Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE

Corso di Laurea (Triennale) in Matematica

Fourier series, Sturm-Liouville problems and the one-dimensional wave equation

Tesi di Laurea in Analisi Matematica

Relatore:
Chiar.ma Prof.ssa
Simonetta Abenda

Presentata da:
Gabriele Calà

Anno Accademico 2023/2024

"Obvious" is the most dangerous word in mathematics.

Introduction

Fourier series, the one-dimensional wave equation, and regular Sturm-Liouville problems have long been fundamental tools in mathematics and physics. Since their inception, these concepts have been intertwined, with each providing insights and applications for the others. This dissertation dwells on the intricate relationships between these three areas, exploring their theoretical foundations and practical implications. By investigating generalized Fourier series, solving the wave equation and analyzing Sturm-Liouville problems, we aim to contribute to a deeper understanding of their functioning.

First, let us linger on the historical context for a while. The concept of Fourier series can be traced back to the ancient Greeks, who studied periodic phenomena like sound waves. However, the systematic development of Fourier series theory began in the 18th century.

The main figures that contributed to the development of this theory are Joseph Fourier and Leonhard Euler. The first one was a French mathematician, who gave life to the idea of representing periodic functions as infinite sums of sine and cosine functions while investigating on heat conduction. The second one was a Swiss mathematician who, among many accomplishments in various field of mathematics, pushed the growth of the mathematical tools necessary for the understanding of Fourier theory, including the concept of complex numbers. Sturm-Liouville problems obviously inherited their name from Jacques C. F. Sturm and from Joseph Liouville, who developed the theory of Sturm-Liouville problems in the 1830s and contributed to the study of their applications. Jean-Baptiste d'Alembert was the one who derived the wave equation in the 18th century while studying the vibrations of a string, while Pierre-Simon Laplace extended the wave equation to three dimensions to describe the propagation of gravitational waves.

But enough about history. This dissertation aims to lay the groundwork for the understanding of generalized Fourier series (chapter 2), then study how to actually build them through the investigation of regular Sturm-Liouville problems (chapter 3) and finally see them in action while solving the one-dimensional wave equation (chapter 4).

More specifically, in chapter 2 we will define what an inner product space is and prove that L^2 is one, and that it is a normed vector space too. After recalling the definitions of orthogonality and orthonormality, we ask ourselves if it is possible to find a set of orthonormal functions $\{\phi_n\}_{n\in\mathbb{N}}$ such that for any $f \in L^2$ one can write it as $f = \sum\langle f, \phi_n\rangle\phi_n$. In order to do that we define and study norm convergence in L^2. Finally, we end up proving that it is possible to find sets of functions as described a few lines back. Those sets are said to be orthonormal bases and the series you can build with them are the so-called (generalized) Fourier series.

In chapter 3 we are going to study how one can actually build orthonormal bases. We will be able to do that thanks to the preliminary study of adjoint linear operators, during which we are going to focus on the second-order differential operator L defined as $L(f) = rf'' + qf' + pf$, and study the so-called Lagrange's identity. The latter will explicitly tell us the difference between formally and actually self-adjoint operators. Afterwards, we define what regular Sturm-Liouville problems are and study two of the most important theorems regarding them. The most important result (and fulcrum of this whole thesis) is the fact that, by solving these kind of boundary value problems, one can and will find enough functions to put together orthonormal bases. We actually build the most important ones, too. Finally, with all these instruments at hand, we investigate further the convergence properties of Fourier series with respect to the regularity of the function they are supposed to be converging to, for it will be fundamental for the study of the one-dimensional wave equation.

During chapter 4, we will first introduce linear partial differential operators and equations, and then take a look at what the superposition principle is and where its usefulness lies. Afterwards, we will dwell on the derivation of the one-dimensional wave equation's model by making reasonable physical assumptions and mathematical deductions. Afterwards, once we have the differential equation at hand, we will try to solve it with the technique of the separation of variables. By doing this, we will find exactly what we were expecting, thanks to the knowledge we acquired about regular Sturm-Liouville problems. Ultimately, we are going to study the global homogeneous Cauchy problem regarding the one-dimensional wave equation, which will lead us to d'Alembert's formula.

