CHAPTER 1

Overview

What are our goals in studying PDEs? To understand the underlying physical (or more generally scientific) phenomena. To understand the behavior of solutions. To devise better numerical methods for solving them or guarantee the accuracy of existing ones.

The goals of mathematicians studying PDEs are to:

: Compute exact solutions (mostly done already, rarely done by mathematicians nowadays).
: Prove that solutions exist.
: Prove that some equation is well-posed (solutions exist, are unique, depend continuously on initial data, i.e. to small changes in the initial/boundary data there correspond small changes in the solutions in finite time; regularity is preserved).
: Understand the long-term behavior of solutions: dissipation, dispersion, solitons or solitary waves (e.g. the interaction between solitons), blow-up and singularity formation (i.e. the solution cannot be continued past a certain time); shocks (i.e. the solution becomes nonsmooth); turbulence (as time goes on, the solution becomes less and less smooth, with energy moving to the higher Fourier modes); stability or instability (e.g. stable or centre-stable manifolds); asymptotic behavior (e.g. does the solution converge to an equilibrium state? in what sense?).
: Devise algorithms for computing the solutions.

Connections with other fields:

**Intersection with Analysis:** Most of the techniques used in the study of PDEs come from Analysis. This will be the main topic of this course.

**Intersection with Geometry:** Equations can be formulated on flat (Euclidean) space, but can just as well be formulated on manifolds or bounded domains, whose geometries or whose boundaries’ geometries one must take into account. In addition, the solutions themselves may have a geometric structure (e.g. vortices).

In addition, one can use PDE techniques to study differential manifolds. Recently Perelman used PDE techniques (the Ricci flow) to prove Thurston’s geometrization conjecture.

**Intersection with Probability:** One is sometimes interested in the stability of solutions under stochastic noise. The domain of the equation can have random features (i.e. a random background). Instead of deterministic initial data, one may consider a random ensemble of initial data and try to figure out properties of the “average” or “typical” ensuing solution.

One very important topic is the connection with Brownian motion. This is given by the Feynman–Kac formula.

There are other connections, too.

**Intersection with other sciences:** Many important scientific problems and theories are formulated in terms of PDEs. This has been especially true since Maxwell’s equations of electromagnetism.

Some important PDEs formulated before that were Fourier’s heat equation, Euler’s equation, the Navier–Stokes equations. Important PDEs formulated after that were Boltzmann’s equation, Einstein’s GR equations, Schrödinger’s equation, Dirac’s equation, etc..

Outside of Physics and Engineering, other sciences have started adopting PDEs mostly since the 70s, as computers made it possible to numerically compute complicated solutions. The most
successful example is the Black–Scholes equation, which is widely used in Finance. Other uses include: in Biology, the study of population dynamics, infections, epidemics, blood flow; in Chemistry, attempting to understand chemical reactions from the ground up, from Schrödinger’s equation.
CHAPTER 2

Introduction

DEFINITION 1. A PDE is an equation involving a function of several variables and its partial derivatives.

The general form of a PDE is

\[ F(x, u, Du, D^2u, \ldots, D^n u) = 0, \]

where \( u = u(x), \ x \in \mathbb{R}^d \ d \geq 2, \) and \( D^k u \) is a schematic way of writing together all partial derivatives of order \( k \) (it’s a \( k \)-dimensional array, in other words a \( k \)-tensor).

If the equation involves a function of just one variable and its ordinary derivatives, it’s called an ODE.

1. What is the domain of a solution?

It can be:

- The Euclidean space \( \mathbb{R}^d \) or \( \mathbb{R}^{d+1} \) (written this way to symbolize \( d \) space and 1 time dimensions)
- Or \( C^n \)
- A manifold such as the sphere \( S^d = \{ x \in \mathbb{R}^{d+1} : |x| = 1 \} \) or torus \( \mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d \)
- Some bounded domain in \( \mathbb{R}^d \) (examples: a ball \( B(x_0, R) \), the interior of a cylinder \( B(x_0, R) \times [0, T] \), etc.)
- The complement of such a domain (e.g. the exterior of a ball); called exterior problems.

2. What kind of values does a solution take?

Usually it’s real-valued, sometimes it’s complex-valued, e.g. Schrödinger’s equation.
More rarely it takes values in some target manifold (e.g. a value is a point on the sphere). See: Yang–Mills.

3. What is a solution?

There exist several kinds of solutions: classical, strong, mild, weak/distributional, viscosity solutions.

Classical solutions are straightforward to define, see below. For the other kinds of solutions, the definition is both more complicated and more particular, depending on the individual equation.


DEFINITION 2. For a general PDE of the form (1)

\[ F(x, u, Du, D^2u, \ldots, D^n u) = 0, \]

\( u \) is said to be a classical solution if it and all its partial derivatives up to order \( n \) are continuous and the identity (1) holds at each point in the domain of \( u \).

DEFINITION 3. \( C^k(D) \) is the space of functions on some domain \( D \) such that the function itself and all its derivatives up to order \( k \) are continuous on \( D \). When \( k = 0 \) the notation for the set of continuous functions on \( D \) is \( C(D) \). \( C^{\infty}(D) \) means infinitely many continuous derivatives on \( D \).
Consider for example one of the simplest possible PDEs: $u : \mathbb{R}^2 \rightarrow \mathbb{R}, u = u(x,t)$, 

$$u_t - u_x = 0.$$ 

(2)

A classical solution for equation (2) is a function $u \in C^1(\mathbb{R}^2)$ (or $\mathbb{R}^{1+1}$, to distinguish time and space) such that for every $(x,t) \in \mathbb{R}^{1+1}$

$$\frac{\partial u}{\partial t}(x,t) - \frac{\partial u}{\partial x}(x,t) = 0.$$ 

Example: $x + t$ is a classical solution. Also, $g(x + t)$ is a classical solution for any $g \in C^1(\mathbb{R})$. As we’ll see below, these are all the classical solutions for this equation.

3.2. Weak solutions. Now we set out to weaken the definition of a solution in order to make it more general. In particular, we won’t require a weak solution to be differentiable or even continuous (or even a function, eventually).

The general principle is that a weak/distributional solution fulfills the equation when tested against a test function.

**Definition 4.** The space $\mathcal{D}(\mathbb{R}^d)$ of test functions is the space of smooth (i.e. $C^\infty$, with infinitely many continuous derivatives) functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ with compact support within $\mathbb{R}^d$, i.e. which are zero outside some bounded set.

$$\mathcal{D}(\mathbb{R}^d) = \{ f \in C^\infty(\mathbb{R}^d) : \exists r \geq 0 \forall x \in \mathbb{R}^d (|x| \geq r \implies f(x) = 0) \}.$$ 

Test functions are also called bump functions. Nonzero test functions exist. Particularly important for us are smooth cutoff functions, which are a special class of test functions. See this Appendix for more details on both.

Formally, we replace the equation with the following condition: for every $f \in \mathcal{D}(\mathbb{R}^{1+1})$

$$\int_{\mathbb{R}^{1+1}} (u_t - u_x)f \, dx \, dt = 0,$$

so

$$\int_{\mathbb{R}^{1+1}} u_t f \, dx \, dt - \int_{\mathbb{R}^{1+1}} u_x f \, dx \, dt = 0.$$ 

Still formally, we do an integration by parts under each integral: (supposing that $f$ is supported inside some rectangle $[a,b] \times [c,d]$)

$$- \int_{\mathbb{R}^{1+1}} u f_t \, dx \, dt + \int_{\mathbb{R}^{1+1}} u f_x \, dx \, dt = 0.$$ 

Up to this point this was a formal computation, because $u$ need not be a differentiable function. However, now $f$ is a test function, so the last expression we got makes sense anyway (note that if $f$ is a test function then its derivatives are test functions as well). We take it as the definition:

**Definition 5.** A locally integrable function $u \in L^1_{loc}(\mathbb{R}^{1+1})$ is called a weak solution of equation (2) if and only if for every test function $f \in \mathcal{D}(\mathbb{R}^{1+1})$

$$\int_{\mathbb{R}^{1+1}} u(-f_t + f_x) \, dx \, dt = 0.$$ 

(3)

Locally integrable = integrable on any bounded domain. Example: $e^x$.

An example of a weak solution which is not a classical solution:

$$u(x,t) = \begin{cases} 
1, & x + t > 0 \\
0, & x + t < 0.
\end{cases}$$ 

The values of $u$ along the diagonal $x + t = 0$ don’t really matter.
3. WHAT IS A SOLUTION?

PROOF. Let $u$ be as above. We have to check that for any $f \in \mathcal{D}$
\[
\int_{\mathbb{R}^{1+1}} u(-f_t + f_x) \, dx \, dt = 0 \iff \int_{x+t>0} (-f_t + f_x) \, dx \, dt = 0 \iff \int_{x+t>0} f_x \, dx \, dt = \int_{x+t>0} f_t \, dx \, dt.
\]
The first integral is
\[
\int_{-\infty}^{\infty} \int_{-t}^{\infty} f_x(x, t) \, dx \, dt = -\int_{-\infty}^{\infty} f(-t, t) \, dt.
\]
The second integral is
\[
\int_{-\infty}^{\infty} \int_{-x}^{\infty} f_t(x, t) \, dt \, dx = -\int_{-\infty}^{\infty} f(x, -x) \, dx.
\]
After the change of variable $x = -t$ it’s clear they are equal. □

Important fact: If a weak solution possesses sufficient regularity, i.e. if $u \in C^1$ in this case, then it is in fact a classical solution.

PROOF. Suppose $u \in C^1$ is a weak solution of (2), in the sense of (3).

Then we can reverse the formal integration by parts through which we arrived at (3) and get that for every $f \in \mathcal{D}$
\[
\int_{\mathbb{R}^{1+1}} (u_t - u_x) f \, dx \, dt = 0.
\]
Here $u_t - u_x \in C$. Suppose $u_t - u_x \neq 0$. With no loss of generality there exists $(x_0, t_0) \in \mathbb{R}^{1+1}$ such that $(u_t - u_x)(x_0, t_0) > 0$. By continuity, for some $\epsilon > 0$, $(u_t - u_x)(x_0, t_0) > 0$ for every $(x, t) \in B((x_0, t_0), \epsilon)$.

Now consider a positive test function $f \in \mathcal{D}$ supported on this neighborhood of $(x_0, t_0)$. We get a contradiction. Hence $u_t - u_x \equiv 0$, so $u$ is a classical solution. □

Conclusions: It is natural to consider weak solutions because discontinuities are natural. Discontinuities are called shocks and appear in the study of certain physical phenomena, such as actual shock waves.

It is important to keep track of the regularity of weak solutions. Usual procedure: construct weak solutions; prove that if the initial data are smooth then the solutions remain sufficiently smooth; upgrade them to classical solutions.

More about various solution types later.

PROBLEM 1. Consider the equation on $\mathbb{R}^{1+1}$
\[
 u_t - 2xu_x = 0.
\]
We define $u$ to be a weak solution of this equation if for all $f \in \mathcal{D}$
\[
\int_{\mathbb{R}^{1+1}} u(-f_t + 2xf_x + 2f) \, dx \, dt = 0.
\]
Prove that if $u \in C^1$ is a weak solution then it is also a classical solution.

PROBLEM 2. Consider the one-dimensional wave equation on $\mathbb{R}^{1+1}$
\[
 u_{tt} - u_{xx} = 0.
\]
We define $u(x, t)$ to be a weak solution of this equation on the domain $\mathbb{R} \times (0, \infty)$ (i.e. $x \in \mathbb{R}$, $t > 0$) if for any test function $f \in \mathcal{D}(\mathbb{R} \times (0, \infty))$
\[
\int_{0}^{\infty} \int_{\mathbb{R}} u(f_{tt} - f_{xx}) \, dx \, dt = 0.
\]
Prove that the function \( u : \mathbb{R} \times (0, \infty) \to \mathbb{R} \),

\[
u(x, t) = \begin{cases} 
3, & |x| < t \\
0, & |x| > t,
\end{cases}
\]
is such a solution.

4. Classification of PDEs

4.1. By order. By definition, the order of a PDE is the order of the highest-order derivative present in the equation.

1: First-order: also known as transport equations. Examples:
Equation (2): \( u_t - uu_x = 0 \)
Burgers’ equation: \( u_t - uu_x = 0 \)
Euler’s equation: (quite hard)

\[
\partial_t \vec{v} + (\vec{v} \cdot \nabla) \vec{v} = -\frac{1}{\rho} \nabla p.
\]
Here \( \vec{v} \) is the fluid velocity, \( \rho \) is the density, and \( p \) is the pressure.

2: Second-order:

A general linear second-order equation: \( u : \Omega \to \mathbb{R} \),

\[
\sum_{j,k=1}^{d} a_{jk}(x) \partial_j \partial_k u + \sum_{j=1}^{d} b_j(x) \partial_j u + c(x)u = f, \ u |_{\partial\Omega} = g.
\]
The coefficients may be: merely measurable functions, continuous functions, smooth functions, etc..

With constant coefficients:

\[
\sum_{j,k=1}^{d} a_{jk} \partial_j \partial_k u + \sum_{j=1}^{d} b_j \partial_j u + cu = f, \ u |_{\partial\Omega} = g.
\]

Even with variable coefficients, it matters to first understand the constant coefficient case. The reason is that, after zooming in onto a very small neighborhood of any given point in the domain of the solution, the equation reduces to constant coefficients + small perturbation and can sometimes be treated as such. This is called the “freezing coefficients” method and works best for elliptic and parabolic equations.

A very important operator is the Laplacian, more generally called the Laplace–Beltrami operator: in \( \mathbb{R}^d \)

\[
\Delta = \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_d^2} = \text{div } \nabla.
\]
For this reason physicists usually denote it by \( \nabla^2 \).

In polar coordinates:

\[
\Delta = \partial_r^2 + \frac{d-1}{r} \partial_r + \frac{1}{r^2} \Delta_\omega.
\]
The last term is the spherical Laplacian, depending only on the tangential derivatives; e.g. in \( \mathbb{R}^2 \) it is \( \partial_\omega^2 \). For radially/spherically symmetric functions, the last term vanishes.

Main types: elliptic, parabolic, hyperbolic.

Elliptic: Some examples:
Laplace: \(-\Delta u = 0\)
Poisson: \(-\Delta u = f\)
Yamabe: \(-\Delta u = u^{\frac{d+2}{2}}\)
Time-independent Schrödinger equation: \(-\Delta u + V(x)u = 0\).
Here \( u : D \to \mathbb{R} \), where \( D = \mathbb{R}^d \), \( \Omega \subset \mathbb{R}^d \), \( S^d \), etc. (domains of all shapes and sizes).

A general linear elliptic equation with constant coefficients:

\[
\sum_{j,k=1}^{d} a_{jk} \partial_j \partial_k u + \sum_{j=1}^{d} b_j \partial_j u + cu = f, \ u |_{\partial \Omega} = g.
\]

The condition that ensures this equation is elliptic is that for any \( \xi \in \mathbb{R}^d \) with \( |\xi| = 1 \)

\[
\sum_{j,k=1}^{d} a_{jk} \xi_j \xi_k > 0.
\]

Nonconstant:

\[
\sum_{j,k=1}^{d} a_{jk}(x) \partial_j \partial_k u + \sum_{j=1}^{d} b_j(x) \partial_j u + c(x)u = f, \ u |_{\partial \Omega} = g
\]

The condition that ensures this equation is elliptic is that for any \( \xi \in \mathbb{R}^d \) with \( |\xi| = 1 \) and any \( x \in D \)

\[
\sum_{j,k=1}^{d} a_{jk}(x) \xi_j \xi_k > 0.
\]

**Heat/diffusion:** \( u_t = \Delta u \).

**Black–Scholes:**

\[
\frac{\partial v}{\partial t} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} + r s \frac{\partial v}{\partial s} - rv = 0.
\]

Here \( r \) is the interest rate and \( \sigma \) is related to the volatility.

**Navier–Stokes:**

\[
\partial_t \vec{v} + \vec{v} \cdot \nabla \vec{v} - \Delta \vec{v} = -\frac{1}{\rho} \nabla p.
\]

**Wave:**

\[
u_{tt} - \Delta u = F.
\]

**Klein–Gordon:**

\[
u_{tt} - \Delta u + u = F.
\]

**Harder:** Einstein’s GR equations are usually hyperbolic equations. In vacuum:

\[
R_{\mu\nu} = 0.
\]

In general:

\[
R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8 \pi G}{c^4} T_{\mu\nu}.
\]

However, there are second-order equations that fall into none of the above categories.

**Examples:**

Schrödinger: \( u_t = -i \Delta x u \) or \( u_t = i(-\Delta + V)u \). Here \( i^2 = -1 \) and \( u \) is complex-valued: \( u : \mathbb{R}^{d+1} \to \mathbb{C} \) (or \( u : \Omega \times \mathbb{R} \to \mathbb{C} \)).

Examples with real coefficients: \( u_t + u_{xx} - u_{yy} = 0 \).

In even more dimensions: ultra-hyperbolic.

**3:** Third-order: Korteweg–de Vries (KdV): \( u_t + u_{xxx} - 6uu_x = 0 \).

**4:** Fourth-order: The beam/plate equation: \( c^2 u_t + \Delta^2 u = F \). Beam in \( \mathbb{R}^{1+1} \), plate in \( \mathbb{R}^{2+1} \). \( \Delta^2 \) is called the bilaplacian. Schrödinger’s equation implies the beam/plate equation.
Burgers’ equation and Euler’s equation are also quasilinear.

Example: \( u \)

Quasilinear: same as above, but the coefficients of the highest-order derivatives are also allowed to be functions of the lower-order ones or of the solution.

Semilinear: The equation is linear as a function of the highest-order derivatives, but contains a nonlinear term depending only on derivatives of strictly lower order or on the solution.

Examples: Einstein’s equations of General Relativity

\[ R_{\mu\nu} = 0. \]

Burgers’ equation and Euler’s equation are also quasilinear.
3: Fully nonlinear: The equation is nonlinear in the highest-order derivatives also. Example: The Monge–Ampère equation is a fully nonlinear equation of elliptic type:
\[
\det D^2 f - K(x)(1 + |\nabla u|^2)^{(d+2)/2} = 0.
\]

**Problem 7.** Classify the following equations as linear, semilinear, quasilinear, or fully nonlinear:

- \( a_t + e^x u_x - y u_y = 0 \)
- \( u_{tt} - u^2 u_x = 0 \)
- \( u + u_t u_x = 0 \)
- \( u_{tt} - u^2 u_{xx} = 0 \)

4.3. Evolution equations versus static/time-independent equations: Evolution equations describe a system that changes with time, i.e. there is one special direction (time) which has different properties from the other directions (space). Example: everything except elliptic equations.

Time-independent equations describe an equilibrium state. Example: elliptic equations.

For elliptic equations we have boundary value problems. Example of a boundary value problem: Poisson equation with Dirichlet boundary data: for \( u : \Omega \to \mathbb{R} \)
\[
-\Delta u = f, \quad u \mid_{\partial\Omega} = g.
\]

For evolution equations we have initial value problems / Cauchy problems. Lateral boundary data are also important if a lateral boundary is present. Example of an initial value problem: Heat equation on a cylinder: for \( u : \Sigma = \Omega \times [0, T] \to \mathbb{R}, \)
\[
u_t - \Delta u = F, \quad u(x,0) = u_0(x), \quad u \mid_{\Sigma = \partial\Omega \times [0, T]} = g.
\]

Observation: \( \Sigma = \Omega \times [0, T] \) is called a (right) cylinder regardless of the shape of its base. If \( \Omega \) is a disk/ball, then it’s called a circular right cylinder.

If there is no lateral boundary, e.g. if \( \Omega = \mathbb{R}^d \), then the lateral boundary data are replaced with some condition at infinity, typically asking that the solution goes to zero at infinity. Example: for \( u : \mathbb{R}^{d+1} \to \mathbb{R} \)
\[
u_t - \Delta u = F, \quad u(x,0) = u_0(x), \quad u \text{ vanishes at spatial infinity}.
\]

This is only an empirical classification. The corresponding rigorous question is: for each given equation, what sort of initial/boundary data does one need to specify for the equation to be well-posed?

4.4. Constant versus non-constant coefficients. A general linear second-order equation: \( u : \Omega \to \mathbb{R}, \)
\[
\sum_{j,k=1}^d a_{jk}(x) \partial_j \partial_k u + \sum_{j=1}^d b_j(x) \partial_j u + c(x) u = f, \quad u \mid_{\partial\Omega} = g.
\]

The coefficients may be: merely measurable functions, continuous functions, smooth functions, etc..

With constant coefficients:
\[
\sum_{j,k=1}^d a_{jk} \partial_j \partial_k u + \sum_{j=1}^d b_j \partial_j u + cu = f, \quad u \mid_{\partial\Omega} = g.
\]

Even if we are dealing with variable/nonconstant coefficients, it’s still worth it to first understand the nonconstant coefficient case first; see the coefficient freezing method.

Equations with nonconstant coefficients are in general significantly harder than the corresponding equations with constant coefficients. **Example:** Poisson equation with Dirichlet boundary data: for \( u : \Omega \to \mathbb{R} \)
\[
-\Delta u = f.
\]
2. INTRODUCTION

A general elliptic equation with nonconstant coefficients:

$$\sum_{j,k=1}^{d} -a_{jk}(x) \partial_{j} \partial_{k} u + \sum_{j=1}^{d} b_{j}(x) \partial_{j} u + c(x) u = f.$$  

The condition that ensures ellipticity (i.e. makes the second equation like the first) is that for any $\xi \in \mathbb{R}^{d}$ with $|\xi| = 1$ and $x \in \mathbb{R}^{d}$

$$\sum_{j,k=1}^{d} a_{jk}(x) \xi_{j} \xi_{k} > 0.$$  

5. Types of boundary conditions

This list is non-exhaustive, but the following are the four most studied types.

**Dirichlet:** $u \mid_{\partial D} = f$. The most straightforward boundary condition. The physical interpretation is obvious: e.g. we measure and specify the boundary data as measured.

**Neumann:** prescribing the normal derivative on the boundary:

$$\frac{du}{d\nu} = g.$$  

Here $\frac{du}{d\nu} = \frac{du}{dn}$ is the normal derivative, i.e. the derivative in the normal/perpendicular direction on the boundary.

The Neumann condition indicates the flow of the solution through the boundary, e.g. the boundary being completely insulated corresponds to a zero/homogenous Neumann boundary condition.

Note that this Neumann is not the same as the famous von Neumann (father of cybernetics, among others).

**Robin:** On $\partial D$, $a(x)u + b(x) \frac{du}{d\nu} = c(x)$. This is a mixture of the previous two conditions and also has some sort of physical significance. It’s mathematically interesting.

**Periodic:** This condition is more important mathematically, as it lends itself to a specific solving method: Fourier series.

Periodic boundary conditions are those that allow the solution to be extended by periodicity to the whole of $\mathbb{R}^{d}$.

Examples of periodic boundary conditions: The heat equation on a line segment: for $u : [a, b] \times [0, \infty) \to \mathbb{R}$, the equation is

$$u_t - u_{xx} = 0, \quad u(x, 0) = u_0(x).$$

The periodic boundary condition is $u(a, t) = u(b, t), \quad u_x(a, t) = u_x(b, t)$.

On a rectangle: To make it more concrete, consider the rectangle $D = [0, a] \times [0, b]$. The heat equation on $D \times [0, \infty)$ is

$$u_t - \Delta u = 0, \quad u(x, 0) = u_0(x).$$

The periodic boundary condition in this case is

$$u(0, y, t) = u(a, y, t), \quad u_x(0, y, t) = u_x(a, y, t), \quad u(x, 0, t) = u(x, b, t), \quad u_y(x, 0, t) = u_y(x, b, t).$$

If this is true, then we can extend the solution by periodicity to $\mathbb{R}^{2+1}$.

Physically, periodic boundary conditions can correspond to e.g. an equation set up on a circle or on the surface of a cylinder or an equation set up inside some large, regular lattice, such as a crystal.
CHAPTER 3

Separation of variables

Caveat: There is almost no general algorithm for solving PDEs. More general methods, classifications, and rules are less useful when dealing with specific equations. Each equation is studied in a specific way. There are many different methods for studying and solving PDEs. This goes double for nonlinear equations.

For linear equations with constant coefficients, on sufficiently regular domains, separation of variables always works.

1. Examples

1.1. Heat equation on a line. The heat equation on $\mathbb{R}^{1+1}$:

$$u_t - u_{xx} = 0, \ u(x, 0) = u_0(x) = e^{-\frac{x^2}{2}}.$$ (4)

Here we’ll look for a solution, without worrying whether it’s unique or not. In fact, solutions are not unique: there are other classical solutions to (4) than the one we’ll get below.

Fundamental idea: separation of variables. Suppose that we could write the solution in the form $u(x, t) = f(x)g(t)$. Then

$$u_t = u_{xx}, \ u_t = f(x)g'(t), \ u_{xx} = f''(x)g(t).$$

So

$$f(x)g'(t) = f''(x)g(t), \ \frac{g'(t)}{g(t)} = \frac{f''(x)}{f(x)}.$$ 

Since these expressions are equal, but one is a function only of $x$ and the other one only of $t$, this means they are both constant.

So we have to solve the equations $g'(t) = \lambda g(t)$ and $f''(x) = \lambda f(x)$.

The first equation has the solution $g(t) = g_0e^{\lambda t}$. From physical considerations (since heat decreases; this is just avoiding a more complicated mathematical answer), we expect $\lambda \leq 0$, so set $\lambda = -\mu^2$.

Then the $x$ equation has solutions

$$f(x) = e^{\pm i\mu x}$$

for $\lambda < 0$, $\mu \neq 0$ and $f(x) = 1, x$ for $\lambda = \mu = 0$.

Another way of writing the solution in the first case would have been

$$f(x) = A \sin(\mu x) + B \cos(\mu x).$$

These two representations are equivalent, due to Euler’s formula:

Thus, for $\mu \in \mathbb{R}$, $e^{-\mu^2 t}e^{i\mu x}$ are solutions to the heat equation. If we could decompose the initial data into some combination of $e^{i\mu x}$, then we could also solve the equation: for

$$u_0(x) = \int_{-\infty}^{\infty} \hat{u}_0(\mu) e^{i\mu x} \, d\mu$$
the corresponding solution will be (check)

\[ u(x, t) = \int_{-\infty}^{\infty} \hat{u}_0(\mu) e^{-\mu^2 t} e^{i \mu x} d\mu. \]

Note: there are several conventions for (ways of defining) the Fourier transform; this is one of them. They only differ by minor details, such as factors of 2π.

At this point, we can retroactively justify why we made the assumption \( \lambda \leq 0 \) and why we discarded the second solution \( f(x) = x \) for \( \lambda = 0 \). The reason is that they were not needed. Those separable solutions we kept are sufficient to represent the initial data.

The function \( \hat{u}_0 \) is called the Fourier transform of \( u_0 \) and is given by the formula

\[ \hat{u}_0(\mu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\mu} u_0(x) dx. \]

The formula that expresses \( u_0 \) in terms of \( \hat{u}_0 \) is called the Fourier inversion formula. The transformation that gives \( u_0 \) in terms of \( \hat{u}_0 \) is called the inverse Fourier transform.

At this point, let’s state the formal definition of the Fourier transform:

**Definition 6.** For a function \( f : \mathbb{R} \to \mathbb{C} \)

\[ \hat{f}(\mu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)e^{-ix\mu} dx. \]

The inverse Fourier transform, which allows one to get \( f \) back from \( \hat{f} \), is given by the formula

**Proposition 1.1 (Fourier inversion formula).**

\[ f(x) = \int_{-\infty}^{\infty} \hat{f}(\mu)e^{ix\mu} d\mu. \]

Standard notations: \( \hat{f} \) for the Fourier transform, \( \check{g} \) for the inverse Fourier transform. If \( \hat{f} = g \), then \( \check{g} = f \). Other standard notations for the Fourier transform: \( \mathcal{F} f = g, \mathcal{F}^{-1} g = f \).

Note that the formulas defining the Fourier transform and the inverse Fourier transform are very similar. Hence, the following connection between them:

\[ \check{\hat{f}}(x) = 2\pi \hat{f}(-x). \]

**Fact:** The Fourier transform of \( e^{-\frac{x^2}{2}} \) is \( \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \). In general, the Fourier transform of a Gaussian is a Gaussian:

\[ \overline{e^{-ax^2}} = \frac{1}{\sqrt{4a\pi}} e^{-\frac{\mu^2}{4a}}. \]

Explicitly, this is stating that for each \( \mu \in \mathbb{R} \) (in fact, also for complex \( \mu \))

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ax^2} e^{-ix\mu} dx = \frac{1}{\sqrt{4a\pi}} e^{-\frac{\mu^2}{4a}}. \]

We’ll prove it later.

The inverse Fourier transform is then

\[ \mathcal{F}^{-1} e^{-ax^2} = 2\pi \frac{1}{\sqrt{4a\pi}} e^{-\frac{(x-a)^2}{4a}} = 2\pi \frac{1}{\sqrt{4a\pi}} e^{-\frac{\mu^2}{4a}}. \]

Finally, we can now solve the equation (4):

1. We take the Fourier transform of the initial data, \( \hat{u}_0(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2}}. \)
2. We multiply it by \( e^{-\mu^2 t} \), getting \( \frac{1}{\sqrt{2\pi}} e^{-\mu^2 \left( \frac{1}{2} + t \right)} \).

3. We (inverse) Fourier transform it back, getting

\[
u(x, t) = \frac{2\pi}{\sqrt{2\pi \sqrt{(2 + 4t)^2}}} e^{-\frac{|\mu|^2}{2+4t}} = \frac{1}{\sqrt{1+2t}} e^{-\frac{|\mu|^2}{2+4t}}.
\]

**The general formula** for solving equation (4) that we got from this computation is

\[ u(x, t) = \mathcal{F}^{-1} [e^{-i\mu^2 \mu_0(\mu)}] . \]

**Problem 8.** Using the same method, find a solution of the one-dimensional heat equation on \( \mathbb{R}^{1+1} \)

\[ u_t = u_{xx}, \quad u(x, 0) = u_0(x) \]

for the following initial data:

a) \( u_0(x) = e^{-x^2} \)

b) \( u_0(x) = e^{-\alpha x^2}, \alpha > 0 \). What happens when \( \alpha = 0 \)?

**1.2. The heat equation in higher dimensions.** Using the same method, we solve the heat equation in \( d \) dimensions

\[ u_t = \Delta_x u, \quad u(x, 0) = u_0(x) = e^{-|x|^2/2}. \]

For concreteness, let \( d = 2 \). Let’s start with separation of variables. We look for solutions of the form

\[ u(x_1, x_2, t) = f_1(x_1) f_2(x_2) g(t) . \]

Plugging this \( u \) into the equation, we get

\[ f_1 f_2 g' = f_1'' f_2 g + f_1' f_2'' g. \]

Dividing by \( f_1 f_2 g \), we get

\[ \frac{g'}{g}(t) = \frac{f_1''}{f_1}(x_1) + \frac{f_2''}{f_2}(x_2). \]

Since all three variables \( x_1, x_2, \) and \( t \) are independent of each other, we get that all three fractions are constant:

\[ \frac{g'}{g}(t) = \lambda, \quad \frac{f_1''}{f_1}(x_1) = \lambda_1, \quad \frac{f_2''}{f_2}(x_2) = \lambda_2, \]

where \( \lambda = \lambda_1 + \lambda_2 \).

Thus, in order to find separable solutions, we need to solve the system

\[ \frac{g'}{g} = \lambda g, \quad \frac{f_1''}{f_1} = \lambda_1 f_1, \quad \frac{f_2''}{f_2} = \lambda_2 f_2, \quad \lambda = \lambda_1 + \lambda_2 . \]

As before, we are not interested in enumerating all possible separable solutions, but only in getting enough of them to express the initial data as a combination of them. With the assumptions

\[ \lambda_1 = -\mu_1^2 \leq 0, \quad \lambda_2 = -\mu_2^2 \leq 0 \],

we get that

\[ g = g_0 e^{-\left(\mu_1^2 + \mu_2^2\right)t}, \quad f_1 = A e^{i\mu_1 x_1} + B e^{-i\mu_1 x_1}, \quad f_2 = C e^{i\mu_2 x_2} + D e^{-i\mu_2 x_2} \]

with the necessary modifications when one of the \( \mu \)'s is zero.

We retain the following family of separable solutions:

\[ u(x_1, x_2, t) = e^{i(\mu_1 x_1 + \mu_2 x_2) - (\mu_1^2 + \mu_2^2)t} . \]

Again, this is not a complete list of separable solutions; see previous example.

If we could represent the initial data as a superposition of them, i.e. at time \( t = 0 \) for some coefficients \( c(\mu_1, \mu_2) \)

\[ u_0(x_1, x_2) = \int_{\mathbb{R}^2} c(\mu_1, \mu_2) e^{i(\mu_1 x_1 + \mu_2 x_2)} d\mu_1 d\mu_2. \]
then the same superposition would give us a formula valid for all $t$:

$$ u(x_1, x_2, t) = \int_{\mathbb{R}^2} c(\mu_1, \mu_2)e^{i(\mu_1 x_1 + \mu_2 x_2) - (\mu_1^2 + \mu_2^2)t} \, d\mu_1 \, d\mu_2. $$

These coefficients $c(\mu_1, \mu_2)$ are, by definition, again nothing but the Fourier transform of $u_0$: $c(\mu_1, \mu_2) = \hat{u}_0(\mu_1, \mu_2)$.

In $\mathbb{R}^d$, the Fourier transform has almost exactly the same definition as in one dimension:

**Definition 7.** For a function $f : \mathbb{R}^d \to \mathbb{C}$

$$ \hat{f}(\vec{\mu}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} f(\vec{x})e^{-i\vec{x} \cdot \vec{\mu}} \, d\vec{x}. $$

Here $\vec{x} = (x_1, \ldots, x_d)$, $\vec{\mu} = (\mu_1, \ldots, \mu_d)$.

The Fourier inversion formula in $\mathbb{R}^d$ is

$$ f(\vec{x}) = \int_{-\infty}^{\infty} \hat{f}(\vec{\mu})e^{i\vec{x} \cdot \vec{\mu}} \, d\vec{\mu}. $$

Going through the computations, we see that

$$ u(\vec{x}, t) = \mathcal{F}^{-1}[e^{-t/|\vec{x}|^2}\hat{u}_0(\vec{\mu})]. $$

**Fact:** The Fourier transform of a Gaussian in $\mathbb{R}^d$ is

$$ [\mathcal{F}e^{-a|\vec{x}|^2}](\mu) = \frac{1}{(4\pi a)^{d/2}} e^{-\frac{|\mu|^2}{4a}}. $$

Using this formula, we again go through the same steps as in the previous example:

1. Take the Fourier transform of the initial data:

$$ \hat{u}_0(\vec{\mu}) = \frac{1}{2\pi} e^{-\frac{|\vec{\mu}|^2}{2}}. $$

2. Multiply this by $e^{-t/|\vec{\mu}|^2}$, getting

$$ \frac{1}{2\pi} e^{-\frac{|\vec{\mu}|^2}{2}(1+t)} $$

3. Take the inverse Fourier transform, getting

$$ \frac{(2\pi)^2}{4(1/2 + t)^\pi} \frac{1}{2\pi} e^{-\frac{|\vec{x}|^2}{4(1/2 + t)^2}} = \frac{1}{1 + 2t} e^{-|\vec{x}|^2/(1 + 2t)}. $$

In general, the solution to the heat equation in $\mathbb{R}^d \times [0, \infty)$

$$ u_t = \Delta_x u, \ u(x, 0) = u_0(x) = e^{-|\vec{x}|^2/2} $$

is given by

$$ u(x, t) = \frac{1}{(1 + 2t)^d} e^{-\frac{|\vec{x}|^2}{2 + 4t}}. $$

**Problem 9.** Redo the proof above in dimension $d = 3$, i.e. compute a solution of the heat equation on $\mathbb{R}^{3+1}$

$$ u_t = \Delta_x u, \ u(x_1, x_2, x_3, 0) = e^{-(x_1^2 + x_2^2 + x_3^2)/2}. $$

Start with the separation of variables. Include as much detail as you can.

