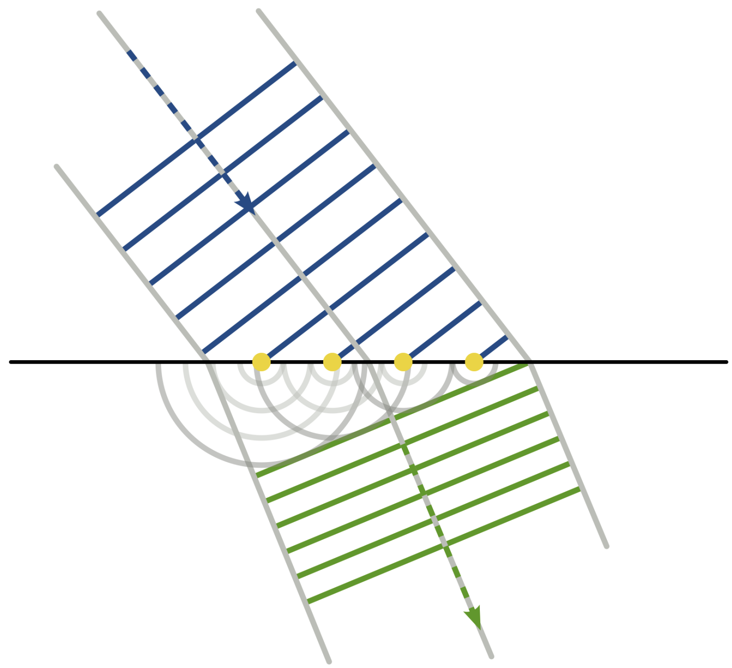
Introduction to PDEs and Fourier Analysis[1]

Matías Raja

[1] Curso de Grado en Matemáticas en la Universidad de Murcia.



Refraction explained with Huygens principle (picture taken from Wikipedia).

Contents

Preface

These notes on Partial Differential Equations (PDEs) and Fourier Analysis were prepared as a support to my teaching of this subject to third (eventually fourth) year students of the degree of Mathematics at University of Murcia. Actually, the writing of the notes was simultaneous with the teaching[1], so it was the choice of the material and the approach to several problems too. As supporting notes for my teaching, I have excused myself of include none of the drawings that I usually perform on the blackboard.

The choice of topics is motivated the requirements of the degree in Mathematics, the background of the students, the duration of the course and my personal taste. Now I will briefly describe the context and the contents.

In our syllabus, the subject that covers some basic Functional Analysis is studied during the last year, after PDEs. For that reason, the approach to PDEs is the most classic one, avoiding, among many other things, the operator point of view, semigroup language, distributional solutions and, therefore, also Sobolev spaces. On the other hand, the multiple connections with other areas of Analysis make this a very interesting subject plenty of useful ideas beyond the mere solution of particular PDEs. This subject is also the first approach of our students to Fourier Analysis, covering series and a bit of the transform.

Chapter 1 shows a variety of PDEs that appear in several disciplines. In particular, the three main second order PDEs coming form Physics (wave, Laplace, heat-diffusion) are obtained here, casting some light of the meaning of the Laplacian operator. The motivation to most of the examples included here is that context matters. Indeed, knowing how a particular PDE arises provides some intuition for the mathematical development around it.

[1] The material is still subject to corrections and minor changes.

Chapter 2 covers some elementary material on first order PDEs. The reason is twofold: on one hand, to provide to the students a bit of general theory before the great second order PDEs, thus the first order PDEs come in handy in that sense; on the other hand, first order PDEs admit systematic solution methods, keeping so some analogy with the ODEs course. The quasi-linear equation is fully discussed. We have included Frobenius theorem on integration of differential forms in three variables. The treatment of the general first order PDE is reduced to the practical method of Lagrange-Charpit, which requires some understanding of envelopes of families of surfaces.

Chapter 3 deals with Fourier series. We begin with the Hilbert theory, that clearly explains the formula of the coefficients, but is quite unsatisfactory as to convergence. Then we study the uniform convergence of Fourier series, which depends on the decay of the coefficients. More difficult is the analysis of pointwise convergence based on the Dirichlet kernel, that will allow us to study the convergence of the series around discontinuities. The section on summation methods has a strong interaction with following chapters.

Chapter 4 is devoted to the wave equation. For the scalar case, the comparison between the d'Alembert method and the separation of variables method (based on Fourier series) provides a first intuition about generalized solutions. The explicit formulas for the solution of the initial value problems in three and two dimensions are given. Boundary problems in several dimensions are also discussed.

Chapter 5 covers son material on harmonic functions, Laplace equation and eigenvalues of the Laplacian. We discuss the representation of solutions with the help of Green functions, applying it to the Laplace equation on the half-plane and the Euclidean ball. We show the connection of the Poisson kernel to complex analysis in the 2-dimensional case. We finish by discussing the solvability of the Dirichlet problem.

Chapter 6, the last one, addresses the heat-diffusion equation. We start by computing the "reasonable solutions" in one dimension. The informal choice of words is related to the lack of uniqueness for solutions, that makes the theorem in this chapter more delicate somehow. Fourier transform is introduced here because it provides a theorem of existence and uniqueness for the initial value problem. The last section is devoted to the Brownian motion, as the expression of particular solution to a diffusion process (whose average is represented by means of the heat kernel). The interesting tool here is the use of random Fourier series, which tights together the material of this course even more.

Every chapter has a section at the end called *Rationale and remarks* where the point of view adopted in the chapter is discussed and some further developments are proposed. I have included an Appendix with some required results from the theory of ordinary differential equations (ODEs).

My area of expertise/research in Mathematics has a negligible intersection with Partial Differential Equations Theory. I (want to) believe that my naivety addressing the mathematical problems of this course could have some pedagogical profit. In any case, I have tried to follow a slowly increasing level of abstraction and technical difficulty.

In order to cover the prerequisites in several real variables and vector analysis I (dare to) recommend my own text available in `https://webs.um.es/matias/miwiki/lib/exe/fetch.php?media=fvvr.pdf`.

As I said before, these notes were prepared while I was teaching the subject. I would like to apologise to my former students who suffered me through the creative process.

Murcia, Fall 2023.

Chapter 1

PDEs are everywhere

Here we will show how partial differential equations appear in many different contexts. It turns out that each important example requires additional knowledge outside from Mathematical Analysis, such as Physics or Probability Theory. Sometimes, the origin of problems is consider perfunctory if not it is directly despised. My opinion is that the task of a mathematician is not just to solve problems, but also to pose them. Moreover, the intuition cast by the real model behind the equation can of great help in the mathematical research.

1.1 What is a PDE?

The acronym PDE stands for *partial differential equation.* A partial differential equation is a relation satisfied by a function of several variables and some of its partial derivatives, so it can be as a generalization of an ordinary differential equation where only one independent variable is involved. Typically, a PDE in two variables is an expression of the form

$$F\left(x, y, u(x,y), \frac{\partial u}{\partial x}(x,y), \frac{\partial u}{\partial y}(x,y), \frac{\partial^2 u}{\partial x^2}(x,y), \dots\right) = 0$$

where $u(x,y)$ is an unknown function. The order of the PDE is the highest order of the involved derivatives of u. Depending on the dimension of the space where F takes values, we may have a single PDE or a system of partial differential equations.

It is natural to guess that PDEs would be more difficult that ODEs because of the greater number of variables, but there are other differences. For instance, in the theory of ODEs is possible to transform a differential equation of degree n into a first order equation on $\mathbb{R}^n$, so that the discussion of existence and uniqueness is simplified. With PDEs, a similar reduction is not possible, so the theory of first order PDEs is of not much help in the study of second order PDEs. Another trick that does not work for PDEs is that an ODE can be written as an integral equation, so the differential operator, which is seldom continuous, is transformed into a nice Lipschitz integral operator and the existence of solutions can be derived from Banach's fixed point theorem.

1.2 Three PDEs you already know

If you have been taught a course in Calculus of several (real) variables and an introduction to Complex Analysis, most likely you already know the following examples of PDEs.

Potential of a conservative field. Given n functions f_k for $k = 1, \ldots, n$ defined on $\mathbb{R}^n$ find a function u such that

$$\frac{\partial u}{\partial x_k} = f_k \text{ for every } k$$

You should remember that a necessary and, somehow, sufficient condition for the existence of solution is the equality

$$\frac{\partial f_k}{\partial x_j} = \frac{\partial f_j}{\partial x_k} \text{ for all } k, j.$$

How can be u obtained from the f_k? Instead of $\mathbb{R}^n$ we could have asked the functions to be defined on a smaller domain. If the function u exists, then it is determined up to a constant on any connected domain. Existence of u is guarantied on star-shape domains and it can be obtained as a line integral from the "center point". The problem can be stated with differential forms: consider

$$\omega = f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n.$$

Is there a function u such that $\omega = du$? The condition of solvability (local or for star-shaped domains) provided ω is C^1 can be expressed as $d\omega = 0$ (exterior differential).

Holomorphic companion. Let $f(z)$ be an holomorphic function. If f is written as

$$f(x+iy) = u(x,y) + iv(x,y)$$

where u and v are real functions, then the Cauchy-Riemann equations imply that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0;$$

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

Now, given, u satisfying that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Could you find v such that $u(x,y) + iv(x,y)$ is holomorphic? It is not difficult to check that this problem can be reduced to the previous one.

Euler's characterization of homogeneous functions. A function

$$f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$$

is said to be *homogeneous of degree* m if $f(t\mathbf{x}) = t^m f(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$ and $t > 0$. Euler showed that f is homogeneous of degree m if and only if

$$x_1 \frac{\partial f}{\partial x_1} + \cdots + x_n \frac{\partial f}{\partial x_n} = mf$$

for all $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$. How many homogeneous functions of degree m there exist on $\mathbb{R}^n$? The answer is a typical fact in PDEs: general solutions depend on arbitrary functions, unlike ODEs whose general solutions depend on a set of constants. For the proof of Euler's result just note that

$$x_1 \frac{\partial f}{\partial x_1}(t\mathbf{x}) + \cdots + x_n \frac{\partial f}{\partial x_n}(t\mathbf{x}) = \frac{d}{dt} f(t\mathbf{x}).$$

Multiplying by t and applying the hypothesis it reduces to an ODE.

1.3 The vibrating string

Consider a tensed stretched string fixed by its two endpoints, whose steady shape will be a straight line if we neglect the effect of its weight. Any small deformation caused to the string will be followed by movements aimed to restore its original shape. We will study the dynamics of those movements.

The assumptions we made to simplify the statement of the equations are the following:

- the string is unidimensional and has constant linear density $\rho > 0$;
- the movement will happen in the plane, so we will assume the steady string is contained in the x-axis, say $[0, a]$ with $a > 0$;
- any point (x, y) of the perturbed string comes from the steady point $(x, 0)$, so any point of the rope only moves vertically;
- the tension force at any point of the perturbed string is constant and equal to the one of the steady state, say $\tau > 0$;
- the deformation is smooth and the derivative takes values in a small neighbourhood of 0.

Let $u(x, t)$ for $x \in [0, a]$ and $t \geq 0$ be a function describing the shape of the string at time t (for $t = 0$ we have the initial deformation). Consider two points $x_1 < x_2$ from $[0, a]$ and some time t. By the assumptions, the mass of the string between the points $u(x_1, t)$ and $u(x_2, t)$ is $\rho(x_2 - x_1)$. Two forces act over that piece of string: the tensions at the end points. Those forces are equal in modulus, however their direction depends on $\frac{\partial u}{\partial x}$ at x_1 and x_2. Note that by the assumptions we only care about the vertical component of the tensions, that can be obtained multiplying τ by the sinuses of the angles the tangent lines made with the x-axis. Let α_i be the angle of the tangent line at $(x_i, u(x_i, t)$ for $i = 1, 2$. Again, by our assumptions

$$\sin \alpha_i \sim \alpha_i \sim \tan \alpha_i = \frac{\partial u}{\partial x}(x_i, t)$$

With the help of a picture, it is easy to see that the effective vertical force on the piece of string is

$$\tau(\frac{\partial u}{\partial x}(x_2, t) - \frac{\partial u}{\partial x}(x_1, t)).$$

Now, assume $x_1 = x$ and $x_2 = x + \kappa$ where κ is very small. Newton's second law of the dynamics says that

$$\tau(\frac{\partial u}{\partial x}(x+\kappa, t) - \frac{\partial u}{\partial x}(x, t)) = \rho\kappa\frac{\partial^2 u}{\partial t^2}(x, t).$$

Dividing by κ and taking limits produces the equation

$$\tau\frac{\partial^2 u}{\partial x^2}(x, t) = \rho\frac{\partial^2 u}{\partial t^2}(x, t)$$

that can be written as

$$\frac{\partial^2 u}{\partial t^2}(x, t) - c^2\frac{\partial^2 u}{\partial x^2}(x, t) = 0$$

where $c = \sqrt{\tau/\rho}$. We will see that the constant c can be interpreted as the *longitudinal speed* of the wave.

1.4 Vibrations in elastic media

We will consider a one-dimensional elastic medium with constant linear density ρ. Elasticity means, in this case, that the medium reacts to longitudinal deformations with a force at any point proportional to the displacement of that point with respect of its steady position. Since there are always a continuum of displaced points, the direct statement of the equations is not elementary. Actually, some background on elasticity theory is needed. In order to avoid that problem, we will "discretize" the elastic medium.

We could imagine our medium as a long coil spring with no mass, and pointwise equal masses regularly attached to the spring. In order to better discretize, we shall assume that any two contiguous masses are at distance κ and the masses weight $\rho\kappa$. Hooke's law says that a segment of coil spring of length l reacts to a deformation δ by a force of magnitude $\eta\delta/l$ where η is the elasticity constant of the spring.

Any mass only interacts with the two contiguous masses through the segment of coils joining them. Let $u(n, t)$ the displacement at time t of the n-th mass with respect its steady position. Considering the displacements of the neighborhood masses, we can deduce that the total force on the mass is

$$\eta(u(n+1, t) - u(n, t))/\kappa + \eta(u(n, t) - u(n-1, t))/\kappa$$

$$= \eta(u(n+1,t) - 2u(n,t) + u(n-1,t))/\kappa$$

By Newton's second law, that force should be equal to

$$\rho\kappa\frac{\partial^2 u}{\partial t^2}(n,t).$$

It may seem strange to use partial derivative being on of the variable discrete... In order to pass to limits we should change the index n by the position x. Now, we have

$$\eta(u(x+\kappa,t) - 2u(x,t) + u(x-\kappa,t))/\kappa = \rho\kappa\frac{\partial^2 u}{\partial t^2}(n,t).$$

Since

$$\lim_{\kappa\to 0}\frac{u(x+\kappa,t) - 2u(x,t) + u(x-\kappa,t)}{\kappa^2} = \frac{\partial^2 u}{\partial x^2}(x,t)$$

we get the differential equation

$$\eta\frac{\partial^2 u}{\partial x^2}(x,t) = \rho\frac{\partial^2 u}{\partial t^2}(n,t)$$

which is the wave equation again. Now the speed is $c = \sqrt{\eta/\rho}$.

The elastic medium can be considered in dimension 3, so the deformation would be a spatial one. That would let to a vector PDE, however we can reduce the problem to a scalar equation by considering the a "average deformation" or another related scalar magnitude, as the pressure. For that, let us recall that the Laplace operator or *laplacian* of a function f on $\mathbb{R}^3$ is defined as

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

With the help of the laplacian, the equation of the waves on elastic medium can be written as

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

That applies, for instance, to the pressure waves in air, also called *sound*. It is remarkable that the propagation speed c remains constant despite the fact that the energy is dissipated as a two or three dimensional wave moves away of its source.

1.5 Newtonian potential and the Dirichlet's principle

Given a measurable bounded function ρ with compact support, we will consider the *Newtonian potential* created by the charge $\mu = \rho\, dV$, where dV is the 3-dimensional Lebesgue measure. The potential at $\mathbf{x}$ is the function Φ defined by

$$\Phi(\mathbf{x}) = \iiint \frac{\rho(\mathbf{y})}{\|\mathbf{y} - \mathbf{x}\|}\, dV(\mathbf{y})$$

which is defined everywhere (convergence on the support of ρ can be showed easily). Poisson proved that the density can be recovered from the potential V by the following formula

$$\Delta\Phi(\mathbf{x}) = -4\pi\rho(\mathbf{x}).$$

The fact is that in some important problems neither the charge or the potential is known, however is clear that in the charge-free part of the space, say $\Omega \subset \mathbb{R}^3$ we have

$$\Delta\Phi(\mathbf{x}) = 0$$

and physical conditions allow us to know the value of Φ on $\partial\Omega$.

The Laplace equation consist in solving

$$\Delta u = 0$$

on a domain $\Omega \subset \mathbb{R}^n$ with certain restrictions. Recall that functions that anihilates the Laplacian are called *harmonic*. The so called Dirichlet's problem is the search of solutions u such that $u|_{\partial\Omega} = f|_{\partial\Omega}$ for some $f \in C(\overline{\Omega})$.

The application of Gauss divergence theorem to $v\nabla u$ where u, v are regular enough functions in $\overline{\Omega}$ produces the following equality

$$\iiint_\Omega v\Delta u\, dV + \iiint_\Omega \nabla v \cdot \nabla u\, dV = \iint_{\partial\Omega} v\nabla u \cdot d\mathbf{S}$$

where we kept the 3-dimensional notation despite the argument is valid in any dimension greater than 1. The same formula applied to $u = v$ gives

$$\iiint_\Omega u\Delta u\, dV + \iiint_\Omega \|\nabla u\|^2 dV = \iint_{\partial\Omega} u\nabla u \cdot d\mathbf{S}$$

Assume now that $u|_{\partial\Omega} = v|_{\partial\Omega}$ and u is harmonic on Ω. The difference of (what remain of) the two previous formulas is

$$\iiint_\Omega \nabla v \cdot \nabla u \, dV - \iiint_\Omega \|\nabla u\|^2 dV = \iint_{\partial\Omega} (v-u)\nabla u \cdot d\mathbf{S} = 0.$$

Consider now the inequalities

$$0 \leq \iiint_\Omega \|\nabla u - \nabla v\|^2 dV$$

$$= \iiint_\Omega \|\nabla u\|^2 dV - 2\iiint_\Omega \nabla u \cdot \nabla v \, dV + \iiint_\Omega \|\nabla v\|^2 dV$$

$$= \iiint_\Omega \|\nabla v\|^2 dV - \iiint_\Omega \|\nabla u\|^2 dV.$$

The inequality shows that u, the harmonic function, minimizes the "energy integral" among all the regular functions taking the same values on $\partial\Omega$, that is,

$$\iiint_\Omega \|\nabla u\|^2 dV = \min\{\iiint_\Omega \|\nabla v\|^2 dV : v|_{\partial D} = u|_{\partial D}\}.$$

By Physical considerations, Dirichlet believed that the existence of the minimizer function was evident and therefore the Dirichlet's problem could be solved always. Moreover, the solution is unique by the *maximum principle* for harmonic functions. This was the statement of the so called *Dirichlet's principle*, whose validity depends upon more carefully chosen hypotheses.

1.6 Equations of the electromagnetic field

The electromagnetic field can be described by two vectors fields that acts individually on charges and magnetic materials: the electric field $\mathbf{E}$, and the magnetic field $\mathbf{B}$. There are other elements as charge density ρ, whose eventual movement can be described as current density $\mathbf{J}$. They are related by the equality

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0,$$

also known as the *charge conservation principle*.

The set of equations describing the electromagnetic field (*Maxwell's equations*) are:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0;$$
$$\nabla \cdot \mathbf{B} = 0;$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t};$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}.$$

As an exercise, try to deduce the charge conservation principle from Maxwell's equations.

It is well known that if the charge remains still in time, then $\mathbf{E}$ can be described in terms of the Newtonian potential produced by ρ. However, when the electric field $\mathbf{E}$ varies in time it cannot be longer expressed in those terms. Instead, we have a scalar potential-like function ϕ and a vector potential $\mathbf{A}$. The relation to the previous ones is given by $\nabla \times \mathbf{A} = \mathbf{B}$ and

$$\nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial \phi}{\partial t} = 0 \text{ (\textit{Lorenz gauge condition})}.$$

After some manipulations is possible to get the following decoupled equations for ϕ and $\mathbf{A}$

$$\Delta \phi - \epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0},$$
$$\Delta \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J}$$

Far away form the charges and currents, the equations reduce to

$$\Delta \phi - \epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2} = 0,$$
$$\Delta \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0$$

that have the same structure that the wave equation (the second one is vectorial). A computation shows that the associated wave speed

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$$

is precisely the speed of the light. That allowed Maxwell to recognise light as a waves in the electromagnetic field, proving so that the wave theory of light was the right one against the particle theory... for some time.

1.7 The heat equation

The simplest model for the variation of the temperature of a body was established by Newton. It says that the temperature of a body at temperature T immersed in a illimited medium at temperature T_0 starts changing its temperature at a rate $T_0 - T$. More precisely,

$$\frac{dT}{dt} = k(T_0 - T)$$

where $k > 0$ is a constant which depends on the calorific capacities of the media, the size and shape of the body. Of course, the temperature of the body is not homogeneous after a while, so Newton's law is just a first rough approximation to the problem.

Fourier devised a way to deal with the non-homogeneous changes of temperature. The idea is to consider any point of a body or substance as surrounded by a medium able to modify the point's temperature. Since the temperature is spatially variable, we should work with averages. For that, it is necessary to remind a the role of the laplacian operator: is f is C^2 defined in a neighbourhood of 0 and $\varepsilon > 0$, then

$$\frac{1}{4\pi\varepsilon^2} \iint_{\partial B(0,\varepsilon)} (f - f(0))\, dS = \frac{\varepsilon^2}{6} \Delta f(0) + o(\varepsilon^2).$$

Let $u(\mathbf{x}, t)$ be the temperature at the point $(\mathbf{x})$ at time t. The local application of Newton's law takes the form

$$\frac{\partial u}{\partial t}(\mathbf{x}) = k(\mathbf{x}) \Delta u(\mathbf{x})$$

where $k(\mathbf{x}) \geq 0$ is the thermal diffusivity of the medium at $\mathbf{x}$. If the medium is homogeneous (at least in some region, say a body) then we may assume $k > 0$ constant and the equation takes the form

$$\frac{\partial u}{\partial t} = k\, \Delta u$$

which is the so called *heat equation* first stated by Joseph Fourier.

In practical applications, one may reduce the number of dimensions. For instance, the evolution of the temperature on a rod or a plate. In addition, in most cases some boundary conditions are added as to maintain some parts to a constant temperature artificially, or introducing periodic variations as the heating of the sunshine during the day.

1.8 The diffusion equation

Consider a particle or set of particles whose position is uncertain and moving by a sort of random diffusion. We may think of the position in time t as a random variable $X(t)$ and the diffusion after some time τ as an independent centered random variable D_τ that acts in this way

$$X(t+\tau) = X(t) + D_\tau$$

We shall deal with the one-dimensional case for simplicity and we will assume that $X(t)$ has a density $f(x,t)$ which is fairly regular. Assuming that D_τ is represented by a density ϕ, the density of $X(t+\tau)$ can be written as a *convolution*, namely

$$f(x,t+\tau) = \int_{-\infty}^{\infty} f(x+s,t)\phi(-s)\,ds.$$

For ϕ we have the following properties (a) and (b) and we will make a reasonable assumption (c) involving the duration of the diffusion

(a) $\phi \geq 0$ and $\int_{-\infty}^{\infty} \phi(s)\,ds = 1$;

(b) $\int_{-\infty}^{\infty} s\,\phi(s)\,ds = 0$;

(c) $\int_{-\infty}^{\infty} s^2\,\phi(s)\,ds = 2k\tau$ for some $k > 0$.

Note that we may think that most of the "mass" of ϕ is concentrated around 0 as τ is smaller. In practise, we may assume that ϕ is supported in a narrow neighbourhood of 0, which is the key for the following heuristic computation, together the assumption on the differentiability of f,

$$f(x,t+\tau) = \int_{-\infty}^{\infty} f(x+s,t)\phi(-s)\,ds$$

$$\simeq \int_{-\infty}^{\infty} \left(f(x,t) + \frac{\partial f}{\partial x}(x,t)s + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x,t)s^2 \right) \phi(-s)\,ds$$

$$= f(x,t) + \frac{\partial^2 f}{\partial x^2}(x,t)k\tau.$$

On the other hand

$$f(x,t+\tau) \simeq f(x,t) + \frac{\partial f}{\partial t}(x,t)\tau.$$

We may assume that the discrepancies between terms separated by "$\simeq$" are $o(\tau)$. Since $\tau > 0$ can be taken arbitrarily small, comparing both equations we get

$$\frac{\partial f}{\partial t}(x,t) = k\frac{\partial^2 f}{\partial x^2}(x,t)$$

that is the 1-dimensional diffusion equation. Note that it is formally the heat equation, which is not surprising as the heat transmission is a diffusion process. Actually, the diffusion equation in more dimensions under homogeneity and isotropy assumptions takes the form

$$\frac{\partial f}{\partial t} = k\Delta f.$$

1.9 Some more examples: Quantum Mechanics, Hydrodynamics, Biology, Economy...

In Quantum Mechanics the state of a system is characterized by a function $\Psi : \mathbb{R}^{n+1} \to \mathbb{C}$ where $n \in \mathbb{N}$ is the freedom degree of the system, so the function can be written $\Psi(\mathbf{x}, t)$, being $\mathbf{x}$ the spatial variables. There is normalization assumption

$$\int_{\mathbb{R}^n} |\Psi(\mathbf{x}, t)|^2 d\mathbf{x} = 1$$

and in the case of a particle with wave function Ψ, then $|\Psi(\mathbf{x}, t)|^2$ represents the probability density of the position of the particle at a given time t. From now on we may assume that there is only one particle of mass m so $\mathbf{x} \in \mathbb{R}^3$. The evolution of a quantum system is ruled by *Schrödinger's equation*

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x}, t) = \hat{H}\Psi(\mathbf{x}, t)$$

where $\hat{H}$ is the Hamiltonian operator, that in most cases, can be expressed as

$$\hat{H}\Psi(\mathbf{x}, t) = -\frac{\hbar^2}{2m}\Delta\Psi(\mathbf{x}, t) + V(\mathbf{x})\Psi(\mathbf{x}, t).$$

A fundamental role is played by the *stationary solutions*, for which the time and spatial parts can be decoupled

$$\Psi(\mathbf{x}, t) = \psi(\mathbf{x})e^{-iEt/\hbar}$$

with E being the energy (constant). Note that those solutions are *eigenfunctions* of the energy operator

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t) = E\Psi(\mathbf{x},t)$$

associated to the *eigenvalue* E. However, the spatial part is to be determined. The time-independent Schrödinger's equation is what is left for ψ, that is,

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) = E\psi(\mathbf{x})$$

which can be solved for several potentials $V(\mathbf{x})$ simple enough.

Schrödinger equation does not look like equations in Classical Dynamics because, among other things, derivation with respect time is only of first order. That is not totally right, indeed the Hamilton-Jacobi of an holonomic system is

$$\frac{\partial S}{\partial t} = -H(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}})$$

if the Hamiltonian H does not contain explicitly time, where $\mathbf{q}$ represents the (generalized) coordinates and S is a function that expresses the evolution of the system, somehow. The idea is to bring the Mathematics from Optics into Mechanics. The starting point is that *Fermat's minimum time principle* can be deduced from Huygens geometrical description of waves: light rays follow curves that are orthogonal to the wavefronts. The trajectory accomplished by a mechanical system minimizes the integral of the Lagrangian among all the feasible alternative trajectories. The Hamilton-Jacobi equation expresses the waves for which the trajectories of the system are orthogonal curves. Moreover, the *theorem of Eisenhart* says that the trajectory followed by a mechanic system is a geodesic in some suitable Riemannian manifold associated to the system.

Consider a moving fluid. At every point we have a speed $\mathbf{v}$, a pressure p and a density ρ that also may depend on time. The conservation of the mass can be expressed in the following terms.

$$\nabla \cdot (\rho\,\mathbf{v}) + \frac{\partial \rho}{\partial t} = 0$$

that is the so called *continuity equation* of fluids. Let us assume that ρ is constant (the fluid is a liquid). If we restrict the acting forces on the fluid to the pressure and external forces (as the weight), then the application of Newton's law to a small portion of volume can be transformed into

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} + \frac{1}{\rho}\nabla p - \mathbf{F} = 0$$

which known as *Euler's equation*. If the *viscosity* has to be taken into account, then the equation takes the form

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} - \nu\Delta\mathbf{v} + \frac{1}{\rho}\nabla p - \mathbf{F} = 0$$

which is the Navier-Stokes equation (ν is the kinematic viscosity coefficient).

The *transport equation*

$$\frac{\partial u}{\partial t} = k\Delta u - \mathbf{v} \cdot \nabla u$$

arises as the combination of a diffusion process with a carrying medium whose velocity field $\mathbf{v}$ is supposed stationary for the sake of simplicity.

In Ecology the predator-prey system of Lotka-Volterra can be modified to introduce spatial variables. Let $u(x, y, t)$ and $v(x, y, t)$ represent prey and predator on a given territory $\Omega \subset \mathbb{R}^2$ at the time t. We shall assume that at any point (x, y) the rates satisfy the hypotheses of Lotka-Volterra with a logistic correction for the prey, but we shall also add an additional term coming from diffusion in the sense already discussed above. For simplicity we assume that the diffusion is homogeneous. With all that we have the following system

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \alpha u(1 - \kappa^{-1}u) - \beta uv + \eta\Delta u, \\
\frac{\partial v}{\partial t} &= \gamma uv - \delta v + \zeta\Delta v,
\end{aligned}$$

where $\alpha, \beta, \gamma, \delta > 0$ are the usual parameters of the model, $\kappa > 0$ is the maximum prey population density affordable by the territory, and $\eta, \zeta > 0$ are the diffusion coefficients.

In Financial Mathematics, the so called *Black-Scholes model* deals with the price evolution in time V of an option over some asset, whose price is S. Introducing some parameters whose meaning is clear for experts in stock market, the equation of Black-Scholes is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

1.10 Rationale and remarks

The motivation of this chapter was well explained at the beginning. It is possible to add many more examples, but our main motivation was to estate the three main examples of second order PDEs: wave equation, Laplace equation and heat-diffusion equation. Each of these equations will be discussed further during the course.

As to the transmission of waves in elastic media, it is possible to obtain the equation of pressure waves (sound) from the linearization of small perturbations in Euler's fluids equation. Some additional knowledge on *adiabatic* changes of volume is necessary to get the right speed of sound at current atmospheric pressure.

To write the equations for the vibrations of elastic bars or plates require some knowledge on elasticity theory. Computations would lead to the *Kirchhoff-Love equation* $\Delta^2 u + k\partial^2 u/\partial t^2 = 0$ and the beautiful *Chladni patterns.*

The Dirichlet problem and principle illustrates a global approach to the existence of solutions: instead of building the function by adding small pieces, like in Euler's method for ODEs, the required solution appears as the minimizer of some *variational problem* whose feasibility could be established by compactness or completeness. Historically, that kind of investigation lead to the birth of *Functional Analysis.*

The change of variable $t \to -t$ on the wave equation leaves it the same, whereas the same change in the heat equation does not. For that reason, we say that vibrations of the string, or any process obeying Newton's laws, is *reversible.* A movie played backwards may seem awkward, but it makes sense from the point of view of Physical laws... provided that all the movements are simple enough. However, there are things that cannot happen backwards: imagine that all the pieces and debris of bullet shot to a wall, eventually gather together to compose one piece of metal that goes back into the gun and all the smoke goes back down the barrel (have you watch Nolan's movie *Tenet*). We say that it is an *irreversible* process. That is the same that happen with the heat equation: one end of a bar does not get hotter spontaneously. Irreversibly is related to the complexity of a system: every particle in it may have a Newtonian reversible behavior, but all together or statistically speaking have an irreversible evolution. That is the spirit of Boltzmann's theory of *entropy.*

Another interesting aspect of transformations, or *symmetries*, that preserve formally the equations is that they are related to *conservation laws* after a remarkable theorem of Emmy Noether. For instance, the fact that $t \to t+a$ does not change the equations of the movement implies the conservation of energy. Homogeneity of the space implies the conservation of linear momentum, and the isotropy of the space implies the conservation of angular momentum.

The simple models we have discussed can be combined once we know the meaning of each term. For instance, we could combine spatial Lotka-Volterra with transport equation to get a model of how chemotherapy acts on a tumor. The applications of PDEs to Oncology have proven to be useful to adjust the right dose of medication.

Nonlinear reaction-diffusion equations could produce discrete patterns that imitates natural animal prints such as leopard's spots or zebra's stripes. This theory of morphogenesis was proposed by A. Turing in 1952, see [29].

1.11 Exercises

1. Consider the function defined as

$$u(x, y) = \frac{xy}{x^2 + y^2}$$

for $(x, y) \neq (0, 0)$ and $u(0, 0) = 0$. Show that u is not continuous at $(0, 0)$. However, u has partial derivatives on $\mathbb{R}^2$ and moreover it satisfies the PDE

$$xu_x + yu_y = 0.$$

2. Solve on a connected domain of $\mathbb{R}^2$ the equation

$$u_x^2 + u_y^2 = 0.$$

3. Given a harmonic function u on the plane $\mathbb{R}^2$ show that it is possible to find another function v such that

$$z = x + iy \to u(x, y) + iv(x, y)$$

is holomorphic.

4. Find the solution on the first quadrant of the equation

$$(1+x)u_{xy} + u_y = x - 2y$$

satisfying $u(x,0) = 1$, $u(0,y) = \cos y$.

5. Let f and g be scalar functions, $\mathbf{F}$ and $\mathbf{G}$ be vectorial fields, all defined on $\mathbb{R}^3$. Prove the following formulas:

   (a) $\nabla(fg) = g\,\nabla f + f\,\nabla g$.

   (b) $\nabla \cdot (f\mathbf{F}) = \nabla f \cdot \mathbf{F} + f\,\nabla \cdot \mathbf{F}$.

   (c) $\nabla \times (f\mathbf{F}) = \nabla f \times \mathbf{F} + f\,\nabla \times \mathbf{F}$.

   (d) $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = (\nabla \times \mathbf{F}) \cdot \mathbf{G} + \mathbf{F} \cdot (\nabla \times \mathbf{G})$.

6. Prove that the expression the Laplacian in polar (2-dimensional), cylindric and spherical (3-dimensional) coordinates are respectively

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2};$$

$$\Delta u = \frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial u}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2};$$

$$\Delta u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial u}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial \varphi^2}.$$

Chapter 2

First order equations

In this chapter we will study partial differential equations of first order, and we will mainly deal with two independent variables. Therefore, the most general form of PDE of first order we will consider is

$$F(x, y, u, u_x, u_y) = 0 \tag{2.1}$$

where $u(x, y)$ is the unknown function, $u_x = \frac{\partial u}{\partial x}$ and $u_y = \frac{\partial u}{\partial y}$. Eventually, some of the theory clearly extends to more variables with no difficulty, although the proofs will remain in the two dimensional frame.

2.1 Preliminaries

Simple examples, such as $u_x = 0$ whose solutions are of the form $u(x, y) = f(y)$ with f an arbitrary one-variable real function suggest that the *general solution* of a first order PDE depends on an arbitrary function. That intuition can be reached from another point of view. Note that the any solution u of (2.1) defines a surface $z = u(x, y)$, and thus the initial data for the *Cauchy problem* is a curve that should be contained in the surface. Later we will see that the Cauchy problem may no have unique solution for fairly regular equations and sometimes not all the solutions of (2.1) are contained in the general solution.

Example 2.1.1. *Find the PDE satisfied by the functions of the form*

$$u(x, y) = xy + f(x^2 + y^2)$$

being f an arbitrary one-variable differentiable function.

Take partial derivatives with respect to x and y

$$u_x(x, y) = y + 2xf'(x^2 + y^2),$$
$$u_y(x, y) = x + 2yf'(x^2 + y^2).$$

We can get rid of f' by multiplying by x and y respectively and taking the difference

$$yu_x(x, y) - xu_y(x, y) = y^2 - x^2$$

that leads to the equation

$$yu_x - xu_y - y^2 + x^2 = 0.$$

■

Sometimes, especially in old literature, the solutions of a PDE are referred as *integrals*, mostly with "labels", e.g.: general integral, complete integral, singular integral...

Along the this chapter, the geometrical intuition plays a fundamental role. For that aim, the classic notation for the first order PDE

$$F(x, y, z, p, q) = 0 \tag{2.2}$$

where $z = u$, $p = u_x$ and $q = u_y$, allows a more flexible interpretation. Indeed, we look for surfaces passing at (x, y, z) where they are perpendicular to the vector $(p, q, -1)$, provided that the quintuple (x, y, z, p, q) satisfies (2.2). Lets us remind that an ordinary differential equation has an associated vector field in such a way that the solutions appear likewise the children's puzzle "join the dots" leads to a drawing. In our case, we may think of "small plane elements" and the solutions of (2.2) will appear before us like after a reasonable arrangement of directions. However, here the geometry is more complicated.

2.2 The quasi-linear equation

The aim of this section is to solve the so called *quasi-linear equation*

$$A(x, y, u)\,\frac{\partial u}{\partial x}(x, y) + B(x, y, u)\,\frac{\partial u}{\partial y}(x, y) = C(x, y, u) \tag{2.3}$$

for $\mathcal{C}^1$ functions $A, B, C : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$. This kind of PDE is particularly simple from a geometric point of view Indeed, take a point $\mathbf{p} = (x_0, y_0, z_0)$ in the surface $z = u(x, y)$. The equation of the tangent plane is

$$\frac{\partial u}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial u}{\partial y}(x_0, y_0)(y - y_0) = z - z_0$$

and the normal vector (to the surface) at $\mathbf{p}$ is

$$\left(\frac{\partial u}{\partial x}(x_0, y_0), \frac{\partial u}{\partial y}(x_0, y_0), -1\right).$$

On the other hand, consider the vector field given by the coefficients of 2.3

$$(A(x, y, z), B(x, y, z), C(x, y, z)).$$

The relation 2.3 expresses that the normal vector and the field are orthogonal at $\mathbf{p}$, since it can be written as a scalar product

$$\left(\frac{\partial u}{\partial x}(x_0, y_0), \frac{\partial u}{\partial y}(x_0, y_0), -1\right) \cdot (A(x_0, y_0, z_0), B(x_0, y_0, z_0), C(x_0, y_0, z_0)) = 0.$$

Therefore, since $\mathbf{p}$ was arbitrary, the field (A, B, C) is tangent to the surface $z = u(x, y)$ at each of its points. Now, we will consider the field lines given as solutions of the following the autonomous system

$$\begin{cases} x' = A(x, y, z) \\ y' = B(x, y, z) \\ z' = C(x, y, z) \end{cases} \tag{2.4}$$

It seems plausible that a field line starting at a point of $z = u(x, y)$ cannot escape from the surface. That is exactly the statement of the following theorem.