Contents

Chapter 1

Notations

In this chapter we will clarify the notation used in order not to dive in it later. For starters, the set of natural numbers $\mathbb{N}$ is considered by default without the zero element, $\mathbb{N} \cup \{0\}$ will be denoted as $\mathbb{N}_0$ and the set of integer numbers will be denoted $\mathbb{Z}$, as usual. $\mathbb{R}$ and $\mathbb{C}$ will be, respectively, the real and complex numbers sets, with $\mathbb{R}^n$ and $\mathbb{C}^n$ being the correspondents n-dimensional vector spaces. The elements of these vector spaces are essentially ordered n-tuples, and will be denoted as $\mathbf{x} = (x_1, x_2, ..., x_n)$ for vectors both in $\mathbb{R}^n$ and in $\mathbb{C}^n$. In the application we are going to see, t will represent the time variable and x the spatial variable.

Partial derivatives are denoted as following:

$$u_x = \frac{\partial u}{\partial x}, \quad u_{xx} = \frac{\partial^2 u}{\partial x^2}, \quad u_{xy} = \frac{\partial^2 u}{\partial x \partial y}, \quad \text{etc.}$$

A function f of one real variable is said to be in $\mathcal{C}^k(I)$, where I is an interval, if its derivatives $f', f'', ..., f^{(k)}$ exist and are continuous on I. If a function $f \in \mathcal{C}^k(I)$ for all $k \in \mathbb{N}$ then it is said that $f \in \mathcal{C}^\infty(I)$.

The following is the (classic) notation for intervals:

$$\begin{aligned} [a,b] &= \{x \in \mathbb{R} : a \leq x \leq b\} \quad & [a,b) &= \{x \in \mathbb{R} : a \leq x < b\} \\ (a,b) &= \{x \in \mathbb{R} : a < x < b\} \quad & (a,b] &= \{x \in \mathbb{R} : a < x \leq b\}. \end{aligned}$$

Chapter 2

Complete orthonormal systems and Fourier series

In this chapter we will first give an introduction on inner product spaces and their associated norm. Immediately afterwards, we are going to define the L^2 function spaces and its own inner product. Our main focus becomes finding, for L^2 spaces, the equivalent of an orthonormal basis in $\mathbb{C}^k$. The main problem is that L^2 is an infinite-dimensional vector space, so we have to worry about convergence. With the completeness of L^2 and some theorems about orthonormal systems, we shall have our final answer, together with the (generalized) Fourier series.

2.1 From $\mathbb{C}^k$ to L^2

Definition 2.1. *An **inner product space** (over $\mathbb{C}$) is the pair $(V, \langle\cdot,\cdot\rangle)$, where V is a **vector space** over $\mathbb{C}$, while $\langle\cdot,\cdot\rangle$ is a so-called **inner product**, which is a map*

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

that, for all vectors $\boldsymbol{a}$,$\boldsymbol{b}$,$\boldsymbol{c}$ in V and for all scalars z, w in $\mathbb{C}$, satisfies the followings:

1. ***Conjugate** (or **Hermitian**) **symmetry**:*

$$\langle \boldsymbol{a}, \boldsymbol{b}\rangle = \overline{\langle \boldsymbol{b}, \boldsymbol{a}\rangle}. \tag{2.1}$$

As $a = \overline{a}$ if and only if a is real, $\langle \boldsymbol{a}, \boldsymbol{a}\rangle$ is always a real number;

2. ***Linearity** in the first argument:*

$$\langle z\boldsymbol{a} + w\boldsymbol{b}, \boldsymbol{c}\rangle = z\langle \boldsymbol{a}, \boldsymbol{c}\rangle + w\langle \boldsymbol{b}, \boldsymbol{c}\rangle, \tag{2.2}$$

and this property, combined with conjugate symmetry, implies that the inner product is ***conjugate linear*** *(or* ***antilinear****) in the second argument, which means that*

$$\langle \boldsymbol{a}, z\boldsymbol{b} + w\boldsymbol{c}\rangle = \overline{z}\langle \boldsymbol{a}, \boldsymbol{b}\rangle + \overline{w}\langle \boldsymbol{a}, \boldsymbol{c}\rangle; \tag{2.3}$$