**Problem 10.** a) Assuming that $f_1(x, t)$ and $f_2(x, t)$ are solutions of the one-dimensional heat equation

$$ f_t = f_{xx}, $$

prove that $u(x_1, x_2, t) = f_1(x_1, t)f_2(x_2, t)$ is a solution of the two-dimensional heat equation

$$ u_t = u_{x_1x_1} + u_{x_2x_2}. $$
b) Use this property and previous problems to find a solution of the two-dimensional heat equation
(5) with initial data
\[ u(x_1, x_2, 0) = u_0(x_1, x_2) = e^{-x_1^2-2x_2^2}. \]

More general linear equations with constant coefficients, on the whole space \( \mathbb{R}^d \) or on some domain \( D \subset \mathbb{R}^d \). Equations (4) and (5) fit in the following category:
\[ u_t = L_x u, \quad u(x, 0) = u_0. \]

Other equations of this type: e.g. Schrödinger’s equation, without and with potential; Poisson’s equation on the upper half-plane (not obvious, has to be proved).

Having a second-order time derivative, the wave equation and Poisson’s equation can be written as
\[ u_{tt} = L_x u, \quad u(x, 0) = u_0(x), \quad u_t(x, 0) = u_1(x). \]

**Remark 1.2.** If everything is analytic, we can use a Wick rotation to get from the heat equation to Schrödinger’s equation.

In general, solving an equation of the form
\[ u_t = Lu, \quad u(x, 0) = u_0 \]
can be done as follows:
1. Find eigenfunctions and eigenvalues of the operator \( L \):
\[ Lf_\lambda = \lambda f_\lambda. \]
This is called an eigenvalue equation. \( \lambda \) are eigenvalues, \( f_\lambda \) are eigenfunctions.

The constant function \( 0 \) is an eigenfunction for every eigenvalue, but not an interesting one.

We are only interested in not-constantly-zero eigenfunctions.

Sometimes the eigenvalues are discrete, sometimes they form a continuum; sometimes there is a mixture of both. Always, there’s an infinite number.

In many cases of interest, the spectrum and the eigenvalues are known.

Having such an eigenfunction as initial data, \( u(x, t) = e^{\lambda t} f_\lambda(x) \) is a separable solution of the equation (6).

2. Decompose the initial data \( u_0 \) in terms of eigenfunctions:
\[ u_0(x) = \int_{\sigma(L)} u_0^\#(\lambda) f_\lambda(x) \, d\lambda. \]
This sort of decomposition (or particular cases of it) is sometimes called the skew-Fourier transform.

3. Then the solution at time \( t \) will be
\[ u(x, t) = \int_{\sigma(L)} u_0^\#(\lambda)e^{\lambda t} f_\lambda(x) \, d\lambda. \]

Here \( \sigma(L) \) is called the **spectrum** of the operator \( L \).

Typical examples where we want to do this:
- a circle
- a line segment
- the interior of a rectangle
- the surface of a sphere
- the interior of a sphere
- a disk
- a cylinder
- an atom, e.g. the hydrogen atom.

Some of these examples are discussed below.

Symmetry helps. Typically, the decomposition depends on the boundary conditions.
On an arbitrary domain: much harder; may try discretizing the equation.
If we understand the operator $L$, we can solve not only the heat equation, but many other equations (Schrödinger, wave, Poisson, etc.) involving that operator.

For the Laplacian, this eigenvalue equation is called the Helmhotlz equation:

$$-\Delta u = \lambda u.$$  

1.3. Heat equation on a segment. The basic example we’ll examine in detail is the heat equation on a line segment, $u : [0, a] \times [0, \infty) \to \mathbb{R}$,

$$u_t = u_{xx}, \quad u(x, 0) = u_0(x),$$  

(7)

with various boundary conditions.

1: On a line segment $[0, a]$, $a > 0$, with homogenous Dirichlet boundary conditions we get a Fourier sine series, see below. $\lambda_n = n \in \mathbb{Z}$, $n > 0$. Here 0 is not an eigenvalue.

Computations: The Dirichlet boundary conditions for equation are that $u(0, t) = u(a, t) = 0$.

Then the operator $L$ is $L = \partial_x^2$ with Dirichlet boundary conditions. Boundary conditions are considered to be part of the operator.

Let’s solve the eigenvalue problem, for $f : [0, a] \to \mathbb{C}$,

$$f''(x) = \lambda f(x), \quad f(0) = f(a) = 0.$$  

We distinguish three cases: $\lambda > 0$, $\lambda = 0$, and $\lambda < 0$.

For $\lambda > 0$, the solution is of the form $f_\lambda(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}$.

The boundary conditions read

$$A + B = 0, \quad A e^{a \sqrt{\lambda}} + B e^{-a \sqrt{\lambda}} = 0,$$

so $B = -A$, so $A(e^{a \sqrt{\lambda}} - e^{-a \sqrt{\lambda}}) = 0$. The parenthesis is strictly positive, so $A = 0$, so $B = -A = 0$ as well. There are no nontrivial solutions.

For $\lambda = 0$, the equation is just $f'' = 0$, so the solution is $f(x) = Ax + B$. Again, plugging this $f$ into the boundary conditions, we get $B = 0, Aa + B = 0$, so $A = B = 0$. No nontrivial solutions here either.

For $\lambda < 0$, the solution is of the form

$$f_\lambda = A \sin(\mu x) + B \cos(\mu x),$$

where $-\lambda = \mu^2$. Plugging this into the boundary conditions, we get

$$B = 0, \quad A \sin(\mu a) = 0.$$  

If $A = 0$ as well, then this is a trivial solution. Hence, for a nontrivial solution we assume $A \neq 0$ and get $\sin(\mu a) = 0$. Here $a$ is fixed and solving the equation for $\mu$ we get the solutions

$$\mu = \frac{\pi n}{a}, \quad n \in \mathbb{Z}.$$  

So the eigenvalues are

$$\lambda_n = -\left(\frac{\pi n}{a}\right)^2,$$

with the associated eigenfunctions $f_n(x) = A \sin\left(\frac{\pi n}{a}x\right)$.

Eigenfunctions are only unique up to a multiplicative constant, so we may discard $A$. Also, $f_{-n}(x) = -f_n(x)$, so we can just keep $n \geq 0$. Finally, for $n = 0$ $\lambda_0 = 0$, which is a case we already considered and discarded.
We end up with the following list of separable solutions:

\[(8) \quad u(x,t) = f_n(x)e^{-\lambda_n t} = \sin \left( \frac{\pi n}{a} x \right) e^{-\left( \frac{\pi n}{a} \right)^2 t}, \quad n \in \mathbb{Z}, \quad n > 0.\]

If we can use some combination of these to match the initial data at time zero, this gives us a formula for the solution. Concretely, if for some coefficients \(c_n\)

\[(9) \quad u_0(x) = \sum_{n>0} c_n f_n(x) = \sum_{n>0} c_n \sin \left( \frac{\pi n}{a} x \right),\]

then a solution is given by

\[u(x,t) = \sum_{n>0} c_n f_n(x)e^{-\lambda_n t} = \sum_{n>0} c_n \sin \left( \frac{\pi n}{a} x \right) e^{-\left( \frac{\pi n}{a} \right)^2 t}.\]

The decomposition (9) is called a sine Fourier expansion/decomposition/series for \(u_0\).

By using this expansion, the Dirichlet boundary conditions are automatically satisfied.

2: With homogenous Neumann boundary conditions we get a Fourier cosine series. Unlike in the Dirichlet case, here 0 is an eigenvalue.

The Neumann boundary conditions for equation (7) are

\[u_x(0,t) = u_x(a,t) = 0.\]

Hence the eigenvalue problem is, for \(f : [0,a] \rightarrow \mathbb{C},\)

\[f''(x) = \lambda f(x), \quad f'(0) = f'(a) = 0.\]

Again we distinguish the three cases \(\lambda > 0, \lambda = 0,\) and \(\lambda < 0.\)

For \(\lambda > 0,\) the solution is of the form

\[f(x) = Ae^{x\sqrt{\lambda}} + Be^{-x\sqrt{\lambda}}.\]

From the boundary conditions we get

\[\sqrt{\lambda}(A - B) = 0, \quad A\sqrt{\lambda}e^{a\sqrt{\lambda}} - B\sqrt{\lambda}e^{-a\sqrt{\lambda}} = 0,\]

hence \(A = B, \quad A\sqrt{\lambda}(e^{a\sqrt{\lambda}} - e^{-a\sqrt{\lambda}}) = 0,\) so \(A = 0\) and \(B = 0.\) There are no nontrivial solutions here.

For \(\lambda = 0,\) the solution is \(f(x) = Ax + B\) and by plugging this into the boundary conditions we only get \(A = 0.\) \(B\) can be anything.

Hence in this case any constant function is a solution. In particular, we take \(f_0 = 1.\)

For \(\lambda < 0,\) we get \(f(x) = A\sin(\mu x) + B\cos(\mu x),\) where \(\lambda = -\mu^2.\) Then \(f'(x) = A\cos(\mu x) - B\sin(\mu x).\)

Plugging this into the Neumann boundary conditions leads to

\[A = 0, \quad -B\sin(\mu a) = 0.\]

For a nontrivial solution we need \(B \neq 0,\) so \(\sin(\mu a) = 0.\) We get

\[\mu_n = \frac{\pi n}{a}, \quad \lambda_n = -\left( \frac{\pi n}{a} \right)^2, \quad f_n(x) = B\cos \left( \frac{\pi n}{a} x \right).\]

Again we can discard \(B\) and this time \(f_{-n} = f_n,\) so negative \(n\)'s are redundant. On the other hand, here for \(n = 0\) \(\lambda = 0\) is a valid answer, unlike in the Dirichlet case.

We get the following list of separable solutions:

\[u(x,t) = f_n(x)e^{-\lambda_n t} = \cos \left( \frac{\pi n}{a} x \right) e^{-\left( \frac{\pi n}{a} \right)^2 t}, \quad n \in \mathbb{Z}, \quad n \geq 0.\]
Hence if

\[ u_0(x) = \sum_{n \geq 0} c_n f_n(x) = \sum_{n \geq 0} c_n \cos \left( \frac{\pi n}{a} x \right), \]

then a solution is given by

\[ u(x, t) = \sum_{n \geq 0} \cos \left( \frac{\pi n}{a} x \right) e^{-\left( \frac{\pi n}{a} \right)^2 t}. \]

(10) is called a cosine Fourier decomposition/series/expansion.

3: Random example: With a Dirichlet condition at one end and a Neumann condition at the other end, we get some other kind of Fourier series (no particular name for this one).

The boundary conditions here are

\[ u(0, t) = 0, \quad u_x(a, t) = 0. \]

The eigenvalue problem is, for \( f : [0, a] \to \mathbb{C}, \)

\[ f''(x) = \lambda f(x), \quad f(0) = 0, \quad f'(a) = 0. \]

By doing the computation in the same way as above, we obtain

\[ \mu_k = \frac{\pi(2k + 1)}{2a}, \quad \lambda_k = -\left( \frac{\pi(2k + 1)}{2a} \right)^2, \quad f_k(x) = \sin \left( \frac{\pi(2k + 1)}{2a} x \right), \quad k \in \mathbb{Z}, \quad k \geq 0. \]

4: Periodic boundary conditions: for equation (7), periodic boundary conditions are of the form

\[ u(0, t) = u(a, t), \quad u_x(0, t) = u_x(a, t). \]

They lead to the following eigenvalue/eigenfunction equation: for \( f : [0, a] \to \mathbb{C}, \)

\[ f'' = \lambda f, \quad f(0) = f(a), \quad f'(0) = f'(a). \]

Again we consider the three cases \( \lambda > 0, \lambda = 0, \) and \( \lambda < 0. \) For \( \lambda > 0 \) there is no nontrivial solution and for \( \lambda = 0 \) \( f(x) = 1 \) is a solution.

For \( \lambda = -\mu^2 < 0 \) we get

\[ f(x) = \tilde{A} e^{i\mu x} + \tilde{B} e^{-i\mu x}. \]

Here the computations are simpler if we use complex exponentials, instead of sines and cosines.

For this expression to satisfy the boundary conditions, it is necessary that

\[ \tilde{A} + \tilde{B} = \tilde{A} e^{i\mu a} + \tilde{B} e^{-i\mu a}, \quad i \mu (\tilde{A} - \tilde{B}) = i \mu (\tilde{A} e^{i\mu a} - \tilde{B} e^{-i\mu a}). \]

From the second equation we get \( \tilde{A} - \tilde{B} = \tilde{A} e^{i\mu a} - \tilde{B} e^{-i\mu a}, \) so combining this with the first equation we get

\[ \tilde{A} = \tilde{A} e^{i\mu a}, \quad \tilde{B} = \tilde{B} e^{-i\mu a}. \]

If at least one of \( A \) or \( B \) is nonzero, as is needed for a nontrivial solution, then \( 1 = e^{i\mu a} = \cos(\mu a) + i \sin(\mu a). \) Thus \( \mu a = 2\pi n, \) so \( \mu = \frac{2\pi n}{a}. \)

The separable solutions are, for \( n \in \mathbb{Z}, \)

\[ u(x, t) = e^{\frac{2\pi n}{a} x} e^{-\left( \frac{2\pi n}{a} \right)^2 t}. \]

Thus, if we could represent the initial data as a superposition

\[ u_0(x) = \sum_{n \in \mathbb{Z}} c_n e^{\frac{2\pi n}{a} x}, \]

then a solution would be given by

\[ u(x, t) = \sum_{n \in \mathbb{Z}} c_n e^{\frac{2\pi n}{a} x} e^{-\left( \frac{2\pi n}{a} \right)^2 t}. \]
Robin: For example, consider the Robin-type boundary conditions
\[ u(0, t) + u_x(0, t) = 0, \quad u(a, t) + 2u_x(a, t) = 0. \]

**Problem 11.** a) For \( f : [0, a] \to \mathbb{C} \), solve the eigenvalue problem
\[ f''(x) = \lambda f(x), \quad f'(0) = 0, \quad f(a) = 0. \]
b) For \( f : [0, a] \to \mathbb{C} \), solve the eigenvalue problem
\[ f''(x) = \lambda f(x), \quad f'(0) + f(0) = 0, \quad f'(a) + f(a) = 0. \]

**1.4. Higher dimensions: rectangles and parallelepipeds.** For such higher-dimensional shapes, one uses a combination of the solutions we obtained in one dimension.

**Example:** Consider the heat equation on \([0, a] \times [0, b] \times \mathbb{R}\), with Dirichlet boundary conditions on the top, bottom, and left sides and Neumann boundary conditions on the right side.

Explicitly, the equation is
\[ u_t = \Delta u = u_{xx} + u_{yy}, \quad u(x, y, 0) = u_0(x, y). \]

The boundary conditions are: In the \( x \) direction, there is a Dirichlet condition on the left and a Neumann condition on the right:
\[ u(0, y) = 0, \quad u_x(a, y) = 0, \quad y \in [0, b]. \]

In the \( y \) direction, there are Dirichlet boundary conditions on the top and on the bottom:
\[ u_y(x, 0) = u_y(x, b) = 0, \quad x \in [0, a]. \]

To solve the equation, we again use the method of separation of variables. Separated solutions will then be a product
\[ u = f_1(x)f_2(y)g(t), \]
where \( f_1 \) fulfills conditions (12) and \( f_2 \) fulfills conditions (13). Using the solutions we have already computed to these 1D problems, we see that by (11)
\[ f_1(x) = \sin \left( \frac{\pi(2k + 1)}{2a} x \right), \quad \lambda_1 = -\left( \frac{\pi(2k + 1)}{2a} \right)^2, \quad k \in \mathbb{Z}, \quad k \geq 0 \]
and by (8)
\[ f_2(y) = \sin \left( \frac{\pi n}{a} y \right), \quad \lambda_2 = -\left( \frac{\pi n}{a} \right)^2, \quad n \in \mathbb{Z}, \quad n > 0. \]

The corresponding eigenvalue is \( \lambda = \lambda_1 + \lambda_2 \) and we get the following family of separable solutions:
\[ u(x, y, t) = \sin \left( \frac{\pi(2k + 1)}{2a} x \right) \sin \left( \frac{\pi n}{a} y \right) e^{-\left( \frac{\pi (2k + 1)}{2a} \right)^2 - \left( \frac{\pi n}{a} \right)^2}, \quad k \in \mathbb{Z}, \quad k \geq 0, \quad n \in \mathbb{Z}, \quad n > 0. \]

Now the question is whether any initial data can be decomposed as
\[ u_0(x, y) = \sum_{k \geq 0, n > 0} c_{k, n} \sin \left( \frac{\pi(2k + 1)}{2a} x \right) \sin \left( \frac{\pi n}{a} x \right). \]

and how to find the coefficients \( c_{k, n} \). Once this is done, a solution is
\[ u(x, y, t) = \sum_{k \geq 0, n > 0} c_{k, n} \sin \left( \frac{\pi(2k + 1)}{2a} x \right) \sin \left( \frac{\pi n}{a} x \right) e^{-\left( \frac{\pi (2k + 1)}{2a} \right)^2 - \left( \frac{\pi n}{a} \right)^2}. \]
2. Euclidean space: the Fourier transform

Let’s recall the definitions.

**Definition 8.** Fourier transform:

\[ \hat{f}(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx. \]

**Definition 9.** Inverse Fourier transform:

\[ \tilde{g}(x) = \int_{-\infty}^{\infty} g(\xi) e^{ix\xi} d\xi. \]

These alternative definitions are also often used:

\[ \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx, \quad \tilde{g}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\xi) e^{ix\xi} d\xi \]

or

\[ \hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx, \quad \tilde{g}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\xi) e^{ix\xi} d\xi. \]

or

\[ \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x\xi} dx, \quad \tilde{g}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\xi) e^{2\pi i \xi} d\xi. \]

All these definitions are the same up to factors of 2\(\pi\).

We want to understand:

- \(Q_1\): When does the Fourier transform exist?
- \(Q_2\): What sort of function is \(\hat{f}\)?
- \(Q_3\): When does the Fourier inversion formula hold, i.e. \(\mathcal{F}^{-1}\hat{f} = f\)?

**Definition 10.** A function \(f: \mathbb{R} \to \mathbb{R}\) is said to belong to the Schwartz class \(\mathcal{S}\) if and only if \(f \in C^\infty\) and for any \(m, n \geq 0\) there exists some constant \(C_{m,n}\) such that

\[ |f^{(n)}(x)| \leq \frac{C_{m,n}}{|x|^m}. \]

In other words, all its derivatives decay faster than any polynomial.

Examples: test functions. Gaussians; \(e^{-x^4}\); using cutoff functions we can construct Schwartz-class functions with any prescribed behavior. Example:

\[ f(x) = \begin{cases} 
  e^{-|x|}, & |x| > 1, \\
  \text{smooth}, & |x| \leq 1.
\end{cases} \]

\(\mathcal{S}\) has the following desirable properties: the Fourier transform of a Schwartz-class function exists and is a Schwartz-class function as well; also, the Fourier inversion formula holds.

Let’s prove them one by one.

**First step:** The Fourier transform \(\hat{f}(\xi)\) of a Schwartz-class function \(f \in \mathcal{S}\) exists for every \(\xi \in \mathbb{R}\) and is bounded: for all \(\xi \in \mathbb{R}\)

\[ |\hat{f}(\xi)| \leq \int_{-\infty}^{\infty} |f(x)| dx. \]

The quantity

\[ \|f\|_{L^1} = \int_{-\infty}^{\infty} |f(x)| dx \]

is called the \(L^1\) norm of the function \(f\). Clearly, if \(f \in \mathcal{S}\) then \(f \in L^1\) with e.g.

\[ \|f\|_{L^1} \leq \int_{-\infty}^{0} \frac{C_2.0}{|x|^2} dx + \int_{0}^{1} C_{0.0} dx + \int_{1}^{\infty} \frac{C_2.0}{|x|^2} dx \leq 2(C_{2.0} + C_{0.0}). \]
Proof. We interpret the definition of the Fourier transform as

\[ \hat{f}(\xi) = \frac{1}{2\pi} \lim_{R_1, R_2 \to \infty} \int_{-R_1}^{R_2} f(x)e^{-ix\xi} \, dx, \]

where we use the Riemann integral to compute the integral from \(-R_1\) to \(R_2\). This way of computing the integral is called the improper Riemann integral. Another possible interpretation of (14) is as a Lebesgue integral, but we’ll stick to Riemann integration for now.

First, we show that the limit exists. By Cauchy’s criterion, the limit exists if and only if for every \(\epsilon > 0\) there exists some \(R\) such that whenever \(R_1, R_2, \bar{R}_1, \bar{R}_2 > R\)

\[ \left| \int_{-R_1}^{R_2} f(x)e^{-ix\xi} \, dx - \int_{-\bar{R}_1}^{\bar{R}_2} f(x)e^{-ix\xi} \, dx \right| < \epsilon. \]

But this expression is equal to

\[ \left| \int_{-R_1}^{-\bar{R}_1} f(x)e^{-ix\xi} \, dx + \int_{\bar{R}_2}^{R_2} f(x)e^{-ix\xi} \, dx \right| \leq \left| \int_{-R_1}^{-\bar{R}_1} f(x)e^{-ix\xi} \, dx \right| + \left| \int_{\bar{R}_2}^{R_2} f(x)e^{-ix\xi} \, dx \right|. \]

Since \(e^{i\theta} = \cos \theta + i \sin \theta\), for \(\theta \in \mathbb{R}\) \(|e^{i\theta}| = \sqrt{\cos^2 \theta + \sin^2 \theta} = 1\), so this is further bounded by

\[ \leq \int_{-R_1}^{-\bar{R}_1} |f(x)| \, dx + \int_{\bar{R}_2}^{R_2} |f(x)| \, dx \leq \int_{-R_1}^{-\bar{R}_1} \frac{C_{2,0}}{|x|^2} \, dx + \int_{\bar{R}_2}^{R_2} \frac{C_{2,0}}{|x|^2} \, dx \leq \max \left( \frac{1}{|R_1|}, \frac{1}{|\bar{R}_1|} \right) + \max \left( \frac{1}{|R_2|}, \frac{1}{|\bar{R}_2|} \right) \leq \frac{2}{|R|}. \]

Here we did not keep track of whether e.g. \(R_1\) or \(\bar{R}_1\) is bigger. The bounds come from the fact that \(f \in S\).

Choosing \(R = 3/\epsilon\), we get that this integral fulfills Cauchy’s criterion, so it converges.

Next, we prove the uniform bound. It suffices to prove it over every finite interval \([-R_1, R_2]\) and then take the limit. But

\[ \left| \int_{-R_1}^{R_2} f(x)e^{-ix\xi} \, dx \right| \leq \int_{-R_1}^{R_2} |f(x)e^{-ix\xi}| \, dx = \int_{-R_1}^{R_2} |f(x)| \, dx \leq \int_{-\infty}^{\infty} |f(x)| \, dx \]

and the conclusion follows by taking the limit as \(R_1, R_2 \to \infty\). \(\square\)

In fact, whenever \(f \in L^1\) its Fourier transform is bounded and

\[ \sup_{\xi \in \mathbb{R}} |\hat{f}(\xi)| \leq \|f\|_{L^1}. \]

This is another situation in which the Fourier transform is well-defined.

Second step: If \(f\) is a Schwartz-class function, then so are \(f'\) and \(xf\).

Proof. If \(f \in S\), then for every \(m, n \in \mathbb{Z}, m, n \geq 0\), there exists \(C_{m,n}\) such that

\[ |f^{(n)}(x)| \leq \frac{C_{m,n}}{|x|^m}. \]

Since \((f')^{(n)} = f^{(n+1)}\),

\[ |(f')^{(n)}(x)| \leq \frac{C_{m,n+1}}{|x|^m}, \]

implying that \(f' \in S\).

Likewise, (one can prove by induction that)

\[ (xf)^{(n)} = xf^{(n)} + nf^{(n-1)}, \]
so

\[ |(xf)^{(n)}(x)| \leq \frac{C_{m+1,n} + nC_{m,n-1}}{|x|^m}, \]

proving that \( xf \in \mathcal{S} \).

So if \( f \in \mathcal{S} \), then \( xf \in \mathcal{S} \), so then \( x^2f = x(xf) \in \mathcal{S} \), etc., and by iterating/induction \( x^m f \in \mathcal{S} \) for all \( m \geq 0 \).

Likewise, if \( f \in \mathcal{S} \), then \( f' = (f')' \in \mathcal{S} \), etc., and by induction \( f^{(n)} \in \mathcal{S} \) for all \( n \geq 0 \).

**Problem 12.**

a) If \( f \in \mathcal{S} \), prove that \( \sin(xf) \in \mathcal{S} \).

b) If \( f, g \in \mathcal{S} \), prove that \( fg \in \mathcal{S} \).

c) If \( f \in \mathcal{S} \), prove that \( e^{f(x)} - 1 \in \mathcal{S} \).

Hint: you can use all sorts of methods, including induction.

**Third step:** Compute the Fourier transforms of \( f' \) and of \( xf \):

\[ \mathcal{F}[f'](\xi) = i\xi \hat{f}(\xi), \quad \mathcal{F}[xf](\xi) = \frac{d}{d\xi} \hat{f}(\xi). \]

**Proof.** First, since we already proved that \( f', xf \in \mathcal{S} \), it follows that their Fourier transforms are well-defined and bounded.

To compute the Fourier transform of \( f' \), we integrate by parts:

\[ \mathcal{F}[f'](\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f'(x)e^{-ix\xi} dx = -\frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)(-i\xi)e^{-ix\xi} dx = i\xi \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)e^{-ix\xi} dx = i\xi \hat{f}(\xi), \]

since the boundary terms are zero, due to the rapid decay at infinity of \( f \).

To compute the Fourier transform of \( xf \), we use the fact that

\[ \frac{\partial}{\partial\xi}(e^{-ix\xi}) = -ixe^{-ix\xi}. \]

Then

\[ \mathcal{F}[xf](\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} xf(x)e^{-ix\xi} dx = \frac{1}{-i} \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)(-ixe^{-ix\xi}) dx = \frac{1}{-i} \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \frac{\partial}{\partial\xi} (e^{-ix\xi}) dx = \frac{1}{-i} \frac{d}{d\xi} \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \xi e^{-ix\xi} dx = \frac{d}{d\xi} \hat{f}(\xi). \]

Again, iterating an arbitrary number of times we get

\[ \mathcal{F}[f^{(n)}](\xi) = (i\xi)^n \hat{f}(\xi), \quad \mathcal{F}[x^m f(x)](\xi) = i^m \frac{d^m}{d\xi^m} \hat{f}(\xi). \]
Conclusion: If \( f \in \mathcal{S} \), then \( \hat{f} \in \mathcal{S} \).

**Proof.** Again, note that \( \hat{f} \) exists and is a bounded function. We need to prove that for every \( M \) and \( N \) there exists some constant \( \tilde{C}_{M,N} \) such that

\[
|\hat{f}^{(N)}(\xi)| \leq \frac{\tilde{C}_{M,N}}{|\xi|^M}.
\]

Equivalently, \( |\xi^M(\hat{f})^{(N)}(\xi)| \leq \tilde{C}_{M,N} \).

But using the previous properties of the Fourier transform we get that

\[
(\hat{f})^{(N)}(\xi) = i^{-N} \mathcal{F}[x^N f](\xi)
\]

and further

\[
\xi^M i^{-N} \mathcal{F}[x^N f](\xi) = i^{-M-N} \mathcal{F}[(x^N f)^{(M)}](\xi).
\]

Thus, the statement (15) reduces to proving that \( \mathcal{F}[(x^N f)^{(M)}](\xi) \) is uniformly bounded.

However, since \( f \in \mathcal{S} \), \( x^N f(x) \in \mathcal{S} \) as well. In turn this implies that \( (x^N f(x))^{(M)} \in \mathcal{S} \). Therefore its Fourier transform is uniformly bounded, q.e.d.

**Remark 2.1.** Why not work with test functions? Because the Fourier transform of a test function, as a rule, is not a test function.

**Fourier transform of a Gaussian:** At this stage, we can also prove that the Fourier transform of a Gaussian is a Gaussian:

\[
e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}}.
\]

**Proof.** There are at least two well-known proofs, one involving complex analysis and another one using ODEs.

Using ODEs: Note that

\[
\frac{d}{dx} e^{-\frac{x^2}{2}} = -x e^{-\frac{x^2}{2}}.
\]

This is an ODE for \( e^{-\frac{x^2}{2}} \). Let’s apply the Fourier transform to it: using the rules we have proved,

\[
i\xi e^{-\frac{x^2}{2}}(\xi) = -i \frac{d}{d\xi} e^{-\frac{\xi^2}{2}}(\xi).
\]

Canceling the \( i \), we obtain that the Fourier transform \( g(\xi) = e^{-\frac{\xi^2}{2}}(\xi) \) satisfies the same equation as the original Gaussian:

\[
\frac{d}{d\xi} g(\xi) = -\xi g(\xi).
\]

Now let’s solve it. Assuming \( g \neq 0 \), we can use separation of variables (or without this assumption we can e.g. use the method of the integrating factor) and get that every solution has the form

\[
g(\xi) = g_0 e^{-\frac{\xi^2}{2}}.
\]

Thus, this Fourier transform is the same as the original function, up to a multiplicative constant. We are left with determining \( g_0 \).

Plugging in \( \xi = 0 \), we get \( g(0) = g_0 \). On the other hand,

\[
g(0) = e^{-\frac{\xi^2}{2}}(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-x^2/2} e^{-ix\cdot 0} \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-x^2/2} \, dx = \frac{1}{2\pi} \sqrt{2\pi} = \frac{1}{\sqrt{2\pi}}
\]

where the integral is left as an exercise.
PROBLEM 13. Show that 
\[ \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \, dx = \sqrt{2\pi}. \]

PROBLEM 14. Prove the following properties of the Fourier transform:

a) If \( g(x) = f(ax) \), then 
\[ \hat{g}(\xi) = \frac{1}{a} \hat{f}\left(\frac{\xi}{a}\right). \]

b) If \( h(x) = f(x + b) \), then 
\[ \hat{h}(\xi) = \hat{f}(\xi). \]

c) Use part a and the fact that \( e^{-\frac{x^2}{2}} = 1 \sqrt{2\pi} e^{-\frac{\xi^2}{2}} \) to prove that for \( \alpha > 0 \)
\[ e^{-\alpha x^2} = \frac{1}{\sqrt{4\alpha \pi}} e^{-\frac{\xi^2}{4\alpha}}. \]

Generalize this by also using part b.

PROBLEM 15. Compute the Fourier transform of Gaussians in higher dimensions.

An easy-to-compute Fourier transform:
\[ (16) \quad \widehat{e^{-a|x|}} = \frac{2a}{\xi^2 + a^2}, \quad \widehat{e^{-a|x|}} = \frac{1}{2\pi} \frac{2a}{\xi^2 + a^2}. \]

Neither is a Schwartz function.

PROOF. First, on \([0, \infty)\)
\[ \int_{0}^{\infty} e^{-ax} e^{-ix\xi} \, dx = \int_{0}^{\infty} e^{-x(a+i\xi)} \, dx = -\frac{1}{a + i\xi} e^{-x(a+i\xi)} \Big|_{x=0}^{\infty} = \frac{1}{a + i\xi}. \]

On the other hand, on \((-\infty, 0]\)
\[ \int_{-\infty}^{0} e^{-a|x|} e^{-ix\xi} \, dx = \int_{0}^{\infty} e^{ax - ix\xi} \, dx = \frac{1}{a - i\xi} e^{x(a-i\xi)} \Big|_{x=-\infty}^{0} = \frac{1}{a - i\xi}. \]

Then
\[ \frac{1}{a + i\xi} + \frac{1}{a - i\xi} = \frac{2a}{\xi^2 + a^2}. \]

Rectangular function: Its Fourier transform is a sinc function.

The rectangular function is the indicator function of an interval. The indicator function of a set is the function that indicates whether a certain element belongs to a set or not.

Let
\[ \chi_{[-L,L]}(x) = \begin{cases} 1, & x \in [-L, L] \\ 0, & x \notin [-L, L]. \end{cases} \]

Its Fourier transform is
\[ \hat{\chi}_{[-L,L]}(\xi) = \frac{1}{\pi} \frac{\sin(L\xi)}{\xi}, \]
which is basically a sinc function.

The sinc function is
\[ \text{sinc} x = \frac{\sin x}{x}. \]

Triangular function: also easy to compute.

Next step: Now we can prove the Fourier inversion formula \( F^{-1}\hat{f} = f \), i.e. the inverse Fourier transform of the Fourier transform of a function is the original function. For now, we prove this for Schwartz-class functions.
2. EUCLIDEAN SPACE: THE FOURIER TRANSFORM

PROPOSITION 2.2. If \( f \in \mathcal{S} \), then \( \mathcal{F}^{-1} \hat{f} = f \).

PROOF. Note that if \( f \in \mathcal{S} \), then \( \hat{f} \in \mathcal{S} \), so \( \mathcal{F}^{-1} \hat{f} = 2\pi \mathcal{F}[\hat{f}(-\xi)] \in \mathcal{S} \) as well. So all the functions are well-defined, smooth, rapidly decaying, etc..

Spelling out the definitions, what we want to prove is
\[
f(y) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{iy\xi} d\xi = \int_{-\infty}^{\infty} \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx \right) e^{iy\xi} d\xi.
\]

Rearranging the expression, we get
\[
f(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-i(x-y)\xi} dx d\xi.
\]

If we could change the order of integration and if we somehow knew that
\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(x-y)\xi} d\xi = \delta_0(x-y),
\]

then the conclusion would follow.

In order to flesh out this formal computation, first we test this equality against some Schwartz-class function (or test function, either works) \( g \). Equality (17) is true if and only if for all \( g \in \mathcal{D} \)
\[
\int_{-\infty}^{\infty} f(y)g(y) \, dy = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-i(x-y)\xi} g(y) \, dx \, d\xi \, dy.
\]

The point of this rephrasing is to be able to change the order of integration.

Next, we want to regularize the rather singular integral (18) and replace it with something better behaved. One can show that
\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-i(x-y)\xi} g(y) \, dx \, d\xi \, dy = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-i(x-y)\xi} - \epsilon^2 g(y) \, dx \, d\xi \, dy.
\]

Another commonly used method to regularize this expression is multiplying by \( e^{-\epsilon|\xi|^2} \) instead of \( e^{-\epsilon|\xi|^2} \).

Interlude. The absolutely first thing we want to prove here is (19). I’ll include this one proof to show their general flavor.

Part of the proof is that for each \( y \in \mathbb{R} \)
\[
\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-i(x-y)\xi} - \epsilon^2 dx \, d\xi = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-i(x-y)\xi} dx \, d\xi.
\]

Since the inner integral is just \( \hat{f}(\xi) \), we are to prove that
\[
\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \hat{f}(\xi) e^{iy\xi} - \epsilon^2 d\xi = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{iy\xi} d\xi.
\]

The subsequent proof is slightly different from the one in class.

