ANALYSIS IV FOR PHYSICS

JUHAN ARU

SECTION 0

Introduction and motivation

As motivation, let us consider the mathematical description of heat transmission on a homogeneous circular rod: the heat equation.

The heat equation on an interval $[0, 1)$ (describing the rod) is given by describing the evolution of the temperature profile

$$\frac{\partial u(t,x)}{\partial t} = D\Delta u(t,x)$$

together with some initial condition $u_0(x) = u(0, x)$ and the boundary condition $u(t, 0) = u(t, 1)$ for all $t \geq 0$ to express that the ends of the rod are connected. Recall that in the 1D case $\Delta f := \frac{\partial^2 f}{\partial x^2}$ and $D > 0$ is the diffusion coefficient.

The revolutionary idea of Fourier was as follows. He noticed empirically that the heat profile over time shows spatially oscillatory behaviour, and thus also motivated by the solution of the wave equation using waves, he proposed to write any solution using spatially oscillating functions like $f_n(x) = \sin(2\pi nx)$ and $g_n(x) = \cos(2\pi nx)$. More precisely, one could try to find a solution of the form

$$u(t,x) = \sum_{n\geq 1} s_n(t)\sin(2\pi nx) + \sum_{n\geq 0} c_n(t)\cos(2\pi nx).$$

But now notice that $\Delta f_n = -4\pi^2 n^2 f_n$ and thus if we try a solution of the form $u_n(x, t) = f_n(x)s_n(t)$ with f_n as above, we obtain an equation

$$\frac{\partial s_n(t)}{\partial t} = -4D\pi^2 n^2 s_n(t).$$

This is a well-known ODE that is easily solved: $s_n(t) = \exp(-4D\pi^2 n^2 t)s_n(0)$. Similarly for the cos terms we get $c_n(t) = \exp(-4D\pi^2 n^2 t)c_n(0)$.

We conclude that it would make sense to propose a solution of the form

$$u(t,x) = \sum_{n\geq 1} s_n(0)\exp(-D4\pi n^2 t)\sin(2\pi nx) + \sum_{n\geq 0} c_n(0)\exp(-D4\pi n^2 t)\cos(2\pi nx).$$

Notice that the initial condition then translates to the condition:

$$u_0 = \sum_{n\geq 1} s_n(0)\sin(2\pi nx) + \sum_{n\geq 0} c_n(0)\cos(2\pi nx).$$

If we do find such $(s_n(0), c_n(0))_{n\geq 0}$, then we may have found at least one solution to the heat equation on the circular rod.

Now, this may sound very convincing, but on a closer look there are several questions here:

(1) We have infinite sums - do they even converge? When do they converge and in which sense?
(2) For which functions u_0 does the above-given expansion hold? In other words for which initial conditions can we find a solution by this method?
(3) Are such expansions unique? Are the solutions we find unique?

(4) Can one approximate solutions? For example this is relevant when trying to numerically solve the equation. This is a question about convergence - and further, how does the notion of convergence relate to the coefficients s_n, c_n?
(5) More generally, how should one measure closesness of different initial conditions, different solutions?
(6) What happens for non-circular rods, e.g. rods with endpoints in heat-baths? Or in higher dimensions?
(7) What about more non-homogeneous case where D is no longer a constant in space? Or when we replace Δ with more general (linear) operators, including for example also some outside influences?

The aim of this course is to study the right mathematical framework to ask and answer such questions. This will bring us to study function spaces, the Lebesgue integral and spectral theory of linear operators. To see why some of those aspects might enter let us further consider a simplified model.

0.1 A discrete model

To understand what we may hope to achieve, let us consider the same problem of heat diffusion but on a discretised space. For example we think that the rod instead is decomposed of n small containers which can exchange heat between its neighbours.

The temperature profile is now given by $u(x,t) : \{0,1,\dots,n\} \to \mathbb{R}$, with the periodicity condition $u(0,t) = u(n,t)$ for all $t \geq 0$.

The evolution is still given by

$$\frac{\partial u(t,x)}{\partial t} = K\Delta_d u(t,x)$$

together with some initial condition $u_0(x) = u(0,x)$, only instead of the real Laplacian, we have the discrete Laplacian $\Delta_d f(x) := \frac{1}{d_x}\sum_{y\sim x} f(y) - f(x)$, where $y \sim x$ means that y, x are neighbours in the underlying discrete graph and d_x is the number of neighbours of the vertex x. In our concrete case we have a circular graph with n vertices and thus $\Delta_d f(x) := \frac{f(y)+f(z)-2f(x)}{2}$, where y, z denote the neighbouring vertices.

Now notice that the problem is really a system of n coupled ordinary differential equations of second degree and Δ_d is just a linear operator on $\mathbb{R}^n \to \mathbb{R}^n$. So how do we solve it?

Let use the same steps as above but see that they have a very simple and concrete meaning here:

- Notice that each u_t can be seen as a vector in $\mathbb{R}^n$ with coordinates and Δ_d can be seen as a symmetric linear operator on $\mathbb{R}^n$(check it!)
- As such Δ_d can be diagonalized: there is an orthonormal basis $\phi_1, \dots, \phi_n$ and eigenvalues $\lambda_1, \dots, \lambda_n$ such that $\Delta_d \phi_i = \lambda_i \phi_i$. In particular any function $u : \mathbb{R}^n \to \mathbb{R}^n$ can be uniqueley written as $\sum_{i=1}^n c_i \phi_i$.
- But now if we write $u_i(t) := c_i(t)\phi_i$, then again each $c_i(t)$ satisfies now a decoupled ODE

  $$\frac{\partial c_i(t)}{\partial t} = K\lambda_i c_i(t)$$

  and thus has a solution $c_i(0)\exp(K\lambda_i t)$.

- We conclude a solution by finding $c_i(0)$ by determining the unique expansion $u_0 := \sum_{i=1}^{n} c_i(0)\phi_i$.
- Given the uniqueness of the expansion, this solution is also unique.
- And finally, we can easily compare solutions just using for example the Euclidean norm. For example conclude that if the initial conditions are close, then so will be the solutions at all times $t > 0$. We also know that this distance is equivalently measured using the distances between two sets of coefficients $(c_i)_{i=1\dots n}, (\tilde{c}_i)_{i=1\dots n}$ - and here using the Euclidean norm instead of some other norm is important.

Hence in this set-up all works super well and would work equally well as long as we have a symmetric linear operator L instead of Δ_d.

What did we use here?

- We used the fact that $\mathbb{R}^n$ is finite-dimensional and thus there exist basis that give unique expansions for each vector
- We used the fact that Δ_d is linear and symmetric and by the spectral theorem can be diagonalised and we can find a basis of eigenvectors
- We used implicitly the linearity of the equation

None of these facts are clear in our original set-up as the space of functions from $[0, 1]$ to $\mathbb{R}$ is no longer finite-dimensional!

To address those we will have to look at spaces of functions and try to first see which such spaces have a nice structure. For example, which spaces of functions satisfy linearity? Which can be define a norm and talk about orthonormality? For which spaces do we have orthonormal expansions? Looking for such nice properties brings us for example to also introduce the Lebesgue integral to construct nice basis of functions.

After that, having spent some time understanding function spaces, we briefly look at the study of linear operators on such spaces and in particular find some set-ups where there are similar orthonormal decompositions using eigenfunctions. We then put all this together to rigorously explain solving the inhomogeneous heat equation and other similar problems.

But this is already enough of introduction, let us get going!

Section 1

The space of continuous functions

Let us start with maybe the most intuitive of function spaces - the space of continuous functions. This is partly a recap, as you have been working with continuous functions in Analysis I-III, and we are just putting things in a wider context.

To start off the functions will be taking values on closed boxes $D \subseteq \mathbb{R}^n$, i.e. rectangles $[a_1, b_1] \times \cdots \times [a_n, b_n])$ and taking values in $\mathbb{R}$. At the end of the section we will discuss to what extent we can (and may want to) generalize both of these choices. You may safely just suppose $D = [0, 1]$, as no actual extra difficulty comes from going to higher dimensions.

The set of all continuous functions from $D \to \mathbb{R}$ will be denoted by $C(D, \mathbb{R})$:

$$C(D, \mathbb{R}) := \{f : D \to \mathbb{R}, f \text{ continuous}\}.$$

In what follows we will try to understand the structure of this space.

1.1 Vector space structure of $C(D, \mathbb{R})$

The first observation we can make about the space $C(D, \mathbb{R})$ is that it has a linear structure like for example the vector space $(\mathbb{R}^n, +)$: if $f, g \in C(D, \mathbb{R})$, then also the function $h(x) := f(x) + g(x)$ is in $C(D, \mathbb{R})$, as is $\lambda f(x)$ where $\lambda \in \mathbb{R}$.

Let us quickly check this for the first statement: for every $x \in D$, by continuity of f, g we can choose δ_f, δ_g such that if $y \in D, \|x - y\| < \delta_f$ then $|f(x) - f(y)| \leq \frac{\epsilon}{2}$ and if $y \in D, \|x - y\| < \delta_g$, then $|g(x) - g(y)| \leq \frac{\epsilon}{2}$. But this means that if $\|x - y\| < \min(\delta_f, \delta_g)$, we have that $|h(x) - h(y)| < \epsilon$ by the triangle inequality.

Exercise 1.1. *Show that in fact $C(D, \mathbb{R})$ has also multiplicative structure: i.e. if $f, g \in C(D, \mathbb{R})$, then also the product $h(x) := f(x)g(x)$ is in $C(D, \mathbb{R})$. What about the function $\max(f, g)$?*

In fact, the space $C(D, \mathbb{R})$ with addition satisfies all the axioms of a vector space! Indeed, the identity element would be just the constant zero function, the inverse element of f the function $-f$ and all conditions are nicely met, as you can easily and patiently check.

Exercise 1.2. *Recall the axioms of a vector space and verify them in the case of $(C(D, \mathbb{R}), +)$.*

In what follows we will often also call the vector space just $C(D, \mathbb{R})$.

Now we might be also interested in summing infinitely many functions, i.e. looking at sums $\sum_{n \geq 1} f_n$. But in what sense can we talk about it? More generally, given a sequence of $(g_n)_{n \geq 1}$ in which sense can we talk about its convergence and limit?

The first idea might be to define limits pointwise: for each $x \in D$ the sequence $(g_n(x))_{n \geq 1}$ is just a sequence of real numbers and thus we know what its convergence means. Thus we may want to define the convergence of $(g_n)_{n \geq 1}$ as functions to mean the convergence of $(g_n(x))_{n \geq 1}$ for all $x \in D$. This is called pointwise convergence and as you have already seen it suffers a small drawback:

Exercise 1.3. *For D a closed box in $\mathbb{R}^n$, find a sequence of functions in $C(D, \mathbb{R})$ that converges pointwise to a function that is not continuous.*

It is a good idea to start from the case $D = [0, 1]$ (which we discussed in class), but then think how to do it in general.

We will look for other notions of convergence and to do this will introduce a norm on the set $C(D, \mathbb{R})$.

1.2 The uniform norm on $C(D, \mathbb{R})$

Recall that the vector space $\mathbb{R}^n$ comes also with several natural norms that give a notion of length of a vector and give us a way to measure distances between vectors. It comes out that one can also endow $C(D, \mathbb{R})$ with a natural norm.

Definition 1.1 (The supremum (or uniform) norm)**.** *For $f \in C(D, \mathbb{R})$ we define its supremum (or uniform) norm $\|f\|_\infty := \sup_{x \in D} |f(x)|$.*

In this definition we use the fact that D is closed and bounded - otherwise the supremum might not be finite.

Exercise 1.4. *Find an example of D that is not closed or not bounded, and $f \in C(D, \mathbb{R})$ such that $\|f\|_\infty$ as defined above is infinite.*

We called the expression above a norm, but recall that a norm on a vector space has again a precise mathematical definition and its conditions need to be checked:

Proposition 1.2. *$\|f\|_\infty$ indeed defines a norm on the vector space $C(D, \mathbb{R})$.*

Proof. We need to check the conditions for a norm.

(1) $\|f\|_\infty \geq 0$ with equality if and only if f is equal to the constant zero function. This is clear.
(2) $\|\lambda f\|_\infty = |\lambda| \|f\|_\infty$ is also clear.
(3) Finally, we need to check the triangle inequality $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. We have

$$\|f + g\|_\infty = \sup_{x \in D} |f(x) + g(x)| \leq \sup_{x \in D} (|f(x)| + |g(x)|)$$

by the triangle inequality. But now

$$\sup_{x \in D} (|f(x)| + |g(x)|) \leq \sup_{x \in D} |f(x)| + \sup_{x \in D} |g(x)|$$

and we conclude.

□

Thus $(C(D, \mathbb{R}), +, \|\cdot\|_\infty)$ is a normed vector space pretty much like $\mathbb{R}^n$ with any of these norms. This gives us a way to talk about convergence that is much more natural:

Proposition 1.3. *Let $(f_n)_{n \geq 1}$ be a sequence of $C(D, \mathbb{R})$ functions converging to some function $f : D \to \mathbb{R}$ w.r.t. the uniform norm. Then in fact f is continuous.*

This is a restatement of a result from Analysis I which says that pointwise limits of continuous functions are not continuous.

The proof technique is called the 3ϵ or $\epsilon/3$ argument and you have again seen it already in Analysis I. Let us give the proof just to understand what is now different from the earlier situation

Proof. It suffices to show that for every $x \in D$, we can find $\delta > 0$ such that $|f(x)-f(y)| < 3\epsilon$ whenever $\|x - y\| < \delta$.

We can first choose a fixed $n \in \mathbb{N}$ large enough so that $\|f_n - f\|_\infty < \epsilon$, and in particular $|f_n(x) - f(x)| < \epsilon$ for every $x \in D$ by the definition (these are the first two epsilons).

Further, by continuity of f_n we can choose $\delta > 0$ such that for every $y \in D$ with $\|x-y\| < \delta$, we have that $|f_n(x) - f_n(y)| < \epsilon$ (this is the third epsilon). Putting things together using triangle inequality we obtain:

$$|f(x) - f(y)| = |f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)| \leq \\ \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < 3\epsilon. \tag{1.1}$$

□

Notice that for pointwise convergence the first step fails: we can not necessarily choose an n such that $\sup_{x\in D} |f_n(x) - f(x)| < 3\epsilon$.

Thus using this norm the set $C(D, \mathbb{R})$ is also closed under taking convergent sequences. In fact, it is even nicer than that and there are no gaps at all in the space, e.g. the space is complete - a notion you have met for $\mathbb{R}^n$ and that we recall here.

Definition 1.4 (Completeness of a normed space)**.** *A normed space $(X, \|\cdot\|)$ is called complete if every Cauchy sequence $(x_n)_{n\geq 1}$ (i.e. every sequence such that for every $\epsilon > 0$, there is an n_ϵ with $\|x_n - x_m\| \leq \epsilon$ for all $n, m \geq n_\epsilon$) converges to an element $x \in X$.*

Theorem 1.5. *The space $(C(D, \mathbb{R}), +, \|\cdot\|_\infty)$ is a complete normed vector space.*

The idea is to use completeness of $\mathbb{R}$ to define a potential limiting function and then to verify that it really is that function.

Proof. We only need to check the completeness. So let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $C(D, \mathbb{R})$. As for every $x \in D$, $(f_n(x))_{n\geq 1}$ is Cauchy and $\mathbb{R}$ is complete, we know a limit exists and we can denote this limit by $f(x)$. It remains to see that $f_n \to f$ in the uniform norm and that f is continuous. The latter claim follows from the proposition above, so we need to just prove the convergence w.r.t. the uniform norm. This is left as an exercise on the exercise sheet.

□

Remark 1.6. *Mathematicians call any normed vector space that is complete a Banach space. Such spaces are quite important in setting up quantum field theory.*

The completeness of the space has important application, one of them is finding solutions to ODEs via approximation. The tool used there is the Banach contraction mapping theorem that you have already met in Analysis II according to the course sheets and that is just recalled here:

Theorem 1.7 (Banach contraction mapping theorem)**.** *Let $F : C(D, \mathbb{R}) \to C(D, \mathbb{R})$ be contractive w.r.t. the uniform norm: $\|F(f) - F(g)\|_\infty < C\|f - g\|_\infty$ with $0 < C < 1$. Then there is a unique solution to $F(f) = f$ that can be obtained from the limit $\lim_{n\to\infty} F^{(n)}(f_0)$ starting from any $f_0 \in C(D, \mathbb{R})$.*

1.3 Fourier series for continuous functions

The expansion of a function f on $[0,1]$ to a series of the form

$$f(x) = \sum_{n\geq 1} s_n \sin(2\pi nx) + \sum_{n\geq 0} c_n \cos(2\pi nx) \tag{1.2}$$

is called the Fourier expansion or Fourier series. We saw in the introduction that it could be quite useful, but we didn't see any results on its existence / uniqueness. So let us look at this in the context of continuous functions f now.

In fact we will see that these questions resolve themselves very smoothly once we find the "right functional space", but it is instructive to consider the questions already.

The first question is how should we go about finding the coefficients s_n, c_n? There the key is the following lemma.

Lemma 1.8. *The following orthogonality relations hold for integers $m, n \geq 0$:*

1. Cosine-cosine Orthogonality:

$$\int_0^1 \cos(2\pi nx)\cos(2\pi mx)\,dx = \begin{cases} 1, & \text{if } n = m = 0, \\ \frac{1}{2}, & \text{if } n = m \neq 0, \\ 0, & \text{if } n \neq m. \end{cases}$$

2. Sine-Sine Orthogonality:

$$\int_0^1 \sin(2\pi nx)\sin(2\pi mx)\,dx = \begin{cases} 0, & \text{if } n = 0 \text{ or } m = 0, \\ \frac{1}{2}, & \text{if } n = m \neq 0, \\ 0, & \text{if } n \neq m. \end{cases}$$

3. Sine-Cosine Orthogonality:

$$\int_0^1 \sin(2\pi nx)\cos(2\pi mx)\,dx = 0 \quad \forall n, m.$$

Proof. The proof is a simple consequence of trigonometric identities and their integrals and is left for the exercise sheet. □

The consequence of this observation is that if we expect the representation (1.2) to hold in any nice sense, then the coefficients s_n, c_n should be given by:

- Cosine Coefficients c_n:

$$c_n = 2\int_0^1 f(x)\cos(2\pi nx)\,dx, \quad \text{for } n \geq 1.$$

  For the constant term c_0, we have:

$$c_0 = \int_0^1 f(x)\,dx.$$

- Sine Coefficients s_n:

$$s_n = 2\int_0^1 f(x)\sin(2\pi nx)\,dx, \quad \text{for } n \geq 1.$$

Further notice that if we want it to hold at the endpoints, then we better have $f(0) = f(1)$ as this also holds for every function in the series.

Maybe a bit surprisingly both the existence and uniqueness are really not clear even for continuous functions!

Indeed, the understanding of counterexamples has evolved with time. The first observation is as follows

- There exists a continuous function f satisfying $f(0) = f(1)$ whose Fourier series converges pointwise everywhere but does not converge uniformly.

It is not easy to come up with such a function but once given, it is easy to check (probably on the exercise sheet).

A more stunning claim comes from the second half of 19th century from Du Bois-Reymond:

- There exist continuous functions $f \in C([0,1], \mathbb{R})$ with $f(0) = f(1)$ such that the Fourier series diverges at a point $x \in [0,1]$.

This was then extended by several people, including Kolmogorov to show that

- There are continuous functions $f \in C([0,1], \mathbb{R})$ with $f(0) = f(1)$ where the Fourier series diverges at infinitely many or even dense set of points.

Finally, Katznelson showed in 1970s that in fact

- For every continuous function f and every $\epsilon > 0$, there is some continuous function g with $\|g - f\| < \epsilon$ and the Fourier series of g diverges at some point.

This means that these unpleasant functions are really everywhere!

There are two ways out of this. First, one could just try to restrict the set of functions that one is considering. Second, one could try to weaken further the notion of convergence and maybe give up having pointwise convergence. We will mainly concentrate on the second direction, as the first is too restrictive. But to finish this section let us still show how the first direction can give us some nice results:

Proposition 1.9. *Let $f \in C^2([0,1])$ be twice continuously differentiable and satisfying $f(0) = f(1)$ and $f'(0) = f'(1)$. Then its Fourier series*

$$f(x) = \lim_{N \to \infty} \sum_{0 \le n \le N} (s_n \sin(2\pi n x) + c_n \cos(2\pi n x)),$$

converges w.r.t. $\| \cdot \|_\infty$.