Theorem 2.2.1. *Assume that u is a $\mathcal{C}^1$ solution of 2.3 defined on $D \subset \mathbb{R}^2$ and consider the surface $S = \{(x, y, u(x, y)) : (x, y) \in D\}$. Let $\mathbf{p} \in S$ be a point and $(x(t), y(t), z(t))$ for $t \in I$ a solution of the Cauchy problem that consists of 2.4 and $(x(0), y(0), z(0)) = \mathbf{p}$. Then $(x(t), y(t), z(t)) \in S$ for all $t \in I$.*

Proof. Take $\mathbf{p} = (x_0, y_0, z_0)$ and define functions

$$\tilde{A}(x, y) = A(x, y, u(x, y)),$$

$$\tilde{B}(x,y) = B(x,y,u(x,y)).$$

Consider the two-dimensional autonomous system

$$\begin{cases} x' = \tilde{A}(x,y) \\ y' = \tilde{B}(x,y) \end{cases} \tag{2.5}$$

and let $(\tilde{x}(t), \tilde{y}(t))$ for $t \in J$ the solution of (2.5) such that

$$(\tilde{x}(0), \tilde{y}(0)) = (x_0, y_0)$$

and put $\tilde{z}(t) = u(\tilde{x}(t), \tilde{y}(t))$. Applying the chain rule (we withdraw the variable t to improve the readability of the formulas) we have

$$\begin{aligned} \tilde{z}' =& \frac{\partial u}{\partial x}(\tilde{x},\tilde{y})\,\tilde{x}' + \frac{\partial u}{\partial y}(\tilde{x},\tilde{y})\,\tilde{y}' \\ =& \frac{\partial u}{\partial x}(\tilde{x},\tilde{y})\,\tilde{A}(\tilde{x},\tilde{y}) + \frac{\partial u}{\partial y}(\tilde{x},\tilde{y})\,\tilde{B}(\tilde{x},\tilde{y}) \\ =& \frac{\partial u}{\partial x}(\tilde{x},\tilde{y})\,A(\tilde{x},\tilde{y},u(\tilde{x},\tilde{y})) + \frac{\partial u}{\partial y}(\tilde{x},\tilde{y})\,B(\tilde{x},\tilde{y},u(\tilde{x},\tilde{y})) \\ =& C(\tilde{x},\tilde{y},u(\tilde{x},\tilde{y})) = C(\tilde{x},\tilde{y},\tilde{z}). \end{aligned}$$

Since we already have

$$\tilde{x}' = A(\tilde{x},\tilde{y},u(\tilde{x},\tilde{y})) = A(\tilde{x},\tilde{y},\tilde{z})$$

$$\tilde{y}' = B(\tilde{x},\tilde{y},u(\tilde{x},\tilde{y})) = B(\tilde{x},\tilde{y},\tilde{z})$$

it turns out that $(\tilde{x}(t), \tilde{y}(t), \tilde{z}(t))$ is solution of 2.4 and since

$$(\tilde{x}(0), \tilde{y}(0), \tilde{z}(0)) = (x_0, y_0, u(x_0, y_0)) = (x_0, y_0, z_0),$$

the uniqueness of the solutions of 2.4 implies that

$$(\tilde{x}(t), \tilde{y}(t), \tilde{y}(t)) = (x(t), y(t), z(t))$$

for $t \in I \cap J$. In such a case, if I and J were maximal domains for the solution we would even have $I = J$. ■

In all what follows, we will refer the trajectories of the system 2.4 as *characteristic* lines of the equation 2.3.

Now, we will prove that the equation 2.3 actually has solutions. The idea is to produce a surface made from characteristic lines. For that aim, we will fix a curve $\gamma(s)$ for $s \in I$ contained in Ω and we consider the trajectories of 2.4 that start at the points of $\gamma(s)$. Namely, let $\Gamma(t,s)$ be a function of the parameters (t,s) such that:

(a) for every $s \in I$, $t \to \Gamma(t,s)$ is solution of 2.4,

(b) for every $s \in I$, $\Gamma(0,s) = \gamma(s)$.

Basic theorems in ODE's theory say that the dependence with respect initial conditions is $\mathcal{C}^k$ if the functions A, B, C are $\mathcal{C}^k$ for $k \geq 1$. If we assume γ is also $\mathcal{C}^k$, then Γ is $\mathcal{C}^k$ too. For the time being, we will only require $\mathcal{C}^1$ regularity.

Note that we have say nothing of the domain of Γ. At first, for different $s \in I$, the domain of t, which is a neighborhood of 0, may be different. In ODE's theory it is show that this domain is limited by Ω and the growth behavior of A, B, C near the boundary of Ω. In other words, if the functions A, B, C remain bounded, the domain of t is limited only by the size of Ω. Therefore, we may assume that Γ is defined on a rectangle $J \times I$, although this is not important because the domain of the solution u could be smaller.

Now, we are ready to present the existence and uniqueness theorem for the quasi-linear equation, whose statement evidently reminds of the Cauchy problem for ODEs.

Theorem 2.2.2. *Let $A, B, C : \Omega \to \mathbb{R}$ be $\mathcal{C}^1$ functions, $\gamma : I \to \Omega$ a $\mathcal{C}^1$ curve. Denote $\gamma(s) = (x(s), y(s), z(s))$ and assume that the transversality condition holds*

$$\begin{vmatrix} A(\gamma(s)) & B(\gamma(s)) \\ x'(s) & y'(s) \end{vmatrix} \neq 0.$$

Then there exists a unique $\mathcal{C}^1$ solution U of

$$A(x,y,u)\,\frac{\partial u}{\partial x}(x,y) + B(x,y,u)\,\frac{\partial u}{\partial y}(x,y) = C(x,y,u) \tag{2.6}$$

defined on a neighborhood of the XY projection of $\gamma(I)$ and whose graph contains γ in the sense

$$U(x(s), y(s)) = z(s).$$

Proof. The uniqueness is a consequence of Theorem 2.2.1, so we will focus on the construction of a solution. For $s \in I$, let $X(t,s), Y(t,s), Z(t,s)$ a solution of the Cauchy problem

$$\begin{cases} x' = A(x,y,z) \\ y' = B(x,y,z) \\ z' = C(x,y,z) \\ (x(0), y(0), z(0)) = \gamma(s) \end{cases}$$

and put $\Gamma(t,s) = (X(t,s), Y(t,s), Z(t,s))$. We will show that Γ defines a parameterized surface on a neighborhood of $(0,s)$. Indeed,

$$\left(\frac{\partial X}{\partial t}, \frac{\partial Y}{\partial t}, \frac{\partial Z}{\partial t}\right)(0,s) = (A(\gamma(s)), B(\gamma(s)), C(\gamma(s)))$$
$$\left(\frac{\partial X}{\partial s}, \frac{\partial Y}{\partial s}, \frac{\partial Z}{\partial s}\right)(0,s) = (x'(s), y'(s), z'(s))$$

The transversality condition implies for the Jacobian given by the two first coordinates

$$\begin{vmatrix} \frac{\partial X}{\partial t} & \frac{\partial X}{\partial s} \\ \frac{\partial Y}{\partial t} & \frac{\partial Y}{\partial s} \end{vmatrix} \neq 0$$

at every point of the form $(0,s)$. By continuity of the functions, that also happens on a neighborhood of the curve. That is enough to have the parameterized surface. Now we will show that the surface can be locally represented as the graph of a function. For that aim, the inverse of the function $(t,s) \to (X(t,s), Y(t,s))$ will be denoted

$$t = T(x,y); \quad s = S(x,y).$$

Define by $U(x,y) = Z(T(x,y), S(x,y))$. We will prove that U satisfies (2.6). From now on we will drop the variables of the functions during computations. First of all, note that

$$\begin{pmatrix} \frac{\partial T}{\partial x} & \frac{\partial T}{\partial y} \\ \frac{\partial S}{\partial x} & \frac{\partial S}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial X}{\partial t} & \frac{\partial X}{\partial s} \\ \frac{\partial Y}{\partial t} & \frac{\partial Y}{\partial s} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

at corresponding points. The chain rule gives

$$\frac{\partial U}{\partial x} = \frac{\partial Z}{\partial t}\frac{\partial T}{\partial x} + \frac{\partial Z}{\partial s}\frac{\partial S}{\partial x};$$

$$\frac{\partial U}{\partial y} = \frac{\partial Z}{\partial t}\frac{\partial T}{\partial y} + \frac{\partial Z}{\partial s}\frac{\partial S}{\partial y}.$$

The substitution in (2.6) leads to

$$A\frac{\partial U}{\partial x} + B\frac{\partial U}{\partial y} = \frac{\partial Z}{\partial t}\left(A\frac{\partial T}{\partial x} + B\frac{\partial T}{\partial y}\right) + \frac{\partial Z}{\partial s}\left(A\frac{\partial S}{\partial x} + B\frac{\partial S}{\partial y}\right)$$

The first bracket equals 1 and the second one 0, as $A = \frac{\partial X}{\partial t}$ and $B = \frac{\partial Y}{\partial t}$. Therefore

$$A\frac{\partial U}{\partial x} + B\frac{\partial U}{\partial y} = \frac{\partial Z}{\partial t} = C,$$

That means equation (2.6) is satisfied. Do not forget that the variables should match too. For instance

$$A = A(x, y, Z(T(x, y), S(x, y))).$$

The solution U was built locally around a point of the form $(0, s)$. The uniqueness, allows to "glue" a finite amount of solutions that overlap and that covers the XY projection of a set of the form $\Gamma([-a, a] \times I)$ with $a > 0$, by a simple compactness argument. ■

Example 2.2.3. *Find the solution of*

$$xy(p - q) = (x - y)z$$

that contains the curve $y^2 + z^2 = x^2$, $z = 1$.

The characteristic system is

$$\begin{cases} x' = xy \\ y' = -xy \\ z' = (x - y)z \end{cases}$$

From the first two equations we get that $x + y = \lambda$ is constant. Using $y = \lambda - x$ in the first equation we have

$$x' = x(\lambda - x)$$

that can be integrated by separation of variables

$$\int dt = \int \frac{dx}{x(\lambda - x)} = \frac{1}{\lambda}\int\left(\frac{1}{x} + \frac{1}{\lambda - x}\right)dx = \frac{1}{\lambda}\log\left(\frac{x}{\lambda - x}\right)$$

leading to

$$\frac{x}{\lambda - x} = \eta\, e^{\lambda t}$$

where η is another constant. Solving the relation for x and then for y we have

$$x = \frac{\lambda\eta\, e^{\lambda t}}{1 + \eta\, e^{\lambda t}}; \quad y = \frac{\lambda}{1 + \eta\, e^{\lambda t}}.$$

Now

$$z' = \left(\frac{\lambda\eta\, e^{\lambda t} - \lambda}{1 + \eta\, e^{\lambda t}}\right) z,$$

that leads to

$$\int \frac{dz}{z} = \int \frac{\lambda\eta\, e^{\lambda t}}{1 + \eta\, e^{\lambda t}} dt - \int \frac{\lambda}{1 + \eta\, e^{\lambda t}} dt$$

$$= \int \frac{\lambda\eta\, e^{\lambda t}}{1 + \eta\, e^{\lambda t}} dt - \int \frac{\lambda\, e^{-\lambda t}}{e^{-\lambda t} + \eta} dt = \log(1 + \eta\, e^{\lambda t}) + \log(e^{-\lambda t} + \eta).$$

We deduce that

$$z = \gamma(1 + \eta\, e^{\lambda t})(e^{-\lambda t} + \eta) = \gamma(1 + \eta\, e^{\lambda t})^2 e^{-\lambda t}.$$

It would be unnecessarily complicated to use the solutions of the characteristic system with the parameter t. Since we just need to build the solution surface from the characteristic curves we could use any other more convenient variable and parameters. Note that

$$xyz = \lambda^2 \eta\gamma$$

which is constant. Actually, we could get this relation by noticing a while before that

$$(xyz)' = x'yz + xy'z + xyz' = xyyz - xxyz + xy(x - y)z = 0.$$

The given curve can be parameterized as $x = \cosh s$, $y = \sinh s$ and $z = 1$. Since $x + y$ should be constant along the characteristic curves, we put

$$X = \cosh s + r \quad \text{and} \quad Y = \sinh s - r,$$

where r is a new and simpler variable to move along the each curve. With all these, we have

$$Z = \frac{\cosh s \sinh s}{(\cosh s + r)(\sinh s - r)}.$$

In order to eliminate the parameters, notice that

$$X + Y = \cosh s + \sinh s = e^s$$

$$XYZ = \cosh s \sinh s = \frac{e^{2s} - e^{-2s}}{4}$$

and thus

$$Z = \frac{1}{4XY}\left((X+Y)^2 - \frac{1}{(X+Y)^2}\right).$$

Therefore, the solution of the problem is

$$u(x,y) = \frac{1}{4xy}\left((x+y)^2 - \frac{1}{(x+y)^2}\right).$$

■

Remark 2.2.4. *If for some reason, we have obtained the solution of the characteristic system 2.4 as the intersecction of two families of surfaces*

$$\begin{cases} f(x,y,z) = \lambda \\ g(x,y,z) = \eta \end{cases}$$

then, for any regular function $\Phi(\lambda, \eta)$, the equation

$$\Phi(f(x,y,z), g(x,y,z)) = 0$$

describes a solution of 2.3. *This method applies to the Cauchy problem too.*

We left the reader to reader the task of solving the problem of Example 2.2.3 following the directions of the remark.

2.3 Differential equations in differential form

Let ω be a 1-differential form in $\mathbb{R}^n$ and consider the equation $\omega = 0$. The general solution can be expressed as the $(n-1)$-dimensional manifolds given by function of n variables that equals an arbitrary constant. Indeed, for $n = 1$ the solution are constants. For $n = 2$ the equation

$$A(x,y)\,dx + B(x,y)\,dy = 0 \tag{2.7}$$

is equivalent to the ordinary differential equation

$$\frac{dy}{dx} = -\frac{A(x,y)}{B(x,y)}$$

whose solution can be implicitly expressed as $F(x,y) = c$ with $c \in \mathbb{R}$ constant. On the other hand, differentiating $F(x,y) = c$ we get

$$\frac{\partial F}{\partial x}(x,y)\,dx + \frac{\partial F}{\partial y}(x,y)\,dy = 0$$

which should be proportional to equation (2.7) (otherwise, $dx = dy = 0$). The proportionality factor is a function $\mu(x,y)$ called integrating factor. We will provide a more general definition.

Definition 2.3.1. *Given a 1-form ω in $\mathbb{R}^n$, an integrating factor is a function μ such that $\mu\,\omega$ is exact, at least, locally. In such a case, if F is a (local) primitive of $\mu\,\omega$, the solutions of $\omega = 0$ can be expressed as manifolds $F = c$, with $c \in \mathbb{R}$ constant.*

In $\mathbb{R}^2$ (dimension 2) the integrating factor always exists, because the equation $\omega = 0$ is equivalent to an ODE, that under very general hypotheses, always has solution. However, for dimensions $n \geq 3$, the existence of the integrating factor depends on a partial differential equation that may have no solution. We will prove a necessary condition.

Proposition 2.3.2. *Let ω be an 1-differential form of class $\mathcal{C}^1$ on $\mathbb{R}^n$ for $n \geq 3$. If ω has an integrating factor, then $\omega \wedge d\omega = 0$.*

Proof. If $\mu\,\omega = dF$ for some nontrivial functions μ, F, we have μ is $\mathcal{C}^1$ and $\mu \neq 0$ except, maybe, a set of empty interior. Differentiating the equality, we get

$$d\mu \wedge \omega + \mu\,d\omega = 0$$

since $d^2F = 0$ for the exterior differential. Multiply exteriorly by ω and observe that

$$(d\mu \wedge \omega) \wedge \omega + \mu\,(d\omega \wedge \omega) = 0.$$

Since $(d\mu \wedge \omega) \wedge \omega = 0$ by the repetition of a factor, we deduce $\mu\,(d\omega \wedge \omega) = 0$, and thus $\omega \wedge d\omega = 0$. ∎

In particular, for an 1-form $Adx+Bdy+Cdz$ on $\mathbb{R}^3$, the condition $\omega\wedge d\omega = 0$ is explicitly written as

$$A\left(\frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}\right) + B\left(\frac{\partial A}{\partial z} - \frac{\partial C}{\partial x}\right) + C\left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\right) = 0, \tag{2.8}$$

that is easier to remember as $\vec{W}\cdot(\nabla\times\vec{W}) = 0$, where $\vec{W} = (A, B, C)$. With that notation a "classic vector analysis" interpretation is possible. Indeed, the solutions of $Adx + Bdy + Cdz = 0$, in case it solvable, is the family of orthogonal surfaces to the lines given by the system

$$\frac{dx}{A} = \frac{dy}{B} = \frac{dz}{C}. \tag{2.9}$$

In order to see that, the best way is to think of (dx, dy, dz) as the allowed displacements within the solution (tangent space). Therefore, the solutions of $Adx + Bdy + Cdz = 0$ are orthogonal to the field $\vec{W}$ meanwhile the solution of the system (2.9) are aligned with it. It is not difficult to obtain the necessary condition Proposition 2.3.2 in dimension 3 from Stokes theorem.

Now we will show that the necessary condition given in Proposition 2.3.2 is also sufficient in dimension 3 (actually, Frobenius proved that for any dimension greater than 2). Moreover, the method of proof will provide a way to solve equation $Adx + Bdy + Cdz = 0$.

Theorem 2.3.3. *If the functions $A(x, y, z), B(x, y, z)$ and $C(x, y, z)$ are $\mathcal{C}^1$ and satisfy the equality (2.8), then the equation*

$$A(x, y, z)\,dx + B(x, y, z)\,dy + C(x, y, z)\,dz = 0 \tag{2.10}$$

can be solved locally.

Proof. We will use a version of the so called "variation of constants method". Consider z as a constant parameter and find a solution of the equation

$$A(x, y, z)\,dx + B(x, y, z)\,dy = 0$$

of the form $G(x, y, z) = g$ with g constant. Now we claim that it is possible to find a solution of the original equation of the form $G(x, y, z) = g(z)$. In such a case, writing the condition as $G(x, y, z) - g(z) = 0$, differentiation would give the expression

$$\frac{\partial G}{\partial x}(x, y, z)\,dx + \frac{\partial G}{\partial y}(x, y, z)\,dy + \left(\frac{\partial G}{\partial z}(x, y, z) - g'(z)\right)\,dz.$$

We wish to identify that equation with (2.10) the first two coefficients are related by a multiplicative factor. Without loss of generality we may assume that

$$A(x, y, z) = \frac{\partial G}{\partial x}(x, y, z); \quad B(x, y, z) = \frac{\partial G}{\partial y}(x, y, z);$$

that is equivalent to multiply by some integrating factor (with respect to x, y only). Indeed, the validity of (2.8) is not modified by multiplication by a scalar function. Indeed, if ν is a scalar function and ω an 1-form

$$(\nu\,\omega) \wedge d(\nu\,\omega) = \nu\,\omega \wedge (d\nu \wedge \omega) + \nu\,\omega \wedge d\omega = \nu\,\omega \wedge d\omega.$$

so the first term is null if and only if the last one is.
Therefore, once we assume that $A = G_x$ and $B = G_y$, we need the third term to satisfy

$$C(x, y, z) = \frac{\partial G}{\partial z}(x, y, z) - g'(z)$$

which means $G_z - C = g'(z)$ should be a function of z only, that is, it does not contain x, y explicitly. In order to prove that, if we express A and B in terms of G in the equality 2.8 we get

$$\begin{aligned} 0 &= \frac{\partial G}{\partial x}\left(\frac{\partial C}{\partial y} - \frac{\partial^2 G}{\partial y \partial z}\right) + \frac{\partial G}{\partial y}\left(\frac{\partial^2 G}{\partial x \partial z} - \frac{\partial C}{\partial x}\right) + C\left(\frac{\partial^2 G}{\partial y \partial x} - \frac{\partial^2 G}{\partial x \partial y}\right) \\ &= \frac{\partial G}{\partial x}\frac{\partial}{\partial y}(C - G_z) - \frac{\partial G}{\partial y}\frac{\partial}{\partial x}(C - G_z) = \frac{\partial(G, C - G_z)}{\partial(x, y)} \end{aligned}$$

The fact that the last Jacobian equals 0 implies that G and $G_z - C$ are functionally dependent as functions of x, y, that is, when z remains constant, namely $C - G_z = \Phi(G, z)$. Since x, y, z are binded by $G(x, y, z) = g(z)$, the function g should satisfy the differential equation

$$g'(z) = \Phi(g(z), z)$$

whose integration will provide g. With all that, the solutions of the original problem are of the form $G(x, y, z) - g(z) = c$ constant. ■

The direct computation of (2.8) provides this criterion that will be applied later.

Corollary 2.3.4. *Let p, q be C^1 functions of x, y, z. Then the equation*

$$dz = p\,dx + q\,dy$$

is integrable if and only if

$$\frac{\partial q}{\partial x} + p\,\frac{\partial q}{\partial z} = \frac{\partial p}{\partial y} + q\,\frac{\partial p}{\partial z}.$$

Note that the criterion is formally the same that the condition for $p\,dx+q\,dy$ being an exact form provided that z is a function of x and y with

$$\frac{\partial z}{\partial x} = p \quad \text{and} \quad \frac{\partial z}{\partial y} = q.$$

Example 2.3.5. *Prove that the following equation is solvable and find the solution*

$$(x-r)\,dx + (y-r)\,dy + (z-r)\,dz = 0,$$

where $r = (x^2+y^2+z^2)^{1/2}$.

The application of (2.8) and the simplification of a common factor give

$$(x-r)(z-y) + (y-r)(x-z) + (z-r)(y-x) = 0$$

which means after Theorem 2.3.3 that the equation is solvable. We will follow the ideas of the proof. Consider the differential form

$$(x-r)\,dx + (y-r)\,dy$$

which is not exact. However, when divided by r becomes exact

$$(\frac{x}{r}-1)\,dx + (\frac{y}{r}-1)\,dy$$

as it can be easily checked. Clearly, $r-x-y$ is primitive. Now we will look for a solution of the given equation of the form

$$r-x-y = g(z).$$

Differentiation gives

$$(\frac{x}{r}-1)\,dx + (\frac{y}{r}-1)\,dy - (\frac{z}{r}-g'(z))\,dz = 0$$

being the to first term anything but a surprise. Multiplying by r leads to

$$(x-r)\,dx + (y-r)\,dy + (z-rg'(z))\,dz = 0,$$

and the comparison with the given equation implies that the equality

$$z - rg'(z) = z - r$$

have to be fulfilled. Note that it reduces to $g'(z) = 1$, disappearing explicitly x and y. Taking $g(z) = z$ we arrive to the solution of the given equation

$$r - x - y - z = c$$

where c is constant. ■

Sometimes, to follow the steps of the proof of Frobenius Theorem 2.3.3 to integrate a differential expression is not so obvious. For that, we provide a second example.

Example 2.3.6. *During the resolution of a first order PDE we arrive to the following integrable differential expression*

$$dz = \frac{z\,dx}{\sqrt{1+\lambda^2}} + \frac{\lambda z\,dy}{\sqrt{1+\lambda^2}},$$

where λ is a parameter. Find the solution.

If z was constant, then we would have

$$\frac{z\,dx}{\sqrt{1+\lambda^2}} + \frac{\lambda z\,dy}{\sqrt{1+\lambda^2}} = 0$$

whose evident solution is

$$\frac{z(x+\lambda y)}{\sqrt{1+\lambda^2}} = g$$

being g constant. If we put $g(z)$ instead of g and differentiate, we arrive to

$$\frac{(x+\lambda y)\,dz}{\sqrt{1+\lambda^2}} + \frac{z\,dx}{\sqrt{1+\lambda^2}} + \frac{\lambda z\,dy}{\sqrt{1+\lambda^2}} = g'(z).$$

Comparison with the original expression, leads to

$$g'(z) - \frac{x+\lambda y}{\sqrt{1+\lambda^2}} = 1$$

that seems to be not solvable since the theorem predicts that it should remain a function that depends only on z. The non solvability is just an appearance. Indeed, using the definition of g,

$$g(z) = \frac{z(x + \lambda y)}{\sqrt{1 + \lambda^2}},$$

we obtain by substitution that

$$g'(z) - \frac{g(z)}{z} = 1,$$

whose particular solution $g(z) = z \log z$ leads to

$$\log z = \frac{x + \lambda y}{\sqrt{1 + \lambda^2}} + \eta$$

with η a constant. Note that we would arrive to the same solution by dividing the original expression by z

$$\frac{dz}{z} = \frac{dx}{\sqrt{1 + \lambda^2}} + \frac{\lambda\, dy}{\sqrt{1 + \lambda^2}},$$

that turns out to be exact. ■

2.4 Parametric families of surfaces and their envelopes

Here we will consider families of surfaces that depends on one or two parameters. We will assume that the surfaces are given in implicit form by

$$f(x, y, z, \lambda) = 0 \text{ or } f(x, y, z, \lambda, \eta) = 0,$$

with standard regularity assumptions on f. The first type of family is called 1-parametric, the second one 2-parametric. The *envelope* of a parameterized family of surfaces is a surface that is is tangent at every of its point to a surface from the family, however it does not belong to the parametric family. Whilst members from a 2-parametric family are tangent to its envelope at only one point, 1-parametric families are tangent to its envelope on a curve. We shall assume that the assignation $(x, y, z) \to \lambda$ or $(x, y, z) \to (\lambda, \eta)$ that takes every point of the envelope to the corresponding parameter values of the tangent member of the family is smooth.

By analogy for the envelopes of curves in the plane, the envelope can be obtained by eliminating the parameter λ from the system

$$\begin{cases} f(x,y,z,\lambda) = 0, \\ \dfrac{\partial f}{\partial \lambda}(x,y,z,\lambda) = 0; \end{cases} \tag{2.11}$$

for a 1-parametric family: Indeed, consider any point (x_0, y_0, z_0) of the envelope. For some λ_0 that point also belongs to the surface $f(x,y,z,\lambda_0) = 0$, which is tangent to the envelope. Assume that you can parameterize a smooth curve $(x(\lambda), y(\lambda), z(\lambda))$ contained passing by our point for $\lambda = \lambda_0$ and such that $(x(\lambda), y(\lambda), z(\lambda))$ belongs to $f(x,y,z,\lambda) = 0$ (this can be justified using the smooth dependence with respect to λ). Derivation with respect to λ gives

$$\frac{\partial f}{\partial x}\frac{dx}{d\lambda} + \frac{\partial f}{\partial y}\frac{dy}{d\lambda} + \frac{\partial f}{\partial z}\frac{dz}{d\lambda} + \frac{\partial f}{\partial \lambda} = 0.$$

On the other hand, the vector

$$\left(\frac{dx}{d\lambda}(\lambda), \frac{dy}{d\lambda}(\lambda), \frac{dz}{d\lambda}(\lambda)\right)$$

is contained in the tangent plane at $(x(\lambda), y(\lambda), z(\lambda))$ to the envelope, and so to the tangent plane to $f(x,y,z,\lambda) = 0$. Therefore,

$$\frac{\partial f}{\partial x}\frac{dx}{d\lambda} + \frac{\partial f}{\partial y}\frac{dy}{d\lambda} + \frac{\partial f}{\partial z}\frac{dz}{d\lambda} = 0.$$

In particular, for $\lambda = \lambda_0$ we have

$$\frac{\partial f}{\partial \lambda}(x_0, y_0), z_0), \lambda_0) = 0.$$

That justifies the given method 2.11 to find the envelope.

Analogously, for a 2-parametric family $f(x,y,z,\lambda,\eta) = 0$, the envelope, if it exists, could be obtained by eliminating λ, η from the system

$$\begin{cases} f(x,y,z,\lambda,\eta) = 0, \\ \dfrac{\partial f}{\partial \lambda}(x,y,z,\lambda,\eta) = 0, \\ \dfrac{\partial f}{\partial \eta}(x,y,z,\lambda,\eta) = 0. \end{cases} \tag{2.12}$$

Note that the expression of (x,y,z) in terms of (λ,η) also gives the envelope as a parameterized surface.

Example 2.4.1. *Find the envelope of the* 2*-parametric family of planes*

$$\frac{x}{\lambda} + \frac{y}{\eta} + \frac{z}{\xi} = 1$$

where $\lambda^2 + \eta^2 + \xi^2 = 1$.

Firstly, we need to express the family explicitly by two parameters

$$\frac{x}{\lambda} + \frac{y}{\eta} + \frac{z}{\sqrt{1-\lambda^2-\eta^2}} = 1.$$

Partial derivation with respect to λ and η gives

$$-\frac{x}{\lambda^2} + \frac{\lambda z}{(1-\lambda^2-\eta^2)^{3/2}} = 0, \quad \text{and} \quad -\frac{y}{\eta^2} + \frac{\eta z}{(1-\lambda^2-\eta^2)^{3/2}} = 0.$$

Using ξ again, those equalities can be written

$$\frac{x}{\lambda^3} = \frac{y}{\eta^3} = \frac{z}{\xi^3} = \gamma$$

where γ is new. We deduce

$$\gamma = \gamma(\lambda^2 + \eta^2 + \xi^2) = \frac{x}{\lambda} + \frac{y}{\eta} + \frac{z}{\xi} = 1,$$

and so

$$\lambda^3 = x; \quad \eta^3 = y; \quad \xi^3 = z.$$

The substitution in $\lambda^2 + \eta^2 + \xi^2 = 1$ gives the equation of the envelope

$$x^{2/3} + y^{2/3} + z^{2/3} = 1$$

as wished. ∎

The following result will be useful later in order to build solutions of first order PDEs containing a given curve.

Proposition 2.4.2. *Let $f(x, y, z, \lambda) = 0$ be a 1-parametric family of surfaces and $\gamma(s) = (x(s), y(s), z(s))$ a curve such that for every s_0 there is a unique λ_0 so γ is tangent to $f(x, y, z, \lambda_0) = 0$ at $\gamma(s_0)$. If the dependence between s and λ is bijective and smooth, then γ is contained in the envelope of $f(x, y, z, \lambda) = 0$.*

Proof. Since the assignation $s \to \lambda$ is one-to-one and smooth, without loss of generality we may assume that λ parameterizes γ. Differentiating the equality

$$f(x(\lambda), y(\lambda), z(\lambda), \lambda) = 0$$

we get

$$\nabla f(x(\lambda), y(\lambda), z(\lambda), \lambda) \cdot \gamma'(\lambda) + \frac{\partial f}{\partial \lambda}(x(\lambda), y(\lambda), z(\lambda), \lambda) = 0.$$

By the hypothesis, the first term is 0. Therefore, the curve γ satisfies the system of equations

$$\begin{cases} f(x(\lambda), y(\lambda), z(\lambda), \lambda) = 0, \\ \dfrac{\partial f}{\partial \lambda}(x(\lambda), y(\lambda), z(\lambda), \lambda) = 0; \end{cases}$$

that defines the envelope. We conclude that γ is contained in the envelope of the family $f(x, y, z, \lambda) = 0$ as claimed. ∎

Note that under the assumptions of Proposition 2.4.2 the system 2.11 has non trivial solution. Then, it is quite easy to add conditions to guarantee the actual existence of the envelope, at least locally.

The surfaces of a 2-parametric family satisfies a PDE of first order. Indeed, it is enough to eliminate the parameters from the following system

$$\begin{cases} f(x, y, z, \lambda, \eta) = 0, \\ f_x(x, y, z, \lambda, \eta) + p\, f_z(x, y, z, \lambda, \eta) = 0, \\ f_y(x, y, z, \lambda, \eta) + q\, f_z(x, y, z, \lambda, \eta) = 0; \end{cases} \tag{2.13}$$

Indeed, we have differentiated z as a function of x, y, so its partial derivatives p, q appear. The elimination of λ, η leads to an equation of the form $F(x, y, z, p, q) = 0$. Note that a 1-parametric family also satisfies a first order PDE, but it is no uniquely determined. Therefore, if a 2-parametric family is contained in the set of solutions of a PDE, we may recover the PDE from the family. In other words, a 2-parametric family of solutions of a PDE contains the same information that the PDE or its *general solution*.

Example 2.4.3. *Find the* 2*-parametric family of planes which are tangent to the sphere* $x^2 + y^2 + z^2 = 1$*. Then find the partial differential equation the family satisfies and the envelope of the* 1*-parametric family made up of those planes that tangent to the sphere on any parallel line "equator" (z constant).*

We will use the spheric coordinates θ, ϕ, so the point in the sphere and the unitary normal vector to the plane, which are the same, can be expressed

$$(\cos\theta\cos\phi, \sin\theta\cos\phi, \sin\phi).$$

The tangent plane at such a point is

$$\cos\theta\cos\phi\, x + \sin\theta\cos\phi\, y + \sin\phi\, z = 1.$$

Derivation with respect to x and y gives

$$\cos\theta\cos\phi + \sin\phi\, p = 0,$$

$$\sin\theta\cos\phi + \sin\phi\, q = 0.$$

These three equations together can considered a linear system with unknown variables $\cos\theta\cos\phi, \sin\theta\cos\phi, \sin\phi$. Cramer's rule gives the solution

$$\cos\theta\cos\phi = \frac{-p}{z - xp - yq},$$

$$\sin\theta\cos\phi = \frac{-q}{z - xp - yq},$$

$$\sin\phi = \frac{1}{z - xp - yq}.$$

Since these three expressions to the square add up to 1, we get the equality

$$p^2 + q^2 + 1 = (z - xp - yq)^2$$

that is the differential equation satisfied by the family of planes. Observe that the sphere $x^2 + y^2 + z^2 = 1$ satisfies also the equation. It can be checked that it is actually the envelope of the whole family.
The admisible values of z to define a parallel line are those corresponding to $\phi \in [-\pi/2, \pi/2]$ in spheric coordinates. Consider the equation

$$\cos\theta\cos\phi\, x + \sin\theta\cos\phi\, y + \sin\phi\, z = 1$$

and derive it with respect to the parameter we want to eliminate θ. We get

$$-\sin\theta\cos\phi\, x + \cos\theta\cos\phi\, y = 0,$$

form which we have

$$\tan\theta = \frac{y}{x}.$$

Therefore,

$$\cos\theta = \frac{x}{\sqrt{x^2+y^2}}, \qquad \sin\theta = \frac{y}{\sqrt{x^2+y^2}}.$$

The substitution in the first equation gives

$$\frac{\cos\phi\, x^2}{\sqrt{x^2+y^2}} + \frac{\cos\phi\, y^2}{\sqrt{x^2+y^2}} + \sin\phi\, z = 1,$$

that is equivalent to

$$\cos\phi\sqrt{x^2+y^2} = 1 - \sin\phi\, z$$

which says that the envelope is the half of a revolution cone touching the sphere at the parallel line, as intuitively was expected. ■

Consider a PDE

$$F(x, y, z, p, q) = 0. \tag{2.14}$$

A 2-parametric family of solutions $f(x, y, z, \lambda, \eta) = 0$ is called a *complete solution* of 2.14. Indeed, we have seen before that the family of solutions contains all the information of the PDE. The role of complete solutions in producing the general solution is explained by the following result.

Proposition 2.4.4. *Assume that a family of surfaces, either* 1*-parametric or* 2*-parametric, satisfy 2.14. Then the envelope of the family satisfies 2.14 too.*

Proof. Note that the relation 2.14 binds a point (x, y, z) with a set of feasible directions (p, q) of the tangent plane at it. Since the envelope is tangent at any of its points (x, y, z) to a surface from the family, they share the quintuple (x, y, z, p, q) and so relation 2.14 is fulfilled by the envelope too. ■

As we will see later, particular solutions of the Cauchy problem are obtained as envelopes of 1-parametric families selected from a complete solution. However, the envelope of the whole complete solution as a 2-parametric family, is also a solution of the PDE normally not contained in the general solution. For that reason, we call it a *singular solution*.

2.5 Geometric meaning of first order equations

The general form of the first order PDE is

$$f(x, y, u, u_x, u_y) = 0.$$

We will go on with the geometrical notation introduced for the quasi-linear equation. Let $u(x, y)$ a solution and put $z_0 = u(x_0.y_0)$, $p_0 = u_x(x_0, y_0)$ and $q_0 = u_y(x_0, y_0)$. The equation of the tangent plane to the surface $z = u(x, y)$ at (x_0, y_0, z_0) is

$$z - z_0 = p_0(x - x_0) + q_0(y - y_0).$$

Now we have to walk in the opposite direction. We wish to construct a "piece" of solution around a point (x_0, y_0, z_0), so we need to chose coefficients of the tangent plane among those numbers p and q that satisfies the equation

$$f(x_0, y_0, z_0, p, q) = 0.$$

Since p and q are binded by the equation, we have a 1-parametric family of planes. Without loss of generality we may assume q is function of p, so the planes are

$$z - z_0 = p(x - x_0) + q(p)(y - y_0).$$

Since all these planes share the point (x_0, y_0, z_0), typically, the envelope of the family will be a "cone", named after *Gaspard Monge*. In the particular case of the quasi-linear equation the Monge cone collapses into a line.