By taking the difference, this is equivalent to
\[
\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \hat{f}(\xi) e^{iy\xi} (1 - e^{-\epsilon^2}) \, d\xi = 0.
\]

By taking the absolute value, this last statement is implied by
\[
\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} |\hat{f}(\xi)| (1 - e^{-\epsilon^2}) \, d\xi = 0.
\]

We next use the mean value theorem: \( f(b) - f(a) = (b - a)f'(c) \). Here we set \( b = 0, a = -\epsilon \), and \( f(z) = e^{z^2} \). Then for some \( c \in (-\epsilon, 0) \)
\[
e^0 - e^{-\epsilon^2} = \epsilon^2 e^{-\epsilon^2} \leq \epsilon^2.
\]
Then the integral is bounded by

$$\int_{-\infty}^{\infty} |\hat{f}(\xi)|(1 - e^{-\epsilon|\xi|^2}) \, d\xi \leq \int_{-\infty}^{\infty} |\hat{f}(\xi)|\epsilon^2 \, d\xi = \epsilon \int_{-\infty}^{\infty} |\hat{f}(\xi)| \xi^2 \, d\xi.$$  

But $\hat{f}$ is a Schwartz-class function, so $|\hat{f}| \leq C_{0,0}$, $|\hat{f}(\xi)| \leq \frac{C_{4,0}}{|\xi|^4}$, and

$$\int_{-\infty}^{\infty} |\hat{f}(\xi)| \xi^2 \, d\xi \leq \int_{-1}^{1} C_{0,0} \xi^2 \, d\xi + \int_{|\xi|>1} C_{4,0} \xi^{-2} \, d\xi \leq \frac{2C_{0,0}}{3} + 2C_{4,0} = \tilde{C} < \infty.$$  

So as $\epsilon \to 0$ the difference, being less than $\tilde{C}\epsilon$, also goes to zero, q.e.d.

**End interlude.**

After testing against $g$ and regularizing the expression, we can now change the order of integration and compute the $\xi$ integral, since it’s just the Fourier transform of a Gaussian:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} e^{-i(x-y)\xi - \epsilon|\xi|^2} \, d\xi \right) f(x)g(y) \, dx \, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\epsilon\pi}} e^{-\frac{(x-y)^2}{4\epsilon}} f(x)g(y) \, dx \, dy.$$  

We further recognize a convolution product here: by definition

$$\left[ f(x) * \frac{1}{\sqrt{4\epsilon\pi}} e^{-\frac{x^2}{4\epsilon}} \right](y) = \int_{-\infty}^{\infty} f(x) \frac{1}{\sqrt{4\epsilon\pi}} e^{-\frac{(y-x)^2}{4\epsilon}} \, dx.$$  

This is just what we have, so we rewrite (20) as

$$\int_{-\infty}^{\infty} \left[ f(x) * \frac{1}{\sqrt{4\epsilon\pi}} e^{-\frac{x^2}{4\epsilon}} \right](y)g(y) \, dx \, dy.$$  

The expression

$$H_{\epsilon}(x) = \frac{1}{\sqrt{4\epsilon\pi}} e^{-\frac{x^2}{4\epsilon}}$$  

is called the heat kernel. If we could show that

$$\lim_{\epsilon \to 0} [f * H_{\epsilon}](y) = f(y)$$  

in some suitable sense, then this would finish the proof.

Here $H_{\epsilon}(x) = \frac{1}{\sqrt{4\epsilon\pi}} e^{-\frac{(x-y)^2}{4\epsilon}}$ is called a good kernel or an approximation to the identity, because it has the following properties:

**Definition 11.** A kernel $(K_{\epsilon}(x))_{\epsilon>0}$ is called a good kernel if it has the following three properties:
1. For each $\epsilon > 0$,

$$\int K_{\epsilon}(x) \, dx = 1.$$  

2. For each $r > 0$, $\lim_{\epsilon \to 0} \int_{|x|>r} |K_{\epsilon}(x)| \, dx = 0$.
3. There exists some constant $M$ such that for all $\epsilon > 0$

$$\int |K_{\epsilon}(x)| \, dx < M.$$  

**Proposition 2.3.** Let $f : \mathbb{R} \to \mathbb{R}$ be continuous at some point $x_0 \in \mathbb{R}$. If $K_{\epsilon}$ is a good kernel, then

$$\lim_{\epsilon \to 0} [f * K_{\epsilon}](x_0) = f(x_0).$$  

If $f$ is uniformly continuous, then $f * K_{\epsilon} \to f$ uniformly.

Since the heat kernel is a good kernel, this is enough to conclude the proof.  \[\square\]
One can also regularize this expression with \( e^{-\epsilon|\xi|} \) instead of \( e^{-|\xi|^2} \).

**Next step:** Now we can prove Plancherel’s identity

\[
\int_{-\infty}^{\infty} |f(x)|^2 \, dx = 2\pi \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 \, d\xi.
\]

**Proof.** We define the dot product

\[
\langle f, g \rangle := \int_{-\infty}^{\infty} f(x)\overline{g(x)} \, dx.
\]

This is a dot product on the Schwartz space, \( \langle \cdot, \cdot \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{C} \), meaning that it has the following properties:

1. It is linear in both arguments: in the first argument

\[
\langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle, \quad \langle \alpha f, g \rangle = \alpha \langle f, g \rangle.
\]

For the second argument there is a twist, when pulling out a constant it gets conjugated:

\[
\langle f, g_1 + g_2 \rangle = \langle f, g_1 \rangle + \langle f, g_2 \rangle, \quad \langle f, \alpha g \rangle = \overline{\alpha} \langle f, g \rangle.
\]

The proper descriptor for this property is “sesquilinear”.

2. \( \langle f, g \rangle = \overline{\langle g, f \rangle} \). See above.

3. \( \langle f, f \rangle \geq 0 \); \( \langle f, f \rangle = 0 \) if and only if \( f = 0 \).

The very last condition has to be properly interpreted: when \( \langle f, f \rangle = 0 \), \( f \) is not necessarily zero everywhere, but almost everywhere, i.e. except on a set of measure zero.

Then we see that on one hand

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)e^{-ix\xi}g(\xi) \, dx \, d\xi = 2\pi \int_{-\infty}^{\infty} \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)e^{-ix\xi} \, dx \right) \overline{g(\xi)} \, d\xi = 2\pi \langle \hat{f}, g \rangle
\]

and on the other hand

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)e^{-ix\xi}\overline{g(\xi)} \, dx \, d\xi = \int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} e^{ix\xi} \overline{g(\xi)} \, d\xi \, dx = \langle f, \hat{g} \rangle.
\]

So

\[
2\pi \langle \hat{f}, g \rangle = \langle f, \hat{g} \rangle.
\]

Setting \( g = \hat{G} \), then \( \hat{g} = \mathcal{F}^{-1}\hat{G} = G \), so we get

\[
2\pi \langle \hat{f}, \hat{G} \rangle = \langle f, G \rangle.
\]

In particular, for \( f = G \)

\[
\langle f, f \rangle = 2\pi \langle \hat{f}, \hat{f} \rangle.
\]

This is Plancherel’s identity. \( \square \)

The quantity

\[
\|f\|_{L^2} = \sqrt{\langle f, f \rangle} = \left( \int_{-\infty}^{\infty} |f(x)|^2 \, dx \right)^{1/2}
\]

is called the \( L^2 \) norm of \( f \). Thus Plancherel’s identity states that

\[
\|\hat{f}\|_{L^2} = \frac{1}{\sqrt{2\pi}} \|f\|_{L^2}.
\]

More generally, \( L^p \)-Lebesgue spaces (\( L \) stands for Lebesgue here) are the spaces of functions that have a finite \( L^p \) norm, defined as

\[
\|f\|_{L^p} := \left( \int_{-\infty}^{\infty} |f(x)|^p \, dx \right)^{1/p}, \quad 1 \leq p < \infty.
\]
\( L^\infty \) is defined differently, as the space of essentially bounded functions, meaning they are bounded except on a measure zero set. The \( L^\infty \) norm is the \textit{essential} least upper bound for the absolute value.

Out of all \( L^p \) spaces, three are most important: \( L^1 \) is the space of absolutely integrable functions, \( L^2 \) is the space of square integrable functions, and \( L^\infty \) is the space of essentially bounded functions.

The dot product is naturally defined on \( L^2 \), since by the Cauchy–Schwartz inequality
\[
|\langle f, g \rangle| \leq \|f\|_{L^2} \|g\|_{L^2}.
\]
Only \( L^2 \) admits a dot product; for \( p \neq 2 \), the other Lebesgue spaces \( L^p \) don’t.

So far we have shown that
\[
\|\hat{f}\|_{L^\infty} \leq \frac{1}{2\pi} \|f\|_{L^1} \tag{21}
\]
and, by Plancherel’s identity,
\[
\|\hat{f}\|_{L^2} = \frac{1}{\sqrt{2\pi}} \|f\|_{L^2},
\]
but only for Schwartz-class functions so far.

By interpolating we get the more general Hausdorff–Young inequality: for \( f \in \mathcal{S} \)
\[
\|\hat{f}\|_{L^{p'}} \leq \frac{1}{(2\pi)^{1/p}} \|f\|_{L^p}, \quad 1 \leq p \leq 2, \quad \frac{1}{p'} + \frac{1}{p} = 1.
\]

Inequalities, identities, etc. that are proved under the assumption that the functions involved are smooth, decay fast, etc. are called \textit{a priori} bounds. Ideally, we want to remove these extraneous assumptions.

Suppose we have accomplished this (in fact, the Hausdorff–Young inequality is actually true for all \( f \in L^p, 1 \leq p \leq 2 \)). Then we shall have found several more instances where the Fourier transform is well-defined, namely for \( f \in L^p, 1 \leq p \leq 2 \).

On the other hand, if \( f \in L^p \) for \( p > 2 \), then its Fourier transform need not be a function.

**Problem 16.** Give an example of a function \( f \in \mathcal{S} \) for which \( \|\hat{f}\|_{L^\infty} < \frac{1}{2\pi} \|f\|_{L^1} \), i.e. the inequality \eqref{eq:21} is strict.

Extra credit: Formulate a necessary and sufficient condition for equality.

**Convolution:** For Schwartz-class functions, the following statements are true:
\[
\widetilde{f * g} = \frac{1}{2\pi} \tilde{f} * \tilde{g}; \quad \widehat{f * g} = \tilde{f} \tilde{g}.
\]

and
\[
\widehat{f * g} = 2\pi \tilde{f} \tilde{g}; \quad \tilde{f} \tilde{g} = \hat{f} \hat{g}.
\]

**Proof.** Start from
\[
\widehat{f * g} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f(x-y)g(y) \, dy \right) e^{-ix\xi} \, dx = 2\pi \tilde{f} \tilde{g}.
\]

Letting \( f = \tilde{F}, \ g = \tilde{G} \), this also implies that
\[
\tilde{F} \star \tilde{G} = 2\pi \tilde{F} \tilde{G}.
\]

The other two properties are proved similarly. \( \square \)

**Problem 17.** Prove that, for \( f, g \in \mathcal{S} \), \( \hat{f} g = \hat{f} * \tilde{g} \).

And with this we have proved most of the important properties of the Fourier transform that we are going to use.
2.1. Solving PDEs on the whole space using the Fourier transform. Finally, we now have all the important properties of the Fourier transform available, as well as the concept itself. Let’s look again at the heat equation in this light.

Heat equation in one dimension:

\[ u_t = u_{xx}, \quad u(x, 0) = u_0(x). \]

Taking the Fourier transform in the \( x \) variable of \( u \),

\[ \hat{u}(\xi, t) = \mathcal{F}_x u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(x, t) e^{-ix\xi} \, dx, \]

we get

\[ \frac{\partial}{\partial t} \hat{u}(\xi, t) = \frac{1}{2\pi} \frac{\partial}{\partial \xi} \int_{-\infty}^{\infty} u(x, t) e^{-ix\xi} \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \frac{\partial}{\partial t} u(x, t) \right] e^{-ix\xi} \, dx = \frac{\partial u}{\partial t}(\xi, t). \]

Using the equation we then obtain

\[ \frac{\partial}{\partial t} \hat{u}(\xi, t) = \frac{i\xi^2}{2\pi} \hat{u}(\xi, t) = -\xi^2 \hat{u}(\xi, t). \]

So for each individual value of \( \xi \) we get the following ODE, in the \( t \) variable only:

\[ \frac{d}{dt} \hat{u}(\xi, t) = -\xi^2 \hat{u}(\xi, t). \]

This equation has the solution

\[ \hat{u}(\xi, t) = e^{-t\xi^2} c(\xi) \]

and setting \( t = 0 \) we get

\[ c(\xi) = \hat{u}(\xi, 0) = \hat{u}_0(\xi). \]

So \( \hat{u}(\xi, t) = e^{-t\xi^2} \hat{u}_0(\xi) \) and reversing the Fourier transform we get

\[ u(x, t) = \mathcal{F}^{-1}[e^{-t\xi^2} \hat{u}_0(\xi)]. \]

We can express this fact in the equivalent form

\[ u(x, t) = \frac{1}{2\pi} \mathcal{F}^{-1}[\hat{u}_0] \ast \mathcal{F}^{-1}[e^{-t\xi^2}] = \frac{1}{2\pi} u_0(x) \ast \left[ \frac{2\pi}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \right] = u_0(x) \ast \left[ \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \right]. \]

All these are solution formulas for the equation.

The quantity \( H_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \) is called the heat kernel and, also, the fundamental solution of the heat equation (in one dimension).

Observation: This is only true for \( t > 0 \). For \( t < 0 \), the Fourier transform of \( e^{-t\xi^2} = e^{a\xi^2}, \ a > 0 \), is not defined.

Not only is this not a Schwartz-class function, it falls outside any class of functions for which the Fourier transform is defined.

Why take the Fourier transform only in \( x \)? In fact, it turns out that sometimes we want to take it only in \( x \), sometimes only in \( t \), sometimes in both; whatever works.

Schrödinger’s equation: We can do exactly the same for Schrödinger’s equation: if

\[ u_t = iu_{xx}, \quad u(x, 0) = u_0(x), \]

then

\[ \frac{\partial}{\partial t} \hat{u}(\xi, t) = -i\xi^2 \hat{u}(\xi, t). \]

Solving this, we get

\[ \hat{u}(\xi, t) = e^{-it\xi^2} \hat{u}_0(\xi), \quad u(\xi, t) = \mathcal{F}^{-1}[e^{-it\xi^2} \hat{u}_0(\xi)], \]
and
\[ u(x, t) = \frac{1}{2\pi} F^{-1}[\widehat{u}_0] * F^{-1}[e^{-it\xi^2}] = u_0(x) * \left[ \frac{1}{\sqrt{4it\pi}} e^{-\frac{x^2}{4it}} \right]. \]

The quantity
\[ S_t(x) = \frac{1}{\sqrt{4it\pi}} e^{-\frac{x^2}{4it}} \]

is called the fundamental solution of Schrödinger’s equation and sometimes the Schrödinger kernel. The function itself is also called a complex Gaussian.

Although it looks similar to the heat kernel, this one behaves very differently: The heat kernel has Gaussian, superexponential, decay, \( e^{-x^2} \), as \( x \to \infty \). The Schrödinger kernel has no decay at all as \( x \to \infty \), just very rapid oscillation.

**Problem 18.** Using the Fourier transform, in the same manner as above, find the general solution of the equation
\[ u_t = u_{xx} + 2iu_x, \quad u(x, 0) = u_0. \]

**The wave equation:** The wave equation in 1D is
\[ u_{tt} - u_{xx} = 0, \quad u(x, 0) = u_0(x), \quad u_t(x, 0) = u_1(x). \]

Taking the Fourier transform in \( x \), here we get
\[ \hat{\varepsilon}^2 \hat{u}(\xi, t) + \xi^2 \hat{u}(\xi, t) = 0. \]

This has solutions
\[ \hat{u}(\xi, t) = A(\xi) \cos(\xi t) + B(\xi) \sin(\xi t). \]

Taking the Fourier transform of initial data,
\[ \hat{u}(\xi, 0) = \hat{u}_0(\xi), \quad \hat{\varepsilon} \hat{u}(\xi, 0) = \hat{u}_1(\xi). \]

From matching the initial data we obtain
\[ A(\xi) = \hat{u}_0(\xi), \quad \xi B(\xi) = \hat{u}_1(\xi), \]

so
\[ \hat{u}(\xi, t) = \hat{u}_0(\xi) \cos(\xi t) + \hat{u}_1(\xi) \frac{\sin(\xi t)}{\xi}. \]

We then get that
\[ u(x, t) = F^{-1} \left[ \hat{u}_0(\xi) \cos(\xi t) + \hat{u}_1(\xi) \frac{\sin(\xi t)}{\xi} \right]. \]

Formally, at least, then
\[ u(x, t) = \frac{1}{2\pi} u_0(x) * F^{-1}[\cos(\xi t)] + \frac{1}{2\pi} u_1(x) * F^{-1}[\frac{\sin(\xi t)}{\xi}]. \]

But
\[ F^{-1}[\cos(\xi t)] = \frac{1}{2} (\delta_t(x) + \delta_{-t}(x)), \quad F^{-1}\left[ \frac{\sin(\xi t)}{\xi} \right] = \frac{1}{\pi} \chi_{[-t,t]}(x). \]

Here \( \delta \) is the Dirac delta function, which is actually not a function; it’s a measure and a distribution. To avoid distributions in this proof, here is a way that doesn’t use them: by Euler’s formula
\[ \cos(\xi t) = \frac{1}{2} (e^{-i\xi t} + e^{i\xi t}), \]

so
\[ F^{-1}[\hat{u}_0(\xi) \cos(\xi t)] = \frac{1}{2} F^{-1}[e^{-i\xi t} \hat{u}_0(\xi)] + \frac{1}{2} F^{-1}[e^{i\xi t} \hat{u}_0(\xi)]. \]
Reversing the Fourier transform, we get in one dimension
\[ u(x, t) = \frac{1}{2} (u_0(x - t) + u_0(x + t)) + u_1 \star \frac{1}{2} \chi[-t, t]. \]
This is called d’Alembert’s formula.

**The Laplace equation:** The Laplace equation in the upper half-plane \( \{(x, t) : t \geq 0\} \) is
\[ u_{tt} + u_{xx} = 0, \ u(x, 0) = u_0(x), \ u \to 0 \text{ as } t \to \infty. \]
This becomes
\[ \hat{u}(\xi, t) = A(\xi)e^{t|\xi|} + B(\xi)e^{-t|\xi|} \]
and we eliminate the exponentially growing part, so
\[ \hat{u}(\xi, t) = \hat{u}_0(\xi)e^{-t|\xi|}. \]
We get the solution
\[ u(x, t) = u_0(x) \ast P_t(x), \]
where by (16)
\[ P_t(x) = \frac{1}{2\pi} F^{-1}[-e^{-t|\xi|}] = \frac{1}{\pi} \frac{t}{t^2 + \xi^2}. \]
The expression \( P_t(x) \) is called the Poisson kernel and also the fundamental solution of Poisson’s equation in the upper half-plane.

**Stokes’ equation:** Stokes’ equation (one of several that bears his name), also known as linear or linearized KdV, is
\[ u_t + u_{xxx} = 0, \ u(0) = u_0. \]
Again, taking the Fourier transform we get
\[ \hat{u}_t = -(i\xi)^3 \hat{u} = i\xi^3 \hat{u}. \]
Same as for the other equations above, this leads to
\[ \hat{u}(\xi, t) = \hat{u}_0(\xi)e^{i\xi^3t}, \]
so
\[ u(x, t) = \frac{1}{2\pi} u_0(x) \ast F^{-1}e^{i\xi^3}. \]
Let’s compute the inverse Fourier transform of \( e^{i\xi^3} \):
\[ F^{-1}e^{i\xi^3} = \int_{-\infty}^{\infty} e^{ix\xi}e^{it\xi^3} \, d\xi = \int_{-\infty}^{\infty} e^{i(x\xi + t\xi^3)} \, d\xi. \]
In this formula, we don’t want a separate dependence on \( t \) and \( \xi \), so let’s eliminate one of the two variables: setting \( y = (3t)^{1/3} \xi \), so \( t\xi^3 = \frac{y^3}{3} \), we get
\[ F^{-1}e^{i\xi^3} = \int_{-\infty}^{\infty} e^{i\left(\frac{x}{(3t)^{1/3}}y + \frac{y^3}{3}\right)} \frac{dy}{(3t)^{1/3}} = \frac{1}{(3t)^{1/3}} \int_{-\infty}^{\infty} e^{i\left(\frac{x}{(3t)^{1/3}}y + \frac{y^3}{3}\right)} \, dy \]
\[ = \frac{2}{(3t)^{1/3}} \int_{-\infty}^{\infty} \cos \left( \frac{x}{(3t)^{1/3}}y + \frac{y^3}{3} \right) \, dy \]
since by Euler’s formula \( \cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \).
This integral turns out to be the **Airy function**. Getting here is the reason for that particular substitution above.
3. SEPARATION OF VARIABLES

DEFINITION 12. Airy’s function of the first kind, also called Airy’s function, is given by

\begin{equation}
\text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos \left( \frac{y^3}{3} + xy \right) \, dy.
\end{equation}

Airy’s function is a so-called special function. Its properties are very well understood, but it has no simple formula in terms of elementary functions (sine, cosine, exponential, powers, roots, etc.).

Then the solution to Stokes’ equation is

\begin{equation}
u(x,t) = u_0(x) + \frac{1}{(3t)^{1/3}} \text{Ai} \left( \frac{x}{(3t)^{1/3}} \right).
\end{equation}

2.2. More about the Airy function. We can integrate by parts to prove the integral (23) converges and to get various properties and more precise estimates for the Airy function.

The Airy function is rather well understood, with many thoroughly studied properties.

In particular, it has the following behavior: see the proof in the Appendix.

\[ \text{Ai}(x) \leq C; \quad \text{Ai}(x) \leq Cx^{-1/4}e^{-\frac{2}{3}x^{3/2}}, \quad x > 0; \quad |\text{Ai}(x)| \leq |x|^{-1/4}, \quad x < 0. \]

Consequently, we get the following result about the solution of Stokes’ equation:

\[ \|u(x,t)\|_{L^\infty_x} \leq C t^{-1/3} \|u_0\|_{L^1}. \]

Another such inequality is

\[ \|u(x,t)\|_{L^2_x} = \|u_0\|_{L^2}. \]

Then we can interpolate between these two inequalities and get

\[ \|u(x,t)\|_{L^p_x} \leq C t^{-\frac{1}{4}(\frac{2}{p}-\frac{1}{2})} \|u_0\|_{L^p}. \]

The Airy functions can be expressed in terms of the Bessel functions for \( x < 0 \) and to the modified Bessel functions for \( x > 0 \). For example, for \( x > 0 \)

\[ \text{Ai}(x) = \frac{1}{\pi} \sqrt{\frac{x}{3}} K_{\frac{1}{3}} \left( \frac{2}{3}x^{3/2} \right), \]

\[ \text{Bi}(x) = \sqrt{\frac{x}{3}} \left( I_{\frac{1}{3}} \left( \frac{2}{3}x^{3/2} \right) + I_{-\frac{1}{3}} \left( \frac{2}{3}x^{3/2} \right) \right). \]

Here \( I_{\pm1/3} \) and \( K_{1/3} \) are modified Bessel functions, see below.

2.3. Estimating fundamental solutions: Method of stationary phase. The main idea in estimating Airy’s function and in several other previous computations is that we can integrate by parts in the integral

\begin{equation}
\int e^{it\phi(x)} \psi(x) \, dx,
\end{equation}

wherever \( \phi'(x) \neq 0 \), and gain something by it.

Here \( \psi \) is called the amplitude and \( \phi \) is called the phase of the integrand. Such an integral is called an oscillatory integral.

The phase may or may not be given by an explicit formula.

Without knowing something about \( \phi \), at most we can say that the integral (24) is bounded: e.g. for \( \psi \in D \)

\[ \left| \int e^{it\phi(x)} \psi(x) \, dx \right| \leq \int |\psi(x)| \, dx < \infty. \]

Then there is no decay as a function of \( t \).

For example, if \( \phi(x) = c \), this is the best possible result.
At the other extreme, if $\phi' \neq 0$ on the domain of $\psi$, then we can integrate by parts and prove the expression decays:

$$
\int e^{it\phi(x)} \psi(x) \, dx = \int it\phi'(x)e^{it\phi(x)} \frac{1}{it\phi'(x)} \psi(x) \, dx = -\int e^{it\phi(x)} \frac{d}{dx} \left( \frac{1}{it\phi'(x)} \psi(x) \right) \, dx,
$$

which is bounded by

$$
\frac{1}{t} \int \left| \frac{d}{dx} \left( \frac{1}{it\phi'(x)} \psi(x) \right) \right| \, dx.
$$

But in fact we can keep integrating by parts, infinitely many times, and gain arbitrarily many powers of decay: $\leq C_N t^{-N}$ for each $N \geq 0$. Boundary terms change this.

Example:

$$
f_2(t) = \int_{-1}^{1} e^{it\sin x} \, dx.
$$

Check for stationary points.

**PROBLEM 19.** Integrating by parts, show that for $t > 0$

$$
f(t) = \int_{0}^{1} e^{it(3x-x^2)} \, dx
$$

is bounded by $Ct^{-1}$ for some constant $C$. Can one prove faster decay?

**PROBLEM 20.** Integrating by parts, show that for $t > 0$

$$
g(t) = \int_{-\infty}^{\infty} e^{ite^x} e^{-e^x} \, dx
$$

is bounded by $Ct^{-2}$ for some constant $C$.

If $\phi' = 0$ at some points, but not everywhere, then we divide the domain of integration into two parts: some small interval near the points where $\phi' = 0$ and some larger interval where we can integrate by parts.

For simplicity, let’s assume that $\phi$ has only one stationary point $c$, i.e. where $f'(c) = 0$. If $\phi''(c) \neq 0$, then in some neighborhood of $c$

$$
\phi(x) = \phi(c) + \frac{\phi''(c)}{2}(x-c)^2 + h.o.t..
$$

and

$$
\phi'(x) = \phi''(c)(x-c) + h.o.t..
$$

If instead there are finitely many stationary points, then we can split the domain, using cutoff functions, so that each piece contains at most one of them.

We subdivide the domain into three pieces, $(-\infty, c - \delta)$, $(c - \delta, c + \delta)$, and $(c + \delta, \infty)$.

Then on the first and third piece we get through integration by parts a bound of size $\frac{1}{t\phi''(c)}$ and on the middle piece, where integration by parts doesn’t gain anything, we have a bound of size $\delta$:

$$
\leq C \left( \frac{1}{\delta\phi''(c)} + \delta \right).
$$

Setting them equal, we get an overall bound of $C(t\phi''(c))^{-1/2}$.

This is called the method of stationary phase. Stationary points are those where the derivative is zero. It is extremely widespread in the study of PDEs and in harmonic analysis.
3. SEPARATION OF VARIABLES

**Problem 21.** Prove that the improper Riemann integral

$$\int_{-\infty}^{\infty} e^{ix^2} \, dx$$

converges, as follows: divide it into three portions, \((-\infty, -1], [-1, 1], [1, \infty)\), and use integration by parts on the first and last portions, as follows:

$$\int_{1}^{\infty} e^{ix^2} \, dx = \int_{1}^{\infty} \frac{1}{2ix} [2ixe^{ix^2}] \, dx = \frac{1}{2ix} e^{ix^2} |_{x=1} - \int_{1}^{\infty} \left( \frac{\partial}{\partial x} \frac{1}{2ix} \right) e^{ix^2} \, dx.$$

Hint: After this integration by parts, the integrals become absolutely convergent.

**Example:** The modified Bessel function of the second kind

$$K_\alpha(x) = \int_{0}^{\infty} e^{-x \cosh t} \cosh(\alpha t) \, dt.$$

For small \(x\), this has size \(x^{-1/2}\). For large \(x\), this has exponential decay, but we can’t (easily) prove that just using stationary phase.

**Example:** Schrödinger’s equation in one dimension: the fundamental solution is

$$\mathcal{F}^{-1} e^{it\xi^2}$$

of size \(t^{-1/2}\).

The method also works in higher dimensions and/or if \(f''(c) = 0\). In general, if the first nonvanishing derivative is the \(N + 1\)-th-order one, then we get a term of size \(t^{-d/N}\), where \(d\) is the dimension.

If all derivatives vanish at \(c\), then decay is harder to prove and it’s more likely logarithmic, not polynomial.

**Examples:**

- The inverse Fourier transform of \(e^{it\xi^N}\) is of size \(t^{-1/N}\), so this says something about the decay of solutions to evolution equations.

- We can compute in this manner the decay rates of solutions to every equation of the form

$$u_{tt} - P(D)u = 0,$$

where \(P\) is a polynomial. It all depends on the roots of \(P'\).

**2.4. The Klein–Gordon equation:** The **Klein–Gordon equation** in one dimension is

$$u_{tt} - u_{xx} + u = 0.$$

The solution is

$$u(x,t) = u_0 * \mathcal{F}^{-1} [\cos(t\sqrt{\xi^2 + 1})] + u_1 * \mathcal{F}^{-1} [\frac{\sin(t\sqrt{\xi^2 + 1})}{\sqrt{\xi^2 + 1}}].$$

Let’s call the first one the cosine kernel and the second one the sine kernel. Here

$$\partial_t S_t = C_t.$$

Let’s compute

$$\int_{-\infty}^{\infty} e^{ix\xi} e^{it\sqrt{\xi^2 + 1}} \, d\xi.$$

Setting \(\xi = \sinh u\), we get

$$\int_{-\infty}^{\infty} e^{i(x \sinh u + t \cosh u)} \cosh u \, du.$$
Let
\[ \frac{t}{\sqrt{t^2 - x^2}} = \cosh v, \quad \frac{x}{\sqrt{t^2 - x^2}} = \sinh v. \]
We get
\[ \int_{-\infty}^{\infty} e^{i\sqrt{t^2 - x^2} \cosh(u + v)} \cosh u \, du = \int_{-\infty}^{\infty} e^{i\sqrt{t^2 - x^2} \cosh u} \cosh(u - v) \, du. \]
We'll compute the cosine kernel (later?).

2.5. Bessel functions. Bessel functions are solutions of the ODE
\[ \frac{d^2 y}{dx^2} + \frac{1}{x} \frac{dy}{dx} + \left(1 - \frac{\alpha^2}{x^2}\right)y = 0. \]

For each \( \alpha \) this ODE has two linearly independent solutions, called Bessel’s functions of the first, \( J_\alpha \), and second, \( Y_\alpha \), kinds.

Closely related are the modified Bessel functions \( I_\alpha \) and \( K_\alpha \), which are solutions of
\[ \frac{d^2 y}{dx^2} + \frac{1}{x} \frac{dy}{dx} + \left(1 + \frac{\alpha^2}{x^2}\right)y = 0. \]

The relation between the Bessel functions and the modified Bessel functions is the same as the one between trigonometric functions and the exponential.

In this course the most important Bessel functions will be \( J_0 \) and \( J_1 \).

Same as most other special functions, Bessel functions can be characterized in several ways. They have a particularly simple power series expansion:
\[ J_\alpha(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m + \alpha + 1)} \left(\frac{x}{2}\right)^{2m+\alpha}, \]

They also have integral representations: for integer \( n \)
\[ J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n\tau - x \sin \tau)} \, d\tau. \]

In particular
\[ J_0(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ix \sin \tau} \, d\tau = \frac{1}{\pi} \int_{-1}^{1} \frac{e^{ixu}}{\sqrt{1 - u^2}} \, du. \]

So \( J_0 \) is (up to a constant) the Fourier transform of \((1 - u^2)^{-1/2} \). Likewise, \( J_1 \) is the Fourier transform of \((1 - u^2)^{1/2} \).

2.6. The two-dimensional wave equation. We get
\[ \int_{0}^{\infty} \int_{0}^{2\pi} e^{i\rho \cos \theta} \sin(t\rho) \, d\theta \, d\rho = 2\pi \int_{0}^{\infty} J_0(t\rho) \sin(t\rho) \, d\rho = \frac{2\pi}{t} \int_{0}^{\infty} J_0\left(\frac{r}{t}\right) \sin(r) \, dr. \]

\[ J_\nu(z) = \frac{(\frac{z}{2})^\nu}{\Gamma(\nu + \frac{1}{2}) \sqrt{\pi}} \int_{-1}^{1} e^{isz} (1 - s^2)^{\nu - \frac{1}{2}} \, ds, \]
\[ = \frac{2}{(\frac{z}{2})^\nu \cdot \sqrt{\pi} \cdot \Gamma\left(\frac{1}{2} - \nu\right)} \int_{1}^{\infty} \frac{\sin(zu)}{(u^2 - 1)^{\nu + \frac{1}{2}}} \, du, \]
2.7. Inhomogenous equations. Duhamel’s formula.

\[ u_t = u_{xx} + F, \ u(0) = u_0. \]

Then in the same way we get

\[ \frac{d}{dt} \hat{u} = -\xi^2 \hat{u} + \hat{F}. \]

Solving this ODE, we see that

\[ \hat{u}(t) = e^{-t\xi^2} \hat{u}_0 + \int_0^t e^{-(t-s)\xi^2} \hat{F}(s) \, ds. \]

This is called Duhamel’s formula. Reversing the Fourier transform, we get another version of Duhamel’s formula:

\[ u(x, t) = u_0 * H_t + \int_0^t F(s) * H_{t-s} \, ds. \]

The source term \( F(s) \) has the same effect as initial data would have when placed at time \( s \), but the initial data occur once at \( t = 0 \), while the forcing term keeps intervening in an ongoing manner, at all times, hence the need for an integral.

2.8. Summary.

2.8.1. First-order equations. For a first order in time equation on \( \mathbb{R}^d \times \mathbb{R} \) of the form

\[ u_t - P(D)u = F, \ u(x, 0) = u_0(x), \]

we perform a Fourier transform and the solution has the form

\[ \hat{u}(\xi, t) = e^{iP(i\xi)} \hat{u}_0(\xi) + \int_0^t e^{i(t-s)P(i\xi)} \hat{F}(\xi, s) \, ds. \]

Then, reverting the Fourier transform, we get

\[ u(x, t) = \left[ \frac{1}{(2\pi)^d} e^{-i|x|^2} \right] * u_0(x) + \int_0^t \left[ \frac{1}{(2\pi)^d} e^{-i(t-s)|\xi|^2} \right] * F(x, s) \, ds. \]

The convolution kernel

\[ \frac{1}{(2\pi)^d} e^{-iP(i\xi)} \]

is then called the fundamental solution of the equation.

Examples: the heat equation, Schrödinger’s equation, and Stokes’ equation.

The heat equation:

\[ u(x, t) = \left[ \frac{1}{(2\pi)^d} e^{-t||\xi||^2} \right] * u_0(x) + \int_0^t \left[ \frac{1}{(2\pi)^d} e^{-(t-s)||\xi||^2} \right] * F(x, s) \, ds. \]

For the heat equation, the fundamental solution / heat kernel is given by

\[ H_t(\vec{x}) = \frac{1}{(2\pi)^d} e^{-t||\vec{x}||^2} = \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|\vec{x}|^2}{4t}}. \]

Using the method of stationary phase, we would be getting that the solution decays like \( t^{-d/2} \), which is true, but does not capture the most important fact about the heat kernel, namely the fact that it’s given by a Gaussian.