Remark 1.10. *In fact the result holds under much less stringent conditions, e.g. when the functions are so-called Holder continuous, i.e. satisfying $|f(x) - f(y)| < |x - y|^a$ for some $a > 0$. Just the proof then needs a bit more care and is out of the scope for us.*

The key ingredient is the following lemma, which we observed when guessing the solution to the heat equation and that really explains why Fourier series are so useful:

Lemma 1.11. *Suppose that $f \in C([0,1], \mathbb{R})$ is k times continuously differentiable and satisfies $f^j(0) = f^j(1)$ for all $j = 0 \dots k-1$ [1] Then there is some $C > 0$ such that for all $n \ge 1$ $|c_n| \le Cn^{-k}$ and $|s_n| \le Cn^{-k}$.*

The full proof is on the exercise sheet, but let's see the case $k = 1$.

[1] Here by $f^j(x)$ we mean the $j-th$ derivative of f at x, the 0−th derivative being the function itself.

- We have by integration by parts that

$$\int_{[0,1]} \sin(2\pi nx) f(x) dx = \frac{1}{2\pi n} \int_{[0,1]} \cos(2\pi nx) f'(x) \leq \frac{1}{2\pi n} \|f'\|_\infty.$$

Let us proceed to the proof of the proposition.

Proof of Proposition 1.9. By the lemma we have that $|c_n|, |s_n| \leq Cn^{-2}$. Hence the Fourier series is Cauchy in the uniform norm. Indeed, if we denote by S_M the partial series

$$S_M(f) = \sum_{n \geq 0}^{N} (s_n \sin(2\pi nx) + c_n \cos(2\pi nx)),$$

then by the triangle inequality for all $M > N$:

$$\|S_M(f) - S_N(f)\|_\infty \leq \sum_{N < n \leq M} (\|s_n \sin(2\pi nx)\|_\infty + \|c_n \cos(2\pi nx)\|_\infty),$$

but $\|\sin(2\pi nx)\|_\infty = \|\cos(2\pi nx)\|_\infty = 1$ and hence we can bound the sum by

$$2C \sum_{N < n \leq M} n^{-2} \leq 2CN^{-1},$$

which goes to 0 as $N \to \infty$. Hence as $C([0,1], \mathbb{R})$ is complete for the uniform norm, we obtain the convergence to some continuous function g.

To conclude the theorem, we still need to argue that $f = g$. To do this we observe first (this is on the exercise sheet) that for all $n \geq 0$

$$\int_0^1 (f - g) \sin(2\pi nx) dx = \int_0^1 (f - g) \cos(2\pi nx) dx = 0.$$

It then follows from the next proposition that $g = f$. □

Proposition 1.12. *Suppose f is a continuous function on $[0,1]$. Then $s_n = 0, c_n = 0$ for all $n \geq 0$ if and only if $f(x) = 0$ for all $x \in [0,1]$.*

In particular, if the Fourier series of a function converges uniformly, then it is equal to the function itself and each function has at most one expansion in Fourier series.

Before proceeding further, you should pause and think why this is not immediate.

In fact proving this proposition giving the means we have is not completely straightforward. We will later see how it becomes slick and swift once we have found the right functional space for the Fourier series, where each function has a unique series converging exactly w.r.t. to the norm of the space.

We will prove here the proposition modulo a key construction, that is given on the example sheet.

Proof. We want to show that $s_n = 0, c_n = 0$ for all $n \geq 0$ gives $f = 0$. We will argue by contradiction and show that if for some $x_0 \in (0,1)$ it holds that $f(x_0) \neq 0$, then there is a contradiction with the hypothesis of the proposition.

The idea is to construct an approximations T_{m,x_0} of the identity, or if you wish an approximations of the Dirac delta function δ_{x_0} using finite sums of sines and cosines and to argue that 1) on the one hand $\int_0^1 f(x) T_{m,x_0} dx \approx f(x_0)$ for m large and 2) on the other hand by

hypothesis $\int_0^1 f(x)T_{m,x_0}dx = 0$ for all $m \geq 1$. The construction is recorded in the following lemma.

Lemma 1.13. *For each $x_0 \in (0,1)$ one can construct a series of functions $T_{m,x_0}(x)$ as a linear sum of $\sin(2\pi xn), \cos(2\pi xm)$ with $m, n \leq N$, i.e. by setting*

$$T_{m,x_0}(x) = \sum_{n \leq N} (a_{n,x_0} \sin(2\pi nx) + b_{n,x_0} \cos(2\pi nx)$$

such that the following points hold.

(1) For every $m \geq 1$, $x \in [0,1]$ we have $T_{m,x_0}(x) \geq 0$
(2) For every $m \geq 1$ we have $\int_0^1 T_{m,x_0}(x) = 1$.
(3) For all $\delta > 0$: $\int_0^1 1_{|x-x_0|>\delta} T_{m,x_0}(x)dx \to 0$ as $N \to \infty$.

Given such a sequence of $(T_{m,x_0})_{m \geq 1}$, we obtain the contradiction as follows.
On the one hand by the hypothesis for all $m \geq 1$ we have

$$\int_0^1 f(x)T_{m,x_0}(x)dx = \sum_{n \leq m} \left(a_{n,x_0} \int_0^1 f(x) \sin(2\pi nx)dx + b_{n,x_0} \int_0^1 f(x) \cos(2\pi nx)dx \right) = 0.$$

On the other hand suppose $f(x_0) \neq 0$, say WLOG $f(x_0) > 0$. Then there is some $\delta > 0$ such that $f(x) > f(x_0)/2$ in some region $[-\delta + x_0, x_0 + \delta]$. Write

$$\int_0^1 f(x)T_{m,x_0}(x)dx = \int_{-\delta+x_0}^{\delta+x_0} f(x)T_{m,x_0}dx + \int_0^1 1_{|x-x_0|>\delta} f(x)T_{m,x_0}dx.$$

We can bound the second term in absolute value by

$$\|f\|_\infty \int_0^1 1_{|x-x_0|>\delta} T_{m,x_0}(x)dx,$$

which goes to zero by the lemma. The first term however can be bounded from below by $f(x_0)/2 \int_{-\delta+x_0}^{\delta+x_0} T_{m,x_0}(x)dx$. Combining the conditions of Lemma, we see that for m large enough this integral is larger than say $1/2$ and thus the whole term is larger than $f(x_0)/4$ for m large enough. And in particular we conclude that $\int_0^1 f(x)T_{m,x_0}(x)dx \neq 0$ for m large enough! This gives a contradiction. But our assumption was that $f(x_0) \neq 0$, so this cannot hold and we conclude the proposition. □

This was in the end not hard, but quite a fiddly proof and moreover also the existence of Fourier series for continuous functions had several delicate points. We would prefer if the existence and uniqueness would be simple consequences of a good set-up, like in the case of $\mathbb{R}^n$. With this in mind, we will go towards larger function spaces.

Section 2

Lebesgue measure and Lebesgue integral on $\mathbb{R}^n$

We will continue our aim of constructing a convenient / appropriate function space for the Fourier expansions. Motivated by the finite-dimensional example, we would want to construct a space of functions with an inner product of the type $\int f(x)g(x)dx$ and then see the Fourier series as an orthogonormal basis of this space.

To do this, we will have to make a detour and renew our understanding of two intimately linked notions: 1) the integral of a function 2) the size / measure of subsets of $\mathbb{R}^n$.

But let us start off by discussing why the Riemann integral does not suffice.

2.1 An issue with the Riemann integral

One way to define the Riemann integral of a function $f : [0, 1] \to \mathbb{R}$ is as follows.

(1) We subdivide $[0, 1]$ into 2^n equal disjoint intervals $D_{n,i} = [i2^{-n}, (i + 1)2^{-n}]$ each of size 2^{-n};
(2) We call a function Riemann integrable if $U_n := 2^{-n} \sum_{i=0}^{2^n-1} \sup_{x\in D_{n,i}} f(x)$ (which is decreasing) and $L_n := 2^{-n} \sum_{i=0}^{2^n-1} \inf_{x\in D_{n,i}} f(x)$ (which is increasing) both converge to the same limit.
(3) We define the Riemann integral of f, that from now on we denote for clarity by $\text{®} \int_0^1 f(x)dx$ to be equal to this limit.

It is easy to see that Riemann integral satisfies some nice properties:

Exercise 2.1. *Show that the Riemann integral satisfies some desirable properties:*

- *All continuous functions on* $[0, 1]$ *are Riemann integrable*
- *Every function* f *that changes value finitely many times is Riemann integrable*
- *Linearity: If* f, g *are Riemann integrable on* $[0, 1]$*, then so is their sum and the integral is equal to the sums.*

However, the Riemann integrability does not behave well under limits or infinite sums. Indeed, consider a enumeration $q_1, q_2, \dots$ of all rational numbers in $[0, 1]$ (can you give a concrete one?) and define $f_n(x) = 1$ if $x \in \{q_1, \dots, q_n\}$ and $f_n(x) = 0$ otherwise. Then each f_n is Riemann-integrable (with $\text{®} \int_0^1 f_n(x) = 0$) by the exercise above, but the limit is not Riemann integrable as in every interval the sup is equal to 1 and inf is equal to 0 and thus $U_n = 1$ for all $n \geq 1$ and $L_n = 0$ for all $n \geq 1$.

We will see how this is remedied with the notion of Lebesgue integral.

2.2 The Lebesgue measure

We start however by revisiting the notion of size / volume / measure of subsets of $\mathbb{R}^n$. This is directly related to the integral as even in the case of Riemann integral, if the set $A \subseteq \mathbb{R}$ is nice enough then $\int_{\mathbb{R}} 1_A(x)dx = \text{size}(A)$. What should nice enough be is one of the main questions.

As said, the Lebesgue measure on $\mathbb{R}$ generalizes the notion of length and assigns each permissible subset of $\mathbb{R}$ a size. More formally, the Lebesgue measure is a function $L : \mathcal{F} \to$

$[0, \infty)$, where $\mathcal{F}$ is some collection of subsets of $\mathbb{R}$ satisfying some collection of properties. What should such natural properties be?

(1) First, in the case of $\mathbb{R}$, we would like the length / measure of each interval $[a, b]$, (a, b), $[a, b)$ or $(a, b]$ to be just $b - a$. In particular each point $\{x\}$ should have length 0
(2) We certainly would want also $L(\emptyset) = 0$ and $L(A) \geq 0$ for all $A \in \mathcal{F}$.
(3) Second, we would like measure to satisfy some additivity properties. For example the size of the union of two disjoint sets should clearly be just the sum of their sizes: i.e. in symbols $L(A_1 \cup A_2) = L(A_1) + L(A_2)$. By induction this should hold for any finite number of disjoint intervals: $L(A_1 \cup \cdots \cup A_n) = L(A_1) + \ldots L(A_n)$.
(4) Further, it might make sense for this additivity to hold also if we have countably many disjoint sets? But attention! We cannot ask it for all infinite unions: indeed, for example $[0, 1]$ can be seen as a disjoint union of all points $\{x\}$ in $[0, 1]$, but the sum of their lengths would be 0 whereas the measure of $[0, 1]$ has to be clearly 1!

Observe that only the first property has something specific to do with $\mathbb{R}$, all the others are of very abstract nature. A big breakthrough by Lebesgue was to understand that combining these properties gives the right mathematical framework to talk of size / measure on any set! This is encapsulated in the following general definition:

Definition 2.1 (Measure space, Borel 1898, Lebesgue 1901-1903)**.** *A measure space is a triple $(\Omega, \mathcal{F}, \mu)$, where*

- *Ω is a set, called the sample space or the universe.*
- *$\mathcal{F}$ is a set of subsets of Ω, satisfying:*
  - *$\emptyset \in \mathcal{F}$;*
  - *if $A \in \mathcal{F}$, then also $A^c \in \mathcal{F}$;*
  - *If $A_1, A_2, \cdots \in \mathcal{F}$, then also $\bigcup_{n \geq 1} A_n \in \mathcal{F}$.*

  $\mathcal{F}$ is called a σ-algebra and any $A \in \mathcal{F}$ is called a measurable set.
- *And finally, we have a function $\mu : \mathcal{F} \to [0, \infty]$ satisfying $\mu(\emptyset) = 0$ and countable additivity for disjoint sets: if $A_1, A_2, \cdots \in \mathcal{F}$ are pairwise disjoint,*

$$\mu(\bigcup_{n \geq 1} A_n) = \sum_{n \geq 1} \mu(A_n).$$

This function μ is called a measure. If $\mu(\Omega) < \infty$, we call μ a finite measure.

Geometrically we interpret:

- Ω as our space of points
- $\mathcal{F}$ as the collection of subsets for which our notion of volume can be defined
- μ our notion of volume: it gives each measurable set its volume.

We can define a measure on any set of points, finite or infinite. Some telling examples are:

Example 2.2 (Counting measure)**.** *On any set Ω one can define the counting measure μ_c: we set $\mathcal{F} := \mathcal{P}(\Omega)$ (the set of all subsets), and $\mu_c(\{\omega\}) := 1$ for any $\omega \in \Omega$. For any finite set E, $\mu_c(E)$ gives its number of elements. If E is infinite, then so is $\mu_c(E)$. In particular, if Ω is an infinite set, then $\mu_c(\Omega) = \infty$, so this is a measure, but not a finite measure.*

Notice that on a space with finite number of points it gives the natural uniform measure - each point is treated in the same way. However, it is not the natural measure of size on say

$\mathbb{R}$ as the size of say $[0,1]$ would be infinite. The natural uniform measure on $[0,1]$ or $\mathbb{R}$ will be called the Lebesgue measure, but its existence is already mathematically non-trivial - we will come to this in a bit.

Example 2.3 (Delta measure)**.** *The (Dirac) delta function that you have seen mentioned in the courses, is actually a measure, not a function and can be defined on any space and for any $\sigma-$algebra that contains points. On any set Ω one can define the Dirac delta measure μ_x at the point x as follows: suppose $\mathcal{F}$ contains points and we set $\mu_x(\{x\}) = 1$ and more generally $\mu_x(F) = 1$ if $x \in F$ and $\mu_x(F) = 0$ otherwise, for every $F \in \mathcal{F}$. In particular, rigorously the delta function is defined as a measure and not a function.*

We will come back to this and its connection to the 'Dirac delta function' you have seen before later on.

Finally, a both nice and important aspect of the framework of measure spaces is that it also gives the mathematical basis for probability theory - this was observed by A. Kolmogorov some 30 years after the introduction of measure spaces! A probability space is a measure space with total mass equal to 1, i.e. $\mu(\Omega) = 1$. In that case we often use the notation of $\mathbb{P}$ for the measure μ. The framework of probability is used for observing / measuring what's going on in the world:

- Ω as the space of all microstates / all possible outcomes; e.g. the states of the atmosphere
- $\mathcal{F}$ is the collection of observable events / outcomes: i.e. subsets of microstates, whose happening or not happening can be observed; for example we can maybe only measure macroscopic parameters like temperature, or the amount of rain over an hour
- The measure $\mathbb{P}$ will assign a number in $[0,1]$, called probability, to each observable event. Those events that surely happen, get probability 1.

Example 2.4. *The probability space for describing a fair coin toss would be*

$$(\{H,T\}, \{\emptyset, \{H\}, \{T\}, \{H,T\}\}, \mathbb{P}),$$

where $\mathbb{P}(\{H\}) = \mathbb{P}(\{F\}) = 1/2$.

The probability space for describing a fair dice would be

$$(\{1,2,3,4,5,6\}, \mathcal{P}(\{1,2,3,4,5,6\}), \mathbb{P}),$$

where we define $\mathbb{P}(F) = |F|/6$. If instead we paint all the faces 1,2,3,4,5 black so they become indistinguishable, we can modify our model by taking $\mathcal{F} = \{\emptyset, \{1,2,3,4,5,6\}, \{1,2,3,4,5\}, \{6\}\}$ and using the probability measure $\widetilde{\mathbb{P}}$ defined only on these subsets, still with the same formula as above.

Exercise 2.2. *Find a measure space to describe two unrelated fair coin tosses. What assumptions are you making in giving the description? Define a sigma-algebra suitable for studying the situation where one can only ask if the two coins have the same side up, or different sides up.*

Finally, as mentioned not all natural measure spaces are simple to define. We already mentioned that the natural "uniform" measure on $[0,1]$ or $\mathbb{R}$ needs some work. But one would actually also want to define natural measures on more complicated structures like the space of all continuous functions - indeed, this gives one way to formalize path-integrals in

quantum mechanics. This was achieved by Wiener in the beginning of 20th century; the similar task for string theory, i.e. defining probability measures over surfaces with different metric structures has been partially resolved only in the recent years.

2.2.1 Basic properties of measure spaces

Before discussing the Lebesgue measure let us play aroud a bit with the notion of a measure space.

First, the following lemma helps to see which other sets would be measurable:

Lemma 2.5 (Constructing more measurable sets)**.** *Consider a set Ω with a σ-algebra $\mathcal{F}$.*

(1) If $A_1, A_2, \ldots, \in \mathcal{F}$, then also $\bigcap_{n\geq 1} A_n \in \mathcal{F}$.
(2) Then also $\Omega \in \mathcal{F}$ and if $A, B \in \mathcal{F}$, then also $A \setminus B \in \mathcal{F}$.
(3) For any $n \geq 1$, if $A_1, \ldots, A_n \in \mathcal{F}$, then also $A_1 \cup \cdots \cup A_n \in \mathcal{F}$ and $A_1 \cap \cdots \cap A_n \in \mathcal{F}$.

Proof of Lemma 2.5. By de Morgan's laws for any sets $(A_i)_{i\in I}$, we have that

$$\bigcap_{i\in I} A_i = (\bigcup_{i\in I} A_i^c)^c.$$

Property (1) follows from this, as if $A_1, A_2, \cdots \in \mathcal{F}$, then by the definition of a σ-algebra also $A_1^c, A_2^c, \cdots \in \mathcal{F}$ and hence

$$(\bigcup_{i\geq 1} A_i^c)^c \in \mathcal{F}.$$

For (3), again by de Morgan laws, it suffices to show that $A_1 \cup \cdots \cup A_n \in \mathcal{F}$. But this follows from the definition of a σ-algebra, as $A_1 \cup \cdots \cup A_n = \bigcup_{i\geq 1} A_i$ with $A_k = \emptyset$ for $k \geq n+1$. Finally, for (2) we can just write $\Omega = \emptyset^c$.

The fact that $A \setminus B \in \mathcal{F}$ is left as an exercise. □

The statements are also very intuitive at least in the context of probability: e.g. the first one says that if we can observe if some events $A_1, A_2, \ldots$ happen, then we can observe if they all happen at once; the second property says that if two events can be observed, then we can always also observe if one of them happened but not the other one.

In a similar vein, the basic conditions on the measure, give rise to several natural properties too:

Proposition 2.6 (Basic properties of a measure and a probability measure)**.** *Consider a measure space $(\Omega, \mathcal{F}, \mu)$. Let $A_1, A_2, \cdots \in \mathcal{F}$. Then*

(1) For any $n \geq 1$, and $A_1, \ldots, A_n$ disjoint, we have finite additivity

$$\mu(A_1) + \cdots + \mu(A_n) = \mu(A_1 \cup \cdots \cup A_n).$$

In particular if $A_1 \subseteq A_2$ then $\mu(A_1) \leq \mu(A_2)$.
(2) If for all $n \geq 1$, we have $A_n \subseteq A_{n+1}$, then as $n \to \infty$, it holds that $\mu(A_n) \to \mu(\bigcup_{k\geq 1} A_k)$.
(3) We have countable subadditivity (also called the union bound): $\mu(\bigcup_{n\geq 1} A_n) \leq \sum_{n\geq 1} \mu(A_n)$.

If in fact $\mu(\Omega)$ is finite (e.g. a probability measure), we further also have the following properties:

(4) For any $A \in \mathcal{F}$, we have that $\mu(A^c) = \mu(\Omega) - \mu(A)$.

(5) If for all $n \geq 1$, we have $A_n \supseteq A_{n+1}$, then as $n \to \infty$, it holds that $\mu(A_n) \to \mu(\bigcap_{k\geq 1} A_k)$.

Again, please do check that all these properties also make sense intuitively!

Proof of Proposition 2.6. Finite additivity follows from countable additivity by taking $A_k = \emptyset$ for $k \geq n+1$.

(2), (3) are left as exercises.

For (4), we just notice that A and A^c are disjoint and $A \cup A^c = \Omega$. Thus by disjoint additivity $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$. Finally, for (5), define $B_n = A_n^c$. Then $\mathbb{P}(A_n) = \mathbb{P}(B_n^c) = 1 - \mathbb{P}(B_n)$. Similarly $\mathbb{P}(\bigcap_{k\geq 1} A_k) = 1 - \mathbb{P}(\bigcup_{k\geq 1} B_k)$. Thus the result follows from (2). The rest is left as an exercise □

2.2.2 The Lebesgue measure

The Lebesgue measure is the right notion uniform measure on the spaces $\mathbb{R}^n$ (or say a unit cube $[0,1]^n$ or a ball). This measure is called uniform because it is isotropic, i.e. it treats all the points in the set equally. More formally, it is up to a multiplicative constant the measure μ such that $\mu(A) = \mu(\lambda + A)$, where A is some measurable set and $\lambda + A := \{a + \lambda : a \in A\}$.

