt{1+|\nabla z|^2} = \frac{t}{(t^2-|x-y|^2)^{1/2}}. \quad \text{(check)}$$

Thus

$$\frac{t}{4\pi t^2}\int_{\partial \bar{B}_t(\bar{x})} \bar{g}\, dS(\bar{y}) = \frac{1}{2\pi}\int_{B_t(x)} \frac{g(y)}{(t^2-|x-y|^2)^{1/2}}\, dy.$$

and

$$\frac{t}{4\pi t^2}\int_{\partial \bar{B}_t(\bar{x})} \bar{f}\, dS(\bar{y}) = \frac{t^2}{2\pi t^2}\int_{B_t(x)} \frac{f(y)}{(t^2-|x-y|^2)^{1/2}}\, dy = \frac{t}{2\pi}\int_{B_1(0)} \frac{f(x+tz)}{(1-|z|^2)^{1/2}}\, dz$$

Hence

$$\begin{aligned}\frac{\partial}{\partial t}(...) &= \frac{1}{2\pi}\int_{B_1(0)} \frac{f(x+tz)}{(1-|z|^2)^{1/2}} + \frac{t}{2\pi}\int_{B_1(0)} \frac{\nabla f(x+tz)\cdot z}{(1-|z|^2)^{1/2}}\, dy \\ &= \frac{1}{2\pi t^2}\int_{B_t(x)} \frac{tf(y) + t\nabla f(y)\cdot(y-x)}{(t^2-|x-y|^2)^{1/2}}\, dy\end{aligned}$$

Putting it all together: the solution to (IVP) in dimension $d = 2$ is given by

$$u(t,x) = \frac{1}{2\pi t^2}\int_{B_t(x)} \frac{tf(y) + t\nabla f(y)\cdot(y-x) + t^2 g(y)}{(t^2-|x-y|^2)^{1/2}}\, dy$$

(This is **Poisson's formula**).

**Remark 5.1.3.** (Weak Huygens principle) The value of $u$ at $(t, x)$ depends only on the initial data in the ball $B_t(x)$. This is true for all even dimensions.

So far we have solved the wave equation in dimensions $d \in \{1, 2, 3\}$. For $d = 1$ we can solve the PDE by changing variables $\xi = x+t$ and $\eta = x-t$. For $d = 3$ we reduced the problem to the $1d$ case by taking spherical means.For $d = 2$ we used the method of descent. In general, for odd dimensions one can use spherical means and for even dimensions one can use the method of descent. We do not pursue the general case here.

## 5.2 Energy Methods

There is no maximum principle for the wave equation:

*Example* 5.2.1. Consider

$$\begin{cases} u_{tt} - u_{xx} = 0, \\ u(0, x) = 0, \ u_t(0, x) = \chi_{(-1,1)}(x). \end{cases}$$

The solution is

$$u(t, x) = \tfrac{1}{2} \int_{x-t}^{x+t} \chi_{(-1,1)}(y)\, dy = \tfrac{1}{2}\text{length}\big[(x - t, x + t) \cap (-1, 1)\big].$$

Draw the solution at $t = 1$, $t = 2$, $t = 3$.

We do have uniqueness for the wave equation. To prove it we will use **energy methods**.

Suppose $\Box u = 0$ and $u$ and its derivatives are integrable. We define the **energy** of a solution $u$ by

$$E[u(t)] = \int_{\mathbb{R}^d} \tfrac{1}{2}[u_t(t, x)]^2 + \tfrac{1}{2}[\nabla u(t, x)]^2\, dx.$$

Notice that

$$\tfrac{d}{dt} E[u(t)] = \int_{\mathbb{R}^d} u_{tt} u_t + \nabla u \cdot \nabla u_t\, dx = \int_{\mathbb{R}^d} u_{tt} u_t - \Delta u u_t\, dx = \int_{\mathbb{R}^d} (\Box u) u_t = 0.$$

Thus $E[u(t)] \equiv E[u(0)]$. This is **conservation of energy**.

Let $\Omega \subset \mathbb{R}^d$ be an open, bounded set with smooth boundary $\partial\Omega$.

As before, for $T > 0$ we let $\Omega_T = (0, T] \times \Omega$ and $\Gamma_T = \overline{\Omega_T} \backslash \Omega_T$.

**Proposition 5.2.1** (Uniqueness)**.** *Suppose $u_1, u_2$ are smooth solutions to*

$$\begin{cases} \Box u = F & \text{on } \Omega_T \\ u = f & \text{on } \Gamma_T \\ u_t = g & \text{on } \{t = 0\} \times \Omega \end{cases}$$

*Then $u_1 \equiv u_2$.*

*Proof.* Define $w = u_1 - u_2$. Then $w$ solves

$$\begin{cases} \Box w = 0 & \text{on } \in \Omega_T \\ w = 0 & \text{on } \Gamma_T \\ w_t = 0 & \text{on } \{t = 0\} \times \Omega \end{cases}$$

Define

$$E_\Omega[w(t)] = \int_\Omega \tfrac{1}{2}[w_t(t,x)]^2 + \tfrac{1}{2}[\nabla w(t,x)]^2 \, dx.$$

Then

$$\begin{aligned} \tfrac{d}{dt} E_\Omega[w(t)] &= \int_\Omega w_t w_{tt} + \nabla w \cdot \nabla w_t \, dx \\ &= \int_\Omega w_t w_{tt} - w_t \Delta w \, dx + \int_{\partial\Omega} w_t \nabla w \cdot \mathbf{n} \, dS \\ &= \int_\Omega (\Box w) w_t = 0. \end{aligned}$$

Thus $E_\Omega[w(t)] \equiv E_\Omega[w(0)] = 0$. This implies that $w_t \equiv 0$ and $\nabla w \equiv 0$ on $\Omega_T$. As $w(0,x) = 0$, we conclude $w \equiv 0$. That is, $u_1 \equiv u_2$. □

**Remark 5.2.2.** 'Energy methods' may also be applied to Laplace/heat equations.

**Proposition 5.2.3** (Finite speed of propagation.)**.** *Suppose $\Box u = 0$. Fix $(t_0, x_0) \in (0, \infty) \times \mathbb{R}^d$ and define the cone*

$$C = \{(t,x) : 0 \le t \le t_0, \quad |x - x_0| \le t_0 - t\} \quad .$$

*If $u(0,x) = u_t(0,x) = 0$ for $x \in B_{t_0}(x_0)$, then $u \equiv 0$ for $(t,x) \in C$.*

*Proof.* For dimensions $d \in \{1, 2, 3\}$, we can deduce this from the explicit formulas we derived.

However, energy methods provide a simpler proof (in all dimensions).

For $t \in [0, t_0]$ we define

$$e(t) = \int_{B_{t_0-t}(x_0)} \tfrac{1}{2}[u_t(t,x)]^2 + \tfrac{1}{2}|\nabla u(t,x)|^2 \, dx$$

In your homework, you show

$$\frac{\partial}{\partial r} \int_{B_r(x)} f(y)\, dy = \int_{\partial B_r(x)} f(y)\, dS(y).$$

Similarly,

$$\frac{\partial}{\partial r} \int_{B_r(x)} f(r,y)\, dy = \int_{B_r(x)} \partial_r f(r,y)\, dy + \int_{\partial B_r(x)} f(r,y)\, dS(y).$$

Thus

$$\begin{aligned}
\tfrac{d}{dt}e(t) &= \int_{B_{t_0-t}(x_0)} u_t u_{tt} + \nabla u \cdot \nabla u_t \, dy - \int_{\partial B_{t_0-t}(x_0)} \tfrac{1}{2}u_t^2 + \tfrac{1}{2}|\nabla u|^2 \, dS \\
&= \int_{B_{t_0-t}(x_0)} (\Box u) u_t \, dy + \int_{\partial B_{t_0-t}(x_0)} u_t \nabla u \cdot \mathbf{n} - \tfrac{1}{2}u_t^2 - \tfrac{1}{2}|\nabla u|^2 \, dS \\
&= \int_{\partial B_{t_0-t}(x_0)} u_t \nabla u \cdot \mathbf{n} - (\tfrac{1}{2}u_t^2 + \tfrac{1}{2}|\nabla u|^2)\, dS.
\end{aligned}$$

Now, we have

$$|u_t \nabla u \cdot \mathbf{n}| \leq |u_t|\,|\nabla u| \leq \tfrac{1}{2}u_t^2 + \tfrac{1}{2}|\nabla u|^2,$$

thus continuing from above we see

$$\tfrac{d}{dt}e(t) \leq 0.$$

But notice that $e(t) \geq 0$ by definition, while $e(0) = 0$. Thus $e(t) \equiv 0$ for each $t$.

As before, this implies $u \equiv 0$ on each $B_{t_0-t}(x_0)$. Thus $u \equiv 0$ on $C$. $\square$

From these computations we can also prove the **energy flux identity**.

This states that if a region of space moves at the speed of light then the energy trapped in the (time-dependent) region is non-increasing.

In particular, the energy flux through the boundary is non-negative.

Indeed, we proved that for $\Box u = 0$, we have

$$\tfrac{d}{dt}e(t) = -\int_{\partial B_{t_0-t}(x_0)} \tfrac{1}{2}u_t^2 + \tfrac{1}{2}|\nabla u|^2 - u_t \nabla u \cdot \mathbf{n}\, dS,$$

so that

$$e(t_2) = e(t_1) - \int_{t_1}^{t_2} \int_{\partial B_{t_0-t}(x_0)} \tfrac{1}{2}u_t^2 + \tfrac{1}{2}|\nabla u|^2 - u_t \nabla u \cdot \mathbf{n}\, dS.$$

# Chapter 6

# Separation of Variables

## 6.1 Linear Algebra Review

**Linear algebra review 1:** Let $X$ be a vector space and $A : X \to X$ a linear transformation. A nonzero vector $v \in X$ is an **eigenvector** of $A$ if there exists $\lambda \in \mathbb{C}$ such that $Av = \lambda v$. We call $\lambda$ the **eigenvalue** associated to $v$. $\square$

**Linear algebra review 2:** Let $X$ be a vector space over $\mathbb{C}$.

An **inner product** on $X$ is a function $b : X \times X \to \mathbb{C}$ such that

- $b(v, w) = \overline{b(w, v)}$
- $b(\alpha v + \beta w, z) = \alpha b(v, z) + \beta b(w, z)$
- $b(v, v) \geq 0$
- $b(v, v) = 0 \implies v = 0$

We use the notation $\langle v, w \rangle = b(v, w)$.

We call $v, w$ **orthogonal** if $\langle v, w \rangle = 0$. We write $v \perp w$.

A set $\{v_n\} \subset X$ is called **orthonormal** if $\langle v_n, v_m \rangle = \begin{cases} 1 & n = m \\ 0 & n \neq m. \end{cases}$

A set $\{v_n\} \subset X$ is called an **orthonormal basis** if it is orthonormal and forms a basis.

If $\{v_n\}$ is an orthonormal basis then for any $x \in X$ we have $x = \sum_n \langle x, v_n \rangle v_n$. $\square$

**Linear algebra review 3:**

If $\langle \cdot, \cdot \rangle$ is an inner product on $X$ then we can define the **norm** of a vector $x \in X$ by

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

We have the following important facts:
(i) triangle inequality: $\|x+y\| \le \|x\| + \|y\|$
(ii) Cauchy–Schwarz inequality: $|\langle x, y\rangle| \le \|x\|\|y\|$
(iii) Pythagorean theorem: if $x \perp y$ then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. $\square$

## 6.2 Separation of Variables

We turn to the study of initial value problems with boundary conditions: $(t, x) \in (0, \infty) \times \Omega$.

The problems have three components:
(i) the equation (PDE)
(ii) the initial conditions (IC) $(t = 0)$
(iii) the boundary conditions (BC) $(x \in \partial\Omega)$
We focus on the following boundary conditions:

- Dirichlet: $u(t, x)$ is specified for $x \in \partial\Omega$
- Neumann: $\nabla u(t, x) \cdot \mathbf{n}(x)$ is specified for $x \in \partial\Omega$

We focus on **homogeneous** BCs (that is, $u = 0$ or $\nabla u \cdot \mathbf{n} = 0$).

The idea of separation of variables is to look for solutions of the form $u(t, x) = p(t)q(x)$.

**Example.** Consider the heat equation (with some BCs):

$$u_t = \Delta u \iff p'(t)q(x) = p(t)\Delta q(x) \iff \frac{-p'(t)}{p(t)} = \frac{-\Delta q(x)}{q(x)}.$$

Thus we look for non-trivial $p, q$ so that

$$\frac{-p'(t)}{p(t)} = \lambda = \frac{-\Delta q(x)}{q(x)} \quad \text{for some} \quad \lambda.$$

In particular, we look for $\lambda$ and $q$ such that

$$-\Delta q = \lambda q.$$

This is the **eigenfunction/eigenvalue** problem for $-\Delta$.

However there is an additional twist: the eigenfunction must satisfy the BCs for $u$ to be a solution.

Depending on the BCs you may find different eigenvalues/eigenfunctions.

Now suppose $q_\lambda$ is an eigenfunction of $\Delta$ with eigenvalue $\lambda$ (satisfying the BCs).

Then we get a solution

$$u_\lambda(t,x) = p_\lambda(t)q_\lambda(x)$$

where $p_\lambda(t) = e^{-\lambda t}$. Note $u_\lambda$ satisfies the PDE and BCs.

What about the IC, say $u(0) = f$ ? In general we won't have $f = q_\lambda$ for any $\lambda$.

**Key observation:** $\sum_\lambda c_\lambda u_\lambda$ still solves the PDE and BCs.

Thus **IF** we could write

$$f(x) = \sum_\lambda c_\lambda q_\lambda(x),$$

then we could solve the PDE/BC/IC by writing

$$u(t,x) = \sum_\lambda c_\lambda p_\lambda(t) q_\lambda(x).$$

We are led to the following problems:

**1.** Solve the eigenvalue problem for $-\Delta$ with prescribed BCs.

**2.** Write functions satisfying BCs as a linear combination of eigenfunctions.

If we can solve 1 and 2, then we can use separation of variables to solve the PDE/BC/IC.

We restrict attention to the case $\Omega = [0, L] \subset \mathbb{R}$.

You will address Problem 1 (homework) and find:

(i) (Dirichlet) eigenvalues $\lambda_n = (\frac{n\pi}{L})^2$, eigenfunctions $q_n(x) = \sin(\frac{n\pi}{L}x)$ $(n > 0)$

(ii) (Neumann) eigenvalues are $\lambda_n = (\frac{n\pi}{L})^2$, eigenfunctions $q_n(x) = \cos(\frac{n\pi}{L}x)$ $(n \geq 0)$

## 6.3 Fourier Series

We discuss Problem 2. In fact, we will show:

**Theorem 6.3.1.** *If $F : [-L, L] \to \mathbb{C}$ is periodic (that is, $F(-L) = F(L)$) and smooth, then we can write $F(x)$ in a **Fourier series:***

$$F(x) = \sum_{n=-\infty}^{\infty} c_n e^{\frac{in\pi}{L}x}.$$

In fact, this will imply (see homework):

**Corollary 6.3.2.**

*(i) if $f : (0, L) \to \mathbb{R}$ satisfies homogeneous Dirichlet BCs, then* $f(x) = \sum_{n=1}^{\infty} \tilde{c_n} \sin(\frac{n\pi}{L}x)$

*(ii) if $f : (0, L) \to \mathbb{R}$ satisfies homogeneous Neumann BCs, then* $f(x) = \sum_{n=0}^{\infty} \tilde{c_n} \cos(\frac{n\pi}{L}x)$.

For (i)/(ii), extend $f$ to $(-L, L)$ by odd/even reflection and show that the Fourier expansion reduces to a sine/cosine series.

We turn to the Theorem.

We rephrase the Theorem in these terms. Defining the vector space

$$X = \{F : [-L, L] \to \mathbb{C} \mid F \text{ is periodic and smooth}\},$$

Theorem 1 says: the set $\{e^{\frac{in\pi}{L}x}\}_{n\in\mathbb{Z}}$ forms a basis for $X$.

(In fact we will see that they form an orthonormal basis if we define the right inner product.)

**Notation:** $e_n(x) := e^{\frac{in\pi}{L}x}$.

With

$$X = \{f : [-L, L] \to \mathbb{C} \mid f \text{ is periodic and smooth}\},$$

the Theorem says: the set $\{e_n\}_{n\in\mathbb{Z}}$ forms a basis for $X$.

Our first task is to figure out what the $c_n$ 'should' be.

**Key observation:** $\{e_n\}$ forms an orthonormal set with respect to the inner product

$$\langle f, g\rangle = \tfrac{1}{2L}\int_{-L}^{L} f(x)\overline{g(x)}\,dx, \qquad f, g \in X.$$

Indeed, for $n \neq m$ we have

$$\langle e_n, e_m\rangle = \tfrac{1}{2L}\int_{-L}^{L} e^{\frac{i(n-m)\pi}{L}x}\,dx = \tfrac{1}{2L}\left[\frac{L}{i\pi(n-m)}e^{\frac{i(n-m)\pi}{L}x}\right]_{-L}^{L} = 0,$$

since $e^{i(n-m)\pi} = e^{-i(n-m)\pi}$. Moreover

$$\langle e_n, e_n\rangle = \tfrac{1}{2L}\int_{-L}^{L} 1\,dx = 1.$$

Thus $\{e_n\}$ is an orthonormal set.

So **<u>if</u>** $\{e_n\}$ formed a basis, we would have

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e_n(x) \quad \text{with} \quad c_n = \langle f, e_n\rangle = \tfrac{1}{2L}\int_{-L}^{L} f(x)\overline{e_n}(x)\,dx.$$

Motivated by this computation, for $f \in X$ we define $S_N f$ by

$$S_N f(x) = \sum_{n=-N}^{N} \langle f, e_n \rangle e_n(x).$$

Our goal is to show that $S_N f \to f$ as $N \to \infty$. (**Question:** in what sense?)

Let us first study $S_N f$.

$$\begin{aligned} S_N f(x) = \sum_{n=-N}^{N} \langle f, e_n \rangle e_n(x) &= \tfrac{1}{2L} \sum_{n=-N}^{N} \int_{-L}^{L} f(y) \overline{e_n}(y) e_n(x)\, dy \\ &= \int_{-L}^{L} f(y) \left( \tfrac{1}{2L} \sum_{n=-N}^{N} e_n(x-y) \right) dy = f * D_N(x), \end{aligned}$$

where

$$D_N(x) = \tfrac{1}{2L} \sum_{n=-N}^{N} e_n(x) \quad \text{for} \quad x \in [-L, L]. \quad (\text{'Dirichlet kernel'})$$

If $D_N$ formed approximate identities as $N \to \infty$, we would be done. Unfortunately, they don't. (They become very oscillatory for large $N$.)

Thus we don't get a simple proof that $S_N f(x) \to f(x)$ for each $x$.