3. ***Positive-definiteness****: if* $\boldsymbol{a} \neq \boldsymbol{0}$ *then*

$$\langle \boldsymbol{a}, \boldsymbol{a}\rangle > 0.$$

Every inner product space has an **associated norm** defined by

$$\|\mathbf{a}\| = \langle \mathbf{a}, \mathbf{a}\rangle^{1/2}. \tag{2.4}$$

The standard inner product of two complex vectors $\mathbf{a}, \mathbf{b} \in \mathbb{C}^k$ is defined as

$$\langle \mathbf{a}, \mathbf{b}\rangle = a_1\overline{b}_1 + a_2\overline{b}_2 + ... + a_k\overline{b}_k, \tag{2.5}$$

hence the standard norm of a single complex vector is

$$\begin{aligned} \|\mathbf{a}\| &= \langle \mathbf{a}, \mathbf{a}\rangle^{1/2} && (2.6)\\ &= (a_1\overline{a}_1 + ... + a_k\overline{a}_k)^{1/2} && (2.7)\\ &= (|a_1|^2 + ... + |a_k|^2)^{1/2}. && (2.8) \end{aligned}$$

Inspired by these definitions, we want to translate them into the language of function spaces. To be able to do so, we have yet to define the function space in which we are going to work. By using the **Lebesgue integral** and limiting ourselves to **measurable** functions, which is a very weak regularity hypothesis, we allow ourselves to define

$$\mathcal{L}^2[a,b] = \left\{ f : \int_a^b |f(x)|^2 dx < +\infty \right\},$$

the space of **square-integrable** functions. This is, first of all, an **infinite-dimensional** vector space, and this fact alone will be the source of most of our problems. But let us not hesitate further.

For starters, we can define its own inner product. To do that, as we said a few lines back, we can let us be inspired by the definition of the standard inner product in $\mathbb{C}^k$. The key is to imagine that vectors of $\mathbb{C}^k$ are functions defined from the discrete domain $\{1, ..., k\}$ to $\mathbb{C}^k$ and that the vectors of $\mathcal{L}^2[a,b]$ are their continuous version. The same line of reasoning can be applied to discrete sums, of which the continuous version is the integral. Therefore we define

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}dx, \quad \text{and} \quad \|f\| = \left(\int_a^b |f(x)|^2\right)^{1/2}, \tag{2.9}$$

as our inner product and associated norm. This definition, although intuitive, has to be checked, of course.

First, it is not obvious that the inner product of two functions $f, g \in \mathcal{L}^2[a,b]$ is finite.

Remark 2.2. It is known that for any real numbers s and t

$$s^2 + t^2 - 2st = (s-t)^2 \geq 0 \Longrightarrow st \leq \frac{1}{2}(s^2 + t^2).$$

Therefore

$$|f(x)\overline{g(x)}| \leq \frac{1}{2}(|f(x)|^2 + |g(x)|^2),$$

and if $f, g \in \mathcal{L}^2[a,b]$, the integral

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}dx$$

is absolutely convergent. this means that the inner product is well-posed for all functions $f, g \in \mathcal{L}^2[a,b]$.

Second, there is a problem with the positive-definiteness of the inner product (hence with its associated norm too), meaning that both fail to be true, if we consider a function $f \in \mathcal{L}^2[a,b]$ such that $f(x) = 0$ for almost every $x \in [a,b]$. However, this problem is easily fixed: we can consider the equivalence relation

$$f \sim g \Longleftrightarrow f(x) = g(x) \text{ for almost every } x \in [a,b].$$

Conventionally, $\mathcal{L}^2[a,b] /_{\sim}$ is denoted $L^2[a,b]$.