To define the uniform measure on $\mathbb{R}^n$, we should first pick the right $\sigma-$algebra. First, it certainly has to be big enough to contain at least all the boxes. Now, it is an interesting fact that in the standard axiomatization of mathematics [2] one cannot take the $\sigma-$-algebra to be equal to $\mathcal{P}(\mathbb{R}^n)$ - otherwise one runs into contradictions as explained in the non-examinable part of the example sheet. However, there are some $\sigma-$algebras that are big enough to contain all sets we might be interested in and small enough at the same time to create no contradiction.

Definition 2.7 (Borel σ-algebra)**.** *The Borel σ-algebra $\mathcal{F}_B$ on $\mathbb{R}^n$ is defined as the smallest $\sigma-$algebra containing all boxes, i.e. all sets of the form $\Pi_{i=1}^n [a_i, b_i]$ with real numbers $a_i < b_i$.*

This definition hides a claim: the fact that such a smallest $\sigma-$algebra exists. However, it is a simple but not that illuminating exercise to show that an arbitrary intersection of $\sigma-$algebras is a $\sigma-$algebra and thus the smallest has a well-defined meaning. It is maybe more interesting to see what it contains, i.e. what we can measure [3]:

Example 2.8. *The Borel $\sigma-$algebra contains for example all points, i.e. sets of the form $\{x\}$: indeed, we can write*

$$\{x\} = \bigcap_{m\geq 1} (\{x\} + [-m^{-1}, m^{-1}]^n).$$

Exercise 2.3. *Show that the Borel $\sigma-$algebra on $\mathbb{R}^n$ also contains all products of half-lines $\Pi_{i=1}^n (-\infty, a_i]$, all open balls $B(x,r)$ and in fact all open sets of $\mathbb{R}^n$*

The main theorem of this section is then the following result, that we assume without proof:

[2] Meaning that we assume the axiom of choice

[3] It is maybe as interesting to see that there are sets in the power-set of $\mathbb{R}^n$ that do not belong to the Borel $\sigma-$algebra. However, describing them explicitly is not that easy - if interested, see the for fun section on the example sheet.

Theorem 2.9 (Existence and uniqueness of Lebesgue measure)**.** *There is a unique measure* λ *defined on* $(\mathbb{R}^n, \mathcal{F}_B)$ *such that the measure of each box* $\Pi_{i=1}^n[a_i, b_i]$ *is given by* $\Pi_{i=1}^n(b_i - a_i)$*.*

Some other nice properties of the Lebesgue measure follow from this theorem:

- It is translation invariant: for every set $A \in \mathcal{F}_B$, if we denote by $A + b$ the set $\{a + b : a \in A\}$, then the Lebesgue measure λ satisfies $\lambda(A) = \lambda(A + b)$. Indeed, denote by $\widetilde{\lambda}(A) := \lambda(A + b)$. This defines another measure on $(\mathbb{R}^n, \mathcal{F}_B)$ such that $\widetilde{\lambda}$(box) equals the volume of the box. Thus by uniqueness part of the theorem we obtain $\widetilde{\lambda} = \lambda$ and hence $\lambda(A + b) = \lambda(A)$ for all Borel sets A.
- It can be also proved that the Lebesgue measure is rotation invariant: for every set $A \in \mathcal{F}_B$, if we denote by $R(A)$ the set rotated by the rotation matrix R, then the Lebesgue measure λ satisfies $\lambda(A) = \lambda(R(A))$.

Maybe somewhat surprisingly the proof of this natural theorem is not immediate. The problem is the following: it is simple to assign measure to each box, or each finite union of disjoint boxes etc...however, the Borel σ-algebra is much richer than that. Indeed, there are sets in the Borel σ-algebra that one cannot obtain in a finite number of steps by starting with boxes and taking iteratively unions, intersections and complements in any order. Hence the fact that one can assign a measure to all Borel sets in a way that the axioms are satisfied and boxes have the right size is not immediate. Also the statement of uniqueness is non-evident for the same reason - why should equality for all boxes imply it for all Borel sets?

The proof goes beyond the scope of this course, but here is the sketch for one of the possible approaches for those interested (not examinable).

$\star$ *Start of non-examinable section* $\star$

For any rectangle $R = \Pi_{i=1}^n[a_i, b_i]$, let's denote by $|R|$ its natural volume $\Pi_{i=1}^n(b_i - a_i)$.

(1) First, we define for any set $A \subseteq \mathbb{R}^d$ a notion of size called the exterior measure: $m^*(E) := \inf \sum_{i=1}^\infty |R_i|$, where the infimum is over all coverings of the set E using rectangles - this gives a certain approximation of size from above.

Notice that from this definition it is not immediate that even $m^*(R) = |R|$, but that can be argued for both closed and open rectangles. Also, it is important that we allow for countably many rectangles - see exercise sheet.

(2) It comes out that showing all the axioms of the measure for all subsets of $\mathbb{R}^d$ is impossible [4]. So now comes the key idea of choosing a subclass of sets which is large enough to contain Borel sets, but small enough to be able to make everything work: we call a set measurable if for every $\epsilon > 0$, there is some countable collection of rectangles $(R_i)_{i\geq 1}$ such that $E \subseteq \bigcup_{i\geq 1} R_i$ and $m^*(E \triangle (\bigcup_{i\geq 1} R_i)) < \epsilon$. This means that our earlier approximation from above can be chosen to fit well.

(3) It then remains to argue that these sets actually form a $\sigma-$algebra and that all axioms are satisfied for $((\mathbb{R}^d, \mathcal{F}_L, m^*)$. In fact they form a $\sigma-$algebra, called the Lebesgue $\sigma-$algebra $\mathcal{F}_L$, that is even larger than $\mathcal{F}_B$!

This final step doesn't require any big theorems or inputs, but it does require quite a bit of care in setting up the order of the argument. It is then an easy conclusion that $\mathcal{F}_B \subseteq \mathcal{F}_L$, as $\mathcal{F}_B$ can be generated from just rectangles and we can conclude.

[4] as long as one assumes the Axiom of Choice

⋆ *End of non-examinable section* ⋆

Example 2.10. *The Lebesgue measure of a point is zero: indeed for every $\epsilon > 0$, we have that $\lambda(\{x\}) \leq \lambda(\{x\} + [-\epsilon, \epsilon]^n) = (2\epsilon)^n$, which can be made arbitrarily small.*

Hence also the measure of all rational numbers is zero: we have by countably additivity $\mu(\mathbb{Q}) = \sum_{q \in \mathbb{Q}} \mu(\{q\}) = 0$.

Exercise 2.4. *Show that the Lebesgue measure of $\mathbb{R}^n$ is infinite and that the Lebesgue measure of the line segment $[0,1] \times \{0\} \cdots \times \{0\} \subseteq R^n$ is zero.*

Now consider the Lebesgue measure on $\mathbb{R}$. Prove that the measure of irrational numbers contained in $[0, R]$ is equal to R; prove also that the Lebesgue measure of the Cantor set is zero.

2.3 Lebesgue integral

Recall our grand plan was to construct a function space which has a nice inner product of the form $\int f(x)g(x)dx$ and all the nice properties of a function space like linearity, closedness under limits and completeness. With Riemann integral this would never be possible, as we saw it does not behave that well under taking limits. Hence let us see another notion of integral, called the Lebesgue integral. To start off, let's see that defining a measure always gives us a natural space of functions and those will be the candidates for defining the integral for.

2.3.1 Measurable functions

Each measure space comes with a class of natural functions, called measurable functions. These will also form the class of functions for which we aim to define the Lebesgue integral.

We will constrain ourselves to working with functions from $\mathbb{R}^n \to \mathbb{R}$, although the notion of a "measurable" function is quite a bit larger, applying to maps between any two sets together with $\sigma-$algebras; in our case these would be the pairs $(\mathbb{R}^n, \mathcal{F}_B)$ and $(\mathbb{R}, \mathcal{F}_B)$, where in both cases we consider the Borel $\sigma-$algebra.

The simplest measurable functions (on $(\mathbb{R}^n, \mathcal{F}_B)$) are those given by characteristic functions 1_E for some Borel-measurable set $E \in \mathcal{F}_B$, i.e. functions that tell us whether x is in a set - then $1_E(x) = 1$ - or not, in which case $1_E(x) = 0$. Their countable linear combinations are called simple functions:

Definition 2.11 (Simple functions)**.** *Let $E_1, E_2, \ldots$ be disjoint Borel sets in $\mathbb{R}^n$ and $c_1, c_2, \ldots$ real numbers. Then a function of the form $f(x) = \sum_{i \geq 1} c_i 1_{x \in E_i}$ is called a simple function.*

We can then define

Definition 2.12 (Measurable function)**.** *We call a function $f : \mathbb{R}^n \to \mathbb{R}$ measurable if it is a pointwise limit of simple functions.*

This definition is natural, however it is not so easy to work with. So let us start by proving an equivalence with another rather nice definition.

Proposition 2.13. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is measurable if and only if for every $a < b$ the preimage $f^{-1}([a, b)))$ is Borel measurable.*

We will sometimes call this condition the preimage condition.

The proof consists of two lemmas, one for each direction both teaching us something about measurable functions:

Lemma 2.14. *Suppose that the sequence of functions $(f_i)_{i\geq 1}$ from $\mathbb{R}^n$ to $\mathbb{R}$ is such that for every $a < b$ the preimage $f_i^{-1}([a,b))$ is Borel measurable. Suppose also that f_i converge pointwise to f.*

Then f also satisfies the same property, i.e. the preimages $f^{-1}([a,b))$ are Borel measurable.

Lemma 2.15. *Suppose f is such that for every $a < b$ the preimage $f^{-1}([a,b))$ is Borel measurable. Then f is a pointwise limit of simple functions f_n.*

Further, a sequence can be chosen to be pointwise increasing and to converge uniformly.

The proof of proposition follows from these two lemmas.

Proof of Proposition. Lemma 2.15 tells us directly that if f satisfies the preimage condition, then it is measurable.

Let us now show conversely that each measurable function satisfies the preimage condition. Using Lemma 2.14 and the definition of measurable functions it satisfies to show that each simple function satisfies the preimage property.

So, consider a simple function $g = \sum_i c_i 1_{E_i}$ with $c_i \in \mathbb{R}$ and $E_i \in \mathcal{F}_B$. Then $f^{-1}([a,b)) = \cup_{i:c_i\in[a,b)]} E_i$ is a countable union of Borel measurable sets and thus Borel measurable as desired, finishing the proof. □

Let us now prove the two lemmas.

Proof of Lemma 2.14. Our aim is to show that $f^{-1}([a,b))$ is a Borel set and this follows from:

$$f^{-1}([a,b)) = \bigcap_{j\geq 1} \bigcup_{k\geq 1} \bigcup_{n\geq 1} \bigcap_{m\geq n} f_m^{-1}((a-1/j, b-1/k)).$$

The verification of this equality is on the exercise sheet

□

Proof of Lemma 2.15. Consider $f_n : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f_n(x) := 2^{-n}\lfloor 2^n f(x)\rfloor.$$

Each f_n is a simple function as we can write

$$f_n(x) = \sum_{k\in\mathbb{Z}} k2^{-n} 1_{\{f(x)\in[k2^{-n},(k+1)2^{-n})\}}$$

and by assumption the sets $\{f(x) \in [k2^{-n}, (k+1)2^{-n})\}$ are measurable. Further we notice that

$$f_n(x) = 2^{-n}\lfloor 2^n f(x)\rfloor = 2^{-m}2^{m-n}\lfloor 2^{-n} f(x)\rfloor \geq 2^{-m}\lfloor 2^{-m} f(x)\rfloor = f_m(x)$$

proving monotonicity. As also

$$f_n(x) \geq 2^{-n}2^n(f(x) - 2^{-n}) = f(x) - 2^{-n}$$

and thus $\|f(x) - f_n(x)\| \leq 2^{-n}$ and we obtain uniform convergence. □

Several nice properties of the space of measurable functions can be now verified. First, the space of measurable functions again has a linear structure:

Lemma 2.16. *If f, g are measurable, then so are λf for $\lambda \in \mathbb{R}$ and $f + g$.*

Proof. This is on the exercise sheet. □

Second, the space of measurable functions is closed under pointwise limits.

Lemma 2.17. *Let $(f_n)_{n\geq 1}$ be a sequence of measurable functions converging pointwise to a function f. Then f is also measurable.*

Motice that this lemma follows directly from Lemma 2.14 under the equivalence of definitions given by Proposition 2.13. Finally, the space contains all continuous functions.

Lemma 2.18. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous, then f is also measurable.*

Proof. This is also on the exercise sheet

□

To finish this section we remark again that in everything we did above we didn't use at all that the domain of our functions was $\mathbb{R}^n$! We could have equally well worked on any other measure space, laying groundwork for defining integration in a very large generality!

2.3.2 The idea behind Lebesgue integral

Recall that if a function $f : \mathbb{R} \to \mathbb{R}$ is Riemann-integrable then we can calculate its Riemann integral on $[0, 1]$ using the following approximation procedure:

- we subdivide $[0, 1]$ into 2^n equal disjoint intervals D_i each of size 2^{-n};
- we calculate the approximated integral $2^{-n} \sum_{i=0}^{2^n-1} f(i2^{-n})$;
- we take the limit $n \to \infty$.

To calculate the Lebesgue integral (that we will shortly define) for a Lebesgue-integrable function on $[0, 1]$ we can also proceed via an approximation, but rather in the image of the function:

- we take the dyadic approximations from the previous subsection: $f_n := 2^{-n}\lfloor 2^n f(x)\rfloor$;
- we calculate $\sum_{i\in\mathbb{Z}} i2^{-n}\lambda(x \in [0, 1] : f_n(x) = i2^{-n})$;
- and take the limit $n \to \infty$.

So in some sense the difference w.r.t. to the Riemann integral is that we group the values not according to the vicinity in the domain $[0, 1]$, but rather based on the vicinity of the function values. So if you wish, you can think that the Lebesgue integral treats each function in a more personal way, the approximations are based on its behaviour.

Let us now move to the formal definition of the Lebesgue integral, which we do using a slightly wider class of approximations.

2.3.3 Definition of the Lebesgue integral via simple functions

There are several ways to define the Riemann integral.[5] Similarly, there are multiple equivalent approaches to constructing the Lebesgue integral (e.g., Stein–Shakarchi, Kolmogorov–Fomin, and Boccarini all present slightly different versions). Last year we picked a

[5]For instance, one can define it using upper and lower (Darboux) sums with arbitrary partitions or just dyadic ones; or even avoid these altogether and define integrability via convergence of approximating Riemann sums in a suitable sense.

definition via dyadic approximations that is maybe simplest to state and intuitive to grasp; this year we go for a definition that is simplest to work with mathematically.

Although the definition we will give works for measurable functions on any measure space $(\Omega, \mathcal{F}, \mu)$, we will focus on the case $(\Omega, \mathcal{F}, \mu) = (\mathbb{R}^n, \mathcal{F}_B, \lambda)$, i.e., $\mathbb{R}^n$ with its Borel sigma-algebra and Lebesgue measure.

For simple functions, i.e. step functions of the form $f(x) = \sum_i c_i 1_{E_i}$, with E_i are disjoint Borel sets and $c_i \in \mathbb{R}$ the Lebesgue integral is simple to define:

Lemma 2.19 (Lebesgue integral for simple functions)**.** *Let $f(x)$ be a simple function given e.g. by $f(x) = \sum_i c_i 1_{E_i}$. We call f Lebesgue integrable if $\sum_i |c_i| \lambda(E_i) < \infty$ and define its Lebesgue integral by*

$$\int_{\mathbb{R}^n} f(x)\lambda(dx) := \sum_i c_i \lambda(E_i).$$

Further, being integrable and the value of the integral are independent of the chosen representation of f as a simple function.

This is called a lemma and not a definition because of the final part. For example the function $f(x) = 1_{[0,1]}$ could be equivalently written as $f(x) = 1_{[0,1/2)} + 1_{[1/2,1]}$ or even as an infinite sum $f(x) = \sum_i 1_{E_i}$ where $(E_i)_{i \geq 1}$ is any partition of $[0, 1]$ into disjoint Borel sets (can you find one?). Thus, one does need to verify that integrability and the integral do not depend on the choice of the representation. Luckily, this is a simple check.

Proof. Denote by S the set the image of f, i.e. the set $\{f(x) : x \in \mathbb{R}^n\}$. Notice that for a simple function it is always countable.

Then observe that for every $s \in S$, we can define $F_s := \{x : f(x) = s\}$ that depend only on the function f. Further, for any representation $f(x) = \sum_i c_i 1_{E_i}$ we have $F_s = \cup_{i:c_i=s} E_i$ and in particular all F_s are Borel and disjoint for different $s \in S$.

As $\sum_i |c_i| \lambda(E_i) = \sum_{s \in S} |s| \lambda(F_s)$ and $\sum_i c_i \lambda(E_i) = \sum_{s \in S} s\lambda(F_s)$ we conclude that both integrability and the integral are well-defined and independent of the representation. □

Example 2.20. *For example, in contrast to the Riemann integral $f(x) = 1_{\mathbb{Q}}(x)$ is integrable with integral equal to 0. Similarly, and $f(x) = 1_{[0,1]\setminus\mathbb{Q}}(x)$ is integrable with integral equal to 1 - both are themselves simple functions!*

For general measurable functions we will proceed in two steps: first we define the Lebesgue integral for non-negative functions, and then generalise it to all measurable functions by separating into non-negative and positive parts.

Definition 2.21 (Lebesgue integral)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be non-negative and measurable. Then we define*

$$\int_{\mathbb{R}^n} f(x)\, \lambda(dx) := \sup \left\{ \int_{\mathbb{R}^n} g(x)\, \lambda(dx) \;\middle|\; 0 \leq g \leq f,\ g \text{ simple} \right\}.$$

We say that f is integrable if this supremum is finite.

For a general measurable function f, we write $f = f_+ - f_-$, where f_+, f_- are the positive and negative parts of f, given by

$$f_+ = \max(f, 0), \quad f_- = \max(-f, 0).$$

We say that f is integrable if both f_+ and f_- are integrable, and define

$$\int_{\mathbb{R}^n} f(x)\,\lambda(dx) := \int_{\mathbb{R}^n} f_+(x)\,\lambda(dx) - \int_{\mathbb{R}^n} f_-(x)\,\lambda(dx).$$

We will sometimes for the sake of brevity also use the shortcut $\int f\lambda(dx) := \int_{\mathbb{R}^n} f(x)\,\lambda(dx)$ or even $\int f d\lambda$.

Remark 2.22. *One could also try to alternatively define the integral as a limit of integrals of any sequence of uniformly approximating simple functions. This works well when integrating over sets of finite measure (like say $[0,1]$); however, as you see on the example sheet, it would require care when integrating over sets of infinite measure, like $\mathbb{R}$ or $\mathbb{R}^n$.*

Exercise 2.5. *Verify from the definitions that $f(x) = x1_{[0,1]}$ is measurable and integrable. Calculate its integral also from the definition. What about $f(x) = x^{-1}1_{(0,1]}$*

Remark 2.23. *We can further define the integral over any Borel integrable set E, which we denote by $\int_E f(x)\lambda(dx)$ by just considering the integral of $1_E(x)f(x)$, which as a product of measurable functions is nicely measurable.*

Remark 2.24. *We can similarly define an integral over complex-valued functions by just separating the real and imaginary parts, i.e. if $f(x) = r(x) + iq(x)$ we call it integrable if the real functions r, q are and just set $\int f d\lambda = \int r d\lambda + i \int g d\lambda$.*

Whereas it is natural to define the Lebesgue integral via simple functions, as countable collections go well with the measure-theoretic framework, it is technically convenient to observe that one can actually work with simple functions that are given by just finite sums.

Lemma 2.25. *Let us call a simple functions f simple and finite, if it can be written as $f(x) = \sum_{i=1}^n c_i 1_{E_i}(x)$ for some finite disjoint sets $E_1, \dots, E_n$ and some real numbers $c_1, \dots, c_n$. Then we have that*

$$\sup\left\{\int_{\mathbb{R}^n} g(x)\,\lambda(dx) \;\middle|\; 0 \le g \le f,\ g \text{ simple}\right\} = \sup\left\{\int_{\mathbb{R}^n} g(x)\,\lambda(dx) \;\middle|\; 0 \le g \le f,\ g \text{ simple and finite}\right\}$$

and in particular one can equivalently define the Lebesgue integral by just considering simple functions that are given by finite sums.