However, if we define the **Cesáro means**

$$\sigma_N f(x) = \tfrac{1}{N} \sum_{n=0}^{N-1} S_n f(x),$$

then we find

$$\sigma_N f(x) = f * F_N(x), \qquad F_N(x) = \tfrac{1}{N} \sum_{n=0}^{N-1} D_n(x). \qquad (\text{'Fejér kernel'})$$

It turns out that averaging eliminates the oscillation and $F_N$ form approximate identities as $N \to \infty$.

Thus $\sigma_N f(x) \to f(x)$ as $N \to \infty$ (in fact, uniformly in $x$).

**Conclusion 1.** $S_N f \to f$ uniformly in an averaged sense. □

**Remark 6.3.3.** Note

$$\sigma_N f = \tfrac{1}{N} \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \langle f, e_m \rangle e_m = \sum_{n=-N}^{N} b_n e_n \quad \text{for some } b_n \quad (*)$$

As a consequence of Conclusion 1 and $(*)$, we find: for any $\varepsilon > 0$ there exists $N$ and $\{b_n\}_{n=-N}^{N}$ such that

$$\left| f(x) - \sum_{n=-N}^{N} b_n e_n(x) \right| < \varepsilon \quad \text{for} \quad x \in (-L, L).$$

Our next goal is to show '**mean-square** $(L^2)$ **convergence**' of $S_N f$.
We can define a norm on $X$ via

$$\|f\|^2 = \langle f, f \rangle = \tfrac{1}{2L} \int_{-L}^{L} f(x)\overline{f(x)}\, dx = \tfrac{1}{2L} \int_{-L}^{L} |f(x)|^2\, dx.$$

This is called the $L^2$-norm. We will show:

$$\lim_{N \to \infty} \|f - S_N f\| = 0.$$

This says that $S_N f \to f$ in the $L^2$-norm.

For $f : [-L, L] \to \mathbb{C}$ periodic and smooth we defined

$$S_N f(x) = \sum_{n=-N}^{N} \langle f, e_n \rangle e_n(x),$$

where

$$e_n(x) = e^{\frac{in\pi}{L}x}, \quad \langle f, g \rangle = \tfrac{1}{2L} \int_{-L}^{L} f(x)\overline{g(x)}\, dx.$$

The 'Cesáro means' converge to $f$ uniformly, which implies

For any $\varepsilon > 0$ there exists $M$ and $\{b_n\}_{n=-M}^{M}$ such that

$$\left| f(x) - \sum_{n=-M}^{M} b_n e_n(x) \right| < \varepsilon \quad \text{for} \quad x \in (-L, L). \quad (*)$$

**Goal.**

$$\lim_{N \to \infty} \|f - S_N f\| = 0, \quad \text{where} \quad \|f\| = \langle f, f \rangle^{1/2}.$$

*Proof.* Let $\varepsilon > 0$. We need to find $M$ so that $\|f - S_N f\| \leq \varepsilon$ for $N \geq M$.

**Key facts:** for any $M$

A. $(f - S_M f) \perp e_n$ for any $-M \leq n \leq M$

B. $(f - S_M f) \perp \sum_{n=-M}^{M} c_n e_n$ for any $\{c_n\}$

C. $\|f - S_N f\| \leq \|f - S_M f\|$ for $N > M$.

For (A) we note

$$\langle f - S_M f, e_n\rangle = \langle f, e_n\rangle - \langle \sum_{m=-M}^{M} \langle f, e_m\rangle e_m, e_n\rangle = \langle f, e_n\rangle - \sum_{m=-M}^{M} \langle f, e_m\rangle\langle e_m, e_n\rangle$$
$$= \langle f, e_n\rangle - \langle f, e_n\rangle \underbrace{\langle e_n, e_n\rangle}_{=1} = 0.$$

Now (B) follows from (A).
For (C) we note

$$f - S_M f = f - S_N f + \sum_{M<|n|\leq N} \langle f, e_n\rangle e_n.$$

Using (B) and the Pythagorean theorem we have

$$\|f - S_M f\|^2 = \|f - S_N f\|^2 + \|\sum_{M<|n|\leq N} \langle f, e_n\rangle e_n\|^2 \geq \|f - S_N f\|^2. \qquad \square$$

Now let $\varepsilon > 0$. We choose $M$ and $b_n$ so that $(*)$ holds.
Squaring and integrating, $(*)$ implies

$$\|f - \sum_{n=-M}^{M} b_n e_n\| < \varepsilon.$$

But now we write

$$f - \sum_{n=-M}^{M} b_n e_n = f - S_M f + \sum_{n=-M}^{M} [\langle f, e_n\rangle - b_n]e_n,$$

so that using (B) and the Pythagorean theorem we find

$$\varepsilon^2 > \|f - \sum_{n=-M}^{M} b_n e_n\|^2 = \|f - S_M f\|^2 + \|\sum_{n=-M}^{M} [\langle f, e_n\rangle - b_n]e_n\|^2 \geq \|f - S_M f\|^2$$

That is,

$$\|f - S_M f\| < \varepsilon.$$

Using (C), we conclude $\|f - S_N f\| < \varepsilon$ for all $N \geq M$. $\square$

**Remark 6.3.4.** Using the 'key facts' above we can show

$$\|f\|^2 = \|f - S_N f\|^2 + \sum_{n=-N}^{N} |\langle f, e_n\rangle|^2.$$

Thus

$$\sum_{n=-\infty}^{\infty} |\langle f, e_n\rangle|^2 = \|f\|^2 \quad \text{('Parseval's identity')}.$$

In particular,

$$\langle f, e_n\rangle \to 0 \quad \text{as} \quad n \to \pm\infty \qquad \text{('Riemann–Lebesgue lemma')}$$

Using the Riemann–Lebesgue lemma and a more careful analysis of the Dirichlet kernel, one can show that if $f$ is differentiable at $x_0$ then $S_N f(x_0) \to f(x_0)$ pointwise.

However, **beware:** if $f$ is merely continuous then its Fourier series may diverge at a point. (Beyond the scope of this course.)

*Example* 6.3.1. Consider

$$\begin{cases} u_t - u_{xx} = 0 & (t,x) \in (0,\infty) \times (0,L) \\ u(0,x) = f(x) & x \in [0,L] \\ \partial_x u(t,0) = \partial_x u(t,L) = 0. \end{cases} \qquad (*)$$

Writing $u(t,x) = p(t)q(x)$ leads to the eigenvalue problem

$$\tfrac{-p'(t)}{p(t)} = \lambda = \tfrac{-q''(x)}{q(x)}, \qquad q'(0) = q'(L) = 0.$$

The eigenfunctions are $q_n(x) = \cos(\frac{n\pi}{L}x)$ for $n \geq 0$, with eigenvalues $(\frac{n\pi}{L})^2$.

So we get a family of solutions $u_n(t,x) = e^{-(n\pi/L)^2 t}\cos(\frac{n\pi}{L}x)$, $n \geq 0$.

If we could write

$$f(x) = \sum_{n=0}^{\infty} c_n \cos(\frac{n\pi}{L}x) \qquad (**)$$

then we can solve $(*)$ via

$$u(t,x) = \sum_{n=0}^{\infty} c_n e^{-(n\pi/L)^2 t}\cos(\tfrac{n\pi}{L}x).$$

Extending $f$ to $[-L, L]$ by even reflection and using the Fourier series, one can find

$$c_0 = \tfrac{1}{L}\int_0^L f(x)\,dx, \quad c_n = \tfrac{2}{L}\int_0^L f(x)\cos(\tfrac{n\pi}{L}x)\,dx.$$

(See Homework 5 Problem 5.)

Then for reasonable $f$ the series expansion in $(**)$ will be valid.

# Chapter 7

# The Fourier Transform

## 7.1 The Fourier Transform

The theory of **Fourier series** says that for $f : [-L, L] \to \mathbb{C}$ periodic and smooth we can write $f$ as a linear combination of waves of frequencies $\frac{n}{2L}$,

$$f(x) = \sum_{n=-\infty}^{\infty} \langle f, e_n \rangle e^{\frac{in\pi}{L}x}, \qquad (*)$$

where the **Fourier coefficients** are given by

$$\langle f, e_n \rangle = \tfrac{1}{2L} \int_{-L}^{L} f(x) e^{-\frac{in\pi}{L}x}\, dx.$$

The **Fourier transform** extends these ideas to the case $L \to \infty$: for $f : \mathbb{R} \to \mathbb{C}$ we define $\widehat{f} : \mathbb{R} \to \mathbb{C}$ (formally) by

$$\widehat{f}(\xi) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} f(x) e^{-ix\xi}\, dx.$$

Thus $\widehat{f}(\xi)$ is the 'Fourier coefficient' at a frequency $\xi \in \mathbb{R}$.

**Question.** Can we recover $f$ from $\widehat{f}$? Do we have an analogue of $(*)$?

Suppose $f : \mathbb{R} \to \mathbb{C}$ has compact support, say $f(x) = 0$ for $|x| > M$. Choose $L > M$. Then

$$\langle f, e_n \rangle = \tfrac{1}{2L} \int_{-L}^{L} f(x) e^{-\frac{in\pi}{L}x}\, dx = \tfrac{\pi}{L} \tfrac{1}{2\pi} \int_{\mathbb{R}} f(x) e^{-\frac{in\pi}{L}x}\, dx = \tfrac{1}{(2\pi)^{1/2}} \tfrac{\pi}{L} \widehat{f}(\tfrac{n\pi}{L}).$$

Thus for fixed $x$, we have

$$f(x) = \tfrac{1}{(2\pi)^{1/2}} \sum_{n=-\infty}^{\infty} \tfrac{\pi}{L} \widehat{f}(\tfrac{n\pi}{L}) e^{\frac{in\pi}{L}x}.$$

Let us write $\varepsilon = \frac{\pi}{L}$ and $G(y) = \widehat{f}(y)e^{iyx}$ and send $L \to \infty$ (that is, $\varepsilon \to 0$). Then

$$f(x) = \tfrac{1}{(2\pi)^{1/2}} \sum_{n=-\infty}^{\infty} \varepsilon G(\varepsilon n) \to \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} G(\xi)\, d\xi = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} \widehat{f}(\xi) e^{ix\xi}\, dy.$$

We arrive (formally) at the **Fourier inversion formula**

$$f(x) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} \widehat{f}(\xi) e^{ix\xi}\, d\xi, \qquad \widehat{f}(\xi) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} f(x) e^{-ix\xi}\, dx.$$

We turn now to the details.

**Definition 7.1.1.** (Schwartz space)

$$\mathcal{S}(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}) : x^k f^{(\ell)}(x) \text{ is bounded for all } k, \ell \geq 0.\}$$

If $f \in \mathcal{S}$ then $f$ is absolutely integrable, and so $\widehat{f}$ is a bounded function $\widehat{f} : \mathbb{R} \to \mathbb{C}$. In fact we will show $\widehat{f} \in \mathcal{S}$.

**Lemma 7.1.2.** *Let $f \in \mathcal{S}(\mathbb{R})$.*

- *if $g(x) = f'(x)$ then $\widehat{g}(\xi) = i\xi \widehat{f}(\xi)$*
- *if $g(x) = -ixf(x)$ then $\widehat{g}(\xi) = \frac{d}{d\xi}\widehat{f}(\xi)$*

*Proof.* If $g(x) = f'(x)$, then

$$\begin{aligned}
\widehat{g}(\xi) &= \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} e^{-ix\xi} f'(x)\, dx \\
&= -\tfrac{1}{(2\pi)^{1/2}} \left[ \int_{\mathbb{R}} \tfrac{d}{dx}(e^{-ix\xi}) f(x)\, dx + [f(x)e^{-ix\xi}]_{x=-\infty}^{\infty} \right] \\
&= \tfrac{1}{(2\pi)^{1/2}} i\xi \int_{\mathbb{R}} e^{-ix\xi} f(x)\, dx = i\xi \widehat{f}(\xi).
\end{aligned}$$

If $g(x) = -ixf(x)$, then

$$\widehat{g}(\xi) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} -ixf(x)e^{-ix\xi}\, dx = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} f(x) \tfrac{d}{d\xi}(e^{-ix\xi})\, dx = \tfrac{d}{d\xi}\widehat{f}(\xi). \qquad \square$$

**Remark 7.1.3.** We can now see the connection of the $\mathcal{F}$ to PDE: it interchanges derivatives and multiplication by $x$.

**Proposition.** If $f \in \mathcal{S}$ then $\widehat{f} \in \mathcal{S}$.

*Proof.* Using the lemma, for any $f \in \mathcal{S}$ and any $k, \ell \geq 0$ we have

$$\xi^k \left(\tfrac{d}{d\xi}\right)^\ell \widehat{f}(\xi) = \widehat{g}(\xi), \qquad \text{where} \quad g = (\tfrac{1}{i}\tfrac{d}{dx})^k (-ix)^\ell f \in \mathcal{S}.$$

As $g \in \mathcal{S}$, we have $\widehat{g}$ is bounded. We conclude $\widehat{f} \in \mathcal{S}$. □

Thus if we write $\mathcal{F}f = \widehat{f}$, we have $\mathcal{F} : \mathcal{S} \to \mathcal{S}$. In fact:

**Theorem.** $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is a bijection, and the Fourier inversion formula holds.

**Lemma.** (Multiplication formula) If $f, g \in \mathcal{S}$ then

$$\int_{\mathbb{R}} f(x)\widehat{g}(x)\,dx = \int_{\mathbb{R}} \widehat{f}(y)g(y)\,dy.$$

*Proof.*

$$\begin{aligned}\int_{\mathbb{R}} f(x)\widehat{g}(x)\,dx &= (2\pi)^{-1/2} \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)e^{-ixy}g(y)\,dy\,dx \\ &= (2\pi)^{-1/2} \int_{\mathbb{R}} g(y)\left(\int_{\mathbb{R}} f(x)e^{-ixy}\,dx\right) dy = \int_{\mathbb{R}} \widehat{f}(y)g(y)\,dy.\end{aligned}$$

□

**Lemma.**

- if $g(x) = f(-x)$ then $\widehat{g}(\xi) = \widehat{f}(-\xi)$
- if $g(x) = f(x-h)$ then $\widehat{g}(\xi) = e^{-ih\xi}\widehat{f}(\xi)$
- if $g(x) = f(\lambda x)$ then $\widehat{g}(\xi) = \frac{1}{\lambda}\widehat{f}(\frac{\xi}{\lambda})$

*Proof.*

$$\int_{\mathbb{R}} f(-x)e^{-ix\xi}\,dx = \int_{\mathbb{R}} f(y)e^{iy\xi}\,dx = \int_{\mathbb{R}} f(y)e^{-iy(-\xi)}\,dx. \quad (y = -x)$$

$$\int_{\mathbb{R}} f(x-h)e^{-ix\xi}\,dx = \int_{\mathbb{R}} f(y)e^{-i(y+h)\xi}\,dy = e^{-ih\xi}\int_{\mathbb{R}} f(y)e^{-iy\xi}\,dy. \quad (y = x-h)$$

$$\int_{\mathbb{R}} f(\lambda x)e^{-ix\xi}\,dx = \tfrac{1}{\lambda}\int_{\mathbb{R}} f(y)e^{-iy\xi/\lambda}\,dy. \quad (y = \lambda x)$$

□

**Corollary.** If $g(y) = f(x-y)$ then $\widehat{g}(y) = e^{-ixy}\widehat{f}(-y)$.
Recall:
**Schwartz space:**

$$\mathcal{S}(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}) : x^k f^{(\ell)}(x) \text{ is bounded for all } k, \ell \geq 0.\}$$

**Fourier transform:** $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ defined by

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} f(x) e^{-ix\xi}\, dx.$$

Our first goal is the following

**Theorem.** $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is a bijection and the **Fourier inversion formula** holds:

$$f(x) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} \widehat{f}(\xi) e^{ix\xi}\, d\xi.$$

So far we proved:

$$(i) \qquad \int_{\mathbb{R}} f(x)\widehat{g}(x)\, dx = \int_{\mathbb{R}} \widehat{f}(x) g(x)\, dx.$$

$$(ii) \qquad \text{if } g(y) = f(x-y) \text{ then } \widehat{g}(y) = e^{-ixy}\widehat{f}(-y)$$

$$(iii) \qquad \text{if } g(y) = f(\lambda y) \text{ then } \widehat{g}(y) = \tfrac{1}{\lambda}\widehat{f}(\tfrac{y}{\lambda}).$$

**Lemma.** Let $f(x) = e^{-x^2/2}$. Then $f \in \mathcal{S}$ and $\widehat{f} = f$.

*Proof.* (You check that $f \in \mathcal{S}$.) We have

$$f'(x) = -xf(x) = \tfrac{1}{i}(-ixf(x))$$

Thus

$$i\xi\widehat{f}(\xi) = \tfrac{1}{i}\tfrac{d}{d\xi}\widehat{f}(\xi). \implies \tfrac{d}{d\xi}\widehat{f} = -\xi\widehat{f},$$

so

$$\widehat{f}(\xi) = e^{-\xi^2/2}\widehat{f}(0).$$

Now we note

$$\widehat{f}(0) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} e^{-x^2/2}\, dx = 1.$$

□

We define $K(x) = (2\pi)^{-1/2} e^{-x^2/2}$. For $\varepsilon > 0$ we define $K_\varepsilon(x) = \frac{1}{\varepsilon}K(\frac{x}{\varepsilon})$. Then

- $K_\varepsilon$ form approximate identities as $\varepsilon \to 0$,

- Using (iii): if $G_\varepsilon(x) = K(\varepsilon x)$, then $\widehat{G_\varepsilon} = K_\varepsilon$.

*Proof of Theorem.* We first show the inversion formula.

Let $f \in \mathcal{S}$, $x \in \mathbb{R}$, and $\varepsilon > 0$. Let $g(y) = f(x-y)$.

$$\begin{aligned} f * K_\varepsilon(x) &= \int_{\mathbb{R}} f(x-y)K_\varepsilon(y)\,dy = \int_{\mathbb{R}} g(y)\widehat{G_\varepsilon}(y)\,dy \\ &= \int_{\mathbb{R}} \widehat{g}(y)G_\varepsilon(y)\,dy = \int_{\mathbb{R}} e^{-ixy}\widehat{f}(-y)K(\varepsilon y)\,dy \\ &= \int_{\mathbb{R}} \widehat{f}(y)e^{ixy}K(-\varepsilon y)\,dy. \end{aligned}$$

Now send $\varepsilon \to 0$. Then $f * K_\varepsilon(x) \to f(x)$, while $K(-\varepsilon y) \to K(0) = (2\pi)^{-1/2}$.

Thus

$$f(x) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} \widehat{f}(y)e^{ixy}\,dy.$$

We can now see that $\mathcal{F}$ is a bijection: we define $\mathcal{F}^* : \mathcal{S} \to \mathcal{S}$ by

$$\mathcal{F}^* g(x) = \tfrac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} g(\xi)e^{ix\xi}\,d\xi.$$

The Fourier inversion formula says that $\mathcal{F}^* \circ \mathcal{F} = Id$ on $\mathcal{S}$.

