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 $f_n(x) = \sin(2\pi x)$ 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 $\Delta f_n = -4D\pi^2 n^2 f_n$ 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 now 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

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

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 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 boxrd $D \subseteq \mathbb{R}^n$, 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 choosing other closed and bounded domains.

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

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

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 $C(D, R)$ 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.