Proof. It is clear that the LHS is larger than the RHS. So it just remains to show that RHS is at least as big as the LHS. To do this notice that for any non-negative integrable simple function $g(x) = \sum_{i\ge 1} c_i 1_{E_i}(x)$, i.e. a function for which $\sum_{i\ge1} c_i\lambda(E_i) < \infty$, we can associate a simple finite function $g_\epsilon(x) = \sum_{i=1}^{N_\epsilon} c_i 1_{E_i}$, where N_ϵ is chosen such that $\sum_{i>N_\epsilon} c_i\lambda(E_i) < \epsilon$. By definition this guarantees that

$$|\int_{\mathbb{R}^n} g(x)\lambda(dx) - \int_{\mathbb{R}^n} g_\epsilon(x)\lambda(dx)| < \epsilon.$$

Now denote by S_L the supremum on the LHS and by S_R the supremum on the RHS of the equality in the lemma. By definition we can choose g such that $\int_{\mathbb{R}^n} g(x)\lambda(dx) \ge S_L - \epsilon$. But then by construction $\int_{\mathbb{R}^n} g_\epsilon(x)\lambda(dx) \ge S_L - 2\epsilon$ and by taking $\epsilon \to 0$ we see that RHS is also at least as large as the LHS.

We conclude that two suprema agree, and the integral could equivalently have been defined using only finite simple functions. □

2.3.4 Basic properties of the Lebesgue integral

We begin by examining some basic properties of the Lebesgue integral. Notice that several of these natural properties do not hold for the Riemann integral!

Proposition 2.26 (Basic properties of the Lebesgue integral)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be measurable. Then*

(1) if $f \geq 0$ and f is integrable, then $\int f d\lambda \geq 0$
(2) if $|f(x)| \leq C$ for all $x \in \mathbb{R}^n$, then it is integrable over any finite box $[a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n$
(3) if $\lambda(f \neq 0) := \lambda(\{x : f(x) \neq 0\}) = 0$, then f is integrable and $\int f d\lambda = 0$
(4) if $f \geq 0$ and $\int f d\lambda = 0$, then $\lambda(f \neq 0) = 0$.

Notice that even property 2 is rather interesting: it somehow says that for a measurable function only unboundedness can prevent it from being integrable! We have separated the question of "regularity" of the function (carried by the notion of measurability) from that of its size (which governs integrability).

Proof. The first property comes directly from the definition. The others are on the exercise sheet.

□

In particular, applying property (3) to the difference $f - g$ of two measurable f, g we should intuitively obtain:

Corollary 2.27. *Let f, g be two measurable functions such that $\lambda(f \neq g) := \lambda(\{x : f(x) \neq g(x)\}) = 0$. Then f is integrable iff g is integrable and $\int f d\lambda = \int g d\lambda$.*

However, writing down the proof we will see that we are still lacking a tool to show this nicely [6]

Proof. Define $h = g - f$. Then h is measurable and $\lambda(h \neq 0) = 0$. Thus by the proposition h is integrable and $\int h d\lambda = 0$.

We would like to know say that $g = f + h$ and it is integrable because f, h are and further

$$\int g d\lambda = \int (f + h) d\lambda = \int f d\lambda + \int h d\lambda.$$

But here we are already using the linearity of the Lebesgue integral, something we still need to prove and that is stated in the next proposition. □

Proposition 2.28. *Let f, g be Lebesgue integrable functions. We have the following linearity statement. For $a, b \in \mathbb{R}$, then $af + bg$ is integrable and*

$$\int (af + bg) d\lambda = a \int f d\lambda + b \int g d\lambda.$$

It comes out this linearity is not as straightforward to prove as one hopes!

Indeed, it is straight-forward to check that for finite simple functions f, g it holds that $\int (f + g) d\lambda = \int f d\lambda + \int g d\lambda$. To see this set $f = \sum_{i=1}^{n} c_i 1_{E_i}$ and $g = \sum_{j=1}^{m} d_j 1_{F_j}$, then

$$g + f = \sum_{i=0\ldots n} \sum_{j=0\ldots m} (c_i + d_j) 1_{E_i \cap F_j},$$

[6] One can show this particular case also by hand, but it is a bit tiring and not worthwhile.

where we define $c_0 = d_0 := 0$ and $E_0 := \mathbb{R}^n \setminus \cup_{i=1}^n E_i$ and $F_0 = \mathbb{R}^n \setminus \cup_{j=1}^m F_j$. Here we have introduced c_0, d_0 because the function f takes value 0 in the complement of $\cup_{i=1}^N E_i$ and g takes value 0 in the complement of $\cup_{j=1}^M F_j$.

It can be then checked from the definition that $\int (f+g)d\lambda = \int f d\lambda + \int g d\lambda$ for these finite simple functions (on the exercise sheet).

It is also easy to see that for non-negative measurable f, g we have that $\int (f+g)d\lambda \geq \int f d\lambda + \int g d\lambda$. Indeed, whenever $h \geq 0, j \geq 0$ are simple finite functions bounded from above by f, g respectively, the function $k := h + j$ is a simple finite function bounded from above by $f + g$. Thus

$$\sup \left\{ \int_{\mathbb{R}^n} k(x)\, \lambda(dx) \;\middle|\; 0 \leq k \leq f+g,\ k \text{ simple and finite} \right\}$$

is larger than the sum of

$$\sup \left\{ \int_{\mathbb{R}^n} h(x)\, \lambda(dx) \;\middle|\; 0 \leq h \leq f,\ h \text{ simple and finite} \right\}$$

and

$$\sup \left\{ \int_{\mathbb{R}^n} j(x)\, \lambda(dx) \;\middle|\; 0 \leq j \leq g,\ j \text{ simple and finite} \right\}$$

The other inequality, however, requires a few tools. The issue is the following:

- given a simple function k below $f + g$, it is not straight-forward to construct two simple functions, h below f, and j below g with $h + j = k$.

So instead of attempting a direct construction, we will take a detour through some general theorems that will allow us to prove linearity rigorously. In short we will show that under certain conditions, when $f_n \to f$ pointwise, we have $\int f_n d\lambda \to \int f d\lambda$.

More precisely we will use one simple ingredient and one more serious one. The simple ingredient follows directly from the definition of the integral.

Lemma 2.29. *Let $0 \leq g \leq f$ be measurable. Then if f is integrable, so is g and moreover $\int f d\lambda \geq \int g d\lambda$.*

The more substantial theorem is about approximating the integral of f via integrals of approximations f_n.

Theorem 2.30 (Monotone convergence theorem)**.** *Let $0 \leq f_1 \leq f_2 \leq \dots$ be a sequence of integrable functions converging pointwise to some $f = \lim_{n\to\infty} f_n$. Then f is integrable if $\lim_{n\to\infty} \int f_n d\lambda < \infty$ and in this case*

$$\int f d\lambda = \lim_{n\to\infty} \int f_n d\lambda.$$

Before proving this, let us see how linearity follows.

Proof of Proposition 2.28. Let us just prove the more interesting case: that $\int (f+g)d\lambda = \int f d\lambda + \int g d\lambda$. We saw that this linearity holds for simple functions on the exercise sheet. We will now first show using Monotone convergence theorem that it holds for all integrable non-negative functions, and then use the decomposition into positive and negative parts to argue for the general case.

Pick a sequence of simple functions $f_n, g_n \geq 0$ with $f_n \leq f_{n+1}$ and $g_n \leq g_{n+1}$ and $f_n \to f$, $g_n \to g$ pointwise from below. Then also $f_n + g_n \to f + g$ pointwise from below. On the one hand by the linearity of simple functions for all $n \geq 1$

$$\int (f_n + g_n) d\lambda = \int f_n d\lambda + \int g_n d\lambda.$$

On the other hand by the Monotone convergence theorem

$$\int f_n d\lambda \to \int f d\lambda \; ; \; \int g_n d\lambda \to \int g d\lambda.$$

In particular this means that

$$\lim_{n\to\infty} \int (f_n + g_n) d\lambda = \lim_{n\to\infty} \left(\int f_n d\lambda + \int g_n d\lambda \right) = \int f d\lambda + \int g d\lambda$$

is finite and hence by Monotone convergence theorem $f + g$ is integrable and

$$\lim_{n\to\infty} \int (f_n + g_n) d\lambda = \int (f + g) d\lambda.$$

So we obtain that

$$\int (f + g) d\lambda = \int f d\lambda + \int g d\lambda,$$

as desired.

For general integrable f, g let us write $f = f_+ - f_-$, $g = g_+ - g_-$ and $f + g = (f + g)_+ - (f + g)_-$ and recall that all these positive and negative parts are non-negative measurable functions.

Now, notice that $(f+g)_+ \leq f_+ + g_+$ and $(f+g)_- \leq f_- + g_-$ and thus if f, g are integrable then so are $(f + g)_+, (f + g)_-$ by Lemma 2.29 and hence also $f + g$.

Now we can rewrite the pointwise equality

$$(f + g)_+ - (f + g)_- = f_+ - f_- + g_+ - g_-,$$

as

$$(f + g)_+ + f_- + g_- = (f - g)_- + f_+ + g_+.$$

But to this we can apply the first part of the proof on both sides to conclude that

$$\int (f + g)_+ d\lambda + \int f_- d\lambda + \int g_- d\lambda = \int (f + g)_- d\lambda + \int f_+ d\lambda + \int g_+ d\lambda.$$

It now remains to recombine the terms and to use the definition of the integral to see that

$$\int (f + g) d\lambda = \int f d\lambda + \int g d\lambda$$

as desired.

□

2.3.5 Convergence theorems

Let us now look more closely at the statement in the style

- if $f_n \to f$ pointwise, then $\int f_n d\lambda \to \int f d\lambda$.

We will first see some counterexamples, then prove the Monotone convergence theorem and a few other useful convergence results.

The first failure could be that the limiting f is not integrable. Recall that the pointwise limit of measurable functions is measurable. Thus we at least know that f is regular enough to be potentially integrable. It could fail to be integrable because of "size":

Example 2.31. *Consider the functions $(f_n)_{n\geq 1}$ defined on $\mathbb{R}$ by $f_n(x) = 1_{[0,n]} - 1_{[-n,0]}$. These functions are all finite simple functions and their integral is equal to 0. But notice that their pointwise limit $f = 1_{[0,\infty)} - 1_{(-\infty,0]}$ is measurable but not integrable.*

But even if the limiting function is integrable, it's integral is not necessarily equal to the limit of integrals.

Example 2.32. *Consider the functions $(f_n)_{n\geq 1}$ defined on $\mathbb{R}$ by $f_n(x) = n1_{(0,1/n)}$. They are finite simple functions and satisfy $\int f_n d\lambda = 1$ by definition. But notice that $f_n(x)$ converge to the constant 0 function pointwise, as for every $x \in \mathbb{R}$, there is some $n_x \in \mathbb{N}$ such that $f_n(x) = 0$ for all $n \geq n_x$. But the integral of the constant 0 function is just 0 and thus the integrals of f_n do not converge to the integral of their pointwise limit.*

In this example the functions concentrate the mass on a smaller and smaller region, keeping area under the graph equal to 1. Eventually this tiny vertical box somehow moves out of the interval $(0, 1)$ and disappears. But suppose we ask all of the f_n to be bounded?

Example 2.33. *Consider the functions $(f_n)_{n\geq 1}$ defined on $\mathbb{R}$ by $f_n(x) = n^{-1}1_{[0,n]}$. Again they are measurable and bounded, thus integrable with $\int f_n d\lambda = 1$. Also $f_n(x)$ converge to the constant 0 function pointwise too, as for every $x \in \mathbb{R}$, there is some $n_x \in \mathbb{N}$ such that $f_n(x) \leq \epsilon$ for all $n \geq n_x$. But the integral of the constant 0 function is just 0 and thus the integrals of f_n do not converge to the integral of their pointwise limit either.*

In this case the area is kept constant by keeping the box horisontally. But suppose, $\lambda(f_n \neq 0) < C$ for some constant C?

Example 2.34. *Consider the functions $(f_n)_{n\geq 1}$ defined on $\mathbb{R}$ by $f_n(x) = 1_{[n,n+1)}$. Again they are measurable and bounded, thus integrable with $\int f_n d\lambda = 1$. Also $f_n(x)$ converge to the constant 0 function pointwise too, as for every $x \in \mathbb{R}$, there is some $n_x \in \mathbb{N}$ such that $f_n(x) = 0$ for all $n \geq n_x$. But the integral of the constant 0 function is just 0 and thus the integrals of f_n do not converge to the integral of their pointwise limit either.*

Now the functions remain bounded but all of the mass moves away to infinity. In some sense these are the counterexamples to keep in mind and the conditions given in the Monotone convergence theorem rule those cases out. Recall the statement:

Theorem 2.35 (Monotone convergence theorem)**.** *Let $0 \leq f_1 \leq f_2 \leq \ldots$ be a sequence of integrable functions converging pointwise to some $f = \lim_{n\to\infty} f_n$. Then f is integrable if $\lim_{n\to\infty} \int f_n d\lambda < \infty$ and in this case*

$$\int f d\lambda = \lim_{n\to\infty} \int f_n d\lambda.$$

And here is the proof.

Proof of the Monotone convergence theorem, Theorem 2.35. First, as $0 \leq f_n \leq f$, it is clear that

$$\lim_{n\to\infty} \int f_n d\lambda \leq \int f d\lambda.$$

The question is why does the other inequality (and thus also the claim on integrability) hold. So we can now assume $\int f d\lambda < \infty$.

Let us start with the case where the limit $f = \lim_{n\to\infty} f_n$ itself is a simple finite function. In particular it can be represented by $f = \sum_{i=1}^m c_i 1_{E_i}$ for some disjoint Borel sets E_i and distinct $c_i > 0$ and its integral equals $\int f d\lambda = \sum_{i=1}^m c_i \lambda(E_i)$. As the integral is finite by assumption, we have that $\lambda(E_i) < \infty$ for every $i = 1 \dots m$.

For each $\epsilon > 0$ and $n \geq 1$ we can then define the sets $F_n := \{x : f_n(x) \geq (1-\epsilon) f(x)\}$ and further $F_{n,i} = F_n \cap E_i$. These sets are increasing and pointwise convergence of f_n to f guarantees that $\cup_{n\geq 1} F_{n,i} = E_i$ for all $i = 1 \dots m$. But then the properties of the measure λ imply that $\lambda(F_{n,i}) \to \lambda(E_i)$; here we use that $\lambda(E_i) < \infty$. But now by definition $f_n(x) \geq (1-\epsilon) \sum_{i=1}^m c_i 1_{F_{n,i}}$ and thus by Lemma 2.29

$$\int f_n d\lambda \geq (1-\epsilon) \sum_{i=1}^m c_i \lambda(F_{n,i})$$

and we conclude that

$$\lim_{n\to\infty} \int f_n d\lambda \geq (1-\epsilon) \int f d\lambda.$$

As ϵ was arbitrary the claim follows for simple limiting functions f.

But now for general f we can pick any simple function $g \leq f$ and repeat the argument. Even though f_n do not converge to g, they converge to $f \geq g$ and this suffices to conclude that $\cup_{n\geq 1} F_{n,i} = E_i$. Hence we obtain similarly that

$$\lim_{n\to\infty} \int f_n d\lambda \geq \int g d\lambda.$$

As this holds for any simple function g that is bounded above by f, we conclude the claims of the theorem from the definition of the integral: if $\lim_{n\to\infty} \int f_n d\lambda < \infty$, then f is integrable and its integral equals that limit. □

The theorem is usually stated for non-negative functions, but notice that if $f_1 \leq f_2 \dots$ is an increasing sequence of integrable functions, then we could consider $\hat{f}_n := f_n - f_1$ to get a sequence of non-negative functions converging to $f - f_1$ and by linearity also conclude a statement for such functions. Also of course one can obtain a similar statement for increasing functions too (see the exercise sheet).

There is a very useful corollary saying that even if the functions f_i are not increasing, we can still say something about the integral:

Corollary 2.36 (Fatou's lemma)**.** *Let $f_1, f_2, \dots$ be non-negative integrable functions. Then*

$$\liminf_{n\to\infty} \int f_n d\lambda \geq \int \liminf_{n\to\infty} f_n d\lambda,$$

where we allow both sides also to be equal to infinity.

Notice that in particular if the non-negative functions f_n converge, we see the limit of their integrals is larger than the integral of their limit. To remember the direction of the inequality, just think of the example of the travelling interval!

Proof. Observe that $g_n := \inf_{m \geq n} f_m$ is non-negative, non-decreasing and by definition converges to $\liminf_{n\to\infty} f_n$.

Thus by Monotone convergence theorem if $\lim_{n\to\infty} \int g_n d\lambda$ is finite then $\liminf_{n\to\infty} f_n$ is integrable and

$$\int \liminf_{n\to\infty} f_n d\lambda = \lim_{n\to\infty} \int g_n d\lambda.$$

But now $f_m \geq g_n$ for all $m \geq n$ and hence

$$\inf_{m\geq n} \int f_m d\lambda \geq \int g_n d\lambda.$$

We conclude that

$$\liminf_{n\to\infty} \int f_n d\lambda \geq \lim_{n\to\infty} \int g_n d\lambda = \int \liminf_{n\to\infty} f_n d\lambda$$

as desired. □

The other useful convergence theorem deals with an arbitrary convergent sequence f_n of integrable functions. Here the counterexamples from above are avoided by asking there to be a dominating function - an integrable non-negative function g such that $g \geq |f_n|$ for all $n \geq 1$. We will admit this result for the sake of time.

Theorem 2.37 (Dominated convergence theorem)**.** *Let $f_1, f_2, \ldots$ be a sequence of integrable functions converging pointwise to some $f = \lim_{n\to\infty} f_n$. Suppose also that there is some $g \geq 0$ that is integrable and such that $|f_n| \leq g$ for all $n \geq 1$. Then f is integrable and*

$$\int f d\lambda = \lim_{n\to\infty} \int f_n d\lambda.$$

2.3.6 Almost sure / almost everywhere equality

We saw that $\int f d\lambda = \int g d\lambda$ whenever $\lambda(f \neq g) = 0$. This is rather interesting and turns out to be quite important! It says explicitly that the integral is insensitive to changing the function on sets of measure 0 - in particular even modifying the value of the function on any countable set leaves the integral unchanged!

This offers a new perspective on why $\int 1_{\mathbb{Q}} d\lambda = 0$ and motivates the following vocabulary that will be helpful to use.

Definition 2.38 (Almost everywhere / almost sure equality)**.** *We say that two measurable functions f, g are equal almost surely or almost everywhere if $\lambda(f \neq g) := \lambda(\{x : f(x) \neq g(f)\}) = 0$.*

More generally we say that some property or condition $E \in \mathcal{F}_B$ holds almost everywhere (a.e.) or almost surely (a.s.) if $\lambda(E^c) = 0$.[7]

[7]It is called almost everywhere in analysis, almost surely in probability, we will probably not be able to avoid using both simultaneously.

The fact that the integral does not change when we change the function on a zero measure set means that, in the context of integration, when we ask for a property of the function to hold, we should be able to ask it to hold only almost everywhere.

For example, let us look at Lemma 2.29. We should be able to replace $0 \leq g \leq f$ pointwise by an almost everywhere statement: if $\lambda(\{x : g(x) > f(x)\} \cup \{x : g(x) < 0\}) = 0$ and f is integrable, then g is integrable and $\int g d\lambda \leq \int f d\lambda$. This is now a different mathematical statement! But one can easily argue it by just considering $\tilde{f}, \tilde{g}$ defined by setting $\tilde{f}(x) = \tilde{g}(x) := 0$ in the complement of the set $\{x : 0 \leq g(x) \leq f(x)\}$. Then the inequality holds pointwise, one can apply Lemma 2.29 to $\tilde{f}, \tilde{g}$ and conclude using Corollary 2.27.

I will leave it on the example sheet to verify that all the other statements we have seen about the integral hold only under hypothesis given in the form of almost sure properties. In particular all the convergence theorems - Monotone convergence theorem, Fatou's lemma and the Dominated convergence theorem hold under these milder conditions using similar little tricks. [8]

2.3.7 Comparison to the Riemann integral

As an application of convergence theorems let us see how to compare the Lebesgue integral with the Riemann integral.

Recall first that there are many functions that are Lebesgue integrable but not Riemann integrable even on $[0, 1]$.

- The function $1_{\mathbb{Q}\cap[0,1]}$ or say $f + 1_{\mathbb{Q}\cap[0,1]}$ for any Riemann integrable function f
- All Riemann integrable functions on $[0, 1]$ are bounded, but for example the function $f(x) := x^{-1/2} 1_{x\in(0,1]}$ is Lebesgue integrable

However, philosophically all Riemann integrable functions are Lebesgue integrable. This statement would be completely true had we used the Lebesgue σ-algebra instead of the Borel one; in our case we have to add a little extra assumption of measurability, but as mentioned finding non-measurable functions is already not so easy!