Schrödinger’s equation:

\[ u(x, t) = \left[ \frac{1}{(2\pi)^d} e^{it||\xi||^2} \right] * u_0(x) + \int_0^t \left[ \frac{1}{2\pi} e^{i(t-s)||\xi||^2} \right] * F(x, s) \, ds. \]

The fundamental solution / Schrödinger kernel is given by

\[ S_t(\vec{x}) = \frac{1}{(2\pi)^d} e^{it||\vec{x}||^2} = \frac{1}{(4\pi it)^{d/2}} e^{-\frac{|\vec{x}|^2}{4it}}. \]
Stokes’ equation: For Stokes’ equation on $\mathbb{R}^d \times \mathbb{R}$, the solution is given by
\[
x(x, t) = \left[\frac{1}{(2\pi)^d} e^{-i|\xi|^2} \right] * u_0(x) + \int_0^t \left[\frac{1}{2\pi} e^{-i(t-s)|\xi|^2} \right] * F(x, s) \, ds.
\]
The fundamental solution is given by
\[
K_t(x) = \frac{1}{2\pi} F^{-1} e^{ix\xi} = \frac{1}{(3t)^{1/3}} \text{Ai} \left( \frac{x}{(3t)^{1/3}} \right).
\]

**Problem 22.** Prove that a solution of the equation
\[
u_t = -u_{xxxx}, \ u(x, 0) = u_0(x)
\]
is given by
\[
u(x, t) = u_0(x) * \frac{1}{2\pi} F^{-1} e^{-t|\xi|^4}.
\]
Prove that $f(x) = F^{-1} e^{-\xi^4}$ is a Schwartz-class function that fulfills the following ODE:
\[
f'' = \frac{x}{4} f, \ \hat{f}(0) = 1, \ f'(0) = 0, \ f''(0) = 0 \text{ (hint: the last one is obvious)}.
\]
Express the fundamental solution of this equation in terms of $f$.

Extra credit: Find $f(0)$ and $f''(0)$ in terms of the $\Gamma$ function
\[
\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} \, dt.
\]
Prove that $f(x)$ has size (roughly) $e^{-|x|^{1/3}}$.

2.8.2. Second-order equations. For a second-order equation on $\mathbb{R}^d \times \mathbb{R}$ of the form
\[
u_{tt} - P(-iD)u = F, \ u(0) = u_0, \ u_t(0) = u_1,
\]
the solution is given by
\[
\hat{u}(\xi, t) = \hat{u}_0(\xi) \cos(t \sqrt{\text{P}(\xi)}) + \hat{u}_1(\xi) \frac{\sin(t \sqrt{\text{P}(\xi)})}{\sqrt{\text{P}(\xi)}} + \int_0^t \sin((t-s) \sqrt{\text{P}(\xi)}) \hat{F}(s) \, ds.
\]
Then, by reverting the Fourier transform,
\[
u(x, t) = u_0(x) * \frac{1}{2\pi} F^{-1} \cos(t \sqrt{\text{P}(\xi)}) + u_1(x) * \frac{1}{2\pi} F^{-1} \frac{\sin(t \sqrt{\text{P}(\xi)})}{\sqrt{\text{P}(\xi)}} + \int_0^t F(x, s) * \frac{1}{2\pi} F^{-1} \sin((t-s) \sqrt{\text{P}(\xi)}) \, ds.
\]
Second-order in time equations have two fundamental solutions, the sine kernel and the cosine kernel
\[
S_t = \frac{1}{2\pi} e^{-t \sin(t \sqrt{\text{P}(\xi)})}, \ C_t = \frac{1}{2\pi} e^{-t \cos(t \sqrt{\text{P}(\xi)})}.
\]
The cosine kernel is the $t$ derivative of the sine kernel, so it suffices to compute one of them to get the other almost for free.

Equations that can be solved this way: the wave equation, the plate equation, Poisson’s equation to some extent.

The wave equation: For the wave equation on $\mathbb{R}^d \times \mathbb{R}$
\[
u_{tt} - \Delta u = F, \ u(x, 0) = u_0(x), \ u_t(x, 0) = u_1(x),
\]
the solution is given by
\[
\hat{u}(\xi, t) = \hat{u}_0(\xi) \cos(t |\xi|) + \hat{u}_1(\xi) \frac{\sin(t |\xi|)}{|\xi|} + \int_0^t \frac{\sin((t-s)|\xi|)}{|\xi|} \hat{F}(\xi, s) \, ds.
\]
After reverting the Fourier transform, we get
\[ u(x,t) = u_0 * \frac{1}{2\pi} \mathcal{F}^{-1} \cos(t|\xi|) + u_1(x) * \frac{1}{2\pi} \mathcal{F}^{-1} \frac{\sin \left( \frac{t|\xi|}{\xi} \right)}{|\xi|} + \int_0^t F(x,s) * \frac{1}{2\pi} \mathcal{F}^{-1} \frac{\sin \left( (t-s)\sqrt{|\xi|} \right)}{|\xi|}. \]

The two fundamental solutions of this equation are
\[ S_t = \frac{1}{2\pi} \mathcal{F}^{-1} \frac{\sin \left( \frac{t|\xi|}{\xi} \right)}{\sqrt{|\xi|}}, \quad C_t = \frac{1}{2\pi} \mathcal{F}^{-1} \cos(t|\xi|). \]

The precise formulas depend on the dimension, with important differences between even and odd dimensions. See: Huygens principle versus sharp Huygens principle.

**The Poisson equation** For the Poisson equation, the solution is given by
\[ \hat{u}(\xi,t) = e^{-t|\xi|}\hat{u}_0(\xi). \]

After reverting the Fourier transform,
\[ u(x,t) = u_0(x) * \frac{1}{2\pi} \mathcal{F}^{-1} e^{-t|\xi|}. \]

### 3. Fourier series

On bounded domains we also solve linear PDEs by separation of variables. This entails expanding the initial data in an appropriate manner, i.e. in terms of the values of separable solutions at time zero.

In general, for the equation
\[ u_t = Lu, \quad u(0) = u_0 \]
we express the initial data \( u_0 \) in terms of the eigenfunctions of \( L \): \( Le_n = \lambda_n e_n \). The eigenfunctions depend on the shape of the domain and on the boundary conditions.

If the initial data are \( u_0 = \sum_n c_n e_n \), then the solution is given by \( u(t) = \sum_n c_n e^{t\lambda_n} e_n \).

The abstract problem is then as follows: given a fixed set of basis functions \( \{e_n\} \), we want to express the initial data \( u_0 \) as a linear combination of these basis functions.

There are at least two ways to do this and accordingly we call \( \{e_n\} \) an algebraic basis or a Schauder basis.

**Definition 13.** The set \( \{e_j : j \in J\} \) is called an algebraic/Hamel basis for the vector space \( V \) if and only if every \( v \in V \) can be uniquely written as a finite linear combination of elements of this set.

Any vector space is always guaranteed to possess an algebraic basis by Zorn’s lemma / the axiom of choice.

However, algebraic bases are useless in our context, because in general there is no algebraic basis made of eigenfunctions of \( L \). We have to deal with Schauder bases instead. We replace finite with infinite linear combinations. For that, we have to be able to define a notion of convergence in the given vector space.

**Definition 14.** The set \( \{e_n : n \geq 0\} \) is called a Schauder basis for the vector space \( V \) if for every element \( v \in V \) there exist a unique sequence of coefficients \( c_n \) such that
\[ \lim_{N \to \infty} \sum_{n=0}^N c_n e_n = v. \]
When the basis is indexed over all integers, as is often the case, series are usually summed as follows:

\[
\lim_{N \to \infty} \sum_{n=-N}^{N} c_n e_n = v.
\]

In particular, this implies that finite linear combinations of basis elements are dense in the vector space \( V \). The converse is probably not true: a Banach space can be separable (have a dense countable subset) without there existing a (countable) Schauder basis.

When the domain is a line segment, a rectangle, a rectangular box, etc., such an expansion is called a Fourier series. In general, this is more properly called an eigenfunction expansion.

### 3.1. Periodic functions and their Fourier transforms.

**Definition 15.** A function \( f : \mathbb{R} \to \mathbb{C} \) is called periodic if there exists some number \( p \neq 0 \) such that \( f(x + p) = f(x) \) for every \( x \in \mathbb{R} \). All such numbers, potentially including \( p = 0 \) which always fits the definition, are called periods of the function \( f \).

Note that if \( p_1 \) and \( p_2 \) are periods for a given function \( f \) then so is \( p_1 + p_2 \), and if \( p \) is a period so is \( np \) for \( n \in \mathbb{Z} \). Thus, the set of periods of a given function \( f \) forms an abelian group (also called a \( \mathbb{Z} \)-module).

The smallest positive period of a function, if it exists, is called the principal period of that function. If a function has a principal period, then all its other periods are integer multiples of its principal period.

**Problem 23.** Suppose \( f : \mathbb{R} \to \mathbb{R} \) is periodic and has a least positive period \( p_0 \). Prove that all other periods of \( f \) are integer multiples of \( p_0 \).

**Periodic functions without principal periods:** Trivially, constant functions are periodic of every period \( p \in \mathbb{R} \). Less trivially, Dirichlet’s function

\[
\chi_{\mathbb{Q}}(x) = \begin{cases}
1, & x \in \mathbb{Q}, \\
0, & x \notin \mathbb{Q}
\end{cases}
\]

is a periodic function admitting every rational number as a period.

In some sense, this and similar ones are the only counterexamples:

**Problem 24.** Prove that if \( f : \mathbb{R} \to \mathbb{R} \) is a periodic function such that \( f \) is continuous and has no principal period then \( f \) is constant.

In higher dimensions, the situation is more complicated, but in principle “nice” periodic functions on \( \mathbb{R}^d \) have between 1 and \( d \) periods in different directions.

**Problem 25.** If a continuous function is periodic, but does not have a principal period, then it must be constant.

#### 3.1.1. Periodic boundary conditions.** We say that a function \( f \in C^1([a,b]) \) satisfies periodic boundary conditions if

\[
f(a) = f(b), \quad f'(a) = f'(b).
\]

Periodic boundary conditions enable us to extend a \( C^1 \) function defined on a bounded interval to a \( C^1 \) periodic function on \( \mathbb{R} \).

**Proposition 3.1.** Given a function \( f \in C^1([a,b]) \), there exists an extension \( g \in C^1(\mathbb{R}) \) of \( f \) to \( \mathbb{R} \) (i.e. such that \( g(x) = f(x) \) for \( x \in [a,b] \)), periodic of period \( b - a \), if and only if \( f \) satisfies periodic boundary conditions.
Then we have conditions hold for product on $C$

This can be computed using Euler’s formula and the uniqueness of the Fourier expansion (in each equivalent to the first, Schauder basis: $f$

This is called the complex Fourier expansion/series of $f$. Sometimes denoted $\hat{f}(n)$.

In other words, we have employed the Schauder basis $\{e^{2\pi i nx/p} : n \in \mathbb{Z}\}$ to expand $f$.

One can also do a Fourier expansion in terms of sines and cosines. This is a second, and equivalent to the first, Schauder basis:

$$\{\cos(2\pi nx/p) : n \geq 0\} \cup \{\sin(2\pi nx/p) : n > 0\}.$$  

Then we have

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi nx/p) + \sum_{n=1}^{\infty} b_n \sin(2\pi nx/p).$$

This is the real Fourier expansion of $f$, with periodic boundary conditions, and the $a_n$ and $b_n$ are called the real Fourier coefficients.

There is a relation between the two kinds of coefficients: for $n \geq 0$

$$c_n = \frac{1}{2} (a_n - ib_n), \quad c_{-n} = \frac{1}{2} (a_n + ib_n), \quad a_n = c_n + c_{-n}, \quad b_n = i(c_n - c_{-n}).$$

This can be computed using Euler’s formula and the uniqueness of the Fourier expansion (in each basis).

For an even function we only need cosines, for an odd function we only need sines.

How do we compute these coefficients? Using orthogonality. For an interval $I$, define a dot product on $C^1(I)$ by

$$\langle f, g \rangle = \int_I f(x)\overline{g(x)} \, dx.$$ 

For two periodic functions of same period $p$, we can define and compute their dot product by this formula, using any interval of length $|I| = p$; then the dot product is independent of the choice of an interval.

The complex exponentials $e^{2\pi inx/p}$ are orthogonal to each other: for $n \neq m$,

$$\langle e^{2\pi inx/p}, e^{2\pi imx/p} \rangle = \int_0^p e^{2\pi inx/p} e^{2\pi imx/p} \, dx = \int_0^p e^{2\pi i(n-m)x/p} \, dx = \frac{e^{2\pi i(n-m)x/p}}{2\pi i(n-m)} \bigg|_{x=0}^{x=p} = \frac{e^{2\pi i(n-m)} - 1}{2\pi i(n-m)} = 0.$$ 

Since $e^{2\pi i} = 1$, we get 0 on the previous line.

For $n = m$ we get the norm of $e^{2\pi inx/p}$:

$$\|e^{2\pi inx/p}\|^2 = \langle e^{2\pi inx/p}, e^{2\pi inx/p} \rangle = \int_0^p 1 \, dx = p.$$
Next, if
\[ f = \sum_n c_n e_n \]
and the basis elements \( e_n \) are orthogonal to each other, then
\[ c_m = \frac{\langle f, e_m \rangle}{\|e_m\|^2}. \tag{25} \]
Indeed, formally
\[ \langle f, e_m \rangle = \left\langle \sum_n c_n e_n, e_m \right\rangle = \sum_n c_n \langle e_n, e_m \rangle = c_m \langle e_m, e_m \rangle = c_m \|e_m\|^2. \]
In this infinite sum only one term is nonzero, because, whenever \( m \neq n \), \( e_m \) and \( e_n \) are orthogonal, i.e. \( \langle e_n, e_m \rangle = 0 \). The formal part is that we interchanged the infinite sum and the dot product without rigorous justification. The rigorous proof is below.

For the concrete case of periodic functions,
\[ c_n = \frac{1}{p} \int_I f(x)e^{-2\pi inx/p}. \]

### 3.2. Convolution of periodic functions.
Convolution is also defined for functions on a finite interval \( I \in \mathbb{R} \).
However, this only works for periodic functions, not for Dirichlet, Neumann, or Robin boundary conditions.

**Definition 16.** Given two periodic functions of period \( p \),
\[ [f * g](x) = \frac{1}{p} \int_{I, |I|=p} f(x - y)g(y) \ dy. \]

The coefficient in front of the integral is a matter of convention.
The value of this integral does not depend on the chosen interval \( I \).
More generally, convolution makes sense for functions defined on a homogeneous space, such as the surface of a sphere.

**Properties:**
* Convolution is linear in both arguments.
* The Fourier transform interchanges multiplication and convolution:
\[ \hat{f} \ast \hat{g}(n) = \hat{f}(n)\hat{g}(n). \]

Convolution of sequences:
\[ [(a_n) \ast (b_n)]_n = \sum_{m \in \mathbb{Z}} a_m b_{n-m}. \]

Then it is also true that
\[ \hat{f} \ast \hat{g} = \hat{f} * \hat{g}. \]

The inverse Fourier transform + Fourier transform can be expressed in terms of a convolution:
\[ \mathcal{F}^{-1}[\mathcal{F}f] = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2\pi inx/p} = \sum_{n=-\infty}^{\infty} \frac{1}{p} \left( \int_0^{p} f(y)e^{-2\pi iny/p} \ dy \right) e^{2\pi inx/p} = \int_0^{p} f(y)\frac{1}{p} \left( \sum_{n=-\infty}^{\infty} e^{2\pi in(x-y)/p} \right) \ dy. \]
This is just a formal computation, because the sum does not converge in the usual sense. However, partial sums of the Fourier series can be expressed as a convolution:
\[ S_N = \sum_{n=-N}^{N} \hat{f}(n)e^{2\pi inx/p} = \sum_{n=-N}^{N} \frac{1}{p} \left( \int_0^{p} f(y)e^{-2\pi iny/p} \ dy \right) e^{2\pi inx/p} = \int_0^{p} f(y)\frac{1}{p} \left( \sum_{n=-N}^{N} e^{2\pi in(x-y)/p} \right) \ dy. \]
Here this is a finite sum, so there are no convergence issues.
3.3. Convergence of Fourier series. In what sense does the Fourier series converge to (or represent) the original function? As a rule, the smoother a function is, the faster and “better” the convergence of its Fourier series to it.

Here is a standard result about the convergence of Fourier series.

**Proposition 3.2.** If \( f \in C^1 \) is periodic, then its Fourier expansion exists and is unique. Its Fourier coefficients are absolutely summable

\[
\sum_{n=-\infty}^{\infty} |c_n| < \infty
\]

and its Fourier series converges uniformly to \( f \) for all \( x \).

Assuming this is true, it justifies the interchanging of summation and integration above, see Proposition 0.13.

**Proof.** Uniqueness: Using the summability of the Fourier coefficients: Suppose \( f \) admits the Fourier expansion

\[
f(x) = \sum_{n} c_n e^{2\pi inx/p}
\]

and let

\[
S_N(x) = \sum_{n=-N}^{N} c_n e^{2\pi inx/p}.
\]

The sum of the Fourier coefficients controls the Fourier series pointwise, so \( S_N \) converges uniformly to \( f \). Therefore

\[
\int_{0}^{p} f(x) e^{-2\pi inx/p} \, dx = \lim_{N \to \infty} \int_{0}^{p} \left( \sum_{n=-N}^{N} c_n e^{2\pi inx/p} \right) e^{-2\pi inx/p} \, dx
\]

\[
= \lim_{N \to \infty} \sum_{n=-N}^{N} \left( \int_{0}^{p} c_n e^{2\pi inx/p} e^{-2\pi inx/p} \, dx \right)
\]

\[
= pc_n,
\]

because the sequence is eventually constant. In a finite sum we have no problem interchanging the sum and the integral and the uniform convergence takes care of the limit.

So

\[
c_n = \frac{1}{p} \int_{0}^{p} f(x) e^{-2\pi inx/p} \, dx,
\]

which is a uniquely determined value.

Existence: First we prove summability.

If \( f \in C \) (if \( f \in L^1 \) really), then its Fourier coefficients are uniformly bounded:

\[
|c_n| = \left| \frac{1}{p} \int_{0}^{p} f(x) e^{-2\pi inx/p} \, dx \right| \leq \frac{1}{p} \int_{0}^{p} |f(x)| e^{-2\pi inx/p} \, dx \leq \frac{1}{p} \int_{0}^{p} |f(x)| \, dx = \frac{1}{p} \|f\|_{L^1} < \infty.
\]

This computation actually applies to the Fourier coefficients of \( f' \in C \), not only those of \( f \), so those coefficients are bounded: \( |\hat{f}'(n)| \leq \frac{1}{p} \|f'\|_{L^1} \). But

\[
\hat{f}'(n) = \frac{2\pi in}{p} \hat{f}(n),
\]

so this implies that

\[
|\hat{f}(n)| \leq \frac{p}{2\pi n} \frac{1}{p} \|f'\|_{L^1} \leq \frac{C}{|n|}.
\]

However, this still falls short of proving summability, because \( \sum \frac{1}{n} \) diverges logarithmically. The next step is using Bessel’s inequality for the derivative and then Cauchy–Schwartz.
Proposition 3.3 (Bessel’s inequality). Let \( u \in H \), where \( H \) is a Hilbert space. Given a set of orthogonal vectors \( \{e_n : n \in \mathbb{Z}\} \),

\[
\|u\|^2 \geq \sum_n |c_n|^2 \|e_n\|^2, \quad c_n = \frac{\langle u, e_n \rangle}{\|e_n\|^2}.
\]

This inequality holds regardless of whether \( e_n \) span the whole space. If equality actually holds for every \( u \in H \), then \( e_n \) will span the whole space; however, Bessel’s inequality holds in a more general setting.

Proof of Bessel’s inequality. We say that two vector subspaces \( V_1, V_2 \subset H \) are orthogonal, \( V_1 \perp V_2 \), if and only if, for every \( v_1 \in V_1 \) and \( v_2 \in V_2 \), \( \langle v_1, v_2 \rangle = 0 \).

Suppose \( v = v_1 + v_2 \) and \( V_1 \perp v_2 \); then Pythagoras’ theorem holds:

\[
\|v\|^2 = \|v_1 + v_2\|^2 = \langle v_1 + v_2, v_1 + v_2 \rangle = \langle v_1, v_1 \rangle + \langle v_1, v_2 \rangle + \langle v_2, v_1 \rangle + \langle v_2, v_2 \rangle = \|v_1\|^2 + \|v_2\|^2,
\]

since \( \langle v_2, v_1 \rangle = 0 \). More generally, if \( v_1, \ldots, v_n \) are mutually perpendicular, then expanding in the same manner we get

\[
\|v_1 + \ldots + v_n\|^2 = \|v_1\|^2 + \ldots + \|v_n\|^2.
\]

Orthogonal complement: Given a vector subspace \( V \subset H \), we define its orthogonal complement \( V^\perp \) as the set of those vectors perpendicular to all vectors in \( V \):

\[
V^\perp = \{u \in H : \forall v \in V, \langle u, v \rangle = 0\}.
\]

Clearly \( V^\perp \) is a closed vector subspace of \( H \).

Projections. Suppose \( V \) is finite-dimensional (or, more generally, closed).

Claim: Every vector \( u \in H \) can then be uniquely decomposed as \( u = v + v^\perp \), where \( v \in V \) and \( v^\perp \in V^\perp \).

See the proof below.

The maps \( P(u) = v \) and \( \overline{P}(u) = v^\perp \) are linear. Such maps are called orthogonal projections and have the following two properties: if \( P(u) = v \), then \( P^2 = P \) (it is idempotent) and \( P^* = P \) (it is self-adjoint).

Proof: Uniqueness: Suppose

\[
u = v_1 + v_1^\perp = v_2 + v_2^\perp,
\]

with \( v_1, v_2 \in V \) and \( v_1^\perp, v_2^\perp \in V^\perp \). Then

\[
v_1 - v_2 = v_2^\perp - v_1^\perp = v
\]

is both in \( V \) (the left-hand side) and in \( V^\perp \) (the right-hand side). Since by definition \( V^\perp \perp V \) and \( v \) is in both, it follows that \( v \perp v \), so \( \|v\|^2 = \langle v, v \rangle = 0 \), so \( v = 0 \), so \( v_1 = v_2 \) and \( v_1^\perp = v_2^\perp \).

Existence: We’ll only prove it in the case when \( V \) is finite-dimensional, which is all we need. More generally, when \( V \) is closed, the proof uses the Banach–Alaoglu theorem.

Given \( u \in H \), let \( v \in V \) be a vector for which \( \|u - v\| \) (or equivalently \( \|u - v\|^2 \)) is minimized. We need to show two things: that such a vector \( v \) exists and that \( u - v = v^\perp \in V^\perp \).

Existence: For every \( v \in V \), if \( \|v\| > 10\|u\| \), then \( \|u - v\| \geq \|v\| - \|u\| > 9\|u\| \). On the other hand, for \( \|v\| \leq \|u\| \), \( \|u - v\| \leq \|u\| + \|v\| \leq 2\|u\| \). Thus,

\[
\min\{\|u - v\| : v \in V\} = \min\{\|u - v\| : v \in V, \|v\| \leq 10\|u\|\}.
\]

In other words, if a minimum is attained inside the ball of radius \( 10\|u\| \), then it is the minimum on the whole \( V \).

Now \( V \) is finite-dimensional, so the ball of radius \( 10\|u\| \), being closed and bounded, is compact. Therefore the continuous function \( \|u - v\| \) reaches its minimum somewhere inside this ball.

Let \( v \) be a point where the minimum of \( \|u - v\|^2 \) is reached and consider any \( h \in V \). Then

\[
\|u - v - \epsilon h\|^2 = \|u - v\|^2 - 2\epsilon \Re \langle u - v, h \rangle + \epsilon^2 \|h\|^2.
\]
This is a second degree polynomial in $\epsilon$, which reaches a minimum at $\epsilon = 0$, since the distance from $u$ to $v + \epsilon h \in V$ is at least as big as the one from $u$ to $V$. So its first derivative as a function of $\epsilon$ must be zero:

$$0 = \frac{d}{d\epsilon} \|u - v - \epsilon h\|^2 = \frac{d}{d\epsilon} (\|u - v\|^2 - 2\epsilon \text{Re}\langle u - v, h \rangle + \epsilon^2 \|h\|^2) = -2 \text{Re}\langle u - v, h \rangle.$$ 

So we get that $\text{Re}\langle u - v, h \rangle = 0$. Replacing $h$ by $ih \in V$, we also get that $\text{Im}\langle u - v, h \rangle = 0$, so overall $\langle u - v, h \rangle = 0$ — and this is true for any $h \in V$, implying that $u - v \in V^\perp$.

Now, since $u = v + v^\perp$ and $v \perp v^\perp$, it follows that

$$\|u\|^2 = \|v\|^2 + \|v^\perp\|^2 \geq \|v\|^2.$$

Now let’s return to the orthogonal set $\{e_n\}$. Let $V$ be the vector subspace spanned by $e_n$, $-N \leq n \leq N$, for some fixed $N$. Since it’s a finite dimensional vector space, from the proof above we get that $u = v + v^\perp$, $v \in V$, $v^\perp \in V^\perp$, and $\|u\| \geq \|v\|$.

Since $u - v \perp V$, in particular $u - v \perp e_n$, $-N \leq n \leq N$, so for $-N \leq n \leq N$

$$e_n = \frac{\langle u - v, e_n \rangle}{\|e_n\|^2} + \frac{\langle v, e_n \rangle}{\|e_n\|^2} e_n = \frac{\langle v, e_n \rangle}{\|e_n\|^2} e_n.$$

But $v$ is in the space spanned by $e_n$, $-N \leq n \leq N$, so it can be written as

$$v = \sum_{n=-N}^{N} d_ne_n$$

for some coefficients $d_n \in \mathbb{C}$. An immediate computation shows that in fact $d_n = c_n$:

$$\langle v, e_m \rangle = \left\langle \sum_{n=-N}^{N} d_n e_n, e_m \right\rangle = \sum_{n=-N}^{N} \langle d_n e_n, e_m \rangle = \sum_{n=-N}^{N} \langle d_n e_n, e_m \rangle = d_m \|e_m\|^2.$$

Here we have no problem interchanging the sum and the dot product, because it’s a finite sum. So

$$d_n = \frac{\langle v, e_n \rangle}{\|e_n\|^2} = c_n.$$

Now from Pythagoras’ theorem

$$\|u\|^2 \geq \|v\|^2 = \sum_{n=-N}^{N} c_n^2 = \sum_{n=-N}^{N} \|e_n\|^2 = \sum_{n=-N}^{N} \|c_n\|^2 \|e_n\|^2.$$ 

Up to here, everything happened in a finite-dimensional space of dimension $2N + 1$, but the inequality does not depend on $N$. Finally, letting $N \to \infty$, we obtain the conclusion. □

By Bessel’s inequality, applied to $f'$,

$$\sum_{n} |\hat{f}'(n)|^2 \leq \frac{1}{p} \int_0^p |f'(x)|^2 \, dx < \infty.$$ 

Next, we again note the connection between the Fourier coefficients of a function and those of its derivative:

$$\hat{f}'(n) = \frac{2\pi in}{p} \hat{f}(n).$$ 

Then

$$\sum_{n \neq 0} |\hat{f}(n)| = \frac{p}{2\pi} \sum_{n \neq 0} \frac{1}{|n|} |\hat{f}'(n)| \leq \frac{p}{2\pi} \left( \sum_{n \neq 0} \frac{1}{n^2} \right)^{1/2} \left( \sum_{n \neq 0} |\hat{f}'(n)|^2 \right)^{1/2} < \infty.$$

The case $n = 0$ has to be treated separately, but one easily sees that

$$|\hat{f}(0)| \leq C \|f\|_{L^1} \leq \tilde{C} \|f\|_C < \infty.$$ 

This concludes the proof of the summability of the Fourier coefficients.
Problem 26. a) By going through the previous proof, find the value of the constant $C$ such that for every periodic $C^1$ function $f$ of period $p$

$$
\sum_{n=-\infty}^{\infty} |c_n| \leq C \|f\|_{C^1},
$$

where by definition

$$
\|f\|_{C^1} = \max_{x \in [0,p]} |f(x)| + \max_{x \in [0,p]} |f'(x)|.
$$

Is your constant optimal?

b) Prove that there exists a constant $\tilde{C}$, depending on $f$, such that for $N > 0$

$$
\sum_{|n| \geq N} |c_n| \leq \frac{\tilde{C}}{\sqrt{N}} \|f\|_{C^1}.
$$

Hint: use the Cauchy–Schwartz inequality.

What does this imply about the Fourier series of $f$?

Now that we know that the Fourier coefficients $\hat{f}(n)$ are summable, let

$$
\hat{f}(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2\pi inx/p}.
$$

The sum of $|\hat{f}(n)|$ dominates the right-hand side series pointwise, so it is well-defined. By uniform convergence, it follows that $\hat{f}$ is continuous (but not necessarily $C^1$) and has the Fourier coefficients $\hat{f}(n)$, same as $f$.

At this point, though, we still don’t know that $f = \hat{f}$. So let $g = f - \hat{f}$. Since $f$ and $\hat{f}$ are continuous, so is $g$. Furthermore,

$$
\hat{g}(n) = \hat{f}(n) - \hat{\hat{f}}(n) = 0.
$$

If all the Fourier coefficients of a continuous function are zero, does it imply that the function is zero?

This is not true in general. If the function is not continuous, then one can easily give an example where it is not zero, even though its Fourier coefficients are all zero.

Example: A function which is nonzero at a finite number of points has zero Fourier coefficients. For continuous functions, however, this is true:

**Proposition 3.4.** For a continuous periodic function $f : \mathbb{R} \to \mathbb{C}$, if all its Fourier coefficients are zero, then the function is zero.

**Proof.** If the Fourier coefficients of $f$ are zero, then its dot product with each element in the Fourier basis is zero:

$$
\int_0^p f(x)e^{-2\pi inx/p} \, dx = 0.
$$

Then the same is true for the dot product of $f$ and any finite linear combination of such complex exponentials:

$$
\int_0^p f(x) \sum_{n=-N}^{N} c_n e^{2\pi inx/p} \, dx = 0.
$$

These finite linear combinations are called trigonometric polynomials. When a function $g$ is uniformly approximated by trigonometric polynomials, we can take the uniform limit under the integral and get that

$$
\int_0^p f(x)g(x) \, dx = 0.
In particular, if we did this for \( g = f \) itself, then we would get that
\[
\int_0^p |f(x)|^2 \, dx = 0,
\]
so \( f \) has to be zero almost everywhere. Since \( f \) is continuous, this implies that \( f \) is zero. \( \square \)

Then, it suffices to prove that every continuous function on a closed interval can be uniformly approximated by trigonometric polynomials:

**Proposition 3.5.** Every continuous function \( f : [0, a] \to \mathbb{C} \) can be uniformly approximated by trigonometric polynomials.

**Proof.** Let \( S_N \) be the \( N \)-th partial sum of the Fourier series:
\[
S_N(x) = \sum_{n=-N}^{N} \hat{f}(n)e^{2\pi inx/p}.
\]

Since
\[
\hat{f}(n) = \frac{1}{p} \int_0^p f(y)e^{-2\pi icy/p} \, dy,
\]
we have that
\[
S_N(x) = \frac{1}{p} \sum_{n=-N}^{N} \int_0^p f(y)e^{-2\pi icy/p}e^{2\pi inx/p} \, dy = \frac{1}{p} \int_0^p f(y) \sum_{n=-N}^{N} e^{2\pi i(x-y)/p} \, dy = f * D_N,
\]
where
\[
D_N(x) = \sum_{n=-N}^{N} e^{2\pi inx/p}
\]
is called the Dirichlet kernel. Here we can interchange the sum and the integral with no problem, since it’s a finite sum.

This being the sum of a geometric progression, we can compute it explicitly:
\[
\sum_{n=-N}^{N} e^{2\pi inx/p} = e^{-2\pi iNx/p} \sum_{n=0}^{2N} e^{2\pi inx/p} = e^{-2\pi iNx/p} \frac{e^{2\pi i(2N+1)x/p} - 1}{e^{2\pi i/p} - 1} = \frac{e^{2\pi i(N+1/2)x/p} - e^{-2\pi i(N+1/2)x/p}}{e^{\pi i/p} - e^{-\pi i/p}} = \frac{\sin(2\pi (N + 1/4)x/p)}{\sin(\pi x/p)}.
\]

The subsequent proof (and others) would be simpler if \( D_N \) were a good kernel. However, \( D_N \) satisfies only one out of the three conditions in the definition of a good kernel:
1. Its integral is constantly equal to \( p \):
\[
\int_0^p D_N(x) \, dx = p.
\]

On the other hand,
2. It does not concentrate near zero: for every \( \delta > 0 \) and large \( N \)
\[
\int_{\delta}^{p-\delta} |D_N(x)| \, dx \sim \frac{2}{\pi} \ln \delta.
\]

Also,
3. The integral of its absolute value is not uniformly bounded, but increases logarithmically:
\[
\int_0^p |D_N(x)| \, dx \sim C \ln N.
\]
We use the Fejér kernel instead:

\[ F_N = \frac{1}{N+1} (S_0 + \ldots + S_N). \]

Convolution with the Fejér kernel corresponds to another method of summing the Fourier series, called Cesàro summation:

\[ \tilde{S}_N = f * F_N = \frac{1}{N+1} (S_0 + \ldots + S_N). \]

The Féjer kernel can also be computed explicitly as the sum of a geometric progression:

\[
\sum_{n=1}^{N} \sin(z(N + \frac{1}{2})) = \frac{1}{2i} \left( \sum_{n=1}^{N} e^{iz(N+1/2)} - e^{-iz(N+1/2)} \right)
\]

\[
= \frac{1}{2i} \left( \frac{e^{iz(N+3/2)} - e^{iz/2}}{e^{iz} - 1} - \frac{e^{-iz(N+3/2)} - e^{-iz/2}}{e^{-iz} - 1} \right)
\]

\[
= \frac{1}{2i} \frac{e^{iz(N+3/2)} - 2e^{iz/2} + e^{-iz(N+1/2)}}{e^{iz} - 1}
\]

\[
= \frac{1 - \cos(z(N + 1))}{2 \sin(z/2)}.
\]

So the Féjer kernel is

\[ F_N(x) = \frac{1 - \cos(2\pi(N + 1)x/p)}{\sin^2(\pi x/p)}. \]

In particular, unlike the Dirichlet kernel, which has oscillating sign, the Féjer kernel is nonnegative! It is also a good kernel, so for every continuous function \( f : [0, p] \to \mathbb{R} \) (which is therefore also uniformly continuous, by Cantor’s theorem) \( F_N * f \to f \) uniformly.

But \( \tilde{S}_N = F_N * f \) is, for each \( N \), a trigonometric polynomial, so \( f \) is uniformly approximated by trigonometric polynomials.

Thus \( f \) is equal to the sum of its own Fourier series, concluding the proof.

### 3.4. Parseval’s identity.

**Theorem 3.6.** Let \( f \in C^1([0, p]) \) be a function satisfying periodic boundary conditions, with the Fourier series

\[ f(x) = \sum_{n \in \mathbb{Z}} c_n e^{2\pi inx/p}. \]

Then

\[ \frac{1}{p} \int_0^p |f(x)|^2 \, dx = \sum_{n \in \mathbb{Z}} |c_n|^2. \]

**Proof.** We replace \( f \) by its Fourier series. Since \( c_n \) are absolutely summable, the Fourier series of \( f \) converges uniformly to \( f \). Thus we can interchange the integration and the summation to get

\[ \frac{1}{p} \int_0^p |f(x)|^2 \, dx = \sum_{m,n \in \mathbb{Z}} c_m \overline{c_n} \frac{1}{p} \int_0^p e^{2\pi imx/p} \overline{e^{2\pi inx/p}} \, dx = \sum_{m \in \mathbb{Z}} c_m \overline{c_m} = \sum_{m \in \mathbb{Z}} |c_m|^2. \]

Note that the right-hand side is always less than or equal to the left-hand side, by Bessel’s inequality. However, for \( C^1 \) functions, one can prove that they are equal.

In fact, Parseval’s identity is true for all \( L^2 \) functions, as we’ll see next.
3.5. Extending the domain by periodicity or reflection. Instead of using Fourier series, in some special cases we can just use the Fourier transform. Consider the heat equation, with periodic boundary conditions in space, on $[0, a] \times [0, \infty)$:

$$u_t = u_{xx}, \quad u(x, 0) = u_0(x), \quad u(0, t) = u(a, t), \quad u_x(0, t) = u_x(a, t),$$

where $u_0$ is the $C^1$ initial data, also satisfying periodic boundary conditions.