Combining this with the fact that $\mathcal{F}^* f(y) = \mathcal{F} f(-y)$, we can see that $\mathcal{F} \circ \mathcal{F}^* = Id$.

We conclude that $\mathcal{F}^* = \mathcal{F}^{-1}$ and $\mathcal{F}$ is a bijection on $\mathcal{S}$. □

We next prove the **Plancherel theorem** for the Fourier transform.

As before, we introduce the $L^2$ inner product and norm on $\mathcal{S}(\mathbb{R})$ by defining

$$\langle f, g\rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}\,dx \quad \text{for } f, g \in \mathcal{S},$$

$$\|f\| = \sqrt{\langle f, f\rangle} = \left(\int_{\mathbb{R}} |f(x)|^2\,dx\right)^{1/2}.$$

**Theorem 7.1.4** (Plancherel)**.** *For $f, g \in \mathcal{S}$ we have $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$. In particular $\|f\| = \|\widehat{f}\|$.*

**Remark 7.1.5.** This says that $\mathcal{F}$ is a **unitary** transformation.

**Lemma.** $\mathcal{F}(f * g)(\xi) = (2\pi)^{1/2}\widehat{f}(\xi)\widehat{g}(\xi)$.

*Proof.*

$$\begin{aligned}\mathcal{F}(f*g)(\xi) &= (2\pi)^{-1/2}\int_{\mathbb{R}}(f*g)(x)e^{-ix\xi}\,dx\\ &= (2\pi)^{-1/2}\int_{\mathbb{R}}g(y)e^{-iy\xi}\bigg(\int_{\mathbb{R}}f(x-y)e^{-i(x-y)\xi}\,dx\bigg)\,dy\\ &= (2\pi)^{1/2}\widehat{f}(\xi)\widehat{g}(\xi). \quad \square\end{aligned}$$

*Proof of Theorem.* Let $f, g \in \mathcal{S}$. Define $G(x) = \overline{g(-x)}$. Then $\widehat{G}(\xi) = \overline{\widehat{g}(\xi)}$. Then

$$\begin{aligned}\int_{\mathbb{R}} f(x)\overline{g}(x)\,dx = f*G(0) &= (2\pi)^{-1/2}\int_{\mathbb{R}}\mathcal{F}(f*G)(\xi)\,d\xi\\ &= \int_{\mathbb{R}}\widehat{f}(\xi)\widehat{G}(\xi)\,d\xi = \int_{\mathbb{R}}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi. \quad \square\end{aligned}$$

## 7.2 Higher Dimensions

We next discuss the Fourier transform in higher dimensions. Let $d \geq 1$. A **multi-index** $\alpha$ is an element of $\mathbb{N}^d$, where $\mathbb{N} = \{0, 1, 2, \dots\}$.

For $\alpha = (\alpha_1, \dots \alpha_d)$, we define

$$\begin{aligned}|\alpha| &= \sum_{i=1}^{d}|\alpha_i|,\\ x^\alpha &= x_1^{\alpha_1}\cdots x_d^{\alpha_d},\\ \partial^\alpha f &= \frac{\partial^{|\alpha|}f}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}.\end{aligned}$$

**Definition 7.2.1** (Schwartz space)**.** $\mathcal{S}(\mathbb{R}^d) = \{f \in C^\infty(\mathbb{R}^d) : x^\alpha\partial^\beta f \text{ is bounded for all } \alpha, \beta\}$

For $f \in \mathcal{S}$ we define

$$\widehat{f}(\xi) = \tfrac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} f(x)e^{-ix\cdot\xi}\,dx.$$

Then $\mathcal{F}$ is a bijection on $\mathcal{S}$ (in fact unitary) and the Fourier inversion formula holds:

$$f(x) = \tfrac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d}\widehat{f}(\xi)e^{ix\cdot\xi}\,d\xi.$$

We have the **Fourier transform** $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$:

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \tfrac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) e^{-ix\cdot\xi}\, dx.$$

This is a bijection, and $\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is given by

$$\mathcal{F}^{-1}f(x) = \check{f}(x) = \tfrac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) e^{ix\cdot\xi}\, dx.$$

In particular we have the Fourier inversion formula:

$$f(x) = \tfrac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{ix\cdot\xi}\, d\xi.$$

We have also the **Plancherel theorem**

$$\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle, \qquad \langle f, g\rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx$$

and the convolution identities

$$\mathcal{F}(f * g)(\xi) = (2\pi)^{d/2} \widehat{f}(\xi)\widehat{g}(\xi), \qquad \mathcal{F}^{-1}(\widehat{f}\ \widehat{g})(x) = (2\pi)^{-d/2}(f * g)(x).$$

## 7.3 Application to PDEs

Let us finally discuss the application of the Fourier transform to PDEs.

Similar to the $1d$ case we have the following:

- if $g(x) = \partial^\alpha f(x)$ then $\widehat{g}(\xi) = (i\xi)^\alpha \widehat{f}(\xi)$
- if $g(x) = (-ix)^\alpha f(x)$ then $\widehat{g}(\xi) = \partial^\alpha \widehat{f}(\xi)$

Here $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, with $|\alpha| = \sum_{i=1}^d |\alpha_i|$, $\quad x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$, $\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$.

*Example* 7.3.1. Consider $\alpha = (1, 0, \ldots, 0)$. We have

$$\mathcal{F}(\tfrac{\partial f}{\partial x_1})(\xi) = i\xi_1 \widehat{f}(\xi).$$

Similarly with $\alpha = (2, 0, \ldots, 0)$ we get

$$\mathcal{F}(\tfrac{\partial^2 f}{\partial x_1^2})(\xi) = -\xi_1^2 \widehat{f}(\xi).$$

Thus we find

$$\mathcal{F}(-\Delta f)(\xi) = |\xi|^2 \widehat{f}(\xi).$$

*Example* 7.3.2 (Laplace/Poisson equation). Consider the equation $-\Delta u = f$ on $\mathbb{R}^d$. Taking the Fourier transform, we find

$$\mathcal{F}(-\Delta u)(\xi) = \widehat{f}(\xi) \quad \Longleftrightarrow \quad |\xi|^2 \widehat{u}(\xi) = \widehat{f}(\xi) \quad \Longleftrightarrow \quad \widehat{u}(\xi) = \widehat{f}(\xi)\tfrac{1}{|\xi|^2}.$$

Thus the solution is given by

$$u(x) = \mathcal{F}^{-1}[\widehat{f}(\xi)\tfrac{1}{|\xi|^2}](x) = (2\pi)^{-d/2}[f * \mathcal{F}^{-1}(\tfrac{1}{|\xi|^2})](x).$$

Evidently the fundamental solution is given by

$$\Phi(x) = (2\pi)^{-d/2}\mathcal{F}^{-1}(\tfrac{1}{|\xi|^2})(x).$$

However, notice that $\frac{1}{|\xi|^2}$ is not a Schwartz function. Thus to make sense of this we need to extend our theory of the Fourier transform.

*Example* 7.3.3 (Heat equation).

$$\begin{cases} u_t - \Delta u = 0 & (t,x) \in (0,\infty) \times \mathbb{R}^d \\ u(0,x) = f(x) & x \in \mathbb{R}^d. \end{cases}$$

We apply the Fourier transform **in the $x$ variables only**. We find

$$\widehat{u}_t(t,\xi) = \mathcal{F}(\Delta u)(t,\xi) \quad \Longleftrightarrow \quad \widehat{u}_t(t,\xi) = -|\xi|^2 \widehat{u}(t,\xi).$$

For each $\xi$, this is an ODE in $t$ that we can solve:

$$\widehat{u}(t,\xi) = \widehat{u}(0,\xi)e^{-t|\xi|^2} = e^{-t|\xi|^2}\widehat{f}(\xi).$$

Thus

$$u(t,x) = \mathcal{F}^{-1}[\widehat{f}\; e^{-t|\xi|^2}](x) = (2\pi)^{-d/2}[f * \mathcal{F}^{-1}(e^{-t|\xi|^2})](x).$$

Evidently $(2\pi)^{-d/2}\mathcal{F}^{-1}(e^{-t|\xi|^2})(x)$ is the fundamental solution, and hence we must have

$$(2\pi)^{-d/2}\mathcal{F}^{-1}(e^{-t|\xi|^2})(x) = (4\pi t)^{-d/2}e^{-|x|^2/2t}. \qquad (*)$$

In fact, recall that we showed $\mathcal{F}(e^{-x^2/2})(\xi) = e^{-\xi^2/2}$ in $d = 1$.

In the same way we have $\mathcal{F}(e^{-|x|^2/2})(\xi) = e^{-|\xi|^2/2}$ in dimension $d$.

Thus $\mathcal{F}(e^{-|x|^2/2t})(\xi) = t^{d/2}e^{-t|\xi|^2/2}$ (see Homework 6 Problem 5)

We conclude $\mathcal{F}^{-1}(e^{-t|\xi|^2/2})(x) = t^{-d/2}e^{-|x|^2/2t}$, which gives $(*)$.

*Example* 7.3.4 (Wave equation).

$$\begin{cases} u_{tt} - \Delta u = 0 & (t,x) \in (0,\infty) \times \mathbb{R}^d \\ u(0,x) = f(x),\ u_t(0,x) = g(x) & x \in \mathbb{R}^d. \end{cases}$$

Taking the Fourier transform in the $x$ variables:

$$\widehat{u}_{tt}(t,\xi) = -|\xi|^2 \widehat{u}(t,\xi).$$

For each $\xi$ this is an ODE that we can solve:

$$\widehat{u}(t,\xi) = A(\xi)\cos(t|\xi|) + B(\xi)\sin(t|\xi|).$$

Imposing the initial conditions:

$$\widehat{f}(\xi) = \widehat{u}(0,\xi) = A(\xi),$$

while

$$\widehat{u}_t(t,\xi) = -|\xi|A(\xi)\sin(t|\xi|) + |\xi|B(\xi)\cos(t|\xi|),$$

so

$$\widehat{g}(\xi) = \widehat{u}_t(0,\xi) = |\xi|B(\xi).$$

Thus

$$\widehat{u}(t,\xi) = \widehat{f}(\xi)\cos(t|\xi|) + \widehat{g}(\xi)\frac{\sin(t|\xi|)}{|\xi|} = \widehat{f}(\xi)\partial_t\left[\frac{\sin(t|\xi|)}{|\xi|}\right] + \widehat{g}(\xi)\frac{\sin(t|\xi|)}{|\xi|}.$$

Defining

$$W(t,x) = (2\pi)^{-d/2}\mathcal{F}^{-1}\big[\tfrac{\sin(t|\xi|)}{|\xi|}\big](x),$$

we get

$$u(t,x) = [f * \partial_t W(t,\cdot)](x) + [f * W(t,\cdot)](x).$$

Evidently the the fundamental solution is given by $W$.

Once again, however, $\frac{\sin(t|\xi|)}{\xi}$ is not a Schwartz function, and hence to make sense of this we need to extend our theory of the Fourier transform. □

## 7.4 Tempered Distributions

We now extend the theory of the Fourier transform.

**Definition 7.4.1.** A **tempered distribution** is a functional $u : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ that is

(i) linear, i.e. $u(\alpha f + \beta g) = \alpha u(f) + \beta u(g)$

(ii) continuous, i.e. if $f_k \to f$ then $u(f_k) \to u(f)$.

We denote the space of tempered distributions by $\mathcal{S}'(\mathbb{R}^d)$. In this context elements of $\mathcal{S}$ are called test functions. The space $\mathcal{S}'$ is the 'dual space' of $\mathcal{S}$.

As before, we can embed $\mathcal{S}$ into $\mathcal{S}'$: for $u \in \mathcal{S}$ we can define $Tu \in \mathcal{S}'$ by

$$Tu(f) = \int_{\mathbb{R}^d} u(x) f(x)\, dx.$$

The map $T$ is injective (cf. Homework 1 Problem 8). Thus we need not even distinguish between $u$ and $Tu$.

In fact, we can define $Tu \in \mathcal{S}'$ for any 'locally integrable' function $u$, which includes many more functions than just Schwartz functions.

As before, we use integration by parts to motivate the definition of **derivatives of distributions**:

$$\int_{\mathbb{R}^d} \partial^\alpha u(x) f(x)\, dx = (-1)^{|\alpha|} \int_{\mathbb{R}^d} u(x) \partial^\alpha f(x)\, dx \quad \text{for } u, f \in \mathcal{S}.$$

So we **define:**

$$\partial^\alpha u(f) := (-1)^{|\alpha|} u(\partial^\alpha f).$$

Similarly we can define $f * u \in \mathcal{S}'$ for $f \in \mathcal{S}$ and $u \in \mathcal{S}'$ as follows.

First, if $f, u, g \in \mathcal{S}$ then

$$\begin{aligned}
\int_{\mathbb{R}^d} (f * u)(x)\, g(x)\, dx &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} u(y) f(x-y) g(x)\, dy\, dx \\
&= \int_{\mathbb{R}^d} u(y) \left( \int_{\mathbb{R}^d} \tilde{f}(y-x) g(x)\, dx \right) dy \qquad (\tilde{f}(x) := f(-x)) \\
&= \int_{\mathbb{R}^d} u(y) (\tilde{f} * g)(y)\, dy.
\end{aligned}$$

Thus we **define**

$$(f * u)(g) = u(\tilde{f} * g).$$

In fact, we can define $f * u$ as a function for $f \in \mathcal{S}$ and $u \in \mathcal{S}'$ via the formula

$$(f * u)(x) = u(\widetilde{\tau_x} f), \quad \tau_x f(y) = f(y - x).$$

Finally for $u \in \mathcal{S}'$ and a function $f$ we define $fu \in \mathcal{S}'$ by $[fu](g) = u(fg)$.

We now turn to the theory of the Fourier transform on tempered distributions.

Recall that for $u, f \in \mathcal{S}$ we have the multiplication identity:

$$\int_{\mathbb{R}^d} \widehat{u}(x) f(x)\, dx = \int_{\mathbb{R}^d} u(x) \widehat{f}(x)\, dx.$$

Moreover for $f \in \mathcal{S}$ we have $\widehat{f} \in \mathcal{S}$.

This motivates the following:

**Definition 7.4.2.** For $u \in \mathcal{S}'$ we define $\widehat{u} : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ by $\quad \widehat{u}(f) = u(\widehat{f})$.

We also write $\widehat{u} = \mathcal{F}u$.

As $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is linear and continuous, we find that $\widehat{u} \in \mathcal{S}'$. Thus $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$.

Similarly we can define $\mathcal{F}^* : \mathcal{S}' \to \mathcal{S}'$ by $\mathcal{F}^* u(f) = u(\mathcal{F}^{-1} f)$. Then

$$\mathcal{F}^* \widehat{u}(f) = \widehat{u}(\mathcal{F}^{-1} f) = u(\mathcal{F}\mathcal{F}^{-1} f) = u(f) \quad \Longrightarrow \quad \mathcal{F}^* \widehat{u} = u.$$

Similarly $\mathcal{F}\mathcal{F}^*(u) = u$. Thus $\mathcal{F}^* = \mathcal{F}^{-1}$ and $\mathcal{F}$ is a bijection on $\mathcal{S}'$.

Moreover, all of the nice algebraic properties of $\mathcal{F}$ continue to hold on $\mathcal{S}'$.

A. $\widehat{\partial^\alpha u} = (i\xi)^\alpha \widehat{u}$

B. $\widehat{(-ix)^\alpha u} = \partial^\alpha \widehat{u}$

C. $\mathcal{F}(f * u) = (2\pi)^{d/2} \widehat{f}\widehat{u}$

**Moral.** You can safely do 'formal' computations with the Fourier transform even if you are not working with Schwartz functions.

*Proof:* for (A) we have:

$$\begin{aligned} \widehat{\partial^\alpha u}(f) &= \partial^\alpha u(\widehat{f}) = (-1)^{|\alpha|} u(\partial^\alpha \widehat{f}) \\ &= (-1)^{|\alpha|} u(\widehat{(-ix)^\alpha} f) = \widehat{u}((ix)^\alpha f) \\ &= (ix)^\alpha \widehat{u}(f) \end{aligned}$$

For (B) we have:

$$\begin{aligned} \widehat{(-ix)^\alpha} u(f) &= (-ix)^\alpha u(\widehat{f}) = (-1)^{|\alpha|} u((ix)^\alpha \widehat{f}) \\ &= (-1)^{|\alpha|} u(\widehat{\partial^\alpha f}) = (-1)^{-|\alpha|} \widehat{u}(\partial^\alpha f) \\ &= \partial^\alpha \widehat{u}(f). \end{aligned}$$

For (C) we first note that for $f, g \in \mathcal{S}$:

$$\underbrace{\mathcal{F}^{-1}(\widehat{f} * \widehat{g}) = (2\pi)^{d/2} fg}_{\text{as before}} \implies \widehat{f} * \widehat{g} = (2\pi)^{d/2} \mathcal{F}(fg).$$

We also note that for $f \in \mathcal{S}$:

$$\mathcal{F}^{-1}\tilde{f} = \widehat{f} \implies \tilde{f} = \widehat{\widehat{f}}.$$

Thus

$$\begin{aligned} \mathcal{F}(f * u)(g) &= (f * u)(\widehat{g}) = u(\tilde{f} * \widehat{g}) \\ &= u(\widehat{\widehat{f}} * \widehat{g}) = (2\pi)^{d/2} u(\mathcal{F}(\widehat{f} g)) \\ &= (2\pi)^{d/2} \widehat{u}(\widehat{f} g) = (2\pi)^{d/2} [\widehat{f}\widehat{u}](g). \quad \square \end{aligned}$$

*Example* 7.4.1. Consider $\delta_0 \in \mathcal{S}'$ defined by $\delta_0(f) = f(0)$. Then $\widehat{\delta_0} = (2\pi)^{-d/2}$ as a distribution:

$$\widehat{\delta_0}(f) = \delta_0(\widehat{f}) = \widehat{f}(0) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x)\, dx.$$

So the Fourier transform of $\delta_0$ is a constant. More generally, using (A) we see that the polynomials are given by the Fourier transforms of $\delta_0$ and its derivatives.

## 7.5 Duhamel's Principle

We turn to **inhomogeneous** problems.

*Example* 7.5.1 (Inhomogeneous heat equation).

$$\begin{cases} u_t - \Delta u = F & (t, x) \in (0, \infty) \times \mathbb{R}^d \\ u(0, x) = f(x) & x \in \mathbb{R}^d \end{cases}$$

for some $F = F(t, x)$.