Proposition 2.39. *Suppose $f : [a, b] \to \mathbb{R}$ is measurable and Riemann integrable. Then f is also Lebesgue integrable and both integrals agree.* [9]

In particular this means that to calculate the Lebesgue integral of any reasonable function f you can use the same tricks you have learned for the Riemann integral as you are calculating the same numbers! Just a few more functions are Lebesgue integrable and importantly the structure of the space of Lebesgue integrable is much nicer - a point that will be illustrated even better with the next section.

The proposition is not entirely straightforward because the piecewise approximations of f used to define the approximate Riemann integrals converge to f only in the sense that the

[8] For enthusiasts. There is only one little crux that comes rather from measurability - when $f_1, f_2, \ldots$ converge almost everywhere, then their pointwise limit is not necessarily measurable (might not even exist!)...yet there exists always a measurable function $\tilde{f}$ that is an almost everywhere limit of $f_1, f_2, \ldots$. I will leave this confusion for the enthusiasts and the rest can just think that everything works well.

[9] For enthusiasts. In fact one does not need to assume the measurability. It also holds that for each Riemann integrable function f there is a measurable function $\tilde{f}$ such that for some Borel set E of zero measure $\{f \neq \tilde{f}\} \subseteq E$ and such that the Lebesgue integral of $\tilde{f}$ agrees with that of f.

limits of upper and lower Riemann integrals will agree. But to apply any of our convergence theorems we would need to have almost everywhere convergence.

Proof. For simplicity let's take $[a,b] = [0,1]$. We use the fact that for a Riemann-integrable f on $[0,1]$ we know that the upper and lower approximations of the integral over dyadics both converge to the integral.

More precisely, if we set

$$L_n(f) := 2^{-n} \sum_{i=0}^{2^n-1} \inf_{x \in [i2^{-n},(i+1)2^{-n}]} f(x)$$

and

$$U_n(f) := 2^{-n} \sum_{i=0}^{2^n-1} \sup_{x \in [i2^{-n},(i+1)2^{-n}]} f(x)$$

then Riemann integrability implies that L_n is increasing and U_n is decreasing both to the same limit that is equal to the Riemann integral $\int_0^1 f(x)dx$.

But now L_n is also the Lebesgue integral of the simple function

$$g_n(x) := \sum_{i=0}^{2^n-1} 1_{x \in [i2^{-n},(i+1)2^{-n})} \inf_{x \in [i2^{-n},(i+1)2^{-n}]} f(x)$$

and U_n the Lebesgue integral of the simple function

$$h_n(x) := \sum_{i=0}^{2^n-1} 1_{x \in [i2^{-n},(i+1)2^{-n})} \sup_{x \in [i2^{-n},(i+1)2^{-n}]} f(x).$$

Further g_n is pointwise increasing and bounded from above by f, and h_n is pointwise decreasing and bounded from below by f. Thus both converge pointwise to measurable functions denoted f_L, f_U and satisfying $f_U \geq f_L$ pointwise. But now by Monotone Convergence Theorem and its cousin proved on the example sheet we see that

$$\int_{[0,1]} f_L d\lambda = \int_{[0,1]} f_U d\lambda = \int_0^1 f(x)dx$$

and thus $\int_{[0,1]} (f_U - f_L)d\lambda = 0$. As $f_U - f_L$ is non-negative we conclude that $\lambda(f_U \neq f_L) = 0$. But also $f_U \geq f \geq f_L$ pointwise, so in particular $\lambda(f_L \neq f) = 0$, where we use the fact that f is measurable by assumption to see that this event is measurable. Hence also the Lebesgue integral of f agrees with $\int_0^1 f(x)dx$ as desired.

□

2.3.8 Fubini's theorem

To finish off the chapter on Lebesgue integral, we still have to see one key result which is the integration analogue to the following result on infinite double sums:

Proposition 2.40. *Let $(a_{i,j})_{i,j \geq 1}$ be an array of real numbers. If $\sum_{i=1}^n \sum_{j=1}^n |a_{i,j}|$ converges as $n \to \infty$, then for all $i \geq 1$, $\sum_{j=1}^n |a_{i,j}|$ converges and for all $j \geq 1$, $\sum_{i=1}^n |a_{i,j}|$ converges.*

Moreover, then

$$\lim_{n\to\infty}\sum_{i=1}^{n}\sum_{j=1}^{n}a_{i,j}=\lim_{N\to\infty}\sum_{n=1}^{N}\sum_{j\geq 1}a_{n,j}=\lim_{N\to\infty}\sum_{n=1}^{N}\sum_{i\geq 1}a_{i,n}.$$

The result for integrals that we state properly in a bit is very similar - we assume that $f:E_1\times E_2\to\mathbb{R}$ is integrable, and we want to say that for all fixed $x_1\in E_1$, the function $f(x_1,y)$ is integrable over E_2; that the function $x_1\to\int_{E_2}f(x_1,y)d\lambda(y)$ is then integrable over E_1 and that finally by doing these two integrations we obtain the Lebesgue integral of f.

Now there are some extra questions that pose themselves in comparison to sums:

(1) First, why is the function $f(x_1,y):E_2\to\mathbb{R}$ for every fixed x_1 measurable?
(2) Second, why is the integral for every x_1 well-defined?
(3) Third, why is the function $x_1\to\int_{E_2}f(x_1,y)d\lambda(y)$ measurable, integrable?

All these aspects require some thought, although it comes out that in our setting where we work with Borel measurable functions the first one is simple to verify (on the sheet). Both the second and the third property do require thought: whereas in the case of sums the finiteness of the double sum clearly implies the those of all individual sums, in the case of integrals this is actually not the case!

Example 2.41. *Consider the function $f:(-1,1)^2\to\mathbb{R}$ defined by*

$$f(x,y)=1_{\mathbb{Q}\cap(-1,1)}(x)\frac{1}{y}1_{(-1,1)}(y).$$

Then $\lambda(f\neq 0)=0$ and thus f is integrable with integral equal to 0. However for $x\in\mathbb{Q}\cap(-1,1)$, the function $f_x(y):=f(x,y)=\frac{1}{y}1_{(-1,1)}(y)$ is measurable but not integrable.

Theorem 2.42 (Fubini)**.** *Let E_1,E_2 be Borel sets of non-zero measure and suppose that $f:E_1\times E_2\to\mathbb{R}$ is Lebesgue integrable. Then*

- *for almost every $x_1\in E_1$, the function $f(x_1,y)$ is integrable over E_2;*
- *the function $F(x_1):=\int_{E_2}f(x_1,y)d\lambda(y)$ if the integral is finite, and $F(x_1)=0$ otherwise is measurable and integrable over E_1;*
- *$\int_{E_1\times E_2}f(x,y)d\lambda(x,y)=\int_{E_1}F(x_1)d\lambda(x_1)=\int_{E_1}(\int_{E_2}f(x,y)d\lambda(y))d\lambda(x)$;*
- *and the same holds if we swap the order of integration.*

Conversely, if either for almost every x_1 the function $|f(x_1,y)|$ is integrable over E_2 and further the function $F(x_1):=\int_{E_2}|f(x_1,y)|d\lambda(y)$ is integrable over E_2 or the same holds when x_1 and x_2 are swapped, then $f(x,y)$ is integrable over $E_1\times E_2$.

We will not prove this theorem, but we will see several applications very soon. For now let us mention that it can actually even just help to do some calculations.

A simple example of that would be

Example 2.43. *Calculate $\int_{[0,1]\times[0,1]}xyd\lambda(x,y)$. Here we don't really have good means to directly calculate the 2d integral, but using Fubini we can write*

$$\int_{[0,1]\times[0,1]}xyd\lambda(x,y)=\int_{[0,1]}\left(\int_{[0,1]}xyd\lambda(y)\right)d\lambda(x)$$

and now we can easily first calculate $\int_{[0,1]} xyd\lambda y = x/2$ *and then calculate* $\int_{[0,1]} x/2d\lambda(x) = 1/4$ *to obtain* $\int_{[0,1]\times[0,1]} xyd\lambda(x,y) = 1/4$.

Exercise 2.6. *The aim is to calculate* $I = \int_{(0,\infty)} \exp(-x)\frac{\sin^2(x)}{x} d\lambda(x)$. *To do this, we define* $f(x,y) = \exp(-x)\sin(2xy)$ *and use Fubini:*

- *Show that* $f(x,y)$ *is integrable over* $(0,\infty)\times[0,1]$
- *Show that when first integrating* y *over* $[0,1]$ *we obtain exactly* I.
- *On the other hand, calculate explicitly the integral by first integrating over* x. *Integration by parts might be useful.*

SECTION 3

The spaces of integrable and square-integrable functions

The aim of this section is to define the spaces of integrable functions and square-integrable functions on some Borel set E, typically either a box $[0,1]^n \subseteq \mathbb{R}^n$ or $\mathbb{R}^n$ itself. Heuristically, these spaces contain respectively all measurable functions f with either $\int_E |f| d\lambda < \infty$ or $\int_E |f|^2 d\lambda < \infty$. However, to make them into nice Banach spaces we will in fact need to be a bit careful.

3.1 The spaces of p-integrable functions $\mathcal{L}^p(E)$

Let us define two sets of functions mentioned above:

$$L^1(E) := \{f : E \to \mathbb{R} \text{ measurable and } \int_E |f| d\lambda < \infty\}$$

$$L^2(E) := \{f : E \to \mathbb{R} \text{ measurable and } \int_E |f|^2 d\lambda < \infty\}.$$

More generally one can define for any $p \geq 1$

$$L^p(E) := \{f : E \to \mathbb{R} \text{ measurable and } \int_E |f|^p d\lambda < \infty\}.$$

As one can similarly define the sets of functions taking values in $\mathbb{C}$, one sometimes makes the notation more precise by using $L^p(E, \mathbb{R})$ and $L^p(E, \mathbb{C})$.

It is not hard to see that these sets come with linear structure - they are closed under sums or multiplying by a constant.

Lemma 3.1. *Let $p \geq 1$ and f, g be in the set $L^p(E)$. Then also for any $a, b \in \mathbb{R}$, the function $af + bg \in L^p(E)$.*

Proof. For $p = 1$ it follows from pointwise triangle inequality. We have $|f(x) + g(x)| \leq |f(x)| + |g(x)|$ and thus $|f + g|$ is integrable, i.e. in L^1

For $p = 2$ we have $|f(x) + g(x)|^2 = f(x)^2 + g(x)^2 + 2f(x)g(x)$ and we use the pointwise inequality $2f(x)g(x) \leq f(x)^2 + g(x)^2$ to conclude $f(x) + g(x)$ is square-interable.

For general p the result follows from the pointwise inequality, called the generalized means inequality, which says that for all positive a, b and some constant $c_p > 0$ we have $(a + b)^p \leq c_p(a^p + b^p)$.

□

If we want to endow them with a norm, however, we encounter a problem. Indeed, for example the natural norm for the set $L^1(E)$ would be $\|f\|_1 := \int_E |f| d\lambda$. And indeed, it is non-negative and one can check that it satisfies the linear scaling property and the triangle inequality. However, recall Proposition 2.26 - if $\int_E |f| d\lambda = 0$ we only know that $\lambda(\{x \in E : |f(x)| \neq 0\}) = 0$. Thus a priori there are many functions of norm zero!

This might sound unpleasant, but keeping in mind the context of Fourier series and the fact that their limits don't necessarily behave well pointwise, it might also be a blessing!

3.1.1 The mathematical definition of $\mathcal{L}^p(E)$

In order to turn L^p into a Banach space (i.e. a complete normed vector space) we will have to deal with this issue of many 0 norm functions. The simple and natural way to do this is to just consider functions that are almost everywhere equal as equal! Mathematically this means that we define the space of equivalence classes of functions where two functions f, g are considered equivalent when $f = g$ almost everywhere.

Lemma 3.2. *Let $f, g \in L^p(E)$. Then the relation $f \sim g$ if $f = g$ almost everywhere, is an equivalence relation.*

Proof. One has to just recall and verify the properties of an equivalence relation:

- reflexivity: $f \sim f$ is clear
- symmetry: $f \sim g$ if and only if $g \sim f$ is also clear
- transitivity: if $f \sim g$, $g \sim h$, then $f \sim h$ follows from the fact that $\lambda(f \neq h) \leq \lambda(f \neq g) + \lambda(g \neq h)$ by the union bound.

□

Definition 3.3 (The space $\mathcal{L}^p(E)$)**.** *For all $p \geq 1$ we define $\mathcal{L}^p(E)$ as the set of all equivalence classes of $L^p(E)/\sim$. We denote the points in $\mathcal{L}^p(E)$, i.e. equivalence classes of functions by $[f], [g], \ldots$.*

For example the functions $f(x) = 0$ or $f(x) = 1_{\mathbb{Q}}$ belong to the same equivalence class are both representatives of $[0]$. To be crystal clear:

- We write $L^p(E)$ for the set of measurable functions with finite integral $\int_E |f|^p d\lambda < \infty$
- and we write $\mathcal{L}^p(E)$ for the quotient space modulo almost-everywhere equality.

We want to see that $\mathcal{L}^p(E)$ has a nice vector space structure, and in fact can be turned into a normed vector space. But we first need to redefine addition and multiplication by constant.

- We define $[f]+[g] := [f+g]$, i.e. by taking two representatives f, g of the equivalence classes $[f], [g]$, adding them up and then our sum is defined as the equivalence class of this sum. Huh! This may sound a bit complicated but really is the natural thing to do and importantly the result does not depend on the representative s that we choose: if $f = \hat{f}$ a.e. and $g = \hat{g}$ a.e. then also $f + g = \hat{f} + \hat{g}$ a.e.
- Similarly we set $c[f] := [cf]$.

Second, we need to define the norm. We first define

- For any function $f \in L^p(E)$ we set
$$\|f\|_p := \left(\int_E |f|^p d\lambda \right)^{1/p}$$
- And by a slight abuse of notation we set
$$\|[f]\|_p := \|f\|_p,$$
where we have picked some representative f of the equivalence class. It is again clear that the choice of the representative of the equivalence class plays no role in the definition of the norm as if $f = g$ almost everywhere, then also $|f|^p = |g|^p$ almost everywhere.

These two definitions reflect a general principle: although the elements of the spaces $\mathcal{L}^p(E)$ are equivalence classes of functions, we can and will usually work directly with representatives. In fact we will go as far as sometimes to also use f to denote the equivalence class $[f]$. This comes out to be completely harmnless as long as one is careful that our questions are well-defined in the context of equivalence classes.

Yet some care is needed! For instance, it does not make sense to ask for the value of an element of $\mathcal{L}^p(E)$ at a specific point x, because changing the value of a function on a set of measure zero—including just one point—does not change its equivalence class. At first this may seem unnatural, but in fact $\mathcal{L}^p$ spaces are the natural mathematical setting for quantum mechanics, where the state of a particle is described by a wavefunction in $\mathcal{L}^2(\mathbb{R}^n)$. And crucially, no experiment contradicts this modeling choice: we do not—and in practice cannot—observe the value of a wavefunction at a single point.

We are now ready to state the first proposition

Proposition 3.4. *For every $p \geq 1$, for every Borel set $E \subseteq \mathbb{R}^n$, we have that $\mathcal{L}^p(E)$ with addition defined above and the norm given by $\|[f]\|_p$ is a normed vector space.*

Here, the $1/p$-th power is somehow necessary to make the norm scale linearly. The triangle inequality is clear for $p = 1$ from the usual triangle inequality for the absolute value, but needs some work for the general p. As we are only interested in $p = 1, 2$, let us argue it in the case $p = 2$. Again, as above it suffices to argue it for concrete representatives of the equivalence classes.

Lemma 3.5 (Triangle inequality for $p = 2$)**.** *For any $f, g \in L^2(E)$, we have that*

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

Proof. By writing out the definitions and squaring both sides we are left to prove

$$\int_E |f + g|^2 d\lambda \leq \int_E (|f|^2 + |g|^2) d\lambda + 2 \left(\int_E |f|^2 d\lambda \int_E |g|^2 d\lambda \right)^{1/2}.$$

By opening the square on the left hand side this reduces to

$$\int_E fg d\lambda \leq \left(\int_E |f|^2 d\lambda \int_E |g|^2 d\lambda \right)^{1/2}.$$

This is the Cauchy-Schwarz inequality for square-integrable functions. □

Theorem 3.6 (Cauchy-Schwarz inequality)**.** *Let $f, g \in L^2(E)$. Then*

$$\int_E |fg| d\lambda \leq \left(\int_E |f|^2 d\lambda \int_E |g|^2 d\lambda \right)^{1/2}.$$

There are tens of proofs of this inequality, which is the cousin of the same inequality in $\mathbb{R}^n$ and has similar interpretations too - the product of the lengths of two vectors is always larger than the length of their inner product. We will also give a proof in this style.

Proof. By multiplying f, g by a constant we can suppose $\int_E |f|^2 d\lambda = 1$ and $\int_E |g|^2 d\lambda$.

Now for each $x \in E$, we have that $|f(x)g(x)| \leq \frac{1}{2}(f(x)^2 + g(x)^2)$ by the arithmetic-geometric mean inequality. Hence

$$\int_E |fg| d\lambda \leq \frac{1}{2} \int_E (|f|^2 + |g|^2) d\lambda = 1 = \left(\int_E |f|^2 d\lambda \int_E |g|^2 d\lambda \right)^{1/2},$$

where we used the normalization choice in the last equality. □

3.2 The Banach space structure of $\mathcal{L}^p(E)$

As we saw with the space of continuous functions, as soon as we have a norm we can start talking about the limits of the elements in the space and about properties like completeness. The key theorem of this section is the following.

Theorem 3.7. *For every $p \geq 1$, for every Borel set $E \subseteq \mathbb{R}^n$, we have that $(\mathcal{L}^p(E), \|\cdot\|_p)$ is a Banach space.*

Recall that the additional point here was completeness - the fact that every Cauchy sequence converges.

We will not prove this theorem, but let us discuss some ingredients that are interesting and important also on their own. The first obvious question that required already a bit of thought in the case of $C([0,1],\mathbb{R})$ was: given a Cauchy sequence how do we identify the limiting function?

As a first naive thought, one could imagine that maybe being Cauchy w.rt. $\|\cdot\|_1$ will tell us something about convergence pointwise or at least convergence almost everywhere. This is unfortunately not the case.

- First, recall that pointwise convergence does not imply convergence of integrals and in particular does not imply convergence in L^p. For example, the sequence $f_n = n1_{(0,1/n)}$ converges pointwise to $f = 0$ but $\|f_n - f\|_1 = 1$ for all n and thus f_n does not converge to f w.r.t the norm $\|\cdot\|_1$.
- Conversely, it is clear that if f_n converges to f w.r.t. $\|\cdot\|_1$, then we can at most hope for almost everywhere convergence of $f_n(x)$ to $f(x)$. Indeed, if f_n to f w.r.t. $\|\cdot\|_1$, then also f_n to g w.r.t. $\|\cdot\|_1$ for any g that is only almost everywhere equal to f.
- However, what is maybe a bit surprising is that we can find a sequence of functions converging w.r.t. $\|\cdot\|_1$ but converging nowhere pointwise - i.e. a sequence of functions f_n such that $\|f_n - f\|_1 \to 0$ and yet $f_n(x)$ does not converge for any x!

Let us look at an example illustrating this second point.

Example 3.8. *The idea is to construct a 'travelling diminishing interval' that firstly visits each point of the domain infinitely many times, such that the values of $f_n(x)$ keep on fluctuating as n changes, and that secondly gets smaller and smaller to obtain convergence w.r.t. $\|\cdot\|_1$.*

Let us formalise this idea. For $n \in \mathbb{N}$ let k_n be such that $2^{k_n} \leq n < 2^{k_n+1}$. We define on $[0,1]$ a sequence of functions by $f_n(x) := 1_{[2^{-k_n}(n-2^{k_n}),2^{-k_n}(n+1-2^{k_n})]}$. Then f_n are all integrable and observe that $\|f_n\| = 2^{-k_n}$, which goes to zero as $n \to \infty$. On the other hand for every $x \in [0,1]$ there are infinitely many n such that $f_n(x) = 0$ and infinitely many n such that $f_n(x) = 1$ and hence $f_n(x)$ cannot converge.

Next time we will see why $\mathcal{L}^p(E)$ is nevertheless complete.

To understand why $\mathcal{L}^p(E)$ still has changes to be complete despite this failure, let us first look at a very useful lemma, which says that functions that are close in the $\|\cdot\|_1$ norm are still also close pointwise in the sense of measure.

Lemma 3.9 (Markov inequality)**.** *Suppose f is integrable and non-negative. Then $\lambda(\{x : f(x) > c\}) \leq \frac{\int f d\lambda}{c}$.*

In particular, if $f, g \in L^p(E)$ satisfy $\|f-g\|_p \leq \epsilon$, then $\lambda(\{x : |f(x)-g(x)| > c\}) \leq (\epsilon/c)^p$.