Therefore, to build a piece of solution of the first order PDE around a point (x_0, y_0, z_0) we have to chose a plane tangent to the Monge cone. As we did with the quasi-linear equation, we may ask the solution to contain a curve $(x(s), y(s), z(s))$. However, this not enough because it could be several coherent choices for the tangent plane coefficients $(p(s), q(s))$. Indeed, If the solution of $F(x, y, z, p, q) = 0$ contains the curve $(x(s), y(s), z(s))$, then the normal vector at $(x(s), y(s), z(s))$ has the form $(p, q, -1)$ (recall. the five numbers $(x(s), y(s), z(s), p, q)$ tied by the PDE) and should be normal to the tangent vector to the curve $(x'(s), y'(s), z'(s))$. That is, p and q should satisfy the system of equations

$$\begin{cases} F(x(s), y(s), z(s), p, q) = 0, \\ x'(s)p + y'(s)q - z'(s) = 0; \end{cases}$$

that may have one, several or none solutions for every s.

The parameterized object

$$(x(s), y(s), z(s), p(s), q(s))$$

satisfying

$$f(x(s), y(s), z(s), p(s), q(s)) = 0$$

is called a *strip*. Given an initial strip, the solution of the Cauchy problem is unique locally for f regular enough. The geometric method to address the general first order equation, with some more computations, would lead to the autonomous system

$$\begin{cases} x' = F_p, \\ y' = F_q, \\ z' = pF_p + qF_q, \\ p' = -F_x - pF_z, \\ q' = -F_y - qF_z. \end{cases}$$

whose solutions are called *characteristic strips*. Note that for the quasi-linear equation it reduces to the system 2.4. The method of solution that we will follow in next section use the notion of complete solution defined in the previous section.

2.6 Solution of the first order equation

Consider the first order equation written in geometrical implicit form

$$F(x, y, z, p, q) = 0. \tag{2.15}$$

Here we will explain the method of *Lagrange-Charpit* to solve (2.15), that consist roughly speaking in the reduction to the equations studied in section 3

$$dz = p\,dx + q\,dy.$$

Eventually, this method could be formulated as a local existence theorem, like we did with the quasi-linear equation. However, the construction is more delicate and the uniqueness needs some particular considerations. Therefore, we will interested only in the practical aspects of the method, that we will lead to explicit complete solution of (2.15) and to the solution of the Cauchy problem with one more step. We will implicitly assume all the necessary regularity in the functions involved to justify the applied operations. The method of Lagrange-Charpit consists of the following steps:

1. Assume that we have found a function G different from F such that the quantity
$$G(x, y, z, p, q)$$
remains constant on every solution of (2.15). How to find G is explained later.

2. Then solve the system
$$\begin{cases} F(x, y, z, p, q) = 0, \\ G(x, y, z, p, q) = \lambda, \end{cases} \tag{2.16}$$
in order to get p, q as functions of x, y, z, λ.

3. Now, solve the equation
$$dz = p\,dx + q\,dy, \tag{2.17}$$
so its solution depends on another parameter η and can be expressed as
$$f(x, y, z, \lambda, \eta) = 0.$$
By construction, for every value of the parameters λ, η, the equation defines implicitely z as a function of x, y, that is, a complete solution.

4. The trick to find the function G is that (2.17) has to be integrable. The Frobenius criterion Corollary 2.3.4 applied to (2.17) will lead to an expression of G as the solution of a quasi-linear equation in five variables. The associated autonomous system can be written as
$$\frac{dx}{F_p} = \frac{dy}{F_q} = \frac{dz}{pF_p + qF_q} = \frac{dp}{-F_x - pF_z} = \frac{dq}{-F_y - qF_z}.$$
Note that the system depends on five constants, but we only need one nontrivial dependence among the variables x, y, z, p, q.

5. At this point, we already have a complete solution of (2.15). If we wish, moreover, to solve the Cauchy-like problem making the solution to contain a given curve
$$\gamma(s) = (x(s), y(s), z(s)),$$
then the solution is obtained as the envelope of the 1-parametric family obtained as
$$f(x, y, z, \lambda(s), \eta(s)) = 0$$
where $\eta(s), \lambda(s)$ are determined as follows: firstly, we should have
$$f(x(s), y(s), z(s), \lambda(s), \eta(s)) = 0,$$
that is the point of the curve meets the surface for the corresponding values of the parameters; secondly, at every point of the curve $\gamma(s)$, the tangent vector $\gamma'(s)$ must be perpendicular to the normal (to the surface) vector $\nabla f(x(s), y(s), z(s))$. That is,
$$\frac{\partial f}{\partial x}(x(s), y(s), z(s), \lambda, \eta)x'(s) + \frac{\partial f}{\partial y}(x(s), y(s), z(s), \lambda(s), \eta(s))y'(s)$$
$$+\frac{\partial f}{\partial z}(x(s), y(s), z(s), \lambda(s), \eta(s))z'(s) = 0.$$
Now, using both equations we should get a smooth determination of the pair $(\lambda(s), \eta(s))$.

Now we will justify the steps. The equation in step 3 has to be solvable. That determines some conditions that G must satisfy. Indeed, recall that the Frobenius criterion Corollary 2.3.4 implies that the following equality must hold
$$\frac{\partial q}{\partial x} + p\,\frac{\partial q}{\partial z} = \frac{\partial p}{\partial y} + q\,\frac{\partial p}{\partial z}.$$
However, the functions p, q are given implicitly. Since we need their partial derivatives, we can get them by implicit derivation of F and G with respect to x, y, z. For instance, derivation with respect x gives
$$\begin{cases} F_x + F_p\,\dfrac{\partial p}{\partial x} + F_q\,\dfrac{\partial q}{\partial x} = 0, \\ G_x + G_p\,\dfrac{\partial p}{\partial x} + G_q\,\dfrac{\partial q}{\partial x} = 0. \end{cases} \tag{2.18}$$

Now we can obtain $\frac{\partial q}{\partial x}$ by Cramer's rule

$$\frac{\partial q}{\partial x} = \frac{F_x G_p - F_p G_x}{F_p G_q - F_q G_p}.$$

Analogously, by derivation with respect to y and z and solving we shall obtain

$$\frac{\partial p}{\partial y} = \frac{F_q G_y - F_y G_q}{F_p G_q - F_q G_p},$$

$$\frac{\partial p}{\partial z} = \frac{F_q G_z - F_z G_q}{F_p G_q - F_q G_p},$$

$$\frac{\partial q}{\partial z} = \frac{F_z G_p - F_p G_z}{F_p G_q - F_q G_p}.$$

Going back to Frobenius criterion, the following identity must hold

$$\frac{F_x G_p - F_p G_x}{F_p G_q - F_q G_p} + p\frac{F_z G_p - F_p G_z}{F_p G_q - F_q G_p} = \frac{F_q G_y - F_y G_q}{F_p G_q - F_q G_p} + q\frac{F_q G_z - F_z G_q}{F_p G_q - F_q G_p}.$$

We can get rid of the common denominator, and having in mind that the unknown function is G, we group the terms this way (we also change the notation for the partial derivatives of G in order to highlight them)

$$F_p\frac{\partial G}{\partial x} + F_q\frac{\partial G}{\partial y} + (pF_p + qF_q)\frac{\partial G}{\partial z} - (F_x + pF_z)\frac{\partial G}{\partial p} - (F_y + qF_z)\frac{\partial G}{\partial q} = 0.$$

That equality is a quasi-linear equation in 5 variables x, y, z, p, q that our mysterious function G should satisfy. That leads to the following autonomous system in $\mathbb{R}^5$

$$\begin{cases} x' = F_p, \\ y' = F_q, \\ z' = pF_p + qF_q, \\ p' = -F_x - pF_z, \\ q' = -F_y - qF_z. \end{cases} \tag{2.19}$$

Note that we only need a 1-parametric family of solutions, so it is enough to find a relation among x, y, z, p, q different from F, but this system is actually the same that appears when the equation (2.15) is solved by the method of characteristic strips.

The justification of the last part to solve the Cauchy problem is a consequence of Proposition 2.4.2. Indeed, the choice of parameters $\lambda(s), \eta(s)$ implies that the curve γ is tangent to the surface

$$f(x, y, z, \lambda(s), \eta(s)) = 0$$

at $\gamma(s)$ and, therefore, γ is contained in the envelope.

Example 2.6.1. *Find the solution of $pq = z$ that contains the line $(s, 1-s, 1)$.*

Note that $F(x, y, z, p, q) = pq - z$. The autonomous system written in differential form is

$$\frac{dx}{q} = \frac{dy}{p} = \frac{dz}{2pq} = \frac{dp}{p} = \frac{dq}{q} = dt.$$

We easily find the solution

$$\begin{cases} p = \alpha\, e^t, \\ q = \beta\, e^t, \\ z = \alpha\beta\, e^{2t} + \delta, \\ x = \beta\, e^t + \epsilon, \\ y = \alpha\, e^t + \lambda. \end{cases}$$

where $\alpha.\beta, \gamma, \delta, \epsilon$ are constants. Since we need a relation among the variables that allows express p, q in terms of the other variables, we choose $p = y - \lambda$, and so $q = z/(y - \lambda)$. Now, we have to solve

$$dz = (y - \lambda)\, dx + \frac{z\, dy}{y - \lambda}.$$

That can be done easily, leading to this solution

$$z = (x - \eta)(y - \lambda)$$

that depends on two parameters. Therefore, we got a complete solution. For the Cauchy problem, take the normal vector

$$(y - \lambda, x - \eta, -1)$$

that should be perpendicular to the curve $(s, 1 - s, 1)$ at each of its points. That is

$$0 = (1 - s - \lambda)(1) + (s - \eta)(-1) + (-1)(0) = (1 - s - \lambda) - (s - \eta).$$

On the other hand, $1 = (1-s-\lambda)(s-\eta)$. Therefore,

$$1-s-\lambda = s-\eta = 1, \text{ or}$$

$$1-s-\lambda = s-\eta = -1.$$

In the first case, we get $\lambda = -s$, $\eta = s-1$. The 1-parametric family of surfaces whose envelope will give the solution is

$$z = (x-s+1)(y+s).$$

Derivation with respect to the parameter s gives

$$0 = -(y+s) + (x-s+1) = x-y+1-2s.$$

We can get rid of s by putting $s = (x-y+1)/2$ in the parametric equation

$$z = (x - \frac{x-y+1}{2} + 1)(y + \frac{x-y+1}{2})$$

$$= \frac{1}{4}(x+y+1)(x+y+1) = \frac{1}{4}(x+y+1)^2.$$

Note that the choice -1 above leads to a different solution for the Cauchy problem

$$z = \frac{1}{4}(x+y-3)^2.$$

That shows that uniqueness for the general first order equation is more delicate than in the quasi-linear case: it is necessary to chose values for p, q, among the feasible ones, along the initial curve. ■

The method we just have illustrated shows how to build a solution of the first order PDE as the envelope of a suitable 1-parametric family issued from a complete solution $f(x, y, z, \lambda, \eta) = 0$. It may happen that the complete solution has an envelope $\phi(x, y, z) = 0$ as a 2-parametric family given by the system 2.13. We already proved in Proposition 2.4.4 that the double envelope satisfies the same PDE that the family, whereas it cannot be expressed as the envelope of an 1-parametric family. For that reason, the envelope $\phi(x, y, z) = 0$ was called a singular solution. For instance, the double envelope of the complete solution $z = (x-\eta)(y-\lambda)$ of the previous example can be easily computed and it turns out to be $z = 0$.

Example 2.6.2. *Find the solution of*

$$p^2 + q^2 = z^2$$

that contains the curve $x^2 + y^2 = 1$, $z = 1$.

The autonomous system written in differential form is

$$\frac{dx}{2p} = \frac{dy}{2q} = \frac{dz}{2z^2} = \frac{dp}{zp} = \frac{dq}{zq} = dt.$$

We do not need to solve the system. Observe that

$$\frac{dp}{p} = \frac{dq}{q}$$

implies $q = \lambda p$, with λ a parameter. Put

$$p = \frac{z}{\sqrt{1+\lambda^2}}; \qquad q = \frac{\lambda z}{\sqrt{1+\lambda^2}}.$$

The differential expression $dz = p\,dx + q\,dy$ for these values was solved in Example 2.3.6 giving

$$\log z = \frac{x + \lambda y}{\sqrt{1+\lambda^2}} + \eta.$$

It would be very convenient to change the parameter λ by another one more suitable to the geometry of the problem. Put

$$\frac{1}{\sqrt{1+\lambda^2}} = \cos\xi; \qquad \frac{\lambda}{\sqrt{1+\lambda^2}} = \sin\xi.$$

With that, the complete solution turns into

$$(\cos\xi)x + (\sin\xi)y - \log z = \eta, \tag{2.20}$$

with normal vector $(\cos\xi, \sin\xi, -1/z)$. The solution must contain the curve $(\cos\theta, \sin\theta, 1)$, with tangent vector $(-\sin\theta, \cos\theta, 0)$. The orthogonality condition

$$0 = (\cos\xi, \sin\xi, -1/z)\cdot(-\sin\theta, \cos\theta, 0) = \sin(\theta - \xi)$$

implies that we can take $\xi = \theta$. The dependence between θ and η appears when we introduce this information in 2.20. Indeed,

$$\eta = \cos^2\theta + \sin^2\theta - \log 1 = 1,$$

(η is actually a constant with respect to θ) therefore, the 1-parametric family of solutions is

$$(\cos\theta)x + (\sin\theta)y - \log z = 1.$$

The envelope is computed easily: derivation with respect to θ gives

$$-(\sin\theta)x + (\cos\theta)y = 0.$$

Therefore

$$\cos\theta = \frac{x}{\sqrt{x^2+y^2}}, \quad \sin\theta = \frac{t}{\sqrt{x^2+y^2}},$$

and now we can get rid of θ. We arrive to

$$\sqrt{x^2+y^2} - \log z = 1,$$

or, equivalently, the explicit solution

$$z = e^{\sqrt{x^2+y^2}-1}$$

as wished. ■

Remark 2.6.3. *Given a complete solution $f(x, y, z, \lambda, \eta) = 0$ of a first order PDE and a smooth curve $(x(s), y(s), z(s))$, the explained envelope method may provide finitely many, infinitely many or none solutions depending on the geometry of the available directions for the tangent plane. A curious situation happens when, by application of the method, the obtained parameters remain constant with respect to s, that is, $\lambda(s) = \lambda_0, \eta(s) = \eta_0$. In that case, the solution is the envelope of a "0-parametric" family, that is, the surface*

$$f(x, y, z, \lambda_0, \eta_0) = 0$$

from the complete solution, in itself.

2.7 Some special equations

There are some kind of first order PDEs that can be solved in a more simpler way. Since linear equations with constant coefficients will be discussed in next section, we will consider only the following two types.

Separable variables. A first order equation that can be written as

$$F_1(x,p) - F_2(y,q) = 0$$

is said to have separated variables. In this way, the characteristic system splits. In particular, for the pair x, p we have

$$x' = \frac{\partial F_1}{\partial p}(x,p); \quad p' = -\frac{\partial F_1}{\partial x}(x,p).$$

We claim that $F_1(x,p)$ remains constant along the trajectories of the system. Indeed,

$$\frac{d}{dt}(F_1(x,p)) = \frac{\partial F_1}{\partial x}(x,p)\, x' + \frac{\partial F_1}{\partial p}(x,p)\, p' = 0.$$

Evidently, the same happens with $F_2(y,q)$. Put

$$F_1(x,p) = F_2(y,q) = \lambda$$

and assume that we can resolve both equations, so we get that $p = P(x,\lambda)$ and $q = Q(y,\lambda)$. Finally, a complete solution of the equation is obtained by integrating

$$dz = P(x,\lambda)dx + Q(y,\lambda)dy.$$

Clairaut. A equation of the form

$$z = px + qy + f(p,q)$$

is called of *Clairaut*. The most remarkable fact is that the family of planes

$$z = \alpha x + \beta y + f(\alpha,\beta)$$

is a complete integral. Conversely, if a 2-parametric family of planes is given in that fashion, then its differential equation is obtained likewise. Typically, a Clairaut equation is obtained from the family of tangent planes to a given surface, which turns to be its singular integral. Note that envelopes of 1-parametric families of planes are *ruled surfaces*, that is, union of straight lines.

Example 2.7.1. *The family of tangent planes to the unit sphere obtained in Example 2.4.3 can be written in Clairaut form, for upper hemisphere, as*

$$z = xp + yq + \sqrt{p^2 + q^2 + 1}.$$

Multipliers for the autonomous system. To solve a first order PDE it is enough to find to two integrals of the characteristic system (3 variables for the quasi-linear, 5 for the general case). Sometimes, the symmetry of the functions helps in finding the integrals. For instance, consider

$$(y^2 - z^2)p + (z^2 - x^2)q = x^2 - y^2.$$

The characteristic system can be written as

$$\frac{dx}{y^2 - z^2} = \frac{dy}{z^2 - x^2} = \frac{dz}{x^2 - y^2}. \tag{2.21}$$

The sum of the three denominators is 0, so $x + y + z$ remains constant along characteristics. However, that can be expressed as

$$\frac{dx + dy + dz}{(y^2 - z^2) + (z^2 - x^2) + (x^2 - y^2)} = \frac{d(x + y + z)}{0}.$$

Now, if we multiply the fractions in 2.21 by x^2, y^2 and z^2 respectively we would get

$$\frac{x^2dx + y^2dy + z^2dz}{x^2(y^2 - z^2) + y^2(z^2 - x^2) + z^2(x^2 - y^2)} = \frac{3^{-1}d(x^3 + y^3 + z^3)}{0}$$

meaning that $x^3 + y^3 + z^3$ is another constant of the system. The *multipliers* are the functions $1, 1, 1$ and x^2, y^2, z^2 that allow the reduction of the system. In particular, the general solution of the original quasi-linear equation is

$$\Phi(x + y + z, x^3 + y^3 + z^3) = 0$$

where Φ is an arbitrary differentiable function of two variables. ∎

2.8 The symbolic method

Here we will consider linear PDEs, that is, equations of the form

$$A(x, y)u_x + B(x, y)u_y + C(x, y) = D(x, y).$$

The first term can be considered from a functional point of view using the partial differentiation operators

$$\partial_x := \frac{\partial}{\partial x}, \quad \partial_y := \frac{\partial}{\partial y}$$

in this way

$$A(x,y)u_x + B(x,y)u_y + C(x,y)u = (A\partial_x + B\partial_y + C)u.$$

Likewise in ODEs, the general solution can be obtained form a particular solution just by adding the general solution of the homogeneous equation

$$(A\partial_x + B\partial_y + C)u = 0.$$

From now on we will assume that the coefficients are constant. In that case, the homogeneous equation has a general solution of the form

$$u(x,y) = e^{-Cx/A} f(Bx - Ay)$$

being f an arbitrary differentiable function. The interest of this method is that, eventually, higher order linear equations can be reduced to first order. Indeed, consider a differential operator $\mathcal{D}$ of the form

$$\mathcal{D} = \partial_x^2 + a\partial_x\partial_y + b\partial_y^2 + c\partial_x + d\partial_y + d$$

which is associated to a second order linear PDE. Clearly, D can be seen as a quadratic form on ∂_x and ∂_y. If the quadratic form factorizes as a product of to linear forms

$$\mathcal{D} = (\partial_x + \alpha\partial_y + \beta)(\partial_x + \gamma\partial_y + \delta)$$

then any solution of

$$(\partial_x + \alpha\partial_y + \beta)u = 0, \text{ or}$$

$$(\partial_x + \gamma\partial_y + \delta)u = 0,$$

is a solution of $\mathcal{D}u = 0$, and also their sum. In other words,

$$u(x,y) = e^{-\beta x} f(x - \alpha y) + e^{-\delta x} g(x - \gamma y)$$

for f and g arbitrary functions is a solution of $\mathcal{D}u = 0$. It can be showed that, actually, all the solutions are of that form.

That applies to the wave equation

$$u_{tt} - c^2 u_{xx} = 0$$

since

$$\partial_t^2 - c^2\partial_x^2 = (\partial_t + c\partial_x)(\partial_t - c\partial_x).$$

Therefore, the solutions are of the form

$$u(x,t) = f(x - ct) + g(x + ct).$$

The method also applies when the coefficients are complex, but the arbitrary functions must be holomorphic. The Laplace equation

$$u_{xx} + u_{yy} = 0$$

has a differential operator that factorizes as

$$\partial_x^2 + \partial_y^2 = (\partial_x - i\partial_y)(\partial_x + i\partial_y)$$

giving the solution

$$u(x,t) = f(x + iy) + g(x - iy)$$

for f and g arbitrary holomorphic functions. That, is the solution can be expressed as the sum of an holomorphic function $f(z)$ and a conjugate holomorphic function $g(\overline{z})$.

In general, by a suitable linear change of variables, an homogeneous second order differential operator (in two variables) can be reduced to the simplest forms $\partial_x\partial_y$ or ∂_x^2 which of great help in case the independent term of the PDE does not reduces to zero. Unfortunately, for the heat-diffusion equation $u_t = ku_{xx}$ we cannot apply these ideas. However, the study of the quadratic form associated to a second order linear PDE provides some insight about how to deal with it. Indeed, the study of second order linear PDEs, not necessarily with constant coefficients, is classified according to that behaviour in three groups: *hyperbolic* if they are similar to the wave equation; *elliptic* if they are similar to the Laplace equation; and *parabolic* if they behave as the heat equation.

2.9 Rationale and remarks

The study of first order PDEs strongly relies on their geometrical interpretation. It is, therefore, necessary the student to master related analogous notions on the plane: autonomous systems, orthogonal family of curves, envelopes, singular solutions... Unfortunately, some of those notions are considered relics of the past and their teaching outdated.

For the general first order PDE, we have followed the method of Lagrange-Charpit, instead of working more on Monge cones, so any particular solution appears as the envelope of a suitable 2-parameterized family of surfaces. The drawback is that we had to deal with equations in differential form (Pfaffian systems) and Frobenius characterization of *complete integrability* which is a subject of independent interest.

The first order equation also play an important role in the reduction of equations of second order to *canonical forms*, as a first step for their classification.

2.10 Exercises

1. Obtain the first order PEDs satisfied by the following families of surfaces (f stands for an arbitrary differentiable function):

   (a) $z = (x + a)(y + b)$

   (b) $2z = (ax + y)^2 + b$

   (c) $ax^2 + by^2 + z^2 = 1$, $z > 0$

   (d) $z = x + y + f(xy)$

2. Consider the linear equation

$$u_x + u_y = u,$$

   where $u = u(x, y)$.

   (a) Find the solution containing the line $x + y = 0, z = 1$.

   (b) Find the solution containing the circle $x^2 + y^2 = 1, z = 1$.

   (c) The solution found in b) depends on the choice of a sign. Is it a violation of the uniqueness in Theorem 2.2.2?

3. Find the solution of the quasi-linear equation

$$-yu_x + xu_y = x^2 + y^2,$$

   containing the line $y = 0, z = 0$. Could you explain the result in terms of Theorem 2.2.2?

4. Find the solution of the following equations meeting the required condition:

   (a) $(x+2)p + 2yq = 2z$, with $u(-1,y) = \sqrt{y}$;
   (b) $xp + yq = 2(x^2+y^2)z$, containing the curve $x = 1$, $z = e$;
   (c) $pz + q = 1$, with $x = y$, $z = x/2$ for $0 \le x \le 1$;
   (d) $p + q = z^2$, passing through $(x, 0, h(x))$;
   (e) $xq - yp = z$, with $u(x,0) = h(x)$;
   (f) $q = xzp$, containing the line $(x, 0, x)$.

5. Write the general solution of

$$xp - yq = z$$

   in the most simple way.

6. Find the PDE satisfied by the family of tangent planes to the surface

$$z = xy.$$

   Without solving the PDE you have obtained, could you provide at least two more solutions different from xy or a plane? Do the same with the surface $z = x^2 + y^2$.

7. Find the envelope of the following families of curves and surfaces

   (a) $(x-\lambda)^2 - 2y = 0$;
   (b) $(x-2\lambda)^2 + y^2 - \lambda^2 = 0$;
   (c) $x^2/\lambda^2 + y^2/\eta^2 = 1$ being the product $\lambda\eta$ constant.
   (d) $x/\lambda + y/\eta + z/\xi$, with $\lambda + \eta + \xi = 1$.

8. Consider the family of planes

$$\lambda x + \eta y + \xi z = 1$$

   with the restriction $\lambda\eta\xi = 1$.

   (a) Find the 2-parametric envelope.
   (b) Find a Clairaut type equation satisfied by the given family.

9. Find a complete solution of $(1 - x^2)yp^2 + x^2q = 0$.

10. Find a complete solution of $z = xp + yq + p^2 - q^2$. Is there any singular solution?

11. Find a complete solution and a singular one of $(xp+yp-z)(p^2+q^2) = pq$.

Chapter 3

Fourier series

In this chapter we will able to proof fancy formulas as the following one

$$\frac{x}{2} = \sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} + \dots$$

for $x \in (-\pi, \pi)$. However, a moment's reflection shows that the right hand-side member, when convergent, defines a 2π-periodic function whereas the left hand-side term is a monotone function. Actually, the equality already fails at the butts of the interval. Around two hundred years ago, the fact that two *analytic expressions* could agree on an interval but not on the whole set of numbers was nothing less than catastrophe. The aftermath was a deep revision of the concept of function and real numbers, giving birth to the modern Mathematical Analysis.

3.1 Introduction

The finite linear combinations of the functions $\sin nx$, $\cos nx$, for $n \in \mathbb{N}$ and the constants are called *trigonometric polynomials*. A motivation for the term "polynomials" is that they are stable by products thanks to the formulas

$$\begin{aligned}
\sin(\alpha)\sin(\beta) &= 2^{-1}(\cos(\alpha-\beta) - \cos(\alpha+\beta)) \\
\cos(\alpha)\cos(\beta) &= 2^{-1}(\cos(\alpha+\beta) + \cos(\alpha-\beta)) \\
\sin(\alpha)\cos(\beta) &= 2^{-1}(\sin(\alpha+\beta) + \sin(\alpha-\beta))
\end{aligned}$$

that allow the reduction of the terms.

A *trigonometric series* is a formal series of the way

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx). \tag{3.1}$$

We say formal because its convergence or the lack of it can be discussed from several points of view. Moreover, we say that the previous series is the *Fourier series* of an integrable function f if

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\, dx \text{ for } n \geq 0,$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\, dx \text{ for } n \geq 1.$$

The motivation for the formulas is the *orthogonality* of the trigonometric system, that is to be discussed in the next section. Nevertheless, we will show the computation with an example.

Example 3.1.1. *Calculate the Fourier series of the function $f(x) = x/2$ as defined on $[-\pi, \pi]$.*

Since the function f is odd we easily deduce that $a_n = 0$ for $n \geq 0$. Therefore, only b_n's are relevant for us. Consider

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{x}{2}\sin nx\, dx = \frac{1}{\pi}\int_{0}^{\pi} x \sin nx\, dx$$

$$= \frac{-1}{\pi n} x\cos nx \,\Big|_0^{\pi} + \frac{1}{\pi n}\int_0^{\pi}\cos nx\, dx = \frac{(-1)^{n+1}}{n}$$

Therefore the Fourier series of f is

$$\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin nx = \sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \dots$$

Note that the series for $x = 0$ coincides with f, which is not great surprise. A little less trivial, for $x = \pi/2$, we have the Leibniz series

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \dots$$

As we have promised, we will investigate in this chapter when a function agrees with its Fourier series.

The theory of trigonometric and Fourier series can be developed in the frame of complex valued real functions. In such a case, we can use the "basic" functions e^{inx} for $n \in \mathbb{Z}$. In that case, the series looks like

$$\sum_{n\in\mathbb{Z}} a_n e^{inx}$$

and the coefficients, in the Fourier case, come from the formula

$$a_n = \hat{f}(n) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx \tag{3.2}$$

(the notation $\hat{f}$ will be used to stress the dependence on f when necessary). Since the real case can be derived always from the complex case, we will develop the theory for series with complex coefficients, meanwhile the examples will refer to the real case.

3.2 The Hilbert theory

Along this section we will consider only continuous functions and the integral can be understand in the sense of Riemann.

Recall that an hermitian product is the suitable notion of scalar product for complex spaces X, that is, a function

$$\langle\ ,\ \rangle : X \times X \to \mathbb{C}$$

such that:

(a) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$;

(b) $\langle y, x\rangle = \overline{\langle x, y\rangle}$;

(c) $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.

Recall that the norm on X is said to be *strictly convex* if whenever $\|x+y\| = \|x\|+\|y\|$, then x and y are linearly dependent trough a real nonnegative coefficient (other equivalent definitions of strict convexity are proposed in exercises). The reader most likely is familiar with the real version of the following result.

Proposition 3.2.1. *Let $\langle\ ,\ \rangle$ be an hermitian product on a complex space X. Then the Cauchy-Schwarz inequality holds*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle,$$

for every $x, y \in X$. The equality happens if and only if x and y are linearly dependent. As a consequence, the formula

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

defines a strictly convex norm on X.

Proof. Consider the following inequality

$$0 \leq \langle x - \lambda y, x - \lambda y\rangle = \|x\|^2 - 2\Re(\lambda\langle y, x\rangle) + \lambda^2\|y\|^2,$$

and set $\lambda = \frac{\langle x,y\rangle}{|\langle x,y\rangle|}t$ with $t \in \mathbb{R}$. We get then

$$0 \leq \|x\|^2 - 2|\langle x, y\rangle|t + \|y\|^2 t^2.$$

Since the polynomial remains positive, for its discriminant we have

$$4|\langle x, y\rangle|^2 - 4\|x\|^2\|y\|^2 \leq 0$$

which is exactly the desired inequality. The equality is trivial whenever x, y are linearly dependent. On the other hand, if $|\langle x, y\rangle|^2 = \langle x, x\rangle\langle y, y\rangle$ put $t = \frac{|\langle x,y\rangle|}{\|y\|^2}$ (we may assume $x, y \neq 0$) and substitute into the polynomial. That leads to

$$0 \leq \langle x - \lambda y, x - \lambda y\rangle = \|x\|^2 - 2\frac{|\langle x, y\rangle|^2}{\|y\|^2} + \frac{|\langle x, y\rangle|^2}{\|y\|^2} = 0.$$

Therefore, $x = \lambda y$ as wished.
To show that $\|\cdot\|$ is a norm, everything is trivial except the triangle inequality. We have

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + 2\Re(\langle x, y\rangle) + \|y\|^2$$

$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

In case the equality $\|x+y\| = \|x\|+\|y\|$ holds, then $\Re(\langle x, y\rangle) = \|x\|\|y\|$. Since the inequality

$$\Re(\langle x, y\rangle) \leq |\langle x, y\rangle| \leq \|x\|\|y\|$$

is always true, we deduce by the first part that x, y are dependent and, moreover, if $x = \lambda y$ then $\lambda \in \mathbb{R}$ and $\lambda \geq 0$. ∎

In order to deal with the complex valued 2π-periodic function on the real line we will follow this convection

$$C(\mathbb{T}) = \{f \in C(\mathbb{R}, \mathbb{C}) : f(x + 2\pi) = f(x) \ \forall \, x \in \mathbb{R}\}$$

where $C(\mathbb{R}, \mathbb{C})$ is the set of complex functions of real variable. From a strict topological point of view, we can think of $\mathbb{T}$ as the topological compact space resulting of identifying $-\pi$ and π. As a topological group, $\mathbb{T}$ is the quotient $\mathbb{R}/2\pi\mathbb{Z}$. It is also useful to consider $\mathbb{T}$ as the set the complex numbers having modulus one, so $f \in C(\mathbb{T})$ is just a function defined on $\partial D(0, 1)$.

Theorem 3.2.2. *$C(\mathbb{T})$ is a complex pre-Hilbert space with the hermitian product*

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)} \, dx.$$

Proof. It is left to the reader. ∎

Now we will discuss the notion of best approximation.

Theorem 3.2.3. *Let X be a normed space and $F \subset X$ a finite dimensional subspace. Then:*

1. *For every $x \in X$ there is a best approximation $y \in F$, that is,*

$$\|x - y\| = d(F, x) := \inf\{\|x - z\| : z \in F\}.$$

2. *If X is moreover strictly convex, then for every $x \in X$ its best approximation in F is unique.*

3. *If the norm of X comes from an inner product, the best approximation $y \in F$ of an element $x \in X$ is its orthogonal projection. Namely, y is characterized by the property*

$$\langle x - y, z \rangle = 0 \text{ for all } z \in F.$$

Proof. Let $R = \|x\|$. The set $B[x, R] \cap F$ is nonempty, since $0 \in F$, and the infimum $d(F, x)$ can be calculated on it by the very definition. Since the set $K = B[x, R] \cap F$ is closed, bounded and finite-dimensional it is compact, the function $z \to \|x - z\|$ attains its minimum on K at some $y \in K$, which is a global minimizer on F by the definition of K. If there are two minimizing points $y_1, y_2 \in F$, that is,

$$\|x - y_1\| = \|x - y_2\| = d(F, x)$$

By the convexity of the norm, we have

$$\|x - \frac{y_1 + y_2}{2}\| \leq \frac{\|x - y_1\| + \|x - y_2\|}{2} = d(F, x).$$

Since the infimum cannot be improved, we have

$$\|x - \frac{y_1 + y_2}{2}\| = \frac{\|x - y_1\| + \|x - y_2\|}{2}.$$

If the norm of X were strictly convex, that would imply $y_1 = y_2$.
Assume now that the norm comes from an inner product $\langle \ , \ \rangle$, which is a real differentiable function on $X \times X$. For any point $x \in X$, the function $y \to \langle x - y, x - y \rangle$ is differentiable (that is a particular case of the differentiability of bounded bilinear or hermitic forms), being its differential

$$z \to \langle z, x - y \rangle + \langle x - y, z \rangle = 2\Re(\langle x - y, z \rangle).$$

If there is a minimum at $y \in F$ among the points of F, then the differential restricted to F has to be 0. Namely,

$$\Re(\langle x - y, z \rangle) = 0$$

for every $z \in F$. Changing z by iz we get also that

$$\Im(\langle x - y, z \rangle) = 0$$

and therefore $\langle x - y, z \rangle = 0$ for every $z \in F$. That is exactly the so called orthogonal projection, which is unique as it is easy to check. ■

The characterization of the best approximation can be done by purely geometrical arguments instead of appealing to differential calculus.

Next result is the key to understand Fourier series from the Hilbert point of view.

Theorem 3.2.4. *Let X be a pre-Hilbert space and let $\{e_n : n \in \mathbb{N}\}$ be an orthonormal system. Then:*

(a) the orthogonal projection of x onto $F_n = \operatorname{span}\{e_k : 1 \le k \le n\}$ is

$$\sum_{k=1}^{n} \langle x, e_k \rangle e_k$$

and it holds Bessel's inequality

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \le \|x\|^2;$$

(b) assume moreover that $\operatorname{span}\{e_n : n \in \mathbb{N}\}$ is moreover dense in X, then

$$\lim_n \|x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k\| = 0,$$

that is, the series $\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ converges to x in X, and it holds Parseval's equality

$$\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 = \|x\|^2.$$

Proof. It is easy to prove that $y = \sum_{k=1}^{n} \langle x, e_k \rangle e_k$ fulfils the criterion of orthogonal projection of the previous theorem. Note that $x - y$ and y are orthogonal, therefore the Pythagoras identity holds

$$\|x\|^2 = \|x - y\|^2 + \|y\|^2$$

and, in particular,

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 = \|y\|^2 \le \|x\|^2.$$

If the orthonormal system is dense and y_n is the orthogonal projection on F_n, then $\lim_n \|x - y_n\| = 0$. Since

$$y_n = \sum_{k=1}^{n} \langle x, e_k \rangle e_k$$

we deduce that the series is convergent to x in X

$$\sum_{k=1}^{\infty}\langle x, e_k\rangle e_k = x$$

and the same happens for the series obtained from the equality

$$\sum_{k=1}^{n} |\langle x, e_k\rangle|^2 = \|y_n\|^2$$

■

An orthonormal system $\{e_n : n \in \mathbb{N}\}$ (sometimes indexed by $\mathbb{Z}$ for convenience) satisfying statement (b) above is called a *Hilbert basis*.

Remark 3.2.5. *The Hilbert sequence space is defined as*

$$\ell_2 = \{(a_n) : (a_n) \subset \mathbb{C}, \sum_{n=1}^{\infty} |a_n|^2 < +\infty\}$$

with the scalar product

$$\langle (a_n), (b_n)\rangle = \sum_{n=1}^{\infty} a_n \overline{b_n}.$$

Note that the mapping $x \to (\langle x, e_k\rangle)_{n=1}^{\infty}$ takes elements from the pre-Hilbert space to ℓ_2 preserving the norm and the scalar product.

From this moment on we forget about a general Hilbert basis because our interest is the *trigonometric system*, that is, the set of functions

$$\{e^{int} : n \in \mathbb{Z}\}.$$

We already know that these functions are orthogonal and we will prove that, up to a multiplicative constant ($\sqrt{2\pi}$) they made up a Hilbert basis. The *real trigonometric system*, that is,

$$\{1\} \cup \{\cos nt : n \in \mathbb{N}\} \cup \{\sin nt : n \in \mathbb{N}\}$$

can be deduced by arrangement in conjugate pairs.

We need the following complex version of the Stone-Weierstrass theorem. We assume that the real version is known (see [33]).

Theorem 3.2.6. *Let K be a compact space and let $A \subset C(K, \mathbb{C})$ be a subalgebra with the following properties:*

(a) A contains the constants;

(b) A distinguishes points of K;

(c) $\overline{f} \in A$ whenever $f \in A$.

Then, A is dense in $(C(K, \mathbb{C}), \| \cdot \|_\infty)$.

Proof. We will prove firstly that the real valued functions form A satisfies the real Stone-Weierstrass theorem. Indeed, they are a real subalgebra that contains the constant. Moreover, given $f \in A$ arbitrary, then

$$\Re(f) = \frac{f + \overline{f}}{2} \quad \text{and} \quad \Im(f) = \frac{f - \overline{f}}{2i}$$

are real valued. Given two points $t, s \in K$ with $t \neq s$ if $f \in A$ disguises the points, that is, $f(t) \neq f(s)$, then either $\Re(f)$ or $\Im(f)$ disguises the points too. All this implies the real valued functions from A are dense in $C(K)$. Now, given $f \in C(K, \mathbb{C})$, the functions $\Re(f)$ and $\Im(f)$ can be uniformly approximated by functions from A. Therefore, the function $f = \Re(f) + i\Im(f)$ can be approximated uniformly as well by functions from A. ■

Now, we can put all the pieces of the puzzle together in order to get the most important result in this section.