We first revisit the homogeneous problem, that is, $F \equiv 0$. We solved this via the Fourier transform:

$$\widehat{u}(t, \xi) = e^{-|\xi|^2 t} \widehat{f}(\xi) \implies u(t) = f * (2\pi)^{-d/2} \mathcal{F}^{-1}(e^{-|\xi|^2 t}) = f * \Phi(t)$$

In particular, if we define the operator $H(t)$ by

$$[H(t)f](x) = [f * \Phi(t,\cdot)](x).$$

then we get that $u(t,x) = [H(t)f](x)$ solves the heat equation with $u(0) = f$.

We call $H(t)$ a **convolution operator**.

Notice that

$$\widehat{H(t)f}(\xi) = e^{-|\xi|^2 t}\widehat{f}(\xi), \quad \text{i.e.} \quad H(t) = \mathcal{F}^{-1}e^{-|\xi|^2 t}\mathcal{F}. \quad (*)$$

Recalling that " $\mathcal{F}(\Delta) = -|\xi|^2$ " we introduce the notation

$$H(t) = e^{t\Delta} = \mathcal{F}^{-1}e^{-|\xi|^2 t}\mathcal{F}. \quad (*).$$

We also call $H(t)$ a **Fourier multiplier operator**.

This is actually great notation, since $\partial_t[e^{t\Delta}f] = \Delta[e^{t\Delta}f]$.

We turn to the case $F \neq 0$. We use 'Duhamel's principle' (i.e. variation of parameters).

We look for a solution of the form $u(t) = e^{t\Delta}[v(t)]$. In this case by the "chain rule"

$$u_t = \Delta e^{t\Delta}v + e^{t\Delta}v_t = \Delta u + e^{t\Delta}v_t.$$

Thus $u_t = \Delta u + F$ provided

$$e^{t\Delta}v_t = F, \quad \text{that is,} \quad v_t = e^{-t\Delta}F$$

and $u(0) = f$ if and only if $v(0) = f$. So we should let

$$v(t,x) = v(0,x) + \int_0^t e^{-s\Delta}F(s,x)\,ds = f(x) + \int_0^t e^{-s\Delta}F(s,x)\,ds.$$

And so

$$\begin{aligned} u(t,x) = e^{t\Delta}v(t,x) &= e^{t\Delta}f(x) + e^{t\Delta}\int_0^t e^{-s\Delta}F(s,x)\,ds \\ &= e^{t\Delta}f(x) + \int_0^t e^{t\Delta}e^{-s\Delta}F(s,x)\,ds. \\ &= \underbrace{e^{t\Delta}f(x)}_{\text{homogeneous part}} + \underbrace{\int_0^t e^{(t-s)\Delta}F(s,x)\,ds.}_{\text{inhomogeneous part}} \end{aligned}$$

In particular

$$e^{t\Delta}f(x) = (4\pi t)^{-d/2}\int_{\mathbb{R}^d} e^{-|x-y|^2/4t}f(y)\,dy,$$

$$\int_0^t e^{(t-s)\Delta} F(s,x)\,ds = \int_0^t \int_{\mathbb{R}^d} (4\pi(t-s))^{-d/2} e^{-|x-y|^2/4(t-s)} F(s,y)\,dy.$$

*Example* 7.5.2 (Inhomogeneous wave equation).

$$\begin{cases} u_{tt} - \Delta u = F & (t,x) \in (0,\infty) \times \mathbb{R}^d \\ u(0,x) = f(x), & x \in \mathbb{R}^d, \\ u_t(0,x) = g(x) & x \in \mathbb{R}^d. \end{cases}$$

For the homogeneous case $F \equiv 0$ we have

$$\partial_t \begin{pmatrix} u \\ u_t \end{pmatrix} = \begin{pmatrix} u_t \\ \Delta u \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix}}_{:=A} \begin{pmatrix} u \\ u_t \end{pmatrix}.$$

Then the solution is given by

$$\begin{pmatrix} u \\ u_t \end{pmatrix} = e^{tA} \begin{pmatrix} f \\ g \end{pmatrix},$$

where

$$e^{tA} := \begin{pmatrix} \mathcal{F}^{-1} & 0 \\ 0 & \mathcal{F}^{-1} \end{pmatrix} e^{tA(\xi)} \begin{pmatrix} \mathcal{F} & 0 \\ 0 & \mathcal{F} \end{pmatrix}, \quad A(\xi) = \begin{pmatrix} 0 & 1 \\ -|\xi|^2 & 0 \end{pmatrix}.$$

We turn to the case $F \neq 0$. We look for a solution of the form

$$\begin{pmatrix} u \\ u_t \end{pmatrix} = e^{tA} \begin{pmatrix} v \\ w \end{pmatrix}$$

Then

$$\begin{aligned} \partial_t \begin{pmatrix} u \\ u_t \end{pmatrix} &= e^{tA} \begin{pmatrix} v_t \\ w_t \end{pmatrix} + A e^{tA} \begin{pmatrix} v \\ w \end{pmatrix} = e^{tA} \begin{pmatrix} v_t \\ w_t \end{pmatrix} + A \begin{pmatrix} u \\ u_t \end{pmatrix} \\ &= e^{tA} \begin{pmatrix} v_t \\ w_t \end{pmatrix} + \begin{pmatrix} u_t \\ \Delta u \end{pmatrix} \underbrace{=}_{\text{want}} \begin{pmatrix} u_t \\ \Delta u + F \end{pmatrix} \end{aligned}$$

Thus we want

$$e^{tA} \begin{pmatrix} v_t \\ w_t \end{pmatrix} = \begin{pmatrix} 0 \\ F \end{pmatrix} \implies \partial_t \begin{pmatrix} v \\ w \end{pmatrix} = e^{-tA} \begin{pmatrix} 0 \\ F \end{pmatrix}$$

Therefore we should take

$$\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} = \begin{pmatrix} v(0) \\ w(0) \end{pmatrix} + \int_0^t e^{-sA} \begin{pmatrix} 0 \\ F(s) \end{pmatrix} ds$$

Imposing the initial conditions and applying $e^{tA}$ we find

$$\begin{pmatrix} u(t) \\ u_t(t) \end{pmatrix} = \underbrace{e^{tA} \begin{pmatrix} f \\ g \end{pmatrix}}_{\text{homogeneous part}} + \underbrace{\int_0^t e^{(t-s)A} \begin{pmatrix} 0 \\ F(s) \end{pmatrix} ds}_{\text{inhomogeneous part}}.$$

One can compute that

$$e^{tA(\xi)} = \begin{pmatrix} \cos(t|\xi|) & |\xi|^{-1} \sin(t|\xi|) \\ -|\xi| \sin(t|\xi|) & \cos(t|\xi|) \end{pmatrix}.$$

Defining $|\nabla| := \mathcal{F}^{-1}|\xi|\mathcal{F}$, we write

$$e^{tA} = \begin{pmatrix} \cos(t|\nabla|) & |\nabla|^{-1} \sin(t|\nabla|) \\ -|\nabla| \sin(t|\nabla|) & \cos(t|\nabla|) \end{pmatrix}.$$

Using these expressions one can write down a formula for the solution to the inhomogeneous wave equation.

# Chapter 8

# The Method of Characteristics

So far we have only discussed **linear** PDE. We introduce the **method of characteristics**, which is capable of treating some nonlinear PDE as well.

## 8.1 The Method of Characteristics

The idea is to try to solve the PDE along certain curves ('characteristics'). This converts the PDE into a system of ODEs.

**Setup.** A general first order PDE in a domain $\Omega \subset \mathbb{R}^d$ can be written

$$F(x, u, \nabla u) = 0 \quad \text{in } \Omega,$$

with some boundary condition

$$u = g \quad \text{on } \Gamma \subset \partial\Omega.$$

**Plan.** For $x \in \Omega$ we look for a curve $x(s) \subset \Omega$ joining some $x_0 \in \Gamma$ to $x$ such that we can compute $z(s) := u(x(s))$.

### 8.1.1 The Linear Case.

First we consider $F$ of the form

$$\begin{cases} F(u, \nabla u) = a(x)u(x) + b(x) \cdot \nabla u(x) = 0 & \text{on } \Omega, \\ u = g & \text{on } \Gamma \subset \partial\Omega. \end{cases}$$

Suppose $u$ is a solution and let $x(s)$ be a curve. Define

$$z(s) = u(x(s)), \quad p(s) = \nabla u(x(s)).$$

Notice that

$$\dot{z}(s) = \nabla u(x(s)) \cdot \dot{x}(s) = p(s) \cdot \dot{x}(s).$$

The PDE along $x(s)$ reads

$$p(s) \cdot b(x(s)) = -a(x(s))z(s).$$

Thus if we let $x(s)$ solve

$$\dot{x}(s) = b(x(s)), \qquad x(0) \in \Gamma \qquad (1)$$

we get

$$\begin{cases} \dot{z}(s) = -a(x(s))z(s), \\ z(0) = g(x(0)). \end{cases} \qquad (2)$$

Equations (1) and (2) are an 'upper triangular' system of ODEs for $x(s), z(s)$.

*Example* 8.1.1. (Transport equation) Consider the equation

$$\begin{cases} u_t(t,y) + c \cdot \nabla u(t,y) = 0, \quad c \in \mathbb{R}^d \\ u(0,y) = g(y) \end{cases}$$

Write

$$\begin{cases} x = (t,y), \quad \nabla u = (u_t, \nabla_y u), \qquad b = (1,c), \\ \Omega = \mathbb{R}^{d+1}_+, \quad \Gamma = \{(0,y) : y \in \mathbb{R}^d\}. \end{cases}$$

The PDE then reads

$$b \cdot \nabla u(x) = 0.$$

That is, $a(x) = 0$ and $b(x) \equiv b$. Then

$$\dot{x}(s) = b \implies x(s) = x(0) + sb, \quad x(0) = (0, y_0) \in \Gamma,$$

$$\dot{z}(s) = 0 \implies z(s) = z(0), \quad \text{that is} \quad u(x(s)) = g(y_0).$$

Now given $x = (t,y)$ we wish to find $x(0) = (0,y_0) \in \Gamma$ and $s$ such that

$$x(s) = x = (t,y), \quad \text{that is} \quad (t,y) = (0,y_0) + s(1,c).$$

Thus we should take

$$s = t \quad \text{and} \quad y = y_0 + ct, \quad \text{i.e.} \quad y_0 = y - ct.$$

Hence

$$u(t,y) = u(x(t)) = g(y_0) = g(y - ct).$$

### 8.1.2 The Quasilinear Case.

A first-order PDE is **quasilinear** if it is 'linear in the highest order term':

$$\begin{cases} F(x,u,\nabla u) = a(x,u(x)) + b(x,u(x))\cdot\nabla u(x) = 0 & \text{on } \Omega \\ u = g & \text{on } \Gamma \subset \partial\Omega. \end{cases}$$

Again we define

$$z(s) = u(x(s)), \quad p(s) = \nabla u(x(s)).$$

Then

$$\dot{z}(s) = p(s)\cdot\dot{x}(s)$$

as before. Then the PDE along $x(s)$ reads

$$p(s)\cdot b(x(s),z(s)) = -a(x(s),z(s)).$$

Thus if we let

$$\dot{x}(s) = b(x(s),z(s)), \quad x(0)\in\Gamma \quad (1)$$

then we get

$$\dot{z}(s) = -a(x(s),z(s)), \quad z(0) = g(x(0)). \quad (2)$$

Equations (1) and (2) are a coupled system of ODEs for $x(s)$ and $z(s)$.

*Example* 8.1.2. Let $\Omega = \{(t,x) : t>0,\ x\in\mathbb{R}\}$ and $\Gamma = \partial\Omega$. Consider

$$\begin{cases} u_t + u_y - u^2 = 0 & (t,y)\in\Omega \\ u(0,y) = g(y) & y\in\mathbb{R}. \end{cases}$$

As before we write $x=(t,y)$. Then $a(u) = -u^2$ and $b = (1,1)$.

We have

$$\dot{x}(s) = b \implies x(s) = x(0) + sb, \quad x(0) = (0,y_0)\in\Gamma.$$

We can also solve

$$\dot{z}(s) = [z(s)]^2, \quad z(0) = g(y_0).$$

We find

$$z(s) = \frac{z(0)}{1-sz(0)} = \frac{g(y_0)}{1-sg(y_0)}.$$

Now given $(t,x)$ we look for $x(0) = (0,y_0)\in\Gamma$ and $s$ such that

$$x(s) = x = (t,y), \quad \text{that is} \quad (t,y) = (0,y_0) + s(1,1).$$

Thus we should take

$$s = t \quad \text{and} \quad y = y_0 + t, \quad \text{i.e.} \quad y_0 = y - t.$$

Thus

$$u(t,y) = u(x(s)) = z(s) = \frac{g(y_0)}{1 - sg(y_0)} = \frac{g(y-t)}{1 - tg(y-t)}.$$

(Note that this only works as long as $tg(y-t) \neq 1$.

It is possible to proceed in the 'fully nonlinear' case but we will not pursue it here. In this case one needs to derive an ODE for $p(s)$ as well and work with the full system of 'characteristic ODEs'.

## 8.2 Scalar Conservation Laws

We continue to discuss the method of characteristics in the setting of "scalar conservation laws". These are PDE of the form

$$\begin{cases} u_t + [F(u)]_y = 0 & (t,y) \in (0,\infty) \times \mathbb{R} \\ u(0,y) = g(y) & y \in \mathbb{R}. \end{cases} \tag{1}$$

The name refers to the fact that solutions formally satisfy $\partial_t \int u(t,y)\, dy = 0$.

In particular we discuss some potential issues that arise when applying the method of characteristics.

We consider **Burger's equation**:

$$\begin{cases} u_t + \frac{1}{2}(u^2)_y = 0 & (t,y) \in (0,\infty) \times \mathbb{R} \\ u(0,y) = g(y) & y \in \mathbb{R}. \end{cases}$$

This equation appears as a simplified model in fluid mechanics.

We rewrite the equation as

$$0 = u_t + uu_y = (1,u) \cdot (u_t, u_y).$$

We let $x = (t,y)$ and $\nabla u = (u_t, u_y)$. We also write $\Omega = (0,\infty) \times \mathbb{R}$ and $\Gamma = \{t = 0\} \times \mathbb{R}$.

This is a quasilinear PDE:

$$\underbrace{(1, u(x))}_{b(u) \text{ from above}} \cdot \nabla u = 0.$$

We define

$$z(s) = u(x(s)), \quad p(s) = \nabla u(x(s))$$

and need to solve the characteristic ODEs

$$\dot{x}(s) = (1, z(s)), \quad x(0) \in \Gamma \quad \text{and} \quad \dot{z}(s) = 0, \quad z(0) = g(x(0)).$$

We find:

$$z(s) = z(0) = g(x(0)),$$
$$\dot{x}(s) = (1, g(x(0))) \implies x(s) = x(0) + s(1, g(x(0))).$$

Given $x = (t, y)$ we look for $x(0) = (0, y_0) \in \Gamma$ and $s$ such that

$$x(s) = (t, y), \quad \text{that is} \quad (t, y) = (0, y_0) + s(1, g(y_0)).$$

Thus we should take $s = t$, and we need to find $y_0$ such that

$$y = y_0 + tg(y_0).$$

In particular we will have

$$u(t, y) = g(y_0), \quad \text{provided} \quad y = y_0 + tg(y_0). \qquad (*)$$

The initial value $g(y_0)$ is carried along the characteristic $(t, y_0 + tg(y_0))$. (Draw picture.)

*Example* 8.2.1 (Formation of shocks). Let

$$g(y) = \begin{cases} 1 & y \le 0 \\ 1 - y & 0 \le y \le 1 \\ 0 & y \ge 1. \end{cases}$$

(Draw picture, $y$ versus $g(y)$ and $t$ versus $y$.)

There is a problem at $t = 1$, since the characteristics cross.

Until $t = 1$ we can solve as follows:

$$u(t, y) = \begin{cases} 1 & y < t \\ \frac{1-y}{1-t} & t \le y \le 1 \\ 0 & y > 1 \end{cases} \qquad \text{(check!)}$$

Notice that the characteristics meet along the curve $t = 2y - 1$, that is, $y(t) = \frac{1}{2}(t + 1)$. This curve is called a **shock**.

To extend the solution past $t = 1$ we can define

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

This is a discontinuous solution, but there is a sense in which it is the most 'physical' solution.

*Example* 8.2.2 (Rarefaction waves). Let

$$g(y) = \begin{cases} 0 & y < 0 \\ 1 & y \geq 0 \end{cases}$$

In this case we find (draw picture):

$$u(t,y) = \begin{cases} 0 & y < 0, \\ 1 & y > t \end{cases}$$

Note that the characteristics give information for $0 < y < t$.

We could try defining a shock:

$$u(t,y) = \begin{cases} 0 & y < t/2, \\ 1 & y > t/2. \end{cases}$$

However, there is a way that we can define a **continuous** solution, namely

$$u(t,y) = \begin{cases} 0 & y < 0, \\ 1 & y > t \\ \frac{y}{t} & 0 < y < t \end{cases} \qquad \text{(check!)}$$

This is called a **rarefaction wave**, and it is a 'better' solution than the shock solution.

Thus we have two potential problems.

**1.** crossing characteristics

**2.** incomplete characteristics

In the first case a 'shock' forms and we will have discontinuous solutions. What conditions should be satisfied along the shock?

In the second case we have a lack of uniqueness. How should we define the solution?

**Definition 8.2.1.** A bounded function $u : (0,\infty) \times \mathbb{R} \to \mathbb{R}$ is an **integral solution** to (1) if

$$\int_0^\infty \int_{\mathbb{R}} u\varphi_t + F(u)\varphi_y \, dy \, dt + \int_{\mathbb{R}} g(y)\varphi(0,y) \, dy = 0$$

for all $\varphi \in C_c^\infty([0,\infty) \times \mathbb{R})$.

Any smooth bounded solution is an integral solution (homework), but this definition allows for solutions that are not necessarily smooth.

If $u$ is a **smooth** integral solution on $\Omega \subset (0,\infty) \times \mathbb{R}$ then $u$ is an actual solution to (1) on $\Omega$.

We continue studying scalar conservation laws and discuss some conditions regarding 'shock' solutions.

### 8.2.1 Rankine–Hugoniot Condition for Shocks.

Suppose $u$ is an integral solution to

$$\begin{cases} u_t + [F(u)]_y = 0 & (t,y) \in (0,\infty)\times\mathbb{R} \\ u(0,y) = g(y) & y \in \mathbb{R}. \end{cases} \tag{1}$$

That is, $u$ is bounded and

$$\int_0^\infty \int_{\mathbb{R}} u\varphi_t + F(u)\varphi_y \, dy\, dt + \int_{\mathbb{R}} g(y)\varphi(0,y)\, dy = 0$$

for all $\varphi \in C_c^\infty([0,\infty)\times\mathbb{R})$.

Suppose $\Omega \subset (0,\infty)\times\mathbb{R}$ and

$$C = \{(t,y) : y = s(t)\}$$

is a curve (maybe a 'shock') that divides $\Omega$ into $\Omega_\ell$ and $\Omega_r$.