Therefore $(L^2[a,b], \langle \cdot, \cdot \rangle, \| \cdot \|)$ is a normed inner product space.

Definition 2.3. *A set $\{f_n\}_{n\in\mathbb{N}}$ is said to be an **orthogonal system** (or **set**) when*

$$\int_a^b f_i(x)\overline{f_j(x)}dx = 0 \quad for\ i \neq j.$$

*An **orthonormal system** (or **set**) is an orthogonal system of which the elements are functions f with $\|f\| = 1$.*

We are now able to utilize the Cauchy-Schwarz inequality, the triangle inequality and the Pythagorean theorem, which explicitly work as follows:

$$\left| \int_a^b f(x)\overline{g(x)}dx \right| \leq \sqrt{\int_a^b |f(x)|^2 dx} \sqrt{\int_a^b |g(x)|^2 dx},$$

$$\sqrt{\int_a^b |f(x) + g(x)|^2 dx} \leq \sqrt{\int_a^b |f(x)|^2 dx} + \sqrt{\int_a^b |g(x)|^2 dx},$$

and

$$\int_a^b \left| \sum_{i=1}^n f_i(x) \right|^2 dx = \sum_{i=1}^n \int_a^b |f_i(x)|^2 dx$$
$$when \quad \int_a^b f_i(x)\overline{f_j(x)}dx = 0 \quad for\ i \neq j.$$

Given the definitions of orthogonality and orthonormality, we want to know if there is an infinite-dimensional analogue of the following theorem:

Theorem 2.4. *Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_k\} \subset V$ be an orthonormal set of vectors, with V a k-dimensional vector space (and inner product space). Then, for any $\boldsymbol{a} \in V$, the following formula holds:*

$$\boldsymbol{a} = \langle \boldsymbol{a}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \langle \boldsymbol{a}, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2 + ... + \langle \boldsymbol{a}, \boldsymbol{u}_k \rangle \boldsymbol{u}_k.$$

Moreover,

$$\|\boldsymbol{a}\|^2 = \left|\langle \boldsymbol{a}, \boldsymbol{u}_1 \rangle\right|^2 + \left|\langle \boldsymbol{a}, \boldsymbol{u}_2 \rangle\right|^2 + ... + \left|\langle \boldsymbol{a}, \boldsymbol{u}_k \rangle\right|^2.$$

That is, we aim to study if there is a way to write any function $f \in L^2[a,b]$ as $\sum \langle f, \phi_n \rangle \phi_n$, where $\{\phi_n\}_{n\in\mathbb{N}}$ is an orthonormal set of functions. The main problem, as anticipated earlier, is the dimension of $L^2[a,b]$ as a vector space, which is infinite. This fact has two main implications:

1. we cannot tell if there are "enough" functions in $\{\phi_n\}_{n\in\mathbb{N}}$ by just counting the linear independent ones, because there can be an infinite number;

2. $\sum \langle f, \phi_n \rangle \phi_n$ is an infinite series, therefore we have to analyze its convergence.

2.2 Norm convergence and orthonormal bases

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions in $L^2[a,b]$. We say that $f_n \to f$ **in norm** if $\|f_n - f\| \to 0$, or more precisely,

$$f_n \xrightarrow[n\to+\infty]{} f \ in\ norm \quad \Longleftrightarrow \quad \int_a^b |f_n(x) - f(x)|^2 dx \xrightarrow[n\to+\infty]{} 0.$$

Definition 2.5. *A sequence $\{\boldsymbol{a}_n\}_{n\in\mathbb{N}}$ of vectors is called a* ***Cauchy sequence*** *if $\|\boldsymbol{a}_m - \boldsymbol{a}_n\| \to 0$ as $m, n \to +\infty$. This means that the terms of the sequence get progressively closer as one goes further out in the sequence. A normed vector space V is said to be* ***complete*** *if every Cauchy sequence in V also has its limit in V.*

It is only natural to ask ourselves if $L^2[a,b]$ is **complete**. The answer is "yes", as we are about to witness with the next theorem.