Theorem 3.2.7. *If $f \in C(\mathbb{T})$ then its Fourier series $\sum_{n \in \mathbb{Z}} \hat{f}(n) e^{int}$ converges to f with respect to the norm $\| \cdot \|_2$.*

Proof. The span$\{e^{int} : n \in \mathbb{Z}\}$ is dense in $C(\mathbb{T})$ with the uniform norm $\| \cdot \|_\infty$, and thus it is dense with respect to the $\| \cdot \|_2$ norm. Indeed,

$$\|f\|_2 = \left(\int_{-\pi}^{\pi} |f(x)|^2 dx \right)^{1/2} \leq (2\pi \|f\|_\infty^2)^{1/2} = \sqrt{2\pi} \|f\|_\infty,$$

meaning that approximation in norm $\| \cdot \|_\infty$ implies approximation in norm $\| \cdot \|_2$. Now the result follows straight from Theorem 3.2.4. ■

3.3 Spaces of integrable functions

From now on we will work with the Lebesgue integral. Let us remind the spaces of integrable functions

$$\mathcal{L}^p(\mu) = \{f \text{ measurable} : \int |f|^p \, d\mu < \infty\};$$

$$\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p} \quad \text{if } f \in \mathcal{L}^p(\mu) \text{ and } 1 \le p < \infty.$$

Consider $\sim$ the equivalence relation $f \sim g$ if $f = g$ almost everywhere. The quotient spaces $L^p(\mu) = \mathcal{L}^p(\mu)/\sim$ are still vector spaces and the functional $\|\cdot\|_p$ is well defined on them. Note that $L^1(\mu)$ is the largest of Lebesgue spaces provided that μ is finite. We assume the following fact is already proven.

Theorem 3.3.1. $(L^p(\mu), \|\cdot\|_p)$ *is Banach space for* $1 \le p \le \infty$*.*

We will work mainly with $L^1(\mathbb{T})$ and $L^2(\mathbb{T})$, meaning in practice that $f \in L^p(\mathbb{T})$ if $f : \mathbb{R} \to \mathbb{C}$ is measurable, 2π-periodic and $f|_{[-\pi,\pi]} \in L^p[-\pi,\pi]$. Note that we stress the domain $\mathbb{T}$, instead of the measure μ, because in this context Lebesgue measure is understood. The density of the continuous functions among the integrable ones is established for the Lebesgue measure. In particular, $C(\mathbb{T})$ is dense in $L^2(\mathbb{T})$, either with real or complex values. That implies the following improvement of Theorem 3.2.7.

Theorem 3.3.2. *For every function* $f \in L^2(\mathbb{T})$*, its Fourier series converges to* f *with respect to the norm* $\|\cdot\|_2$*, that is, the equality*

$$f(x) = \sum_{n\in\mathbb{Z}} \hat{f}(n)\, e^{inx}, \qquad \text{where } \hat{f}(n) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt,$$

is interpreted as elements of $L^2(\mathbb{T})$*. In particular, Parseval's equality holds*

$$\|f\|_2^2 = 2\pi \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2$$

Proof. Just apply (b) of Theorem 3.2.4. ■

The previos result is often referred as that, up to a multiplicative constant, the trigonometric system $\{e^{inx} : n \in \mathbb{Z}\}$ is a Hilbert basis of $L^2(\mathbb{T})$. It is very important to remark that the equality in the previous theorem has no pointwise meaning. The pointwise and, eventually uniform, convergence of the Fourier series will be investigated in the remaining part of this chapter.

Example 3.3.3. *Apply Parseval's equality to Example 3.1.1.*

Since we are using the real Fourier series trigonometric, the coefficients should be adapted (note that $\|\sin nt\|_2^2 = \pi$ or use exercise 2)

$$\frac{\pi^3}{6} = \int_{-\pi}^{\pi} \left(\frac{t}{2}\right)^2 dt = \pi \sum_{n=1}^{\infty} \left|\frac{1}{\pi} \int_{-\pi}^{\pi} \frac{t}{2} \sin nt\right|^2 dt = \pi \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

We deduce so the famous equality

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Now, note that the Fourier coefficients can be defined by functions in $L^1(\mathbb{T})$. Indeed,

$$|a_n| = \left|\int_{-\pi}^{\pi} f(t)e^{-ixt}dt\right| \leq \int_{-\pi}^{\pi} |f(t)|dt = \|f\|_1.$$

In general we cannot expect that $\sum_n |a_n|^2 < +\infty$, but at least we can prove that the coefficients go to 0. More generally, we have the following result known as the *Riemann-Lebesgue lemma*.

Theorem 3.3.4. *Let $f \in L^1(\mathbb{R})$ (in particular, if $f \in L^1(\mathbb{T})$). Then*

$$\lim_{|x|\to\infty} \int_{-\infty}^{\infty} f(t)e^{ixt}dt = 0.$$

As a consequence, $\lim_{|x|\to\infty} \int_{-\infty}^{\infty} f(t)\sin(xt)dt = 0$.

Proof. Firstly, we will prove that the statement is true for the characteristic function of an interval $[a, b] \subset \mathbb{T}$. Indeed,

$$\lim_{|x|\to\infty} \int_{-\infty}^{\infty} \chi_{[a,b]}e^{ixt}dt = \lim_{|x|\to\infty} \int_{a}^{b} e^{ixt}dt = \lim_{|x|\to\infty} \frac{e^{ibx} - e^{iax}}{ix} = 0,$$

because $|e^{ibx} - e^{iax}| \leq 2$. The statement is obviously true for linear combinations of characteristic functions of intervals too. Note that those functions are dense dense in $L^1(\mathbb{R})$ with respect to the norm $\|\cdot\|_1$. Take $f \in L^1(\mathbb{R})$. Given $\varepsilon > 0$, we can find g a linear combination of characteristic functions of intervals such that $\|f - g\|_1 < \varepsilon/2$. Take $M > 0$ such that $|\int_{-\infty}^{\infty} f(t)e^{ixt}dt|\varepsilon/2$ for $|x| > M$. Then we have

$$\left|\int_{-\infty}^{\infty} f(t)e^{ixt}dt\right| \leq \left|\int_{-\infty}^{\infty} g(t)e^{ixt}dt\right| + \|f - g\|_1 < \varepsilon,$$

which implies that f satisfies the statement as $\varepsilon > 0$ was arbitrary. ■

The use of the *convolution* will allow us to simplify the formulas to come and it will help to better understand the limit processes. Recall that the convolution of two functions $f_1, f_2 \in L^1(\mathbb{T})$ (the definition is also valid for $L^1(\mathbb{R})$ or $L^1(\mathbb{R}^n)$ with small changes) is defined almost everywhere as

$$(f_1 * f_2)(x) = \int_{-\pi}^{\pi} f_1(t) f_2(x - t)\, dt$$

where the difference $x - t$ is understood in the group $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. The convolution has the following properties.

Proposition 3.3.5. *Let $f_1, f_2, f_3 \in L^1(\mathbb{T})$. Then:*

(a) $f_1 * f_2 \in L^1(\mathbb{T})$ *and* $\|f_1 * f_2\|_1 \leq \|f_1\|_1 \cdot \|f_2\|_1$.

(b) $(f_1 + f_2) * f_3 = f_1 * f_3 + f_2 * f_3$ *(linearity).*

(c) $(f_1 * f_2) * f_3 = f_1 * (f_2 * f_3)$ *(associativity).*

(d) $f_1 * f_2 = f_2 * f_1$ *(commutativity).*

(e) If $f \in C^k(\mathbb{T})$ *with* $k \geq 0$, *then* $f * f_2 \in C^k(\mathbb{T})$ *and* $(f * f_2)^{(k)} = f^{(k)} * f_2$.

(f) if $f_1, f_2 \in L^2(\mathbb{T})$, *then* $f_1 * f_2 \in C(\mathbb{T})$ *and* $\|f_1 * f_2\|_\infty \leq \|f_1\|_2 \cdot \|f_2\|_2$.

Proof. Statement (a) follows from

$$\int_{-\pi}^{\pi} |(f_1 * f_2)(x)|\, dx \leq \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |f_1(t) f_2(x - t)|\, dt\, dx = \|f_1\|_1 \|f_2\|_1$$

where the last equality follows by an elementary change of variable. Properties (b), (c) and (d) are not difficult and left to the reader.
To prove (e), we will start by $k = 0$. Note that uniform continuity implies that

$\|f - f_h\|_\infty$ can be arbitrarily small, where $f_h(x) := f(x+h)$. Now, observe that

$$|(f * f_2)(x+h) - (f * f_2)(x)| \le \int_{-\pi}^{\pi} |(f(x+h-t) - f(x-t))f_2(t)|\,dt$$

$$\le \int_{-\pi}^{\pi} |(f_h(x-t) - f(x-t))f_2(t)|\,dt \le \|f_h - f\|_\infty \|f_2\|_1$$

which implies the continuity of $f * f_2$. For $k = 1$, note that

$$\frac{(f * f_2)(x+h) - (f * f_2)(x)}{h} = \left(\frac{f - f_h}{h} * f_2\right)(x).$$

Since $h^{-1}(f_h - f)$ converges uniformly to f', we deduce

$$\lim_{h\to 0} h^{-1}((f * f_2)(x+h) - (f * f_2)(x)) = (f' * f_2)(x).$$

Therefore, $f * f_2 \in C^1(\mathbb{T})$ and the rest of the proof follows by induction. Statement (f) follows from Cauchy-Schwarz, and the proof of the continuity is done using density of $C(\mathbb{T})$. ■

The corresponding result for $L^1(\mathbb{R})$ can be proved using $C_{00}^k(\mathbb{R})$ and the density of $C_{00}(\mathbb{R})$ in $L^1(\mathbb{R})$.

3.4 Uniform convergence

The best way to guarantee the uniform convergence of a the Fourier series of a function f is to control the decay of its Fourier coefficients $|a_n|$ as $|n| \to \infty$. Note that if the Fourier coefficients (a_n) of a given continuous function $f \in C(\mathbb{T})$ made up a absolutely convergent series, that is, if

$$\sum_{n\in\mathbb{Z}} |a_n| < +\infty,$$

then the Fourier series is absolutely convergent and

$$f(x) = \sum_{n\in\mathbb{Z}} a_n e^{inx}.$$

Indeed, the Fourier series is uniformly convergent by Weierstrass criterion, to some continuous function $g \in C(\mathbb{T})$. The functions f and g have the same

Fourier coefficients by the commutativity of series and integral. Therefore $f = g$ as elements of $L^2(\mathbb{T})$, so they are equal almost everywhere. The continuity implies that they are equal everywhere.

We already know that for $f \in L^1(\mathbb{T})$ it is true that $\lim_{|n|} |a_n| = 0$ (see Theorem 3.3.4), that can be expressed with Landau's notation as $a_n = o(1)$. We also have seen that if $f \in L^2(\mathbb{T})$ then $\sum_n |a_n|^2 < +\infty$, which is optimal for the $\|\cdot\|_2$-convergence, but not for the uniform one. Indeed, there are examples of $f \in C(\mathbb{T})$ for which the decay of $|a_n|$ for is not enough for the uniform convergence of its Fourier series.

However, if we ask for some more regularity we can get faster decay. Indeed, assume that $f \in C^1(\mathbb{T})$. Applying the integration by parts we get

$$\int_{-\pi}^{\pi} f(t)e^{-int}dt = f(t)\frac{e^{-int}}{-in}\bigg|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(t)\frac{e^{-int}}{-in}dt = \frac{-i}{n}\int_{-\pi}^{\pi} f'(t)e^{-int}dt.$$

Therefore, if (b_n) denotes the sequence of Fourier coefficients of f', we have $a_n = -ib_n/n$ and thus $a_n = o(1/n)$. Although this is not enough for the uniform convergence, one more iteration of the method would give $a_n = o(1/n^2)$ provided that $f \in C^2(\mathbb{T})$. Actually, that proves even more.

Proposition 3.4.1. *Let $f \in C^k(\mathbb{T})$ for some $k \in \mathbb{N}$. Then $\hat{f}(n) = o(|n|^{-k})$.*

If we only need uniform convergence of the Fourier series, we can get it with weaker hypotheses. Recall that a function $f \in C(\mathbb{T})$ is said *absolutely continuous* if for every $\varepsilon > 0$ there is $\delta > 0$ such that for any choice of points

$$-\pi \le a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n \le \pi$$

with $\sum_{k=1}^n (b_k - a_k) < \delta$ then $\sum_{k=1}^n |f(b_k) - f(a_k)| < \varepsilon$. The relevance of absolutely continuous functions is that they are differentiable almost everywhere and it is possible to retrieve the function from the derivative but a constant. In particular, Barrow's rule and integration by parts are valid for absolutely continuous functions.

Theorem 3.4.2. *Let $f \in C(\mathbb{T})$ be an absolutely continuous function such that $f' \in L^2(\mathbb{T})$. Then the Fourier series of f converges uniformly to f.*

Proof. Let (a_n) and (b_n) denote the Fourier coefficients of f and f' respectively. It is easy to prove that the product of two (bounded) absolutely continuous functions is absolutely continuous too. In particular, that applies to

$f(t)e^{-int}$ and implies the validity of this computation

$$a_n = \int_{-\pi}^{\pi} f(t)e^{-int}dt = f(t)\frac{e^{-int}}{-in}\bigg|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(t)\frac{e^{-int}}{-in}dt = \frac{-ib_n}{n}.$$

Now we have

$$\sum_{k=-n,k\neq 0}^{n} |a_k| = \sum_{k=-n,k\neq 0}^{n} |k^{-1}b_k| \leq (\sum_{k=-n,k\neq 0}^{n} k^{-2})^{1/2}(\sum_{k=-n}^{n} |b_k|^2)^{1/2} \leq K\,\|f'\|_2$$

thanks to the Cauchy-Schwarz inequality, where $K > 0$ can be computed and it turns out that $K = \sqrt{\pi/6}$. Now the uniform convergence follows from Weierstrass criterion. ■

The idea of using integration by parts also provides the following.

Proposition 3.4.3. *Let $f \in L^1(\mathbb{T})$ be such that $\hat{f}(0) = 0$ and let F be a primitive of F. Then, for $n \neq 0$,*

$$\hat{F}(n) = \frac{-i}{n}\hat{f}(n).$$

Proof. Note $\hat{f}(0) = 0$ is the same that saying $\int_{-\pi}^{\pi} f(t)\,dt = 0$. Therefore, $F(-\pi) = F(\pi)$ and so $F \in C(\mathbb{T})$, which is the key for the computation. ■

3.5 Pointwise convergence

First of all, we need an explicit "compact" form for the sum of the partial sums of the Fourier series. Recall that

$$S_n(f)(x) = \sum_{|k|\leq n} \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt\right) e^{ikx}.$$

Since we can exchange the finite sum with the integral, we get

$$S_n(f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\left(\sum_{|k|\leq n} e^{ik(x-t)}\right) dt.$$

The sum inside he bracket is actually a geometric progression, and its sum can be computed with this trick

$$(e^{it/2} - e^{-it/2}) \left(\sum_{|k|\le n} e^{ikt} \right) = \sum_{k=-n}^{n} e^{i(k+1/2)t} - \sum_{k=-n}^{n} e^{i(k-1/2)t}$$

$$= \sum_{k=-n}^{n} e^{i(k+1/2)t} - \sum_{k=-n-1}^{n-1} e^{i(k+1/2)t} = e^{i(n+1/2)t} - e^{i(-n-1/2)t}.$$

Using the complex exponential form for the sinus we get

$$\sum_{|k|\le n} e^{ikt} = \frac{\sin((n+1/2)t)}{\sin(t/2)}$$

Let us introduce the so called *Dirichlet kernel*

$$D_n(t) = \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)}.$$

Then we have

$$S_n(f)(x) = \int_{-\pi}^{\pi} f(t) D_n(x-t)dt = (D_n * f)(x)$$

using the convolution.

The analysis of the pointwise convergence depends on the behavior of D_n as n goes to ∞. The first consequence is quite amazing.

Proposition 3.5.1. *The convergence of the Fourier series of $f \in L^1(\mathbb{T})$ at some $x \in \mathbb{T}$ depends only on the values of f over a neighbourhood of x.*

Proof. Without loose of generality we may take $x = 0$. Fix $\varepsilon \in (0, \pi)$ and write

$$S_n(f)(0) = \int_{-\varepsilon}^{\varepsilon} f(t) \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)} dt + \int_{\mathbb{T}\setminus[-\varepsilon,\varepsilon]} f(t) \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)} dt.$$

The second integral can be written as

$$I_n = \int_{-\pi}^{\pi} g(t) \sin((n+1/2)t)dt$$

where

$$g(t) = \frac{f(t)}{2\pi \sin(t/2)} \quad \text{if} \quad t \in \mathbb{T} \setminus [-\varepsilon, \varepsilon], \quad \text{and} \quad g(t) = 0 \quad \text{otherwise.}$$

Obviously, $g \in L^1(\mathbb{T})$, so Theorem 3.3.4 applies to get that $\lim_n I_n = 0$. Therefore, the behaviour of $S_n(f)(0)$ depends only of the values of f in $[-\varepsilon, \varepsilon]$. ■

The precedent result implies the pointwise convergence of the Fourier series at any point having a neighbourhood where the function is piecewise C^1 on $\mathbb{T}$. Indeed, the function can be modified out that neighbourhood to be C^1 piecewise C^1 . However, we wish to find a more precise condition for the convergence of the Fourier series at a given point. This is a version of the so called *Dini test*.

Theorem 3.5.2. *Let $f \in L^1(\mathbb{T})$ and $x \in \mathbb{T}$ be such that $\diamondsuit_x f \in L^1(\mathbb{T})$, where*

$$(\diamondsuit_x f)(t) = \frac{f(t) - f(x)}{t - x}.$$

Then the Fourier series of f converges at x to $f(x)$.

Proof. First of all, since $\int_{-\pi}^{\pi} D_n(t)dt = 1$, we have

$$S_n(f)(x) - f(x) = \int_{-\pi}^{\pi} (f(t) - f(x)) \frac{\sin((n+1/2)(x-t))}{2\pi \sin((x-t)/2)}\, dt.$$

Note that the function

$$g(t) = \frac{f(t) - f(x)}{2\pi \sin((x-t)/2)}$$

is integrable by the hypothesis and the equivalence $\sin((x-t)/2) \sim (x-t)/2$. Therefore,

$$\begin{aligned} \lim_n (S_n(f)(x) - f(x)) &= \lim_n \int_{-\pi}^{\pi} g(t) \sin((n+1/2)(x-t))\, dt \\ &= \lim_n \int_{-\pi}^{\pi} g(x-t) \sin((n+1/2)(t))\, dt = 0 \end{aligned}$$

by Theorem 3.3.4. ■

We say that a function satisfies an α-Hölder condition at x, with $\alpha \in (0, 1]$, if there is $c > 0$ such that $|f(x+h) - f(x)| \leq c|h|^\alpha$ for t in a neighborhood of 0. For $\alpha = 1$ this is just the Lipschitz condition at x.

Corollary 3.5.3. *Let $f \in L^1(\mathbb{T})$ satisfy an α-Hölder condition at $x \in \mathbb{T}$ for some $\alpha \in (0,1]$ (e.g. f is differentiable at x). The the Fourier series of f converges to $f(x)$ at x.*

Example 3.5.4. *Study the convergence of the Fourier series obtained in Example 3.1.1.*

The previous result implies the convergence of the series

$$\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \dots$$

to $x/2$ on $(-\pi, \pi)$. Obviously, on $\pi = -\pi$ the sum of the series is 0, so it is convergent too. We can show that the convergence is uniform on intervals of the form $[a,b] \subset (-\pi,\pi)$. Indeed, using the complex exponential it is easy to check the bound

$$\left|\sum_{k=1}^{n} (-1)^k \sin kx\right| \leq \frac{2}{|1+e^{ix}|}$$

that implies those partial sums are uniformly bounded on $[a,b]$. As $(1/n)$ is monotonic with limit 0, The Dirichlet's test (see the Exercises) implies the uniform convergence of the Fourier series on $[a,b]$. ■

The last part of this section is devoted to the convergence of the Fourier series of functions having jump type discontinuities. We will use this notation

$$f(a^-) := \lim_{x \to a^-} f(x) \text{ and } f(a^+) := \lim_{x \to a^+} f(x).$$

Proposition 3.5.5. *Let $f \in L^1(\mathbb{T})$ be a function which is absolutely continuous out a finite subset $F \subset \mathbb{T}$ where f has jump discontinuities and $f' \in L^2(\mathbb{T})$. Then the Fourier series of f converges uniformly on compact sets $K \subset \mathbb{T} \setminus F$.*

Proof. Let $h(x) = x/2\pi$ on $(-\pi,\pi)$ considered as a function on $\mathbb{T}$. Define a function as

$$g(x) = \sum_{a \in F} (f(a^+) - f(a^-)) h(x + \pi - a).$$

Note that $f + g$ is continuous at F, and thus absolutely continuous on $\mathbb{T}$. The Fourier series of $f + g$ uniformly converges on $\mathbb{T}$ by Theorem 3.4.2. On the other hand, the Fourier series of g uniformly converges on any compact subset K such that $K \cap F = \emptyset$ after Example 3.5.4. ■

Now we will study what happens at the jump discontinuity. We say that a function f satisfies a Lipschitz condition at every side of $a \in \mathbb{T}$ if

$$|f(x) - f(a^-)| \leq c|x| \text{ for } x < a;$$

$$|f(x) - f(a^+)| \leq c|x| \text{ for } x > a$$

(since the definition matters for x close to a, the meaning of $<$ and $>$ offers no confusion). The definition extends in the obvious way for Hölder.

Proposition 3.5.6. *Assume that $f \in L^1(\mathbb{T})$ has a jump discontinuity (the side limits exist) at $x \in \mathbb{T}$ and a Lipschitz condition (or more generally, Hölder) is satisfied at every side of x (e.g. there exist the side limits of the derivative). Then the Fourier series at x converges to*

$$\frac{f(x^+) + f(x^-)}{2}.$$

Proof. Without loss of generality, we will work around 0 for simplicity, where we assume the existence of the side limits. The argument used in the proof of Theorem 3.5.2 shows the existence of the limits

$$\lim_n \int_{-\pi}^{0} (f(t) - f(0^-)) \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)}\, dt = 0;$$

$$\lim_n \int_{0}^{\pi} (f(t) - f(0^+)) \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)}\, dt = 0.$$

Therefore

$$S_n(f)(0) = \lim_n \int_{-\pi}^{\pi} f(t) \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)}\, dt$$

$$= \lim_n \int_{0}^{\pi} f(0^+) \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)}\, dt + \lim_n \int_{-\pi}^{0} f(0^-) \frac{\sin((n+1/2)t)}{2\pi \sin(t/2)}\, dt$$

$$= \frac{1}{2} f(0^+) + \frac{1}{2} f(0^-)$$

that is the desired result. ■

However, despite this regularity of the limits at jump discontinuities, the behavior of the partial sums of the series is highly irregular since there are oscillations above and below the graph of f about $\frac{1}{12}|f(x^+) - f(x^-)|$. This is the so called *Gibbs phenomenon*.

3.6 Summation methods

The lack of convergence (pointwise or uniform) of the Fourier series of a continuous function (to be proved in the next section) can be amended by considering weaker definitions of convergence. Here we will consider the convergences in the senses of Cesàro and Abel applied to series. A main tool in this section is the following, that can be stated in $\mathbb{T}$, $\mathbb{R}$ or in more dimensions.

Definition 3.6.1. *A sequence (K_n) of functions is called a good kernels sequence if satisfies the following properties:*

(a) $K_n \geq 0$*;*

(b) $\int K_n = 1$*;*

(c) $\lim_n K_n(t) = 0$ *uniformly out of every neighbourhood of* 0*;*

(d) $\lim_n \int_V K_n = 1$ *for every V a neighbourhood of* 0.

The definition can be adapted with the suitable changes for families depending on a continuous parameter.

Note that $(b) + (c) \Rightarrow (d)$ in case we are in $\mathbb{T}$.

Theorem 3.6.2. *Let (K_n) be a good kernels sequence on $\mathbb{R}$ or $\mathbb{T}$ and $f \in L^1$. Then the following statements hold:*

(a) $\lim_n (K_n * f)(x) = f(x)$ *if f is continuous at x;*

(b) $\lim_n (K_n * f) = f$ *uniformly if f is uniformly continuous;*

(c) $\lim_n (K_n * f) = f$ *with respect to the norm* $\|\cdot\|_1$.

Proof. (a) Since the result does not change by constant products, without loss of generality we may assume $0 \leq f(x) \leq 1$. Find a symmetric neighborhood V of 0 such that

$$f(x) - \varepsilon \leq f(x-t) \leq f(x) + \varepsilon$$

whenever t is such that $t \in V$. For n large enough we may ask $\int_V K_n \geq 1 - \varepsilon$ and $|K_n(t)| < \varepsilon / \|f\|_1$ for $t \in V^c$. With those choices we have

$$f(x) - 2\varepsilon \leq \int_V f(x-t) K_n(t) dt < f(x) + \varepsilon$$

$$\left|\int_{V^c} f(x-t)K_n(t)dt\right| < \varepsilon.$$

Therefore

$$|(K_n * f)(x) - f(x)| \leq 3\varepsilon.$$

(b) Just notice that the proof of (a) works for V a neighbourhood of 0 witnessing the uniform convergence.
(c) Finally, let us assume that the functions (K_n) share a bounded support. If $f \in L^1$, for any $\varepsilon > 0$ we can take g continuous with bounded support such that $\|f - g\|_1 < \varepsilon$. We have

$$\|K_n * f - K_n * g\|_1 = \|K_n * (f - g)\|_1 \leq \|K_n\|_1 \|f - g\|_1 \leq \varepsilon.$$

Since the supports of the functions $K_n * g$ are contained in the same set, the uniform convergence implies the $\|\cdot\|_1$-convergence. Take n large enough such that $\|K_n * g - g\|_1 \leq \varepsilon$. All these inequalities together impliy$\|K_n * f - f\|_1 \leq 3\varepsilon$. In case the (K_n) do not have a common support, take

$$K_n^\varepsilon = \max\{0, K_n - \varepsilon\}$$

for some $\varepsilon > 0$. The functions (K_n^ε) made up a good kernels sequence up to a factor converging to 1. Using that $\|K_n^\varepsilon - K_n\|_1$ can be made arbitrarily small as n increases, the general result follows now from the one with uniformly bounded supports. ■

Cesàro summation. Recall that a sequence (x_n) is said convergent in the sense of Cesàro if the sequence of arithmetic means

$$\frac{x_1 + \cdots + x_n}{n}$$

is convergent, and its limit is called the limit in the sense of Cesàro. The method of Cesàro extends usual convergence since a convergent sequence is always convergent in sense of Cesàro. In case of a (formal) series $\sum_{n=0}^{\infty} x_n$, we say that the series converges in sense of Cesàro if the partial sums do. Namely, if $s_n = \sum_{k=0}^{n} x_n$ then the limit of

$$\frac{s_0 + \cdots + s_{n-1}}{n}$$

is the Cesàro sum of the series.

In the case of the Fourier series of a function f, the Cesàro means can be expressed this way

$$\frac{S_0(f)+\cdots+S_{n-1}(f)}{n}=\left(\frac{D_0+\cdots+D_{n-1}}{n}\right)*f$$

which means that we just need to study what happens with the Cesàro operation applied to the Dirichlet kernel. For that, we will start by computing

$$\sin(t/2)(D_0+\cdots+D_{n-1})=\frac{1}{2\pi}(\sin(t/2)+\sin(3t/2)+\cdots+\sin((n-1/2)t))$$

$$=\frac{1}{2\pi}\Im(e^{it/2}+e^{3it/2}+\cdots+e^{i(n-1/2)t})=\frac{1}{2\pi}\Im\left(e^{it/2}\frac{1-e^{int}}{1-e^{it}}\right)$$
$$=\frac{1}{2\pi}\Im\left(\frac{1-e^{int}}{e^{-it/2}-e^{it/2}}\right)=\frac{1}{2\pi}\Im\left(\frac{1-e^{int}}{-2i\sin(t/2)}\right)=\frac{1}{2\pi}\left(\frac{1-\cos(nt)}{2\sin(t/2)}\right)$$
$$=\frac{1}{2\pi}\frac{\sin^2(nt/2)}{\sin(t/2)}.$$

Therefore, if we define the *Fejér kernel* as

$$F_n(t):=\frac{D_0+\cdots+D_{n-1}}{n}=\frac{1}{2\pi n}\left(\frac{\sin(nt/2)}{\sin(t/2)}\right)^2$$

the Cesàro means of the Fourier series reduce to

$$\frac{S_0(f)+\cdots+S_{n-1}(f)}{n}=F_n*f.$$

The main difference between the Fejér kernel and the Dirichlet kernel is that the first one is positive. That makes of (F_n) a good sequence of kernels, what implies that

$$\lim_n(F_n*f)(t)=f(t)$$

at any $t\in\mathbb{T}$ where f is continuous. We will state the definitions and result.

Theorem 3.6.3. *If $f\in C(\mathbb{T})$, then F_n*f converges uniformly to f.*

Proof. We have to prove that (F_n) is a sequence of good kernels, but if follows easily from the very definition since $\int_{\mathbb{T}}F_n=1$ and

$$F_n(t)\le\frac{1}{2\pi n(\sin\delta/2)^2}$$

for $t \notin (-\delta, \delta)$ for any $0 < \delta\pi$. ∎

We can recover Weierstrass theorem on approximation of 2π-periodic continuous functions by trigonometric polynomials from the previous result. We also get the uniqueness of Fourier coefficients.

Corollary 3.6.4. *If two functions $f, g \in L^1(\mathbb{T})$ have the same Fourier series (coefficients), then they are equal as elements of $L^1(\mathbb{T})$, that is, $f = g$ almost everywhere.*

Abel summation. Abel proved the following remarkable theorem.

Theorem 3.6.5. *Let $\sum_{n=0}^{\infty} a_n$ be a convergent series of complex numbers. Then the power series*

$$\sum_{n=0}^{\infty} a_n x^n$$

converges uniformly on $[0, 1]$*. As a consequence*

$$\lim_{x \to 1^-} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n.$$

Sketch of proof. Use the Abel transformation for series, which is a sort of integration by parts. ∎

Let $\sum_{n\in\mathbb{Z}} a_n e^{inx}$ be the Fourier series of some $f \in L^1(\mathbb{T})$. For $r \in [0, 1)$ we may consider the modified series

$$S(f, r) = \sum_{n\in\mathbb{Z}} a_n r^{|n|} e^{inx}.$$

Note that the series is absolutely convergent and, therefore, the manipulation which follow are justified

$$S(f, r)(x) = \sum_{n\in\mathbb{Z}} \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} f(t) e^{-int} dt \right) r^{|n|} e^{inx} = \int_{-\pi}^{\pi} f(t) \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} r^{|n|} e^{in(x-t)} dt$$

so it is enough to study the kernel

$$\sum_{n\in\mathbb{Z}} r^{|n|} e^{in\theta} = \sum_{n=0}^{\infty} r^n e^{in\theta} + \sum_{n=1}^{\infty} r^n e^{-in\theta} = \frac{1}{1 - re^{i\theta}} + \frac{re^{-i\theta}}{1 - re^{-i\theta}}$$

$$= \frac{1 - re^{-i\theta} + re^{-i\theta} - r^2}{1 - re^{i\theta} - re^{-i\theta} + r^2} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}.$$

where the change of variable to θ is motivated by the convenience of interpreting the pair (r, θ) as polar coordinates in the plane .The function

$$P_r(\theta) = \frac{1}{2\pi} \cdot \frac{1 - r^2}{1 - 2r\cos\theta + r^2},$$

with r playing the role of parameter, is the so called *Poisson kernel*. It is not difficult to check that $P_r(\theta)$ is a good family of kernels as $r \to 1^-$. Therefore

$$\lim_{r \to 1^-} (P_r * f)(x) = f(x)$$

at any $x \in \mathbb{T}$ where f is continuous. For $f \in C(\mathbb{T})$ we will state a result that we will use later.

Theorem 3.6.6. *Let $f \in C(\mathbb{T})$, then $\lim_{r \to 1^-} P_r * f = f$ uniformly.*

Hint of proof. Find the best bounds to show that the Poisson kernel is a good family. ■

It is also interesting to note that

$$P_r(\theta) = \frac{1}{2\pi} \Re \left(\frac{1 + re^{i\theta}}{1 - re^{i\theta}} \right).$$

3.7 The failure of convergence of Fourier series

This section, that could be omitted in a first reading, requires some notions on bounded linear operators on Banach spaces. Note that the Dirichlet kernel expresses an orthogonal projection in the Hilbert space $L^2(\mathbb{T})$, and so its has norm 1 when we measure with $\| \cdot \|_2$. However, considered on other function spaces that is not longer true.

Proposition 3.7.1. *The norm of the operator $\mathcal{D}_n$ induced by the convolution against D_n, either on $(C(\mathbb{T}), \| \cdot \|_\infty)$ or $(L^1(\mathbb{T}), \| \cdot \|_1)$, is $\|D_n\|_1$.*

Proof. The function $\operatorname{sign}(D_n(t))$ can be approximated arbitrarily in $\| \cdot \|_1$ by continuous functions with values in $[-1, 1]$, thus

$$\|\mathcal{D}_n\| \geq \sup\{|(D_n * f)(0)| : f \in C(\mathbb{T}), |f| \leq 1\} = \int_{-\pi}^{\pi} |D_n(t)| dt = \|D_n\|_1.$$

On the other hand

$$\|\mathcal{D}_n(f)\|_\infty = \|D_n * f\|_\infty \leq \int_{-\pi}^{\pi} \|f\|_\infty |D_n(t)| dt = \|f\|_\infty \|D_n\|_1.$$

Therefore $\|\mathcal{D}_n\| = \|D_n\|_1$.
For the $L^1(\mathbb{T})$ case, consider the functions $K_m(t) = 2n$ if $-1/m \leq t \leq 1/m$ and $K_m(t) = 0$ otherwise, for $m \in \mathbb{N}$. It easy to see that (K_n) is a good kernel sequence. Therefore,

$$\lim_m \mathcal{D}_n(K_m) = \lim_m D_n * K_m = D_n$$

in $\|\cdot\|_1$. We deduce $\|\mathcal{D}_n\| \geq \|D_n\|_1$ since (K_m) is a norm one sequence. On the other hand,

$$\|\mathcal{D}_n(f)\|_1 = \|D_n * f\|_1 \leq \|D_n\|_1 \|f\|_1.$$

Therefore $\|\mathcal{D}_n\| = \|D_n\|_1$. ■

Proposition 3.7.2. *The following estimation holds*

$$\|D_n\|_1 \geq \frac{8}{\pi} \sum_{k=1}^{n} \frac{1}{k}.$$

In particular, $\lim_n \|D_n\|_1 = +\infty$.

Proof. Since $|\sin t| \leq t$ we have

$$\|D_n\|_1 = \int_{-\pi}^{\pi} |D_n(t)| dt \geq 4 \int_0^{\pi} |\sin((n+1/2)t)| t^{-1} dt$$

$$= 4 \int_0^{\pi(n+1/2)} |\sin t| t^{-1} dt \geq 4 \sum_{k=1}^{n} \int_{\pi(k-1)}^{\pi k} |\sin t| (\pi k)^{-1} dt = \frac{8}{\pi} \sum_{k=1}^{n} \frac{1}{k}$$

as claimed. ■

We need the following classic result of Banach and Steinhaus.

Theorem 3.7.3. *Let X and Y be Banach space and $(T_n) \in \mathcal{L}(X, Y)$ a sequence of bounded operators such that $\lim_n T_n(x)$ exists for every $x \in X$. Then* $\sup\{\|T_n\| : n \in \mathbb{N}\} < +\infty$ *and $T(x) := \lim_n T_n(x)$ defines a bounded operator.*

Hint of proof. Assuming the pointwise convergence, then every sequence $(\|T_n(x)\|)$ is bounded. We can write as a countable union of closed sets

$$X = \bigcup_{k=1}^{\infty} \{x \in X : \sup\{\|T_n(x)\| : n \in \mathbb{N}\} \leq k\}.$$

Baire's theorem implies that there is some $k \in \mathbb{N}$ such that

$$\{x \in X : \sup\{\|T_n(x)\| : n \in \mathbb{N}\} \leq k\}$$

has nonempty interior, that is, it contains a ball of positive radius. Without loss of generality, the ball is cnetered at 0 and by scaling we deduce

$$\sup\{\|T_n(x)\| : n \in \mathbb{N}\}$$

is uniformly bounded for $x \in B_X$. That implies the boundedness of the set $\sup\{\|T_n\| : n \in \mathbb{N}\}$ and the limit operator T. ■

Corollary 3.7.4. *The Fourier series does not converge in* $(L^1(\mathbb{T}), \|\cdot\|_1)$ *or pointwise in* $C(\mathbb{T})$ *in general.*

Proof. The combination of the previous results give the convergence in $(L^1(\mathbb{T}), \|\cdot\|_1)$. For the pointwise, just consider the operator $f \to (D_n * f)(0)$ from $C(\mathbb{T})$ to $\mathbb{R}$, whose norm is $\|D_n\|_1$ too. ■

Remark 3.7.5. *The non fulfillment of the hypothesis in Theorem 3.7.3 implies that the points* $x \in X$ *where* $\lim_n T_n(x)$ *does not exists is dense. Moreover, it is residual set, that implies nonempty intersection with other residual sets in* X*, even countably many. Using that, we can deduce the existence of function* $f \in C(\mathbb{T})$ *such that its Fourier series diverges at a dense set of* $C(\mathbb{T})$*.*

3.8 Series in several variables

On $L^2(\mathbb{T}^2)$ we can define an hermitian product by

$$\langle f, g\rangle = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(x,y)\,\overline{g(x,y)}\,dx\,dy$$

that makes it a Hilbert space. It is not difficult to show that the set of functions

$$\{e^{inx+imy} : n, m \in \mathbb{Z}\}$$

is orthogonal. The corresponding Fourier coefficients are obtained by

$$a_{n,m} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x,y)\, e^{-inx-imy}\, dx\, dy$$

and the Fourier series is

$$\sum_{n,m\in\mathbb{Z}} a_{n,m}\, e^{inx+imy}.$$

The Hilbert theory extends without trouble since there is no difference between simple and double indexing for unconditionally convergent series. The denseness of the functions can be deduced again from Stone-Weierstrass theorem. On the other hand, the theory for pointwise and uniform convergence requires some development: the formulas for the kernels vary, but at the end the behavior is similar.