Then we can either use Fourier series to solve the equation, as above, getting

$$u(x, t) = \sum_{n=-\infty}^{\infty} \hat{u}_n(n) e^{-(2\pi n/p)^2 t} e^{2\pi inx/p} = u_0(x) \ast \sum_{n=-\infty}^{\infty} e^{-(2\pi n/p)^2 t} e^{2\pi inx/p}.$$

Or we can extend $u_0$ by periodicity to $\mathbb{R}$ and solve the equation by the Fourier transform on $\mathbb{R}$. The solution with periodic initial data that extends $u_0$ to $\mathbb{R}$, with period $a$, will be

$$u(x, t) = \int_{-\infty}^{\infty} u_0(x-y) \frac{1}{\sqrt{4\pi t}} e^{-\frac{|y|^2}{4t}} dy = \int_0^a u_0(x-y) \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} e^{-\frac{|y-na|^2}{4t}} dy.$$

This can be written in terms of Jacobi theta functions.

Extension to twice the interval: Take a function $f : [0, a] \rightarrow \mathbb{R}$ that satisfies Dirichlet boundary conditions and extend it to twice the interval so as to make it odd:

$$\tilde{f}(x) = \begin{cases} f(x), & x \in [0, a], \\ -f(-x), & x \in [-a, 0] \end{cases}$$

If $f$ satisfies Dirichlet boundary conditions, then $\tilde{f}$ satisfies periodic boundary conditions, so we can extend it by periodicity to the whole real line. So the solution to the heat equation with $f$ initial data and Dirichlet boundary conditions is

$$u(x, t) = \int_{-\infty}^{\infty} \tilde{f}(y) \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}} dy = \int_0^a \tilde{f}(y) \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} e^{-\frac{|x-y-2na|^2}{4t}} dy$$

$$= \int_0^a \tilde{f}(y) \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} \left( e^{-\frac{|x-y-2na|^2}{4t}} - e^{-\frac{|x+y-2na|^2}{4t}} \right) dy.$$ 

The final expression is not a convolution product, though we started from one.

Take a function $f : [0, a] \rightarrow \mathbb{R}$ that satisfies Neumann boundary conditions and extend it to twice the interval so as to make it odd:

$$\tilde{f}(x) = \begin{cases} f(x), & x \in [0, a], \\ f(-x), & x \in [-a, 0] \end{cases}$$

If $f$ satisfies Neumann boundary conditions, then $\tilde{f}$ satisfies periodic boundary conditions, so we can extend it by periodicity to the whole real line. The solution of the heat equation with these initial data is then given by

$$u(x, t) = \int_0^a \tilde{f}(y) \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} \left( e^{-\frac{|x-y-2na|^2}{4t}} + e^{-\frac{|x+y-2na|^2}{4t}} \right) dy$$

Conclusion: When the solution hits the boundary, under Dirichlet boundary conditions it is reflected back with the opposite sign, while under Neumann boundary conditions it is reflected back with the same sign.

**Problem 27.** Suppose the function $f : [0, a] \rightarrow \mathbb{R}$ satisfies Dirichlet boundary conditions at $0$ and Neumann boundary conditions at $a$. Extend it in some suitable way to a function defined on an interval four times larger and prove that the extension satisfies periodic boundary conditions.
Use this to compute the solution to the heat equation with initial data and these boundary conditions.

The same is true in higher dimensions, literally so in the case when the boundary or some portion of it is straight. For example, when the domain is a rectangle or a rectangular box, then we can enlarge the domain to the whole space $\mathbb{R}^d$ by paving $\mathbb{R}^d$ with mirrored copies of the initial data (with the same sign or with the opposite sign, according to whether).

After we solve the problem on the whole space, using the Fourier transform, in particular this will also be a solution back on the original rectangle/rectangular box.

What if the shape of the domain is more complicated?

Suppose only some portion of the boundary is straight and has Dirichlet or Neumann boundary conditions; then we can reflect across this portion of the boundary only, double the domain, and solve the equation on the new doubled domain.

When the (portion of) the boundary is curved, then the same principle applies, but the reflection is a distorted, funhouse mirror reflection.

An exact formula for the reflection is known when the domain is a disk/sphere/etc. or a half-space (or a half-disk etc.).

Physicists call this the method of mirror charges.

3.6. Error estimates. Knowing the rate of convergence of Fourier series allows us to estimate the error when truncating them, making them more useful for concrete computations.

**Example:** Let $u : [0, 1] \times [0, \infty) \rightarrow \mathbb{R}$ be a solution of the heat equation

$$u_t = u_{xx}, \quad u(x, 0) = 1, \quad u(-1, t) = u(1, t) = 0.$$

The eigenfunctions are $\sin(n \pi x)$, $n > 0$. In this basis, the Fourier coefficients of the initial data are

$$b_n = \int_{-1}^{1} \sin(n \pi x) \, dx = -\frac{2}{\pi} \cos(n \pi) \bigg|_{x=-1}^{1} = \frac{2}{\pi} (1 - (-1)^n).$$

Then $u$ is given by the Fourier series

$$u(x, t) = \sum_{n>0 \atop n \text{ odd}} e^{-\pi^2 n^2 t} \frac{4}{\pi n} \sin(n \pi x).$$

Suppose that we want to know how many terms of this infinite series are sufficient to approximate $u$ with an error of at most $10^{-9}$, for all $t > 0.1$.

We aren’t concerned with the precise number of terms, but just wish to get a rough estimate of it.

Suppose we retain the $n \leq N$ terms in the series. Then the error is given by

$$e(x, t) = \sum_{n>N \atop n \text{ odd}} e^{-\pi^2 n^2 t} \frac{4}{\pi n} \sin(n \pi x).$$

We roughly estimate it as follows:

$$|e(x, t)| = \left| \sum_{n>N \atop n \text{ odd}} e^{-\pi^2 n^2 t} \frac{4}{\pi n} \sin(n \pi x) \right| \leq \sum_{n>N \atop n \text{ odd}} |e^{-\pi^2 n^2 t} \frac{4}{\pi n} \sin(n \pi x)| \leq \frac{4}{\pi (N + 1)} \sum_{n>N \atop n \text{ odd}} e^{-\pi^2 n^2 t}.$$

Since we don’t need an exact formula for this sum, we could estimate it as follows (this is not the only way):

$$\sum_{n>N \atop n \text{ odd}} e^{-\pi^2 n^2 t} \leq \sum_{n>N \atop n \text{ odd}} e^{-\pi^2 n(N+1)^2 t} \leq \frac{e^{-\pi^2 (N+1)^2 t}}{1 - e^{-\pi^2 (N+1)^2 t}}.$$
If \( N \geq 1 \) and \( t \geq 0.1 \), then the denominator is at most 0.5, so we get an overall bound of
\[
\frac{8}{\pi(N+1)}e^{-\pi^2(N+1)^2t} < 10^{-9}.
\]
Roughly, this means that we require \( \pi^2(N+1)^2(0.1) > 9\ln 10 \) or \((N+1)^2 > 10\ln 10 \sim 23 \). So \( N = 5 \) is good enough.

Thus, the sum converges really fast, the tail is very small, and retaining the first \( N = 5 \) terms suffices to approximate the solution with an error of at most \( 10^{-9} \) for all \( t > 0.1 \).

The error gets worse as \( t \) approaches 0, because the series does not converge uniformly to the initial data.

**Problem 28.**

### 4. Eigenfunction expansions on other domains

**4.1. The Laplacian on a disk.** Consider the eigenvalue problem for the Laplacian on a disk: find \( f : D(0,R) \to \mathbb{C} \)
\[
-\Delta f = \lambda f, \quad f|_{\partial D} = 0.
\]
In polar coordinates
\[
-(\partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2)f = \lambda f.
\]
We solve this by separation of variables: assume \( f(r,\theta) = g(r)h(\theta) \). Then the eigenfunction equation becomes
\[
-(g''h + \frac{1}{r}g'h + \frac{1}{r^2}gh'') = \lambda gh.
\]
Dividing by \( gh \) we get
\[
\frac{g''+r^{-1}g'}{g} - \frac{1}{r^2}\frac{h''}{h} = \lambda
\]
or equivalently
\[
\frac{h''}{h} = r^2\left(\lambda + \frac{g''+r^{-1}g'}{g}\right) = \omega.
\]
Here the left-hand side is a function only of \( \theta \) and the right-hand side is a function only of \( r \), so they must both be constant.

Solving the equation for \( h \) first, we get the following eigenvalue problem with periodic boundary conditions:
\[
h'' = -\omega h, \quad h(0) = h(2\pi), \quad h'(0) = h'(2\pi).
\]
The solutions are \( h = e^{in\theta}, \omega = n^2, n \in \mathbb{Z} \), and if we prefer real solutions then the pair \( \{e^{in\theta}, e^{-i\theta}\} \) can be replaced with the pair \( \{\cos(n\theta), \sin(n\theta)\}, n > 0 \) (and 1 is present in both bases).

Plugging this back into the right-hand-side equation, we get that
\[
n^2 = r^2\left(\lambda + \frac{g''+r^{-1}g'}{g}\right)
\]
or equivalently
\[
g'' + \frac{g'}{r} + \left(\lambda - \frac{n^2}{r^2}\right)g = 0. \tag{26}
\]
This equation is very similar to the Bessel equation
\[
\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(1 - \frac{\alpha^2}{x^2}\right)y = 0
\]
for \( \alpha = n \), except that \( \lambda \) needs to be replaced by 1. We do this by a change of variable,
\[
\tilde{g}(r) = g(r/\sqrt{\lambda}), \quad g(r) = \tilde{g}(r\sqrt{\lambda}).
\]
Then the equation becomes

$$\lambda \ddot{g}(r\sqrt{\lambda}) + \frac{\sqrt{\lambda} \dot{g}(r\sqrt{\lambda})}{r} + \left(\lambda - \frac{n^2}{r^2}\right) \ddot{g}(r\sqrt{\lambda}) = 0$$

and after dividing by $\lambda$

$$\ddot{g}(r\sqrt{\lambda}) + \frac{\dot{g}(r\sqrt{\lambda})}{r\sqrt{\lambda}} + \left(1 - \frac{n^2}{(r\sqrt{\lambda})^2}\right) \ddot{g}(r\sqrt{\lambda}) = 0.$$ 

So $\ddot{g}$ fulfills Bessel’s equation for $\alpha = n$, hence must be given by

$$\ddot{g}(r\sqrt{\lambda}) = aJ_n(r) + bY_n(r).$$

How do we know it’s only $J$ and no $Y$? Because $J$ is continuous at zero, while $Y$ diverges logarithmically as $r \to 0$. Thus, if we are looking for smooth eigenfunctions, then necessarily $b = 0$.

So a solution to this ODE (26) is $g = J_n(r/\sqrt{\lambda})$.

Finally, we impose the Dirichlet boundary conditions. Since $h \neq 0$, it follows that $g(R) = J_n(R/\sqrt{\lambda}) = 0$. So

$$R/\sqrt{\lambda} = z_{m,n},$$

where $z_{m,n}$ is the $m$-th zero (some zero) of the Bessel function $J_n$.

The zeros $z_{m,n}$ of Bessel functions are numbers of interest and have been known and computed for a long time, just like $\pi$ or $e$, but are not given by any especially nice formula (in particular, they do not occur at regular intervals, like the zeros of trigonometric functions), unlike $\pi$ and $e$.

Since $R$ is fixed, we solve for $\lambda$ and get $\lambda = R^2/z_{m,n}^2$. Thus, the eigenfunctions are of the form

$$f_{m,n}(r, \theta) = J_n(rz_{m,n}/R)e^{in\theta},$$

with corresponding eigenvalues $\lambda_{m,n} = R^2/z_{m,n}^2$.

Again, these Laplacian eigenfunctions are mutually orthogonal in $L^2(D)$ with respect to the dot product

$$\langle f, g \rangle = \int_D f\overline{g} \, dA = \int_0^{2\pi} \int_0^R f(r, \theta)\overline{g(r, \theta)} r \, dr \, d\theta.$$ 

This can be checked check by an explicit computation, involving the properties of Bessel functions, and also follows from the general theory.

Note: we are not proving here that this list contains all the eigenfunctions of the Laplacian and that they form a basis for $L^2(D)$, i.e. that the list above is complete. However, both facts are true. The proof is easier to do in a more abstract setting, see below.

A decomposition in terms of these eigenfunctions is called a Fourier–Bessel series:

$$f = \sum c_{m,n} f_{m,n}.$$ 

The coefficients are again given by (25), where the dot product is (27).

The fact that zeros of Bessel functions are not distributed regularly and are not multiples of some fundamental frequency is probably the main reason why drums do not sound as harmonious as string instruments.

In order to decompose the initial data of the equation in terms of these eigenfunctions,

**PROBLEM 29.** Let $y$ be a solution of the ODE

$$\frac{d^2 y}{dx^2} + \frac{1}{x} \frac{dy}{dx} + \left(1 - \frac{\alpha^2}{x^2}\right) y = 0,$$ 

such that $y$ is continuous at $x = 0$, and let $z_m$ and $z_{m'}$ be two different zeros of $y$.

$$\int_0^1 J_n(r/a_m)J_n(r/a_{m'})r \, dr = 0.$$
4.2. Harmonic potential well. Here we solve the eigenfunction equation for \( L = -\Delta + \| \vec{x} \|^2 \) in \( \mathbb{R}^d \).

We use the following principle: a commuting family of operators can be simultaneously diagonalized.

Here is the statement in the case when the operators are compact:

**Proposition 4.1.** Let \( H \) be a Hilbert space and \( A_1, \ldots, A_d : H \to H \) be bounded, compact, and commuting: \( A_j A_k = A_k A_j \) for \( 1 \leq j, k \leq d \). Then there exists a Schauder base \( (e_n)_n \) of \( H \) in which \( A_j e_n = \lambda_{nj} e_n \).

**Proof.** (by a different proof).

However, the general principle also holds when the operators aren’t bounded or compact (with a different proof).

We use this in the following context: let \( L = -\Delta + \| \vec{x} \|^2 \) and \( L_j = -\partial_x^2 + x_j^2 \), \( 1 \leq j \leq d \). Since the operators \( L_j \) commute with each other, there exists some basis in which they are simultaneously diagonalized: \( L_j e_n = \lambda_{nj} e_n \).

They also commute with their sum, which is \( L \), and
\[
L e_n = \left( \sum_j \lambda_{nj} \right) e_n.
\]

So the eigenvalues of \( L \) on \( \mathbb{R}^d \) are sums of any \( d \) eigenvalues of the one-dimensional operator \(-\partial_x^2 + x^2\).

Next, let’s analyze this one-dimensional operator.

**Proposition 4.2.** The one-dimensional operator \( L = -\partial_x^2 + x^2 \) has eigenvalues \( 2n + 1 \), \( n \in \mathbb{Z} \), \( n \geq 0 \). For each eigenvalue \( 2n + 1 \), the corresponding eigenspace has dimension 1 and is spanned by an Hermite function \( H_n = e^{-x^2/2} h_n(x) \), where \( h_n(x) \) is a polynomial of degree \( n \), called the Hermite polynomial.

**Proof.** The lowest energy eigenfunction, also called the ground state, is the Gaussian \( H_0 = e^{-x^2/2} \), for \( h_0(x) = 1 \).

The proof is based on the factorization
\[
-\partial_x^2 + x^2 = (x - \partial_x)(x + \partial_x) + 1 = (x + \partial_x)(x - \partial_x) - 1.
\]
Let \( L_+ = x + \partial_x \), \( L_- = x - \partial_x \). So
\[
L = L_+ L_- + I = L_- L_+ - I.
\]

Note that \( L_- e^{-x^2/2} = 0 \), so \( L e^{-x^2/2} = e^{-x^2/2} \).

We also compute their adjoints:
\[
\langle L_- f, g \rangle = \int_{-\infty}^{\infty} (x + \partial_x) f(x) \overline{g(x)} \, dx = \int_{-\infty}^{\infty} f(x)(x - \partial_x) g(x) = \langle f, L_+ g \rangle \, dx,
\]
where we used integration by parts. So \( L_-^* = L_+ \) and similarly \( L_+^* = L_- \).

Suppose \( f \in L^2 \) is an eigenfunction of \( L \) for the eigenvalue \( \lambda \): \( L f = \lambda f \). Then
\[
L_- \lambda f = L_- L f = L_- (L_+ L_- - I) f = L_- L_+ L_- - L_- f = (L_- L_+ - I) L_- f = (L_+ - 2I) L_- f = LL_- f - 2L_- f.
\]
So
\[
LL_- f = (\lambda - 2) L_- f.
\]

Similarly, \( LL_+ f = (\lambda + 2) L_+ f \).

This is why \( L_+ \) is called the raising/creation operator and \( L_- \) is called the lowering/annihilation operator.
4. EIGENFUNCTION EXPANSIONS ON OTHER DOMAINS 57

Thus, starting from the so-called ground state $H_0 = e^{-x^2/2}$, which is an eigenfunction for the eigenvalue $\lambda = 1$, if we apply $L_+ n$ times we get the following infinite sequence of eigenvalues and eigenfunctions.

$$L(L_+^n e^{-x^2/2}) = (2n + 1)L_+^{n} e^{-x^2/2}.$$  
These are precisely the Hermite functions, $H_n = L_+^n e^{-x^2/2}$ (though sometimes they are defined to be $L^2$-normalized).

So if $Lf = \lambda f$ then $L L_-^n f = (\lambda - 2n) L_n^2 f$. By elliptic regularity (see below) and using Agmon-type decay estimates, one can show that $L_n^2 f \in L^2$ for any $n \geq 0$. Thus, if $\lambda$ is an eigenvalue of $L$, then $\lambda - 2n$ must also be one, unless $L_n^2 f = 0$.

However, $L$ cannot have arbitrarily small eigenvalues; any eigenvalues of $L$ must be at least $1$. Indeed, if $Lf = \lambda f$, then

$$\lambda \|f\|^2 = \langle \lambda f, f \rangle = \langle Lf, f \rangle = \langle (L_+ L_- + I) f, f \rangle = \langle f, f \rangle + \langle L_- f, L_- f \rangle \geq \|f\|^2.$$  
So if $f \neq 0$ then $\lambda \geq 1$.

Thus, if $f$ is an eigenfunction of $L$ then $L_n^2 f$ are also eigenfunctions of $L$ for all $n \geq 0$, with eigenvalues $\lambda - 2n$, and $L_n^2 f = 0$ for any $n > \frac{\lambda - 1}{2}$.

Let $n$ be the smallest positive integer for which $L_n^2 f = 0$. In particular, $L_n^{-1} f$ must be a nonzero solution of the ODE $L_- [L_n^{-1} f] = 0$.

Solving this ODE, we get that

$$L_n^{-1} f = Ce^{-x^2/2}.$$  
In particular, then $\lambda - 2(n - 1) = 1$, so $\lambda = 2n - 1$. So there are no eigenvalues other than positive odd integers.

Applying $L_+$ to the equality,

$$L_+ L_- L_n^{-2} f = CL_+ e^{-x^2/2}.$$  
For any eigenfunction $g$ of $L$, if $Lg = \mu g$, then $L_+ L_- g = (L - 1) g = (\mu - 1) g$.

Then, since $L_n^{-2} f$ is an eigenfunction of $L$ for the eigenvalue $\mu = \lambda - 2(n - 2) = 3$, we get

$$L_n^{-2} f = \frac{1}{3} CL_+ e^{-x^2/2}.$$  
Iterating this, in the end we get

$$f = \frac{1}{(2n - 1)!!} CL_n^{-1} e^{-x^2/2}.$$  
This is exactly the sequence of Hermite functions, up to the constant $C$. So there are no other eigenfunctions. \hfill \square

PROBLEM 30. Prove by induction that $L_+^n e^{-x^2/2} = e^{-x^2/2} h_n(x)$, where $h_n$ is a monic polynomial of degree $n$.

This problem, called the quantum harmonic oscillator, has a lot of symmetry. In particular, if $f$ is an eigenfunction for $L$,

$$Lf = \left(- \frac{d^2}{dx^2} + x^2\right) f = \lambda f,$$

then so is $\hat{f}$: as $\frac{d}{dx} = i \xi$, $\hat{x} = i \frac{d}{d\xi}$,

$$\hat{L}\hat{f} = \left(\xi^2 - \frac{d^2}{d\xi^2}\right) \hat{f} = \lambda \hat{f}.$$  
As we saw, in one dimension each eigenspace of $L$ has dimension 1. Therefore, $\hat{f}$ must be a multiple of $f$:

$$\hat{f} = kf.$$
Repeating the Fourier transform four times brings back the original function, meaning
\[ \mathcal{F}^2 f(x) = \frac{1}{2\pi} f(-x), \quad \mathcal{F}^4 f(x) = \frac{1}{(2\pi)^2} f(x). \]
It follows that \( k^4 = \frac{1}{(2\pi)^2} \), so \( k = \pm 1, \pm i \times \frac{1}{\sqrt{2\pi}} \).

So Hermite functions form a basis in which the Fourier transform is diagonalized. The eigenvalues repeat with period four, i.e.
\[ \mathcal{F}H_0 = \frac{1}{\sqrt{2\pi}} H_0, \quad \mathcal{F}H_1 = -\frac{i}{\sqrt{2\pi}} H_1, \quad \mathcal{F}H_2 = -\frac{1}{\sqrt{2\pi}} H_2, \quad \mathcal{F}H_3 = \frac{i}{\sqrt{2\pi}} H_3 \]
and henceforth it keeps repeating.

**Problem 31.** Prove that if \( \mathcal{F}f = \lambda f \), then \( \mathcal{F}L_+ f = -i\lambda L_+ f \).

### 4.3. Rectangular boxes.
Let’s solve the eigenvalue equation for \(-\Delta\) on a rectangular box \( D = [0, a_1] \times \ldots \times [0, a_d] \subset \mathbb{R}^d \), with some specified homogenous boundary condition on each face.

The operators \(-\Delta_{x_j}\) on \([0, a_j]\), with the specified boundary conditions, commute with each other and with their sum \(-\Delta\).

The importance of having a rectangular box lies in the fact that the intervals \([0, a_j]\) are constantly the same throughout the domain, which wouldn’t be true for any other shape (or even a non-rectangular box).

The eigenfunctions are then of the form
\[ f(x) = f_1(x_1) \ldots f_d(x_d), \]
where each \( f_j \) is an eigenfunction of \(-\Delta_{x_j}\) on \([0, a_j]\), with the corresponding eigenvalue
\[ \lambda = \lambda_1 + \ldots + \lambda_d. \]
Suppose all the boundary conditions are Dirichlet boundary conditions. Then \( \lambda_j = (2\pi n_j/a_j)^2 \), for some \( n_j > 0 \). So
\[ \lambda = (2\pi n_1/a_1)^2 + \ldots + (2\pi n_d/a_d)^2. \]

Finding the distribution and multiplicity of eigenvalues for such a box is an interesting combinatorics and/or number theory problem. If the side lengths \( a_j \) are incommensurate, then all the eigenvalues are simple (no multiplicity). At the other extreme, if all the sides have the same size, then one arrives at the Diophantine equation
\[ n_1^2 + \ldots + n_d^2 = \frac{a^2}{4\pi^2} \lambda. \]
This is an extremely well-studied (but still not completely understood) Diophantine equation.

### 4.4. The Bunimovich stadium.
Going even slightly beyond these very well understood shapes — spheres, disks, rectangular boxes, harmonic potentials, and several others — it becomes very difficult to find exact formulas for Laplacian eigenfunctions and eigenvalues.

Usually, in such cases they are computed numerically, though the general theory lets us state some results without explicit computations.

The Bunimovich stadium \( B \subset \mathbb{R}^2 \) is a rectangle with two half-disks fitted to two opposite sides:
\[ B = ([-a, a] \times [-b, b]) \cup B((a, 0), b) \cup B((-a, 0), b). \]
Although rectangles and half-disks are well understood, the eigenfunction equation for the Laplacian on the Bunimovich stadium is poorly understood in most important respects. And no exact formulas for the Laplacian eigenfunctions and eigenvalues are known in this case.

### 5. General facts about eigenvalues and eigenfunctions

#### 5.1. Dot product spaces and Hilbert spaces.
5. GENERAL FACTS ABOUT EIGENVALUES AND EIGENFUNCTIONS

5.1.1. Dot product spaces.

Definition 17. A dot product space is a (real or complex) vector space $V$ endowed with a sesquilinear operation $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ such that:

1. It’s sesquilinear: linear in the first factor

$$\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle, \quad \langle \alpha f, g \rangle = \alpha \langle f, g \rangle$$

and conjugate linear for the other factor:

$$\langle f, g + h \rangle = \langle f, g \rangle + \langle f, h \rangle, \quad \langle f, \alpha g \rangle = \overline{\alpha} \langle f, g \rangle.$$

2. $\langle f, g \rangle = \overline{\langle g, f \rangle}$.

By setting $f = g$, this implies that $\langle f, f \rangle = \overline{\langle f, f \rangle}$, so $\langle f, f \rangle \in \mathbb{R}$. But in fact more must be true:

3. It’s positive definite: $\langle f, f \rangle \geq 0$ and $\langle f, f \rangle = 0$ if and only if $f = 0$.

Examples: $\mathbb{R}^n$, $\mathbb{C}^n$ with the dot product

$$\langle (a_1, \ldots, a_n), (b_1, \ldots, b_n) \rangle = a_1 \overline{b_1} + \ldots + a_n \overline{b_n}.$$

These are finite-dimensional spaces. Then there is

$$\ell^2 = \{ (x_n) : \sum_n |x_n|^2 < \infty \}$$

with the dot product

$$\langle (x_n), (y_n) \rangle = \sum_n x_n \overline{y_n}.$$

Finally, on the set of continuous (or $C^1$, or smooth, ...) and bounded functions on an interval $I$ we can define a dot product

$$\langle f, g \rangle = \int_I f(x) \overline{g(x)} \, dx.$$

This turns e.g. $C^1(I)$ or $C(I)$ into a dot product space.

On every dot product space we can define a norm by

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

For example, on $C^1(I)$ the norm induced by the dot product will be

$$\|f\| = \sqrt{\int_I |f(x)|^2 \, dx}.$$ 

Hence we also get a notion of convergence: $f_n \to f$ if $\lim_{n \to \infty} \|f_n - f\| = 0$.

A norm must always have the following three properties:

1. It’s homogenous: $\|\alpha f\| = |\alpha| \|f\|$.

2. It satisfies the triangle inequality: $\|f + g\| \leq \|f\| + \|g\|$.

3. It’s positive definite: $\|f\| \geq 0$ and $\|f\| = 0$ if and only if $f = 0$.

A vector space endowed with a norm is called a normed vector space. More than one norm can be defined on a vector space. They may or may not be equivalent (i.e. a sequence converges in one if and only if it convergences in the other).

All $L^p$ norms, $1 \leq p \leq \infty$, satisfy these properties. However, for the third property, $\|f\|_{L^p} = 0$ if and only if $f$ is zero almost everywhere, i.e. $f$ can still be nonzero on a sufficiently small set of points, e.g. at finitely many points, and still have norm zero.

To still have $L^p$ be normed spaces, the convention is to consider elements of $L^p$ to be not functions, but equivalence classes of functions equal to each other almost everywhere. Thus, from this point of view, a function which is zero almost everywhere is (in the equivalence class of) zero. We won’t insist here on this technical distinction between zero and zero almost everywhere.
In addition, in a dot product space, the norm and the dot product are connected by the Cauchy–Schwartz inequality:

\[ |\langle f, g \rangle| \leq \|f\| \|g\|. \]

5.1.2. Hilbert spaces.

**Definition 18.** A Hilbert space is a complete dot product space. A space is called complete if every Cauchy sequence converges. A sequence \( f_n \) is called Cauchy if for every \( \epsilon > 0 \) there exists \( N \) such that, whenever \( m, n \geq N \), then \( \|f_m - f_n\| < \epsilon \).

So in a complete (metric) space Cauchy's convergence test is valid; in an incomplete space there exists at least a Cauchy sequence that does not converge.

All the dot product spaces mentioned above are complete, hence Hilbert spaces, with the exception of \( C^k(I) \). However, every dot product space can be completed into a Hilbert space, by assigning a limit to Cauchy sequences that don’t have one (and such the limit is the same for sequences that “converge together”). The completion of \( C^k(I) \) is \( L^2(I) \), the space of square integrable functions.

**Example:** For a smooth cutoff function \( \chi \), \( f_n(x) = (1 - \chi(nx))x^{-1/4} \) is a Cauchy sequence which does not converge in \( C^1([0, 1]) \) with the norm; this is because the limit is \( x^{-1/4} \), which is not in \( C^1 \).

**Definition 19.** A linear operator \( L : V \to V \) is called bounded if and only if there exists some constant \( C \) for which \( \|Lv\| \leq C \|v\| \) for all \( v \in V \).

**Definition 20.** An operator \( L : U \to V \) is called continuous if and only if \( u_n \to u \) implies that \( L(u_n) \to L(u) \).

**Problem 32.** For linear operators on normed spaces, bounded is equivalent to continuous.

**Example of a bounded operator:** Consider \( H = L^2([a, b]) \) and \( L : H \to H \),

\[ [Lf](x) = \int_0^x f(t) \, dt. \]

Claim: This is a bounded operator.

**Proof.** By Cauchy–Schwartz,

\[ |[Lf](x)| \leq \left( \int_0^x dt \right)^{1/2} \left( \int_0^x |f(t)|^2 \, dt \right)^{1/2} \leq x^{1/2} \|f\|_{L^2}. \]

Then

\[ \|Lf\|_{L^2} \leq \|f\|_{L^2} x^{1/2} \|L^2\| = \frac{b - a}{\sqrt{2}} \|f\|_{L^2}. \]

This is the best possible constant, because one has equality for \( f(t) = 1 \). \( \square \)

**Problem 33.** Prove that \( Lf = xf \) is bounded on \( L^2([a, b]) \).

**Example of an unbounded operator:** Over the same space \( H = L^2([a, b]) \), \( L = \partial_x : H \to H \) is an unbounded operator.

Unbounded operators are not defined over the whole space. In the case, the domain of \( L \) is the set of \( L^2 \) functions whose derivative is also in \( L^2 \).

**Problem 34.** Prove that the operator \( L \) is unbounded.

**Importance of completeness:** Why does completeness matter? Because in many proofs we can construct the object we want, such as the solution to some equation, only indirectly, as the presumptive limit of some Cauchy sequence.

**Example:** The Fourier transform of an \( L^2 \) function is an \( L^2 \) function.
Proof. We truncate and approximate the original function by Schwartz functions, whose Fourier transforms are also Schwartz-class functions. However, these form a Cauchy sequence in $L^2$. Since $L^2$ is complete, this Cauchy sequence has an $L^2$ limit.

Example: For a square summable sequence $(c_n)_n \in \ell^2$, there exists a unique $L^2([0,p])$ function with these coefficients:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/p}.$$ 

Proof. The sequence of partial sums is a Cauchy sequence in $L^2$.

5.1.3. Adjoint of an operator.

Definition 21. Let $H_1$ and $H_2$ be two Hilbert spaces and $T : H_1 \to H_2$ be a bounded linear operator. Its adjoint $T^*$ is defined to be the unique operator $T^* : H_2 \to H_1$ such that

$$\langle Tu, v \rangle_{H_2} = \langle u, T^* v \rangle_{H_1}$$

for any $u \in H_1, v \in H_2$.

The existence of an adjoint is guaranteed by the Riesz representation theorem, which is valid only on Hilbert spaces (as opposed to incomplete dot product spaces).

Example: If $H_1$ and $H_2$ are finite dimensional, in any given basis $T$ is represented by a matrix. Then $T^* = T^t$ is the transpose of $T$ in the real case and is the conjugate of the transpose for complex Hilbert spaces: $T^* = \overline{T^t}$.

Example: Consider $H = L^2([a,b])$ and $L : H \to H$,

$$(Lf)(x) = \int_a^x f(t) \, dt.$$ 

Its adjoint is

$$(L^* g)(x) = \int_x^b g(t) \, dt.$$ 

Definition 22. A bounded operator $T : H \to H$ equal to its own adjoint, $T = T^*$, is called self-adjoint.

The situation is more complicated for unbounded operators. An unbounded operator that formally satisfies the definition of self-adjointness is called symmetric:

Definition 23. Let $H$ be a Hilbert space and $A : V \to H$ be an unbounded operator defined on some dense vector subspace $V \subset H$. If for all $f, g \in V$

$$\langle Af, g \rangle = \langle f, Ag \rangle$$

then $A$ is called symmetric.

An unbounded symmetric operator is called self-adjoint if it satisfies the additional condition that the operator and its adjoint should have the same domain.

Example: Differential operators: Consider a differential operator on $L^2(\mathbb{R}^d)$ given by

$$Tu = \sum_{|\alpha| \leq n} c_\alpha(x) D^\alpha u.$$ 

Its adjoint is

$$T^* u = \sum_{|\alpha| \leq n} (-1)^{|\alpha|} D^\alpha (c_\alpha(x) u).$$

Thus, if the coefficients are constants, then $T$ is symmetric if and only if $c_\alpha \in \mathbb{R}$ whenever $|\alpha|$ is even and $c_\alpha$ are purely imaginary whenever $|\alpha|$ is odd.
Proposition 5.1. Eigenfunctions of symmetric operators, corresponding to distinct eigenvalues, are orthogonal.

Proof. First, eigenvalues of symmetric operators must be real numbers. Indeed, suppose $Tv = \lambda v$. Then
\[ \langle v, v \rangle = \langle Tv, v \rangle = \overline{\langle v, Tv \rangle} = \overline{\lambda \langle v, v \rangle}. \]
If $v \neq 0$, then $\langle v, v \rangle = \|v\|^2 \neq 0$, so $\lambda = \overline{\lambda}$, so $\lambda \in \mathbb{R}$.

Suppose $Tv_1 = \lambda_1 v_1$ and $Tv_2 = \lambda_2 v_2$. Then
\[ \lambda_1 \langle v_1, v_2 \rangle = \langle Tv_1, v_2 \rangle = \langle v_1, Tv_2 \rangle = \overline{\lambda_2} \langle v_1, v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle. \]
If $\lambda_1 \neq \lambda_2$, then it must be that $\langle v_1, v_2 \rangle = 0$. \qed

The real Fourier coefficients are Fourier bases: $e^{2\pi inx}$, $\sin(\pi nx)$, $\cos(\pi nx)$.

It is always possible to go from one such basis to another. Each is useful for solving PDEs for a different boundary condition.

Problem 35. Suppose $f : \mathbb{R} \to \mathbb{R}$ is given by
\[ f(x) = \sum_{n \in \mathbb{Z}} c_n e^{2\pi inx} \]
as well as
\[ f(x) = \sum_{n > 0} b_n \sin(\pi nx). \]
Compute the $b_n$’s as a function of the $c_n$’s.

5.2. Completeness of the Fourier basis. What do we want for a basis? It should be independent and should span the whole space of functions.

Example: The usual Fourier basis $\{e^{2\pi in/p}\}$ is a basis for $L^2([0,p])$.

Can we omit any of these functions? No. Suppose we omit any of these functions. Then it cannot be represented by the remaining ones.

Orthogonality: Helps with independence and makes computing coefficients simple.

Are these functions enough to span the whole space? Then called complete orthonormal sequence / Fourier basis. If not linearly independent, called overcomplete.

Can we choose an altogether different set of frequencies? Yes. There is nothing preventing us from extending $f$ in some arbitrary manner to some wider interval and computing the Fourier series there; or to the whole real line and taking the Fourier transform there.

The appropriate choice of Fourier expansion is determined by the boundary conditions.