Theorem 2.6. *The followings stand:*

1. $L^2[a,b]$ *is complete with respect to the convergence in norm;*

2. $\forall f \in L^2[a,b]$ *there exists a sequence of continuous functions* $\{f_n\}_{n\in\mathbb{N}}$ *such that* $f_n \to f$ *in norm; moreover, the functions* f_n *can be taken to be restrictions to the interval* $[a,b]$ *of (b-a)-periodic functions in* $\mathcal{C}^\infty(\mathbb{R})$.

Now onto the main question: under what circumstances do infinite series converge? To answer this question we must first prove a lemma:

Lemma 2.7. *For any* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *in* V*, where* V *is an inner product space,*

$$\|\boldsymbol{a}+\boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + 2\Re\langle \boldsymbol{a}, \boldsymbol{b}\rangle + \|\boldsymbol{b}\|^2,$$

where $\Re$ *is the function that associates to any complex number its real part.*

Proof.

$$\begin{aligned}
\|\mathbf{a}+\mathbf{b}\|^2 &= \langle \mathbf{a}+\mathbf{b}, \mathbf{a}+\mathbf{b}\rangle \\
&= \langle \mathbf{a}, \mathbf{a}\rangle + \langle \mathbf{a}, \mathbf{b}\rangle + \langle \mathbf{b}, \mathbf{a}\rangle + \langle \mathbf{b}, \mathbf{b}\rangle \\
&= \langle \mathbf{a}, \mathbf{a}\rangle + \langle \mathbf{a}, \mathbf{b}\rangle + \overline{\langle \mathbf{a}, \mathbf{b}\rangle} + \langle \mathbf{b}, \mathbf{b}\rangle \\
&= \|\mathbf{a}\|^2 + 2\Re\langle \mathbf{a}, \mathbf{b}\rangle + \|\mathbf{b}\|^2
\end{aligned}$$

□

Theorem 2.8 (Bessel's Inequality)**.** *Let* $f \in L^2[a,b]$ *and let* $\{\phi_n\}_{n\in\mathbb{N}} \subset L^2[a,b]$ *be an orthonormal set, then*

$$\sum_{n=1}^{+\infty} \left|\langle f, \phi_n\rangle\right|^2 \leq \|f\|^2$$

Proof. First we notice that

$$\left\langle f, \langle f, \phi_n\rangle \phi_n\right\rangle = \overline{\langle f, \phi_n\rangle}\langle f, \phi_n\rangle = \left|\langle f, \phi_n\rangle\right|^2,$$

and thanks to the Pythagorean theorem we know that

$$\left\|\sum_{n=1}^{N} \langle f, \phi_n\rangle \phi_n\right\|^2 = \sum_{n=1}^{N} \left|\langle f, \phi_n\rangle\right|^2.$$

Hence, for any $N \in \mathbb{N}$, by lemma 2.7,

$$
\begin{aligned}
0 &\leq \left\| f - \sum_{n=1}^{N} \langle f, \phi_n \rangle \phi_n \right\|^2 \\
&= \|f\|^2 - 2\Re \left\langle f, \sum_{n=1}^{N} \langle f, \phi_n \rangle \phi_n \right\rangle + \left\| \sum_{n=1}^{N} \langle f, \phi_n \rangle \phi_n \right\|^2 \\
&= \|f\|^2 - 2\sum_{n=1}^{N} \big| \langle f, \phi_n \rangle \big|^2 + \sum_{n=1}^{N} \big| \langle f, \phi_n \rangle \big|^2 \\
&= \|f\|^2 - \sum_{n=1}^{N} \big| \langle f, \phi_n \rangle \big|^2,
\end{aligned}
$$

and by letting $N \to +\infty$ we prove the statement. □

This theorem tells us, that whenever $f \in L^2[a,b]$, the series $\sum |\langle f, \phi_n \rangle|^2$, of the coefficients of $\sum \langle f, \phi_n \rangle \phi_n$, converges.