Suppose $u$ is smooth on $\Omega_\ell$ and $\Omega_r$ (not necessarily on $C$).

In particular, $u$ solves (1) pointwise on $\Omega_\ell$ and $\Omega_r$.

Denote the limits of $u$ on $C$ from the left/right by $u_\ell, u_r$.

Let $\varphi \in C_c^\infty(\Omega)$. As $u$ is an integral solution, we have

$$0 = \iint_{\Omega_\ell} u\varphi_t + F(u)\varphi_y \, dy\, dt + \iint_{\Omega_r} u\varphi_t + F(u)\varphi_y \, dy. \quad (*)$$

We will now integrate by parts — the boundary terms will tell us what is happening on $C$.

We need to compute the normal vector $\mathbf{n} = (n_t, n_y)$ on $C$ (pointing from $\Omega_\ell$ into $\Omega_r$, say).

As the tangent vector $\mathbf{T} = (T_t, T_y)$ is given by $(1, \dot{s}(t))$, we have that

$$\mathbf{n} = (n_t, n_y) = \frac{1}{\sqrt{1+[\dot{s}(t)]^2}}(-\dot{s(t)}, 1).$$

Now we integrate by parts:

$$\iint_{\Omega_\ell} u\varphi_t + F(u)\varphi_y \, dy\, dt = \underbrace{-\iint_{\Omega_\ell} u_t\varphi + [F(u)]_y\varphi \, dy\, dt}_{=0} + \int_C u_\ell\varphi\, n_t + F(u_\ell)\varphi\, n_y \, dS$$

Similarly

$$\iint_{\Omega_\ell} u\varphi_t + F(u)\varphi_y \, dy\, dt = -\int_C u_r\varphi\, n_t + F(u_r)\varphi\, n_y \, dS.$$

The minus sign comes from the definition of $\mathbf{n}$.

Using $(*)$ we find

$$\int_C [u_\ell n_t + F(u_\ell)n_y]\varphi \, dS = \int_C [u_r n_t + F(u_r)n_y]\varphi \, dS.$$

Since this holds for any $\varphi \in C_c^\infty(\Omega)$ we conclude

$$u_\ell n_t + F(u_\ell)n_y = u_r n_t + F(u_r)n_y \quad \text{on} \quad C,$$

that is:

$$[F(u_\ell) - F(u_r)]n_y + (u_\ell - u_r)n_t = 0 \quad \text{on} \quad C.$$

Recalling $(n_t, n_y) \propto (-\dot{s}, 1)$ we get

$$F(u_\ell) - F(u_r) = \dot{s}(u_\ell - u_r) \quad \text{on} \quad C = \{(t, y) : y = s(t)\}.$$

This is the **Rankine–Hugoniot condition**.

### 8.2.2 Entropy Solutions

We know that as $t$ increases, we may encounter shocks/discontinuities. We would like to impose a condition that guarantees that we don't encounter shocks if $t$ **decreases**. The **entropy condition** is the following: if $C = \{(t, y) : y = s(t)\}$ is a curve of discontinuities along which $u$ has left and right limits $u_\ell, u_r$ then

$$F'(u_\ell) > \dot{s} > F'(u_r).$$

To make sense of this, we notice that the characteristic ODEs for

$$u_t + [F(u)]_y = u_t + F'(u)u_y = (1, F'(u)) \cdot (u_t, u_y) = 0 \quad \text{are}$$

$$\dot{x}(s) = (1, F'(z(s))), \quad x(0) = (0, y_0), \quad \dot{z}(s) = 0$$

So

$$x(s) = (s, y_0 + sF'(g(y_0))), \quad u(x(s)) \equiv g(y_0)$$
$$\implies \dot{x}(s) = (1, F'(g(y_0))),$$

The speed of the characteristic hitting $C$ from the left is $F'(u_\ell)$. Similarly $F'(u_r)$ from the right.

Thus the picture is as follows, which rules out the 'non-physical' case. (Draw picture.)

**Exercise:** check whether or not the Rankine–Hugoniot and entropy conditions are satisfied for the solutions to Burger's equation we discussed last lecture.

*Example* 8.2.3. Previously we solved Burger's equation $u_t + \frac{1}{2}(u^2)_y = 0$ with

$$u(0,y) = g(y) = \begin{cases} 1 & y \le 0 \\ 1-y & 0 \le y < 1 \\ 0 & y \ge 1 \end{cases}$$

We found that a shock formed at $t = 1$, and for $t > 1$ defined

$$u(t,y) = \begin{cases} 1 & y < (t+1)/2 \\ 0 & y > (t+1)/2. \end{cases}$$

Note $\dot{s} = \frac{1}{2}$, $u_\ell = 1$ and $u_r = 0$, and $F(u_\ell) - F(u_r) = \frac{1}{2}$. Thus the Rankine–Hugoniot condition is satisfied.

**Definition 8.2.2.** An integral solution satisfying the Rankine–Hugoniot and entropy conditions is called a **entropy solution.**

It turns out that (under reasonable assumptions on $F$) one can show existence/uniqueness of entropy solutions.

To pursue this in detail requires a study of "Hamilton–Jacobi equations". (**Buzzwords:** calculus of variations*, Lagrangian, Euler–Lagrange equations*, Hamiltonian, Legendre transform, Hopf–Lax formula, Lax–Oleiknik formula)

We will not pursue this direction, but see your homework for the computation of some solutions in the case of Burger's equation.

# Chapter 9

# Calculus of Variations

We begin a brief study of the 'calculus of variations'. The idea is that some PDE arise naturally when considering certain 'optimization' problems. In particular, solving a PDE may be equivalent to minimizing an appropriate 'functional'. This provides a different approach for finding solutions to ODEs/PDEs.

## 9.1 Several Examples

*Example* 9.1.1. What is the shortest (smooth) curve joining $(0,0)$ to $(1,1)$ in $\mathbb{R}^2$ ?

To pose this problem more precisely, we define

$$\mathcal{A} = \{v \in C^\infty([0,1]) : v(0) = 0,\ v(1) = 1\},$$

$$\ell[v] = \int_0^1 \sqrt{1 + (v')^2}\, dt \qquad \text{(arc-length)}.$$

**Goal.** Find necessary and sufficient conditions for $u \in \mathcal{A}$ to satisfy

$$\ell[u] = \min_{v \in \mathcal{A}} \ell[v]. \quad (*)$$

*Theorem* 9.1.1. *If $u \in \mathcal{A}$ satisfies $(*)$, then $u'' = 0$. (Hence $u$ is a line, as expected!)*

*Remark* 9.1.2.

1. The equation $u'' = 0$ is called the **Euler–Lagrange equation** for $(*)$.
2. With the theorem, have we attained the goal?

*Proof of Theorem.* Suppose that $u \in \mathcal{A}$ satisfies $(*)$.

Take any $\varphi \in C_c^\infty((0,1))$. In particular, $\varphi(0) = \varphi(1) = 0$. Consider the function

$$i(\varepsilon) = \ell[u + \varepsilon\varphi].$$

Note that $u + \varepsilon\varphi \in \mathcal{A}$ for each $\varepsilon$.

Thus by $(*)$ we have that $i$ has a minimum at $\varepsilon = 0$. In particular $i'(0) = 0$.

We now compute:

$$i(\varepsilon) = \int_0^1 \sqrt{1 + (u' + \varepsilon\varphi')^2}\, dt$$

$$i'(\varepsilon) = \int_0^1 \tfrac{(u'+\varepsilon\varphi')\varphi'}{\sqrt{1+(u'+\varepsilon\varphi')^2}}\, dt$$

$$0 = i'(0) = \int_0^1 \frac{u'\varphi'}{\sqrt{1+(u')^2}}\, dt = -\int_0^1 \left(\frac{u'}{\sqrt{1+(u')^2}}\right)' \varphi\, dt + \underbrace{\frac{u'\varphi}{\sqrt{1+(u')^2}}\bigg|_{t=0}^{1}}_{\varphi(0)=\varphi(1)=0}$$

Since this holds for all $\varphi \in C_c^\infty((0,1))$ we conclude that

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

Since $1 + (u')^2 > 0$ we conclude $u'' = 0$. □

*Example* 9.1.2. Let $\Omega \subset \mathbb{R}^2$. What is the 'flattest membrane' over $\Omega$ with prescribed boundary conditions on $\partial\Omega$ ? That is, given $g$ what is the 'flattest' $v : \overline{\Omega} \to \mathbb{R}$ such that $v = g$ on $\partial\Omega$ ?

The deviation of $v$ from being flat is measured by $\nabla v$. Inspired by this, we introduce the following **'Dirichlet' energy functional**

$$E[v] = \tfrac{1}{2}\int_\Omega |\nabla v|^2\, dx,$$

$$\text{where} \quad v \in \mathcal{A} = \{v \in C^\infty(\overline{\Omega}) : v = g \text{ on } \partial\Omega\}.$$

**Goal.** Find necessary and sufficient conditions for $u \in \mathcal{A}$ to satisfy

$$E[u] = \min_{v\in\mathcal{A}} E[v]. \quad (*)$$

**Theorem.** (Dirichlet's principle) A function $u \in \mathcal{A}$ satisfies $(*)$ if and only if $u$ solves

$$\begin{cases} -\Delta u = 0 & \text{on } \Omega \\ u = g & \text{on } \partial\Omega. \end{cases} \qquad (**)$$

*Remark* 9.1.3. Laplace's equation is the 'Euler–Lagrange' equation for $(*)$.

*Proof of Theorem.* Suppose $u$ solves $(**)$. Let $v \in \mathcal{A}$. Then

$$0 = \int_\Omega \Delta u(u-v)\,dx = -\int_\Omega \nabla u \cdot \nabla(u-v)\,dx + \underbrace{\int_{\partial\Omega} (u-v)\nabla u \cdot \mathbf{n}\,dS}_{u=v=g \text{ on } \partial\Omega}.$$

Thus

$$\int_\Omega |\nabla u|^2\,dx = \int_\Omega \nabla u \cdot \nabla v\,dx \le \int_\Omega \tfrac{1}{2}|\nabla u|^2 + \tfrac{1}{2}|\nabla v|^2\,dx \quad (|a\cdot b| \le |a|\,|b| \le \tfrac{1}{2}|a|^2 + \tfrac{1}{2}|b|^2)$$

and so

$$\tfrac{1}{2}\int_\Omega |\nabla u|^2\,dx \le \tfrac{1}{2}\int_\Omega |\nabla v|^2\,dx, \qquad \text{that is,} \qquad E[u] \le E[v].$$

Thus $u$ satisfies $(*)$.

Next, suppose $u \in \mathcal{A}$ satisfies $(*)$. We wish to show that $-\Delta u = 0$ in $\Omega$.

As $u$ is smooth it suffices to show that for any $\varphi \in C_c^\infty(\Omega)$, we have

$$-\int_\Omega \Delta u(x)\,\varphi(x)\,dx = 0.$$

Let $\varphi \in C_c^\infty(\Omega)$. In particular, $\varphi = 0$ on $\partial\Omega$. Consider the function

$$i(\varepsilon) = E[u + \varepsilon\varphi].$$

Note that $u + \varepsilon\varphi \in \mathcal{A}$ for each $\varepsilon$.

By $(*)$ we have that $i$ has a minimum at $\varepsilon = 0$. In particular $i'(0) = 0$.

We compute:

$$i(\varepsilon) = \tfrac{1}{2}\int_\Omega |\nabla u|^2 + 2\varepsilon\nabla u\cdot\nabla\varphi + \varepsilon^2|\nabla\varphi|^2\,dx,$$

$$i'(\varepsilon) = \int_\Omega \nabla u\cdot\nabla\varphi + \varepsilon|\nabla\varphi|^2\,dx$$

$$0 = i'(0) = \int_\Omega \nabla u\cdot\nabla\varphi\,dx = -\int_\Omega \Delta u\,\varphi\,dx + \underbrace{\int_{\partial\Omega} \nabla u\,\varphi\cdot\mathbf{n}\,dS}_{\varphi=0 \text{ on } \partial\Omega} = -\int_\Omega \Delta u\,\varphi\,dx.$$

□

**Remark 9.1.4.** The dimension $d = 2$ played no role in the arguments above. It simply allowed for a convenient physical motivation.

One can show similarly that Poisson's equation $-\Delta u = f$ is the Euler–Lagrange equation for

$$\tilde{E}[v] = \int_\Omega \tfrac{1}{2}|\nabla v|^2 - vf\,dx.$$

(See homework.)

*Example* 9.1.3. We consider a 'constrained' minimization problem for the Dirichlet energy functional. In particular we let $\Omega \subset \mathbb{R}^d$,

$$\mathcal{A} = \{v \in C^\infty(\overline{\Omega}) : v = 0 \text{ on } \partial\Omega\}, \quad E[v] = \tfrac{1}{2}\int_\Omega |\nabla v|^2\,dx,$$

and we wish to solve the problem:

$$\begin{aligned} &\text{minimize :} \quad E[v] \quad \text{over} \quad v \in \mathcal{A} \\ &\text{subject to :} \quad \underbrace{\int_\Omega |v|^2\,dx}_{:=\|v\|^2} = 1. \end{aligned} \tag{$*$}$$

**Theorem.** If $u$ is a minimizer of $(*)$, then $u$ solves

$$-\Delta u = \lambda u, \quad \lambda = \|\nabla u\|^2 = 2E[u].$$

That is, $u$ is an eigenfunction of $-\Delta$ with (lowest) eigenvalue $\lambda$. (We call $u$ a **ground state** eigenfunction.)

*Proof.* Suppose $u$ is a minimizer. As before, we let $\varphi \in C_c^\infty(\Omega)$ and consider $u + \varepsilon\varphi$.

We would like to say $E[u + \varepsilon\varphi]$ has a minimum at $\varepsilon = 0$, but we don't know whether $\|u + \varepsilon\varphi\| = 1$.

To remedy this we instead define

$$i(\varepsilon) = E\big[\tfrac{u+\varepsilon\varphi}{\|u+\varepsilon\varphi\|}\big] = \frac{\frac{1}{2}\int_\Omega |\nabla(u+\varepsilon\varphi)|^2\,dx}{\int_\Omega |u+\varepsilon\varphi|^2\,dx} = \frac{\frac{1}{2}\|\nabla(u+\varepsilon\varphi)\|_2^2}{\|u+\varepsilon\varphi\|^2}$$

and it follows that $i'(0) = 0$.

A computation (quotient rule...) shows

$$i'(\varepsilon) = \frac{\|u+\varepsilon\varphi\|_2^2\big[\int_\Omega \nabla u\cdot\nabla\varphi + \varepsilon|\nabla\varphi|^2\,dx\big] - \|\nabla(u+\varepsilon\varphi)\|^2\big[\int_\Omega u\varphi + \varepsilon|\varphi|^2\,dx\big]}{\|u+\varepsilon\varphi\|^4}$$

$$\implies \quad 0 = i'(0) = \int_\Omega \nabla u \cdot \nabla \varphi \, dx - \|\nabla u\|^2 \int_\Omega u\varphi \, dx = -\int_\Omega \Delta u \varphi \, dx - \|\nabla u\|^2 \int_\Omega u\varphi \, dx.$$

As this holds for every $\varphi \in C_c^\infty(\Omega)$, we conclude

$$-\Delta u = \lambda u, \quad \lambda = \|\nabla u\|_2^2 = 2E[u].$$

□

**Remark 9.1.5.** 1. The eigenvalue problem for $-\Delta$ already arose in the setting of separation of variables.

2. From Homework 5 you know that all eigenvalues of $-\Delta$ (with homogeneous Dirichlet BCs) are real. In fact, they are **positive**. To see this, suppose $u$ is a (real-valued) eigenfunction with eigenvalue $\lambda \in \mathbb{R}$. Then

$$\lambda u = -\Delta u \implies \lambda \int_\Omega u^2 \, dx = -\int_\Omega \Delta u \, u \, dx = \int_\Omega |\nabla u|^2 \, dx \implies \lambda = \frac{\int_\Omega |\nabla u|^2 \, dx}{\int_\Omega u^2 \, dx} > 0. \quad \square$$

So far we have characterized the lowest eigenvalue $\lambda_1 > 0$. One can show that there are at most countably many eigenvalues $\lambda_1 \le \lambda_2 \le \cdots$ and $\lambda_k \to \infty$ (beyond the scope of this course)

How can we characterize the other eigenvalues?

We 'project away' the subspace spanned by a ground state eigenfunction $u_1$ and solve the resulting constrained minimization problem.

Let $u_1$ be a ground state eigenfunction, so that $-\Delta u_1 = \lambda_1 u_1$ and $\|u_1\| = 1$.

Recall the notation

$$\langle f, g \rangle = \int_\Omega f \, g \, dx, \quad \|f\| = \sqrt{\langle f, f \rangle}.$$

**Example 4.** Consider the problem

$$\begin{aligned} &\text{minimize :} \quad E[v] \quad \text{over} \quad v \in \mathcal{A} \\ &\text{subject to :} \quad \|v\| = 1, \quad \langle v, u_1 \rangle = 0. \qquad (**) \end{aligned}$$

**Theorem.** If $u$ is a minimizer of $(**)$ then

$$-\Delta u = \lambda u, \quad \lambda = \|\nabla u\|^2 = 2E[u] \ge \lambda_1.$$

*Proof.* Suppose $u$ is a minimizer.

Let $\varphi \in C_c^\infty(\Omega)$ and define $\tilde{\varphi} = \varphi - \langle \varphi, u_1 \rangle u_1$. (This guarantees that $\langle \tilde{\varphi}, u_1 \rangle = 0$.)

As $u$ is a minimizer, we have that the function

$$i(\varepsilon) = E\big[\tfrac{u+\varepsilon\tilde{\varphi}}{\|u+\varepsilon\tilde{\varphi}\|}\big] = \frac{\frac{1}{2}\int_\Omega |\nabla(u+\varepsilon\tilde{\varphi})|^2\,dx}{\int_\Omega |u+\varepsilon\tilde{\varphi}|^2\,dx}$$

has a minimum at $\varepsilon = 0$. We compute exactly as in Example 3 to discover

$$\int_\Omega \nabla u\cdot\nabla\tilde{\varphi}\,dx = \|\nabla u\|^2\int_\Omega u\tilde{\varphi}\,dx.$$

We compute:

$$\begin{aligned}\int_\Omega \nabla u\cdot\nabla\tilde{\varphi}\,dx &= \int_\Omega \nabla u\cdot\nabla\varphi\,dx - \langle\varphi,u_1\rangle\int_\Omega \nabla u\cdot\nabla u_1\,dx\\ &= \int_\Omega -\Delta u\varphi\,dx + \langle\varphi,u_1\rangle\int_\Omega u\Delta u_1\,dx\\ &= \int_\Omega -\Delta u\varphi\,dx - \underbrace{\lambda_1\langle\varphi,u_1\rangle\int_\Omega uu_1\,dx}_{=0}\end{aligned}$$

Similarly

$$\int_\Omega u\tilde{\varphi}\,dx = \int_\Omega u\varphi\,dx - \underbrace{\langle\varphi,u_1\rangle\int_\Omega uu_1\,dx}_{=0}.$$

So once again we have

$$\int_\Omega -\Delta u\varphi\,dx = \|\nabla u\|^2\int_\Omega u\varphi\,dx$$

for all $\varphi\in C_c^\infty(\Omega)$, which implies

$$-\Delta u = \lambda u,\quad \lambda = \|\nabla u\|^2 = 2E[u].$$

Moreover we must have $\lambda\geq\lambda_1$, since we are minimizing the same energy functional over a smaller set of functions. □

**Remark 9.1.6.** 1. Proceeding in this way we can characterize higher eigenvalues.

2. We have left completely open the problem of proving the **existence** of minimizers.

3. Another important issue we won't discuss: 'completeness' of the eigenfunctions.

# Chapter 10

# Numerical Methods

We begin a **very brief** introduction into some numerical methods.