5.3. Eigenfunction expansions in general settings. Proposition 5.1 proved that the eigenvectors are orthogonal. However, how do we know in general that they span the whole space? How do we know that an appropriate Fourier basis exists?

In more general settings, the completeness of the Fourier / Schauder basis made of eigenfunctions is usually proved by the Hilbert–Schmidt theorem.

A set is called compact if every sequence in the set has a convergent subsequence whose limit is also in the set. If the limit is not necessarily in the set, then the set is called precompact.

Compact operators: An operator $A : U \to V$ between two normed vector spaces $U$ and $V$ is compact if, for every bounded sequence $v_n$, the sequence $Av_n$ has a converging subsequence.

Theorem 5.2 (Hilbert–Schmidt). Let $H$ be a Hilbert space and $A : H \to H$ be a compact, self-adjoint operator. Then there exists an orthonormal basis $e_n$ such that $A$ is diagonalized in this basis: $A e_n = \lambda_n e_n$. Furthermore, $\lim_{n \to \infty} \lambda_n = 0$. 

One can also order the eigenvalues by decreasing absolute value.

This theorem will guarantee that the eigenfunctions we found by solving the eigenvalue problem (usually for the Laplacian) span the whole space, so we can solve the associated PDE by separation of variables and one can always find an eigenfunction expansion of the initial data.

We often cannot apply the Hilbert–Schmidt theorem directly because the operator \( L \) we want to study is unbounded, hence certainly not compact. Thus, we apply it to some other directly related operator, e.g. \( L^{-1}, (L + I)^{-1} - I \), etc. — provided these operators are compact.

This method works on bounded domains or with confining potentials.

**Proposition 5.3.** Let \( V \) be a potential on \( \mathbb{R}^d \) such that \( V \) is measurable and bounded from below and \( \lim_{x \to \infty} V(x) = +\infty \).

Then there exists an orthonormal basis of \( L^2(\mathbb{R}^d) \) that diagonalizes \(-\Delta + V\).

In order to use Hilbert–Schmidt theorem, we first need to prove that some operator, such as \( L^{-1} \), is compact. At their most basic, many proofs of compactness in spaces of functions use the Arzelà–Ascoli or the Kolmogorov–Riesz criterion.

**Theorem 5.4 (Arzelà-Ascoli).** Let \( D \subset \mathbb{R}^d \) be an open set and \( (f_n)_n, f_n : D \to \mathbb{C} \) be a sequence of continuous functions on \( D \). Suppose that

1. \( f_n \) are uniformly bounded: there exists some \( M \) such that, for every \( n \) and \( x \in D \), \(|f_n(x)| \leq M\).
2. \( f_n \) are equicontinuous: for each \( \epsilon > 0 \) there exists \( \delta > 0 \) such that whenever \(|x - y| < \delta, x, y \in D\), then \(|f_n(x) - f_n(y)| < \epsilon\).

Then \( f_n \) must have a subsequence \( f_{n_k} \) that converges uniformly on \( D \) to some bounded and uniformly continuous limit function \( f : D \to \mathbb{C} \).

**Problem 36.** Give examples of two sequences of continuous functions that have one of the properties, but not the other, and have no uniformly converging subsequence.

**Problem 37.** Let \( f_n : D \to \mathbb{C} \) be a sequence of uniformly converging functions whose limit \( f \) is uniformly continuous and bounded. Prove that \( f_n \) are equicontinuous and uniformly bounded.

Give counterexamples when \( f \) is not uniformly continuous or not bounded.

**Theorem 5.5 (Kolmogorov–Riesz).** Consider the sequence \((f_n)_n, f_n \in L^p(\mathbb{R}^d), 1 \leq p < \infty\), and assume that

1. Uniform decay at infinity: for every \( \epsilon > 0 \) there exists \( R \) such that for all \( n \)
   \[
   \|\chi_{|x| \geq R} f_n(x)\|_{L^p} < \epsilon.
   \]
2. Equicontinuity under translations: for every \( \epsilon > 0 \) there exists \( \delta > 0 \) such that for all \( y \in \mathbb{R}^d \) with \(|y| < \delta \) and for all \( n \)
   \[
   \|f_n(x - y) - f_n(x)\|_{L^p} < \epsilon.
   \]

Then \( f_n \) must have a subsequence \( f_{n_k} \) that converges in the \( L^p \) norm to some \( L^p \) function \( f \).

**Problem 38.** a) Under the above conditions, prove that the \( L^p \) norm does not concentrate on small sets: for every \( \epsilon > 0 \) there exists \( \delta > 0 \) such that for every \( n \) and \( x_0 \in \mathbb{R}^d \)
   \[
   \|\chi_{B(x_0, \delta)}(x)f(x)\|_{L^p} < \epsilon.
   \]

b) Under the same conditions, prove that the \( L^p \) norms of \( f_n \) are uniformly bounded.

**Example:** On \([0, a]\) with periodic boundary conditions: we construct a compact, self-adjoint operator to which we apply the Hilbert–Schmidt theorem.

Consider the operator \( L = -\partial_x^2 \) on \([0, a]\) with Dirichlet boundary conditions. Its inverse is

\[
L^{-1} g(y) = \frac{y}{a} \int_0^a (a - z) g(z) \, dz - \int_0^y (y - z) g(z) \, dz.
\]

**Proposition 5.6.** \( L^{-1} \) is compact on \( L^2([0, a]) \).
Proof. Let’s reuse the proof of Arzelà–Ascoli (alternatively, we could just invoke one of the two previous compactness criteria). First, we want to check that

Let \( g_n \in L^2 \) be uniformly bounded

Let \( h_n(y) = L^{-1}g_n \).

By diagonalization we can find a subsequence \( h_{n_k} \) which converges pointwise at all rational numbers \( q \in \mathbb{Q} \cap [0,a] \).

Also, the \( h_{n_k} \) are uniformly Hölder continuous.

We now prove that the convergence is uniform. □

Problem 39. a) Find the inverse of \( L = -c_x^2 \) on the \( C^1([0,a]) \), with a Neumann boundary condition at 0 and a Dirichlet boundary condition at \( a \).

b) Find the inverse \( L^{-1} \) of \( L = -c_x^2 \) from the set of periodic \( C^1 \) functions of period \( p \) whose average is zero to itself:

\[
\int_0^p f(x) \, dx = 0.
\]

Explain why the inverse is not well-defined for functions with nonzero average.

Problem 40. a) Find the appropriate Fourier expansion of the function \( f : [0, \pi] \to \mathbb{R}, f(x) = \cos x \), for each of the following boundary conditions:

i. periodic boundary conditions

ii. Dirichlet boundary conditions

iii. Neumann boundary conditions

iv. Robin boundary conditions \( u' + u = 0 \).

b) Write the solution of the heat equation with this function as the initial data for each of the above boundary conditions.

Problem 41. Solve the wave equation with initial data \( u_0 : [0, \pi] \to \mathbb{R}, u_0(x) = \sin x, u_1(x) = 0 \), with

a) Dirichlet boundary conditions

b) Neumann boundary conditions.

Problem 42. In problems 28i) and 29b), estimate the number of terms in the Fourier expansion needed to approximate the solution with \( 10^{-12} \) accuracy for all \( x \) and all times \( t > 0.5 \).

Your answer need not be optimal, but should be correct.

5.3.1. Green–Stokes theorem.

Theorem 5.7. For a domain \( D \subset \mathbb{R}^d \) and \( f \) and \( g : D \to \mathbb{C} \) of class \( C^2 \),

\[
\int_D (\Delta f) g + \nabla f \cdot \nabla g = \int_{\partial D} \frac{df}{dn} g
\]

and

\[
\int_D (\Delta f) g - f(\Delta g) = \int_{\partial D} \frac{df}{dn} g - f \frac{dg}{dn}.
\]

Proof. Gauss’s theorem, also called the divergence theorem, states that

\[
\int_D \text{div} \, \mathcal{G} = \int_{\partial D} \mathcal{G} \cdot \vec{n},
\]

where \( \vec{n} \) is the outward pointing unit normal vector. We apply it to \( \mathcal{G} = g \nabla f \), where \( g \) and \( f \) are as in the theorem’s statement above. Then

\[
\int_D \text{div}(g \nabla f) = \int_{\partial D} g \nabla f \cdot \vec{n}.
\]
A vector calculus product rule states that
\[ \text{div}(g\vec{F}) = \nabla g \cdot \vec{F} + g \text{div} \vec{F}. \]
Here
\[ \text{div}(g\nabla f) = \nabla g \cdot \nabla f + g \text{div} \nabla f \]
and \( \text{div} \nabla = \Delta \). In addition, \( \nabla f \cdot \vec{n} = \frac{df}{dn} \) (it's the directional derivative in the outward direction).

The second identity is obtained by swapping \( f \) and \( g \) and subtracting the two copies of the first identity from each other. □

The Green–Stokes theorem is useful in several ways. First, it can be used to prove the self-adjointness of \( \Delta \) (at least formally, i.e. its symmetry):

**Proposition 5.8.** Suppose \( f, g \in C^\infty(D) \) fulfill the Dirichlet boundary condition \( f|_{\partial D} = g|_{\partial D} = 0 \). Then
\[ \int_D (\Delta f)g = \int_D f(\Delta g). \]

**Proof.** Obvious from the Green–Stokes theorem. □

**Problem 43.**
(a) Show that the same is true when both \( f \) and \( g \) satisfy the Neumann boundary condition.
(b) Show that the same is true when \( f \) and \( g \) satisfy the same Robin boundary condition: on \( \partial D \)
\[ f(x) + \alpha(x) \frac{df}{dn} = 0, \quad g(x) + \alpha(x) \frac{dg}{dn} = 0. \]
Here \( \alpha \in C(\partial D) \) is a continuous function of \( x \) on \( \partial D \), same for both \( f \) and \( g \).

We now prove that each eigenvalue of the Laplacian is real and nonnegative and has a corresponding real-valued eigenfunction.

**Proposition 5.9.** Let \( D \subset \mathbb{R}^d \) be a domain with sufficiently smooth boundary and consider the eigenvalue equation
\[ -\Delta f = \lambda f, \quad f|_{\partial D} = 0 \]
for the Laplacian with Dirichlet boundary conditions. If \( f \) is not constantly zero, then \( \lambda \in \mathbb{R} \) and \( \lambda \geq 0 \).

**Proof.** □

**5.4. The skew-Fourier transform.** In general, however, \( L^{-1} \) etc. may not be compact.
If \( L^{-1} \) etc. are not compact, we cannot always expand the solution into a series. In general, we need to use a series, as in the Fourier series above, plus an integral, like the one in the Fourier transform.

**Example:** Consider the potential well
\[ V(x) = \]

Example of a square potential well that reduces to Robin boundary conditions.
\[ -u'' = \lambda u \]
\[ u = e^{-ix} \]
\[ u = e^{ix} \]
\[ u' = \mp \mu u \text{ at } x = \pm a. \]
\[ u = Ae^{ix} + Be^{-ix}, \quad u' = i\nu(Ae^{ix} - Be^{-ix}). \]
\[ iv(Ae^{ixa} - Be^{-ixa}) = \mu(Ae^{ixa} + Be^{-ixa}). \]
\[ iv(Ae^{-ixa} - Be^{ixa}) = -\mu(Ae^{-ixa} + Be^{ixa}). \]
\[ iv(A + B)(e + e) = \mu(A - B)(e - e) \]
\[ iv(A - B)(e - e) = \mu(A + B)(e + e). \]

So \( e - e = 0 \) or \( e + e = 0 \).

5.5. Limitations of this method. The Fourier transform and series, eigenfunction expansions, and the technique of separation of variables are most useful for solving linear equations with constant coefficients on a few special domains with some special boundary conditions. They are also useful as auxiliary tools and as a starting point in the study of nonlinear equations. However, these techniques won’t usually provide a straightforward answer for equations with nonconstant coefficients or for nonlinear equations.

6. Other function spaces

In what sense does the Fourier series of a function converge? In the \( L^2 \) sense if the function is in \( L^2 \), in \( L^p \) if the function is in \( L^p \):

**Proposition 6.1.** Suppose \( f \in L^p(I) \), \( 1 < p < \infty \). Then

\[
\lim_{n \to \infty} \| S_n - f \|_{L^p} = 0.
\]

**Proof.** The proof is elementary for \( p = 2 \), based on Parseval’s identity, but less elementary for \( p \neq 2 \), based on the Hilbert transform. \( \square \)

However, suppose we are specifically interested in pointwise or uniform convergence. One can do better than \( C^1 \) for uniform convergence: there are wider spaces of functions in which Fourier series converge uniformly and there are much wider spaces of functions in which Fourier series converge, but non-uniformly (i.e. for other meanings of convergence, as above).


**Definition 24.** On a domain \( D \in \mathbb{R}^d \), the Hölder class \( C^\alpha(D) \) is the class of continuous and bounded functions \( f : D \to \mathbb{C} \) that also have the Hölder property

\[ |f(\vec{x}) - f(\vec{y})| \leq C\|\vec{x} - \vec{y}\|^\alpha. \]

\( C^\alpha \) is a Banach space with norm

\[
\| f \|_{C^\alpha} = \sup_{x \in D} |f(x)| + \sup_{x, y \in D, x \neq y} \frac{|f(x) - f(y)|}{\|\vec{x} - \vec{y}\|^\alpha}.
\]

**Example:** \( \sqrt{x} \) is not of class \( C^1 \), but is in \( C^{1/2} \).

**Problem 44.** Show that \( f : [0, 1], f(x) = x^\alpha, 0 < \alpha < 1 \), is in \( C^\alpha \).

For \( \alpha = 1 \), a similar space is called the space of Lipschitz continuous functions.

**Definition 25.** \( Lip(D) = C^{1,0}(D) \) is the space of continuous functions on \( D \) which also have the Lipschitz continuity property

\[ |f(\vec{x}) - f(\vec{y})| \leq C\|\vec{x} - \vec{y}\| \]

for some constant \( C \).
The norm is
\[ \|f\|_{Lip} = \sup_{x \in D} |f(x)| + \sup_{x,y \in D, x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|}. \]

Example: \(|x|\) is not in \(C^1\) but is in \(Lip = C^{0.1}\).

If we want more regularity and formally put \(\alpha > 1\) in the above definition of Hölder spaces \(C^\alpha\), then only constant functions will fit the definition.

Problem 45. Let \(f : [a, b] \to \mathbb{C}\). Prove that if there exist \(\alpha > 1\) and \(C \in \mathbb{R}\) such that for all \(x, y \in [a, b]\)
\[ |f(x) - f(y)| \leq C|x - y|^{\alpha} \]
then \(f\) is constant.

So this is not the right approach. The correct definition for higher-order Hölder spaces is

**Definition 26.**
\[ C^{n,\alpha}(D) = \{f : D \to \mathbb{C} : f \text{ and all its derivatives up to order } n \text{ are in } C^\alpha\}. \]

**Importance of Hölder spaces:**
Hölder spaces are one way of giving meaning to a non-integer number of derivatives. In some sense, functions in \(C^\alpha\) have “\(\alpha\)” derivatives and functions in \(C^{n,\alpha}\) have “\(n + \alpha\)” derivatives.

Example: \(x^{3/2}\) is in \(C^{1.1/2}\).

**Fourier series:** the Fourier series of a periodic \(C^\alpha\) function converges uniformly if \(\alpha > 1/2\):

**Proposition 6.2.** Let \(f : [0, p] \to \mathbb{C}\) be a \(C^\alpha\) function such that \(f(0) = f(p)\). Then the sequence of partial sums in the Fourier series converges uniformly to \(f\).

Hölder spaces are widely used in the study of elliptic and parabolic equations, but not too useful for the study of hyperbolic equations or that of Schrödinger’s equation.

### 6.2. Sobolev spaces.

**Definition 27.** The space of functions such that they and all their derivatives up to order \(n\) are in \(L^p\) is called the Sobolev space \(W^{n,p}\).

Here we are not referring to derivatives in the classical sense, but to weak derivatives:

**Definition 28.** We say that \(f' = g\) in the weak sense if for every \(h \in \mathcal{D}\)
\[ -\int f(x)h'(x) \, dx = \int g(x)h(x) \, dx. \]

**Problem 46.** Prove that \(|x'| = \text{sgn } x\) in the weak sense.

A different notion is

**Definition 29.** We say that \(f' = g\) in the strong sense if \(f' = g\) almost everywhere (i.e. except on a set of Lebesgue measure zero).

Example: \(|x'| = \text{sgn } x\) almost everywhere as well.
In fact, the two notions are pretty much equivalent: see absolute continuity on lines.

For \(p = 2\), we also denote \(W^{n,2}\) by \(H^n\). We can also give the following equivalent definition of Sobolev spaces:
\[ H^s = \{f | \int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 (1 + |\xi|)^{2s} \, d\xi < \infty \}. \]

Note that this definition also makes sense when the number of derivatives \(s\) is not an integer, as well as when \(s\) is negative.
3. SEPARATION OF VARIABLES

As always, $p \neq 2$ is more complicated, but one can also define fractional and negative order Sobolev spaces when $p \neq 2$.

When $s = 0$, $W^{0,p} = L^p$ and $H^0 = L^2$. For positive $s$, $W^{s,p} \subset L^p$ and $h^s \subset L^2$. For negative $s$, it’s the opposite: $L^p \subset W^{-s,p}$ and $L^2 \subset H^{-s}$.

$H^s$ are always Hilbert spaces, unlike $W^{n,p}$ for $p \neq 2$, which never are Hilbert spaces. A dot product on $H^s$ is

$$\langle f, g \rangle_{H^s} = \int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}(1 + |\xi|^2)^{s}\,d\xi.$$ 

Example: $H^{-1}$.

In general, first-order derivatives of an $L^2$ function are in $H^{-1}$.

In $\mathbb{R}^d$, $L^{2d/(d+1)} \subset H^{-1}$ and in any dimension $L^2 \subset H^{-1}$.

Finally, in one dimension $\delta_0 \in H^{-1}$, see below.

Importance of Sobolev spaces:

Frequency cutoff, generalized Leibniz rule.

Sobolev spaces can be used to solve every type of PDE.

Problem 47. Prove that if $f \in H^s$ then its translate $g(x) = f(x + a)$ is also in $H^s$.

6.3. Tempered distributions. One of the widest spaces useful for solving PDEs is that of tempered distributions, $S'$.

Definition 30. A sequence of functions $f_k : \mathbb{R} \to \mathbb{C}$ converges in the Schwartz space $\mathcal{S}$ if and only if it converges in each of its defining seminorms: for each $m, n \geq 0$

$$\lim_{k \to \infty} \sup_x |x|^m |f_k^{(n)}(x) - f^{(n)}(x)| = 0.$$ 

Definition 31. Tempered distributions are the linear and continuous functionals on $\mathcal{S}$, meaning linear functionals $g : \mathcal{S} \to \mathbb{C}$ such that whenever $f_n \to f$ in $\mathcal{S}$ then $g(f_n) \to g(f)$ in $\mathbb{C}$.

The space of tempered distributions is the dual of the Schwartz space and is denoted by $\mathcal{S}'$.

The pairing between $f \in \mathcal{S}$ and $g \in \mathcal{S}'$ is often denoted by $\overline{\pi}(f) = \langle f, g \rangle$, same as a dot product.

The space of tempered distributions is the widest space we have defined so far, as it contains all the other ones we discuss: Sobolev spaces, Hölder spaces, Lipschitz spaces, spaces of continuous and/or differentiable functions, the Schwartz space, etc.. They are all embedded in it in the following manner: each function $g$ defines a linear and continuous functional on the Schwartz space by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)\,dx.$$ 

Two such functionals are equal if and only if their defining functions are equal almost everywhere (i.e. except on a set of measure zero).

Not all functions are tempered distributions, but only those that don’t grow too fast at infinity.

The space of tempered distributions contains more than just functions. In general, its elements are called generalized functions.

Example: Dirac’s delta function. The most important one is Dirac’s $\delta$, defined by

$$\delta_a(f) = \langle f, \delta_a \rangle = \int_{-\infty}^{\infty} f(x)\delta_a(x) = f(a).$$

Dirac’s delta function, which is in fact not a function, but a generalized function, corresponds to the physical notion of point masses/charges/etc.c.

Operations with tempered distributions: If $g \in \mathcal{S}'$ is a tempered distribution, then so are:

* Its conjugate $\overline{g}$
* Its derivative $g'$
* $xg$
* Its Fourier transform $\hat{g}$.

The related tempered distributions are defined by duality, as follows.

**Definition 32.** For a tempered distribution $g \in \mathcal{S}'$, its Fourier transform is defined as the unique tempered distribution $\hat{g} \in \mathcal{S}'$ such that for every $f \in \mathcal{S}$

$$\langle f, g \rangle = 2\pi \langle \hat{f}, \hat{g} \rangle.$$ 

The proof of existence uses the following fact about the Schwartz class:

**Proposition 6.3.** $f \mapsto \overline{f}$, $f \mapsto \hat{f}$, $f \mapsto f'$, and $f \mapsto xf$ are continuous operations on the Schwartz class $\mathcal{S}(\mathbb{R})$.

**Problem 48.** Prove that $f \mapsto \hat{f}$ is a continuous operation on $\mathcal{S}$; in other words, if for a sequence of Schwartz-class functions $(f_n)$ one has that $f_n \to f$ in $\mathcal{S}$, then $\hat{f}_n \to \hat{f}$ in $\mathcal{S}$.

**Problem 49.** Prove that $f \mapsto xf$ is a continuous operation on $\mathcal{S}$.

**Problem 50.** Prove that if $g \in \mathcal{S}'$ is a tempered distribution, then so is $(\sin x)g(x)$.

Dirac’s delta function belongs to negative-order Sobolev spaces:

**Proposition 6.4.** In $\mathbb{R}^d$, $\delta_0 \in H^s$ for any $s < -d/2$.

**Proof.** The Fourier transform of the Dirac delta function $\delta_0$ is a constant: $\square$
CHAPTER 4

Second-order PDEs

1. Classification and canonical form

1.1. Definitions. We say that two PDEs are of the “same type” if they can be transformed into one another by a coordinate change.

The general form of a linear constant-coefficient second-order PDE is

\[ Lu = \sum_{k,\ell=1}^{d} a_{k\ell} \frac{\partial^2 u}{\partial x_k \partial x_\ell} + \sum_{k=1}^{d} b_k \frac{\partial u}{\partial x_k} + cu = f(x). \]

Taking the Fourier transform of both sides, we get

\[- \sum_{k,\ell=1}^{d} a_{ij} \xi_i \xi_j \hat{u} + i \sum_{k=1}^{d} b_k \xi_k \hat{u} + c \hat{u} = \hat{f}.\]

Thus, at least formally, the equation always admits the solution

\[ \hat{u}(\xi) = \frac{\hat{f}(\xi)}{- \sum_{k,\ell=1}^{d} a_{ij} \xi_i \xi_j + i \sum_{k=1}^{d} b_k \xi_k + c} = \frac{\hat{f}(\xi)}{P(i\xi)}, \]

where the polynomial in \( d \) variables

\[ P(i\xi) = - \sum_{k,\ell=1}^{d} a_{ij} \xi_i \xi_j + i \sum_{k=1}^{d} b_k \xi_k + c \]

is called the symbol of the operator \( L \).

Reverting the Fourier transform, we get

\[ u = f \ast \left( \frac{1}{2\pi} \mathcal{F}^{-1} \frac{1}{P(i\xi)} \right) = f \ast S, \]

where \( S \) is called the fundamental solution of the equation.

If we are able to make sense of them — usually by taking some limit — these distributions are weak solutions. The fundamental solution \( S \) is itself the solution corresponding to \( f = \delta_0 \).

Understanding these solutions (existence, uniqueness, regularity) is closely tied to understanding level curves/surfaces for \( P(i\xi) \)

\[ \Sigma = \{ \xi \mid P(i\xi) = c \}. \]

For a second-order equation, all the level surfaces are non-degenerate quadrics of the same shape, possibly except for a couple of exceptional degenerate ones from the same family.

The shape of the level curves/surfaces of the symbol gives the name of the equation. If the shape is an ellipsoid, then the equation is called elliptic; if the shape is a paraboloid, then the equation is called parabolic; if the shape is a hyperboloid, then the equation is called hyperbolic.

These shapes are invariant under linear coordinate changes, as we’ll see below.

Example: Laplace’s equation \(-\Delta u = f\). The symbol is \( P(i\xi) = \|\xi\|^2 = \xi_1^2 + \ldots + \xi_d^2 \). Its level surfaces

\[ \|\xi\|^2 = c \]
are spheres (which fall under ellipsoids) for $c > 0$, a point for $c = 0$, and the empty set for $c < 0$. This is an elliptic equation.

**Example:** The wave equation $u_{tt} - \Delta_x u = f$. Its symbol is $-\tau^2 + \|\xi\|^2$. Its level surfaces

$$-\tau^2 + \|\xi\|^2 = c$$

are one-sheeted hyperboloids for $c > 0$, a cone for $c = 0$, and two-sheeted hyperboloids for $c < 0$. This is a hyperbolic equation.

### 1.2. Canonical form

In fact, one can always bring an equation such as (28) to the canonical form for its type, by linear transformations.

Let $A = (a_{i,j})_{1 \leq i, j \leq d}$ be the matrix of second-order coefficients and $\vec{b} = (b_i)_{1 \leq i \leq d}$ be the vector of first-order coefficients.

In the most common two settings, classical solutions and weak solutions, mixed partial derivatives are equal. So, with no loss of generality, we can assume that $A$ is symmetric, because if $a_{ij} \neq a_{ji}$ we can replace both by their half-sum.

In addition, we assume all coefficients are real. This excludes Schrödinger’s equation from our classification.

Using the (real) dot product of two matrices

$$\langle A, B \rangle = tr(AB) = \sum_{i,j=1}^{n} a_{ij} b_{ij},$$

we rewrite (28) as

$$A \cdot \nabla^2 u + \vec{b} \cdot \nabla u + cu = f(x).$$

For (28), it suffices to understand the effect of linear coordinate changes. For an invertible matrix $C \in M_d(\mathbb{R})$, det $C \neq 0$, let $y = Cx$, so $x = C^{-1}y$.

So

$$y_i = \sum_{j=1}^{d} c_{ij} x_j, \quad x_j = \sum_{i=1}^{n} (C^{-1})_{ji} y_i.$$

Then for any function $F$

$$\frac{\partial F}{\partial y_i} = \sum_{j=1}^{d} \frac{\partial F}{\partial x_j} \frac{\partial x_j}{\partial y_i} = \sum_{j=1}^{d} \frac{\partial F}{\partial x_j} (C^{-1})_{ji}.$$

So

$$(\nabla_y F)^T = (\nabla_x F)^T C^{-1} \quad \Rightarrow \quad (\nabla_y F)^T = (\nabla_y F)^T C \quad \Rightarrow \quad \nabla_x F = C^T \nabla_y F.$$

Likewise,

$$\frac{\partial^2 F}{\partial y_i \partial y_j} = \sum_{k, \ell=1}^{d} \frac{\partial^2 F}{\partial x_k \partial y_i} \frac{\partial x_k}{\partial y_j} \frac{\partial^2 F}{\partial x_\ell \partial y_j} \frac{\partial x_\ell}{\partial y_j},$$

meaning that

$$\nabla^2_y F = (C^{-1})^T (\nabla^2_x F) C^{-1}$$

or equivalently

$$\nabla^2_x F = C^T FC.$$

Thus in the new coordinates the equation becomes

$$(CAC^T) \cdot \nabla^2 u + (\vec{C}b) \cdot \nabla u + cu = f(x).$$

So $A$ transforms to $CAC^T$ and $\vec{b}$ transforms to $\vec{C}b$.

The relation $A \sim \tilde{A}$ defined by

$$A \sim \tilde{A} \equiv \exists C \in M_d(\mathbb{R}), \text{ det } C \neq 0, \text{ s.t. } \tilde{A} = CAC^T$$
is an equivalence relation, both on the set of matrices and on the set of symmetric matrices, so it partitions both $M_d(\mathbb{R})$ and the set of symmetric matrices into equivalence classes.

We need a way of distinguishing and understanding these equivalence classes. This is given by Sylvester’s law of inertia:

**Proposition 1.1.** Each equivalence class of symmetric matrices contains at least one diagonal matrix having only 1s, −1s, and zeros on the diagonal. All such matrices in the same equivalence class have the same numbers of 1s, −1s, and zeros.

The triple $(\sigma_+, \sigma_-, \sigma_0)$ is called the **signature** of the matrix. $\sigma_+ + \sigma_-$ is the rank of the matrix, with $\sigma_+$ being the number of positive and $\sigma_-$ being the number of negative eigenvalues.

**Proof.** Diagonalize the matrix, then apply a (possibly anisotropic) dilation. □

Thus, by a coordinate change the equation can always be made into

$$\frac{\partial^2 u}{\partial x_j^2} - \sum_{\sigma_{i+1}}^{\sigma_+} \frac{\partial^2 u}{\partial x_j^2} + \bar{b} \cdot \nabla u + cu = f(x).$$

In these coordinates, the matrix of second-order coefficients $A$ is a diagonal matrix with only $1, -1,$ or $0$ on the diagonal.

If $\sigma_0 > 0$, i.e. $\det A = 0$, the matrix $A$ is not of full rank and there are zeros on the diagonal, then the corresponding directions are called degenerate directions of the matrix $A$.

The next step to simplify the equation is to get rid of first-order terms. We can only do it in the non-degenerate directions. Thus, if $A$ has degenerate directions, the first-order terms in those directions cannot be disposed of and influence the type of the equation.

Let $x = (x_+, x_-, x_0)$, $x_+ \in \mathbb{R}^{\sigma_+}$, $x_- \in \mathbb{R}^{\sigma_-}$, $x_0 \in \mathbb{R}^{\sigma_0}$, be the decomposition of $x$ according to the directions of $A$ and same for $b$ and $\nabla = (\nabla_+, \nabla_-, \nabla_0)$. We write the equation as

$$\Delta_+ u - \Delta_- u + (\bar{b}_+ \cdot \nabla_+ + \bar{b}_- \cdot \nabla_- + \bar{b}_0 \cdot \nabla_0) u + cu = f(x).$$

We use the following change of variable:

$$v = e^{\bar{w}_+ \cdot \bar{x}_+ + \bar{w}_- \cdot \bar{x}_-} u,$$

for some suitably chosen $\bar{w}_+$ and $\bar{w}_-$. If $u = e^{-\bar{w}_- \cdot \bar{x}} v$, then

$$\nabla u = e^{-\bar{w}_- \cdot \bar{x}} (\nabla v - v \bar{w})$$

and

$$\Delta u = e^{-\bar{w}_- \cdot \bar{x}} (\Delta v - 2\nabla v \cdot \bar{w} + v \bar{w} \| \bar{w} \|^2).$$

Doing this for both the positive and the negative eigenvalues, in general the equation becomes

$$\Delta_+ v - \Delta_- v + (\bar{b}_+ - 2\bar{w}_+) \cdot \nabla_+ v + (\bar{b}_- + 2\bar{w}_-) \cdot \nabla_- v + \bar{b}_0 \cdot \nabla_0 v + (c + \| \bar{w}_+ \|^2 - \| \bar{w}_- \|^2) v = e^{\bar{w}_- \cdot \bar{x}} f.$$  

The same computation shows that this does nothing in the degenerate directions.

If we set $\bar{w}_+ = \frac{1}{2} \bar{b}_+$ and $\bar{w}_- = \frac{1}{2} \bar{b}_-$, then we can get rid of the first-order terms in non-degenerate directions.

Note, however, that in general we cannot get rid of the zeroth-order term in this manner, since that would necessitate

$$c + \frac{1}{4} \| \bar{b}_+ \|^2 - \frac{1}{4} \| \bar{b}_- \|^2 = 0.$$

Thus, every equation (28) can be reduced to

$$\Delta_+ u - \Delta_- u + \bar{b}_0 \cdot \nabla_0 u + cu = f.$$

In practice, though, we may not always want to get rid of first-order terms in this manner, so this latest transformation is more useful for classification purposes than in practice.
Hence we classify them as follows:
1. First, assume that \( A \) has no degenerate directions, \( \sigma_0 = 0 \). Then by a further change of variables we can set \( \bar{b} = 0 \), as explained above.

**1.3. Elliptic equations.** If all eigenvalues of \( A \) have the same sign (all positive or all negative, doesn’t matter which one, because we can get from one to the other by multiplying the equation by \( -1 \)), then modulo a change of variables the equation is equivalent to Laplace’s equation

\[-\Delta u + \bar{b} \cdot \nabla u + cu = f(x)\]

Then the equation is called (nondegenerate) elliptic.

**Definition 33.** A symmetric matrix whose eigenvalues are all

i. positive is called positive definite
ii. positive or zero is called positive semidefinite
iii. negative is called negative definite
iv. negative or zero is called negative semidefinite.

Otherwise it is called “of indefinite sign”.

A symmetric matrix \( A \in M_d(\mathbb{R}) \) is positive semidefinite if and only if for any \( \bar{v} \in \mathbb{R}^d \) \( \langle A\bar{v}, \bar{v} \rangle \geq 0 \). If, in addition, \( \langle A\bar{v}, \bar{v} \rangle = 0 \) implies \( \bar{v} = 0 \), then \( A \) is positive definite.

**Problem 51.** Show that a symmetric matrix is positive definite if and only if all its eigenvalues are positive.

**Problem 52.** Prove that if a matrix \( A \) is positive definite then there exists some constant \( k > 0 \) such that \( \langle Av, v \rangle \geq k\|v\|^2 \).

An elliptic equation can thus be reduced to

\[-\Delta u + cu = f(x)\]

Then, by a dilation, we can also make \( c \) be \( \pm 1 \) or 0, i.e. its absolute value doesn’t matter, only its sign does. These three cases actually have slightly different properties, so no further reduction is possible.

**1.4. Hyperbolic equations.** If all eigenvalues of \( A \) have the same sign except for one, which has the opposite sign, then we call the special direction corresponding to that eigenvalue \( t \) (time) and the remaining \( d - 1 \) directions \( x \) (space). The equation is then equivalent to

\[ \frac{\partial^2 u}{\partial t^2} - \Delta_x u + \bar{b} \cdot \nabla_{t,x} u + cu = f(x, t) \]

Such equations are called (nondegenerate) hyperbolic.

A hyperbolic equation can be reduced to

\[ u_{tt} - \Delta u + cu = f(x) \]

Then, by a dilation, we can also make \( c \) be \( \pm 1 \) or 0. If \( c = 0 \), this is called the wave equation, while if \( c = 1 \) it is called the Klein–Gordon equation.

These three cases again have quite different properties — except in \( \mathbb{R}^{1+1} \), where one can go from \( c = 1 \) to \( c = -1 \) by multiplying the whole equation by \( -1 \).
1.5. **Ultrahyperbolic equations.** If there are at least two positive and two negative eigenvalues, then the equation is called ultrahyperbolic. There is no preferred time direction. The canonical form is

$$\Delta_x u - \Delta_y u + \vec{b} \cdot \nabla_{x,y} u + cu = f(x, y),$$

where $x$ has $\sigma_+$ coordinates and $y$ has $\sigma_-$ coordinates.

Ultrahyperbolic equations are less widely studied than elliptic or hyperbolic ones.

2. If $A$ has null eigenvalues, $\sigma_0 \neq 0$, then the analysis has to take into account the first-order terms as well.

By the same method described above, we first reduce the equation to

$$\Delta_x u - \Delta_y u + \vec{b}_0 \cdot \nabla_{0} u + cu = f,$$

where $\vec{b}_0$ points only in the degenerate direction(s) of $A$.

If at this stage $\vec{b}_0 = 0$, then the equation is called degenerate, in the last remaining $\sigma_0$ directions, and is classified according to its type in the other directions.