Now our main concern is to find out if, given any $f \in L^2[a,b]$ and an orthonormal set $\{\phi_n\}_{n \in \mathbb{N}} \subset L^2[a,b]$, we are allowed to say that

$$
f = \sum_{n=1}^{+\infty} \langle f, \phi_n \rangle \phi_n. \tag{2.10}
$$

But before that, we have got to check that the right-hand side of the equation (2.10) does make sense. Hence we state and prove the following:

Lemma 2.9. *Let $f \in L^2[a,b]$ and let $\{\phi_n\}_{n \in \mathbb{N}} \subset L^2[a,b]$ be an orthonormal set. Then the series $\sum \langle f, \phi_n \rangle \phi_n$ converges in norm and*

$$
\left\| \sum_{n=1}^{+\infty} \langle f, \phi_n \rangle \phi_n \right\| \leq \|f\|
$$

Proof. Thanks to Bessel's inequality we know that $\sum \big| \langle f, \phi_n \rangle \big|^2$ converges, so, by the Pythagorean theorem,

$$
\left\| \sum_{i=n}^{m} \langle f, \phi_i \rangle \phi_i \right\|^2 = \sum_{i=n}^{m} \big| \langle f, \phi_i \rangle \big|^2 \xrightarrow[m,n \to +\infty]{} 0.
$$

Hence $\sum \langle f, \phi_n \rangle \phi_n$ is Cauchy and thus convergent, thanks to the completeness of $L^2[a,b]$. Through another use of the Pythagorean theorem and Bessel's inequality, we prove the statement, because

$$\begin{aligned}\left\|\sum_{n=1}^{+\infty}\langle f,\phi_n\rangle\phi_n\right\|^2 &= \lim_{N\to+\infty}\left\|\sum_{n=1}^{N}\langle f,\phi_n\rangle\phi_n\right\|^2\\ &= \lim_{N\to+\infty}\sum_{n=1}^{N}\left|\langle f,\phi_n\rangle\right|^2\\ &= \sum_{n=1}^{+\infty}\left|\langle f,\phi_n\rangle\right|^2 \le \|f\|^2\end{aligned}$$

□

The holding of (2.10) for all functions $f \in L^2[a,b]$ implies a couple of facts:

1. if $\langle f,\phi_n\rangle = 0\ \forall n \in \mathbb{N}$ then $f = 0$; one can read into this that the set $\{\phi_n\}_{n\in\mathbb{N}}$ has to be "complete" in some way, that there have to be "enough" orthonormal vectors belonging to this collection;
2. if the Pythagorean theorem extends to infinite sums, then the Bessel's inequality becomes an equality.

Bearing these thoughts in mind, we enunciate the following:

Theorem 2.10. *Let $\{\phi_n\}_{n\in\mathbb{N}} \subset L^2[a,b]$ be an orthonormal set. The followings are equivalent:*

1. *if $\langle f,\phi_n\rangle = 0\ \forall n \in \mathbb{N}$ then $f = 0$;*
2. *$\forall f \in L^2[a,b]$ we have that the series $\sum\langle f,\phi_n\rangle\phi_n$ converges to f in norm;*
3. *$\forall f \in L^2[a,b]$ we have **Parseval's equation**:*

$$\|f\|^2 = \sum_{n=1}^{+\infty}\left|\langle f,\phi_n\rangle\right|^2$$

Proof. (1)⇒(2): thanks to lemma 2.9, we know that $f \in L^2[a,b]$ implies that $\sum\langle f,\phi_n\rangle\phi_n$ converges. We must show that it conveges to f. Let $g = f - \sum\langle f,\phi_n\rangle\phi_n$, then for all $m \in \mathbb{N}$,

$$\langle g,\phi_m\rangle = \langle f,\phi_m\rangle - \sum_{n=1}^{+\infty}\langle f,\phi_n\rangle\langle\phi_n,\phi_m\rangle = \langle f,\phi_m\rangle - \langle f,\phi_m\rangle = 0.$$

Therefore if (1) holds, $g = 0$.