Proof. On the exercise sheet □

In particular, if two functions are close in the sense of $\|\cdot\|_1$ they can differ pointwise even macroscopically, but only on a small set.

This helps us prove the following key observation, which says that if a sequence of functions converges sufficiently fast to its limit w.r.t. $\|\cdot\|_p$, then it does converge almost everywhere! The theorem is thereafter a simple exercise that I leave for the very motivated.

Lemma 3.10. *Let $p \geq 1$ and f_n be a sequence in $L^p(E)$ converging to $f \in L^p(E)$ w.r.t $\|\cdot\|_p$ such that $\|f_n - f\|_p \leq \epsilon_n$ with $\sum_{n\geq 1} \epsilon_n^p < \infty$. Then f_n converges to f almost everywhere.*

Remark 3.11. *This is a very close cousin of a result called the Borel-Cantelli lemma in probability theory.*

Proof. We write

$$\{x : f_n(x) \nrightarrow f(x)\} \subseteq \bigcup_{k\geq 1} \bigcap_{m\geq 1} \bigcup_{n\geq m} \{x : |f_n(x) - f(x)| \geq 1/k\}$$

From the union bound

$$\lambda(\bigcup_{k\geq 1} \bigcap_{m\geq 1} \bigcup_{m\geq n} \{x : |f_m(x) - f(x)| \geq 1/k\}) \leq \sum_{k\geq 1} \lambda(\bigcap_{m\geq 1} \bigcup_{n\geq m} \{x : |f_n(x) - f(x)| \geq 1/k\}).$$

Further for each k

$$\lambda(\bigcap_{m\geq 1} \bigcup_{n\geq m} \{x : |f_m(x) - f(x)| \geq 1/k\} \leq \lambda(\bigcup_{n\geq m_k} \{x : |f_n(x) - f(x)| \geq 1/k\}),$$

which we can again by union bound $\sum_{n\geq m_k} \lambda(\{x : |f_n(x) - f(x)| \geq 1/k\})$. By the Lemma 3.9 and the hypothesis we have that for any fixed n, k:

$$\lambda(\{x : |f_n(x) - f(x)| \geq 1/k\}) \leq k^p \epsilon_n^p.$$

Fix $\delta > 0$. As $\sum_{n\geq 1} \epsilon_n^p < \infty$, we can for each k choose m_k such that the tail is as small as we wish:

$$\sum_{n\geq m_k} k^p \epsilon_n^p < \delta k^{-2}.$$

Putting all together this gives us

$$\lambda(\{x : f_n(x) \nrightarrow f(x)\}) \leq \delta \sum_{k\geq 1} k^{-2},$$

which can be made arbitrarily small by taking $\delta > 0$ arbitrarily small. This means that in fact $\lambda(\{x : f_n(x) \nrightarrow f(x)\}) = 0$ and we have proved the lemma. □

3.3 The space of square-integrable functions

So far everything we did was true for $\mathcal{L}^p(E)$ with any $p \geq 1$. But in fact the spaces do not behave equivalently as comparing the following two results, stated informally, shows:

- Kolmogorov, when he was 19 years old, showed that there exists a function $f \in L^1([0,1])$ such that its Fourier series diverges everywhere and in particular does not converge w.r.t. $\|\cdot\|_1$.
- However, for every $p > 1$ and every $f \in L^p([0,1])$, the Fourier series converges almost everywhere and w.r.t. $L^p([0,1])$.

Both are highly non-trivial results, hinting at potential subtle differences between L^p spaces. We will only prove a part of the second result, in particular that for every $f \in L^2([0,1])$ its Fourier series converges w.r.t. L^2.

3.3.1 The scalar product on $\mathcal{L}^2(E)$

Recall the following facts that are common to all $\mathcal{L}^p(E)$ spaces:

- $L^2(E) := \{f : E \to \mathbb{R} : f^2 \text{ integrable}\}$
- $\mathcal{L}^2(E)$ the set of equivalence classes $L^2(E)/\sim$ under $f \sim g$ iff $f = g$ almost everywhere
- With the norm:

$$\|f\|_2 := \sqrt{\int_E |f|^2 d\lambda},$$

  $\mathcal{L}^2(E)$ is a Banach space.

Now, compared to $\mathcal{L}^1(E)$ (or indeed any $\mathcal{L}^p(E)$ with $p \neq 2$), the space $\mathcal{L}^2(E)$ has even more structure. In addition to a norm, one can also define a notion of an inner product / angle, very much similar to the inner product on $\mathbb{R}^n$. This works as follows.

For any $f, g \in L^2(E)$ we set

$$\langle f, g\rangle_2 := \int_E fg d\lambda,$$

[10]. This assumes that fg is integrable, which is clear from $|f(x)g(x)| \leq |f(x)|^2 + |g(x)|^2$ or from Cauchy-Schwartz.

One can verify that on $L^2(E)$ (and also $\mathcal{L}^2(E)$, the definition of $\langle f, g\rangle_2$ above satisfies the axioms of an inner product:

Definition 3.12 (Inner product)**.** *Let V be a vector space. We call $\langle v, w\rangle$ a real inner product if the following conditions hold:*

- *Real-valued: $\langle v, w\rangle \in \mathbb{R}$*
- *Symmetry: $\langle v, w\rangle = \langle w, v\rangle$*
- *Linearity: for all $a, b \in \mathbb{R}$, we have $\langle av + bu, w\rangle = a\langle v, w\rangle + b\langle u, w\rangle$*
- *Non-negativity: $\langle v, v\rangle \geq 0$ with equality if and only if $v = 0$.*

We call it a complex inner product if it is real-valued, conjugate symmetric and linearity holds only for $a, b \in \mathbb{C}$. A vector space V endowed with an inner product is called an inner product space.

[10] For $L^2(E, \mathbb{C})$ the inner product is $\int_E f\bar{g} d\lambda$, where $\bar{g}$ denotes the complex conjugate

In fact each inner product gives rise to a norm.

Exercise 3.1. *Consider an inner product $\langle \cdot, \cdot \rangle$ defined on a vector space V. Then $\|v\| := \sqrt{\langle v, v \rangle}$ defines a norm.*

The key is to notice that the Cauchy-Schwartz inequality holds for every norm that stems from an inner product: $\langle v, w \rangle^2 \leq \langle v, v \rangle \langle w, w \rangle$!

By our definitions for $f \in L^2(E)$ we have that $\|f\|_2^2 = \langle f, f \rangle$, i.e. the inner product gives rise to our previously defined norm.

An important thing to notice is that not every norm can stem from an inner product. Indeed, coming from an inner product forces some algebraic conditions on the norm. For example, the so called parallelogram law has to be always satisfied for a norm compatible with a real inner product:

Lemma 3.13 (Parallelogram law)**.** *Suppose on a vector space the norm $\| \cdot \|$ is compatible with a real inner product $\|f\|^2 = \langle f, f \rangle_2$. Then*

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$$

Proof. This comes from a direct computation using the linearity of $\langle f, f \rangle_2$. □

Now we can test this for $\| \cdot \|_p$ on say $L^2([0,1])$. Consider $v = 1_{[0,1/2]}$ and $w = 1_{[1/2,1]}$. Then we get for the LHS 2 and for RHS $2^{-2/p+2}$, These are equal if and only if $p = 2$! This shows that no other $\| \cdot \|_p$ on $[0,1]$ can stem from an inner product.

3.3.2 Orthonormal bases of $\mathcal{L}^2([0,1])$ and Fourier series

The key result that makes Fourier series work so nicely for functions in $L^2([0,1])$ is the following statement. Informally we could just say:

Theorem 3.14. *The set of functions $(\sqrt{2}\sin(2\pi nx))_{n\geq 1}, (\sqrt{2}\cos(2\pi nx))_{n\geq 1}$ together with the constant function 1 form an orthonormal basis of $L^2([0,1])$.*

This would be philosophically correct, just one would need to just interpret everything under the flag of almost surely. A formally precise statement is:

Theorem 3.15. *The set of equivalence classes $([\sqrt{2}\sin(2\pi nx)])_{n\geq 1}, ([\sqrt{2}\cos(2\pi nx)])_{n\geq 1}$ together with the equivalence class of the constant function $[1]$ form an orthonormal basis of the vector space $(\mathcal{L}^2([0,1]), \langle \cdot, \cdot \rangle_2)$.*

Here, the orthonormal basis for a (potentially infinite-dimensional) inner product space is a direct generalisation from the finite-dimensional case.

Definition 3.16 (Orthonormal basis)**.** *Let $(V, \langle \cdot, \cdot, \rangle)$ be an inner product space. We call $(v_i)_{i\geq 1}$ an orthonormal basis of V if the following two conditions hold*

- *Orthogonality: for all $i, j \geq 1$ we have that$\langle v_i, v_j \rangle = 1_{i=j}$*
- *Basis condition: Each $w \in V$ can be written as $\sum_{i\geq 1} c_i v_i$, where $c_i \in \mathbb{R}$ and the convergence is w.r.t. the norm $\|v\| := \langle v, v \rangle^{1/2}$.*

Orthogonal basis have several useful properties mirroring those of the finite-dimensional setting:

Lemma 3.17. *Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and $(v_i)_{i\geq 1}$ an orthonormal basis. Then*

- *If $w \in V$ has a writing $w = \sum_{i\geq 1} a_i v_i$, then $\|w\|^2 = \sum_{i\geq 1} a_i^2$.*
- *The writing $w = \sum_{i\geq 1} a_i v_i$ is unique and each coefficient is given by $a_i = \langle w, v_i\rangle$.*

Proof. For the first property, let $w_N := \sum_{i=1}^N a_i v_i$. Then

$$\|w_n\|^2 = \langle w_N, w_N\rangle = \langle \sum_{i=1}^N a_i v_i, \sum_{i=1}^N a_i v_i\rangle = \sum_{i\geq 1}^N a_i^2,$$

where in the last equality we used linearity and orthogonality. But now $\|w_n\|^2 \to \|w\|^2$ - indeed,

$$|\|w_n\|^2 - \|w\|^2| = |\|w_n\| - \|w\||(\|w_n\| + \|w\|),$$

which goes to zero as $\|w_n - w\| \to 0$ by assumption.

In particular, the first part directly implies that if $\sum_{i\geq} a_i v_i = 0$ then $a_i = 0$ for all $i \geq 1$. The uniqueness is on the exercise sheet. □

Notice that in particular this means that Theorem 3.15 rather directly implies the following!

Theorem 3.18 (Fourier series in L^2)**.** *Let $f \in L^2([0,1]$. If we define the partial Fourier series by*

$$S_N(f) := \langle f, 1\rangle_2 + \sum_{n=1}^N 2\langle f, \sin(2\pi nx)\rangle_2 \sin(2\pi nx) + \sum_{n=1}^N 2\langle f, \cos(2\pi nx)\rangle_2 \cos(2\pi nx),$$

then $\|f - S_N(f)\|_2 \to 0$ as $N \to \infty$, i.e. the partial Fourier series of f converge to f w.r.t. to the L^2-norm.

Let us now discuss the proof of Theorem 3.15. First, notice that all of $\sin(2\pi nx), \cos(2\pi mx)$ are continuous and bounded on $[0,1]$ and thus in $L^p([0,1])$ for all $p \geq 1$.

Second, orthogonality follows from Lemma 1.8 as firstly it suffices to show it for chosen representatives of the equivalence classes and second, everything is Riemann integrable and we can thus use the computation we did also for the Lebesgue integral.

To make use of this orthogonality one further needs the following, that mirrors the lemma above about complete orthogonal systems. The proof of this is on the exercise sheet

Lemma 3.19. *Let $v_1, v_2, \ldots$ be orthonormal vectors in a complete inner product space V. Then for any $w \in V$, we have that $\widehat{w} := \sum_{i\geq 1}\langle v_i, w\rangle$ is well-defined and satisfies 1) $\|\widehat{w}\| \leq \|w\|$ and 2) $\langle w - \widehat{w}, v_i\rangle = 0$ for all $i \geq 1$.*

What we obtain from this lemma is the following. Let $w \in L^2([0,1])$ and set

$$\widehat{w} := \langle w, 1\rangle + \sum_{n\geq 1} 2\langle w, \sin(2\pi nx)\rangle \sin(2\pi nx) + \sum_{n\geq 1} 2\langle w, \cos(2\pi nx)\rangle \cos(2\pi nx).$$

Then by the lemma above this is well defined and $w - \widehat{w}$ is orthogonal to all the $\sin(2\pi nx), \cos(2\pi nx)$ and the constant function. If we are able to show that in fact $w - \widehat{w}$ is zero almost everywhere, then have proved the completeness of the basis, i.e. the spanning property of the basis. Thus it remains to show the following lemma.

Lemma 3.20. *Suppose $w \in L^2([0,1])$ is orthogonal to all the $\sin(2\pi nx), \cos(2\pi nx)$ and the constant function w.r.t. $\langle \cdot, \cdot \rangle_2$, then w is the zero function.*

Recall that we already proved such a result in case f was a continuous function! So this is a generalisation, but one can again use the Féjer kernel to conclude (we leave it in the non-examinable part of the example sheet).

In fact there are many useful orthonormal bases on $[0,1]$. For example:

Lemma 3.21. *The set of functions $(\sqrt{2}\sin(\pi nx))_{n\geq 1}$ forms an orthonormal basis of $L^2([0,1])$ as does the set of functions $(\sqrt{2}\cos(\pi nx))_{n\geq 1}$ together with the constant function 1.*

This might look a bit odd - we had a basis using both sin and cos functions, and now we say that we can use just one of the two. But observe that we are now using more of them - instead of $2\pi nx$ we now have in the argument πnx.

To see the link between the two clearly let us first observe that by scaling and translation we have that:

Lemma 3.22. *Let $L \in \mathbb{N}$, then an orthonormal basis of $L^2([-L/2, L/2))$ is given by the set of functions: $(\sqrt{\frac{2}{L}}\sin(\frac{2}{L}\pi nx))_{n\geq 1}, (\sqrt{\frac{2}{L}}\cos(\frac{2}{L}\pi nx))_{n\geq 1}$ together with the constant function $\frac{1}{\sqrt{L}}$.*

For the complex $L^2([-L/2, L/2], \mathbb{C})$ they are given in a nice compact form by $(\frac{1}{\sqrt{L}}\exp(2\pi inx/L)_{n\in\mathbb{Z}}$.

Then looking at the functions over $[-1,1]$ gives us the link between the different basis and proves the lemma.

Proof of Lemma 3.21. We have that $(\sin(\pi nx))_{n\geq 1}, (\cos(\pi nx))_{n\geq 1}$ together with the function $1/2$ form an orthonormal basis of $L^2([-1,1])$.

But each $f \in L^2([-1,1])$ can be written as a sum $f = f_o + f_e$ of an odd function $f_o(x) = -f_o(-x)$ and even function $f_e(x) = f_e(-x)$. Concretely, $f_o(x) = \frac{f(x)-f(-x)}{2}$ and $f_e(x) = \frac{f(x)+f(-x)}{2}$.

Now, notice that each even function is orthogonal to an odd function over $L^2([-1,1])$! Moreover, for each n, $\sin(\pi nx)$ is odd and $\cos(\pi nx)$ is even. Thus any even function in $L^2([-1,1])$ can be expanded only using cos and the constant function, and every odd function just using sin.

On the other hand, every $g \in L^2([0,1])$ can be extended to either an even function on $[-1,1]$ by defining $g_e(x)$ defined as $g(x)$ on $[0,1]$ and as $g(-x)$ on $[-1,0]$, or an odd function if we define $g_o(x)$ as $g(x)$ on $[0,1]$ and as $-g(-x)$ on $[-1,0]$. Their expansions in $L^2([-1,1])$ then descend to the sin or cos expansions on $[0,1]$. □

A nice consequence is that we can now fully solve the heat equation on $[0,1]$.

Theorem 3.23. *The heat equation*

$$\frac{\partial u(t,x)}{\partial t} = D\Delta u(t,x)$$

on $[0,1]$ with the initial condition $u(0,x) = u_0(x) \in L^2([0,1])$ has a unique solution when we either fix:

- *Periodic boundary conditions: $u_t(0) = u_t(1)$ and $u_t'(0) = u_t'(1)$*
- *Dirchlet boundary conditions: $u_t(0) = u_t(1) = 0$,*

- *Neumann boundary conditions:* $u_t'(0) = u_t'(1) = 0$.

In each cases the solution is explicit and is given by

$$u(t,x) = \langle u_0, 1\rangle + \sum_{n\geq 1} \exp(-D4\pi^2 n^2 t)\left(\sin(2\pi nx)\langle u_0, 2\sin(2\pi nx)\rangle + \cos(2\pi nx)\langle u_0, 2\cos(2\pi nx)\rangle\right)$$

in the case of periodic conditions. By

$$u(t,x) = \sum_{n\geq 1} \exp(-D\pi^2 n^2 t)\sin(\pi nx)\langle u_0, 2\sin(\pi nx)\rangle$$

in the case of Dirichlet conditions. And by

$$u(t,x) = \langle u_0, 1\rangle + \sum_{n\geq 1} \exp(-D\pi^2 n^2 t)\cos(\pi nx)\langle u_0, 2\cos(\pi nx)\rangle$$

in the case of Neumann boundary conditions.

Proof. As $u_0 \in L^2([0,1])$ we know that any of the given expansions is true at $t = 0$ by the fact that that all three series form ON basis.

Also as $\exp(-D4\pi^2 n^2 t) \leq 1$ or $\exp(-\pi^2 n^2 t) \leq 1$ we also see that in all cases $u_t(x)$ in the statement is in $L^2([0,1])$ for every $t > 0$. It is then easy to check that every proposed solution is twice-differentiable in x and satisfies the equation.

Thus it just remains to argue that the solution is unique. To do this suppose there are two solutions $u(t,x)$ and $v(t,x)$ from the same initial condition and same boundary conditions. Then also $w(t,x) = u(t,x) - v(t,x)$ is a solution starting from the zero function at $t = 0$.

Now observe that $\|u_t(x)\|^2$ is decreasing in time. Indeed, we write

$$\frac{1}{2}\frac{\partial \|u_t(x)\|^2}{\partial t} = \int_{[0,1]} u_t(x)\frac{\partial u(t,x)}{\partial t} d\lambda(x)$$

which using heat equation can be written as

$$D\int_{[0,1]} u_t(x)\Delta u_t(x) = -D\int_{[0,1]} |\nabla u_t(x)|^2 + u_t(1)u_t'(1) - u_t(0)u_t'(0).$$

Using one of the three boundary conditions we conclude that $\frac{1}{2}\frac{\partial \|u_t(x)\|^2}{\partial t} \leq 0$ and thus indeed $\|u_t(x)\|^2$ is decreasing in time.

This is called an energy estimate and here the justification of all the steps will be left for the exercises sheet, but in the first step we just changed the order of integration and derivation and wrote out the time-derivative of $u_t(x)^2$ and thereafter we used integration by parts. This in particular means that if $u_0(x) = 0$ then $u_t(x) = 0$ for all $t > 0$ giving uniqueness. □

Remark 3.24. *In fact one can solve also the heat equation using so called mixed or Robin boundary conditions, we will try to find a more general conceptual framework that also fits those soon enough.*

3.4 Fourier transform

We have now seen how to solve the heat equation on any interval using the Fourier series, that gave us a convenient orthonormal basis. This generalises nicely to boxes of the form $[-L/2, L/2]^n \subseteq \mathbb{R}^n$. But what about $\mathbb{R}$ and $\mathbb{R}^n$?

First notice that when we considered Fourier basis on $[-L/2, L/2]$ then in the complex form they were given by $(\frac{1}{\sqrt{L}} \exp(2\pi inx/L)_{n\in\mathbb{Z}}$. So in some sense as $L \to \infty$ the set of frequencies we consider becomes more and more dense. And indeed, over $\mathbb{R}$ the Fourier series becomes an integral, called the Fourier transform:

Definition 3.25 (Fourier transform)**.** *Let f be a real or complex-valued function $f \in L^1(\mathbb{R})$. We can define the Fourier transform $\mathcal{F}(f)(k) : \mathbb{R} \to \mathbb{C}$ by*

$$\hat{f}(k) := \mathcal{F}(f)(k) := \int_{\mathbb{R}} f(x) \exp(-i2\pi kx) d\lambda(x).$$

This is well defined for each k again as $f(x)$ is integrable and $|f(x)\exp(-i2\pi kx)| = |f(x)|$.