We have used a series of functions of two variables in relation with Poisson summation method. In the solution, theoretical and practical, of PDEs power series in several variables are a valuable tool. Here we will consider power series in several variables, but in practise two variables are enough to show the general behavior. A general power series in two variables looks like

$$\sum_{n,m} a_{nm}(x-x_0)^n(y-y_0)^m$$

where we understand that the summation is over all integers $n, m \geq 0$.

Theorem 3.8.1. *The power series is uniformly convergent on the set*

$$\{(x,y) : |x-x_0| \leq r, |y-y_0| \leq r\}$$

if r satisfies

$$0 \leq r < \frac{1}{\limsup_{n,m} \sqrt[n+m]{|a_{nm}|}}$$

Proof. Without loss of generality we may assume $x_0 = y_0 = 0$. The inequality above implies the existence of some $0 < \lambda < 1$ such that

$$|a_{n,m}| r^{n+m} \leq \lambda^{n+m}$$

for all but finitely many indices (n,m). Now note that the series

$$\sum_{n,m} \lambda^{n+m} = (\sum_n \lambda^n)^2 < +\infty$$

is convergent and so
$$\sum_{n,m} |a_{n,m}| r^{n+m} < +\infty.$$
Therefore, the uniform convergence follows now from Weierstrass criterion since
$$|a_{nm}||x|^n |y|^m \leq |a_{n,m}| r^{n+m}.$$
■

The functions that admit a development as power series at any point of their domain are called *analytical.*

3.9 Some other orthogonal functions

Legendre polynomials. We know that the sequence $\{x^n : n \geq 0\}$ has dense linear span in $(C[a,b], \|\cdot\|_\infty)$ for any interval $[a,b] \subset \mathbb{R}$. Fix the space of continuous functions $C[-1,1]$ and consider the scalar product
$$\langle f, g\rangle = \int_{-1}^{1} f(x)\, g(x)\, dx.$$
The Gram-Schmidt method applied to the sequence of powers cast a sequence of polynomials, named *Legendre polynomials*, whose first terms are
$$P_0(x) = \frac{1}{\sqrt{2}}, \quad P_1(x) = \sqrt{\frac{3}{2}}\, x, \quad P_2(x) = \sqrt{\frac{5}{8}}\,(3x^2-1), \ldots$$
As the algorithm is getting complicated as the degree increases, there are alternative methods to deal with the sequence of Legendre polynomials. In particular, we have the remarkable Rodrigues formula which provides an explicit expression
$$P_n(x) = \frac{\sqrt{n+1/2}}{2^n\, n!}\, \frac{d^n}{dx^n}\left((x^2-1)^n\right) = \frac{\sqrt{n+1/2}}{2^n} \sum_{k=0}^{n} \binom{n}{k}^2 (x+1)^{n-k}(x-1)^k.$$

Weighted orthogonality. We may consider a continuous function $\rho(x)$ such that $\rho(x) > 0$ for some $I \subset \mathbb{R}$. On the space of those continuous functions f such that
$$\int_I f(x)^2\, \rho(x)\, dx < +\infty$$

we may consider the scalar product

$$\langle f, g \rangle = \int_I f(x)\, g(x)\, \rho(x)\, dx.$$

Important particular cases are the following: if $I = [-1, 1]$ and

$$\rho(x) = \frac{1}{\sqrt{1 - x^2}}$$

the orthogonalization of $\{x^n : n \geq 0\}$ produces the so called *Chebyshev polynomials* that can be obtained explicitly from

$$T_n(\cos\theta) = \cos(n\theta).$$

Now, taking $I = \mathbb{R}$ and the weight

$$\rho(x) = e^{-x^2}$$

the orthogonalization of $\{x^n : n \geq 0\}$ produces the *Hermite polynomials* which are very important in the discussion of solutions of the Schrödinger equation.

Sturm-Liouville systems. A generalization of the properties of the trigonometric system can be obtained by generalizing the differential equation they satisfy. Consider a the following second order ODE

$$\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) - q(x)u + \lambda\rho(x)u = 0$$

where the functions p, q, ρ are defined on $[0, 1]$ where $p, \rho > 0$ and $q \geq 0$. We are interested in the values of $\lambda \in \mathbb{R}$ for which the equation has a non trivial solution u such that $u(0) = u(1) = 0$. Those values are called *eigenvalues* and the corresponding solutions *eigenfunctions* by similarity to linear operators. Actually, there is more than similarity if we could take the point of view from Functional Analysis. Every eigenvalue has, essentially one eigenfunction up to multiplicative constant. The eigenvalues are positive and made up an increasing sequence

$$0 < \lambda_1 < \lambda_2 < \cdots \to \infty.$$

Two eigenfunctions corresponding to different eigenvalues are orthogonal with respect to the weighted scalar product

$$\langle f, g \rangle = \int_0^1 f(x)\, g(x)\, \rho(x)\, dx.$$

The normalized sequence of eigenfuctions (u_n) is a Hilbert basis and thus any $f \in C[0,1]$ has a representation as "Fourier series"

$$f \sim \sum_n a_n \, u_n.$$

The actual uniform convergence of the series to f depends on additional conditions of regularity of f.

Bessel functions. Unfortunately, there are interesting second order ODEs that do not fulfill all the requirements of a Sturm-Liouville system. Consider the equation

$$\frac{d}{dx}\left(x\frac{du}{dx}\right) - \frac{m^2}{x}u + \lambda x u = 0$$

where $m \geq 0$ is an integer constant. Evidently, the greatest problem is to deal with the singularity at $x = 0$. Instead of looking for a solution that vanishes at $x = 0$, we will look for u bounded defined on $(0,1]$ such that $u(1) = 0$ and

$$\lim_{x \to 0} x u'(x) = 0.$$

The change of variable $t = \sqrt{\lambda}\, x$ transforms the equation into

$$\frac{d}{dt}\left(t\frac{du}{dt}\right) - \frac{m^2}{t}u + t u = 0$$

or equivalently

$$t^2\frac{d^2u}{dt^2} + t\frac{du}{dt} + (t^2 - m^2)u = 0$$

that it is know as the *Bessel equation*. The boundary condition at $x = 1$ is now $u(\sqrt{\lambda}) = 0$. The unique solution, up to a multiplicative constant, of the differential equation satisfying being bounded at $t = 0$ is developable into a power series as

$$J_m(t) = \sum_{n=0}^{\infty} \frac{(-1)^n (t/2)^{2n+m}}{n!(n+m)!}$$

which is called the m-th *Bessel function*. All the Bessel functions have infinitely many zeroes on $\mathbb{R}^+$, which means that our original "almost Sturm-Liouville" problem has an infinite set of eigenvalues. As before, it is feasible to develop any continuous function on $[0,1]$ as a Bessel-Fourier series. For some future use of the Bessel functions we need to define J_m for $m < 0$: just take the series above and take ∞ as the value of the factorial of a negative integer. With that definition, it is easy to check that

$$J_{-m}(t) = (-1)^m J_m(t).$$

3.10 Rationale and remarks

The Hilbert theory (orthogonality) is the main motivation for the notion of Fourier series, however it is not enough for the representation of a function as a series, which requires pointwise convergence. It is quite surprising the fact that the convergence of the Fourier series of a function f at some point depends on the local behavior of f at x. Let us mention a historic curiosity: the early investigations of George Cantor on the uniqueness of Fourier coefficients lead to the invention of Set Theory.

It is possible to prove a *uniform Dini condition* that gives a more accurate description of the sets of uniform convergence of the Fourier series of a quite regular but discontinuous function (that forbids global uniform convergence). The proof depends on a quantitative version of the Riemann-Lebesgue lemma 3.3.4.

Carleson proved a deep result establishing that the Fourier series of any function in $L^2(\mathbb{T})$, in particular piecewise continuous, is convergent almost everywhere.

The proof we provided of the existence of continuous functions having a non convergent Fourier series was non constructive and relies in Baire's theorem, through Banach-Steinhaus. Mathematics are plenty of nonconstructive existence proofs. Besides Baire arguments, there are fixed point, combinatorial and probabilistic arguments.

Analytic functions of several variables is another tool to investigate the existence of solutions of a PDE. The key result in this area is the celebrated Cauchy-Kowalewski theorem.

3.11 Exercises

1. Prove that the strict convexity of the norm is equivalent to the following condition: whenever
$$\|x\| = \|y\| = \left\| \frac{x+y}{2} \right\|$$
then $x = y$. Equivalently, the unit sphere does not contain nontrivial segments.

2. Show that the coefficients of the real Fourier expansion 3.1 are related to the complex ones 3.2 by
$$\hat{f}(0) = \frac{a_0}{2};\ \hat{f}(n) = \frac{a_n - ib_n}{2};\ \hat{f}(-n) = \frac{a_n + ib_n}{2} \quad \text{for} \quad n > 0.$$
Deduce, as a consequence, the real version of Parseval formula
$$\|f\|_2^2 = \pi a_0^2/2 + \pi \sum_{n=1}^{\infty} (a_n^2 + b_n^2).$$

3. Find the Fourier expansions on $[-\pi, \pi]$ of the following functions:
   (a) $f_1(x) = (\pi/4)\mathrm{sign}(x)$
   (b) $f_2(x) = (\sin x)^3$
   (c) $f_3(x) = \pi/2 - |x|$
   (d) $f_4(x) = x^2$
   (e) $f_5(x) = x^4$
   (f) $f_6(x) = |\sin x|$
   (g) $f_7(x) = \cos(x/2)$
   (h) $f_8(x) = 1 - |x|/\delta$ for $|x| \le \delta$
   (i) $f_9(x) = \sinh x$

4. Find the Fourier series of the function $\sin(x/2)$ on $[-\pi, \pi]$ and describe its convergence. Which sets is the convergence uniforme on?

5. Find the Fourier series of the function defined by

$$f(x) = \begin{cases} -x - \pi, & \text{si } x \in [-\pi, -\pi/2] \\ x, & \text{si } x \in (-\pi/2, \pi/2] \\ \pi - x, & \text{si } x \in (\pi/2, \pi] \end{cases}$$

and describe its convergence.

6. Find the Fourier series of the function defined by

$$f(x) = \begin{cases} x(\pi + x), & \text{si } x \in [-\pi, 0] \\ x(\pi - x), & \text{si } x \in (0, \pi] \end{cases}$$

and describe its convergence.

7. Assuming the fact that $\sum_{n=1}^{\infty} n^{-2} = \pi^2/6$, compute $\sum_{n=1}^{\infty} (-1)^{n+1} n^{-2}$.

8. Use Fourier theory to get the following sums:

   (a) $\sum_{n=1}^{\infty} n^{-4} = \pi^4/90$;
   (b) $\sum_{n=1}^{\infty} n^{-6} = \pi^6/945$;
   (c) $\sum_{n=1}^{\infty} n^{-8} = \pi^8/9450$.

9. Denote by $\{x\}$ the fractional part of a number $x \in \mathbb{R}$. Prove that for every $f \in C[0,1]$ and $\alpha \in [0,1] \setminus \mathbb{Q}$ we have

$$\int_0^1 f(t)\,dt = \lim_n \frac{1}{n} \sum_{k=1}^{n} f(\{k\alpha\}).$$

   Prove that the result extends to ruled functions, and therefore characteristic functions of intervals.

10. Prove the pointwise convergence of the following trigonometric series

$$\sum_{n=2}^{\infty} |\log n|^{-1} \sin nx,$$

   however it is the Fourier series of no function from $L^1(\mathbb{T})$.

11. Compute the value of the series

$$\sum_{n=1}^{\infty} 2^{-n} \sin nx.$$

12. Prove the following form of the Dirichlet test for series: Let $(x_n) \subset X$ be a sequence in a (real or complex) Banach space X such that the partial sums $\sum_{k=1}^{n} x_k$ are uniformly bounded, then the series $\sum_{n=1}^{\infty} a_n x_n$ converges in X for every sequence (a_n) of positive real numbers that converges monotonically to 0.

13. Let X be a Hilbert space.

    (a) Prove the generalized parallelogram identity

    $$\sum_{k=1}^{n} \|x_k\|^2 = 2^{-n} \sum_{(\epsilon_k)_{k=1}^{n} \in \{-1,1\}^n} \|\sum_{k=1}^{n} \epsilon_k x_k\|^2$$

    (b) Let $(x_n) \subset X$ be a sequence of pairwise orthogonal vector. Prove that the series $\sum_{n=1}^{\infty} x_n$ converges if and only if $\sum_{n=1}^{\infty} \|x_n\|^2 < +\infty$.

    (c) Show moreover that under the previous assumptions the series is unconditionally convergent, that is, the convergence and limit are stable by reordering of the terms.

    (d) Let $\sum_{n=1}^{\infty} x_n$ be an unconditionally convergent in X, not necessarily orthogonal. Prove that, in such a case, $\sum_{n=1}^{\infty} \|x_n\|^2 < +\infty$.

Chapter 4

The wave equation

After the discussion in the first chapter, we call wave equation in n dimensions to

$$u_{tt} - c^2 \Delta u = 0,$$

where $u = u(\mathbf{x}, t)$ with $t \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ and the Laplace operator Δ is taken with respect to the spatial coordinates. Obviously, the real constant c can be taken to be positive and for reasons that will be clear in a moment it can be interpreted as the speed of the wave.

4.1 D'Alembert's solution in one dimension

For $n = 1$, the wave equation reduces to

$$u_{tt} = c^2 u_{xx}.$$

In that particular case, d'Alembert discovered that any function of the form

$$u(x,t) = \phi(x - ct) + \psi(x + ct),$$

where ϕ, ψ are second order differentiable real functions, is a solution of the equation (we have seen that as an application of the symbolic method). Actually, every C^2-solution of the 1-dimensional wave equation is of that form. Indeed, the change of variable $r = x - ct, s = x + ct$ reduces the equation to the form $w_{rs} = 0$ that can be solved by direct computation. We will see later that this fact leads to a "controversy" when compared with a different method of solution.

Note that the value of $u(x,t)$ for $t>0$ depends on $u(x-ct,0)$ and $u(x+ct,0)$. For that reason, the interval $[x-ct,x+ct]$ is called the *domain of dependence* for (x,t). Analogously, the value $u(x,0)$ affects to the values $u(x-ct,t)$ and $u(x+ct)$, so the interval $[x-ct,x+ct]$ is called the *domain of influence* of $(x,0)$. It is not difficult to prove that the solution of the one-dimensional wave equation satisfies the following *functional equation*

$$u(x,t)-u(x+c\alpha,t+\alpha)-u(x-c\beta,t+\beta)+u(x+c\alpha-c\beta,t+\alpha+\beta)=0$$

which has a nice geometrical interpretation: on the space-time plane, the four points are the vertices of a parallelogram made of *characteristic lines* (straight lines with slope $\pm c$).

Example 4.1.1. *Suppose that the solution u of the equation $u_{tt}=u_{xx}$ is known for $(x,0)$ and (x,x) with $0\le x\le a$. Find the domain where we could deduce the values of u.*

The formula of d'Alembert gives us that

$$u(x,t)=\phi(x-t)+\psi(x+t).$$

The information we have about u implies

$$u(x,0)=\phi(x)+\psi(x);\quad u(x,x)=\phi(0)+\psi(2x).$$

Without loss of generality we may assume $\phi(0)=0$. Therefore, ψ is determined for $x\in[0,2a]$, whereas ϕ is for $x\in[0,a]$. Now observe that we can calculate $u(x,t)$ whenever $0\le x-t\le a$ and $0\le x+t\le 2a$. Clearly, those equations delimitate a parallelogram. ∎

Now, we will deduce the solution to the initial value problem. We wish to find a function $u(x,t)$ that satisfies $u(x,0)=f(x)$ and $u_t(x,0)=g(x)$, for f,g enough regular functions. The initial value problem means here that at $t=0$ we know the position and speed of the particles subdued to the dynamical equation (remember that the wave equation comes from Newton's second law), therefore future positions and speeds are determined. The following is called the *d'Alembert formula for initial values.*

Theorem 4.1.2. *Let $f\in C^2(\mathbb{R})$ and $g\in C^1(\mathbb{R})$. Then there is a unique solution of $u_{tt}=c^2u_{xx}$ on $\mathbb{R}^2$ such that $u(x,0)=f(x)$ and $u_t(x,0)=g(x)$ given by*

$$u(x,t)=\frac{1}{2}(f(x+ct)+f(x-ct))+\frac{1}{2c}\int_{x-ct}^{x+ct}g(s)\,ds.$$

Proof. D'Alembert general solution of the 1-dimensional wave equation applied in this case tell us that

$$f(x) = u(x,0) = \phi(x) + \psi(x)$$

$$g(x) = u_t(x,0) = -c\phi'(x) + c\psi'(x)$$

Differentiating the first equation with respect to x gives

$$f'(x) = \phi'(x) + \psi'(x),$$

and solving the linear system we have

$$\phi'(x) = \frac{cf'(x) - g(x)}{2c}, \qquad \psi'(x) = \frac{cf'(x) + g(x)}{2c}.$$

Therefore

$$\phi(x) = \frac{f(x)}{2} - \frac{1}{2c}\int_0^x g(s)\,ds + cte,$$

$$\psi(x) = \frac{f(x)}{2} + \frac{1}{2c}\int_0^x g(s)\,ds + cte,$$

and thus

$$u(x,t) = \frac{1}{2}(f(x+ct) + f(x-ct)) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\,ds$$

where the sum of the added unknown constants turns out to be 0. Uniqueness of the solution comes from the complete determination of ϕ and ψ by f and g.■

Note that d'Alembert formula represents any C^2 solution of the wave equation in terms of the initial values, however it makes sense for less regular initial data, namely $f \in C^1(\mathbb{R})$ and $g \in C(\mathbb{R})$. We will develop this argument further later. Also, the formula allows to establish the continuous dependence on the initial conditions f and g over a finite time interval $[0, T]$. For zero speed at $t = 0$, the continuous dependence extends to $[0, +\infty)$.

4.2 The vibrating string with fixed endpoints

One of the physical cases leading to the wave equation was the vibrating string whose endpoints are fixed, we may assume that at $x = 0$ and $x = L$. Since the solution should be contained in d'Alembert's expression, we have

$$\phi(0 - ct) + \psi(0 + ct) = 0,$$

$$\phi(L-ct)+\psi(L+ct)=0,$$

for all t. Clearly we can change ct by x. The first one, obviously implies $\psi(x)=-\phi(-x)$, that applied to the second equation gives

$$\psi(L+x)=-\phi(L-x)=\psi(x-L)$$

which means that ψ (and, thus ϕ) is $2L$-periodic. The solution, therefore, can be written as

$$u(x,t)=\phi(x-ct)-\phi(-x-ct).$$

One can think of that as if the traveller waves were reflected in "antisymmetric mirrors" at $x=0$ and $x=L$. Since the periodicity of the function affects both in space and time, an easy computation gives that the vibration has period $T=2L/c$.

The application of the formula given in Theorem 4.1.2 requires the extension of the initial conditions f and g to $\mathbb{R}$ in a suitable way to $[-L,0]$ expressed by the fact that $\psi(x)=-\phi(-x)$ must be fulfilled. That is easier if we assume, for instance, that $u_t(x,0)=0$.

Example 4.2.1. *Obtain the solution of the vibrating string with an arbitrary* $u(x,0)=f(x)$ *and* $u_t(x,0)=0$ *for* $x\in[0,L]$.

We have, in general, $u(x,t)=\phi(x-ct)-\phi(-x-ct)$. The fact $u_t(x,0)=0$ easily implies $\phi'(x)=\phi'(-x)$ so ϕ differs from an odd function in a constant that we may assume it is 0 without loss of generality. Now

$$f(x)=u(x,0)=\phi(x)-\phi(-x)=2\phi(x).$$

That implies we should define f on $[-L,0]$ by $f(x)=-f(-x)$ and then on all $\mathbb{R}$ by $2L$-periodicity. Therefore we have

$$u(x,t)=\frac{1}{2}(f(x-ct)-f(-x-ct))=\frac{1}{2}(f(x-ct)+f(x+ct)),$$

that is the desired formula. This expression agrees with d'Alembert formula but the fact that we did not know how to extend f to $\mathbb{R}$. ■

We will show a curious behaviour of some particular solutions of the wave equation. In order to simplify, assume $c=1$ and $L=\pi$ in what follows. Since sin is 2π-periodic we can use it to define an explicit solution

$$u(x,t)=\sin(x-t)-\sin(-x-t)=\sin(x-t)+\sin(x+t)=2\sin x\cos t.$$

That kind of solution is called *stationary* because it does not give the impression of moving to any of the sides. In general, a stationary solution is of the form

$$u(x,t) = \phi(x)\psi(t).$$

The existence of stationary solutions and their application to the general problem will be discussed in relation with the *separation of variables method*.

4.3 Non-differentiable solutions

It may seem strange to discuss non-differentiable solutions of differential equations but there are several motivations, practical and theoretical. For instance, consider the wave equation

$$u_{tt} = u_{xx}$$

with fixed endpoints $u(0,t) = u(1,t) = 0$ and with initial condition

$$u(x,0) = \min\{x, 1-x\}$$

and initial speed $u_t(x,0) = 0$ for $x \in [0,1]$. According to the differential equation model, the string should not move at all because the initial speed is 0 and there is no force towards the X-axe. However, nobody would not deny that an elastic string with that initial shape will move. Of course, one could argue that the real string is not the mathematical model: the real string has thickness and the shape with a non differentiable point is just an approximation.

Let's make an experiment: we will use d'Alembert's solution with the previous initial data. For that, define

$$f(x) = \min\{x, 1-x\}$$

on $[0,1]$. Extend f to $[-1,0]$ by $f(x) = -f(-x)$. Finally, extend f to all $\mathbb{R}$ as a 2-periodic function. D'Alembert formula (Theorem 4.1.2 or Example 4.2.1) for initial null speed produces

$$u(x,t) = \frac{1}{2}(f(x-t) + f(x+t)).$$

Actually, we can get a more simplified expression for $t \in [0,1]$

$$u(x,t) = \min\{x, 1-x, t\}.$$

Therefore, d'Alembert solution moves and the differential equation is satisfied in an awkward way: note that the central horizontal segment that appears for $t > 0$ moves with uniform speed, so $u_{tt} = 0$ on it, as well as $u_{xx} = 0$ trivially.

We can motivate the validity of that solution as follows. Let $\varepsilon > 0$ and consider $g \in C^2(\mathbb{R})$ a 2-periodic function such that $\|f - g\|_\infty < \varepsilon$. If $v(x,t)$ is the solution of

$$v_{tt} = v_{xx}$$

with fixed endpoints $v(0,t) = v(1,t) = 0$ and with initial condition $v(x,0) = g(x)$ and $v_t(x,0) = 0$ for $x \in [0,1]$, then we have

$$v(x,t) = \frac{1}{2}(g(x-t) + g(x+t))$$

and $\|u(x,t) - v(x,t)\|_\infty < \varepsilon$ for all $x \in [0,1]$ and $t \in \mathbb{R}$. No one could deny that v follows the mathematical model. Therefore, u starting from similar initial conditions just follows a similar pattern.

There is a deeper reason to accept non-differentiable solutions that is out of the scope of this course. Differential equations seems to be applied to differentiable functions, however the *differential operators* they involve act in a more general set of objects that apparently includes non-differentiable functions. The way a differential operator acts over non necessarily differentiable functions is through the *distributional derivative*, which is an indirect method of assigning derivatives when classical differentiation does not work. Sometimes, the solutions cannot even be represented as functions but they exists as distributions. Those are the so called *generalized solutions*, that allow to deal with the existence of solutions in a more general frame. The drawback is that sometimes it is necessary to know if a distributional solution provided by some existence theorem is an actual function, that is, a *classic solution*.

To finish this section, note that in the example above, the *singularities* or point where the solution is not differentiable travel by the lines $x = t$ and $x = 1 - t$. These lines, that depend on the PDE, are called *characteristics* and play an important role in the study of second order equations.

4.4 The separation of variables method

The idea is to find a suitable sequence of solutions of the equation

$$u_{tt} = c^2 u_{xx}$$

of the form $X_n(x)Y_n(t)$ and to expect that any other solution could be expressed as a series

$$u(x,t) = \sum_n a_n X_n(x) T_n(t).$$

The suitable functions will appear adapted to the boundary problem. To see that, we will look for the solution of the vibrating string on $[0, \pi]$ such that $u(x, 0) = f(x)$ and $u_t(x, 0) = 0$. Obviously, $f(0) = f(\pi) = 0$. The wave equation applied to $X(x)T(t)$ gives

$$X(x)T''(t) = c^2 X''(x)T(t)$$

that implies the following ratio should be constant

$$\lambda = \frac{X''(x)}{X(x)} = \frac{1}{c^2}\frac{T''(t)}{T(t)}$$

as it represents the equality between functions of independent variables. The eventual vanishing of one denominator is not a problem as the limit of the quotient exists. The boundary condition leads to look for functions X such that $X(0) = X(\pi) = 0$. The unique no null solutions of

$$X''(x) - \lambda X(x) = 0$$

that can take the value 0 twice appear when $\lambda < 0$. Moreover, $X(0) = 0$ if $X(x)$ is a multiple of $\sin(\sqrt{-\lambda}x)$ and $X(\pi) = 0$ if and only if $\lambda = -n^2$ for some $n \in \mathbb{N}$. In other words, we have a sequence of solutions

$$X_n(x) = \sin nx, n \in \mathbb{N}.$$

The corresponding T-functions satisfy the differential equation

$$T_n''(t) + c^2 n^2 T_n(t) = 0.$$

The solutions are combinations of $\sin cnt$ and $\cos cnt$. The hypothesis zero speed at $t = 0$ implies we chose cosines. Therefore

$$X_n(x)T_n(t) = \sin nx \cos cnt.$$

Now we look for a solution of the wave equation of the form

$$u(x,t) = \sum_n a_n \sin nx \cos cnt.$$

At this moment, we do not care much about the differentiability of the series since it depend on the decay of (a_n). To match the initial condition, extend f to $[-\pi, 0]$ by antisymmetry $f(x) = -f(x)$. Doing this, f admits a development in sine series on $[-\pi, \pi]$ that is uniformly convergent if f is, for instance, Lipschitz. Thus we may write

$$f(x) = \sum_n a_n \sin nx.$$

Therefore, the formal solution of our problem is

$$u(x,t) = \sum_n a_n \sin nx \cos cnt.$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(s) \sin ns \, ds.$$

As we have said before, the differentiability (in classical sense) of this solution will depend on additional hypotheses of regularity about f, but we may accept this solution in a generalized sense. For instance, the example discussed in the previous section could be expressed in this way. We will adapt it to the interval $[0, \pi]$ in order to simplify the trigonometric part.

Example 4.4.1. *Solve the problem $u_{tt} = u_{xx}$ with $u(0,t) = u(\pi,t) = 0$ for all $t \in \mathbb{R}$ and $u(x,0) = \min\{x, \pi - x\}$, $u_t(x,0) = 0$ for $x \in [0,\pi]$.*

The extension of $f(x) = \min\{x, \pi - x\}$ to $[-\pi.\pi]$ by antisymmetry is

$$f(x) = \min\{\pi - x, \max\{x, -\pi - x\}\},$$

so the Fourier series contains only terms in $\sin nx$. It is not difficult to see that coefficients for n even terms vanish. For n odd we have

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx = 4 \int_0^{\pi/2} x \sin nx \, dx$$

$$4 \left. \frac{-x \cos nx}{n} \right|_0^{\pi/2} + 4 \int_0^{\pi/2} \frac{\cos nx}{n} dx = \left. \frac{\sin nx}{n^2} \right|_0^{\pi/2} = \frac{\pm 4}{n^2}$$

where the sign depends on the remainder class of n modulo 4. So, the computation leads to a uniformly convergent series that can be written as

$$f(x) = 4\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2} \sin((2n+1)x).$$

Therefore, the solution of the wave equation for zero speed at $t = 0$ is given by

$$u(x,t) = 4\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2} \sin((2n+1)x)\cos((2n+1)t)$$

that can be considered a satisfactory solution in terms of Fourier series. ■

The separation of variables method not only has a practical value. It shows that any solution of the vibrating string can be expressed as a "sum" (series, actually) of stationary solutions, or *standing waves*. In the case the string is placed in a violin, standing waves are the notes and their suitable combination is called *music*.

4.5 Uniqueness for the boundary problem

We will prove an argument for the wave equation with still boundary that covers the case of the vibrating string. Consider the problem

$$\begin{cases} u_{tt} = c^2 \Delta u \text{ on } \Omega, \\ u_t = 0 \text{ on } \partial\Omega, \end{cases} \tag{4.1}$$

where $\Omega \subset \mathbb{R}^n$ is bounded. We will prove the following *conservation of energy principle*.

Proposition 4.5.1. *For the problem 4.1, the following quantity remains constant in time*

$$E(t) = \iiint_{\Omega} (u_t^2 + c^2\|\nabla u\|^2)\, dV.$$

Recall that the laplacian and volume integrals applies only to spatial variables, not to t.

Proof. We will show that $E'(t) = 0$. Firstly, note that

$$E'(t) = \iiint_{\Omega} (2u_t u_{tt} + 2c^2 \nabla u \cdot \nabla u_t)\, dV.$$

Now apply Gauss theorem to the field $u_t \nabla u$

$$0 = \iint_{\partial\Omega} u_t \nabla u \cdot d\mathbf{S} = \iiint_{\Omega} (\nabla u \cdot \nabla u_t + u_t \Delta u)\, dV.$$

Using this in the previous formula we get

$$E'(t) = \iiint_{\Omega} (2u_t u_{tt} - 2c^2 u_t \Delta u)\, dV = 2 \iiint_{\Omega} u_t (u_{tt} - c^2 \Delta u)\, dV = 0$$

as wanted. ∎

We may think of $E(t)$ like some kind of energy by unit of mass. The term with u_t^2 corresponds to the kinetic energy, whereas the term $\|\nabla u\|^2$ measures the deformation and so corresponds to the "elastic" energy.

Corollary 4.5.2. *Let u, v be to solutions of the wave equation $u_{tt} = c^2 \Delta u$ on a bounded domain Ω such that $u(\mathbf{x}, t) = v(\mathbf{x}, t)$ for $\mathbf{x} \in \partial\Omega$ and all t, $u(\mathbf{x}, 0) = v(\mathbf{x}, 0)$, $u_t(\mathbf{x}, 0) = v_t(\mathbf{x}, 0)$ for $\mathbf{x} \in \Omega$. Then $u = v$ on Ω.*

Assuming some decay of the derivatives of u, we could "push" the boundary to infinity so the energy principle would still hold. That is not necessary because uniqueness in the boundary free problem will be obtained as a consequence of a explicit formula for the solution.

Remark 4.5.3. *Since the energy integral represents the sum of squared Hilbert norms of derivatives of $u(\mathbf{x}, t)$, it can be computed by Parseval's equality when there is an orthogonal decomposition. That is the case for the vibrating string when solved by the separation of variables method, implying that the total energy is the sum of the energies corresponding to the stationary components.*

4.6 The wave equation in three dimensions

Now we will consider the wave equation

$$u_{tt} - c^2 \Delta u = 0,$$

on $\mathbb{R}^3$ without boundary conditions. The idea is to prove that this case can be reduced to the one-dimensional case. Why we are doing this in three dimensions instead of two or a general n will be explained later.

For the proof we will use the so called *spherical averages* or means. Given a continuous function h on $\mathbb{R}^3$, consider

$$\mathcal{A}(h, \mathbf{x}, r) = \frac{1}{4\pi} \iint_{\partial B(0,1)} h(\mathbf{x} + r\mathbf{y})\, dS(\mathbf{y}),$$

where $r \in \mathbb{R}$. Note that $\mathcal{A}(h, \mathbf{x}, r) = \mathcal{A}(h, \mathbf{x}, -r)$ and $\mathcal{A}(h, \mathbf{x}, 0) = h(\mathbf{x})$. Moreover, if h is differentiable at $\mathbf{x}$, then

$$\mathcal{A}(h, \mathbf{x}, r) = h(\mathbf{x}) + o(|r|).$$

That implies that $\frac{\partial}{\partial r}\mathcal{A}(h, \mathbf{x}, 0) = 0$. For the next computations some more regularity is assumed, namely h is C^2. It is quite evident that averages and derivations, either with respect spatial variables or parameters, commute. In particular, we have

$$\Delta_{\mathbf{x}} \mathcal{A}(h, \mathbf{x}, r) = \mathcal{A}(\Delta h, \mathbf{x}, r),$$
$$\frac{\partial}{\partial t}\mathcal{A}(h, \mathbf{x}, r) = \mathcal{A}(\frac{\partial h}{\partial t}, \mathbf{x}, r),$$

in case h depends on some t too for the second formula.

Now, we are going to relate partial derivation with respect to r with the Laplacian. Starting by

$$\frac{\partial}{\partial r}\mathcal{A}(h, \mathbf{x}, r) = \frac{1}{4\pi} \iint_{\partial B(0,1)} \nabla_{\mathbf{x}} h(\mathbf{x} + r\mathbf{y}) \cdot \mathbf{y}\, dS(\mathbf{y})$$

$$= \frac{1}{4\pi} \iiint_{B(0,1)} \nabla_{\mathbf{y}} \cdot \nabla_{\mathbf{x}} h(\mathbf{x} + r\mathbf{y})\, dV(\mathbf{y}) = \frac{r}{4\pi} \iiint_{B(0,1)} \Delta_{\mathbf{x}} h(\mathbf{x} + r\mathbf{y})\, dV(\mathbf{y})$$

$$= \frac{1}{4\pi r^2} \Delta_{\mathbf{x}} \iiint_{B(\mathbf{x},r)} h(\mathbf{y})\, dV(\mathbf{y}).$$

On the other hand

$$\frac{1}{4\pi}\iiint_{B(\mathbf{x},r)} h(\mathbf{y})\,dV(\mathbf{y}) = \frac{1}{4\pi}\iiint_{B(0,r)} h(\mathbf{x}+\mathbf{y})\,dV(\mathbf{y})$$

$$= \frac{1}{4\pi}\int_0^r \iint_{\partial B(0,1)} h(\mathbf{x}+\rho\mathbf{y})\,\rho^2 dS(\mathbf{y})d\rho = \int_0^r \mathcal{A}(h,\mathbf{x},\rho)\,\rho^2\,d\rho.$$

Now, we can combine the previous computations as follows

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\mathcal{A}(h,\mathbf{x},r)\right) = \frac{\partial}{\partial r}\left(\frac{1}{4\pi}\Delta_{\mathbf{x}}\iiint_{B(\mathbf{x},r)} h(\mathbf{y})\,dV(\mathbf{y})\right)$$

$$= \Delta_{\mathbf{x}}\left(\frac{\partial}{\partial r}\left(\int_0^r \mathcal{A}(h,\mathbf{x},\rho)\,\rho^2\,d\rho\right)\right) = r^2\Delta_{\mathbf{x}}\mathcal{A}(h,\mathbf{x},r).$$

Therefore, we get that the following equation is satisfied

$$\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)\mathcal{A}(h,\mathbf{x},r) = \Delta_{\mathbf{x}}\mathcal{A}(h,\mathbf{x},r).$$

This also known as Kirchhoff's formula. So far, the argument would work in any dimension (number 2 in the formula above would be replaced by $n-1$). Now, we are going to take advantage from the 3-dimensional case. Indeed, put $H(\mathbf{x},r) = r\mathcal{A}(h,\mathbf{x},r)$ and then we have

$$\frac{\partial^2}{\partial r^2}(H(\mathbf{x},r)) = \left(r\frac{\partial^2}{\partial r^2}+2\frac{\partial}{\partial r}\right)\mathcal{A}(h,\mathbf{x},r) = \Delta_{\mathbf{x}}(H(\mathbf{x},r)) \tag{4.2}$$

for all $\mathbf{x}\in\mathbb{R}^3$. That simplification will play a role in the proof of the following result.