(2)⇒(3): if $f = \sum\langle f,\phi_n\rangle\phi_n$, then because of the Pythagorean theorem,

$$\|f\|^2 = \lim_{N\to+\infty}\left\|\sum_{n=1}^{N}\langle f,\phi_n\rangle\phi_n\right\|^2 = \lim_{N\to+\infty}\sum_{n=1}^{N}\left|\langle f,\phi_n\rangle\right|^2 = \sum_{n=1}^{+\infty}\left|\langle f,\phi_n\rangle\right|^2.$$

(3)⇒(1): if (3) holds and $\langle f, \phi_n \rangle = 0$ for all $n \in \mathbb{N}$, then $\|f\| = 0$, which means that $f = 0$. □

Definition 2.11. *If an orthonormal set $\{\phi_n\}_{n\in\mathbb{N}} \subset L^2[a,b]$ satisfies any of the properties (and therefore all of them) enumerated in theorem* (2.10)*, then it is called a **complete orthonormal set** or an **orthonormal basis**. Moreover, the coefficients $\langle f, \phi_n \rangle$ are said to be the (generalized) **Fourier coefficients** of the series $\sum \langle f, \phi_n \rangle \phi_n$, which is, not very surprisingly, said to be the (generalized) **Fourier series**.*

The only difference between orthogonal and orthonormal sets is that one has unit vectors as elements, while the other does not. Hence sometimes it is more manageable to require a set of vectors $\{\psi_n\}_{n\in\mathbb{N}}$ to be orthogonal, instead of orthonormal. Afterwards, one can obtain an orthonormal set by taking each vector that belongs to the orthogonal one and normalizing it.

In conclusion, all we need to do now, is to actually find explicit orthogonal bases, which we will do in the next chapter, through the definition and resolution of **regular Sturm-Liouville problems**.

Chapter 3

Regular Sturm-Liouville problems

In this chapter we will investigate the properties of a large class of boundary value problems on an interval $[a, b]$, whose peculiarity is the fact that their solutions form orthogonal bases for $L^2[a, b]$. In order to do this, we will study the meaning of adjoint operators and some of their properties. We are aiming for an equivalent of the spectral theorem for linear differential operators working on the space $L^2[a, b]$. Once we find it, we will be able to build the most important orthonormal bases with it too. Finally, we study "how well" can these Fourier series converge, relatively to the regularity of the functions they are supposed to be converging to.

3.1 Adjoint operators and the Lagrange's identity

Definition 3.1. *Let $S : \mathcal{D}_S \to L^2[a, b]$ and $T : \mathcal{D}_T \to L^2[a, b]$ be linear operators, where both $\mathcal{D}_S$ and $\mathcal{D}_T$ are subspaces of $L^2[a, b]$. We say that S and T are* ***adjoint*** *to each other if, for all $f \in \mathcal{D}_S$ and $g \in \mathcal{D}_T$ we have*

$$\langle S(f), g\rangle = \langle f, T(g)\rangle.$$

We also say that S is ***self-adjoint*** *if, for all $f, g \in \mathcal{D}_S$,*

$$\langle S(f), g\rangle = \langle f, S(g)\rangle.$$

We shall now consider a linear differential operator L such that

$$L(f) = rf'' + qf' + pf,$$

where $r, p, q \in \mathcal{C}^2[a, b]$ and $r(x) \neq 0$ for all $x \in [a, b]$ (which means that either $r > 0$ or $r < 0$ on $[a, b]$). For the sake of our goal we do not have to linger on the domain's choice of the operators we are going to use, so from now on we will just take $\mathcal{C}^2[a, b]$.