## 10.1 Finite Difference Schemes

We next discuss **finite difference schemes**. For example, consider a PDE for $(t, x) \in [0, \infty) \times [0, L]$. In order to use the computer, we need to discretize the domain.

For $J \in \mathbb{N}$ we can approximate $[0, L]$ by $\{x_j\}_{j=0}^{J}$, where

$$x_j = j\Delta x, \qquad J\Delta x = L.$$

Similarly given $N \in \mathbb{N}$ and $\Delta t > 0$ we can approximate $[0, N\Delta t]$ by $\{t_n\}_{n=0}^{N}$, where

$$t_n = n\Delta t.$$

We approximate a solution $u(t, x)$ on $[0, N\Delta t] \times [0, L]$ by the $(N + 1) \times (J + 1)$ matrix

$$\{u(t_n, x_j) : n = 0, \ldots, N, \quad j = 0, \ldots J\}$$

In order to "solve" the PDE, we also need to approximate the derivatives of $u$. Several options are available, all of which come from the Taylor series expansion for $u$:

$$\begin{aligned}
\partial_x u(x_j) \approx \delta_x^- u(x_j) &:= \frac{u(x_j) - u(x_{j-1})}{\Delta x} \qquad \text{(backward difference)} \\
\approx \delta_x^+ u(x_j) &:= \frac{u(x_{j+1}) - u(x_j)}{\Delta x} \qquad \text{(forward difference)} \\
\approx \delta_x u(x_j) &:= \frac{u(x_{j+1}) - u(x_{j-1})}{2\Delta x} \qquad \text{(centered difference).}
\end{aligned}$$

The first two have errors like $\Delta x$, while the third has an error like $(\Delta x)^2$.

Similarly,

$$\partial_{xx}u(x_j) \approx \delta_x^2 u(x_j) := \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{(\Delta x)^2} \quad \text{(centered second difference)}$$

with errors like $(\Delta x)^2$.

*Example* 10.1.1 (Heat equation). Consider

$$\begin{cases} u_t = u_{xx} & (t,x) \in (0,\infty) \times (0,L) \\ u(0,x) = f(x) & x \in [0,L], \\ u(t,0) = u(t,L) = 0 & t \in (0,\infty) \end{cases} \qquad (*)$$

Given $J, \Delta x$ and $N, \Delta t$ we can approximate this problem for $(t,x) \in [0, N\Delta t] \times [0,L]$ as follows:

$$\begin{cases} \delta_t^+ u(t_n, x_j) = \delta_x^2 u(t_n, x_j) & n = 0, \dots, N-1, \quad j = 1, \dots J-1 \\ u(t_0, x_j) = f(x_j) & j = 0, \dots, J \\ u(t_n, 0) = u(t_n, x_J) = 0 & n = 0, \dots, N. \end{cases}$$

We need to compute $u(t_n, x_j)$ for $n = 0, \dots, N$ and $j = 0, \dots, J$.

The initial condition gives us $u(t_0, x_j) = f(x_j)$ for $j = 0, \dots, J$.

The boundary condition gives us $u(t_n, 0) = u(t_n, x_J) = 0$ for $n = 0, \dots N$.

Defining $r = \frac{\Delta t}{(\Delta x)^2}$ the equation gives (after rearranging):

$$u(t_{n+1}, x_j) = (1-2r)u(t_n, x_j) + ru(t_n, x_{j+1}) + ru(t_n, x_{j-1}). \qquad (**)$$

for $n = 0, \dots, N-1$ and $j = 1, \dots, J-1$.

We can compute $u(t_{n+1}, \cdot)$ directly from $u(t_n, \cdot)$ (i.e. the scheme is **explicit**).

Moreover we know $u(t_0, \cdot)$, so by induction we can compute $u(t_n, \cdot)$ for each $n$.

Furthermore we expect the solution to converge to a solution to $(*)$ as $\Delta t, \Delta x \to 0$.

**Problem.** On a computer we cannot actually compute $u(t_n, x_j)$ to perfect accuracy! Instead we can only compute $\tilde{u}(t_n, x_j) = u(t_n, x_j) + e(t_n, x_j)$, where $e$ measures the error. We can only trust the computer if we can bound $e(t_n, \cdot)$ as $n$ increases. This is the question of the **numerical stability** of the scheme.

## 10.2 Numerical Stability

We try to answer this via **von Neumann stability analysis**: First we note that if $u$ and $\tilde{u}$ satisfy $(**)$, then so does $e$. (The problem is **linear**!)

**Idea.** Expand $e$ in a Fourier series:

$$e(t_n, x_j) = \sum_k p_k(t_n) \underbrace{e^{i\frac{k\pi}{L}x_j}}_{:=q_k(x_j)}.$$

**Goal.** Bound each $p_k(t_n)$.

After some rearranging, $(**)$ yields the following for the coefficients:

$$\frac{p_k(t_{n+1})}{p_k(t_n)} = \frac{(1-2r)q_k(x_j) + rq_k(x_{j+1}) + rq_k(x_{j-1})}{q_k(x_j)} = \lambda_k \quad \text{for some constant } \lambda_k.$$

In particular

$$p_k(t_{n+1}) = \lambda_k p_k(t_n) \implies p_k(t_n) = \lambda_k^n p_k(t_0),$$

and hence we find that stability requires $|\lambda_k| \leq 1$ for each $k$.

However, using the equation for the $q_k$'s (and Euler's formula $e^{i\theta} = \cos(\theta) + i\sin(\theta)$) we find

$$\lambda_k = 1 - 2r(1 - \cos(\tfrac{k\pi}{L}\Delta x)).$$

It follows that stability holds if and only if

$$r = \tfrac{\Delta t}{(\Delta x)^2} \leq \tfrac{1}{2}. \quad \square$$

We will next discuss another scheme (an implicit scheme called the Crank–Nicolson scheme) that features **unconditional stability**.

(Demonstrate)

We continue our discussion of finite difference schemes.

Recall the notation:

$$\delta_t^+ u(t_n) := \frac{u(t_{n+1}) - u(t_n)}{\Delta t},$$

$$\delta_x^2 u(x_j) := \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{(\Delta x)^2}.$$

**Example 2.** (Crank–Nicolson) We revisit the heat equation:

$$\begin{cases} u_t = u_{xx} & (t,x) \in (0,\infty) \times (0,L) \\ u(0,x) = f(x) & x \in [0,L] \\ u(t,0) = u(t,L) = 0 & t \in (0,\infty). \end{cases}$$

Previously we considered the explicit scheme

$$\delta_t^+ u(t_n, x_j) = \delta_x^2 u(t_n, x_j)$$

and saw that it was stable if and only if

$$r := \tfrac{\Delta t}{(\Delta x)^2} \leq \tfrac{1}{2}.$$

Now let $\theta \in (0,1)$ and instead consider the scheme

$$\delta_t^+ u(t_n, x_j) = (1-\theta)\delta_x^2 u(t_n, x_j) + \theta \delta_x^2 u(t_{n+1}, x_j).$$

(Setting $\theta = 0$ recovers the previous scheme.)

For $\theta > 0$ this scheme is **implicit**, since $u(t_{n+1}, \cdot)$ appears on the RHS.

This makes it harder to compute, since we have to solve systems of linear equations.

On the other hand, we will show that for $\theta \geq \frac{1}{2}$ the scheme is stable regardless of the value of $r$. We call such a scheme **unconditionally stable**.

The case $\theta = \frac{1}{2}$ is called the **Crank–Nicolson scheme**.

**Proof of stability.** As before, we use von Neumann stability analysis and expand the numerical error $e$ in a Fourier series:

$$e(t_n, x_j) = \sum_k p_k(t_n) q_k(x_j), \qquad q_k(x) = e^{i\frac{\pi k}{L}x}.$$

Applying the scheme to $e(t_n, x_j)$ and rearranging leads to

$$\frac{p_k(t_{n+1}) - p_k(t_n)}{(1-\theta)p_k(t_n) + \theta p_k(t_{n+1})} = r\frac{q_k(x_{j+1}) - 2q_k(x_j) + q_k(x_{j-1})}{q_k(x_j)} = \lambda_k \quad \text{for some constant } \lambda_k.$$

In particular one finds

$$\frac{p_k(t_{n+1})}{p_k(t_n)} = 1 + \frac{\lambda_k}{1 - \lambda_k\theta} \implies p_k(t_n) = \big(1 + \frac{\lambda_k}{1-\lambda_k\theta}\big)^n p_k(t_0) = \mu_k^n p_k(t_0),$$

and as before we need $|\mu_k^n| \leq 1$ for stability.

Now, the equation for the $q_k$ implies

$$\lambda_k = -2r(1 - \cos(\tfrac{k\pi}{L}\Delta x)) \leq 0.$$

This implies $\mu_k \leq 1$.

However, we still need $\mu_k \geq -1$. Rearranging, one finds that this is equivalent to the condition

$$|\lambda_k|(1 - 2\theta) \leq 2.$$

Notice that if $\theta \geq \frac{1}{2}$ then this condition is satisfied, since the LHS$\leq 0$.

**Conclusion.** For $\theta \geq \frac{1}{2}$, the scheme is stable (regardless of the value of $r = \frac{\Delta t}{(\Delta x)^2}$.) $\square$

## 10.3 Finite Element Method

**Example 3.** (finite element method) Let $\Omega \subset \mathbb{R}^d$ and consider Poisson's equation

$$\begin{cases} -\Delta u = f & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \qquad (*)$$

One can attempt to use finite difference schemes to solve this (see Strauss 8.4).

It turns out that for irregular/curved domains $\Omega$, this is not very efficient.

The finite element method provides an alternative.

The idea is to use a 'weak' formulation of $(*)$. In particular, $u$ solves $(*)$ if and only if we have

$$\int_\Omega -\Delta u\, \varphi \, dx = \int_\Omega f\varphi \, dx \quad \text{for all 'test functions'} \quad \varphi.$$

Equivalently:

$$\int_\Omega \nabla u \cdot \nabla \varphi \, dx = \int_\Omega f\varphi \, dx \quad \text{for all} \quad \varphi. \quad (**)$$

If we could construct $u$ satisfying $(**)$, then we would have $-\Delta u(x_0) = f(x_0)$ by the usual approximate identity argument, that is, by taking a sequence $\varphi_n$ converging to $\delta_{x_0}$.

The idea is therefore to choose a set of test functions $\{\varphi_j\}_{j=1}^N$ such that "$\varphi_j \approx \delta_{x_j}$" for some $x_j \in \Omega$. We will then approximate the solution by

$$u(x) \approx \tilde{u}(x) = \sum_{j=1}^N u_j \varphi_j(x) \quad \text{for some } \{u_j\}_{j=1}^N.$$

**Key idea.** We will choose the $u_j$ such that $(**)$ is satisfied when tested against each $\varphi_k$. Then $\tilde{u}$ should (hopefully) be a good approximate solution to $(*)$.

In particular, imposing $(**)$ with $u = \tilde{u}$ and $\varphi = \varphi_k$ gives

$$\sum_{j=1}^N u_j \underbrace{\int_\Omega \nabla\varphi_j \cdot \nabla\varphi_k \, dx}_{:=m_{jk}} = \underbrace{\int_\Omega f\varphi_k \, dx}_{:=f_k}$$

This is a linear equation of the form $M\vec{u} = \vec{f}$, and hence we determine $u_j$ by solving

$$\vec{u} = M^{-1}\vec{f},$$

where $M, \vec{f}$ are given in terms of the $\varphi_j$ and the data $f$.

The problem has now been reduced to linear algebra! (And the computation of $m_{jk}$, $f_k$.)

In practice, one must decide how to choose the functions $\varphi_j$.

If $\Omega \subset \mathbb{R}^2$ one commonly approximates $\Omega$ by a "triangulated" domain $\tilde{\Omega}$. Letting $\{x_j\}_{j=1}^N$ denote the interior vertices, one defines $\varphi_j$ to be a linear function such that $\varphi_j(x_j) = 1$ but $\varphi_j = 0$ at every other vertex.

# Chapter 11

# Topics from Physics

## 11.1 Classical Mechanics

At the heart of classical mechanics is **Newton's law**: $F = ma$. To be more precise, we let $x(t) \in \mathbb{R}^d$ denote the **position** of a particle $P$ of mass $m > 0$ at time $t$. Then $\dot{x}(t)$ denotes the **velocity** of $P$, and $\ddot{x}(t)$ denotes its **acceleration**. Newton's law reads $m\ddot{x}(t) = F(x(t))$, where $F$ describes the **force** acting on $P$.

Defining the **momentum** of $P$ by $p(t) = m\dot{x}(t)$, we get the first-order system of ODEs

$$\begin{cases} \dot{x}(t) = \frac{1}{m}p(t), \\ \dot{p}(t) = F(x(t)). \end{cases}$$

The force $F$ is called **conservative** if $\int_C F = 0$ for any closed loop $C \subset \mathbb{R}^d$. It turns out that this is equivalent to saying $F(x) = -\nabla V(x)$ for some $V$ called the **potential energy function.**

We call the force **attractive** if $F(x) \cdot x < 0$ and **repulsive** if $F(x) \cdot x > 0$.

So for a particle under the influence conservative forces the classical picture is the following:

$$(*) \quad \begin{cases} \dot{x}(t) = \frac{1}{m}p(t), \quad x(0) = x_0 \\ \dot{p}(t) = -\nabla V(x(t)), \quad p(0) = p_0. \end{cases}$$

The pair $(x(t), p(t)) \in \mathbb{R}^d \times \mathbb{R}^d$ gives us a complete description of the particle. (We call $\mathbb{R}^d \times \mathbb{R}^d$ the **phase space** or **state space** of the particle). In particular, every physical "observable" $Q$ is some function of $x(t)$ and $p(t)$.

The most basic examples are the position and momentum themselves. Perhaps the most important example is the **energy** (or **Hamiltonian**) of the particle, which is given by the sum of the **kinetic energy** and the **potential energy**:

$$H(x(t), p(t)) = \tfrac{1}{2m}|p(t)|^2 + V(x(t)).$$

The Hamiltonian 'drives the dynamics'. Indeed $(*)$ is equivalent to **Hamilton's equations**

$$\begin{cases} \dot{x}(t) = +\frac{\partial H}{\partial p}, & x(0) = x_0, \\ \dot{p}(t) = -\frac{\partial H}{\partial x}, & p(0) = p_0. \end{cases}$$

Notice that we have **conservation of energy:**

$$\begin{aligned} \partial_t H(x(t), p(t)) &= \tfrac{1}{m} p(t) \cdot \dot{p}(t) + \nabla V(x(t)) \cdot \dot{x}(t) \\ &= -\dot{x}(t) \cdot \nabla V(x(t)) + \nabla V(x(t)) \cdot \dot{x}(t) = 0, \end{aligned}$$

so that

$$H(x(t), p(t)) \equiv H(x_0, p_0) =: E.$$

Note $E < 0$ is certainly possible.

If $d = 3$ one can also consider the **angular momentum** of the particle:

$$L(t) = x(t) \times p(t) \qquad \text{(cross product)}.$$

*Example* 11.1.1 (Free particle). If $V \equiv 0$ then $p(t) \equiv p_0$ and $x(t) = x_0 + \frac{p_0}{m} t$.

In this case the energy is pure kinetic energy.

Notice that (in general) $\lim_{t\to\infty} |x(t)| = \infty$. We say the particle **scatters**.

*Example* 11.1.2 (Harmonic oscillator). $V(x) = \frac{1}{2}k|x|^2$ for some $k > 0$ $\quad (F(x) = -kx)$

For $d = 1$ one can work out the solutions explicitly (sines and cosines).

Solutions are oscillatory, transferring energy back and forth between kinetic and potential energy.

Solutions are **bound**: Indeed, since $V(x) \to \infty$ as $|x| \to \infty$ it would require **infinite** energy to have $\lim_{t\to\infty} |x(t)| = \infty$.

*Example* 11.1.3 (Finite well). Consider

$$V(x) = \begin{cases} \frac{1}{2}k|x|^2 & |x| \le r_0 \\ \\ \frac{1}{2}k|r_0|^2 & |x| > r_0. \end{cases}$$

The behavior of the particle depends on its energy.

If $E > \frac{1}{2}k|r_0|^2$ then the particle will eventually 'escape the well', at which point it will behave like a free particle and **scatter**. In particular $\lim_{t\to\infty} |x(t)| = \infty$.

If $E < \frac{1}{2}k|r_0|^2$ then if the particle ever finds itself in the well, it will not be able to escape. In particular it will be **bound**.

In general, one has the following:

- If $E > \lim_{|x|\to\infty} V(x)$, then the particle scatters.
- Otherwise the particle may be bound.

**Example 4.** (gravity) Suppose we fix a "point mass" $P'$ of mass $M$ at the origin. The Newtonian model of gravity models the effect of $P'$ on $P$ via

$$V(x(t)) = -GMm\frac{1}{|x(t)|}, \quad \text{i.e.} \quad F(x(t)) = -GMm\frac{x(t)}{|x(t)|^3},$$

where $G$ is the "gravitational constant". The function $V(x) = -\frac{1}{|x|}$ is called the **Newtonian potential**. Note that $F$ is an attractive force.

**Question.** (escape velocity) Fix $x_0 \neq 0$. How large does $|p_0|$ need to be in order for $x(t)$ to scatter?

(Answer: choose $p_0$ large enough that $E > 0$.)

## 11.2 Quantum Mechanics

Recall: classical mechanics models a particle $P$ under conservative forces ($F = -\nabla V$) by states $(x(t), p(t)) \in \mathbb{R}^d \times \mathbb{R}^d$, with dynamics governed by Newton's law/Hamilton's equations.

Classical mechanics works well in many settings, but breaks down in many limiting regimes.