Theorem 4.6.1. *The unique solution $u\in C^2(\mathbb{R}^3)$ of the initial values problem*

$$u_{tt} - c^2\Delta u = 0,$$

$u(\mathbf{x},0) = f(\mathbf{x}), u_t(\mathbf{x},0) = g(\mathbf{x})$ *is given by the formula*

$$u(\mathbf{x},t) = \frac{\partial}{\partial t}\left(\frac{1}{4\pi c^2 t}\iint_{\|\mathbf{y}-\mathbf{x}\|=ct} f(\mathbf{y})\,dS(\mathbf{y})\right) + \frac{1}{4\pi c^2 t}\iint_{\|\mathbf{y}-\mathbf{x}\|=ct} g(\mathbf{y})\,dS(\mathbf{y}).$$

Proof. The fact that the formula above gives a solution of he problem can be established by direct computation taking advantage that we know the derivative of averages with respect to the radius of the sphere. Uniqueness will come from the way to deduce the formula: any solution of the Cauchy problem $u(\mathbf{x}, t)$ could be expressed so. We add the variable t to the spherical averages as a parameter, so put $U(\mathbf{x}, r, t) = r\mathcal{A}(u, \mathbf{x}, r, t)$. As said before, the average commutes with derivations, thus U satisfies the wave equation. Indeed,

$$\frac{\partial^2}{\partial t^2}(U(\mathbf{x}, r, t)) = r\frac{\partial^2}{\partial t^2}(\mathcal{A}(u, \mathbf{x}, r, t)) = r(\mathcal{A}(\frac{\partial^2 u}{\partial t^2}, \mathbf{x}, r, t))$$

$$= r(\mathcal{A}(c^2\Delta u, \mathbf{x}, r, t)) = c^2\Delta_{\mathbf{x}}(r\mathcal{A}(u, \mathbf{x}, r, t)) = c^2\Delta_{\mathbf{x}}U(\mathbf{x}, r, t)).$$

Using formula 4.2 we deduce, as follows, that $U(\mathbf{x}, r, t))$ as function of (r, t) with $\mathbf{x}$ as a parameter satisfies the 1-dimensional wave equation

$$\frac{\partial^2}{\partial t^2}(U(\mathbf{x}, r, t)) = c^2\Delta_{\mathbf{x}}U(\mathbf{x}, r, t)) = c^2\frac{\partial^2}{\partial r^2}(U(\mathbf{x}, r, t)).$$

D'Alembert formula Theorem 4.1.2 applies to get

$$U(\mathbf{x}, r, t) = \frac{1}{2}(U(\mathbf{x}, r + ct, 0) + U(\mathbf{x}, r - ct), 0) + \frac{1}{2c}\int_{r-ct}^{r+ct} U_t(\mathbf{x}, \rho, 0)d\rho$$

If we write that formula in terms of the averages we obtain

$$r\mathcal{A}(u, \mathbf{x}, r, t)$$

$$= \frac{1}{2}((r + ct)\mathcal{A}(f, \mathbf{x}, r + ct) + (r - ct)\mathcal{A}(f, \mathbf{x}, r - ct)) + \frac{1}{2c}\int_{r-ct}^{r+ct} \rho\mathcal{A}(g, \mathbf{x}, \rho)d\rho$$

$$= \frac{1}{2}((ct + r)\mathcal{A}(f, \mathbf{x}, ct + r) - (ct - r)\mathcal{A}(f, \mathbf{x}, ct - r)) + \frac{1}{2c}\int_{ct-r}^{ct+r} \rho\mathcal{A}(g, \mathbf{x}, \rho)d\rho,$$

using twice the evenness of $\mathcal{A}(u, \mathbf{x}, r, t)$ (with respect to r) in the last equality. Dividing by r and taking limits as $r \to 0$ we get

$$u(\mathbf{x}, t) = \frac{\partial}{\partial t}(t\mathcal{A}(f, \mathbf{x}, ct)) + t\mathcal{A}(g, \mathbf{x}, ct)$$

$$= \frac{\partial}{\partial t}\left(\frac{1}{4\pi c^2 t}\iint_{\|\mathbf{y}-\mathbf{x}\|=ct} f(\mathbf{y})\, dS(\mathbf{y})\right) + \frac{1}{4\pi c^2 t}\iint_{\|\mathbf{y}-\mathbf{x}\|=ct} g(\mathbf{y})\, dS(\mathbf{y}) \quad (4.3)$$

as claimed. ■

Note that, as in the one-dimensional case, we can establish domains of influence and dependence with respect the initial data. A particularly interesting case are *spherical waves*, that is, the solutions have radial symmetry. Put

$$r = \sqrt{x^2 + y^2 + z^2}$$

and $u(r,t) = u(x,y,z,t)$. The wave equation transforms into

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \right) u(r,t) = 0.$$

The same trick as before with the averages led to this equivalent equation

$$\frac{\partial^2(ru)}{\partial t^2} - c^2 \frac{\partial^2(ru)}{\partial r^2} = 0.$$

Therefore, the solutions are of the form

$$u(r,t) = \frac{1}{r}\Phi(r - ct) + \frac{1}{r}\Psi(r + ct)$$

that shows that the amplitude decay is proportional to $1/r$. Somehow, formula 4.3 could be interpreted as a superposition of spheric waves. On the other hand, the wave equation also have solutions of the type *traveling plane wave* given by

$$u(\mathbf{x}, t) = \Phi(\mathbf{x} \cdot \mathbf{n} - ct)$$

where $\mathbf{n}$ is a unitary vector pointing the direction of the wave.

Theorem 4.6.1 and these examples illustrates quantitatively the behaviour of wave transmission postulated by *Christiaan Huygens*. The idea is that in an isotropic medium perturbations propagated in concentric circles at uniform speed, and any point reached by the perturbation can be considered as a the origin of a new generation of waves. That allowed him to describe what happen with waves when they reach obstacles or another medium with a different propagation speed. Reflection, refraction and diffraction of waves can be qualitatively, and even partially quantitatively, explained by Huygens principle. The mathematical development based in solving the wave equation in particular situations was carried out by *Augustin Fresnel*

4.7 The wave equation in two dimensions

We will use the solution of the Cauchy problem for the wave equation in three dimensions to solve it for two dimensions. A main reason to proceed so is that the method of spherical averages does not work in even dimensions. Consider the wave equation

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$$

with initial condition $u(x_1, x_2, t) = f(x_1, x_2)$ and $u_t(x_1, x_2, t) = g(x_1, x_2)$. The idea is that a 3-dimensional solution of the wave equation that does not depend on the variable x_3 is, as well, a 2-dimensional solution of the wave equation. Define functions $\tilde{f}$ and $\tilde{g}$ on $\mathbb{R}^3$ by

$$\tilde{f}(\mathbf{x}) = f(x_1, x_2); \quad \tilde{g}(\mathbf{x}) = g(x_1, x_2);$$

where $\mathbf{x} = (x_1, x_2, x_3)$. From a physical point of view is evident that the solution of the 3-dimensional Cauchy problem does not depend on x_3. Such an intuition, can be corroborated by applying the uniqueness result to a solution $u(x_1, x_2, x_3, t)$ and $u(x_1, x_2, x_3 + \xi, t)$. The solution can be expressed as

$$u(x_1, x_2, t) = \frac{\partial}{\partial t}\left(\frac{1}{4\pi c^2 t}\iint_{\|\mathbf{y}-\mathbf{x}\|=ct} f(\mathbf{y})\, dS(\mathbf{y})\right) + \frac{1}{4\pi c^2 t}\iint_{\|\mathbf{y}-\mathbf{x}\|=ct} g(\mathbf{y})\, dS(\mathbf{y}),$$

however, it would be desirable to have the solution expresses "2-dimensionally". Note that we only care for spheres with centres at the XY-plane, so the sphere where the integrals are taken is

$$\|\mathbf{y} - \mathbf{x}\| = \sqrt{(y_1 - x_1)^{+}(y_2 - x_2) + y_3^2} = ct$$

and

$$dS(\mathbf{y}) = \sqrt{1 + \left(\frac{\partial y_3}{\partial y_1}\right)^2 + \left(\frac{\partial y_3}{\partial y_2}\right)^2}\, dy_1 dy_2.$$

A computation shows that this can be simplified to

$$dS(\mathbf{y}) = \frac{ct}{|y_3|} dy_1 dy_2 = \frac{ct}{\sqrt{c^2t^2 - r^2}}$$

being $r^2 = y_1^2 + y_2^2$. Therefore, the integrals can be transformed to

$$u(x_1, x_2, t) = \frac{1}{2\pi c}\frac{\partial}{\partial t}\left(\iint_{r\leq ct} \frac{f(y_1, y_2)}{\sqrt{c^2t^2 - r^2}} dy_1 dy_2\right) + \frac{1}{2\pi c}\iint_{r\leq ct} \frac{g(y_1, y_2)}{\sqrt{c^2t^2 - r^2}} dy_1 dy_2$$

where the 4 at the denominators has been replaced by 2 because the symmetric contribution of the two caps of the sphere. Therefore, we have the following result.

Theorem 4.7.1. *The unique solution $u \in C^2(\mathbb{R}^2)$ of the initial values problem*

$$u_{tt} - c^2 \Delta u = 0,$$

$u(\mathbf{x}, 0) = f(\mathbf{x}), u_t(\mathbf{x}, 0) = g(\mathbf{x})$ *is given by the formula*

$$u(\mathbf{x}, t) = \frac{1}{2\pi c} \frac{\partial}{\partial t} \left(\iint_{r \le ct} \frac{f(y_1, y_2)}{\sqrt{c^2 t^2 - r^2}} dy_1 dy_2 \right) + \frac{1}{2\pi c} \iint_{r \le ct} \frac{g(y_1, y_2)}{\sqrt{c^2 t^2 - r^2}} dy_1 dy_2$$

where $r = \sqrt{y_1^2 + y_2^2}$.

Proof. As we have seen, it reduces to Theorem 4.6.1. ■

The trick to pass from 3 to 2 dimensions is called *Hadamard's method of descent*, that can be used also for some other problems related to the wave equation. Both formulas, in dimensions 2 and 3, show a phenomenon that does not happen in one dimension: the solutions of the Cauchy problem could be less regular that the initial data. The eventual apparition of some kind of singularity at a later time $t > 0$ is called the *focussing effect*.

The comparison between Theorem 4.6.1 and Theorem 4.7.1 tell us that the domains of influence and dependence in 2 and 3 dimensions are different. Whereas in 3 dimensions we have spheres, in 2 dimensions we get solid balls. That has a physical consequence that all we have experienced: when a firecracker explodes we only hear one short bang, but when a stone is throw to a pond, circular waves are persistent over all the surface encompassed by the wavefront. For that reason, thanks to the fact we live in three (spatial) dimensions, we can enjoy music.

4.8 Vibrations of a rectangular drum

We will use the method of separation of variables to address the wave equation

$$u_{tt} = c^2 u_{xx} + c^2 u_{yy}$$

with boundary conditions on a rectangle $[0, a] \times [0, b]$. We look for solutions of the form

$$X(x)Y(y)T(t),$$

thus
$$X(x)Y(y)T''(t) = c^2X''(x)Y(y)T(t) + c^2X(x)Y''(y)T(t)$$
and so
$$\lambda = \frac{1}{c^2}\frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)}$$
is constant. As in the one-dimensional case, $\lambda < 0$ can take only certain values. Since
$$X(0) = X(a) = Y(0) = Y(b) = 0$$
we get that
$$X_n(x) = \sin(\pi nx/a), \quad Y_m(y) = \sin(\pi my/b),$$
with $n, m \in \mathbb{N}$. Therefore $\lambda = -\pi^2(n^2/a^2 + m^2/b^2)$ and so
$$T_{nm}(t) = \cos(c\pi\sqrt{n^2/a^2 + m^2/b^2}\ t)$$
in case of zero speed at $t = 0$. Now, the coefficients of the series
$$u(x, y, t) = \sum_{n,m} a_{nm} \cos(c\pi\sqrt{n^2/a^2 + m^2/b^2}\ t) \sin(\pi nx/a) \sin(\pi my/b)$$
must be adjust so the initial condition is fulfilled, as convergent double Fourier series.

For this solution, on the contrary that the one-dimensional, the frequencies do not keep rational ratios. That has consequences from the point of view of the type and "musical quality" of the sounds can be obtained from the square drum. Another phenomenon is that one same frequency could have several different vibration patterns. For instance, $(n, m) = (1, 7)$ and $(n, m) = (5, 5)$ for $a = b = 1$.

4.9 Rationale and remarks

The wave equation admits many variations. For instance, the problem of a vibrating string with variable density leaded to the Sturm-Liouville problems that we mentioned in Chapter 2.

Historically, the seeming difference between the solution of the wave equation with arbitrary functions (d'Alembert) and the solution with Fourier series (by L. Euler and D. Bernoulli, before J. Fourier) was a stone in the shoe of Mathematical Analysis for more than a century that lead to the modern notion of *function*. Before that, it was inconceivable that a discontinuous graph had an analytical representation.

We have used the vibrating string to illustrate that a PDE "admits" non-differentiable solutions. The idea is to show the students that maybe we are not using the right frame for our problems, like the ancient algebraists who found polynomial equations requiring the use of negative or imaginary quantities to reach a positive real solution. The students are not ready at this stage to face the technical difficulties of distribution theory, however is posible to provide them with some intuition. For dimension $n \geq 2$, the regularity of the solution could be worse than the one of the initial data, so non differentiability points may appear spontaneously.

The wave equation does not model sea waves, which is somehow a disappointment. The waves on shallow liquid surfaces are described by the *Korteweg-de Vries equation*, a third order non linear PDE, that admits a particular kind of traveling solutions called *solitons*.

4.10 Exercises

1. Study the solutions of the vibrating string with only one butt fixed.

2. An elastic string whose butts are fixed on $x = 0$ and $x = \pi$ is deformed by pulling it from a point $a \in (0, \pi)$. Describe analytically the evolution of the string once it is released.

3. Show that d'Alembert formula 4.1.2 can be applied to the vibrating string on $[0, L]$ with $u(x, 0) = f(x)$, $u_t(x, 0) = g(x)$ for $x \in [0, L]$, just taking the odd extension of f and g to $[-L, 0]$ and then the extension to all $\mathbb{R}$ by $2L$-periodicity.

4. An elastic string of length L with its butts are fixed is taken to a symmetric parabolic shape. Describe analytically the evolution of the string once it is released. How does the frequency depend on the length?

5. Find the solution of the problem on $\mathbb{R} \times [0, +\infty)$

$$\begin{cases} u_{tt} = u_{xx}, \\ u(x, x) = f(x), \\ u_x(x, -x) - u_t(x, -x) = g(x). \end{cases}$$

6. Find the solution of the problem on $[0, \pi] \times [0. + \infty)$

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

7. Find the solution of the problem on $[0, 1] \times [0. + \infty)$

$$\begin{cases} u_{xx} = u_{tt}, \\ u(x, 0) = (1 - x)^3, \\ u_t(x, 0) = (1 - x)^2, \\ u_x(0, t) = u_x(1, t) = 0. \end{cases}$$

   and compute $u(1, 1)$.

8. Let $u(x, t)$ where $x \in \mathbb{R}^n$ a solution of the n-dimensional wave equation, that is,

$$u_{tt} - \Delta u = 0$$

   that verifies the conditions $u(x, 0) = 0$, $u_t(x, 0) = f(x)$ for some function f. Prove that the function $v = u_t$ also satisfies the wave equation and verifies the "switched conditions" $v(x, 0) = f(x)$ and $v_t(x, 0) = 0$.

9. Write the energy of the vibrating string in terms of the Fourier coefficients as expressed in Remark 4.5.3

10. Find the solution of the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \sin(\alpha x - \omega t)$$

   that satisfies $u(x, 0) = 0$, $u_t(x, 0) = 0$.

11. Consider the Schrödinger's equation on $\mathbb{R}^3$

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t) = -\frac{\hbar^2}{2m}\Delta\Psi(\mathbf{x},t) + V(\mathbf{x})\Psi(\mathbf{x},t).$$

Find a condition for the existence of stationary solutions

$$\Psi(\mathbf{x},t) = \psi(\mathbf{x})e^{-iEt/\hbar}$$

with E a constant that represents the energy.

Chapter 5

Laplace equation and related problems

The main goal of this chapter is to study the solvability of the Laplace equation

$$\Delta u = 0$$

with boundary conditions, the so called Dirichlet problem. Explicit solutions can be found for domains with relatively simple geometry, however in the general case will be glad with a mere existence proof. By the way, we will show some properties of the Laplacian that preludes the methods of Functional Analysis to deal with that problem.

5.1 General facts about harmonic functions

In this section we will remember the mean value property for harmonic functions and its important consequences. The arguments are true in any dimension, however we will work in the usual 3-dimensional space for the sake of intuition.

Let start by the straightforward application of the Gauss-Ostrogradsky theorem: if f is harmonic in a domain that contains the ball $B[\mathbf{x}, R]$ with $R > 0$ then

$$\iint_{\partial B[\mathbf{x},R]} \nabla f \cdot d\mathbf{S} = \iiint_{B[\mathbf{x},R]} \Delta f \, dV = 0.$$

We can rewrite this equality

$$\iint_{\partial B[\mathbf{x},R]} \nabla f \cdot \mathbf{N}\, dS = 0$$

and the term $\nabla f \cdot \mathbf{N}$ can be interpreted as a *normal derivative*, usually denoted by $\frac{\partial f}{\partial n}$ (in this particular case is a radial derivative $\frac{\partial f}{\partial r}$). We may parameterize the sphere $\partial B[\mathbf{x}, r]$ by means of the unit sphere $\partial B[0,1]$ as $\mathbf{x} + r\mathbf{y}$. In such a case we have $\mathbf{y} = \mathbf{N}$ as well. Since the sizes of the spheres differ in a r^2 factor we have

$$\iint_{\partial B[0,1]} \nabla f(\mathbf{x} + r\mathbf{y}) \cdot \mathbf{y}\, dS(\mathbf{y}) = \frac{1}{r^2} \iint_{\partial B[\mathbf{x},r]} \nabla f \cdot \mathbf{N}\, dS = 0.$$

Then, integrating with respect to r we get

$$0 = \int_0^R \iint_{\partial B[0,1]} \nabla f(\mathbf{x} + r\mathbf{y}) \cdot \mathbf{y}\, dS(\mathbf{y})\, dr =$$

$$\iint_{\partial B[0,1]} \int_0^R \frac{d}{dr}\Big(f(\mathbf{x} + r\mathbf{y})\Big)\, dr\, dS(\mathbf{y}) = \iint_{\partial B[0,1]} f(\mathbf{x} + r\mathbf{y})\Big|_{r=0}^{R} dS(\mathbf{y})$$

$$= \iint_{\partial B[0,1]} (f(\mathbf{x} + R\,\mathbf{y}) - f(\mathbf{x}))\, dS(\mathbf{y})$$

which implies after rescaling (integration over the ball of radius R) that

$$0 = \iint_{\partial B[\mathbf{x},R]} (f - f(\mathbf{x}))\, dS = \iint_{\partial B[\mathbf{x},R]} f\, dS - 4\pi R^2 f(\mathbf{x})$$

and so

$$f(\mathbf{x}) = \frac{1}{4\pi R^2} \iint_{\partial B[\mathbf{x},R]} f\, dS.$$

This remarkable identity is the so called *mean value property* of the harmonic functions, that is the value at any point can be expressed as an average of the values over any sphere centered at that point, provided that the corresponding ball is contained into the domain.

This result is true in any dimension. For instance, in dimension 1 it is obvious because harmonic functions are affine. In dimension 2 comes from and adequate use of the Green-Riemann formula. However, for dimension greater than 3 it is necessary to prove an adequate version of the Gauss-Ostrogradsky formula, whose great complication is the discussion on the hypotheses for the domain. Therefore, we can state the following result.

Theorem 5.1.1. *The value of a harmonic function at an inner point of its domain is the average of the values on Euclidean spheres centred at the point provided that the ball is contained in the domain.*

Theorem 5.1.2. *Let $D \subset \mathbb{R}^n$ be a connected domain and f a harmonic function defined on D. Then*

(a) f does not have relative strictly extremum values;

(b) if f attains an absolute extreme value on D then f is constant;

(c) if D is bounded and f can be extended continuously to $\overline{D}$ then f attains its extreme values on ∂D.

Proof. We will argue with maximums, being the argument with minimums similar. Recall that we keep assuming dimension 3 to keep a better intuition. (a) Assume that f has a relative strict maximum at $\mathbf{x}$. Then there is $\varepsilon > 0$ such that $f|_{\partial B(\mathbf{x},\varepsilon)} < f(\mathbf{x})$. By the continuity of f (a strict inequality at a particular point integrated remains strict) we get that

$$\frac{1}{4\pi\varepsilon^2} \iint f \, dS < f(\mathbf{x})$$

which is a contradiction.
(b) Now, if the function attains a maximum on Ω, the previous argument shows that actually we have $f|_{\partial B(\mathbf{x},\varepsilon)} = f(\mathbf{x})$ for any $\mathbf{x}$ where the maximum is attained and any $\varepsilon > 0$ such that $B[\mathbf{x}, \varepsilon] \subset D$. That easily implies that the set

$$\{\mathbf{x} \in D : f(\mathbf{x}) = \max(f)\}$$

is open. As it is clearly closed, then it must be all D by connection.
(c) Finally, the last statement is a consequence that the maximum has to be attained somewhere. If attained on D, then the function is constant and so the maximum is also attained on ∂D. ∎

Remark 5.1.3. *It is natural to ask if (b) in Theorem 5.1.2 holds with relative extreme values. The positive answer is evident in the one-dimensional case. For the 2-dimensional case we could use the identity theorem for holomorphic functions. For dimension $n \geq 3$, we could use an identity theorem for real analytic functions of several variables, but it is necessary to know that harmonic functions are real analytic that will be proved later (Theorem 5.4.2).*

A consequence of the previous result is the uniqueness of the solution of the Dirichlet's problem on bounded domains taking the connected components. In case, of unbounded domains we have the following.

Corollary 5.1.4. *A harmonic function defined on $\mathbb{R}^n$ which vanishes at ∞ must be null.*

This result can be applied to prove that a function vanishing at ∞ such that its derivatives also vanishes at ∞ with a suitable rate of decay is actually a potential with charge given by Poisson's equation.

5.2 Solutions with spherical symmetry

When looking for solutions of the Laplace equation (or related) with some symmetries it is convenient to express the laplacian in the suitable form in order to reduce the number of variables. For instance, in polar coordinates the laplacian becomes

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2},$$

that essentially already appeared in the previous chapter, and in cylindrical coordinates

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2}.$$

We will not write the formula for spherical coordinates which is more complicated and it will not be used.

In order to find the solutions of the Laplace equation with spherical symmetry we can proceed straightforward as follows. The function must be depend only on the distance to the origin, that is, $u(\mathbf{x}) = \phi(\|\mathbf{x}\|)$ being $\|\cdot\|$ the Euclidean norm. Taking

$$r = \|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

the partial derivatives are

$$\frac{\partial r}{\partial x_k} = \frac{x_k}{r}.$$

Applying that to u we get

$$\frac{\partial u}{\partial x_k} = \phi'(r)\frac{x_k}{r},$$

$$\frac{\partial^2 u}{\partial x_k^2} = \phi''(r)\frac{x_k^2}{r^2} + \phi'(r)\frac{r - x_k^2/r}{r^2}$$
$$= \phi''(r)\frac{x_k^2}{r^2} + \phi'(r)\frac{r^2 - x_k^2}{r^3}$$

Summing up for $k = 1, \ldots, n$ we get

$$\Delta u = \phi''(r)\frac{r^2}{r^2} + \phi'(r)\frac{nr^2 - r^2}{r^3} = \phi''(r) + \frac{n-1}{r}\phi'(r)$$

The Laplace equation $\Delta u = 0$ led to ϕ has to satisfy the differential equation

$$\phi''(t) + \frac{n-1}{t}\phi'(t) = 0$$

which leads to solutions of the form

$$\alpha|x| + \beta \text{ if } n = 1;$$

$$\alpha \log \|\mathbf{x}\| + \beta \text{ if } n = 2;$$
$$\frac{\alpha}{\|\mathbf{x}\|^{n-2}} + \beta \text{ if } n \geq 3.$$

These are the unique solutions with spherical symmetry defined everywhere but the origin. These functiosn will play an important role in the theory that follows.

5.3 Green's formulas and functions

Assume that f, g are fairly regular on $\overline{\Omega}$ where $\Omega \subset \mathbb{R}^n$ is a bounded domain It is not difficult to prove that

$$\nabla \cdot (v\nabla u) = \nabla v \cdot \nabla u + v\Delta u.$$

The integration of the previous equation over Ω using the Gauss-Ostrogradsky theorem gives us

$$\iiint_\Omega v\Delta u\, dV + \iiint_\Omega \nabla v \cdot \nabla u\, dV = \iint_{\partial\Omega} v\nabla f \cdot d\mathbf{S}.$$

(we are using the 3-dimensional notation in order to stress the difference between the domain and its boundary). We may exchange the roles of u and v leading to

$$\iiint_{\Omega} u\Delta v\, dV + \iiint_{\Omega} \nabla v \cdot \nabla u\, dV = \iint_{\partial\Omega} u\nabla v \cdot d\mathbf{S}.$$

The subtraction of one another gives

$$\iiint_{\Omega} (u\Delta v - v\Delta u)\, dV = \iint_{\partial\Omega} (u\nabla v - v\nabla u) \cdot d\mathbf{S}$$

This last formula will be interesting for other purposes later. For the time being we will use with harmonic functions in order to get a formula in terms of the boundary values. Let $\Omega \subset \mathbb{R}^3$ (the argument is general for $\mathbb{R}^n$ but some parameters have to be changed) be a domain, $\mathbf{x} \in \Omega$ and $B \subset \Omega$ a closed ball of center $\mathbf{x}$ and radius $r > 0$. We already know that the function $v(\mathbf{y}) = \|\mathbf{y} - \mathbf{x}\|^{-1}$ is harmonic in $\mathbb{R}^3 \setminus \{\mathbf{x}\}$. Let u be a harmonic function on Ω. Since both u and v are harmonic on $\Omega \setminus B$, we have

$$\begin{aligned} 0 &= \iint_{\partial(\Omega\setminus B)} (u\nabla v - v\nabla u) \cdot d\mathbf{S} \\ &= \iint_{\partial\Omega} (u\nabla v - v\nabla u) \cdot d\mathbf{S} - \iint_{\partial B} (u\frac{-\mathbf{N}}{r^2} - \frac{1}{r}\nabla u) \cdot d\mathbf{S}. \end{aligned}$$

Using the directional derivative with respect the (outer) normal $\mathbf{N}$, the formula can be written in terms of scalar surface integrals

$$\iint_{\partial\Omega} (v\frac{\partial u}{\partial \mathbf{N}} - u\frac{\partial v}{\partial \mathbf{N}})\, dS = \iint_{\partial B} (\frac{u}{r^2} + \frac{1}{r}\frac{\partial u}{\partial \mathbf{N}})\, dS$$

Taking limits as $r \to 0$ on the right-hand term produces $4\pi u(\mathbf{x})$. Indeed, the first member behaves like an averaging of u, meanwhile the second one vanishes as it is $O(r^{-1})O(r^2)$. Thus we have

$$u(\mathbf{x}) = \frac{1}{4\pi} \iint_{\partial\Omega} (v_{\mathbf{x}}\frac{\partial u}{\partial \mathbf{N}} - u\frac{\partial v_{\mathbf{x}}}{\partial \mathbf{N}})\, dS$$

where we have put $v = v_{\mathbf{x}}$ to stress the dependence upon that point. This formula is valid in higher dimensions just changing 4π by the "area" of the unit ball.

The formula would be more satisfactory if it just depended on the values of u on the boundary, not is derivative. To get rid of the normal derivative $\frac{\partial u}{\partial \mathbf{N}}$ assume that we are able to find a function $w_{\mathbf{x}}$ harmonic on Ω which coincides with $v_{\mathbf{x}}$ on $\partial\Omega$. For that function we will have

$$0 = \frac{1}{4\pi} \iint_{\partial\Omega} (w_{\mathbf{x}} \frac{\partial u}{\partial \mathbf{N}} - u \frac{\partial w_{\mathbf{x}}}{\partial \mathbf{N}})\, dS.$$

Taking the difference with the other formula we get

$$u(\mathbf{x}) = \frac{-1}{4\pi} \iint_{\partial\Omega} u\, \frac{\partial(v_{\mathbf{x}} - w_{\mathbf{x}})}{\partial \mathbf{N}}\, dS$$

The function

$$G(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi}(w_{\mathbf{x}}(\mathbf{y}) - v_{\mathbf{x}}(\mathbf{y}))$$

is called the Green function associated to the domain Ω, and it allows this expression for the harmonic function u as

$$u(\mathbf{x}) = \iint_{\partial\Omega} u(\mathbf{y})\, \frac{\partial G}{\partial \mathbf{N}_{\mathbf{y}}}(\mathbf{x}, \mathbf{y})\, dS(\mathbf{y}). \tag{5.1}$$

Sometimes the normal derivative of G is also referred as the Green function. Let us recall that the construction of the Green function depends on finding a suitable function $w_{\mathbf{x}}$.

In case of $\Omega \subset \mathbb{R}^2$, the "area of the unit sphere" is 2π and we have to take

$$v_{\mathbf{x}}(\mathbf{y}) = -\log(\|\mathbf{y} - \mathbf{x}\|)$$

(the minus in the logarithm is to make the function decreasing as $\mathbf{y}$ goes far from $\mathbf{x}$). Despite the fact that the arguments needed a bounded domain Ω, the method works in unbounded domains as a limit case.

Example 5.3.1. *Dirichlet problem in the upper half-plane by Green's function.*

Take $\Omega = \{(x, y) : y > 0\}$. Let $\mathbf{x} = (x, y)$ the point to evaluate u and $\mathbf{z} = (r, s)$ the auxiliary variables. Let $\mathbf{x}^*$ be the symmetric point to $\mathbf{x}$ with respect to the X-axis It is evident that any point $\mathbf{z} \in \partial\Omega$ is equidistant from $\mathbf{x}$ and $\mathbf{x}^*$. Therefore, the function

$$w_{\mathbf{x}}(\mathbf{z}) = -\log(\|\mathbf{z} - \mathbf{x}^*\|)$$

satisfies the conditions for the construction of the Green function. Thus, the Green function is

$$G(\mathbf{x},\mathbf{z}) = \frac{1}{2\pi}(\log(\|\mathbf{z}-\mathbf{x}\|) - \log(\|\mathbf{z}-\mathbf{x}^*\|))$$

$$= \frac{1}{4\pi}\log((r-x)^2+(s-y)^2) - \frac{1}{4\pi}\log((r-x)^2+(s+y)^2).$$

The normal derivative is just the partial derivative with respect to s, but it is necessary to change the sign because the normal points downside. We get

$$\frac{\partial G}{\partial \mathbf{N_z}}(\mathbf{x},\mathbf{z}) = \frac{1}{2\pi}\left(\frac{s+y}{(r-x)^2+(s+y)^2} - \frac{s-y}{(r-x)^2+(s-y)^2}\right)$$

In particular, if we assume $\mathbf{z} \in \partial\Omega$ then

$$\frac{\partial G}{\partial \mathbf{N_z}}(\mathbf{x},\mathbf{z}) = \frac{1}{\pi}\frac{y}{(r-x)^2+y^2}$$

Therefore, a harmonic function u on the upper half-space can be written as

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{y\,u(r,0)}{(r-x)^2+y^2}\,dr.$$

Let $f(x)$ a continuous function on $\mathbb{R}$. If the Dirichlet problem $\Delta u = 0$, $u(x,0) = f(x)$ has solution in the upper half-space, then it can be expressed as

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{y\,f(r)}{(r-x)^2+y^2}\,dr.$$

The solvability of the Dirichlet problem for a given function f depends on the convergence and twice differentiability of the parametric integral. Note that the boundedness of f is enough for that. The uniqueness of the solution can be guarantied asking u to vanish at ∞.

5.4 Solution of the Dirichlet problem for balls

As we have seen in Example 5.3.1 the symmetries of the domain can be of great help to build the Green function. Here we will take advantage of an old notion in Euclidean Geometry: the symmetries with respect to a circle. We will consider the unit ball $\mathbf{B} = B[0,1]$ in $\mathbb{R}^n$. The *symmetric* (with respect the unit ball) of a point $\mathbf{x} \neq 0$ is the point

$$\mathbf{x}^* = \frac{\mathbf{x}}{\|\mathbf{x}\|^2}$$

The remarkable property of this symmetry is that

$$\|\mathbf{y} - \mathbf{x}^*\| \|\mathbf{x}\| = \|\mathbf{y} - \mathbf{x}\|$$

for all $\mathbf{y}$ with $\|\mathbf{y}\| = 1$. Indeed,

$$\begin{aligned} &\|\mathbf{x}\|^2 \|\mathbf{y} - \mathbf{x}^*\|^2 \\ &= \|\mathbf{x}\|^2 (\|\mathbf{y}\|^2 + \|\mathbf{x}^*\|^2 - 2\mathbf{y} \cdot \mathbf{x}^*) = \|\mathbf{x}\|^2 (1 + \|\mathbf{x}\|^{-2} - 2\|\mathbf{x}\|^{-2} \mathbf{y} \cdot \mathbf{x}) \\ &= \|\mathbf{x}\|^2 + 1 - 2\mathbf{y} \cdot \mathbf{x} = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{y} \cdot \mathbf{x} \\ &= \|\mathbf{y} - \mathbf{x}\|^2. \end{aligned}$$

We will use this property to build a Green function for the Dirichlet problem on the unit ball of $\mathbb{R}^3$ so we can use 5.1. Evidently, the following will do the work

$$G(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \left(\frac{\|\mathbf{x}\|^{-1}}{\|\mathbf{y} - \mathbf{x}^*\|} - \frac{1}{\|\mathbf{y} - \mathbf{x}\|} \right).$$

In order to compute the derivative it is necessary to have in mind that

$$d(\|\mathbf{y}\|)(\mathbf{z}) = \frac{\mathbf{y} \cdot \mathbf{z}}{\|\mathbf{y}\|}$$

for $\mathbf{y} \neq 0$. Thus we have

$$\frac{\partial G}{\partial \mathbf{N_y}}(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \left(\frac{-\|\mathbf{x}\|^{-1}}{\|\mathbf{y} - \mathbf{x}^*\|^3} (\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{N_y} - \frac{-1}{\|\mathbf{y} - \mathbf{x}\|^3} (\mathbf{y} - \mathbf{x}) \cdot \mathbf{N_y} \right)$$

We are interested only in the values of the normal derivative at the points of the sphere $\|\mathbf{y}\| = 1$, $\mathbf{N_y} = \mathbf{y}$ and the equalities $\|\mathbf{x}\| \|\mathbf{y} - \mathbf{x}^*\| = \|\mathbf{y} - \mathbf{x}\|$ and $\|\mathbf{x}\|^2 \mathbf{x}^* = \mathbf{x}$ hold. Having all that in mind, we perform the computations

$$\begin{aligned} \frac{\partial G}{\partial \mathbf{N_y}}(\mathbf{x}, \mathbf{y}) &= \frac{1}{4\pi} \left(\frac{-\|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3} (\mathbf{y} - \mathbf{x}^*) \cdot \mathbf{y} - \frac{-1}{\|\mathbf{y} - \mathbf{x}\|^3} (\mathbf{y} - \mathbf{x}) \cdot \mathbf{y} \right) \\ &= \frac{1}{4\pi} \left(\frac{-\|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3} (1 - \mathbf{x}^* \cdot \mathbf{y}) - \frac{-1}{\|\mathbf{y} - \mathbf{x}\|^3} (1 - \mathbf{x} \cdot \mathbf{y}) \right) \end{aligned}$$

$$= \frac{1}{4\pi} \cdot \frac{1 - \|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3}$$

which is called the *Poisson kernel* (same name that in Abel summation of Fourier series). The corresponding integral formula for the solution of Dirichlet problem on $B[0,1]$ can be stated as a theorem (Poisson's formula).

Theorem 5.4.1. *Let $f \in C(\partial\mathbf{B})$, then there exists a unique $u \in C^2(\mathbf{B})$ such that $\Delta u = 0$ and $u|_{\partial\mathbf{B}} = f$ and it is given by*

$$u(\mathbf{x}) = \frac{1}{4\pi} \iint_{\partial\mathbf{B}} \frac{1 - \|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3} f(\mathbf{y})\, dS(\mathbf{y}).$$

Hint of Proof. Firstly, check that the integral kernel is harmonic as a function of $\mathbf{x}$. The commutation of the Laplacian and the integral shows that u is harmonic too. Now, prove that

$$\frac{1}{4\pi} \frac{1 - \|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3} f(\mathbf{y})$$

is good kernel family as $\mathbf{x} \to \mathbf{y} \in \partial\mathbf{B}$ (use 5.1 with constant function 1 to show normalization of integrals). ■

Obviously, this formula can be scaled and translated. Moreover, Poisson's formula has important theoretical consequences.

Theorem 5.4.2. *A harmonic function is real analytic, that is, it can be developed into a power series of nontrivial radius at any point of its domain. In particular, harmonic functions are C^∞.*

Hint of proof. In other words, a harmonic function inherits the regularity properties of the Poisson's kernel. Try to follow the qualitative part of the proof for the development into a complex power series of an holomorphic function based in *Cauchy's formula.* ■

We shall use an alternative technique in the 2-dimensional case that links the Dirichlet problem and Fourier series. We shall use the standard notation in Complex Analysis. Any harmonic function u defined in the unit disc $D(0,1)$ is the real part of an holomorphic function defined also on $D(0,1)$ and expressible by a power series

$$\sum_{n=0}^{\infty} c_n z^n.$$

Therefore, for u we have

$$u(z) = \frac{1}{2}\left(\sum_{n=0}^{\infty} c_n z^n + \sum_{n=0}^{\infty} \overline{c_n z}^n\right).$$

The same function expressed in polar form becomes

$$u(r,\theta) = \frac{1}{2}\left(\sum_{n=0}^{\infty} c_n r^n e^{in\theta} + \sum_{n=0}^{\infty} \overline{c_n} r^n e^{-in\theta}\right).$$

That equality means that we should just look for our solution in the form

$$u(r,\theta) = \sum_{n\in\mathbb{Z}} a_n r^{|n|} e^{in\theta}. \tag{5.2}$$

If we wish that $u(1,\theta) = f(\theta)$ the substitution leads to an equality with a trigonometric series

$$f(\theta) = \sum_{n\in\mathbb{Z}} a_n e^{in\theta}.$$

If $f \in C(\mathbb{T})$ and (a_n) are its Fourier coefficients, then equality 5.2 becomes Poisson's summation formula and using Theorem 3.6.6 we get that

$$u(r,\theta) = (P_r * f)(\theta),$$

where the Poisson kernel was defined as

$$P_r(\theta) = \frac{1}{2\pi} \cdot \frac{1-r^2}{1-2r\cos\theta + r^2}.$$

Note that if $f \in L^1(\mathbb{T})$ merely, then $u(r,\theta)$ may not be continuous on $\overline{D(0.1)}$ but we still have

$$\lim_{r\to 1^-} (P_r * f)(\theta) = v(\theta)$$

at any θ where f is continuous and $\|\cdot\|_1$-convergence in general.