We would hope that the Fourier transform would work as nicely on $L^2(\mathbb{R})$ as the Fourier series on $L^2([0, 1])$, but notice some differences:

- First, there is a question question of integrability. Whereas on $[0, 1$ each square-integrable function is also integrable, its Fourier series is well-defined because $f(x)\exp(-2\pi ikx)$ is integrable for any k. But $L^2(\mathbb{R})$ is not a subset of $L^1(\mathbb{R})$: there are functions in $L^2(\mathbb{R})$ that are not integrable (for example the function $f(x) = (|x|+1)^{-1}$) and hence there is no reason that $f(x)\exp(-2\pi ikx)$ should be integrable for $f \in L^2(\mathbb{R})$.
- Second, contrary to the case of intervals, we are using a non-countable number of waves over $\mathbb{R}$. Moreover notice that none of the wave functions $\exp(i2\pi xk)$ itself is square-integrable over $\mathbb{R}$ (or integrable for that sake!)

Still, things work again nicely for $\mathcal{L}^2$ spaces, explaining why they are so important:

Theorem 3.26 (Fourier transform on square-integrable functions)**.** *The Fourier transform*

$$\mathcal{F}(f)(k) := \hat{f}(k) := \int_{\mathbb{R}} f(x) \exp(-2\pi ikx) d\lambda(x)$$

defined on $\mathcal{L}^2(\mathbb{R}) \cap \mathcal{L}^1(\mathbb{R})$ has a unique extensions to an operator $\tilde{\mathcal{F}}$ on the whole of $\mathcal{L}^2(\mathbb{R})$ such that $\tilde{\mathcal{F}} : \mathcal{L}^2(\mathbb{R}) \to \mathcal{L}^2(\mathbb{R})$ is an isomorphism of complete vector spaces with an inner product.

Moreover, the inverse operator is given similarly by an extension of

$$\mathcal{F}^{-1}(f)(x) := \int_{\mathbb{R}} \hat{f}(k) \exp(2\pi ikx) d\lambda(k)$$

from the subset $\hat{f} \in \mathcal{L}^2(\mathbb{R}) \cap \mathcal{L}^1(\mathbb{R})$.

We will often abuse the notation and denote the-above mentioned extensions also by $\mathcal{F}$ and $\mathcal{F}^{-1}$.

Already the statement of this result looks a bit unpleasant with this extension and so! But this is necessary as for a general function $f \in \mathcal{L}^2(\mathbb{R})$ the integral just doesn't make directly sense.

There are several ways to prove this theorem, We will discuss a way that goes via approximations from larger and larger intervals and thus also helps to connect the Fourier transform back to the Fourier series.

In fact we will start checking that everything works well for a subset of functions.

Proposition 3.27. *Let $f \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ be such that $\hat{f}$ has a decay $|\hat{f}(k)| \leq Ck^{-2}$ for some constant C and $f(x) = 0$ outside some interval $[-M, M]$. Then we have the following Fourier inverse transform: for almost every $x \in \mathbb{R}$ we can write*

$$f(x) := \int_{\mathbb{R}} \hat{f}(k) \exp(2\pi i x k) d\lambda(k). \tag{3.1}$$

Moreover, we have the Plancherel formula:

$$\|f\|_2^2 := \int_{\mathbb{R}} |f(x)|^2 d\lambda(x) = \int_{\mathbb{R}} |\hat{f}(k)|^2 d\lambda(k) =: \|\hat{f}\|_2^2.$$

First, using the orthonormal basis $(\sqrt{\frac{1}{L}} \exp(2\pi i L^{-1} n x))_{n \in \mathbb{Z}}$ we can write

$$f(x) = L^{-1} \sum_{n \in \mathbb{Z}} \hat{f}_L(n/L) \exp(2\pi i L^{-1} n x), \tag{3.2}$$

where the summing is absolute for any $x \in [-L/2, L/2]$ and in the sense L^2, and

$$\hat{f}_L(n/L) := \int_{[-L/2, L/2]} f_L(x) \exp(-2\pi i L^{-1} n x) d\lambda(x). \tag{3.3}$$

Further, this Fourier series can be extended to a function on $\mathbb{R}$ by setting $\hat{f}_L(x) := f_L(L^{-1} \lfloor xL \rfloor)$.

Proof. Consider f as in the statement and let $L \geq 2M$. Notice that $\hat{f}_L(n/L) = \hat{f}(n/L)$ for all $n \in \mathbb{Z}$. Further, as $|f(x) \exp(-2\pi L^{-1} n x)| \leq |f(x)|$ we can first use the Dominated convergence theorem to show that in fact for every $k \in \mathbb{R}$ we have that $\hat{f}_L(k) \to \hat{f}(k)$ as $L \to \infty$, and thus the Fourier coefficient on $\mathbb{R}$ corresponds to the approximated Fourier coefficients.

We will now check that under our conditions (3.2) becomes in the limit $L \to \infty$ the inverse Fourier transform (3.1). First, given the decay of $\hat{f}$ we can write the RHS of (3.2) as the Lebesgue integral of the $\hat{f}_L$ extended to $\mathbb{R}$, i.e.

$$L^{-1} \sum_{n \in \mathbb{Z}} \hat{f}_L(n/L) \exp(2\pi i L^{-1} n x) = \int_{\mathbb{R}} \hat{f}_L(k) \exp(2\pi i L^{-1} \lfloor kL \rfloor x) d\lambda(k).$$

Now $\hat{f}_L(k) \exp(2\pi i L^{-1} \lfloor kL \rfloor x) \to \hat{f}(k) \exp(2\pi i k x)$ pointwise, and we can use the Dominated convergence theorem and the assumed decay of $\hat{f}$ to conclude that

$$\int_{\mathbb{R}} \hat{f}_L(k) \exp(2\pi i L^{-1} \lfloor kL \rfloor x) d\lambda(k) \to \int_{\mathbb{R}} \hat{f}(k) \exp(2\pi i k x) d\lambda(k).$$

Finally, the Plancherel identity for f on the interval $[-L/2, L/2]$ implies that

$$\|f\|_2^2 = L^{-1} \sum_{n \in \mathbb{Z}} |\hat{f}_L(n/L)|^2.$$

It remains to notice that the RHS converges to $\|\hat{f}\|_2^2$ - to do this we can argue like above using pointwise convergence and then dominated convergence that we obtain from the assumed decay of $\hat{f}(k)$. $\square$

To finish the proof of Theorem 3.26 it remains to do two things:

(1) Check that this special set of functions we consider in the proposition is dense in $\mathcal{L}^2(\mathbb{R})$ w.r.t. the $\|\cdot\|_2$ norm meaning that for every $f \in L^2(\mathbb{R})$ we can find some sequence f_n such that $f_n \to f$ w.r.t. the $\|\cdot\|_2$ norm and f_n satisfies the conditions of the proposition.

This can be argued as follows. First, observe that we can always approximate f by $f1_{-n,n}$ arbitrarily well.

Thus it suffices to show that we can approximate each $f1_{[-n,n]}$ arbitrarily well by some functions whose Fourier transforms behave well, i.e. decay like Ck^{-2}. But actually we have seen such functions! We just consider the heat equation on the interval $[-n, n]$ starting from $u_0 := f1_{[-n,n]}$ - we saw both that $u_t \to u_0$ in L^2 norm but also that u_t has a very rapidly decreasing Fourier transform! The details of this second part of the argument are non-examinable and are on non-examinable part of the exercise sheet.

(2) Find an argument to extend the Fourier transform using this density. We will see this argument in a more general set-up in the following section.

Section 4

Hilbert spaces and linear operators

We will now try to put the theory of L^2 spaces, their nice interplay with Fourier series and its helpfulness to solve PDEs into a more general and abstract setting. This will allow in particular to generalise our method of solving the heat equation to a wider range of equations.

4.1 Hilbert spaces

The spaces $\mathcal{L}^2$ are a prime example of what one calls Hilbert spaces. A Hilbert space is an inner product space that is complete, or in other words a Banach space with an inner product.

Definition 4.1 (Hilbert space)**.** *Let $(H, \|\cdot\|)$ be a Banach space. If H can be endowed with a compatible inner product $\langle f, g\rangle$, i.e. such that $\|f\|^2 = \langle f, f\rangle$, then we call H a Hilbert space.*

Mostly one encounters only certain class of nice Hilbert spaces that have countably dense subset in the following sense.

Definition 4.2 (Separable Hilbert space)**.** *Let H be a Hilbert space. We call H separable if there is some countable set S such that for every $h \in H$ and every $\epsilon > 0$ there is some $s \in S$ with $\|h - s\| < \epsilon$.*

What makes these Hilbert spaces extremely nice is the fact that they admit a countable basis.

Theorem 4.3. *A Hilbert space H is separable if and only if there exists an orthonormal basis $(h_n)_{n\geq 1}$ of H.*

For the space $\mathcal{L}^2([0,1])$ we showed that the trigonometric functions form an orthonormal basis and thus we know that it is also separable. We could concretely find this countable dense set by looking at all L^2 functions with rational Fourier coefficients.

It then follows that $\mathcal{L}^2([a,b])$ is also separable for any $a < b \in \mathbb{R}$. From there one can conclude that $\mathcal{L}^2(\mathbb{R})$ is separable and also that $\mathcal{L}^2([a_1,b_1] \times \dots [a_n,b_n])$ and $\mathcal{L}^2(\mathbb{R}^n)$ are separable. Some of these considerations are on the example sheet.

Let us argue how to prove the theorem. One part is relatively easy. Namely, given a countable orthonormal basis $(h_n)_{n\geq 1}$, we can consider expansions $\sum_{n\geq 1} a_n h_n$ with $a_n \in \mathbb{Q}$ and $\sum a_n^2 < \infty$. There are counatbly many such functions and they are dense in H.

The other part follows from Gram-Schmidt orthogonalization in two steps as we will explain now.

Step 1: As a first step, one extracts from our dense countable set S, a subset S_B that is linearly independent.

So consider the countable set $S = \{h_1, h_2, \dots\}$ whose closure is equal to H. We can now make it a bit smaller as follows:

(1) we set $S_B^1 := \{h_1\}$;
(2) at step m we include h_m in S_B^m if and only if there is no linear combination of elements in $S_B^{m-1} \cup \{h_m\}$ that gives 0;

(3) we define $S_B = \cup_{i\geq 1} S_B^i$.

This is called a linearly independent spanning set: linearly independent because we cannot find any $h_1, \dots, h_m \in S_B$ and numbers $a_1, \dots, a_m$ such that $\sum_{i=1}^m a_i h_i = 0$; spanning because for any $h \in H$ and any $\epsilon > 0$, we can find $h_1, \dots, h_m$ and $a_1, \dots, a_m$ such that $\| \sum_{i=1}^m a_i h_i - h \| < \epsilon$.

Step 2: The next step is to filter out an orthonormal set. This is done by revisiting an old friend of yours - the Gram-Schmidt orthogonalization procedure.

Proposition 4.4 (Gram-Schmidt orthogonalization)**.** *Suppose we have a linearly independent spanningset $S_B = \{h_1, h_2, \dots\}$. Consider the following recursive procedure:*

(1) We set $H_B^1 := \{h_1/||h_1||\}$

(2) Given H_B^{n-1}, we set $H_B^n := H_B^{n-1} \cup \{\tilde{h}_n\}$ where $\tilde{h}_n$ is defined by normalizing the vector $h_n - \sum_{i=1}^{n-1} h_i \langle h_n, h_i \rangle$ to have unit norm. [11]

(3) We define $H_B = \cup_{i\geq 1} H_B^i$.

Then H_B is an orthonormal basis of H.

Proof. By construction H_B is orthonormal. So it remains to check that it indeed satisfies the basis condition.

Consider any $h \in H$. Then as S_B is a spanning set, and each $s \in S_B$ is a finite combination of some $h_1, \dots, h_n$ in H_B, we can for any $m > 0$ find $h_1, \dots h_{n_\epsilon}$ and $c_1, \dots, c_{n_m} \in \mathbb{R}$ such that $\|h - \sum_{i=1}^{n_m} c_i h_i\| < 1/m$. It is important here to make consistent choices of c_i and this can be done by considering explicitly the orthogonal projection $g_m := \sum_{i=1}^{n_m} h_i \langle h_i, h \rangle$, which satisfies $\|h - g_m\| < 1/m$ as the projection minimizes the distance. We conclude that

$$h = \sum_{i=1}^{\infty} h_i \langle h_i, h \rangle$$

as desired. □

4.2 Bounded and unbounded linear operators on Hilbert spaces

Linear operators on $\mathbb{R}^n$ are both extremely useful and very well behaved - they can be always nicely represented using a matrix, and they are continuous.

When we go to infinite-dimensional vector spaces, things get more tricky. Linear operators are still easy to define.

Definition 4.5 (Linear operator on a Hilbert space)**.** *Let $(\mathcal{H}, \|\cdot\|)$ be a separable Hilbert space. A linear map $T : \mathcal{H} \to \mathcal{H}$, i.e. a map that satisfies $T(af + bg) = aT(f) + bT(g)$ for all $a, b \in \mathbb{C}$ (or in $\mathbb{R}$) and $f, g \in \mathcal{H}$ is called a linear operator on $\mathcal{H}$.*

However, not all of them behave very nicely. We are used to the fact that all linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ are continuous, even Lipschitz continuous. This is no longer the case - only so called bounded operators are continuous.

Definition 4.6 (Bounded linear operator)**.** *Let $(\mathcal{H}, \|\cdot\|)$ be a separable Hilbert space and T a linear operator. We call T bounded if $\sup_{\|f\|\leq 1} \|Tf\| < \infty$. Further, we define this supremum itself as the operator norm of T: $\|T\|_{op} := \sup_{\|f\|\leq 1} \|Tf\|$.*

[11] Here $\sum_{i=1}^{n-1} h_i \langle h_n, h_i \rangle$ is the projection of h_n onto the subspace spanned by $h_1, \dots, h_{n-1}$.

A key result says that continuity and boundedness are the same thing. We are more interested in the part bounded implies continuous, which is on the example sheet.

Theorem 4.7. *A linear operator $T : \mathcal{H} \to \mathcal{H}$ is continuous if and only if it is bounded.*

Here are some examples of bounded linear operators.

Example 4.8 (Bounded linear operators)**.**

- *The identity operator $Id : \mathbb{H} \to \mathbb{H}$ defined by $Id(f) := f$. It has norm equal to* 1
- *Finite rank operators that are the equivalent of linear operators on finite spaces: pick $f_1, \dots, f_n \in \mathcal{H}$, $g_1, \dots, g_n \in \mathcal{H}$ and define $T(f) := \sum_{i=1}^n \langle f, f_i \rangle g_i$.*
- *Fourier transform on $\mathcal{L}^2(\mathbb{R})$ is bounded and has unit norm as we have that $\|\mathcal{F}(f)\|_2 = \|f\|_2$.*
- *Position operator from quantum mechanics on $\mathcal{L}^2([a,b])$ defined by $T(f) := xf$ is bounded: indeed*

$$\|T(f)\|_2^2 = \int_{[a,b]} |x|^2 |f(x)|^2 d\lambda(x) \leq \max(|a|, |b|)^2 \|f\|_2^2.$$

But there are natural unbounded ones too!

Exercise 4.1. *Show that the following operators are not bounded:*

- *The position operator on $\mathcal{L}^2(\mathbb{R})$ defined by $T(f) := xf$ is not defined on the whole of $\mathcal{L}^2(\mathbb{R})$ and is not bounded;*
- *The momentum operator on $\mathcal{L}^2(\mathbb{R})$ defined by $T(f) := -i\frac{df}{dx}$ is neither defined on the whole of $\mathcal{L}^2(\mathbb{R})$ and is not bounded.*

The mathematical theory of unbounded operators is more subtle and not touched upon here - as we saw natural properties like continuity get lost and in fact one cannot even define the operator on the whole space. This is contrast with bounded operators that can always be nicely extended to the whole space, even if sometimes we manage to directly define them only on a subspace. The following theorem in particular fills in the remaining part of the proof for the Fourier transform - it shows that the Fourier transform that we defined using the Lebesgue integral for a subclass of nice functions can be extended to all of $\mathcal{L}^2(\mathbb{R})$.

Theorem 4.9 (Extension of bounded operators)**.** *Let T be a bounded operator on a subspace $D \subseteq \mathcal{H}$ of a separable Hilbert space $\mathcal{H}$. Then there is a extension of the operator T to an operator $\tilde{T} : \mathcal{H} \to \mathcal{H}$ such that $\tilde{T} = T$ when restricted to D and $\|\tilde{T}\|_{op} = \|T\|_{op}$.*

Moreover, if D is dense, this extension is unique.

Proof. Let us look at the proof in the case where D is dense. Then the extension works using completeness: let $h \in H$ and suppose $T(H)$ not defined. As D is dense, we have $h_1, h_2, \dots \in D$ such that $\|h - h_i\| \to 0$. But then $\|T(h_n - h_m)\| \leq \|T\|_{op} \|h_n - h_m\|$ and in particular as $(h_n)_{n \geq 1}$ is Cauchy because it converges, so is $(Th_n)_{n \geq 1}$. But H is complete, so $(Th_n)_{n \geq 1}$ converges to some $\tilde{h}$. We then define $\tilde{T}h := \tilde{h}$. This gives a way to extend T to the whole of H and one can verify that the norm of the extension doesn't change.

To see uniqueness, consider two such extensions $\tilde{T}$ and $\widehat{T}$ both agreeing on D. Then $S := \tilde{T} - \widehat{T}$ is a bounded linear operator defined on H and equal to 0 on D. But then for every $h \in H$ we have a sequence h_n converging to h with $Sh_n = 0$, it follows that $Sh = 0$ and thus $\tilde{T} = \widehat{T}$. □

4.3 Compact operators and the spectral theorem

Among all bounded operators, the compact operators are the very well-behaved ones. They behave very much like the friendly linear operators on $\mathbb{R}^n$.

Definition 4.10 (Compact operator)**.** *Let $(\mathcal{H}, \|\cdot\|)$ be a Hilbert space and T a bounded linear operator. We call T compact if for every sequence $(f_n)_{n\geq 1}$ with $\|f_n\| \leq 1$ there is a subsequence $(f_{n_k})_{k\geq 1}$ such that $(T(f_{n_k}))_{k\geq 1}$ converges.*

Remark 4.11. *Another equivalent way to define compactness is just to say that the image of the unit ball under T has to be compact. But often the criteria via sequences is what one checks.*

In particular, notice that compact operators are bounded by definition. Yet not all the bounded operators that one would hope or expect to be compact are compact:

Example 4.12 (Non-compact and compact operators)**.**

- *The identity operator is compact if and only if the Hilbert space $(\mathcal{H}, \|\cdot\|)$ is finite dimensional (and hence isomorphic to $(\mathbb{R}^n, \|\cdot\|_2)$.*

  Indeed, it is an application of Bolzano Weierstrass to see it is compact in case it the dimension is finite.

  In the infinite-dimensional case consider an orthonormal basis $(\phi_n)_{n\geq 1}$. But $Id(\phi_n) = \phi_n$ and $\|\phi_n - \phi_m\|_2 = \sqrt{2}$ for all $m \neq n$. Hence $(Id(\phi_n))_{n\geq 1}$ has no subsequence that is Cauchy and hence no subsequence that is convergent.
- *The Fourier operator $\mathcal{F}$ on $\mathcal{L}^2(\mathbb{R})$ is bounded but not compact - indeed a orthonormal sequence $(\phi_n)_{n\geq 1}$ is mapped to an orthonormal sequence $(\hat{\phi}_n)_{n\geq 1}$ which has no convergent subsequence.*
- *However, finite rank operators are compact, again it is a matter of Bolzano Weierstrass as you can verify on the example sheet.*

In fact compact operators are close as one can get to finite-rank operators: they are given by their limits.

Proposition 4.13. *Let $(\mathcal{H}, \|\cdot\|)$ be a Hilbert space. Then a bounded operator $T : \mathcal{H} \to \mathcal{H}$ is compact if and only if it is a limit of finite-rank operators in the operator norm, i.e. there exist $(T_n)_{n\geq 1}$ of finite rank with $\|T_n - T\|_{op} \to 0$.*

The key lemma is the following:

Lemma 4.14. *Let $(T_n)_{n\geq 1}, T$ be bounded operators and suppose that $\|T_n - T\|_{op} \to 0$. If $(T_n)_{n\geq 1}$ are compact, then so is T.*

Proof idea. We need to prove that for any sequence $(f_n)_{n\geq 1}$ in $\mathcal{H}$ with $\|f_n\| \leq 1$ we can find a subsequence $(f_{n_k})_{k\geq 1}$ such that $T(f_{n_k})$ converges as $k \to \infty$.

To do this we use an iterative extraction procedure:

(1) As T_1 is compact, we can find a subsequence $(f_{n_{1,k}})_{k\geq 1}$ such that $(T_1(f_{n_{1,k}}))_{k\geq 1}$ converges.

(2) As further T_2 is compact, we can extract a further subsequence $(f_{n_{2,k}})_{k\geq 1}$ from the previous one such that $(T_2(f_{n_{2,k}}))_{k\geq 1}$ also converges (notice that by definition the convergence holds for T_1.