**Example:** Let $u = u(x, y, z)$ satisfy Laplace’s equation in the first two variables only

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y, z).$$

This is a degenerate elliptic equation, degenerate in the $z$ direction.

Here the solution is straightforward: the equation decouples and for each value of $z$ we treat this as a separate elliptic equation, in $x$ and $y$ only.

In general, though, degenerate equations are harder than non-degenerate ones. The main problem is quantifying exactly how degenerate they are and how this affects the solution.

If, on the other hand, $\vec{b} \neq 0$, then we perform a rotation (in the last $\sigma_0$ coordinates only) so that it becomes aligned with one of the coordinate axes. The equation then becomes

$$\Delta_x u - \Delta_y u + \frac{\partial u}{\partial t} + cu = f,$$

where $b = \|\vec{b}\|$.

If $\sigma_0 = 1$, then the above equation involves all variables, hence is nondegenerate, but of a new type.

If $\sigma_0 > 1$, then still are some degenerate directions left even after accounting for the first-order terms, so the equation is degenerate.

1.6. **Parabolic equations.** If $\sigma_+ = 0$ or $\sigma_- = 0$, then we are dealing with a parabolic equation:

$$b \frac{\partial u}{\partial t} \pm \Delta_x u + cu = f.$$

Dividing by $b$ and making the substitution $v = e^{c/b t} u$, we can get rid of the first-order term. By changing $t$ to $-t$ we can also fix the sign. We arrive at

$$\frac{\partial u}{\partial t} - \Delta_x u = f.$$

Thus we can reduce any parabolic equation with constant coefficients to the heat equation.

1.7. **Other equations.** If both $\sigma_+ > 0$, $\sigma_- > 0$, and $\sigma_0 > 0$, the corresponding class of equations does not have a specific name, but could be called parabolic-hyperbolic. No important equation of science or engineering falls into this class, as far as I know.
1.8. Equations with nonconstant coefficients. When the coefficients are not constant

\[ Lu = \sum_{k,\ell=1}^{d} a_{k\ell}(x) \frac{\partial^2 u}{\partial x_k \partial x_\ell} + \sum_{k=1}^{d} b_k(x) \frac{\partial u}{\partial x_k} + c(x) u = f(x), \]

let’s consider more general nonlinear coordinate changes. Let \( y = y(x), x \in \mathbb{R}^d, y \in \mathbb{R}^d \), and \( J \) be the Jacobian matrix of the coordinate change:

\[ J = \nabla_x y = \left( \frac{\partial y_i}{\partial x_j} \right)_{1 \leq i,j \leq d}, \]

where \( i \) is the row and \( j \) is the column. We may assume that \( \det J \neq 0 \).

Then by the chain rule

\[ \frac{\partial u}{\partial x_j} = \sum_{\ell=1}^{d} \frac{\partial u}{\partial y_\ell} \frac{\partial y_\ell}{\partial x_j}, \]

so in matrix notation

\[ \nabla_x u = \nabla_y u \cdot J^T. \]

This is the same as in the constant coefficient case. When taking one more derivative, though, there is an extra term from the product rule:

\[ \frac{\partial^2 u}{\partial x_i \partial x_j} = \sum_{k,\ell=1}^{d} \frac{\partial^2 u}{\partial y_k \partial y_\ell} \frac{\partial y_k}{\partial x_i} \frac{\partial y_\ell}{\partial x_j} + \sum_{k,\ell=1}^{d} \frac{\partial u}{\partial y_\ell} \frac{\partial^2 y_\ell}{\partial x_i \partial x_j}. \]

In matrix notation

\[ (29) \quad \nabla^2 u = J \nabla^2 u J^T + \nabla_y u \cdot \nabla^2 x y. \]

The second-order derivatives change as in the constant coefficient case, except that the Jacobian matrix replaces the change-of-basis matrix.

However, the second term in (29) is an extra first-order term. Thus, in order to compute the first-order coefficients in the new coordinates, we also need to know the second-order coefficients.

In general, the highest-order derivatives transform independently of the lower ones, but the lower ones also get a contribution from the higher-order ones.

In particular: if at some point in the domain the equation is nondegenerate, then it remains nondegenerate, if it is elliptic then it remain elliptic, if it is hyperbolic then it remains hyperbolic.

The coordinate change may produce new first-order terms, but only in the non-degenerate directions. Thus, a parabolic equation will remain parabolic too.

This means we can still classify equations with nonconstant coefficients into elliptic, hyperbolic, parabolic, etc., but the classification must be done pointwise. An equation with nonconstant coefficients may be of different types at different points in the domain.

Example:

\[ u_{xx} + y u_{yy} = 0 \]

is elliptic in the upper half-plane and hyperbolic in the lower half-plane (and degenerate elliptic on the boundary, but that doesn’t matter).

Example: Consider the equation

\[ 7u_{xx} + 5u_{yy} + 6u_{zz} = f. \]

The matrix of second-order coefficients is

\[ A = \begin{pmatrix} 7 & 0 & -2 \\ 0 & 5 & -2 \\ -2 & -2 & 6 \end{pmatrix}. \]

Eigenvalues: 3, 6, 9.
Eigenvectors: $(1, 2, 2), (2, -2, 1), (2, 1, -2)$.

**Example:** Let’s bring the following equation to its canonical form:

$$u_{xx} - 4u_{xy} + u_{yy} + 2u_x + 3u_y + 4u = 0.$$

First, we look at the matrix of second-order coefficients

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}.$$

It has eigenvalues $3$ and $-1$. The corresponding eigenvectors are $(1, -1), (1, 1)$. There is no need to normalize them.

Change of variables: $\tilde{x} = x + y, \tilde{y} = x - y.$ We use the chain rule:

$$\frac{\partial}{\partial \tilde{x}} = \frac{\partial}{\partial x} \frac{\partial x}{\partial \tilde{x}} + \frac{\partial}{\partial y} \frac{\partial y}{\partial \tilde{x}} = \frac{\partial}{\partial \tilde{x}} + \frac{\partial}{\partial \tilde{y}}.$$

So

$$\frac{\partial^2}{\partial \tilde{x}^2} = \left(\frac{\partial}{\partial \tilde{x}} + \frac{\partial}{\partial \tilde{y}}\right)^2, \quad \frac{\partial^2}{\partial \tilde{y}^2} = \left(\frac{\partial}{\partial \tilde{x}} - \frac{\partial}{\partial \tilde{y}}\right)^2, \quad \frac{\partial^2}{\partial \tilde{x} \partial \tilde{y}} = \left(\frac{\partial}{\partial \tilde{x}} + \frac{\partial}{\partial \tilde{y}}\right)\left(\frac{\partial}{\partial \tilde{x}} - \frac{\partial}{\partial \tilde{y}}\right).$$

In the new coordinates, the equation then becomes

$$-2u_{\tilde{x}\tilde{x}} + 6u_{\tilde{y}\tilde{y}} + 5u_{\tilde{x}} - u_{\tilde{y}} + 4u = 0.$$

At this step, we can already tell that the equation is hyperbolic.

We make a second change of variables to bring the second-order terms to their canonical form, i.e. get rid of the coefficients $2$ and $6$:

$$x^* = 1/\sqrt{2}\tilde{x}, \quad y^* = 1/\sqrt{6}\tilde{y}.$$

This leads to

$$-u_{x^*x^*} + u_{y^*y^*} + 5/\sqrt{2}u_{x^*} - 1/\sqrt{6}y^* + 4u = 0.$$

Next, in order to get rid of the first-order terms, set

$$\tilde{u} = u(x^*, y^*)e^{ax^* + by^*}.$$

For the appropriate choice of $a = -\frac{5}{2\sqrt{2}}, b = \frac{1}{2\sqrt{6}},$ we get

$$-\tilde{u}_{x^*x^*} + \tilde{u}_{y^*y^*} + \tilde{u}(4 - 25/8 + 1/24 + 25/4 + 1/12) = 0,$$

where $4 - 25/8 + 1/24 + 25/4 + 1/12 = 29/4$.

This is an equation of Klein–Gordon type. We can make one more simplification using by the isotropic dilation

$$x^# = \sqrt{\frac{29}{4}}x^*, \quad y^# = \sqrt{\frac{29}{4}}y^*.$$

Finally, we have obtained the standard form

$$-\tilde{u}_{x^#x^#} + \tilde{u}_{y^#y^#} + \tilde{u} = 0.$$

**Problem 53.** Consider the equation

$$u_{xx} - 2u_{xy} + u_{yy} + u_x + 2u_y + 3u = 0,$$

where $u = u(x, y)$. Bring this equation to its canonical form (as simple as you can make it) and classify it.
Problem 54. Consider the equation
\[ u_{xx} - u_{xy} + u_x - u_y + u = 0, \]
where \( u = u(x, y) \). Bring this equation to its canonical form (as simple as you can make it) and classify it.

2. Elliptic equations

The general form of (nondegenerate) elliptic equations is, for \( u : D \to \mathbb{R} \),
\begin{equation}
Lu = P(D)u = -A(x) \cdot \nabla^2 u + \vec{b}(x) \cdot \nabla u + c(x)u = f,
\end{equation}
where \( A \) is a positive definite matrix: at each \( x \in D \), \( \langle Av, v \rangle \geq k \|v\|^2 \), \( k > 0 \).

If the constant \( k(x) = k \) does not depend on \( x \), then the equation is called uniformly elliptic. Otherwise, obviously, it’s non-uniformly elliptic.

Non-uniformly elliptic equations and degenerate elliptic equations have the same properties as non-degenerate, uniformly elliptic equations, but to a lesser extent.

There are three main approaches in the study of elliptic equations: elliptic regularity, the energy identity, and the maximum principle (also see the related Harnack’s inequality).

Methods for studying parabolic equations are by and large similar to the above. Hyperbolic equations, on the other hand, are quite different: there are no maximum principles or regularizing effects. Instead, one uses characteristics and frequency or microlocal (wave packet) decompositions.

Elliptic regularity: In many situations, the solution \( u \) of a non-degenerate elliptic equation has two more derivatives than the right-hand-side terms.

Examples: If \( u, f \in L^2 \) then \( u \in H^2 \):
\[ \|u\|_{H^2} \leq C(\|u\|_{L^2} + \|f\|_{L^2}). \]

If \( u, f \in L^p \) then \( u \in W^{2,p} \):
\begin{equation}
\|u\|_{W^{2,p}} \leq C(\|u\|_{L^p} + \|f\|_{L^p}).
\end{equation}

If \( u, f \in C^\alpha \) then \( u \in C^{2,\alpha} \):
\begin{equation}
\|u\|_{C^{2,\alpha}} \leq C(\|u\|_{C^\alpha} + \|f\|_{C^\alpha}).
\end{equation}

This last one is called a Schauder-type estimate.

Thus, by elliptic regularity one almost always gains two derivatives, in several rather different settings.

Unfortunately, this does not work in the endpoint settings of \( L^1 \), \( L^\infty \), or \( C \) (the space of continuous functions). In particular, having this estimate be true in \( C \) would have been convenient, because it would have led to \( C^2 \) (classical) solutions.

In general, \( W^{2,p} \) solutions are weak solutions, while \( C^{2,\alpha} \) solutions are classical solutions, for \( \alpha > 0 \). So why not do everything using Schauder estimates and in the setting of Hölder spaces? Because this is only an option for elliptic and parabolic equations; however, Sobolev spaces work for hyperbolic (and Schrödinger) equations too.

2.1. Elliptic equations with constant coefficients. The general form of second-order constant-coefficient elliptic equations is
\begin{equation}
Lu = P(D)u = -A \cdot \nabla^2 u + \vec{b} \cdot \nabla u + cu = f,
\end{equation}
where \( A \) is a positive definite matrix: \( \langle Av, v \rangle \geq k \|v\|^2 \), \( k > 0 \).

Recall that
\begin{equation}
P(i\xi) = \sum a_{ij} \xi_i \xi_j + i\vec{b} \cdot \xi \xi + c = \langle A\xi, \xi \rangle + i\vec{b} \cdot \xi + c
\end{equation}
is called the symbol of this operator. After taking the Fourier transform, a solution is given by
\[ \hat{u} = \frac{1}{P(i\xi)} \hat{f}. \]

**Example:** Consider Laplace’s equation on \( \mathbb{R}^d \)
\[ (-\Delta + 1)u = f. \]
The solution is given by
\[ \hat{u} = \frac{\hat{f}}{\|\xi\|^2 + 1}, \]
so in three dimensions
\[ u = \frac{1}{2\pi} \mathcal{F}^{-1} \frac{1}{\|\xi\|^2 + 1} * f = \frac{e^{-|x|}}{4\pi|x|} * f. \]
We could deduce some of the properties of \( u \) from this formula, using e.g. Young’s inequality.

**Example:** Fundamental solution to Laplace’s equation. Consider Laplace’s equation on \( \mathbb{R}^d \)
\[ -\Delta u = f. \]
A solution is given by
\[ u = f * \frac{1}{2\pi} \mathcal{F}^{-1} \frac{1}{\|\xi\|} \cdot \]
This solution is not unique, but will become so under extra conditions, such as decay at infinity.
In general, the difference of any two solutions must be a harmonic function.
The quantity
\[ \frac{1}{(2\pi)^d} \mathcal{F}^{-1} \frac{1}{\|\xi\|^2} \]
is called the fundamental solution of Laplace’s equation. It is itself a solution of Laplace’s equation with Dirac’s delta on the right-hand side:
\[ -\Delta \frac{1}{(2\pi)^d} \mathcal{F}^{-1} \frac{1}{\|\xi\|^2} = \frac{1}{(2\pi)^d} \mathcal{F}^{-1} 1 = \delta_0. \]
These computations make sense for tempered distributions.
The fundamental solution of Laplace’s equation is then:
\[ G(|x|) = -\frac{1}{2} |x| \]
in one dimension;
\[ G(|x|) = -\frac{1}{2\pi} \ln |x| \]
in two dimensions;
\[ G(|x|) = \frac{1}{4\pi r} \]
in three dimensions; and more generally
\[ G(|x|) = \frac{1}{|\partial B(0,1)|} r^{2-d} \]
in \( \mathbb{R}^d, \ d \geq 3. \) Here \( |\partial B(0,1)| \) is the area/volume/measure of the \( d - 1 \)-dimensional unit sphere.
In particular, these functions are harmonic on \( \mathbb{R}^d \setminus \{0\} \).
We can also do the computation directly on the Fourier side. Let’s consider a slightly more general example, to see how ellipticity works:
**Example:** Consider the equation
\[ LU = -\sum a_{ij} \partial_i \partial_j u + cu = f, \]
where \( c > 0 \). Then
\[
\hat{u} = \sum a_{ij} \xi_i \xi_j + c
\]
Suppose that the operator is elliptic, but possibly degenerate:

(35) 
\[
\sum a_{ij} \xi_i \xi_j \geq 0
\]
for any vector \( \xi \in \mathbb{R}^d \); then
\[
|\hat{u}(\xi)| = \left| \frac{\hat{f}(\xi)}{a_{ij} \xi_i \xi_j + c} \right| \leq \left| \frac{\hat{f}(\xi)}{c} \right|
\]
Since this is true for each \( \xi \in \mathbb{R}^d \), integrating over \( \mathbb{R}^d \) we get
\[
\|u\|_{L^2} = \frac{\sqrt{2\pi}}{c} \|\hat{f}\|_{L^2} \leq \frac{\sqrt{2\pi}}{c} \|\hat{f}\|_{L^2} = \frac{1}{c} \|f\|_{L^2}.
\]
This only works in \( L^2 \), due to Plancherel’s identity, not in any other \( L^p \) spaces; the proof for \( p \neq 2 \) is different and significantly more complicated.

Next, also suppose the operator is non-degenerate, i.e.

(36) 
\[
\sum a_{ij} \xi_i \xi_j \geq k \|\xi\|^2
\]
Then there is a gain of two derivatives: for each \( \xi \in \mathbb{R}^d \)
\[
|(1 + \|\xi\|^2) \hat{u}| = \left| \frac{\hat{f}}{1 + \|\xi\|^2} \right| \leq \left| \frac{\hat{f}}{k \|\xi\|^2 + c} \right| \leq \frac{1}{\min(k, c)} |\hat{f}|,
\]
so
\[
\|u\|_{H^2} = \|(1 + \|\xi\|^2) \hat{u}\|_{L^2} = \left\| \frac{\hat{f}}{k \|\xi\|^2 + c} \right\|_{L^2} \leq \frac{1}{\min(k, c)} \|\hat{f}\|_{L^2} = \frac{1}{\sqrt{2\pi} \min(k, c)} \|f\|_{L^2}.
\]
Note the difference between non-degenerate elliptic (36) and degenerate elliptic (35): in the first case we gain two derivatives, in the latter not necessarily so; or perhaps only in the non-degenerate directions.

The above computation doesn’t quite work:
* if \( c \leq 0 \)
* or if there are first-order terms.

Example: Consider the equation \(-\Delta u = f\) on \( \mathbb{R}^3 \). Then
\[
u = \frac{1}{2\pi} F^{-1} \frac{1}{|\xi|^2} * f = \frac{1}{4\pi |x|} * f.
\]
So even if \( f \) is smooth and compactly supported (i.e. \( f \in \mathcal{D} \)) then \( u \) will decay like \( |x|^{-1} \) at infinity; the first term in its asymptotic expansion will be \( \frac{1}{2\pi |x|} \int f \). This cannot be in \( L^2 \) unless \( \int f = 0 \); so in general it’s not.

Nevertheless, elliptic regularity, with a gain of two derivatives, is expected for second-order elliptic equations.

\( H^2 \) functions are weak solutions to (33), i.e. in the sense of distributions. The derivatives are also taken in the weak sense. If we want to upgrade to classical solutions, one possibility is to use a Sobolev embedding.

**Proposition 2.1.** On \( \mathbb{R}^d \), for \( s > d/2 \), \( H^s \subset C \) and
\[
\max_x |f(x)| \leq C \|f\|_{H^s}.
\]

**Proof.** The proof has two parts: first we show that \( f \) is bounded and then we uniformly approximate it by continuous functions, thus proving it must also be continuous. \( \square \)
Proposition 2.2. Consider the equation (33). If $s > d/2$ and $f \in H^s$, then $u \in C^2$ is a classical solution to the equation.

Proof. Each derivative of $u$ up to second order satisfies the equation
$$LD^\alpha u = D^\alpha f.$$ Since $f \in H^s$, by definition $D^\alpha f \in H^{s-|\alpha|} \subset H^{s-2}$. Due to the elliptic regularity, this implies that $D^\alpha u \in H^s \subset C$. Thus, $u \in C^2$. Since $u$ is a weak solution and is also in $C^2$, it must therefore be a classical solution. □

Much more generally, there are various abstract definitions of “ellipticity”. Here is one of them.

Definition 34. A constant-coefficient operator $L$ is called elliptic if the absolute value of its symbol $P(i\xi)$ is bounded from below by some power of $|\xi|$ outside some bounded set:
$$\exists R \exists C \exists N \forall \xi \in \mathbb{R}^d \|\xi\| \geq R \implies |P(i\xi)| \geq C|\xi|^N.$$ This definition is clearly satisfied by (34).

It is easy to find elliptic operators of any even order, e.g. $(-\Delta)^n$. However, we’ll stick to second-order equations here.

Proposition 2.3. Any $L^2$ solution $u$ of equation (33) must fulfill the inequality
$$\|u\|_{H^2} \leq C(\|f\|_{L^2} + \|u\|_{L^2}).$$ In addition, if the zeroth-order coefficient $c$ is sufficiently positive, then
$$\|u\|_{H^2} \leq C\|f\|_{L^2}.$$ This is no longer true if the equation is degenerate.

Proof. Inside any bounded set in phase space, $\|u\|_{H^2}$ is controlled by $\|u\|_{L^2}$. Outside this bounded region, we use the ellipticity of the equation.

The real part of $P(i\xi)$ is bounded from below, so after adding a sufficiently large constant to it the absolute value of $P(i\xi) + \hat{c}$ becomes uniformly positive:
$$|P(i\xi)| \geq k > 0.$$ Then for the solution $u$ of (37)
$$|\hat{u}(\xi)| = \left| \frac{\hat{f}(\xi)}{|P(i\xi)|} \right| \leq \frac{1}{k} |\hat{f}(\xi)|,$$ hence
$$\|u\|_{L^2} \leq \frac{1}{k} \|f\|_{L^2}.$$ □

2.2. Elliptic equations with non-constant coefficients.

2.2.1. Divergence form. Given a second-order elliptic equation
$$L = \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \ldots$$ it is sometimes advantageous to commute one of the derivatives through the coefficients:
$$- \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} = - \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + \sum_{i,j=1}^d \frac{\partial a_{ij}}{\partial x_i} \frac{\partial u}{\partial x_j}$$ $$= - \text{div}(A \nabla u) + (\text{div } A) \cdot \nabla u,$$ where $A = (a_{ij})$ is the matrix of second-order coefficients. This is called the divergence form.
If the coefficients are nonconstant, this transformation also changes the first-order terms. If the coefficients are constant, then \( \text{div} A = 0 \) and the first-order terms remain unchanged.

Sometimes using the divergence form is preferable, sometimes not. Having the equation in divergence form helps when dealing with very rough coefficients, see below.

Equally helpful is to write the first-order terms \( \vec{b} \cdot \nabla u \) in a symmetric form

\[
\vec{b} \cdot \nabla u = \frac{1}{2} (\vec{b} \cdot \nabla u + \text{div}(\vec{b}u)) - \frac{1}{2} (\text{div} \vec{b}u).
\]

2.2.2. \( H^1 \) solutions. Here is a general existence and uniqueness result.

**Proposition 2.4.** Let

\[
L = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^{d} b_i(x) \frac{\partial u}{\partial x_i} + c(x)u = -\text{div}(A(x) \cdot \nabla u) + \vec{b}(x) \cdot \nabla u + c(x)u
\]

be an elliptic operator on some domain \( D \subset \mathbb{R}^d \) such that:

i. There exists a constant \( k > 0 \) such that \( \langle A(x)v,v \rangle \geq k \|v\|^2 \) for every \( x \in D \) and \( v \in \mathbb{R}^d \) (uniform ellipticity)

ii. \( b_i, c \in L^\infty \).

Let \( u \in H^1 \) be a solution of the equation

\[
Lu = f, \quad u|_{\partial D} = 0.
\]

Then

\[
\|u\|_{H^1} \leq C(\|f\|_{H^{-1}} + \|u\|_{L^2}).
\]

For a sufficiently large constant \( c(x) \geq c_0 \), also assuming \( \text{div} b \in L^\infty \), the equation \( Lu = f \) always possesses a \( H^1 \) solution \( u \), which fulfills

\[
\|u\|_{H^1} \leq C\|f\|_{H^{-1}}.
\]

So again there is a two-derivative gain, from \( H^{-1} \) to \( H^1 \). Neumann boundary conditions are just as good.

When the equation is not in divergence form, the proof involves, directly or indirectly, putting it in divergence form. Thus, in general, we also need a condition of the type

\[
\nabla A \in L^\infty.
\]

**Proof.** Taking the dot product of both sides of the equation with \( u \), we get

\[
\langle Lu, u \rangle = \langle f, u \rangle.
\]

But

\[
\langle Lu, u \rangle = \int_D (A(x) \nabla u) \cdot \nabla u + (\vec{b}(x) \cdot \nabla u)u + c(x)u^2.
\]

Indeed, by Gauss’s theorem,

\[
\int_D \text{div}[u(A(x) \nabla u)] = \int_{\partial D} u(A(x) \nabla u) \cdot \vec{n} = 0.
\]

Note that the boundary terms are zero because \( u|_{\partial D} = 0 \), but \( \nabla u \cdot \vec{n} = \frac{du}{dn} = 0 \) would work just as well. But

\[
\text{div}[u(A(x) \cdot \nabla u)] = u \text{div}(A(x) \nabla u) + \nabla u \cdot (A(x) \nabla u).
\]

This proves the validity of the following integration by parts:

\[
-\int_D u \text{div}(A(x) \nabla u) = \int_D \nabla u \cdot (A(x) \nabla u).
\]
Due to uniform ellipticity, we get that
\[ k \int_D |\nabla u|^2 \leq -\int_D (b(x) \cdot \nabla u) u + c(x) u^2 + \int_D f u \leq \int_D \|b\|_{L^\infty} |\nabla u| |u| + \|c\|_{L^\infty} u^2 + \frac{1}{2} (f^2 + u^2). \]
Here we used the means inequality
\[ ab \leq \frac{1}{2} (a^2 + b^2). \]
However, the next step is absorbing the $\nabla u$ term in the left-hand side. For that, we use the following modified version of the means inequality (obtained from the original by replacing $a$ by $a \sqrt{\epsilon}$ and $b$ by $b / \sqrt{\epsilon}$):
\[ ab \leq \frac{1}{2} (\epsilon a^2 + \frac{1}{\epsilon} b^2). \]
So
\[ \|b\|_{L^\infty} |\nabla u| |u| \leq \frac{1}{2} (\epsilon |\nabla u|^2 + \frac{1}{\epsilon} \|b\|^2_{L^\infty} u^2). \]
Set $\epsilon = k$. We have thus obtained that
\[ \frac{k}{2} \int_D |\nabla u|^2 \leq \int_D \frac{1}{2k} \|b\|^2 u^2 + \|c\| u^2 + \frac{1}{2} (f^2 + u^2) = \left( \frac{\|b\|^2}{2k} + \|c\| \right) \|u\|^2_{L^2} + \frac{1}{2} \|f\|^2_{L^2}. \]
The conclusion follows.

However, here we used the $L^2$ norm of $f$ on the right-hand side, for a merely one-derivative gain. In order to use the $H^{-1}$ norm, we estimate $\langle f, u \rangle$ as follows:
\[ \langle f, u \rangle \leq \|u\|_{H^1} \|f\|_{H^{-1}} \leq \frac{\epsilon}{2} \|u\|^2_{H^1} + \frac{1}{2\epsilon} \|f\|^2_{H^{-1}}. \]
But
\[ \|u\|^2_{H^1} \leq \int \int_D |\nabla u|^2 + u^2. \]
By making $\epsilon$ sufficiently small, $\epsilon = \frac{k}{2}$, we can absorb the $|\nabla u|^2$ term into the left-hand side.

In the second case, we can also prove that the $H^1$ solution exists, using the Riesz representation theorem.

The Riesz representation theorem states that every linear and continuous functional on a Hilbert space is given by an element of the same.

If a solution $u$ to the equation existed, it would have to be the unique function that satisfies the identity
\[ \langle f, v \rangle = \langle Lu, v \rangle. \]
Consider the Sobolev space of functions that vanish on the border
\[ H^1_0 = \{ u \in H^1 | u|_{\partial D} = 0 \}. \]
The actual definition of $H^1_0$ is somewhat different and makes sense rigorously of vanishing on the border.

This is a Hilbert space. One possible choice of dot product on it is
\[ \langle u, v \rangle = \int_D \nabla u \cdot \nabla v. \]
Note that we omitted the zeroth-order terms; this is because they are superfluous for $H^1_0$, being controlled by the first-order derivatives:
\[ \int_D u^2 \leq C \int_D |\nabla u|^2. \]
This is called Poincaré’s inequality.
Consider another dot product, given by
\[
\left< u, v \right>_{H^1_0} = \left< Lu, v \right> = \int_D (- \text{div}(A \nabla u) + \vec{b} \cdot \nabla u + cu)v = \int_D (A \nabla u) \cdot \nabla v - \frac{1}{2}(\text{div} \vec{b})uv + cuv.
\]

Assuming that \( c \) is uniformly large and \( \text{div} \vec{b} \in L^\infty \) (these hypotheses can be modified), it is easy to check that this is a dot product equivalent to the previous one; positive definiteness is the nontrivial part.

Then, it is easy to see that \( v \mapsto Tv = \left< f, v \right> \) is a continuous linear functional on \( H^1_0 \), due to Poincaré’s inequality. By the Riesz representation theorem, it follows that this functional can be represented by
\[
Tv = \left< u, v \right>_{H^1_0}
\]
for some \( u \in H^1_0 \). Since this is true for every \( v \in H^1_0 \), \( u \) must be a solution to the equation.

2.2.3. \( H^2 \) solutions: the method of freezing coefficients. On a domain \( D \in \mathbb{R}^d \), consider the elliptic equation
\[
Lu = - \sum a_{ij}(x) \partial_i \partial_j u + \sum b_i(x) \partial_i u + c(x)u = f.
\]
If the right-hand side is in \( L^2 \), we’ll prove that the solution is in \( H^2 \).

The method of proof is called the coefficient freezing method, introduced by Nirenberg. The idea is that we localize the equation in some small neighborhood of each point in the domain. Then, it reduces to the constant coefficient case plus some small error terms that we can control. This requires continuous coefficients. Finally, we piece it back together and get a conclusion for the whole domain.

This is harder to do near the boundary, especially if the domain’s boundary is insufficiently smooth. In general, interior and boundary regularity are two distinct problems. The exact amount of boundary regularity of solutions depends on the boundary’s geometry and on the boundary conditions. Interior regularity, on the other hand, is universal.

Here we prove interior regularity:

**Proposition 2.5.** On a bounded domain \( D \in \mathbb{R}^d \), let \( u \in H^1 \) be a solution of the elliptic equation
\[
Lu = - \sum a_{ij}(x) \partial_i \partial_j u + \sum b_i(x) \partial_i u + c(x)u = f,
\]
where the operator \( L \) is elliptic, i.e. for and all \( x \in D \) there exists \( k(x) > 0 \) such that \( \left< A(x)v, v \right> \geq k(x)\|v\|^2 \).

In addition, assume that the coefficients \( a_{ij}(x) \) are continuous and that \( b \) and \( c \) are bounded.

Then, for every \( \delta > 0 \), the solution \( u \) is in \( H^2(D_\delta) \), where \( D_\delta \) is the portion of the domain situated at distance at least \( \delta \) from the boundary:
\[
\|u\|_{H^2(D_\delta)} \leq C(\|f\|_{L^2} + \|u\|_{H^1}).
\]

Presumably, we know that the solution \( u \in H^1 \) exists by other means, such as the previous theorem (for which we need to put the equation in divergence form), so here we are only concerned with establishing the extra regularity.

**Proof.** Firstly, since \( u \in H^1 \), we can use this fact to handle all the lower-order terms:
\[
\tilde{L}u = - \sum a_{ij}(x) \partial_i \partial_j u = \tilde{f} = f - \sum b_i(x) \partial_i u - cu,
\]
where
\[
\|\tilde{f}\|_{L^2} \leq \|f\|_{L^2} + \max(\|\vec{b}\|_{L^\infty}, \|c\|_{L^\infty})\|u\|_{H^1}.
\]
Consider some point \( x_0 \) in the interior of \( D \) and let \( \chi : \mathbb{R}^d \to \mathbb{R} \) be a smooth cutoff function with the following properties:

\[
\chi(x) = 1, \quad |x| \leq 1, \quad \chi(x) = 0, \quad |x| \geq 2, \quad \chi \in C^\infty, \quad 0 \leq \chi \leq 1.
\]

Let \( \tilde{u} \) be the solution \( u \) cut off to a small neighborhood of size \( \epsilon \) around \( x_0 \):

\[
\tilde{u}(x) = u(x)\chi(x) = u(x)\chi\left(\frac{x - x_0}{\epsilon}\right).
\]

Let’s compute the equation satisfied by \( \tilde{u} \):

\[
\tilde{L}\tilde{u} = f_\epsilon = \chi_\epsilon f + \sum a_{ij} \partial_i \partial_j \chi_\epsilon u + 2 \sum \partial_i \chi_\epsilon a_{ij} \partial_j u.
\]

The right-hand side may be bounded by

\[
\|f_\epsilon\|_L^2 = \|\chi_\epsilon f + \sum a_{ij} \partial_i \partial_j \chi_\epsilon u + 2 \sum \partial_i \chi_\epsilon a_{ij} \partial_j u\|_L^2 \\
\leq \|f\|_L^2 + C(1 + \epsilon^{-2})(\|A\|_L^\infty + \|\tilde{b}\|_L^\infty + \|c\|_L^\infty)\|u\|_{H^1}.
\]

So far there is no gain from this computation; in fact the \( \epsilon^{-2} \) is a loss. The gain will come from the fact that in a sufficiently small neighborhood of \( x_0 \) the coefficients can be approximated by constants.

Let

\[
\tilde{L}_0 = \sum a_{ij}(x_0) \partial_i \partial_j.
\]

Then

\[
(38) \quad \tilde{L}_0\tilde{u} = f_\epsilon + (\tilde{L}_0 - \tilde{L})u_\epsilon.
\]

Since \( \tilde{L}_0 \) is a constant-coefficient operator, we can apply the standard elliptic regularity estimates and get that

\[
(39) \quad \|\tilde{u}\|_{H^2} \leq C \min(1, 1/k(x_0))\|f_\epsilon + (\tilde{L}_0 - \tilde{L})u_\epsilon\|_L^2 \leq C \min(1, 1/k(x_0))\|f_\epsilon\|_L^2 + \|(\tilde{L}_0 - \tilde{L})u_\epsilon\|_L^2.
\]

However, \( u_\epsilon \) is supported on \( B(x_0, \epsilon) \), so the second term can be bounded by

\[
\|(\tilde{L}_0 - \tilde{L})u_\epsilon\|_L^2 \leq \sup_{x \in B(x_0, \epsilon)} \|A(x) - A(x_0)\|\|\nabla^2 u\|_L^2.
\]

Since the coefficients \( a_{ij} \) are continuous (equivalently \( \lim_{x \to x_0} A(x) = A(x_0) \)), for sufficiently small \( \epsilon < \epsilon_0 \sup_{x \in B(x_0, \epsilon)} \|A(x) - A(x_0)\| \) can be made arbitrarily small and absorbed into the left-hand side:

\[
C \sup_{x \in B(x_0, \epsilon)} \|A(x) - A(x_0)\| \leq \frac{1}{2C \min(1, 1/k(x_0))}.
\]

Thus, for \( \epsilon < \epsilon_0 \),

\[
\|\tilde{u}\|_{H^2} \leq C \frac{\min(1, 1/k(x_0))\|f_\epsilon\|_L^2}{2}.
\]

This is only a formal computation, given that if \( u \notin H^2 \) both sides in (39) are infinite, so one cannot perform a cancellation. The actual, rigorous justification is as follows: instead of (38), consider the equation

\[
\tilde{L}_0 v_n = f_\epsilon + (\tilde{L}_0 - \tilde{L})v_{n-1},
\]

where \( v_0 = 0 \) and each subsequent term is recursively defined in terms of the previous one.

We prove that \( (v_n) \) is a Cauchy sequence in \( H^2 \). Since \( H^2 \) is a complete metric space, the sequence must have a limit, call it \( \tilde{v} \).

By passing to the limit, \( \tilde{v} \) is a solution of (38). Then, we prove that the solution is unique in \( H^1 \), so in fact \( \tilde{u} = \tilde{v} \in H^2 \).

But \( u = \tilde{u} \) on \( B(x_0, \epsilon) \). In conclusion, every point \( x_0 \) in the interior of the domain has some neighborhood \( B(x_0, \epsilon) \) on which \( u \in H^2 \) and

\[
\|u\|_{H^2(B(x_0, \epsilon))} \leq C(x_0, \epsilon)\|u\|_{H^1(B(x_0, 2\epsilon))}
\]
for $\epsilon < \epsilon_0(x_0)$.

Importantly, this estimate is not “self-contained”: we need to know the $H^1$ norm of $u$ on a slightly larger neighborhood of $x_0$ in order to bound its $H^2$ norm on a slightly smaller one.