5.5 Dirichlet problem on a rectangle

We will use the separation of variables, but we have to put ourselves in a very convenient hypotheses. We will look for the a function $u(x,y)$ that satisfies the equation $u_{xx} + u_{yy} = 0$ in the interior of $[0, \pi] \times [0, a]$ (any rectangle can be scaled to that one). However, we will assume firstly that

$$u(0, y) = u(\pi, y) = u(x, 0) = 0$$

and $u(x, a) = f(x)$, for $x \in [0, \pi]$, $y \in [0, a]$. Of course, one has to have $f(0) = f(\pi) = 0$. We write $u(x, y) = X(x)Y(y)$ for the separation of variables. We easily arrive to

$$-\frac{X''(x)}{X(x)} = \frac{Y''(y)}{Y(y)} = \lambda.$$

In order to fulfil the boundary conditions, we deduce that $\lambda > 0$ must be of the form $\lambda = n^2$ with $n \in \mathbb{N}$, thus $X(x) = \sin nx$. On the other hand, Y must vanish for $y = 0$, so we can take $Y(y) = \sinh ny$. In order to adjust the upper side boundary condition we will consider a series of the form

$$u(x, y) = \sum_{n=1}^{\infty} a_n \sin(nx) \sinh(ny).$$

We need, therefore, that

$$f(x) = \sum_{n=1}^{\infty} a_n \sinh(na) \sin(nx).$$

Complete f to be an odd function on $[-\pi, \pi]$ and let b_n be the coefficients of its expansion in sinus series. It is enough to take

$$a_n = \frac{b_n}{\sinh na},$$

that implies that

$$u(x, y) = \frac{2}{\pi} \sum_{n=1}^{\infty} \left(\int_0^{\pi} f(t) \sin nt \, dt \right) \sin(nx) \frac{\sinh(ny)}{\sinh na}.$$

The permutation between series and integral would lead to a sort of Green formula. If the Fourier series of f converges uniformly, so does the series for u because $\sinh ny \leq \sinh na$ for $y \in [0, a]$. Actually, me have a little more convergence. Indeed, we have

$$\frac{\sinh(ny)}{\sinh na} \sim e^{-(a-y)n}$$

which implies that the fraction on the righthand side acts like a summation method: Abel series composed with $e^{-(a-y)}$ that goes to 1^- as $y \to a^-$. In particular, if f is merely continuous, the function $u(x, y)$ defined above by a series extends continuously to $[0, \pi] \times \{a\}$. To ensure differentiability in a general case we may need additional regularity assumptions of f. Anyway, the Dirichlet problem for our special boundary conditions is solved at the theoretical level.

Assume that we have to find a solution on a rectangle that vanishes on all the sides of the rectangle but one. By a simple rotation and scaling, it can be transformed in the previous problem. Assume now we have to solve the Dirichlet problem on a rectangle $[0, a] \times [0, b]$ and we know that

$$f(0, 0) = f(a, 0) = f(a, b) = f(0, b) = 0.$$

In these conditions we can consider four Dirichlet problems with boundary conditions which consist in preserving f for one of the sides and making 0 on the others. Any of those problems reduces to the first on of this section, and their sum gives us the desired solution. Finally, we will remove the condition of vanishing the function on the vertices of the rectangle. For that, assume given a function f on the boundary of the rectangle. Consider the function

$$g(x, y) = f(0, 0) + \frac{x}{a}(f(a, 0) - f(0, 0)) + \frac{y}{b}(f(0, b) - f(0, 0))$$

$$+\frac{xy}{ab}(f(a, b) - f(a, 0) - f(0, b) + f(0, 0)).$$

Note that g is harmonic and takes the same values that f on the vertices of the rectangle. Therefore, the Dirichlet problem with boundary condition $f - g$ can be solved with the method we just described. Once found the solution, add g back.

Obviously, the method can be apply in more dimensions, so we have the following.

Theorem 5.5.1. *The Dirichlet problem can be solved for $\Omega \subset \mathbb{R}^n$ a product of bounded intervals, by reducing it to $2n$ separation of variables problems plus a polynomial.*

5.6 Eigenvalues of the Laplacian

The so called *Helmholtz equation*

$$\Delta u = -ku,$$

together with boundary conditions, can be understood as the search for eigenvalues of the Laplacian. We will provide an elementary mathematical frame for this problem. Let $\Omega \subset \mathbb{R}^n$ be a bounded open domain with piece-wise C^1 boundary. On the space $C^\infty(\Omega)$ we may consider the scalar product defined by

$$\langle u, v\rangle = \iiint_\Omega uv\, dV.$$

Recall that $C_0(\Omega)$ is composed of the continuous functions on $\overline{\Omega}$ vanishing on $\partial\Omega$. Consider now the subspace

$$W_0^2(\Omega) = C_0(\Omega) \cap C^\infty(\Omega).$$

Given $u, v \in W_0^2(\Omega)$, the Green formula implies that

$$\iiint_\Omega (u\Delta v - v\Delta u)\, dV = \iint_{\partial\Omega} (u\nabla v - v\nabla u) \cdot d\mathbf{S} = 0.$$

Therefore

$$\langle \Delta u, v\rangle = \langle u, \Delta v\rangle$$

which means that the Laplace operator is symmetric (also called self-adjoint) when acting on $W_0^2(\Omega)$. Previously we have obtained this formula

$$\iiint_\Omega u\Delta u\, dV + \iiint_\Omega \|\nabla u\|^2 dV = \iint_{\partial\Omega} u\nabla u \cdot d\mathbf{S}$$

that applied to $u \in W^2(\Omega)$ with $u \neq 0$ gives

$$\langle u, \Delta u\rangle = -\iiint_\Omega \|\nabla u\|^2 dV < 0$$

meaning that the Laplacian is a *negative operator*. In particular, if u is an eigenfunction with eigenvalue λ, then

$$\lambda\langle u, u\rangle = \langle u, \Delta u\rangle < 0,$$

thus $\lambda < 0$. If v is another eigenfunction associated to an eigenvalue $\mu \neq \lambda$ we have

$$\lambda\langle u, v\rangle = \langle \Delta u, v\rangle = \langle u, \Delta v\rangle = \mu\langle u, v\rangle$$

implying that

$$\langle u, v\rangle = 0.$$

Our discussion can be summarized in the following theorem.

Theorem 5.6.1. *The Laplacian as an operator, when restricted to the subspace $W_0^2(\Omega)$, is symmetric, all its eigenvalues are negative and eigenfunctions corresponding to different eigenvalues are orthogonal.*

The results can be extended to complex valued functions in the following way. First of all, the Gauss divergence theorem still applies with some changes

$$\iiint_\Omega \Delta u\, \overline{v}\, dV + \iiint_\omega \nabla u \cdot \nabla \overline{v}\, dV = \iint_{\partial\Omega} \overline{v}\, \nabla u \cdot d\mathbf{S}.$$

Switching u and $\overline{v}$ we get

$$\iiint_\Omega u\, \overline{\Delta v}\, dV + \iiint_\omega \nabla u \cdot \nabla \overline{v}\, dV = \iint_{\partial\Omega} u\, \nabla \overline{v} \cdot d\mathbf{S}.$$

The difference of the two formulas gives us

$$\iiint_\Omega \Delta u\, \overline{v}\, dV - \iiint_\Omega u\, \overline{\Delta v}\, dV = \iint_{\partial\Omega} (\overline{v}\, \nabla u - u\, \nabla \overline{v}) \cdot d\mathbf{S}.$$

If the hermitian product is defined now for $u, v \in W^{k,2}(\Omega)$ (complex values) as

$$\langle u, v\rangle = \iiint_\Omega u\, \overline{v}\, dV$$

we have

$$\langle \Delta u, v\rangle = \langle u, \Delta v\rangle$$

as before. However, now we can deduce that any eigenvalue is real. Indeed, if u is an eigenfunction with eigenvalue $\lambda \in \mathbb{C}$, we have

$$\lambda\langle u, u\rangle = \langle \Delta u, u\rangle = \langle u, \Delta u\rangle = \overline{\lambda}\langle u, u\rangle.$$

Therefore, $\lambda = \overline{\lambda}$ and so $\lambda \in \mathbb{R}$.

Example 5.6.2. *Compute the eigenvalues of the 2-dimensional Laplacian on the rectangle* $(0,\pi)^2$.

We will use the separation of variables method for the equation $\Delta u = \lambda u$. Put $u(x,y) = X(x)Y(y)$. Then

$$X''(x)Y(y) + X(x)Y''(y) = \lambda X(x)Y(y),$$

and so

$$\frac{X''(x)}{X(x)} = \lambda - \frac{Y''(y)}{Y(y)}.$$

Since the right-hand side term depends only on x and the left one on y, both must be constant. The zero boundary condition implies

$$X(x) = \sin nx; \qquad Y(y) = \sin my;$$

for some $n, m \in \mathbb{N}$. Therefore $\lambda = -n^2 - m^2$ is the feasible eigenvalue for the eigenfunction is $\sin nx \sin my$. Now note that the series of the form

$$\sum_{n=1}^{\infty} a_{n,m} \sin nx \, \sin my$$

represent any element of $W_0^2((0,\pi)^2)$ with uniqueness of the coefficients. Indeed, use the suitable extension of f to $(-\pi,\pi)^2$ and note that, actually, the series and its formal derivatives uniformly converge for the C^∞ assumption. That implies that eigenfunctions can be obtained only as linear combinations of terms $\sin nx \sin my$ corresponding to the same eigenvalue $-n^2 - m^2$. We deduce that the unique eigenvalues of the Laplacian on $(0,\pi)^2$ are the negative integers of the form $-n^2 - m^2$ with $n, m \in \mathbb{N}$. ■

5.7 Vibrations of a regular drum

The vibration of an "elastic n-dimensional material" on a domain $\Omega \subset \mathbb{R}^n$ fastened by its boundary reduces to the Laplacian eigenvalue problem. Indeed, the equation

$$\Delta u = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}$$

reduces by separation of variables $u(\mathbf{x}, t) = W(\mathbf{x})T(t)$ to

$$\frac{1}{c^2}\frac{T''(t)}{T(t)} = \frac{\Delta W(\mathbf{x})}{W(\mathbf{x})} = -\lambda$$

being $-\lambda$ an eigenvalue of Δ on Ω since $W(\mathbf{x}) = 0$ for $\mathbf{x} \in \partial\Omega$ (we adopt the convention that $\lambda > 0$).

We all understand that a regular drum membrane is 2-dimensional and circular. Without loss of generality we may assume the disc supporting the vibrating membrane is

$$D = \{(x, y) : x^2 + y^2 \leq 1\}.$$

The method applies to this case

$$\frac{1}{c^2}\frac{T''(t)}{T(t)} = \frac{\Delta W(x, y)}{W(x, y)} = -\lambda$$

with $\lambda > 0$ should be an eigenvalue of the Laplacian. Therefore, our next step is to determine the set of eigenvalues of the Laplacian on the unit circle. For that, we will use polar coordinates assuming that our function is $W(r, \theta)$ now

$$0 = (\Delta + \lambda)W = \frac{\partial^2 W}{\partial r^2} + \frac{1}{r}\frac{\partial W}{\partial r} + \frac{1}{r^2}\frac{\partial^2 W}{\partial \theta^2} + \lambda W.$$

We can apply again the method of separation of variables by putting

$$W(r, \theta) = R(r)\Theta(\theta)$$

we have

$$\frac{1}{R}\left(r^2 R'' + rR' + r^2\lambda R\right) = -\frac{\Theta''}{\Theta} = m^2$$

where $m \in \mathbb{Z}$ since Θ should be 2π-periodic. For the radial part we have

$$r^2 R'' + rR' + (r^2\lambda - m^2)R = 0.$$

As we have seen in section 3.9, the change of variables $\tau = \sqrt{\lambda}\, r$ reduces the equation to the form

$$\tau^2 R'' + \tau R' + (\tau^2 - m^2)R = 0.$$

whose solutions are given by the Bessel functions J_m and $\sqrt{\lambda}$ has to be one of its zeroes. Going back to the original equation, that means the frequencies of the standing waves of the drum are of the form $c\beta_{m,n}$ where $\beta_{m,n}$ is the n'th zero of J_m. For the explicit solution of our problem it is convenient to introduce some notation. Up to a rotation, the eigenfunctions can be expresed as

$$w_{m,n}(r, \theta) = J_m(\beta_{|m|,n} r)e^{im\theta}.$$

In case we are interested in real valued functions, we may use the angular part $\sin m\theta$ for $m > 0$ and by $\cos m\theta$ for $m < 0$ (the difference is just a $\pi/2$ rotation, which is irrelevant from a qualitative point of view, the weird numeration is just a trick to simplify the formula). Any solution of the vibrating membrane for zero speed at $t = 0$ can be written as

$$\sum_{n\in\mathbb{N}}\sum_{m\in\mathbb{Z}} a_{m,n} w_{m,n}(r,\theta)\sin(c\beta_{|m|,n}t).$$

5.8 Solvability of the Dirichlet problem

So far far we have shown that the Dirichlet problem can be solved for some types of domains: balls, rectangles and halfspaces. We are far from a general result, however we can explain the failure of the variational methods and give some tips on Perron's approach.

Hadamard's example. Firstly, we will show the failure of the Dirichlet principle we spoke about at the introductory chapter. Let u be a harmonic function on $D(0,1)$ (with complex values) that extends continuously to $\partial D(0,1)$. In a previous section we have proved the following polar representation

$$u(r,\theta) = \sum_{n\in\mathbb{Z}} a_n r^{|n|} e^{in\theta}.$$

We will prove the following formula for the Dirichlet's energy integral

$$\iint_{D(0,1)} \|\nabla u\|^2 dV = 2\pi \sum_{n\in\mathbb{Z}} |n||a_n|^2.$$

For that, firstly note that the coordinates of the gradient is

$$\nabla u = \left(\frac{x}{r}\frac{\partial u}{\partial r} - \frac{y}{r^2}\frac{\partial u}{\partial \theta}, \frac{y}{r}\frac{\partial u}{\partial r} + \frac{x}{r^2}\frac{\partial u}{\partial \theta}\right).$$

Form that we get

$$\|\nabla u\|^2 = \left|\frac{\partial u}{\partial r}\right|^2 + \frac{1}{r^2}\left|\frac{\partial u}{\partial \theta}\right|^2.$$

In our case we have

$$\frac{\partial u}{\partial r}(r,\theta) = \sum_{n\neq 0} a_n |n| r^{|n|-1} e^{in\theta}$$

$$\frac{\partial u}{\partial \theta}(r,\theta) = \sum_n a_n i n r^{|n|} e^{in\theta}$$

Orthogonality gives

$$\int_0^{2\pi} \left| \frac{\partial u}{\partial r}(r,\theta) \right|^2 d\theta = 2\pi \sum_{n\neq 0} |a_n|^2 |n|^2 r^{2|n|-2}$$

$$\frac{1}{r^2} \int_0^{2\pi} \left| \frac{\partial u}{\partial \theta}(r,\theta) \right|^2 d\theta = 2\pi \sum_{n\neq 0} |a_n|^2 |n|^2 r^{2|n|-2}$$

Finally we deduce the caimed formula,

$$\iint_{D(0,1)} \|\nabla u\|^2 dV = 4\pi \int_0^1 \sum_{n\neq 0} |a_n|^2 |n|^2 r^{2|n|-1} dr$$

$$= 4\pi \sum_{n\neq 0} |a_n|^2 |n|^2 \frac{1}{2|n|} = 2\pi \sum_{n\in\mathbb{Z}} |n||a_n|^2,$$

where the Jacobian r was taken into account.

The following example of failure of Dirichlet's principle was found by Hadamard.

Example 5.8.1. *The function defined in polar coordinates by*

$$h(r,\theta) = \sum_{n\neq 0} n^{-2} r^{|n|} e^{in^4\theta}$$

is continuous on $\overline{D(0,1)}$*, harmonic on* $D(0,1)$ *and satisfies*

$$\iint_{D(0,1)} \|\nabla h\|^2 dV = +\infty.$$

Therefore, the solution of Dirichlet problem $\Delta u = 0$ *with*

$$u(1,\theta) = \sum_{n\neq 0} n^{-2} e^{in^4\theta}$$

cannot be found by Dirichlet's principle.

Proof. Just apply the preceding computation. ■

Perron's method. The fact that for some boundary values Hilbert space methods do not work is not an obstacle to alternative different approaches. It can be showed that the Dirichlet problem can be solved on domains of $\mathbb{R}^n$ with a regular enough border. We will sketch the main ideas of Perron's method.

1. Note that a harmonic function is C^∞ by Theorem 5.4.2.

2. The value of the derivatives of a harmonic function on a set is controlled by the values of the function on the border. That is a combination of the maximum principle with the preceding observation (on the plane, it is analogous to Cauchy's estimations of the coefficients of the power series of an holomorphic function).

3. That implies that a set of harmonic functions which is bounded on the border of a set is *equicontinuous*, and therefore any bounded sequence of harmonic functions has a pointwise convergent subsequence to a harmonic function.

4. A function is called *subharmonic* if its value at any point is not greater than its spherical means for radii small enough. Playing with cut-and-paste Poisson solutions on balls it is possible to prove that a function u is harmonic if and only if u and $-u$ are both subharmonics.

5. Let $f \in C(\partial\Omega)$ and consider the set

$$S = \{v \in C(\overline{\Omega}) : v \text{ is subharmonic on } \Omega, v \leq f \text{ on } \partial\Omega\}$$

   and let $u(\mathbf{x}) := \sup\{v(\mathbf{x}) : v \in S\}$. Then u is harmonic.

6. *Barrier property.* If for every $\mathbf{x} \in \partial\Omega$ there exists $h_{\mathbf{x}} \in C(\overline{\Omega})$ harmonic on Ω that attains its maximum on $C(\overline{\Omega})$ exactly at $\mathbf{x}$, then the function u built before satisfies $u(\mathbf{x}) = f(\mathbf{x})$ for all $\mathbf{x} \in \partial\Omega$.

The barrier property is satisfied by domains which are convex or have a piecewise C^2 boundary.

5.9 Rationale and remarks

The notion of Green function is used widely in solving EDOs or PDEs, understood as an integral representation of the solution. From the point of view of Functional Analysis, the *Green operator* is the bounded inverse of the, generally unbounded, differential operator given by the differential equation.

An abstract result of Functional Analysis, based on *Riesz representation theorem*, says that for any bounded domain $\Omega \subset \mathbb{R}^n$ there is a family of Borel measures $\{\mu_x : x \in \Omega\}$ on $\partial\Omega$ such that

$$f(x) = \int_{\partial\Omega} f d\mu_x$$

for every $f \subset C(\overline{\Omega})$ with $\Delta f = 0$ on Ω.

"Can we hear the shape of a drum?" That was a question asked by Marc Kac in 1966. In other words, does the eigenvalues of the Laplacian of a plane region determine its shape. The problem was solved negatively in 1992 by C. Gordon, D. Webb and S. Wolpert.

In relation with the eigenfunctions of the Laplacian on the 3-dimensional sphere, the role of Bessel functions are replaced by the so called *spherical harmonics*. This theory has many applications, but we do not include the details because the computations are rather long and technical.

Despite Hadamard's example, Hilbert space methods play a fundamental role in the study of the Laplace equation and related problems. The space $W^{k,2}(\Omega)$ that we have used is essentially the *Sobolev space* that eases the introduction of differential operators in Hilbert spaces Moreover, variational principles, such as the one of Dirichlet, can adapted for numerical methods (e.g. Galerkin). The idea is to transform the PDE into a variational problem involving a bilinear form. Then, the explicit solutions of the variational problem over finite dimensional spaces can be explicitly computed and they approaches the solution of the original problem as the dimension increases.

5.10 Exercises

1. Show that the following function is harmonic

$$f(x,y,z) = \frac{xy}{(x^2+y^2+z^2)^{5/2}}.$$

2. Consider the following domain in $\mathbb{R}^3$

$$D = \{(x,y,z) : x^2+y^2+z^2 < 4\}$$

and the functions

$$\phi_1(x,y,z) = (x^2+y^2+z^2)^{-1/2},$$

$$\phi_2(x,y,z) = ((x-1)^2+y^2+z^2)^{-1/2}.$$

   (a) Find a harmonic function ψ_1 on D that agrees with ϕ_1 on ∂D.

   (b) Find a harmonic function ψ_2 on D that agrees with ϕ_2 on ∂D.

3. Solve the following problem on $(0,\pi)^2$

$$\begin{cases} \Delta u = 0, \\ u(0,y) = u(\pi,y) = u(x,\pi) = 0, \\ u(x,0) = x^2(\pi - x). \end{cases}$$

4. Solve the following problem on $(0,\pi)\times(0,1)$

$$\begin{cases} \Delta u = 0, \\ u(x,0) = u(x,1) = \sin^3 x, \\ u(0,y) = \sin \pi y, \\ u(\pi,y) = 0. \end{cases}$$

5. Solve the following problem on $(0,1)\times(0,1)$

$$\begin{cases} \Delta u = 0, \\ u(x,0) = (1-x)^2, \\ u(x,1) = u(1,y) = 0, \\ u_x(0,y) = 0. \end{cases}$$

6. Study by the separation of variables method the solutions of the Laplace equation with cylindric symmetry, that is, rotations around the Z axis.

7. Find the solution of the following problem on the rectangle $[0.a] \times [0, b]$, with $A > 0$,

$$\begin{cases} \Delta u = 0, \\ u(0, y) = A, u(a, y) = Ay, \\ u_y(x, 0), u_y(x, b) = 0. \end{cases}$$

8. Solve the following problem on the domain defined by $0 < x + y < 1$, $0 < x - y < 1$,

$$\begin{cases} \Delta u = 0, \\ u(x, -x) = u(x, 1 - x) = u(x, x - 1) = 0, \\ u(x, x) = x(1 - 2x). \end{cases}$$

9. Prove that the solution of the Dirichlet problem on the disc $D(0, R)$ is given by

$$u(r, \theta) = \frac{R^2 - r^2}{2\pi} \int_{-\pi}^{\pi} \frac{f(\theta)\, d\phi}{R^2 - 2Rr\cos(\theta - \phi) + r^2}.$$

10. Let $f(x) = \min\{x, \pi - x\}$.

    (a) Solve the problem $\Delta u = 0$ en $[0, \pi]^2$ with boundary conditions $u(0, y) = u(\pi, y) = 0$, $u(x, 0) = f(x)$, $u(x, \pi) = -f(x)$.

    (b) Prove that the solution obtained takes takes the value 0 on the segment $y = \pi/2$.

11. Let $D \subset \mathbb{R}^3$ be a bounded domain with C^1 border. Show that the integral

$$\iint_{\partial D} \nabla \left(\frac{1}{\rho} \right) \cdot d\overrightarrow{S}$$

where $\rho = \sqrt{x^2 + y^2 + z^2}$ takes the values -4π or 0 depending on 0 being interior or exterior to D. Interpret the integral in terms f the solid angle and make a guess on the values in case $0 \in \partial D$.

12. Proof the following result due to Chasles: Let $D \subset \mathbb{R}^3$ be a bounded domain with C^1 border and let u be a harmonic function defined on an open set containing $\mathbb{R}^3 \setminus D$ that is constant on ∂D, then for any $\mathbf{x}$ outer to $\overline{D}$ we have

$$u(\mathbf{x}) = -\frac{1}{4\pi} \iint_{\partial D} \| \mathbf{x} - \mathbf{y}\|^{-1} \frac{\partial u}{\partial \mathbf{N}}(\mathbf{y})\, dS(\mathbf{y}).$$

13. Prove the following result of Harnack: If u is a nonnegative harmonic function defined in the ball $B(0,R) \subset \mathbb{R}^n$, then for any $\mathbf{x} \in B(0,R)$

$$\frac{R^{n-2}(R - \|\mathbf{x}\|)}{(R + \|\mathbf{x}\|)^{n-1}} u(0) \leq u(x) \leq \frac{R^{n-2}(R + \|\mathbf{x}\|)}{(R - \|\mathbf{x}\|)^{n-1}} u(0).$$

Deduce Liouville's theorem: a harmonic function defined on $\mathbb{R}^n$ that is bounded below (or above) must be constant.

Chapter 6

The heat-diffusion equation

The heat equation in several dimensions

$$u_t - k\Delta u = 0,$$

where $k > 0$, expresses the evolution of a system. The intuition coming from Physics tell us that the system evolves towards some stable situation. Indeed, if the limit $u(\mathbf{x}) = \lim_{t\to+\infty} u(\mathbf{x}, t)$, if it exists, can be understood as an stationary solution of the heat equation, that is, the temperature distribution corresponding to a thermal equilibrium (under the boundary conditions). The solution $u(\mathbf{x})$ still must satisfy the heat equation but as it does not depend on t we have $u_t = 0$ and the equation reduces to Laplace equation

$$\Delta u = 0.$$

Of course, in that case, we could apply the methods of the previous chapter to find the solution of the heat equation.

6.1 Boundary-free 1-dimensional problem

Our aim here is to find a symmetric temperature distribution $\phi(x)$ on $\mathbb{R}$ such that its evolution under the heat equation is affinely equivalent at any time. That means, the distribution after some time is of the form $\alpha\phi(\beta x)$ with $\alpha, \beta > 0$. Note that the energetic interpretation of temperature imposes that the area under the curve must be constant. Therefore, we should take $\alpha = \beta$.

Now, recall that we have the heat equation

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

and we wish a solution of the form

$$u(x,t) = \alpha(t)\phi(\,\alpha(t)x\,)$$

where $\phi \geq 0$ is even and $\alpha > 0$. The physical interpretation suggests that $\alpha' < 0$. Derivation and substitution in 6.1 gives us

$$\alpha'(t)\,\phi(\,\alpha(t)x\,) + \alpha(t)\alpha'(t)x\,\phi'(\,\alpha(t)x\,) = k\alpha(t)^3\phi''(\,\alpha(t)x\,)$$

Fix t and set $a = \alpha(t)$, $b = -\alpha'(t)$ and $g(x) = \phi(ax)$, so we get the ordinary differential equation

$$-b\,g(x) - b\,xg'(x) = ka\,g''(x)$$

whose unique even solutions are multiples of the function

$$g(x) = e^{-\frac{bx^2}{2ka}}$$

Now, the formula for ϕ implies that

$$\phi(x) = e^{-\frac{bx^2}{2ka^3}}$$

Since ϕ cannot depend on t, we can take $\phi(x) = e^{-x^2}$ whenever $\frac{b}{2ka^3} = 1$. In other words, we impose that

$$\alpha'(t) = -2k\,\alpha(t)^3$$

The solutions of the last equation are translations of

$$\alpha(t) = \frac{1}{\sqrt{4kt}}$$

Therefore, the function that we will call the *heat kernel* defined by

$$H(x,t) = \frac{1}{\sqrt{\pi}}\,\alpha(t)\phi(\alpha(t)x) = \frac{1}{\sqrt{4\pi kt}}\,e^{-\frac{x^2}{4kt}}$$

is solution of 6.1, where we divide by $\sqrt{\pi}$ in order to normalize the integral. It is not difficult to check that the function $H(x,t)$, as a family of functions of x depending on a parameter t, is a good family of kernels as $t \to 0^+$. A superposition of translations of that *elementary solution* can be used to prove the following formula for the solution of the heat equation 6.1 with the initial condition $u(x,0) = f(x)$ as

$$u(x,t) = (H(\ ,t) * f)(x) = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{+\infty} e^{-\frac{(x-s)^2}{4kt}}\,f(s)\,ds.$$

6.2 Fundamental solutions in more dimensions

Consider the product of three heat kernels, one for each dimension

$$\mathbf{H}(x, y, z, t) = H(x, t)H(y, t)H(z, t).$$

It is easy to check that it satisfies the heat equation and behave as a good kernel family. In general, the fundamental solution in n-dimensions (whatever it means) will be given by

$$\mathbf{H}(\mathbf{x}, t) = (4\pi kt)^{-n/2}\, e^{-\frac{\|\mathbf{x}\|^2}{4kt}}.$$

The solution for some initial distribution of temperature $\phi(\mathbf{x})$ is given by the convolution

$$u(\mathbf{x}, t) = (\mathbf{H}(\ , t) * \phi)(\mathbf{x}).$$

Example 6.2.1. *Assume that at $t = 0$ the temperature of a plane (with homogeneous thermal diffusivity) is 0, with the exception of of the disc $D(0, 1)$ which is at temperature $T > 0$. Find the temperature at the origin for $t > 0$ and compare with the simpler model with EDOs based in Newton's cooling law.*

Using the formula we have obtained above with the 2-dimensional kernel

$$\mathbf{H}(x, y, t) = \frac{1}{4\pi kt}\, e^{-\frac{x^2+y^2}{4kt}}$$

we get

$$u(0, t) = \iint_{D(0,1)} \frac{T}{4\pi kt}\, e^{-\frac{x^2+y^2}{4kt}} dx\, dy$$

$$= \int_0^{2\pi} \int_0^1 \frac{T}{4\pi kt}\, r\, e^{-\frac{r^2}{4kt}} dr\, d\theta = 2\pi \frac{-T}{2\pi}\, e^{-\frac{r^2}{4kt}}\Big|_0^1 = (1 - e^{-\frac{1}{4kt}})T.$$

The EDOs model would give a solution of the form $u(0, t) = Te^{-kt}$, where the constant k is a different one (2π, which is the perimeter, times the *heat transfer coefficient*). The main point is to compare the shape of the graphs: for instance, the solution given by the heat kernel starts decreasing smoothly, whereas the EDOs model decreases abruptly. For $t \to +\infty$, the PDE model solution behaves $u(0, t) = O(1/t)$, meanwhile the EDO solution has an exponential decay $\sim e^{-kt}$. ■

A nice property of the heat kernels is that they provide an explicit approximation of functions on $\mathbb{R}^n$ by C^∞ ones. Indeed, the application of Theorem 3.6.2 implies that

$$\lim_{t\to 0^+} (\mathbf{H}(\ ,t) * f)(\mathbf{x}) = f(\mathbf{x})$$

uniformly for every f continuous with bounded support and it is relatively easy to prove that $\mathbf{H}(\ ,t) * f$ is C^∞ with a suitable adaptation of the ideas of Proposition 3.3.5 (e). Unfortunately, the support after the convolution is not longer bounded, but it can be amended in several ways.

Physically, the fact that the initial condition has instantaneous influence in all $\mathbb{R}^n$ can be questioned. We should not forget that we are dealing with mathematical models that are merely approximations to reality. There exist some variations of the heat equation, of course more complicate, that introduce the finiteness of heat propagation.

6.3 Boundary problem and uniqueness

Assume $\Omega \subset \mathbb{R}^n$ is a bounded domain. For the heat equation we have the following conservation of energy principle. If

$$E(t) = \iiint_\Omega u(\mathbf{x},t)\, dV,$$

then

$$E'(t) = \iiint_\Omega u_t(\mathbf{x},t)\, dV = k \iiint_\Omega \Delta u(\mathbf{x},t)\, dV = k \iint_{\partial\Omega} \nabla u(\mathbf{x},t)\, d\mathbf{S}$$

which means that the variation of the total energy equals the flux of heat from outside through $\partial\Omega$.

Now we will show that a slight variation of this idea lead to an inequality that can be used for uniqueness of solutions. Consider the problem

$$\begin{cases} u_t(\mathbf{x},t) = k\Delta u \quad x \in \Omega, t > 0, \\ u(\mathbf{x},t) = 0 \quad x \in \partial\Omega, t > 0, \\ u(\mathbf{x},0) = f(x) \quad x \in \Omega, \end{cases}$$

for some continuous $f \in C_0(\Omega)$.

Proposition 6.3.1. *If u is a solution of the previous problem, then for any $t > 0$ we have*

$$\iiint_\Omega u(\mathbf{x},t)^2 dV \le \iiint_\Omega f(\mathbf{x})^2 dV.$$

Proof. We will show that the quantity

$$G(t) = \iiint_\Omega u(\mathbf{x},t)^2\, dV$$

satisfies $G'(t) \le 0$ for $t > 0$ and continuity will do the rest. Indeed,

$$G'(t) = \iiint_\Omega 2uu_t\, dV = 2k \iiint_\Omega u\Delta u\, dV.$$

Apply now Gauss theorem to the field $u\nabla u$ we get

$$0 = \iint_{\partial\Omega} u\nabla u \cdot d\mathbf{S} = \iiint_\Omega (\nabla u \cdot \nabla u + u\Delta u)\, dV.$$

Therefore

$$G'(t) = -2k \iiint_\Omega \|\nabla u\|^2\, dV \le 0$$

as wished. ∎

Corollary 6.3.2. *The solution of the heat equation $u_t = k\Delta u$ is unique on $\Omega \times (0,T)$ with continuous values on the semiboundary*

$$(\Omega \times \{0\}) \cup (\partial\Omega \times [0,T)).$$

The computations in the proof of Proposition 6.3.1 show that, after a while, the average of the squared temperature is not great that the average before. We will show that the temperatures do not reach maximums at any inner point of a vertical strip. More generally, we have the following *maximum principle*.

Proposition 6.3.3. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and let u a C^2 function on the set $\overline{\Omega} \times [0,+\infty)$. Suppose that the following inequality*

$$u_t \le k\Delta u$$

is satisfied on $\Omega \times (0,+\infty)$. Then the maximum of u on $\overline{\Omega} \times [0,T]$ is attained at

$$(\Omega \times \{0\}) \cup (\partial\Omega \times [0,T)).$$

Proof. Suppose at first that u satisfies the stronger inequality $u_t < \Delta u$. Then the statement is true. Indeed, if the maximum is attained at some $(\mathbf{x}_0, t_0) \in \Omega \times (0, T]$, then $\Delta u(\mathbf{x}_0, t_0) \leq 0$ because the Hessian does not have positive eigenvalues. Thus $u_t(\mathbf{x}_0, t_0) < 0$ and $u(\mathbf{x}_0, t) > u(\mathbf{x}_0, t_0)$ for some $t < t_0$, violating the choice of $(\mathbf{x}_0, t_0)$. If u only satisfies the inequality $u_t \leq k\Delta u$, take

$$v(\mathbf{x}, t) = u(\mathbf{x}, t) + \varepsilon \|\mathbf{x}\|^2$$

which satisfies the stronger statement for $\varepsilon > 0$. Therefore,

$$\max\{v(\mathbf{x}, t) : (\mathbf{x}, t) \in \overline{\Omega} \times [0, T]\} = \max\{v(\mathbf{x}, t) : (\mathbf{x}, t) \in (\Omega \times \{0\}) \cup (\partial\Omega \times [0, T))\}.$$

Taking limits when $\varepsilon \to 0^+$ we will get the desired result. ■

6.4 Separation of variables

We will consider firstly the heat equation in one dimension. The separation of variables method works fine if the temperature is constant along time at the butts of the interval, say $[0, \pi]$. Adding a linear function of x, we may assume that $u(0, t) = u(\pi, 0) = 0$ for all $t \geq 0$. Now we write $u(x, t) = X(x)T(t)$, thus we have

$$\frac{T'}{kT} = \frac{X''}{X} = \lambda.$$

As usual, the problem is feasible when $\lambda = -n^2$ with $n \in \mathbb{N}$. That leads to the solutions

$$X_n(x) = \sin nx$$

$$T_n(t) = e^{-kn^2 t}$$

and so, the solution with initial condition $u(x, 0) = f(x)$ is given by

$$u(x, t) = \sum_{n=1}^{\infty} a_n e^{-kn^2 t} \sin nx$$

being the (a_n) the sequence of Fourier coefficients of f when extended to $[-\pi, 0]$ by $f(x) = -f(-x)$. Note that the quick convergence of the series implies that the solution belongs to C^∞ for $t > 0$.

This solution may appear quite different from the ones obtained by convolution with the one dimensional heat kernel, but it is possible to prove that they agree. For instance, for $f(x) = \sin x$ we have

$$\frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{+\infty} e^{-\frac{(x-s)^2}{4kt}} \sin s \, ds = e^{-kt} \sin x$$

that is quite awesome. Instead of a hard computation, we may use the uniqueness of the solution on the vertical strip by showing that the first term of the equality is 0 at $x = 0, \pi$.

Example 6.4.1. *Consider an homogeneous one-dimensional rod of length L at temperature T_0 that is placed at $t = 0$ between two sources of heat having constant temperatures T_0 and T_1 respectively. Compute the temperature on the rod as a function of t and the stationary temperature distribution for $t = +\infty$.*

Assume that the rod is on $[0, L]$. It is easier to start by the stationary temperature. The solution is an affine function given by

$$T_s(x) = T_0 + \frac{(T_1 - T_0)}{L} x$$

for $x \in [0, l]$. The solution $u(x,t)$ of the evolution problem can be reduced to the zero boundary problem by taking

$$v(x,t) = u(x,t) - T_s(x).$$

Without loss of generality we may assume $L = \pi$. Indeed, the change of scale is "compensated" by changing k. However, to make things easier we will assume $k = 1$ that can be understand as a change of time unit. Now, we have

$$v(x,0) = -\frac{(T_1 - T_0)}{\pi} x,$$

and $v(0,t) = v(\pi,t) = 0$ for all $t \geq 0$. The solution can be obtained by developing x as a sinus series on $(-\pi, \pi)$, see Example 3.1.1,

$$x = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nx.$$

Therefore, we have

$$v(x,t) = -\frac{2(T_1 - T_0)}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} e^{-n^2 t} \sin nx,$$

and thus the formula

$$u(x,t) = T_0 + \frac{(T_1 - T_0)}{\pi}\, x - \frac{2(T_1 - T_0)}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\, e^{-n^2 t} \sin nx$$

describes the evolution of the temperature along the rod for $x \in [0, \pi)$. ■

For two or higher dimension, the reduction to value 0 on the boundary can be achieved by adding a suitable harmonic function (solution of a Dirichlet problem). Then, the separation of variables method will lead to the investigation of the eigenvalues and eigenfunctions of the laplacian on the prescribed domain. We could profit the work done in that direction in previous chapters for rectangular and circular domains.