(3) we continue iteratively.

The claim is then that if we consider the diagonal subsequence $(f_{n_{k,k}})_{k\geq 1}$ then $(T(f_{n_{k,k}}))_{k\geq 1}$ converges as $k \to \infty$. To do this, it suffices to verify that the sequence is Cauchy, which one can do by a 3ϵ argument: we first take n large enough such that $\|T_n - T\|_{op}$ is as small as we wish and then go far enough in the sequence such that for $\|T_n(f_{k,k}) - T_n(f_{l,l})\|$ is as small as we wish for all k, l large enough. □

4.3.1 The spectral theorem for compact operators

Maybe the most important result about compact operators is the fact that when they are symmetric, then they can be diagonalised in the very similar way as in the finite-dimensional case, i.e. similary to say real symmetric matrices. The definition of Hermitian / symmetric operator in this setting is as follows.

Definition 4.15 (Hermitian / symmetric operator)**.** *Let $(\mathcal{H}, \|\cdot\|)$ be a Hilbert space and T a bounded operator. We call T Hermitian or symmetric if for all $f, g \in \mathcal{H}$ it holds that $\langle Tf, g\rangle = \langle f, Tg\rangle$.*

The spectral theorem then says that any Hermitian compact operator can be diagonalized. As mentioned, it is the generalisation of diagonalisation of symmetric / Hermitian matrices.

Theorem 4.16 (Spectral theorem for compact operators)**.** *Let $\mathcal{H}$ be a separable Hilbert space and $T : \mathcal{H} \to \mathcal{H}$ a compact Hermitian operator.*

Then we can find an orthonormal basis $(\phi_n)_{n\geq 1}$ and real numbers $(\lambda_n)_{n\geq 1}$, called eigenfunctions and eigenvalues such that

- *$T(f) = \sum_{n\geq 1}\langle f, \phi_n\rangle\lambda_n\phi_n$ and in particular $T(\phi_n) = \lambda_n\phi_n$ for all $n \geq 1$*
- *For each $\epsilon > 0$, there are only finitely many eigenvalues λ_i with $|\lambda_i| > \epsilon$.*

We call this the diagonalization of the operator, because the basis of eigenfunctions $(\phi_n)_{n\geq 1}$ allows us to represent the operator T by a multiplication in each of the coordinates exactly like in the finite-dimensional case: if $f = \sum_{n\geq 1} c_n\phi_n$ then $T(f) = \sum_{n\geq 1} c_n\lambda_n\phi_n$.

The proof of the spectral theorem is quite beautiful itself, but we have to skip it. I have still given the argument in an appendix of the notes.

4.3.2 Heat equation / Fourier series from a compact operator

We will now connect this abstract theorem to the heat equation and its solution via the Fourier series. Our aim is to find a way of solving the heat equation in a way that generalises to the inhomogeneous case and to further PDEs.

So consider from the beginning the inhomogeneous heat equation

$$\frac{\partial u(t,x)}{\partial t} = k(x)\Delta_x u(t,x)$$

with some boundary conditions to be determined. Here the inhomogeneity in space is given by a positive continuous function $k(x)$ - when it is constant we are back to the usual heat equation.

When we solved the heat equation earlier for $k = const$ we observed that trigonometric functions behave well with respect to Δ_x. You can verify that this doesn't help us much in

the case $k(x)$ non-constant, as multiplying a trigonometric function by $k(x)$ mixes up all the frequencies.

So instead, let us make another good guess to start off: we will look for solutions of the form $u(x,t) = v(x)w(t)$. Plugging this form into the heat equation we obtain

$$v(x)\frac{dw(t)}{dt} = k(x)w(t)\frac{d^2v(x)}{dx^2}.$$

Thus we obtain a solution $u(x,t) = v(x)w(t)$ as soon as we find solutions to the two ODEs:

$$(1)\ \ k(x)\frac{d^2v(x)}{dx^2} + \lambda v(x) = 0$$

and

$$(2)\ \ \frac{dw(t)}{dt} = -\lambda w(t).$$

The second of these equations has a straightforward solution given by $w(t) = c\exp(-\lambda t)$. The first equation is what is called a Sturm-Liouville type of equation and can be interpreted as finding the eigenvalues and eigenfunctions of the operator $(Lf)(x) := k(x)\frac{d^2f(x)}{dx^2}$.

Let us see how this Sturm-Liouville type of equation can be solved in the earlier homogeneous case $k = const$ and how it links to Fourier series. In this case the equation becomes just

$$(1)\ \ k\frac{d^2v(x)}{dx^2} = -\lambda v(x),$$

i.e. the equation for eigenfunctions of the Laplacian operator Δ times a constant.

Now the idea is that whereas Δ is a unbounded operator, its inverse - the Green's operator T_G - is actually a compact operator and one can use the spectral theorem for that!

You have already seen the Green's operator, but let us recall it with our vocabulary and mindset. Indeed, consider the Poisson equation $-\Delta f = u$ on say the unit ball $B(0,1) \subset \mathbb{R}^d$ with zero boundary conditions. This can be solved by $f(x) := \int_{B(0,1)} G(x,y)u(y)dy$ where $G(x,y)$ is the Green's kernel of the Laplacian with zero boundary conditions i.e. the symmetric kernel given informally by $(-\Delta)^{-1}$ with zero boundary conditions. In $d = 1, 2$ there is an explicit formula:

- given on $[0,1]$ by $G(x,y) := x(1-y)$ for $0 \leq x \leq y \leq 1$
- and by $G(x,y) := -\log|x-y| + \log|1-x\bar{y}|$ on $B(0,1) \subseteq \mathbb{R}^2$.

Definition 4.17 (Green's operator of the Laplacian)**.** *We call the operator on $\mathcal{L}^2(B)$ defined by $(T_Gf)(x) := \int_B G(x,y)f(y)d\lambda(y)$ the Green's operator (of the Laplacian) with zero boundary conditions.*

Now consider the setting of the unit interval and let us admit for the moment that the (zero boundary). We claim the following.

Claim 4.18. *Green's operator from $\mathcal{L}^2([0,1])$ to $\mathcal{L}^2([0,1])$ is compact and symmetric.*

So in particular the spectral theorem applies and we can diagonalise the Green's kernel. What are the eigenfunctions / eigenvaues? IT comes out that we obtain exactly the sin-Fourier series!

Claim 4.19. *The spectral theorem applied to the Green's operator T_G on $\mathcal{L}^2([0,1])$ gives the collection of eigenfunctions with zero boundary conditions: $\phi_n = \sin(\pi n x)$ and $\lambda_n = -(n\pi)^{-2}$. Further, these are exactly all the zero boundary eigenfunctions of Δ with eigenvalues λ_n^{-1}.*

Proof. To see this observe that by definition every eigenfunction f_i with eigenvalue λ_i of T_G solves $T_G(f_i) = \lambda_i f_i$. But as T_G is the inverse operator of $-\Delta$, we know that $T_G(f_i)$ satisfies $-\Delta(T_G(f_i)) = \lambda_i f_i$. Thus we see that every eigenfunction of T_G is also an eigenfunction of $-\Delta$ and vice-versa. Further the eigenvalues are related just by taking $\lambda_n \to \lambda_n^{-1}$. Moreover, we observe that $T_G(f)(0) = T_G(f)(1) = 0$ for any $f \in L^2([0,1])$.

But now we have already seen a collection of eigenfunctions of the Laplacian with zero boundary conditions on $(0,1)$: the sine series $\sin(\pi n x)$ for $n \geq 1$. We also saw that they form an orthonormal basis, so they must be in fact exactly the complete set of eigenfunctions of the Green's operator, with eigenvalues $\lambda_n := (n\pi)^{-2}$! □

This means that all solutions to the equation (1) are given by $(\sin(\pi n x))_{n\geq 1}$, each with $\lambda_n = n^2\pi^2/k$ because of the extra k factor. You can think of it as follows: the spectral theorem gives us the right orthonormal basis of $\mathcal{L}^2$ that matches with the spatial equation (1) and hence the initial PDE!

The nice thing is that this new strategy generalises! The only thing that remains to do is to understand how and in which set-up we can solve the equation (1) that we obtained from separation of time and space variables. This in turn reduces to being able to write down the inverse operator and verifying this is compact. This in fact can be done in a very large generality with no serious difficulty! This is achieved by the so-called Sturm-Liouville theory discussed this year in the non-examinable section.

This way we obtain hand-crafted series expansions suitable to our initial PDE in a much wider generality. So when in the beginning of the course we found the matching series - Fourier series in the case of the heat equation - just by chance, then now we have a general method for finding such a series, which I think is very nice!

4.4 Sturm-Liouville equations * non-examinable *

In this non-examinable section we discuss how to solve the equation (1) just above in a general set-up. It fits in the framework of so-called Sturm-Liouville equations. They are helpful to solve a large class of PDEs.

The form of the equation encompasses quite a large class of ODEs - in fact, any 2nd order ODE of the form $a(x)u''(x) + b(x)u'(x) + c(x)u(x) + \lambda d(x)u(x) = 0$ with say positive a, d can be brought into this form. Let us state a general (but not the most general) theorem and then discuss how to connect it to the abstract theory developed above.

Theorem 4.20 (Sturm-Liouville problem)**.** *Consider a regular Sturm-Liouville problem on $[a,b]$, i.e. finding pairs $u \in C^2([a,b],\mathbb{C})$ and $\lambda \in \mathbb{C}$ such that*

$$\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u(x) = -\lambda u(x), \tag{4.1}$$

where the functions $p, p', q : [a,b] \to \mathbb{R}$ are continuous and further p is positive throughout $[a,b]$, and that the boundary conditions

$$\alpha_a u(a) + \beta_a u'(a) = 0 \tag{4.2}$$

$$\alpha_b u(b) + \beta_b u'(b) = 0, \tag{4.3}$$

hold for some $\alpha_a, \beta_a, \alpha_b, \beta_b$ so that neiter both α_a, β_a nor both α_b, β_b are zero.

Then there is countable sequence of real eigenvalues $\lambda_1 < \lambda_2 < \dots$ and corresponding solutions u_{λ_i} such that $(\lambda_i, u_{\lambda_i})_{i \geq 1}$ are the only possible pairs of solutions and moreover $(u_{\lambda_i})_{i \geq 1}$ form an orthonormal basis of $L^2([a,b])$.

12

Let us see how to prove this theorem. First there are some consequences from the general theory of ODEs:

- One can always find a unique solution for Equation (4.1) with boundary conditions only at one endpoint of the interval - i.e. either with $u(a) = u_0(a)$, $u'(a) = u_{00}(a)$ or with $u(b) = u_0(b)$, $u'(b) = u_{00}(b)$. In particular we can for any $\lambda \in \mathbb{C}$ find solutions $u_1(x)$, $u_2(x)$ that satisfy respectively either boundary conditions (4.2) or (4.3).
- Moreover for any $\lambda \in \mathbb{C}$ there is at most one solution up to multiplication: indeed, if there were two solutions, one would be able to contradict the uniqueness of solutions for ODEs with boundary conditions $u(a) = u_0(a)$, $u'(a) = u_{00}(a)$.

Second, notice that if we denote $Tf := \frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u(x)$ defined on the subset $C^2([a,b], \mathbb{C})$ of $L^2([a,b], \lambda, \mathbb{C})$ then by an integration by parts identity for any $f, g \in C^2([a,b], \mathbb{C})$ we see that $(Tf, g) = (Tg, f)$. Thus T is Hermitian, and in particular if λ_1, λ_2 are two eigenvalues, then as we saw above the respective eigenfunctions are orthogonal w.r.t. $\| \cdot \|_2$. In particular there can be only countably many eigenvalues because our space is separable. We can also infer already from here that all eigenvalues have to be real. Yet T is not compact, so we cannot conclude just as yet.

But the trick is - exactly as above in the case of the Laplace operator - to construct an inverse operator T_K of T that is compact and to which we can apply the spectral theorem to get an orthonormal basis of eigenfunctions. Showing that every eigenfunction of T_K is also an eigenfunction of T will then conclude the completeness of the eigenfunction basis and thus the theorem exactly as in the case of Fourier.

To do this assume WLOG that 0 is not an eigenvalue of T (otherwise we can take some λ_0 that is not an eigenvalue and relabel $\lambda \to \lambda - \lambda_0$ in the set-up).

Using the remarks above, we can then choose the solutions $u_1(x)$, $u_2(x)$ satisfying boundary conditions either only at a or b respectively to be linearly independent - otherwise they would satisfy both boundary conditions and 0 would be an eigenvalue. Then we have that $W(x) := u_1(x)u_2'(x) - u_2(x)u_1'(x)$ is non-zero but by a direct computation $(pW)'(x) = 0$ and thus $p(x)W(x)$ is a constant we denote by just pW.

[12] You may notice that the inhomogenous heat equation doesn't directly fit in this framework. However, the general Sturm-Liouville theoryalso works when on the RHS one has $-\lambda k(x)u(x)$ for a continuous positive function k, it just changes a bit the orthogonality statements as one would need to look at weighted L^2 spaces - the reason why we kept this coefficient equal to 1. One can also reduce the inhomogenous heat equation to the form in the theorem via a change of variable and by a redefinition of the function u

Now we define

$$K(x,y) = \begin{cases} -u_1(x)u_2(y)/pW & \text{for } a \leq x \leq y \leq b \\ -u_1(y)u_2(x)/pW & \text{for } a \leq y \leq x \leq b \end{cases}$$

This gives a special type of operator:

Definition 4.21 (Hilbert-Schmidt integral operators)**.** *Let $B(0,R) \subseteq \mathbb{R}^d$ be the ball*[13] *of radius $R > 0$.*

If $K(x,y) : B(0,R) \times B(0,R) \to \mathbb{R}$ is square integrable, i.e. $\int_{B(0,R)\times B(0,R)} |K(x,y)|^2 d\lambda(x,y) < \infty$, the integral operator defined on $\mathcal{L}^2(B(0,R))$ and given by

$$T_K(f) := \int_{B(0,R)} K(x,y) f(y) d\lambda(y)$$

is called the Hilbert-Schmidt integral operator.

For example it is easy to see that the Green's operator in 1D, 2D, 3D is Hilbert-Schmidt. The key claim now, stated a bit informally, is the following:

Lemma 4.22. *Hilbert-Schmidt integral operators as defined just above and seen as operators from $\mathcal{L}^2([a,b]) \to \mathcal{L}^2([a,b])$ are compact.*

In particular they can be diagonalised! As one can further verify that all eigenfunctions of T will necessarily be eigenfunctions of T_K, and vice-versa, the solution to the Sturm-Liouville problem follows!

I think it is quite elegant how this slightly abstract looking spectral theorem could help us with such concrete problems. In fact, historically Sturm-Liouville problems were the motivation and inspirations for a part of spectral theory, so the interplay of physics and mathematics is most certainly a fruitful one. ;)

Thank you!

[13] It could also be anything else reasonable like a box of $\mathbb{R}^d$ or similar

Section 5

*non-examinable appendix: the proof of spectral theorem for hermitian compact operators*ee

The proof of the spectral theorem relies on the following lemmas:

Lemma 5.1. *Let T be compact and Hermitian on a Hilbert space $\mathcal{H}$. Then either $\|T\|_{op}$ or $-\|T\|_{op}$ is an eigenvalue, i.e. there is some $f \in \mathcal{H}$ such that $Tf = \|T\|_{op} f$ or $Tf = -\|T\|_{op} f$.*

Lemma 5.2. *Let T be compact and Hermitian on a Hilbert space $\mathcal{H}$. Then all the eigenvalues λ are real, and for each non-zero eigenvalue λ the subspace of $f \in \mathcal{H}$ with $Tf = \lambda f$ is finite-dimensional.*

Lemma 5.3. *Let T be compact and Hermitian on a Hilbert space $(\mathcal{H}, \|\cdot\|)$. Then any two eigenfunctions f, g corresponding to different eigenvalues are orthogonal. Moreover, for every $\epsilon > 0$ there are only finitely many eigenvalues with $|\lambda_i| > \epsilon$.*

Before proving them, let us see how to conclude the theorem.

Proof sketch of Theorem 4.16. Most of the statements in the theorem are given by the lemmas above. Indeed, by the first lemma the set of eigenvalues is non-empty, by the second lemma we can for each eigenvalue λ find orthonormal vectors $\phi_{\lambda,1}, \dots \phi_{\lambda,d_\lambda}$ with $\mathcal{H}\phi_{\lambda,i} = \lambda\phi_{\lambda,i}$. Further by the third lemma we can choose all the eigenfunctions to be orthonormal.

It just suffices to prove that this set of orthonormal eigenfunctions $(\phi_n)_{n\geq 1}$ is complete: i.e. that every f can be written in $\mathcal{H}$ as $f = \sum_{n\geq 1}\langle f, \phi_n\rangle\phi_n$.

Now one can verify that the linear span of all these eigenfunctions is a closed subspace of $\mathcal{H}$. Consider it's orthogonal complement $\mathcal{H}_R \subseteq \mathcal{H}$ that is itself a Hilbert space. Notice that for any $h \in \mathcal{H}_R$ we have that $Th \in \mathcal{H}_R$: indeed $(Th, \phi_n) = (h, T\phi_n) = (h, \lambda_n\phi_n) = 0$ for any eigenfunction ϕ_n. If $\mathcal{H}_R$ is non-empty, then the restriction of T to $\mathcal{H}_R$ is again a compact Hermitian operator. But by Lemma 1, it always has an eigenvalue and an eigenfunction, giving a contradiction with the fact that we already listed all of them. Thus $\mathcal{H}_R$ is empty and thus the set of orthonormal eigenfunctions indeed forms a basis. □

None of the lemmas is too tricky to prove and it is interesting to see how being Hermitian really forces things to be work out.

Proof of Lemma 5.1. We can assume by scaling that $\|T\|_{op} = 1$ and thus we need to show that either 1 or -1 is an eigenvalue.

First we claim that there is some unit norm function g with $\|Tg\| = 1$: indeed, by definition of $\|T\|_{op}$ there is some sequence of unit norm functions g_n with $\|Tg_n\| \to 1$. As T is compact, $\|Tg_n\|$ converges to some g. We claim that $\|Tg\| = 1$: indeed, $T^2 g_n \to Tg$ but

$$\|T^2 g_n\|\|g_n\| \geq (T^2 g_n, g_n) = (Tg_n, Tg_n) = \|Tg_n\|^2 \to 1$$

and hence as $\|g_n\| = 1$, we see that $\|T^2 g_n\| \to 1$, concluding that $\|Tg\| = 1$.

We will now show that either $Tg = g$ and thus -1 is an eigenvalue or $Tg + g$ is an eigenfunction with eigenvalue 1. To do this write $Tg = cg + df$ with f orthogonal to g and $c, d \in \mathbb{C}$ with $|c|^2 + |d|^2 = 1$. Then as

$$1 = (Tg, Tg) = (T^2 g, g) \leq \|T^2 g\|\|g\| = \|T^2 g\|,$$

we conclude that $T^2 g = g$. But now we can conclude: either $Tg+g = 0$ or $T(Tg+g) = g+Tg$ as promised. □

Proof of Lemma 5.2. Let (f, λ) be an eigenfunction, eiganvalue pair. Then

$$\lambda(f, f) = (Tf, f) = (f, Tf) = \bar{\lambda}(f, f),$$

and thus λ is real.

Suppose for contradiction that $\lambda \neq 0$ but the subspace of functions f with $Tf = \lambda f$ is not finite-dimensional. Then we can find unit norm orthogonal functions $f_1, f_2, \ldots$ such that $Tf_n = \lambda f_n$ and in particular $\|Tf_n - Tf_m\| \geq \sqrt{2}|\lambda|$. But this is a contradiction as T is compact and thus Tf_n should admit a convergent subsequence, in particular a subsequence that is Cauchy. □

Proof of Lemma 5.3. First, let $(f, \lambda_1), (g, \lambda_2)$ be two eigenfunction, eigenvalue pairs with $\lambda_1 \neq \lambda_2$. Like in the lemma above: $\lambda_1(f, g) = (Tf, g) = (f, Tg) = \lambda_2(f, g)$, which can only hold when $(f, g) = 0$, i.e. f, g are orthogonal.

Now suppose there are infinitely many eigenvalues with $|\lambda| > \epsilon$. Then we can find orthonormal eigenfunctions $(\phi_n)_{n\geq 1}$ with eigenvalues $(\lambda_n)_{n\geq 1}$ with $|\lambda_n| > \epsilon$. But we can now replicate the proof of the lemma above to obtain a contradiction: as T is compact the sequence $(T\phi_n)_{n\geq 1}$ should admit a convergent subsequence, yet $\|T\phi_n - T\phi_m\| \geq \sqrt{2}\epsilon$ thus there is no subsequence that is Cauchy. □