We next discuss quantum mechanics, where the fundamental law (i.e. Newton's law) is replaced by a new one, namely the **Schrödinger equation**:

$$(*) \quad \begin{cases} i\hbar\Psi_t = -\frac{\hbar^2}{2m}\Delta\Psi + V(x)\Psi & (t,x) \in \mathbb{R}\times\mathbb{R}^d \\ \Psi(0,x) = \Psi_0(x) & x \in \mathbb{R}^d. \end{cases}$$

(Here $i = \sqrt{-1}$ and $\hbar$ is **Planck's constant**.)

To be more precise, suppose $P$ is a particle of mass $m$ in the presence of a potential $V : \mathbb{R}^d \to \mathbb{R}$.

The quantum mechanics model describes $P$ as an element $\Psi \in L^2(\mathbb{R}^d)$, called the **state** of $P$.

If $P$ is in the state $\Psi_0$ at time $t = 0$, then the state of $P$ at later times is determined by solving the Schrödinger equation. We call $\Psi(t, x)$ the **wave function** of $P$.

In this model the particle does not have a well-defined position or momentum. Instead, $\Psi(x)$ is the 'position distribution' of $P$ and $\Phi(\xi) = \hbar^{-d/2}\widehat{\Psi}(\frac{\xi}{\hbar})$ is the 'momentum distribution' of $P$.

In particular we have the following interpretations:

$$\int_\Omega |\Psi(x)|^2\,dx = \text{probability that the position of } P \text{ is in } \Omega,$$
$$\int_{\Omega'} |\Phi(\xi)|^2\,d\xi = \text{probability that the momentum of } P \text{ is in } \Omega'.$$

According to this interpretation, the initial state $\Psi_0$ should be "normalized", that is:

$$\int_{\mathbb{R}^d} |\Psi_0|^2\,dx = 1.$$

(Exercise: show that the Schrödinger equation preserves this property.)

Position and momentum have new interpretations under this model.

In general, an "observable" $Q$ is modelled as a hermitian operator on $L^2$, with the following statistical interpretation: if $P$ is in the state $\Psi$ then the **expected value** of $Q$ is

$$\mathbf{E}(Q) = \langle \Psi, Q\Psi \rangle = \int_{\mathbb{R}^d} \overline{\Psi(x)}[Q\Psi](x)\,dx.$$

(Here **hermitian** means that $\langle \Psi, Q\Psi \rangle = \langle Q\Psi, \Psi \rangle$.)

**Position** and **momentum** are modelled by the operators

$$[x\Psi](x) = x\Psi(x), \quad [p\Psi](x) = -i\hbar\nabla\Psi(x).$$

(Exercise: check that $x, p$ are self-adjoint.)

As in classical mechanics, other observables can be written as combinations of $x$ and $p$.

For example, the energy/Hamiltonian was given by $H(x, p) = \frac{1}{2m}p \cdot p + V(x)$. Similarly in quantum mechanics the Hamiltonian is given by

$$H\Psi = \tfrac{-i\hbar}{2m}\nabla \cdot (-i\hbar\nabla\Psi) + V(x)\Psi = -\tfrac{\hbar^2}{2m}\Delta\Psi + V(x)\Psi.$$

As in the classical setting, the Hamiltonian drives the dynamics. Indeed, $(*)$ is equivalent to

$$i\hbar\partial_t\Psi = H\Psi$$

If $d = 3$ we can also define **angular momentum**. Since we had $L = x \times p$, we define

$$L\Psi = -i\hbar\, x \times \nabla\Psi.$$

**Example 1.** (free particle) Suppose $V \equiv 0$. Then $(*)$ becomes

$$\begin{cases} i\Psi_t = -k\Delta\Psi, \quad k = \frac{\hbar}{2m} \\ \Psi(0,x) = \Psi_0(x). \end{cases}$$

Applying the Fourier transform:

$$i\widehat{\Psi}_t(t,\xi) = k|\xi|^2\Psi(t,\xi).$$

We find

$$\widehat{\Psi}(t,\xi) = e^{-ikt|\xi|^2}\widehat{\Psi}(0,\xi) = e^{-ikt|\xi|^2}\widehat{\Psi_0}(\xi).$$

We therefore write

$$\Psi(t) = e^{ikt\Delta}\Psi_0, \quad e^{ikt\Delta} = \mathcal{F}^{-1}e^{-ikt|\xi|^2}\mathcal{F}.$$

**Question.** How does this compare to the classical picture?

In the classical setting we found $x(t) = x_0 + \frac{p_0}{m}t, \quad p(t) \equiv p_0$.

In the quantum setting we cannot specify both $x_0$ and $p_0$. As a substitute, consider an initial state

$$\Psi_0(x) = ce^{-|x-x_0|^2/2}e^{ix\cdot\xi_0},$$

where $c$ is chosen so that $\|\Psi_0\| = 1$. Then

$$\widehat{\Psi_0}(\xi) = ce^{-|\xi-\xi_0|^2/2}e^{-i\xi\cdot x_0} \implies \Phi_0(\xi) = \hbar^{-d/2}\widehat{\Psi_0}(\tfrac{\xi}{\hbar}) = c\hbar^{-d/2}e^{-|\xi-\hbar\xi_0|^2/2\hbar^2}e^{-i\xi\cdot x_0/\hbar}.$$

Thus $\Phi_0$ is concentrated where $\xi \sim p_0$ provided we choose $\xi_0 = \frac{p_0}{\hbar}$.

From above we have $\widehat{\Psi}(t,\xi) = e^{-ikt|\xi|^2}\widehat{\Psi_0}(\xi)$, so $\widehat{\Psi}$ remains concentrated where $\xi \sim \xi_0$.

Equivalently $\Phi$ remains concentrated near $\xi \sim p_0$.

Taking the inverse Fourier transform we have

$$\Psi(t,x) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{ix\cdot\xi}e^{-ikt|\xi|^2}\widehat{\Psi_0}(\xi)\,d\xi$$

Since $\widehat{\Psi_0}$ is concentrated where $\xi \sim \xi_0$ we use Taylor's theorem and approximate

$$|\xi|^2 \approx |\xi_0|^2 + 2\xi_0 \cdot (\xi - \xi_0).$$

Thus

$$\begin{aligned}
\Psi(t,x) &\approx (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{ix\cdot\xi} e^{-ikt[|\xi_0|^2 + 2\xi_0\cdot(\xi-\xi_0)]} \widehat{\Psi_0}(\xi)\, d\xi \\
&= (2\pi)^{-d/2} e^{-ikt|\xi_0|^2} e^{ix\cdot\xi_0} \int_{\mathbb{R}^d} e^{i(x-2tk\xi_0)\cdot(\xi-\xi_0)} \widehat{\Psi_0}(\xi - \xi_0 + \xi_0)\, d\xi \\
&= (2\pi)^{-d/2} e^{-ikt|\xi_0|^2 + ix\cdot\xi_0} \int_{\mathbb{R}^d} e^{i(x-2tk\xi_0)\cdot(\xi-\xi_0)} \mathcal{F}[e^{-i\xi_0\cdot x}\Psi_0](\xi - \xi_0)\, d\xi \\
&= e^{-ikt|\xi_0|^2 + ix\cdot\xi_0} \big[e^{-i\xi_0(x-2tk\xi_0)} \Psi_0(x - 2tk\xi_0)\big] \\
&= e^{itk|\xi_0|^2} \Psi_0(x - 2tk\xi_0).
\end{aligned}$$

Thus $\Psi$ travels with speed $2k\xi_0 = \frac{\hbar\xi_0}{m} = \frac{p_0}{m}$, in agreement with the classical picture!

Recall the Schrödinger equation

$$(*) \quad \begin{cases} i\hbar\Psi_t = H\Psi \\ \Psi(0) = \Psi_0 \in L^2(\mathbb{R}^d). \end{cases}$$

Here $\Psi(t,x)$ is the wave function of a particle in the presence of a potential $V : \mathbb{R}^d \to \mathbb{R}$, and

$$H = -\tfrac{\hbar^2}{2m}\Delta + V(x)$$

is the Hamiltonian (or 'energy' observable).

Above we used the Fourier transform to study the case of a free particle $V \equiv 0$. We recovered some qualitative information in agreement with the classical picture.

In the case $V \not\equiv 0$, one often uses separation of variables.

Looking for a solution of $(*)$ of the form $\Psi(t,x) = p(t)q(x)$ leads to

$$\frac{i\hbar p'(t)}{p(t)} = \frac{-\frac{\hbar^2}{2m}\Delta q(x) + V(x)q(x)}{q(x)} = \lambda \quad \text{for some} \quad \lambda \in \mathbb{C}.$$

In particular, we are led to the eigenvalue problem for the Hamiltonian:

$$[Hq](x) = \lambda q(x). \quad (**)$$

**Remark 11.2.1.** (i) Operators of the form $H = -\Delta + V$ are called **Schrödinger operators**.

(ii) For any $\lambda$ we can solve $i\hbar p'(t) = \lambda p(t)$ via $p_\lambda(t) = c_\lambda e^{-i\lambda t/\hbar}$, so the interesting problem is $(**)$.

Thus, we have the following

**Possible strategy:**

- Find the eigenvalues $\lambda_n$ and (orthogonal) eigenfunctions $q_n$ for $(**)$.
- Show that any $\Psi_0 \in L^2$ can be written as $\Psi_0(x) = \sum_n c_n q_n(x)$.
- Solve $(*)$ via $\Psi(t,x) = \sum_n c_n p_n(t) q_n(x)$.

In some cases, this strategy works beautifully (e.g. infinite well, harmonic oscillator, Coulomb potential, ...). We will do an example in detail in what follows.

Note that ($L^2$-)eigenfunctions $q_n$ are steady-state solutions to $(*)$. If a particle has state $q_n$ then measurement of the energy will yield $\lambda_n$ (see below). We call such states **determinate states** for the energy.

**Possible issues:**

- What if the 'eigenvalues' are no longer discrete?
- What if the 'eigenfunctions' do not belong to $L^2$?

These issues are really 'two sides of the same coin'. They are connected to the "**spectral theory**" of Schrödinger operators.

To make a long story short, the solution to both issues is to replace the sums by integrals. The "eigenfunctions" are still useful in building up states $\Psi(t,x)$ mathematically, but they no longer represent possible states themselves.

Before moving on to examples, we discuss some general principles of the quantum mechanics model.

Recall that observables are modelled by hermitian operators $Q$ on $L^2$. If $P$ is in state $\Psi$ then the expected value of $Q$ is

$$\mathbf{E}(Q) = \langle \Psi, Q\Psi \rangle = \int_{\mathbb{R}^d} \overline{\Psi}(x)[Q\Psi](x)\, dx.$$

(In fact, this definition makes sense for any operator $Q$.)

If $Q$ is hermitian then

$$\overline{\mathbf{E}(Q)} = \overline{\langle \Psi, Q\Psi \rangle} = \langle Q\Psi, \Psi \rangle = \langle \Psi, Q\Psi \rangle = \mathbf{E}(Q)$$

Thus $\mathbf{E}(Q) \in \mathbb{R}$, as one would expect.

For $\lambda \in \mathbb{R}$ we define $Q - \lambda$ by $(Q-\lambda)\Psi = Q\Psi - \lambda\Psi$. If $Q$ is hermitian, so is $Q - \lambda$. (check!)

The **variance** of $Q$, defined by $\sigma_Q^2 = \mathbf{E}\big((Q - \mathbf{E}(Q))^2\big)$, is computed by

$$\sigma_Q^2 = \langle \Psi, (Q-\mathbf{E}(Q))^2\Psi\rangle = \langle (Q-\mathbf{E}(Q))\Psi, (Q-\mathbf{E}(Q))\Psi\rangle = \|(Q-\mathbf{E}(Q))\Psi\|^2.$$

The **commutator** of two observables $Q, Q'$ is the operator defined by $[Q, Q'] = QQ' - Q'Q$. (Remark: $\mathbf{E}([Q, Q'])$ is pure imaginary. Check!)

For any state $\Psi$ and any two observables $Q, Q'$, the **uncertainty principle** holds:

$$\sigma_Q^2 \sigma_{Q'}^2 \geq \big(-\tfrac{i}{2}\mathbf{E}([Q, Q'])\big)^2.$$

A **determinate state** for an observable $Q$ is a state $\Psi \in L^2$ such that $\sigma_Q^2 = 0$. Equivalently:

$$(Q - \mathbf{E}(Q))\Psi = 0, \quad \text{or} \quad Q\Psi = \mathbf{E}(Q)\Psi.$$

That is, determinate states are $L^2$-eigenvectors; the expected values are the eigenvalues.

The set of $\lambda$ such that $Q - \lambda$ is not (boundedly) invertible is called the **spectrum** of $Q$, denoted $\sigma(Q)$. The set of $\lambda \in \sigma(Q)$ with $L^2$-eigenvectors is the **discrete spectrum**; the set of $\lambda \in \sigma(Q)$ without $L^2$-eigenvectors is the **continuous spectrum**.

### Eigenvalue problem

Consider the eigenvalue problem $Hq = \lambda q$ in three dimensions, where

$$H = -\tfrac{\hbar^2}{2m}\Delta + V(x), \qquad V : \mathbb{R}^3 \to \mathbb{R}.$$

Let us only consider **radial potentials**, i.e. $V = V(|x|)$, and let $\kappa = \frac{\hbar^2}{2m}$.

We introduce **spherical coordinates:**

$$(x_1, x_2, x_3) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$$

Then

$$\Delta = r^{-2}[\partial_r(r^2\partial_r) + (\sin\theta)^{-1}\partial_\theta(\sin\theta\partial_\theta) + (\sin\theta)^{-2}\partial_{\phi\phi}].$$

We use separation of variables and look for a solution of the form

$$q(r, \theta, \phi) = a(r)b(\theta, \phi).$$

Then $Hq = \lambda q$ becomes

$$-\kappa r^{-2}\big[b\partial_r(r^2 a_r) + a\{(\sin\theta)^{-1}\partial_\theta(\sin\theta b_\theta) + (\sin\theta)^{-2}b_{\phi\phi}\}\big] + V(r)ab = \lambda ab.$$

Rearranging and multiplying by $\frac{r^2}{\kappa ab}$:

$$\underbrace{\tfrac{1}{a}\partial_r(r^2 a_r) - \tfrac{r^2}{\kappa}[V(r)-\lambda]}_{\text{radial equation}} = \underbrace{-\tfrac{1}{b}[(\sin\theta)^{-1}\partial_\theta(\sin\theta b_\theta) + (\sin\theta)^{-2} b_{\phi\phi}]}_{\text{angular equation}} = \mu\kappa$$

for some $\mu \in \mathbb{C}$.

We now study the **angular equation**, which does not depend on $V$ (since $V$ is radial).

We use separation of variables (again) and look for a solution of the form

$$b(\theta,\phi) = f(\theta)g(\phi).$$

Multiplying by $-\sin^2\theta$ and rearranging yields

$$\tfrac{1}{f}\sin\theta\partial_\theta(\sin\theta f'(\theta)) + \mu\sin^2\theta = -\tfrac{1}{g}g''(\phi) = \alpha$$

for some $\alpha \in \mathbb{C}$.

We solve $\quad g'' = -\alpha g \quad$ via $\quad g(\phi) = e^{\pm i\sqrt{\alpha}\phi}$.

Since we want $g(\phi + 2\pi) = g(\phi)$, we need $e^{2\pi i\sqrt{\alpha}} = 1$, and hence

$$\sqrt{\alpha} = m \quad \text{for some} \quad m \in \{0\} \cup \mathbb{N}.$$

Rearranging the $f$ equation gives

$$\tfrac{\partial_\theta(\sin\theta f')}{\sin\theta} + [\mu - \tfrac{m^2}{\sin^2\theta}]f = 0.$$

If we define $z = \cos\theta$ and suppose $f(\theta) = h(z)$, this becomes

$$\begin{cases} (1-z^2)h'' - 2zh' + [\mu - \frac{m^2}{1-z^2}]h = 0 \qquad (*)_m \quad \text{(check!)} \\ h(z) \text{ finite at } z = \pm 1. \end{cases}$$

The case $m = 0$ is **Legendre's equation**; the others are **associated Legendre equations**.

We first solve $(*)_0$. We will generate solutions to $(*)_m$ below.

We employ the **power series method** and write

$$h(z) = \sum_{k=0}^{\infty} c_k z^k, \quad h'(z) = \sum_{k=0}^{\infty} k c_k z^{k-1}, \quad h''(z) = \sum_{k=0}^{\infty} k(k-1)c_k z^{k-2}.$$

Then $(*)_0$ becomes (after rearranging)

$$\sum_{k=0}^{\infty} \{(k+2)(k+1)c_{k+2} - [k(k+1) - \mu]c_k\} z^k = 0.$$

Thus $h$ is a solution provided

$$c_{k+2} = \tfrac{k(k+1)-\mu}{(k+2)(k+1)} c_k, \quad c_0, c_1 \text{ arbitrary.}$$

Note that **if** $\mu = \ell(\ell+1)$ for some $\ell \in \{0\} \cup \mathbb{N}$ then $c_{\ell+2} = c_{\ell+4} = \cdots = 0$.

Thus for $\mu = \ell(\ell+1)$ we can (by choosing $c_0 = 0$ or $c_1 = 0$) find a degree $\ell$ polynomial solution $h_0^\ell$, where $h_0^\ell$ is odd for $\ell$ odd and even for $\ell$ even. (These are the **Legendre polynomials**.)

If $\mu$ is **not** of this form then solutions exist but are badly behaved at $z = 1$ or $z = -1$.

We turn to $(*)_m$ for $m > 0$. Recall $\mu = \ell(\ell+1), \quad \ell \in \{0\} \cup \mathbb{N}$.

As $h_0^\ell$ solves $(*)_0$ then the function $h_m^\ell = \frac{\partial^m}{\partial z^m} h_0^\ell$ solves

$$(1-z^2)h_m'' - 2(m+1)zh_m' + [\ell(\ell+1) - m(m+1)]h_m = 0,$$

and it follows that the **associated Legendre function**

$$H_m^\ell := (1-z^2)^{m/2} h_m^\ell$$

satisfies $(*)_m$. (Check these!)

Note that we should restrict to $m \leq \ell$ to guarantee $h_m^\ell \neq 0$ (recall $h_0^\ell$ is a degree $\ell$ polynomial).

**Conclusion.** For $\ell \in \{0\} \cup \mathbb{N}$ and $m \in \{0, \ldots, \ell\}$ we can solve the angular equation with $\mu = \ell(\ell+1)$ with

$$b_{\pm m}^\ell(\theta, \phi) = e^{\pm im\phi} H_m^\ell(\cos(\theta)).$$

These functions (after normalization) are called **spherical harmonics**.

The number $\ell$ is called the **azimuthal quantum number**, while $m$ is called the **magnetic quantum number.**

(The spherical harmonics are determinate states for angular momentum.)

We return to the eigenvalue problem $Hq = \lambda q$, where

$$H = -\kappa\Delta + V(x),$$

with $\kappa = \frac{\hbar^2}{2m}$ and $V : \mathbb{R}^3 \to \mathbb{R}$ radial.