6.5 Non uniqueness = non physical

We have established the uniqueness of the solution for semiboundary conditions and we have found a fancy formula for the boundary-free problem. Therefore, it is quite astonishing that the equation that modelizes a deterministic physical phenomenon does not have unique solution. We will show how to construct a non null solution $u(x,t)$ of the 1-dimensional heat equation that $u(x,0) = 0$ for all $x \in \mathbb{R}$. The idea is to find the solution in the form

$$u(x,t) = \sum_{n=0}^{\infty} f_n(t)\, x^n$$

where f_n are $C^\infty(\mathbb{R})$ non-null functions that vanish for $t < 0$, and such that $(f_n(t))$ goes to 0 for $t > 0$ fast enough to ensure the convergence of the power series. Since u satisfies the heat equation (with $k = 1$) we deduce that taking $f_0 = f$, that we have to find, and $f_1 = 0$, then

$$f_n' = (n+2)(n+1) f_{n+2},$$

so the expression for u becomes

$$u(x,t) = \sum_{n=0}^{\infty} \frac{f^{(n)}(t)}{(2n)!}\, x^{2n}.$$

Now, the choice of f is the delicate part. Fix some $\alpha > 1$ and take

$$f(t) = e^{-t^{-\alpha}} \quad \text{for} \quad t > 0, \quad \text{and} \quad f(t) = 0 \quad \text{otherwise}.$$

This function is known to be $C^\infty(\mathbb{R})$. The convergence of the power series for $t > 0$ is ensured by this tricky inequality: there is some $\theta > 0$ such that

$$|f^{(n)}(t)| < \frac{n!}{(\theta t)^n} e^{-t^{-\alpha}/2}.$$

The inequality can be proved by techniques of Complex Analysis. Now the series that defines u can be majorized by

$$\sum_{n=0}^{\infty} \left| \frac{f^{(n)}(t)}{(2n)!} x^{2n} \right| \leq \sum_{n=0}^{\infty} \frac{|x|^{2n}}{n!(\theta t)^n} e^{-t^{-\alpha}/2} = e^{t^{-1}(x^2/\theta - t^{1-\alpha}/2)}$$

that shows the uniform convergence of the series as $t \to 0^+$ for x bounded. The convergence of the series of the derivatives can be establishes likewise, implying that $u(x,t)$ defined as above is $C^\infty(\mathbb{R}^2)$.

The lack of regularity cannot be a reason to discard that solution of the heat equation. It is clear that physically it has not sense: starting from constant temperature along the line, homogeneity is spontaneously broken. Nevertheless, there are several ways to enforce uniqueness of the solution by appealing to physical plausible hypotheses:

1. Assume that we are working with the *absolute temperature*. In that case, negative values of u are not possible. A deep theorem of Widder gives the uniqueness of the solution under the premise that $u \geq 0$.

2. A real physical process involves a finite amount of energy, that can be translated into the finiteness of integrals of u over x. The precise statement will be given and proven as an application of the *Fourier transform* later in this chapter.

6.6 The Fourier transform

The Fourier transform of a function $f \in L^1(\mathbb{R})$ is defined as the function $\hat{f}$ given by the formula

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{-2\pi i x \xi} dx.$$

Note that the variable of $\hat{f}$ lives in a "different space". We will use often this notation for the transform

$$f(x) \to \hat{f}(\xi),$$

and this other one $\mathcal{F}(f) = \hat{f}$ more suitable to stress the role of the transform as a linear operator defined on $L^1(\mathbb{R})$ (or some other domains as we will see).

Proposition 6.6.1. *Let $f \in L^1(\mathbb{R})$. Then $\hat{f} \in C_0(\mathbb{R})$ and $\|\hat{f}\|_\infty \leq \|f\|_1$.*

Proof. The boundedness of $\hat{f}$ is evident as the inequality $\|\hat{f}\|_\infty \leq \|f\|_1$. Using the denseness of $C_{00}(\mathbb{R})$ in $L^1(\mathbb{R})$ we deduce that $\hat{f}$ for every $f \in L^1(\mathbb{R})$. Finally, $\lim_{\xi\to\pm\infty} \hat{f}(\xi)$ by Theorem 3.3.4. ∎

We will provide a heuristic explanation for the strange definition of the Fourier transform. The trigonometric system adapted to an interval $[-m, m]$ is composed of the functions

$$\{e^{\pi inx/m} : n \in \mathbb{Z}\}$$

and the coefficients of $f \in L^1[-m, m]$ are given by

$$a_n = \frac{1}{2m} \int_{-m}^{m} f(s)e^{-\pi ins/m} ds.$$

Therefore, as the series is pointwise convergent, we have this expression

$$f(x) = \sum_{n\in\mathbb{Z}} \left(\frac{1}{2m} \int_{-m}^{m} f(s)e^{-\pi ins/m} ds \right) e^{\pi inx/m}$$

$$= \sum_{n\in\mathbb{Z}} \frac{1}{2m} \left(\int_{-m}^{m} f(s)e^{-2\pi ins/2m} ds \right) e^{2\pi inx/2m}$$

If $f \in L^1(\mathbb{R})$, now we can take limits with respect to m to have a representation of f whenever there is convergence

$$f(x) = \lim_m \sum_{n\in\mathbb{Z}} \frac{1}{2m} \left(\int_{-\infty}^{\infty} f(s)e^{-2\pi ins/2m} ds \right) e^{2\pi inx/2m}$$

$$= \lim_m \sum_{n\in\mathbb{Z}} \frac{1}{2m} \hat{f}(\frac{n}{2m}) e^{2\pi ixn/2m} = \int_{-\infty}^{\infty} \hat{f}(\xi)e^{2\pi ix\xi} d\xi$$

where the last equality is based in the convergence of a Riemann type sum for the improper integral involving $\hat{f}$, assuming implicitly some regularity of the Fourier transform. All that can be justified *a posteriori* for a certain class of functions.

In order to understand the meaning of the Fourier transform, the following example can cast some light.

Example 6.6.2. *Calculate the Fourier transform of a piece of sinusoidal wave*

$$f(x) = \chi_{[a,b]} e^{2\pi i \omega x},$$

where $a < b$ and ω are reals, and recover the its main features from it.

Put $\delta = b - a$ and $c = (a+b)/2$. Now

$$\hat{f}(\xi) = \int_a^b e^{2\pi i(\omega-\xi)x} dx = \frac{e^{2\pi i(\omega-\xi)b} - e^{2\pi i(\omega-\xi)a}}{2\pi i(\omega-\xi)}$$

$$= e^{2\pi i c} \frac{e^{\pi i(\omega-\xi)\delta} - e^{-\pi i(\omega-\xi)\delta}}{2\pi i(\omega-\xi)} = e^{2\pi i c} \frac{\sin(\pi(\omega-\xi)\delta)}{\pi(\omega-\xi)}.$$

Now, note that the maximum of $|\hat{f}(\xi)|$ is δ and it is attained at $\xi = \omega$. Therefore, the Fourier transform recognizes the duration and the frequency of the wave. Therefore, we can use Fourier transform to analyze a superposition of waves and thus to *signal processing.* ■

In order to justify the heuristic computations that motivated the Fourier transform we need to introduce the so called the *Schwartz class* of functions.

Definition 6.6.3. *The Schwartz class $\mathcal{S}(\mathbb{R})$ is composed of all the functions $f \in C^\infty(\mathbb{R})$ such that*

$$\sup\{|x^n f^{(m)}(x)| : x \in \mathbb{R}\} < +\infty$$

for every $n, m \in \mathbb{N}$.

Note that functions in the Schwartz class are in $L^p(\mathbb{R})$ for all $p \in [1, +\infty]$ together their products by polynomials. Moreover, they made up an algebra of functions stable by differentiation.

Proposition 6.6.4. *The Fourier transform has the following properties for $f \in \mathcal{S}(\mathbb{R})$:*

1. $f(x+t) \to \hat{f}(\xi)e^{2\pi i t\xi}$ *for* $t \in \mathbb{R}$.
2. $f(x)e^{-2\pi i x\tau} \to \hat{f}(\xi+\tau)$ *for* $\tau \in \mathbb{R}$.
3. $f(\alpha x) \to \alpha^{-1}\hat{f}(\alpha^{-1}\xi)$.
4. $f'(x) \to 2\pi i\xi \hat{f}(\xi)$.
5. $-2\pi i x f(x) \to \frac{d}{d\xi}\hat{f}(\xi)$.
6. $(f * g)(x) \to \hat{f}(\xi)\hat{g}(\xi)$.

Proof. All the properties come from standard manipulation of integrals. ∎

Corollary 6.6.5. *$f \in \mathcal{S}(\mathbb{R})$ if and only if $\hat{f} \in \mathcal{S}(\mathbb{R})$.*

With this, the heuristic computation we did before can be justified giving the following result

Theorem 6.6.6. *Let $f \in \mathcal{S}(\mathbb{R})$. Then we can recover f from $\hat{f}$ by the inversion formula*

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(\xi)e^{2\pi i x\xi} d\xi$$

The following example will play an important role in all what follows.

Example 6.6.7. *Let $f(x) = e^{-\pi x^2}$, then $\hat{f}(\xi) = e^{-\pi\xi^2}$.*

Indeed, consider the derivative

$$\frac{d}{d\xi}\mathcal{F}(f)(\xi) = \frac{d}{d\xi}\int_{-\infty}^{+\infty} e^{-\pi x^2}e^{-2\pi i x\xi}dx = \int_{-\infty}^{+\infty} e^{-\pi x^2}(-2\pi i x)e^{-2\pi i x\xi}dx$$

$$= \int_{-\infty}^{+\infty} \frac{d}{dx}(e^{-\pi x^2})ie^{-2\pi i x\xi}dx = i\mathcal{F}(f')(\xi) = -2\pi\xi\mathcal{F}(f)(\xi).$$

That means $\hat{f}$ satisfies the differential equation $\hat{f}'(\xi) = -2\pi\xi\hat{f}(\xi)$. Since

$$\hat{f}(0) = \int_{-\infty}^{+\infty} e^{-\pi x^2}dx = 1,$$

we deduce $\hat{f}(\xi) = e^{-\pi\xi^2}$ as wished. ■

Despite the theory of Fourier transform can be extended beyond the class $\mathcal{S}(\mathbb{R})$, that task is not for us. Nevertheless, we will need the following observation.

Proposition 6.6.8. *The Fourier transform on $L^1(\mathbb{R})$ is injective.*

Proof. From the inversion formula, we know the injectivity of $\mathcal{F}$ on $\mathcal{S}(\mathbb{R})$. Let $f \in L^1(\mathbb{R})$ such that $\hat{f} = 0$. Consider the family of good kernels $K_t(x) = H(x,t)$. We have

$$\mathcal{F}(K_t * f) = \hat{K}_t \hat{f} = 0.$$

Since $K_t * f \in \mathcal{S}(\mathbb{R})$, we get $K_t * f = 0$ for every $t > 0$. But $\lim_t K_t * f = f$ in $\| \cdot \|_1$, thus $f = 0$ as wanted. ■

6.7 Application of Fourier transform to PDEs

The fact that Fourier transform turns derivatives with respecto to x into products by powers of ξ can be used to reduce EDOs to "algebraical problems". In the case of a PDE, the Fourier transform reduces derivatives with respect to spatial variables so the ecuation becomes an ODEs. We will use that technique to establish a result on the uniqueness of the solution of the heat equation, as well as to provide an alternative way to get the formula with the heat kernel.

Theorem 6.7.1. *The initial value problem for the heat equation in one dimension*

$$\begin{cases} u_t(x,t) = u_{xx}(x,t), \\ u(x,0) = f(x) \end{cases}$$

has unique solution among the functions $u(x,t)$ such that

(a) $f, u(\ ,t), u_x(\ ,t), u_{xx}(\ ,t) \in L^1(\mathbb{R})$ for all $t \geq 0$;

(b) for every $T > 0$ there exists $\Phi \in L^1(\mathbb{R})$ such that $|u_t(x,t)| \leq \Phi(x)$ for all $(x,t) \in \mathbb{R} \times [0,T]$.

Proof. Put $v(\xi,t) = \mathcal{F}(u(x,t))$, where the transform is take with respect to the variable x. The second hypothesis allows the derivation of the transform with respect to t, getting so that

$$\mathcal{F}(u_t(x,t)) = v_t(\xi,t).$$

On the other hand, we have

$$\mathcal{F}(u_{xx}(x,t)) = -4\pi^2\xi^2 v(\xi,t).$$

Therefore, the heat equation is transformed into

$$v_t(\xi,t) = -4\pi^2\xi^2 v(\xi,t)$$

with the condition $v(\xi,0) = \hat{f}(\xi)$. This is an EDO with respect to t, being ξ just as a parameter, whose solution is

$$v(\xi,t) = e^{-4\pi^2\xi^2 t}\hat{f}(\xi).$$

Now, note that

$$\mathcal{F}^{-1}(e^{-4\pi^2\xi^2 t}) = \frac{1}{\sqrt{4\pi t}}\, e^{-\frac{x^2}{4t}} = H(x,t)$$

and therefore

$$\mathcal{F}^{-1}(e^{-4\pi^2\xi^2 t}\hat{f}(\xi)) = (H(\ ,t) * f)(x)$$

which is the solution based in the heat kernel found at the beginning of the chapter. Note that we are using the injectivity of the Fourier transform established in Proposition 6.6.8. ∎

6.8 Brownian motion

When we obtained the diffusion equation in the introductory chapter, the aim was to find a the evolution of some probability density. That can be also interpreted as the average behavior of the particles. Now, our aim is to find the individual behavior of *a single* particle under a diffusion process. In order to do that, we will work straight with the random variable rather than its distribution. Remembering the assumptions we did, the properties we should ask to a random variable W_t parameterized with time $t \geq 0$ are the following:

1. $W_0 = 0$;
2. $W_t - W_s$ is independent of W_s for $0 \leq s < t$;
3. $W_t - W_s$ has normal distribution $N(0, t-s)$ for $t > s \geq 0$ ($\sigma^2 = t - s$);
4. the function $t \to W_t(\omega)$ is continuous for every realization ω.

A random variable like W_t is called a *Wiener process* and it is the mathematical model for the *Brownian motion*. We will provide a quite satisfactory construction that gives the continuity for *almost all* ω. Here ℓ_2 will denote the real Hilbert sequence space, see Remark 3.2.5. We need the following technical lemmas.

Lemma 6.8.1. *Let (X_n) be a sequence of independent random variables with distribution $N(0,1)$, and let $(a_n) \in \ell^2$ with $\sigma^2 = \sum_{n=1}^{\infty} a_n^2$. Then the series*

$$\sum_{n=1}^{\infty} a_n X_n$$

taken in L^2 sense is a normal variable with distribution $N(0,\sigma^2)$.

Hint of proof. Independent variables are orthogonal, so the L^2-convergence of the series follows easily. It is well know that the linear combination of independent normal variables is also a normal variable. By the way that could be deduced using the Fourier transform. The result extend to a series taking limits. ∎

Lemma 6.8.2. *Let (X_n) be a sequence of independent random variables with distribution $N(0,1)$ and let $(a_n),(b_n) \in \ell_2$. If (a_n) and (b_n) are orthogonal, then the random variables*

$$\sum_{n=1}^{\infty} a_n X_n \quad \text{and} \quad \sum_{n=1}^{\infty} b_n X_n$$

are independent.

Hint of proof. That is just a generalization of the fact that any rotation on $\mathbb{R}^2$ of the pairing (X,Y) of two independent normal $N(0,1)$ variables define two new independent normal $N(0,1)$ variables. ∎

Theorem 6.8.3. *The Wiener process exists.*

Proof. We will give an explicit construction for $t \in [0,1)$, then it can be taken to larger intervals by re-scaling and thus to $[0,+\infty)$. The re-scaling can be done with the help of the following fact: $\tilde{W}_t = \lambda^{-1} W_{\lambda^2 t}$ is also a Wiener process for any $\lambda > 0$. Now we will consider Fourier series on $L^2[-1,1]$ using the trigonometric adapted orthonormal basis

$$\{1/\sqrt{2}, \cos \pi n x, \sin \pi n x : n \in \mathbb{N}\}.$$

Let $a_n(t)$ for $n \geq 0$ be the cosine Fourier coefficients of $\chi_{[-t/2,t/2]}$ with $t \in [0,2)$.. If $(X_n)_{n=0}^{\infty}$ is a sequence of independent random variables with distribution $N(0,1)$, then the random variable

$$W_t = \sum_{n=0}^{\infty} a_n(t) X_n$$

has variance

$$\sum_{n=0}^{\infty} a_n^2(t) = \|\chi_{[-t/2,t/2]}\|_2^2 = t.$$

Clearly, $W_0 = 0$ and the increments of the form $W_t - W_s$ with $t > 0$ have series coefficients corresponding to the development of

$$\chi_{[-t/2,-s/2]\cup[s/2,t/2]}.$$

That implies the normality, variance equal to $t-s$ and the independence to $\chi_{[-s/2,s/2]}$. In order to study the continuity, we need to compute $a_n(t)$. We have $a_0(t) = t/\sqrt{2}$ and

$$a_n(t) = \int_{-t/2}^{t/2} \cos(\pi n x)\, dx = \frac{2}{\pi n} \sin(\frac{\pi n t}{2}).$$

Therefore,

$$W_t = \frac{t}{\sqrt{2}} X_0 + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin(\pi n t/2)}{n} X_n.$$

Re-scaling, we get this nicer formula valid for $t \in [0,1)$

$$W_t = t X_0 + \sqrt{2} \sum_{n=1}^{\infty} \frac{\sin(\pi n t)}{\pi n} X_n.$$

Finally, for the almost sure continuity we will use a estimation for the tails of the normal. Let X be normal $N(0,1)$ and $b > 1$, then

$$\mathbb{P}(|X| \geq \sqrt{2b}) = \frac{2}{\sqrt{2\pi}} \int_{\sqrt{2b}}^{+\infty} e^{-x^2/2} dx \leq \frac{2}{\sqrt{2\pi}} \int_{\sqrt{2b}}^{+\infty} x e^{-x^2/2} dx = \sqrt{\frac{2}{\pi}}\, e^{-b}.$$

Changing b by bn we get

$$\mathbb{P}(|X_n| \leq \sqrt{2bn}) \geq 1 - \sqrt{\frac{2}{\pi}}\, e^{-bn},$$

and so

$$\mathbb{P}(\{|X_n| \leq \sqrt{2bn} : n \in \mathbb{N}\}) \geq \prod_{n=1}^{\infty}(1 - \sqrt{2/\pi}\, e^{-bn}).$$

The infinite product on the right hand-side is convergent and it can be taken arbitrarily close to 1 as $b \to +\infty$ since the associated series

$$\sum_{n=1}^{\infty} \sqrt{2/\pi}\, e^{-bn} = \sqrt{\frac{2}{\pi}} \frac{e^{-b}}{1 - e^{-b}}$$

can be taken as close to 0 as we wish. Now, if the sample values $X_n(\omega) = a_n$ satisfy $|a_n| \leq \sqrt{2bn}$, then the series

$$\sum_{n=1}^{\infty} \frac{\sin(\pi n t/2)}{\pi n}\, a_n$$

is uniformly convergent in intervals of the form $t \in [0, a]$ with $a < 1$ by Dirichlet criterion. Since that happens for ω form a set of measure arbitrarily close to 1, our claim on the almost sure continuity is proven. ∎

The construction for W_t we have done can be interpreted as a *random Fourier series*. Giving explicitly values to the coefficients (a_n) with normal $N(0, 1)$ we will get the path of a particle with Brownian motion. Typically, the function defined by those series is nowhere differentiable.

We have constructed the one-dimensional Brownian motion. If we wish an n-dimensional Brownian motion it is enough to take n independent Wiener processes. Despite the elemental construction, some characteristics change as the dimension increases. Indeed, the probability of coming back to a bounded set is 0 for $n \geq 3$.

6.9 Rationale and remarks

The heat equation can be discussed from the point of view of *semigroup theory*. Indeed, if $S_t(f) = u(x, t)$ where u is the (physically admissible) solution of the problem $u_t = u_{xx}$ and $u(x, 0) = f(x)$, then one has $S_{t_1+t_2} = S_{t_1} \circ S_{t_2}$ for any $t_1, t_2 \geq 0$. The most representative result in semigroup theory with implications to EDPs is the Yosida-Hille theorem.

The Fourier transform on $\mathcal{S}(\mathbb{R})$ preserves the norm $\|\cdot\|_2$ and that can be used to extend it to an isometry of $L^2(\mathbb{R})$, which is known as the *Plancherel theorem*. The Fourier transform also behaves well with respect the norms $\|\cdot\|_p$ with $1 < p < +\infty$, thanks to *operator interpolation* results, particularly the *Riesz-Thorin theorem*.

The theories of Fourier series and Fourier transform can both considered particular cases of the Fourier analysis on abelian topological groups. According to Pontryagin, the dual of a topological group consist of the *characters*, namely, the continuous group homeomorphisms into $\mathbb{T}$. Notably, the dual of $\mathbb{T}$ itself is $\mathbb{Z}$ and viceversa, meanwhile $\mathbb{R}$ is its own dual. Fourier transform in that setting carries a function f defined on an abelian topological group to another function $\hat{f}$ (the transform) defined on the dual of the group, using integration with respect to the *Haar measure*. Therefore, the Fourier complex coefficients (for a function on $L^1(\mathbb{T})$) and the Fourier transform (for a function on $L^1(\mathbb{R})$) are the same notion from that point of view. That explains, somehow the use of the same notation.

A different extension of the theory is given by *wavelets*, that allow discretization on unbounded domains, $\mathbb{R}$ or $\mathbb{R}^n$. From a suitable single function of bounded support, by scaling and translations (usually dyadic) it is possible to generate a Hilbert basis of $L^2(\mathbb{R})$, the wavelet base.

If we wish to enjoy the *operational calculus*, that is, a functional transform that takes derivatives into polynomials or pumpkins, but without messing with complex numbers we could use *Laplace transform* instead of Fourier's.

The Brownian motion can be applied to the solution of the Dirichlet problem. Roughly speaking, the harmonic function u on Ω that agrees with a given function f on $\partial\Omega$ is taken as follows: $u(x)$ is the average (expectation) on the values of f over the Brownian paths starting at x.

6.10 Exercises

1. Consider a hollow ball B made of some homogeneous material

$$B = \{(x, y, z) : r^2 \leq x^2 + y^2 + z^2 \leq R^2\}$$

with $0 < r < R < +\infty$. Assume that the inner surface is kept at temperature T_1 and the outer surface at temperature, both of them constant T_2 constantes. Find the stationary distribution of temperatures on B. Where is attained the temperature $(T_1 + T_2)/2$?

2. Solve the problem on $[0, \pi] \times [0, +\infty)$

$$\begin{cases} u_t - ku_{xx} = 0, \\ u(0, t) = u(\pi, t) = 0, \\ u(x, 0) = \sin^3 x. \end{cases}$$

3. Solve the problem on $[0, \pi] \times [0, +\infty)$

$$\begin{cases} u_t - ku_{xx} = 0, \\ u(0, t) = u(\pi, t) = 0, \\ u(x, 0) = x(\pi - x). \end{cases}$$

4. Solve the problem on $[0, \pi] \times [0, +\infty)$

$$\begin{cases} u_t - ku_{xx} = 0, \\ u_x(0, t) = u_x(\pi, t) = 0, \\ u(x, 0) = \sin x. \end{cases}$$

5. A cube of homogeneous materia is at temperature T_1 before being sunk into a fluid at constant temperature T_2. Compute the distribution of temperatures a while after.

6. Solve the following problem on $[0, \pi] \times [0, +\infty)$ by the separation of variables method

$$\begin{cases} u_t - t^2 u_{xx} - u = 0, \\ u(0, t) = u(\pi, t) = 0, \\ u(x, 0) = x(\pi - x). \end{cases}$$

7. Compute the Fourier transform of the Cauchy probability density

$$f(x) = \frac{1}{\pi(1 + x^2)}.$$

Deduce that the central limit theorem fails for a sequence of independent variables with the Cauchy distribution.

8. Compute the Fourier transform of the "triangle function"

$$f(x) = \max\{0, 1 - |x|\}.$$

9. Prove that the Fourier transform of $f \in L^1(\mathbb{R})$ takes real values only, if and only if $\overline{f(x)} = f(-x)$ for almost all $x \in \mathbb{R}$.

10. Let $f \in \mathcal{S}$ (Schwartz class). Prove that the $\sum_{n\in\mathbb{Z}} f(x+n)$ converges uniformly on compact subsets of $\mathbb{R}$ and defines C^∞ periodic function.

11. Prove Poisson's summation formula: given $f \in \mathcal{S}$, then

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n).$$

12. The function $e^{-\pi x^2}$ is an eigenfunction of the Fourier transform regarded as an operator on $L_2(\mathbb{R})$ with 1 as eigenvalue. Find more independent eigenfunctions and their corresponding eigenvalues.

Chapter 7

Appendix: ODEs

Existence, uniqueness, regularity. All we need can be summarised in the following result.

Theorem 7.0.1. *Consider the following differential equation*

$$\mathbf{y}' = F(x, \mathbf{y}).$$

Assume that F is defined on some domain $\Omega \subset \mathbb{R} \times \mathbb{R}^n$, where it is continuous. Assume moreover that F is locally Lipschitz with respect to the variable $\mathbf{y}$. Then:

1. *For every $(x_0, \mathbf{y}_0) \in \Omega$ there is a unique solution f of the equation defined on a neighbourhood of x_0 such that $f(x_0) = \mathbf{y}_0$;*
2. *Moreover, if $F \in C^k(\Omega)$ then f is $(k+1)$ times differentiable;*
3. *Assuming the last assumption, if $k \geq 1$, then the assignment*

$$(x_1, \mathbf{y}_1, x) \to f(x)$$

where f satisfies $f(x_1) = \mathbf{y}_1$ is defined in a neighbourhood of $(x_0, \mathbf{y}_0)$, where it is k times differentiable with respect $(x_0, \mathbf{y}_0)$.

A dynamical system does not contain explicitly the independent variable, e.g. $\mathbf{y}' = F(\mathbf{y})$. That implies, obviously, the existence of solutions f such that $f(x_0) = \mathbf{y}_0$ for any $\mathbf{y}_0 \in \Omega$. Moreover, if $\Omega = \mathbb{R}^n$ and F is bounded, then $f(x)$ is defined for every $x \in \mathbb{R}$.

Singular solutions. Existence and uniqueness theorem is established for ordinary differential equations (ODE) of the form

$$y' = F(x, y).$$

Most times ODEs are given as

$$F(x, y, y') = 0 \tag{7.1}$$

that may fail the uniqueness theorem in two ways. Assume that for x_0, y_0 the equation $F(x_0, y_0, z) = 0$ admits several solutions z_1, z_2... and around them the equation can be solved. That implies the existence of several solutions of 7.1 passing at (x_0, y_0) with different angles.
It is more interesting the following scenario: assume that the general solution of 7.1 is given as $f(x, y, \lambda) = 0$, being $\lambda \in \mathbb{R}$ the parameter. If the family of curves $f(x, y, \lambda) = 0$ admits an envelope $g(x, y) = 0$, then the envelope also satisfies 7.1. Indeed, for any point (x, y) such that $g(x, y) = 0$ there is some λ such that $f(x, y, \lambda) = 0$. Since that solution and g are tangent at (x, y), they share a common value for y'. We say that $g(x, y) = 0$ is a *a singular solution* of 7.1 (roughly, a solution not contained in the general solution).
The fact that the uniqueness is violated at any point of the envelope g means that the implicit function theorem fails. Since regularity is not the issue, we should have

$$\frac{\partial F}{\partial y'}(x, y, y') = 0$$

at the points of the singular solution, which provides a way to find it without using the general solution of 7.1.

Chapter 8

Appendix: Probability theory survival kit

A *probability space* is a measure space $(\Omega, \Sigma, \mathbb{P})$ with $\mathbb{P}(\Omega) = 1$. The space Ω represents all the possible outcomes of a random trial and its detailed description is difficult in general. For that reason, a great deal of Probability Theory is to provide tools to work on $\mathbb{R}$ instead of Ω. The elements of Σ are sets of single outcomes for whom the probability is defined. We will use the following convention

$$\{X \leq x\} := \{\omega \in \Omega : X(\omega) \leq x\},$$
$$\{X \in A\} := \{\omega \in \Omega : X(\omega) \in A\}.$$

The probability $\mathbb{P}$ is a measure of plausibility. It has also an statistical interpretation if the trial can be reproduced indefinitely.

The real, complex or vector valued measurable functions defined on (Ω, Σ) are called *random variables* and denoted, usually, by capital letters $X, Y, \ldots$ The integration of random variables with respect to $\mathbb{P}$ is called *expectation* and denoted $\mathbb{E}$, that is

$$\mathbb{E}(X) := \int_\Omega X \, d\mathbb{P},$$

assuming integrability. The expectation operator is linear as it is the integral

$$\mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E}(X) + \beta \mathbb{E}(Y),$$

for any random variables X, Y and $\alpha, \beta \in \mathbb{R}$. The *mean* of a random variable X is simply its expectation $\mathbb{E}(X)$. The *variance* of a real random variable X

is

$$\mathrm{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

We will always assume that the variance is finite unless specified otherwise.

A random variable X is said *continuous* if there exists a distribution *density* f, a positive Borel measurable function, such that

$$\mathbb{P}(\{X \leq x\}) = \int_{-\infty}^{x} f(s)\, ds.$$

In such a case, we also have

$$\mathbb{P}(\{X \in A\}) = \int_A f(s)\, ds$$

for every $A \subset \mathbb{R}$ Borel. Obviously, $\int_{-\infty}^{+\infty} f = 1$. Mean and variance of a continuos random variable X can be represented in terms of its density f by using the following relations

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} sf(s)\, ds;$$

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} s^2 f(s)\, ds;$$

and, in general, for any *moment*, we have

$$\mathbb{E}(X^n) = \int_{-\infty}^{+\infty} s^n f(s)\, ds.$$

The proof is quite easy for f being a simple function. The general case is obtained by approximation and Lebesgue convergence theorems.

Two events $A, B \subset \Sigma$ are said to be *independen*t if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\, \mathbb{P}(B).$$

Two sub-σ-algebras $\Sigma_1, \Sigma_2 \subset \Sigma$ are said independent if A and B are independent for any choice $A \in \Sigma_1$ and $B \in \Sigma_2$. Two real random variables X and Y are independent if the sub-σ-algebras $X^{-1}(\mathrm{Borel}(\mathbb{R}))$ and $Y^{-1}(\mathrm{Borel}(\mathbb{R}))$ are independent. Independence extends to n-tuples of sets in the obvious way. Analytically, the independence of two random variables X and Y implies

$$\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y).$$

Again, the proof is easy for simple functions. For the variance of the sum of two independent variables we have a nice result. Indeed,

$$\mathrm{Var}(X+Y) = \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2$$

$$= \mathbb{E}(X^2 + 2XY + Y^2) - (\mathbb{E}(X)^2 + 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y)^2)$$

$$= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)^2)$$

$$= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \mathrm{Var}(X) + \mathrm{Var}(Y),$$

applying the independence hypothesis in the second to last equality.

We will discuss independence in terms of the density function for continuous random variables. If X, Y are independent with density functions f, g respectively, then

$$\mathbb{P}(\{(X,Y) \in A \times B\}) = \mathbb{P}(\{X \in A\})\mathbb{P}(\{Y \in B\})$$

$$= \int_A f(s)\,ds \int_B g(t)\,dt = \iint_{A\times B} f(s)g(t)\,dsdt$$

Being this true for all the sets of type $A \times B$, it extends to any set of $D \in \Sigma \otimes \Sigma$ as

$$\mathbb{P}(\{(X,Y) \in D\}) = \iint_D f(s)g(t)\,dsdt$$

Now, we will prove that the sum of two continuous independent random variables is continuous and we will provide a formula for its density. We keep the notation from above. We have

$$\mathbb{P}(\{X+Y \leq z\}) = \iint_{D(z)} f(s)g(t)\,dsdt$$

where $D(z) = \{(x,y) : x+y \leq z\}$. We perform a change to variables (s,r) where $r = s+t$ (note that the Jacobian is 1).

$$\iint_{D(z)} f(s)g(t)\,dsdt = \int_{-\infty}^{z} \int_{-\infty}^{+\infty} f(s)g(r-s)\,dsdr = \int_{-\infty}^{z} h(r)\,dr$$

where

$$h(r) = \int_{-\infty}^{+\infty} f(s)g(r-s)\,ds,$$

the convolution product $f * g$ we shaw in Chapter 3. Therefore, the density of $X + Y$ is $f * g$.

The square root of the variance is called the *deviation*, usually denoted by σ. A random variable is said to be *normal* with mean μ and variance σ^2 (referred as $N(\mu, \sigma^2)$) if it has a density of the form

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

It easy to check that the sum of two independent normal variables $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ is normal too with parameters $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Bibliography

[1] T. M, Apostol, *Análisis Matemático*, (1^a ed.), Ed. Reverté, Barcelona, 1960.

[2] V. I. Arnold, *Lectures on Partial Differential Equations*, Springer - Phasis, 2004.

[3] I. Bronshtein, K. Semendiaev, *Manual de Matemáticas para ingenieros y estudiantes*, Editorial MIR, Moscú, 1973.

[4] H. Brezis, *Analyse Fonctionelle: théorie et applications*, Dunod, 2005.

[5] G. Bruhat, *Cours de Physique générale*, (4 vol.), Masson, 1963-1968.

[6] H. Cartan, *Formas Diferenciales*, Omega - Colección Métodos, Barcelona, 1972.

[7] B. Cascales, J. M. Mira, J. Orihuela, M. Raja, *Análisis Funcional*, e-Lectolibris, Murcia, 2018.

[8] G. Chilov, *Analyse Mathématique*, 2 vol., MIR, Moscou, 1978.

[9] G. Choquet, *Cours de Topologie*, (2^{eme} ed.), Dunod, Paris, 2000.

[10] K. L. Chung, *Green, Brown, and Probability & Browninan Motion on the line*, World Scientific, Singapore 2002.

[11] D. L. Cohn, *Measure Theory*, Birkhäuser, 2013.

[12] R. Courant, F. John, *Introducción al Cálculo y al Análisis Matemático*, (2 vol.), Limusa, México, 1982.

[13] R. Courant, D. Hilbert, *Methods of Mathematical Physics*, (2 vol.), Wiley Classics Library, New York, 1989.

[14] E. A. DESLOGE, *Classical Mechanics*, John Wiley & Sons Inc, 1982.

[15] J. DIEUDONNÉ, *Fundamentos de Análisis Moderno*, Ed. Reverté, Barcelona, 1979.

[16] P. DREYFUSS, *Introduction à l'analyse des équations de Navier-Stokes*, Ellipses, Paris, 2012.

[17] C. H. EDWARDS JR., *Advanced Calculus of Several Variables*, Dover Publ. Inc., New York, 1994.

[18] K. FALCONER, *Fractal Geometry : Mathematical Foundations and Applications*, Wiley, 2006.

[19] J. A. FERNÁNDEZ VIÑA, *Análisis Matemático*, (3 vol.), Tecnos, Madrid, 1986.

[20] S. K. GODUNOV, *Ecuaciones de la física matemática*, URSS, Moscú 1994.

[21] E. GOURSAT, *Cours d'Analyse mathématique*, 3 vol., Gauthier-Villars, Paris, 1902-1913.

[22] G. JÄGER, *Física Teórica*, Labor, 1959.

[23] F. JOHN, *Partial Differential Equations*, (4^{th} ed.), Springer, New York 1982.

[24] G. JOOS, I. M. FREEMAN, *Theoretical Physics*, (3^{rd} ed.), Blackie & Son LMT, Glasgow, 1960.

[25] A. N. KOLMOGÓROV, S. V. FOMÍN, *Elementos de la Teoría de Funciones y del Análisis Funcional*, Editorial MIR, Moscú, 1975.

[26] A. LICHNEROWICZ, *Elementos de cálculo tensorial*, Aguilar, Madrid 1965.

[27] A. MISHCHENKO, A. FOMENKO, *A course of Differential Geometry and Topology*, MIR Publishers, Moscow, 1988.

[28] S. MIZOHATA, *The theory of partial differential equations*, Cambridge University Press, 1973.

[29] J. D. MURRAY, *Mathematical Biology*, 3^{rd} edition, 2 vol., Interdisciplinary Applied Mathematics, Springer, 2004.

[30] I. Peral Alonso, *Primer Curso de Ecuaciones en Derivadas Parciales*, Universidad Autónoma de Madrid 2004. `https://matematicas.uam.es/~ireneo.peral/libro.pdf`

[31] I. G. Petrovski, *Ordinary Differential Equations*, Dover Publications, New York, 1973.

[32] P. Puig Adam, *Curso Teórico-Práctico de Cálculo Integral*, Biblioteca Matemática S. L., Madrid, 1975.

[33] M. Raja, *Functions of Several Real Variables*, Apuntes de Clase, Universidad de Murcia, 2021, `https://webs.um.es/matias/miwiki/lib/exe/fetch.php?media=fvvr.pdf`.

[34] J. Rey Pastor, P. Pi Calleja, C. A. Trejo, *Análisis Matemático* (3 vol.) Kapelusz, Buenos Aires, 1968.

[35] W. Rudin, *Análisis Real y Complejo*, (3^a ed.) McGraw-Hill, 1987.

[36] S. Salsa, *Partial Differential Equations in Action. From Modelling to Theory*, Springer, 2010.

[37] L. A. Santaló, *Vectores y Tensores, con sus Aplicaciones*, EUDEBA, Buenos Aires, 1962.

[38] L. Schwartz *Métodos matemáticos para las ciencias físicas*, Selecciones Científicas, Madrid, 1969.

[39] J. H. Shapiro, *A Fixed-Point Farrago*, Springer, 2016.

[40] M. Spivak, *Cálculo en Variedades*, Ed. Reverté, Barcelona, 1988.

[41] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.

[42] E. M. Stein, R. Shakarchi, *Princeton Lectures in Analysis*, (4 vol.), Princeton and Oxford University Press, 2003.

[43] A. N. Tijonov, A. A. Samarsky, *Ecuaciones de la Física Matemática*, MIR, Moscú, 1983.

[44] G. Valiron, *Théorie des Fonctions*, (3^{eme} ed.) Masson, Paris, 1990.

[45] C. E. Weatherburn, *Advanced Vector Analysis*, G. Bell and Sons, LMT, London, 1943.

[46] H. F. Weinberger, *Ecuaciones en derivadas parciales*, Reverté, 2011.