$D_\delta$ is covered by the union of these neighborhoods. The set $D_\delta$ is closed and bounded, therefore compact, so it can be covered by finitely many of them. Adding up the estimates on these finitely many neighborhoods, we obtain a finite $H^2$ norm on $D_\delta$. \hfill \Box

In order to deal with the presence of a boundary, we use a change of coordinates that straightens (a piece of) that region of the domain to a half-space (or to a cone, or to some other standard shape that best represents it). This works best when the boundary is smooth.

Note that we did not use any special feature of $H^2$, so the same proof works for establishing elliptic regularity in $W^{2,p}$ or in $C^{2,\alpha}$, assuming that we can first prove (31) and (32) for constant-coefficient equations.

2.2.4. Harmonic functions.

Definition 35. A function $u : D \to \mathbb{R}$ is called harmonic if $-\Delta u = 0$.

We need not specify how smooth $u$ is, because every locally integrable harmonic function is infinitely smooth (theorem). In fact, even every harmonic distribution must be given by an infinitely smooth function (theorem).

Examples of harmonic functions: in two dimensions: $x^2 - y^2$, $xy$, $e^x \sin y$, etc.

Theorem 2.6. Consider a solution $u : D \to \mathbb{R}, u \in C^2$, of Laplace’s equation $-\Delta u = 0$ and $B(x_0, R) \subset D$. Then

(40) \[ u(x_0) = \int_{\partial B(x_0, R)} u(x) \, dx. \]

Integrating in $R$ with a weight, (40) implies that

(41) \[ u(x_0) = \int_{\partial B(x_0, R)} u(x) \, dx. \]

This implies that every $C^2$ harmonic function is infinitely smooth.

Proof. Without loss of generality, let’s assume $x_0 = 0$. We use the Green–Stokes Theorem on the domain $D_\epsilon = B(0, R) \setminus B(0, \epsilon)$, a spherical shell:

\[ \int_{D_\epsilon} u \Delta v - v \Delta u = \int_{\partial D_\epsilon} u \frac{dv}{dn} - v \frac{du}{dn}. \]

Set $v = G(|x|) - G(R)$ in the above. Then $v = 0$ on $\partial B(0, R)$ and $\Delta u = \Delta v = 0$ on $D_\epsilon$. We get

\[ \int_{\partial B(0, R)} u v_r = \int_{\partial B(0, \epsilon)} u v_r - v u_r. \]

But

\[ v_r = \frac{2 - d}{|\partial B(0, 1)|} r^{1-d} \]

(for $d \geq 3$ — and similar in dimensions 1 and 2), so this becomes

\[ \frac{2 - d}{|\partial B(0, 1)|} R^{d-1} \int_{B(0, R)} u = \frac{2 - d}{|\partial B(0, 1)|} e^{d-1} \int_{B(0, \epsilon)} u - \frac{1}{|\partial B(0, 1)|} e^{d-2} \int_{\partial B(0, \epsilon)} u_r. \]

Note that $|\partial B(0, 1)| R^{d-1} = |\partial B(0, R)|$ and $|\partial B(0, 1)| e^{d-1} = |\partial B(0, \epsilon)|$.

We now take the limit as $\epsilon \to 0$. By the Mean Value Theorem, the first term on the right-hand side converges to $u(0)$. The second term is of size $\epsilon$, so it goes to 0. We obtain the desired conclusion. \hfill \Box
2.2.5. Subharmonic and superharmonic functions. The previous identity (being always true) is necessary and sufficient for a function to be harmonic.

Same happens if we replace the equality with an inequality. Indeed, without assuming anything about $-\Delta u$, we obtain instead the more general identity

\[ \oint_{\partial B(x_0, R)} u = u(x_0) - \frac{1}{2 - d} \int_{B(x_0, R)} v \Delta u, \]

where, importantly, $v \geq 0$ inside the domain for $d \geq 3$.

So if $-\Delta u \geq 0$ then

\[ u(x_0) \geq \oint_{\partial B(x_0, R)} u. \]

This condition, properly interpreted, is actually equivalent to $-\Delta u \geq 0$, if it holds for any ball contained in the domain.

If $-\Delta u \leq 0$, then

\[ u(x_0) \leq \oint_{\partial B(x_0, R)} u. \]

Definition 36. Let $u : D \subset \mathbb{R}^d \to \mathbb{R}$. If $-\Delta u \geq 0$, then $u$ is called superharmonic, while if $-\Delta u \leq 0$ then $u$ is called subharmonic.

Clearly, $u$ is subharmonic if and only if $-u$ is superharmonic.

Things work out quite similarly in dimensions $d = 1$ and $d = 2$. In one dimension, harmonic functions reduce to linear functions: $u'' = 0$ means that $u$ is linear. Likewise, in one dimension the previous identity just says

\[ \frac{1}{2} [u(x + R) + u(x - R)] = u(0) + \frac{1}{2} \int_{x-R}^{x+R} (R - |y - x|)u''(y) dy. \]

So, in one dimension, subharmonic functions are same as convex and superharmonic functions are same as concave.

However, in higher dimensions these are different notions: convex functions are always subharmonic, but subharmonic functions need not always be convex etc..

Subharmonic and superharmonic functions arise naturally, see below, so are interesting in their own right.

These concepts and the related maximum and comparison principles can also be used to define viscosity solutions for elliptic and parabolic PDEs.

Example: Suppose $u$ is harmonic and real-valued. Then

\[ -\Delta (u^2) = -2u\Delta u - 2(\nabla u)^2 \leq 0. \]

So $u^2$ is subharmonic.

Example: Suppose $u$ is harmonic and real-valued. Then

\[ |u(x_0)| = \int_{\partial B(x_0, R)} u \leq \int_{\partial B(x_0, R)} |u|. \]

So $|u|$ is subharmonic.

Example: Suppose $f$ is a holomorphic function. Then

\[ -\Delta |f| = -\Delta (\sqrt{f}) \sqrt{f} = (-\Delta \sqrt{f})\sqrt{f} + \sqrt{f}(-\Delta \sqrt{f}) - 2\nabla \sqrt{f} \cdot \nabla \sqrt{f} = -\frac{1}{2} \frac{|
abla f|^2}{|f|} \leq 0. \]

Indeed, since $f$ is holomorphic, $\sqrt{f}$ is also holomorphic, hence harmonic (and so is its conjugate).

So $|f|$ is subharmonic.

Problem 55. Prove that if $u$ is subharmonic then $e^u$ is subharmonic.
Problem 56. Prove that if $f$ and $g$ are subharmonic then $\max(f, g)$ is also subharmonic. Generalize this to arbitrarily many subharmonic functions.

2.2.6. The maximum principle. The maximum principle is a general principle, true for solutions of elliptic equations and in a modified form also for solutions of parabolic equations. It states that such solutions can only reach their maximum and/or minimum values on the boundary.

There are two versions: the weak and the strong maximum principles.

Theorem 2.7 (Weak maximum principle). Let $D$ be a bounded domain and $u : D \cup \partial D \to \mathbb{R}$ be a continuous subharmonic function, i.e. $-\Delta u \leq 0$. Then

$$\max_{x \in D \cup \partial D} u(x) = \max_{x \in \partial D} u(x).$$

Another way of stating (42) is that for every $x_0 \in D$

$$u(x_0) \leq \max_{x \in \partial D} u(x).$$

Intuitively, this statement is implied by the inequality (41): since the value of $f$ at a point is at most the average of values at surrounding points, a strict maximum cannot be reached inside the domain, but only on the boundary.

If the domain is not bounded, then we must also take into account the behavior of $u$ at infinity. There are several ways to prove this formally. Here is one of them.

Proof. Consider $x_0 \in D$ strictly inside the domain. The proof becomes simpler if $-\Delta u < 0$ (it’s not allowed to be zero). So, instead of $u$, for $\epsilon > 0$ let’s consider $u_\epsilon$ given by

$$u_\epsilon(x) = u(x) + \epsilon \|x - x_0\|^2.$$

Then $u_\epsilon$ is still continuous and

$$-\Delta u_\epsilon = -\Delta u - 2d \epsilon \leq -2d \epsilon < 0.$$

The inequality (41) is replaced by

$$u_\epsilon(x) < \frac{1}{2B(x, R)} \int_{\partial B(x, R)} u_\epsilon.$$

The function $u_\epsilon$, being continuous, must reach its maximum somewhere on $D \cup \partial D$. However, the maximum cannot be inside of $D$.

Indeed, suppose the maximum is reached at some point $x \in D$ such that $B(x, R) \subset D$; but this is made impossible by the inequality (44). If the average on $\partial B(x, R)$ is strictly bigger, then at some point there the value of $u$ must be strictly bigger.

So the maximum cannot be reached inside the domain, so it is reached on the boundary. Therefore, for all $x \in D$ and for $x_0$ in particular,

$$u_\epsilon(x_0) \leq \max_{x \in D \cup \partial D} u_\epsilon = \max_{x \in \partial D} u_\epsilon.$$

By the definition of $u_\epsilon$,

$$u(x_0) \leq \max_{x \in D} \{u + \epsilon \|x - x_0\|^2\} \leq \max_{x \in \partial D} u + \epsilon \text{diam}^2(D),$$

where $\text{diam}(D)$ is the diameter of $D$, the greatest distance between any two points of $D$, and is finite because $D$ is bounded.

This inequality is true for any $\epsilon > 0$, so by letting $\epsilon$ go to 0 we get

$$u(x_0) \leq \max_{x \in \partial D} u.$$

This is what we wanted to prove. □
The strong maximum principle also specifies what happens at points where the maximum is attained. The weak maximum principle is just the inequality (43).

**Theorem 2.8.**

If \( u \) is superharmonic, \(-\Delta u \geq 0\), by the same means we obtain that

\[
\min_{x \in \partial D} u(x) = \min_{x \in \partial D} u(x).
\]

Separately, each half of the maximum principle holds for sub- and superharmonic functions. For a harmonic function, both are true: for any \( x_0 \in \partial D \),

\[
\min_{x \in \partial D} u(x) \leq u(x_0) \leq \max_{x \in \partial D} u(x).
\]

So both the minimum and the maximum are attained on the boundary.

The maximum principle is also related to the maximum modulus principle for holomorphic functions, as follows: if \( f : D \to \mathbb{C} \) is holomorphic, then \( |f| \) is subharmonic, so for any \( z_0 \in \partial D \)

\[
|f(z_0)| \leq \sup_{z_0 \in \partial D} |f(z)|.
\]

So subharmonicity can be used to prove the maximum modulus principle.

The maximum principle can also be used to compare solutions.

**Proposition 2.9.** Consider two \( C^2 \) solutions of Laplace’s equation on the same domain, \( u_1, u_2 : D \to \mathbb{R}, \ u_i|_{\partial D} = f_i, -\Delta u_i = g_i \). If \( f_1 \leq f_2 \) and \( g_1 \leq g_2 \), then \( u_1 \leq u_2 \).

**Proof.** The difference \( v = u_1 - u_2 \) is subharmonic and \( v \leq 0 \) on the boundary, so by the maximum principle \( v \leq 0 \) on the whole domain. \( \square \)

In particular, this implies that any subharmonic function is less than the harmonic function, which is less than any superharmonic function — for the same boundary data and on the same domain. This is where the names subharmonic and superharmonic come from.

This approach can be used to study the uniqueness of solutions.

**Proposition 2.10.** Consider two \( C^2 \) solutions of Laplace’s equation on the same domain, \( u_1, u_2 : D \to \mathbb{R}, \ u_i|_{\partial D} = f_i, -\Delta u_i = g_i \). Then \( u_1 = u_2 \).

**Proof.** Using the comparison criterion above, we get both that \( u_1 \leq u_2 \) and that \( u_2 \leq u_1 \). \( \square \)

The maximum principle is true not only for the Laplacian, but also for general elliptic operators

\[
Lu = -\sum a_{ij}(x) \partial_{ij}u + \sum b_i(x) \partial_iu + c(x)u.
\]

This requires some conditions on the coefficients, the main ones being uniform ellipticity and that \( c(x) \geq 0 \).

**Definition 37.** The function \( u : D \to \mathbb{R} \) is called \( L \)-subharmonic if \( Lu \leq 0 \), \( L \)-superharmonic if \( Lu \leq 0 \), and \( L \)-harmonic if \( Lu = 0 \).

**Theorem 2.11 (Weak maximum principle for general elliptic operators).** Let \( L \) be a uniformly elliptic operator (30) on the bounded domain \( D \subset \mathbb{R}^d \), i.e. such that for some \( k > 0 \) and every \( x \in D \)

\[
\sum a_{ij}(x)\xi_i\xi_j \geq k\|\xi\|^2.
\]

Also assume that \( c(x) \geq 0 \) and that \( b(x) \) is bounded. Let \( u : D \cup \partial D \to \mathbb{R} \) be a \( C^2 \) \( L \)-subharmonic function, i.e. such that \( Lu \leq 0 \).

Then

\[
\max_{x \in D \cup \partial D} u(x) \leq \max_{x \in \partial D} u(x).
\]
Proof. Again, the conclusion is easier to prove if the inequality is strict. We begin by constructing a bounded function \( v \) such that \( Lv \) has constant sign. Let
\[
v(x) = e^{C_1|x|^2} - C_2.
\]
This is a bounded function on \( D \) if \( D \) is bounded.

Due to the uniform ellipticity, \( a_{ii} \geq k \), \( 1 \leq i \leq d \), so \( \sum a_{ij} \partial_i \partial_j v \leq C_1 \|x\|^2 \), and so
\[
\sum a_{ij} \partial_i \partial_j v \leq C_1 |x|^2 \left( 2C_1 \sum a_{ii} + 2C_1 \|b\||x| \right).
\]
For sufficiently large \( C_1 \) this is uniformly negative on the whole domain. Also, by choosing some large \( C_2 \), \( v \) becomes negative, so the last term \( c(x)v \) is also negative.

Thus, we have constructed a continuous and bounded function \( v \), \( |v| \leq M \), such that \( Lv \leq 0 \).

Let
\[
u_\epsilon(x) = u(x) + \epsilon v(x).
\]
Then
\[
Lu_\epsilon = Lu + \epsilon Lv < 0.
\]

Suppose the inequality (45) fails for \( u_\epsilon \), i.e.
\[
\max_{x \in \partial D} u_\epsilon(x) > \max_{x \in \partial D} (0, \max u_\epsilon(x)).
\]
Since \( u_\epsilon \) is continuous, it must reach its maximum somewhere on the bounded domain \( D \cup \partial D \).

By the previous inequality, the maximum must be positive and can only be reached in the interior of \( D \).

Suppose the maximum is reached at some interior point \( x_0 \). Then \( u_\epsilon(x_0) > 0 \), \( \nabla u_\epsilon(x_0) = 0 \), and the second-order derivatives in all directions at \( x_0 \) are negative or zero.

By diagonalizing the second-order coefficient matrix \( A(x_0) \), we then immediately get that
\[
\sum a_{ij} \partial_i \partial_j u_\epsilon = \sum \lambda_i \partial_i^2 u_\epsilon \leq 0.
\]
Thus, all the terms in the expression of \( Lu_\epsilon(x_0) \) are positive or zero, so
\[
Lu_\epsilon(x_0) \geq 0.
\]

This is a contradiction, so in fact the inequality (45) must be true for each \( u_\epsilon, \epsilon > 0 \). Going back to the original function \( u \), this proves that
\[
\max_{x \in \partial D} u(x) \leq \max_{x \in \partial D} u_\epsilon(x) + \epsilon M \leq \max_{x \in \partial D} (0, \max u_\epsilon(x)) + \epsilon M \leq \max_{x \in \partial D} (0, \max u(x)) + 2\epsilon M.
\]

Finally, letting \( \epsilon \) go to zero, this proves (45) for \( u \) itself. \( \square \)

2.3. Energy. Energy is a physically meaningful quantity, which can also be used in the mathematical study of elliptic equations.

For example, consider the divergence form equation
CHAPTER 5

Transport equations

\[ u : \mathbb{R}^2 \to \mathbb{R}, \quad u_t - u_x = 0, \quad u(x, 0) = e^x. \]  Cauchy / initial value problem.

In this equation we see that time and space are interchangeable; not all PDEs have a preferred time direction.

Solution: By the method of characteristics. This is a transport equation, meaning that the solution gets transported along certain trajectories called characteristics. Solving the equation reduces to computing the characteristics (draw picture).

Consider the homogenous equation for \( u \):

\[ u_t + c(x, t) \cdot \nabla_x u = 0. \]  

(46)

If we don’t distinguish one special direction as time, then \( u : \mathbb{R}^d \to \mathbb{R} \),

\[ a(x) \cdot \nabla u(x) = 0. \]

If at least one component of \( a = (a_1, \ldots, a_d) \) is nonzero, we can divide by that and get to the previous form.

A characteristic curve is a curve \( \gamma(s) \) such that any classical solution \( u \in C^1 \) of (46) must be constant along \( \gamma \): \( u(\gamma(s)) = \text{constant} \). Taking a derivative we get

\[
[\nabla u(\gamma(s))] \cdot \gamma'(s) = 0.
\]

However, all we know is that

\[
[\nabla u(\gamma(s))] \cdot a(\gamma(s)) = 0.
\]

The way to ensure this holds for all solutions \( u \) is to set

\[ \gamma'(s) = a(\gamma(s)). \]

Or we could use some multiple of \( a \), but that just amounts to reparametrizing the characteristic curve.

If we distinguish one dimension as per (46), then we can use it as the parameter \( s = t \). The equation of a characteristic curve is still the same:

\[ \gamma'(t) = a(\gamma(t)). \]

Due to the standard ODE theory, for each point there is one characteristic curve passing through it.

The equation is then well-posed if and only if each characteristic crosses the initial line/surface/hypersurface at exactly one point.

Example: For an equation with constant coefficients:

\[ u_t + 2u_x - u_y + u_z = 0. \]

The characteristic curves will have the equation

\[ \gamma'(t) = (2, -1, 1), \]

with solution \( \gamma(t) = \gamma(0) + t(2, -1, 1) \).

Thus solutions must be constant along these lines: \( u(c + t(2, -1, 1), t) = \text{constant} \).
Let’s solve this equation with the initial condition $u(x, y, z, 0) = x^2 + y^2 + z^2$. The solution is

$$u(x, y, z, t) = u(x - 2t, y + t, z - t, 0) = (x - 2t)^2 + (y + t)^2 + (z - t)^2.$$ 

Now with nonconstant coefficients:

$$yu_x - xu_y = 0.$$ 

Characteristic curves $\gamma(s)$ must solve the equation $x' = y, y' = -x$. A basis for the space of solutions is $\{\sin x, \cos x\}$. Hence characteristic curves have the formula

$$\gamma(s) = (A \cos s + B \sin s, -A \sin s + B \cos s).$$

So they are circles centered at the origin, $x^2 + y^2 = A^2 + B^2 = r^2$.

Solutions are functions constant along these circles: $u = u(r)$, in other words. If $u \in C^1$ then it’ll be a classical solution; if $u$ is rougher (but constant along circles centered at the origin), then it’ll only be a weak solution.

In this case we shouldn’t take initial data on a whole line, since it will intersect some characteristic of these circles twice.

Returning to the original example: $u : \mathbb{R}^2 \to \mathbb{R}, u_t - u_x = 0, u(x, 0) = e^x$. Characteristics have the equation $\gamma' = -1$, so are lines $\gamma(t) = \gamma_0 - t$ (in other words, $x + t = \text{constant}$). Along these lines $u(\gamma(t), t) = \text{constant}$. We get

$$u(x, t) = u(x + t, 0) = e^{x+t}.$$ 

If the equation is inhomogenous, then the solution is still transported along the characteristic curves, but need no longer be constant on them. Concrete example:
CHAPTER 6

Guessing

Suppose we somehow guess that the solution to the heat equation

\[ u(x, t) = f(t)e^{-\frac{x^2}{4t}}. \]

Such an “educated guess” is called an ansatz. It can be based on something we know about the equation or about related equations or it can be a pure guess.

Then we can plug it into the equation and get
A more complicated, but still standard, PDE

Consider Burgers’ equation: \( u_t - uu_x = 0, \quad u(x, 0) = u_0(x) \).

The Cauchy–Kovalewskaya method: expand everything into a power series. Only works for a short time/close to the boundary and only for analytic boundary and initial data.
APPENDIX A

Linear maps

A function $f : V \to U$ between two vector spaces $V$ and $U$ over the same field of scalars (usually $\mathbb{R}$ or $\mathbb{C}$) is said to be linear if $f(v_1 + v_2) = f(v_1) + f(v_2)$ and $f(\alpha v) = \alpha f(v)$ for any vectors $v_1, v_2$, and $v \in V$ and scalar $\alpha$.

Suppose the vector spaces $V$ and $U$ are finite-dimensional, $\dim V = m$, $\dim U = n$. After fixing bases for the vector spaces $V$ and $U$, $f$ can be represented by a matrix $F \in M_{m \times n}$ ($m$ lines and $n$ columns).

Matrices with real entries.

Important matrix classes: symmetric, orthogonal.

A matrix $A$ with real entries is called symmetric if $A = A^T$.

A matrix $A$ is called orthogonal if $AA^T = I$.

Orthogonal matrices correspond to rotations.

Problem 57. Let $L$ be a homogenous constant-coefficient second-order differential operator on $\mathbb{R}^2$, i.e.

$$L = a \frac{\partial^2}{\partial x^2} + b \frac{\partial^2}{\partial x \partial y} + c \frac{\partial^2}{\partial y^2}.$$  

Assuming that $L$ commutes with all rotations, prove that $a = c$ and $b = 0$.

In other words, you are to prove that multiples of the Laplacian are the only homogenous second-order differential operators on $\mathbb{R}^2$ invariant under all rotations.
APPENDIX B

Test functions and cutoff functions

In any dimension $d$, the function given by
\[ g(x) = e^{-\frac{1}{\rho^2|x|^2}}, \ |x| < \rho, \ 0, \ |x| \geq \rho \]
is a test function, $g \in \mathcal{D}(\mathbb{R}^d)$, supported on $B(0, \rho)$.

**Definition 38.** A (smooth) cutoff function is a smooth function which is 1 on some specified set and 0 outside some small neighborhood of that set.

Here is a construction for a cutoff function for the interval $[a, b]$.

**Fact:** There exists a function $\chi : \mathbb{R} \to \mathbb{R}$, $\chi \in C^\infty$, such that $\chi(x) = 1$ for $x \in [a, b]$, $\chi(x) = 0$ for $x \in (-\infty, a - \epsilon] \cup [b + \epsilon, \infty)$, and $0 \leq \chi \leq 1$.

**Proof:** To make this more concrete, let’s construct a cutoff function for the interval $[1, 9]$, which vanishes outside the interval $[-2, 12]$, i.e. $\epsilon = 1$. The proof works in the same way in general, after adjusting the numbers.

Let $h(x) = 1$ for $0 < x < 10$ and zero otherwise and let $G = g * h$. Here the convolution product is given by the formula
\[ G(x) = \int_{-\infty}^{\infty} g(y)h(x - y) \, dy. \]

One can check, as we did in class, that this function has all the desired properties, except that $G(x) = constant$ without the constant necessarily being 1.

**Problem 58.** Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = e^{-1/x}$ for $x > 0$ and 0 otherwise. Prove that $f$ is continuous at zero. Prove that $f$ is differentiable at zero and $f'(0) = 0$. Prove that $f'$ is also differentiable at zero and $f''(0) = 0$.

**Problem 59.** For two test functions $f, g \in \mathcal{D}(\mathbb{R})$, define their convolution product as $f * g : \mathbb{R} \to \mathbb{R}$,
\[ (f * g)(x) := \int_{-\infty}^{\infty} f(x - y)g(y) \, dy. \]
a) Prove that $f * g$ is also a test function.
b) Prove that $f * g = g * f$. c) Considering some third test function $h \in \mathcal{D}(\mathbb{R})$, show that $(f * g) * h = f * (g * h)$.

Note: The condition that $f$ and $g$ should be test functions is very strong. Similar results also hold under less stringent conditions.
The Airy function

Here is the detailed proof that the function defined by the integral (23) satisfies a certain ODE, which is often taken as an alternative definition of Airy’s function.

Doing the computations carefully requires the notion of uniform convergence.

**Definition 39.** A sequence of functions \( f_n : I \to \mathbb{C} \) is said to converge uniformly to a limit \( f : I \to \mathbb{C} \) if for every \( \epsilon > 0 \) there exists \( N \) such that whenever \( n \geq N \), \( |f_n - f| \leq \epsilon \).

So the limit is \( f \) and the rate of convergence is uniform over the whole domain \( I \), i.e. at the same overall rate for all points in the domain.

The opposite of uniform convergence is pointwise or non-uniform convergence, for which the sequence of functions converges to its limit in a non-uniform way.

Example: The sequence \( f_n(x) = \frac{1}{n(1+x^n)} \) converges uniformly to 0 on \( \mathbb{R} \).

Example: The sequence \( f_n(x) = x^n \) converges non-uniformly on \([0,1]\) to 0 for \( x < 1 \) and 1 when \( x = 1 \).

**Problem 60.** Let \( f_n : \mathbb{R} \to \mathbb{R}, f_n(x) = \cos(x/n) \). Prove that as \( n \to \infty \), \( f_n(x) \to 1 \), but the convergence is not uniform.

The point of introducing this notion is that uniform convergence allows us to interchange limits, integrals, and derivatives, whereas for non-uniform convergence this fails.

For example

**Theorem 0.12.** Suppose a sequence of functions \( f_n \) converges uniformly to \( f \) and each \( f_n \) is (uniformly) continuous. Then \( f \) must be (uniformly) continuous.

Also, for integrals

**Theorem 0.13.** Suppose a sequence of functions \( f_n : [a,b] \to \mathbb{R} \) converges uniformly to \( f \) and each \( f_n \) is integrable. Then so is \( f \) and

\[
\int_a^b f(x) \, dx = \lim_{n \to \infty} \int_a^b f_n(x) \, dx.
\]

The theorem we’ll use in this section is

**Theorem 0.14.** Suppose a sequence of functions \( f_n \) converges uniformly to \( f \) and \( g_n \) converges uniformly to \( g \). If for each \( n \), \( f_n' = g_n \), then \( f' = g \).

**Problem 61.** Consider the sequences \( f_n, g_n : [0,1] \to \mathbb{R}, f_n(x) = \frac{x^{n+1}}{n+1}, g_n(x) = x^n \). Show that \( f_n \) converges uniformly to 0 and \( g_n \) converges non-uniformly to 0 for \( x < 1 \) and 1 when \( x = 1 \).

We may replace the index \( n \) with other indices, too, such as \( R \) or \( \epsilon \). Then instead of a “sequence” it’s called a “family”, but everything else is the same.

**Proposition 0.15.** The Airy function \( \text{Ai}(x) \) defined by (23) is also a solution of the following ODE:

\[
f''(x) = xf(x).
\]
Since this is a second-order equation (called Airy's or Stokes' equation — another one with the same name), the space of solutions is in fact spanned by two functions, \( \text{Ai}(x) \) and \( \text{Bi}(x) \), called Airy functions of first and second kinds.

In the subsequent integrations by parts,

\[
\frac{\partial}{\partial y} \left[ \frac{y^3}{3} + xy \right] = y^2 + x.
\]

**Proof.** Formally taking two derivatives as a function of \( x \),

\[
\text{Ai}''(x) = -\frac{1}{\pi} \int_0^\infty y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy.
\]

Then

\[
\frac{\partial^2}{\partial x^2} \int_0^R y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy = \int_0^R y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy.
\]

Integrating by parts, we get

\[
\int_0^R y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy = \left[ \frac{\partial}{\partial y} \left( \frac{y^3}{3} + xy \right) \right]_0^R - x \int_0^R \cos \left( \frac{y^3}{3} + xy \right) \, dy = \sin \left( \frac{y^3}{3} + xy \right) \bigg|_0^R - x \int_0^R \cos \left( \frac{y^3}{3} + xy \right) \, dy.
\]

If the sine went to zero as \( R \to \infty \), then we could take the limit and obtain the correct equation.

However, there are two problems with this “reasoning”: obviously, sine doesn’t go to zero as \( R \to \infty \), but keeps oscillating. On the other hand, there is no reason why

\[
\lim_{R \to \infty} \frac{\partial^2}{\partial x^2} \int_0^R \cos \left( \frac{y^3}{3} + xy \right) \, dy = \frac{\partial^2}{\partial x^2} \int_0^\infty \cos \left( \frac{y^3}{3} + xy \right) \, dy
\]

should be true and in fact this statement is false. Therefore, this is a wrong proof, though it points at the correct result.

Now let’s do it more rigorously.

Consider the mollified/regularized versions

\[
\text{Ai}_\epsilon(x) = \frac{1}{\pi} \int_0^\infty e^{-\epsilon y} \cos \left( \frac{y^3}{3} + xy \right) \, dy
\]

and in order to understand these improper integrals also their finite versions

\[
\text{Ai}_\epsilon^R(x) = \frac{1}{\pi} \int_0^R e^{-\epsilon y} \cos \left( \frac{y^3}{3} + xy \right) \, dy.
\]

Then, assuming differentiation works normally on finite intervals,

\[
(\text{Ai}_\epsilon^R)'(x) = -\frac{1}{\pi} \int_0^R e^{-\epsilon y} y \sin \left( \frac{y^3}{3} + xy \right) \, dy
\]

and

\[
(\text{Ai}_\epsilon^R)''(x) = -\frac{1}{\pi} \int_0^R e^{-\epsilon y} y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy.
\]

Fix some \( \epsilon > 0 \). As \( R \to \infty \), the integrals from 0 to \( R \) converge to the corresponding integrals from 0 to \( \infty \) and the difference is bounded by

\[
\left| \int_0^R e^{-\epsilon y} y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy - \int_0^\infty e^{-\epsilon y} y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy \right| \leq \int_0^\infty e^{-\epsilon y} y^2 \cos \left( \frac{y^3}{3} + xy \right) \, dy
\]

\[
= \int_0^\infty e^{-\epsilon y} y^2 \, dy
\]

\[
= e^{-\epsilon R} \left( R^2 / \epsilon + 2R / \epsilon^2 + 2 / \epsilon^3 \right) \to 0
\]
as $R \to \infty$. Same goes for the other two integrals. Since the convergence is uniform (the rate does not depend on $x$), by Theorem 0.14 we get that

$$Ai'(x) = -\frac{1}{\pi} \int_0^\infty e^{-\epsilon y} y \sin \left( \frac{y^3}{3} + xy \right) dy$$

and

$$Ai''(x) = -\frac{1}{\pi} \int_0^\infty e^{-\epsilon y} y^2 \cos \left( \frac{y^3}{3} + xy \right) dy.$$

In this computation it’s essential that $\epsilon > 0$.

So now we can look into what ODE these $Ai_\epsilon(x)$ functions satisfy. Indeed,

$$Ai''(x) = -\frac{1}{\pi} \int_0^\infty e^{-\epsilon y} y^2 \cos \left( \frac{y^3}{3} + xy \right) dy$$

$$= -\frac{1}{\pi} \int_0^\infty e^{-\epsilon y} (y^2 + x) \cos \left( \frac{y^3}{3} + xy \right) dy + x Ai_\epsilon(x)$$

$$= -\frac{1}{\pi} e^{-\epsilon y} \sin \left( \frac{y^3}{3} + xy \right) \bigg|_{y=0}^\infty - \frac{1}{\pi} \int_0^\infty e^{-\epsilon y} \sin \left( \frac{y^3}{3} + xy \right) dy + x Ai_\epsilon(x).$$

The boundary term is now zero (this is what we gained by doing the computation carefully), so schematically we got

$$\text{(47)} \quad Ai''(x) = \text{error}_\epsilon(x) + x Ai_\epsilon(x).$$

So these mollified versions of the Airy functions satisfy some approximate version of Airy’s equation, where the trade-off is that there is an error term.

Next, we take the limit $\epsilon \to 0$. Let’s look at the error term:

$$\int_0^\infty e^{-\epsilon y} \sin \left( \frac{y^3}{3} + xy \right) dy.$$

If we evaluate it in absolute value, then we always get a bound of 1, but we want to prove it goes to zero uniformly.

Assume $x$ belongs to some bounded set, $|x| \leq R_0$ for some fixed $R_0$. We split the domain of integration $[0, \infty)$ into two parts, $[0, 2\sqrt{R_0}]$ and $[2\sqrt{R_0}, \infty)$. On the first part we use a rough bound of $2\epsilon\sqrt{R_0}$ and on the other portion we integrate by parts:

$$\int_{2\sqrt{R_0}}^\infty e^{-\epsilon y} \sin \left( \frac{y^3}{3} + xy \right) dy = \int_{2\sqrt{R_0}}^\infty \frac{e^{-\epsilon y}}{y^2 + x} \frac{d}{dy} \left(- \cos \left( \frac{y^3}{3} + xy \right) \right) dy$$

$$= \int_{2\sqrt{R_0}}^\infty \frac{d}{dy} \frac{e^{-\epsilon y}}{y^2 + x} \cos \left( \frac{y^3}{3} + xy \right) dy.$$

When we evaluate this in absolute value we get at most $\epsilon/R_0$. The reason for using this estimate only from $2R_0^2$ to infinity is that otherwise the denominator $y^2 + x$ may be zero, because $x$ can be negative.

So the error term is overall bounded by $C\epsilon(R_0^{1/2} + R_0^{-1})$ and it goes to zero uniformly for $x$ belonging to the bounded set $|x| \leq R_0$.

Next, using the same sort of integration by parts one can prove that for $|x| \leq R_0$ then $Ai_\epsilon(x) \to Ai(x)$ and $Ai'_\epsilon(x) \to Ai'(x)$ uniformly as $\epsilon \to 0$. Let’s prove the second one, i.e.

$$\int_0^\infty e^{-\epsilon y} y \sin \left( \frac{y^3}{3} + xy \right) dy \to \int_0^\infty y \sin \left( \frac{y^3}{3} + xy \right) dy.$$

Taking the difference, we have to prove that

$$\int_0^\infty (e^{-\epsilon y} - 1)y \sin \left( \frac{y^3}{3} + xy \right) dy \to 0.$$
Again we split the interval into two parts, \([0, R]\) and \([R, \infty)\). On the first interval we use the fact that \(|e^{-\epsilon y} - 1| \leq \epsilon y\), getting a bound of \(\epsilon R^3\). On the second interval we integrate by parts, getting a bound of

\[
\int_R^\infty e^{-\epsilon y} y \sin \left(\frac{y^3}{3} + xy\right) \, dy \leq C \frac{e^{-\epsilon R}}{R^2 - R_0} \leq C \frac{R}{R^2 - R_0}.
\]

For \(R > 2\sqrt{R_0}\) we get an overall bound of \(\epsilon R^3 + \tilde{C}/R\), so by setting \(R = \epsilon^{-1/4}\) we get \(C_1\epsilon^{-1/4}\).

So for \(\epsilon < C_2 R_0^{-2}\) the difference is less than \(C_1 \epsilon^{-1/4}\). This proves uniform convergence for \(x\) in this set.

This argument also works for \(\text{Ai}_\epsilon\), but breaks down when trying to use it for \(\text{Ai}_\epsilon''\). However, going back to the approximate equation (47), we see that the right-hand side converges uniformly to \(x \text{Ai}(x)\). So we have the following uniform limits for \(|x| \leq R_0\):

\[
\text{Ai}_\epsilon \to \text{Ai}, \quad \text{Ai}_\epsilon' \to -\frac{1}{\pi} \int_0^\infty \sin \left(\frac{y^3}{3} + xy\right) \, dy, \quad (\text{Ai}_\epsilon')' \to x \text{Ai}(x).
\]

So, using Theorem 0.14, we get that

\[
\text{Ai}' = -\frac{1}{\pi} \int_0^\infty \sin \left(\frac{y^3}{3} + xy\right) \, dy
\]

and then

\[
(\text{Ai}')' = x \text{Ai},
\]

which is what we wanted to prove. Since \(R_0\) was arbitrary, the conclusion follows. \(\square\)