Employing spherical coordinates and looking for solutions of the form

$$q(r, \theta, \phi) = a(r)b(\theta, \phi)$$

led to the "angular equation" for $b$, which we solved via spherical harmonics $b_{\pm m}^\ell$, and the **radial equation**

$$\begin{cases} r^2 a'' + 2ra' - \frac{r^2}{\kappa}[V(r) - \lambda]a = \ell(\ell+1)a \\ \\ \ell \in \{0\} \cup \mathbb{N}, \quad \lim_{r\to\infty} a(r) = 0, \end{cases}$$

to which we turn now.

Introducing the function $u(r) = r\,a(r)$ and rearranging yields

$$\begin{aligned} u'' - \tfrac{1}{\kappa}[V(r) - \lambda]u &= \tfrac{\ell(\ell+1)}{r^2}u \\ \rightsquigarrow -\kappa u'' + [V(r) + \kappa\tfrac{\ell(\ell+1)}{r^2}]u &= \lambda u \qquad (**) \end{aligned}$$

This is the eigenvalue problem for a one-dimensional Schrödinger operator on $(0, \infty)$ with potential

$$\tilde{V}(r) = V(r) + \kappa\tfrac{\ell(\ell+1)}{r^2}.$$

**Example.** QM models an electron in orbit around a proton with the **Coulomb potential:**

$$V(x) = -\tfrac{e^2}{4\pi\varepsilon_0}\tfrac{1}{|x|}, \quad \text{i.e.} \quad V(r) = -\tfrac{e^2}{4\pi\varepsilon_0}\tfrac{1}{r}.$$

where $e$ is the charge of the electron and $\varepsilon_0$ is the permittivity of vacuum (in SI units).

We are interested in the **discrete spectrum**, corresponding to "bound states". Eigenvalues are real (they are expected values of the Hamiltonian), and because $\lim_{r\to\infty} V(r) = 0$ one can show they are contained in $(-\infty, 0)$. (Beyond the scope of this course.)

With the Coulomb potential, $(*)$ becomes

$$u'' = c^2[1 - \tfrac{e^2}{4c\pi\kappa\varepsilon_0}\tfrac{1}{cr} + \tfrac{\ell(\ell+1)}{(cr)^2}]u,$$

where $c^2 = -\frac{\lambda}{\kappa}$. We define

$$\rho = cr, \qquad \tfrac{d}{d\rho} = \tfrac{1}{c}\tfrac{d}{dr}, \qquad u(r) = w(\rho), \qquad \rho_0 = \tfrac{e^2}{4c\pi\kappa\varepsilon_0}.$$

Then

$$w'' = [1 - \tfrac{\rho_0}{\rho} + \tfrac{\ell(\ell+1)}{\rho^2}]w, \qquad ' = \tfrac{d}{d\rho} \qquad (**)$$

We revisit the **power series method**, but with a twist.

First, as $\rho \to \infty$ this ODE resembles $w'' = w$, whose (decaying) solution is like $e^{-\rho}$.

Next, as $\rho \to 0$ this ODE resembles $w'' = \frac{\ell(\ell+1)}{\rho^2}w$, whose (well-behaved) solution is like $\rho^{\ell+1}$.

Thus we look for $w$ of the form

$$w(\rho) = \rho^{\ell+1}e^{-\rho}v(\rho).$$

We compute

$$w' = \rho^\ell e^{-\rho}[\rho v' + (\ell + 1 - \rho)v],$$
$$w'' = \rho^\ell e^{-\rho}\left\{\rho v'' + 2[\ell + 1 - \rho]v' + [-2\ell - 2 + \rho + \tfrac{\ell(\ell+1)}{\rho}]v\right\}.$$

Then $(**)$ becomes

$$\rho v'' + 2(\ell + 1 - \rho)v' + [\rho_0 - 2(\ell + 1)]v = 0.$$

Now we try to write

$$v(\rho) = \sum_{k=0}^{\infty} c_k \rho^k, \quad v'(\rho) = \sum_{k=0}^{\infty} k c_k \rho^{k-1}, \quad v''(\rho) = \sum_{k=0}^{\infty} k(k-1) c_k \rho^{k-2}.$$

Inserting this into the ODE yields

$$\sum \left\{k(k+1)c_{k+1} + 2(\ell+1)(k+1)c_{k+1} - 2kc_k + [\rho_0 - 2(\ell+1)]c_k\right\}\rho^k = 0.$$

Thus we should take

$$c_{k+1} = \tfrac{2(k+\ell+1)-\rho_0}{(k+1)(k+2\ell+2)} c_k, \quad c_0 \text{ arbitrary.}$$

Now **if** $\rho_0 = 2(k^* + \ell + 1)$ for some $k^* \in \{0\} \cup \mathbb{N}$, then $c_{k^*+1} = c_{k^*+2} = \cdots = 0$.

(If $\rho_0$ is not of this form it turns out that the solution is badly behaved.)

In this case we call $n := k^* + \ell + 1$ the **principle quantum number**.

We denote the degree $n - \ell - 1$ polynomial solution by $v_n^\ell(\rho)$ (called a **Laguerre polynomial**).

Then

$$u_n^\ell(r) = w_n^\ell(\rho) = \rho^{\ell+1} e^{-\rho} v_n^\ell(\rho), \quad \rho = c_n r = \sqrt{\tfrac{|\lambda_n|}{\kappa}}\, r$$
$$\rightsquigarrow \quad a_n^\ell(r) = r^\ell e^{-c_n r} v_n^\ell(c_n r),$$

where we can unravel the constants in

$$2n = \rho_0 = \tfrac{e^2}{4c_n \pi \kappa \varepsilon_0}, \quad c_n^2 = -\tfrac{\lambda_n}{\kappa}, \quad \kappa = \tfrac{\hbar^2}{2m_e}$$

to compute $\lambda_n$. In particular:

**Conclusion.** The determinate energies of the hydrogen atom are

$$\lambda_n = -\left[\left(\tfrac{e^2}{4\pi\varepsilon_0}\right)^2 \tfrac{m_e}{2\hbar^2}\right] \tfrac{1}{n^2} \quad \text{for } n \in \mathbb{N},$$

and for each $n$ there are $2n^2$ (orthogonal) eigenvectors

$$q_{n,\pm m}^\ell(r, \theta, \phi) = a_n^\ell(r) b_{\pm m}^\ell(\theta, \phi),$$

where $\ell \in \{0, 1, \ldots, n-1\}$, $m \in \{0, \ldots, \ell\}$.

## 11.3 Electromagnetism

In classical electromagnetism, the **electric field** $E(t,x) \in \mathbb{R}^3$ and **magnetic field** $B(t,x) \in \mathbb{R}^3$ have the following interpretation: the force on a particle with charge $q$ and velocity $v$ is

$$F = q(E + v \times B),$$

while the behavior of $E, B$ is determined by **Maxwell's equations**:

$$\begin{cases} \nabla \cdot E = \frac{1}{\varepsilon_0}\rho & \text{(Gauss's law)} \\ \nabla \cdot B = 0 & \\ \nabla \times E + \partial_t B = 0 & \text{(Faraday's law)} \\ \nabla \times B - \mu_0\varepsilon_0\partial_t E = \mu_0 J, & \text{(Ampere's law, with Maxwell's correction)} \end{cases}$$

where $\rho = \rho(t,x) \in \mathbb{R}$ is the charge density, $J = J(t,x) \in \mathbb{R}^3$ is the current density, $\varepsilon_0$ is the permittivity of vacuum, and $\mu_0$ is the permeability of vacuum.

Using Gauss's law and Ampere's law we can also deduce the **continuity equation** for $\rho$ and $J$:

$$\partial_t \rho + \nabla \cdot J = 0.$$

Consider a region of space in which there is no charge or current ($\rho, J = 0$). Note

$$\nabla \times (\nabla \times E) = \nabla \times (-\partial_t B) = -\partial_t(\nabla \times B) = -\partial_t(\mu_0\varepsilon_0 \tfrac{\partial E}{\partial t}) = -\mu_0\varepsilon_0\partial_{tt}E$$

and

$$\nabla \times (\nabla \times E) = -\Delta E + \nabla(\nabla \cdot E) = -\Delta E.$$

Rearranging yields

$$\mu_0\varepsilon_0\partial_{tt}E - \Delta E = 0.$$

Similarly (computing $\nabla \times (\nabla \times B)$ we find

$$\mu_0\varepsilon_0\partial_{tt}B - \Delta B = 0.$$

**Conclusion.** In regions of spacetime with no charge or current, the components of $E$ and $B$ satisfy the wave equation, with speed of propagation equal to $c = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$ (which happens to be the speed of light in vacuum).

One can also study Maxwell's equations using the **potential formulation**.

**Lemma 11.3.1** (Helmholtz Theorem)**.** *If $G$ is a (sufficiently nice, decaying) vector field then*

$$G = -\nabla g + \nabla \times H$$

*for some scalar field $g$ and vector field $H$.*

**Corollary.** If $\nabla \cdot G = 0$ then $G = \nabla \times H$ and if $\nabla \times G = 0$ then $G = -\nabla g$.

We apply these results to $E$ and $B$. In particular, $\nabla \cdot B = 0$ implies

$$B = \nabla \times A \quad \text{for some } \textbf{magnetic potential } A.$$

Combining this with Faraday's law yields $\nabla \times (E + \partial_t A) = 0$, so that

$$E + \partial_t A = -\nabla V \quad \text{for some } \textbf{electric potential } V.$$

Plugging this into Gauss's law yields

$$-\Delta V = \nabla \cdot E + \partial_t \nabla \cdot A = \tfrac{1}{\varepsilon_0}\rho + \partial_t \nabla \cdot A.$$

(Note if $\partial_t B = 0$ then $\partial_t A = 0$ and this reduces to Coulomb's law from electrostatics.)

Using Ampere's law and $B = \nabla \times A$ yields

$$\begin{aligned} &\nabla \times (\nabla \times A) + \mu_0 \varepsilon_0 \partial_t [\partial_t A + \nabla V] = \mu_0 J \\ \implies &\mu_0 \varepsilon_0 \partial_{tt} A - \Delta A + \nabla [\nabla \cdot A + \mu_0 \varepsilon_0 \partial_t V] = \mu_0 J. \end{aligned}$$

Thus, given $\rho$ and $J$ the equations for $A$ and $V$ are:

$$(*) \quad \begin{cases} -\Delta V - \partial_t \nabla \cdot A = \frac{1}{\varepsilon_0}\rho \\ \\ \mu_0 \varepsilon_0 \partial_{tt} A - \Delta A + \nabla(\nabla \cdot A + \mu_0 \varepsilon_0 \partial_t V) = \mu_0 J. \end{cases}$$

This looks horrible.

**Key observation.** (gauge invariance) Let $\lambda(t, x)$ be a scalar function. If we define

$$\tilde{A} = A + \nabla \lambda, \quad \tilde{V} = V - \partial_t \lambda,$$

then $\tilde{A}$ and $\tilde{V}$ yield the same fields $E, B$. Indeed,

$$\nabla \times \tilde{A} = \nabla \times A + 0 = B,$$

$$\partial_t \tilde{A} + \nabla \tilde{V} = \partial_t A + \partial_t \nabla \lambda + \nabla V - \partial_t \nabla \lambda = \partial_t A + \nabla V = -E.$$

Thus we have some freedom in how we choose $A, V$, which may simplify $(*)$.

**Example 1.** If we choose $A$ to satisfy the **Coulomb gauge condition**

$$\nabla \cdot A = 0,$$

then $(*)$ becomes

$$\begin{cases} -\Delta V = \frac{1}{\varepsilon_0}\rho, \\ \mu_0\varepsilon_0\partial_{tt}A - \Delta A = \mu_0 J - \mu_0\varepsilon_0\partial_t\nabla V. \end{cases}$$

**Example 2.** If we choose $A$ to satisfy the **Lorentz gauge condition**

$$\nabla \cdot A = -\mu_0\varepsilon_0\partial_t V,$$

then $(*)$ becomes

$$\begin{cases} \mu_0\varepsilon_0\partial_{tt}V - \Delta V = \frac{1}{\varepsilon_0}\rho, \\ \mu_0\varepsilon_0\partial_{tt}A - \Delta A = \mu_0 J. \end{cases}$$

Once again we get the (inhomogeneous) wave equation, with speed of propagation equal to $c$.

It is generally accepted that the laws of physics should be invariant under space translation, space rotation, and time translation.

The laws of Newtonian physics (and its descendants) are also invariant under **Galilean transformations**, such as

$$(t, x_1, x_2, x_3) \mapsto (t, x_1 - vt, x_2, x_3) \quad \text{for } v \in \mathbb{R}.$$

The laws of electromagnetism are **not**. Instead, they are invariant under **Lorentz transformations**, such as

$$(t, x_1, x_2, x_3) \mapsto \left(\frac{t - (vx_1/c^2)}{\sqrt{1-(v/c)^2}}, \frac{x_1 - vt}{\sqrt{1-(v/c)^2}}, x_2, x_3\right), \quad v \in \mathbb{R}, \quad |v| < c.$$

(These arise when one takes as an axiom that the speed of light is the same for any inertial observer.)

Einstein's insight was that the laws of mechanics should be modified to be invariant under Lorentz transformations. (This led to the theory of relativity.)

## 11.4 Elementary Particles

The QM model improves the classical model at microscopic length scales.

Relativity substitutes Galilean invariance with Lorentz invariance, which improves the classical mechanics model at high velocities.

The QM model still respects Galilean invariance.

The modification of QM to a model that respects Lorentz invariance is known as **relativistic quantum mechanics**.

**Classical mechanics:** (Lagrangian version) If $P$ is a particle of (rest) mass $m$ in the presence of a potential $V$, we define the **Lagrangian** by

$$L(x, v) = \tfrac{1}{2}mv \cdot v - V(x).$$

The **momentum** is given by $p = \nabla_v L(x, \dot{x})$ and $\dot{p} = -\nabla V(x)$ is the Euler–Lagrange equation for the functional $\int L(x(t), \dot{x}(t))\, dt$. The **energy** is given by

$$H = p \cdot \dot{x} - L(x, \dot{x}) = \tfrac{1}{2m}|p|^2 + V(x).$$

**Relativistic mechanics:** (for free particle, i.e. $V \equiv 0$) Define the Lagrangian by

$$L(x, v) = -mc^2\sqrt{1 - |v|^2/c^2}.$$

- analogue of arc length integral, but with Lorentz invariance
- constant $mc^2$ chosen to recover classical mechanics if $|v| \ll c$ (cf. Taylor expansion)

The momentum is

$$p = \nabla_v L(x, \dot{x}) = \tfrac{m\dot{x}}{\sqrt{1-|\dot{x}|^2/c^2}},$$

and $\dot{p} = 0$ is the Euler–Lagrange equation for the functional $\int L(x(t), \dot{x}(t))\, dt$. The **energy** is given by

$$H = p \cdot \dot{x} - L(x, \dot{x}) = c\sqrt{|p|^2 + m^2c^2}.$$

**Quantisation:** To get the "quantum version" we make the following replacements:

$$i\hbar\partial_t = H, \quad p = -i\hbar\nabla.$$

Formally this becomes

$$i\hbar\partial_t = c\sqrt{(-i\hbar\nabla) \cdot (-i\hbar\nabla) + m^2c^2} = c\sqrt{-\hbar^2\Delta + m^2c^2}.$$

Squaring both sides:

$$-\hbar^2\partial_{tt} = c^2(-\hbar^2\Delta + m^2c^2).$$

Thus we derive the **Klein–Gordon equation**

$$\tfrac{1}{c^2}\partial_{tt}\Psi - \Delta\Psi + \tfrac{m^2c^2}{\hbar^2}\Psi = 0. \qquad \text{(KG)}$$

Evidently squaring loses something vital, because KG does not accurately model the electron.

(However, it can be used to accurately model "spinless pions".)

**Dirac equation.** The idea of Dirac was to use an equation of the form

$$i\hbar\partial_t\Psi = H\Psi,$$

where $H$ was a **first order** differential operator satisfying

$$H^2 = c^2\big[-\hbar^2\Delta + m^2c^2\big]. \quad (*)$$

Writing

$$H = -ic\hbar(\alpha_1\partial_1 + \alpha_2\partial_2 + \alpha_3\partial_3) + mc^2\beta$$

one finds

$$H^2 = -c^2\hbar^2[\sum_{j=1}^{3}\alpha_j^2\partial_j^2 + (\alpha_1\alpha_2+\alpha_2\alpha_1)\partial_1\partial_2 + (\alpha_1\alpha_3+\alpha_3\alpha_1)\partial_1\partial_3 + (\alpha_2\alpha_3+\alpha_3\alpha_2)\partial_2\partial_3)]$$
$$- ic^3\hbar m[(\alpha_1\beta+\beta\alpha_1)\partial_1 + (\alpha_2\beta+\beta\alpha_2)\partial_2 + (\alpha_3\beta+\beta\alpha_3)\partial_3] + m^2c^4\beta^2.$$

Imposing $(*)$, one finds the following conditions need to be satisfied:

$$\begin{aligned}
&\alpha_j\alpha_k + \alpha_k\alpha_j = 0 \quad (j \neq k),\\
&\alpha_k\beta + \beta\alpha_k = 0,\\
&\alpha_j^2 = 1, \quad \beta^2 = 1.
\end{aligned}$$

While this is impossible to achieve if $\alpha_j, \beta \in \mathbb{C}$, it is possible if they are matrices!

This is now a purely algebraic problem, a solution to which is given by the following $4\times 4$ matrices:

$$\alpha_j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix},$$

where $I_2$ is the $2\times 2$ identity matrix and $\sigma_j$ are the $2\times 2$ **Pauli matrices** are defined by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note that for the PDE to make sense, we must have $\Psi : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{C}^4$.

We arrive at the **Dirac equation** for a free particle of mass $m$:

$$i\hbar\partial_t\Psi = -i\hbar c\sum_{j=1}^{3}\alpha_j\partial_j\Psi + mc^2\beta\Psi. \qquad \text{(D)}$$

An alternate form is follows. Defining the **Dirac matrices**

$$\gamma_0 = \beta, \quad \gamma_j = \beta\alpha_j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} \quad (j = 1, 2, 3)$$

and noting $\beta^2 = I$, we can multiply (D) on the left by $\beta$ and arrive at the following:

$$i\hbar\left[\tfrac{1}{c}\gamma_0\partial_t\Psi + \sum_{j=1}^{3}\gamma_j\partial_j\psi\right] = mc\Psi, \qquad \Psi : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{C}^4.$$

(This is often written more succinctly as $i\hbar\gamma^\mu\partial_\mu\Psi = mc\Psi$, but we have not introduced relativistic tensor notation or the Einstein summation convention.)

The Dirac equation provides one of the best models for certain types of elementary particles (including electrons, quarks, ...).