Now let us investigate the identity of the adjoint of this operator L. It is only natural to start from the definition and then take it from there. It is easy to notice, that the product $\langle L(f), g\rangle$ can be studied term by term, thanks to the linearity of both the inner product (with respect to the first variable) and the operator L, so we will do just that. We want to move the derivatives from f to g, so the key now is the integration by parts. Therefore, regarding the second order term, we obtain

$$\int_a^b (rf'')\overline{g}dx = -\int_a^b f'(r\overline{g})'dx + rf'\overline{g}\Big|_a^b = \int_a^b f(r\overline{g})''dx + \Big[rf'\overline{g} - f(r\overline{g})'\Big]_a^b,$$

while for the first order term we get

$$\int_a^b (qf')\overline{g}dx = -\int_a^b f(q\overline{g})'dx + qf\overline{g}\Big|_a^b.$$

Hence

$$\begin{aligned}\langle L(f), g\rangle &= \int_a^b (rf'' + qf' + pf)\overline{g}dx \\ &= \int_a^b f\big((r\overline{g})'' - (q\overline{g})' + p\overline{g}\big)dx + \Big[rf'\overline{g} - f(r\overline{g})' + qf\overline{g}\Big]_a^b \\ &= \langle f, L^\star(g)\rangle + \Big[r(f'\overline{g} - f\overline{g}') + (q - r')f\overline{g}\Big],\end{aligned}$$

where $L^\star$ is called the **formal adjoint** of L and is defined by

$$L^\star(g) = (rg)'' - (qg)' + pg = rg'' + (2r' - q)g' + (r'' - q' + p)g.$$

L is said to be **formally self-adjoint** when $L = L^\star$, which in this case, one can see it by comparing the coefficients of L and $L^\star$, translates to $2r' - q = q$ and $r'' - q = 0$, thus providing the condition $q = r'$; so L assumes the following shape

$$L(f) = rf'' + r'f' + pf = (rf')' + pf.$$

We can notice that when $q = r'$, the second boundary term vanishes. With that, we proved the following theorem.

Theorem 3.2 (Lagrange's Identity)**.** *If L is formally self-adjoint, then the following holds:*

$$\langle L(f), g\rangle = \langle f, L(g)\rangle + \Big[r(f'\overline{g} - f\overline{g}')\Big]_a^b.$$

This identity tells us exactly where to look at for our next step. The difference between the formal self-adjoint and the "actual" self-adjoint is $\Big[r(f'\overline{g} - f\overline{g}')\Big]_a^b$, which can be eliminated by imposing suitable boundary conditions. Usually, for a second-order differential operator, it is befitting to have two independent boundary conditions, which in this case take the form

$$\begin{aligned} B_1(f) &= \alpha_1 f_1(a) + \alpha_1' f_1'(a) + \beta_1 f_1(b) + \beta_1' f_1'(b) = 0 \\ B_2(f) &= \alpha_2 f_2(a) + \alpha_2' f_2'(a) + \beta_2 f_2(b) + \beta_2' f_2'(b) = 0 \end{aligned} \tag{3.1}$$

These two boundary conditions are said to be **self-adjoint** (with respect to the operator L) if, for all $f, g \in \mathcal{C}^2[a, b]$ that satisfy (3.1),

$$\Big[r(f'\overline{g} - f\overline{g}')\Big]_a^b = 0.$$

Example 3.3. Here are two important examples of self-adjoint boundary conditions (with respect to L):

1. let f and g be two functions that fit some separated boundary conditions. Then, for example at a (it is analogous at b) we have

$$\alpha f(a) + \alpha' f'(a) = 0 \quad \text{and} \quad \alpha g(a) + \alpha' g'(a) = 0, \tag{3.2}$$

   then $r(f'\overline{g} - f\overline{g}') = 0$ at $x = a$. As a matter of fact, when $\alpha' = 0$ the case is trivial, since it means that $f(a) = g(a) = 0$, whereas if $\alpha \neq 0$ we are able to rewrite (3.2) as

$$f'(a) = \frac{\alpha}{\alpha'} f(a), \quad g'(a) = \frac{\alpha}{\alpha'} g(a),$$

   so that

$$r(a)\big(f'(a)\overline{g(a)} - f(a)\overline{g'(a)}\big) = \frac{\alpha}{\alpha'} r(a)\big(f(a)\overline{g(a)} - f(a)\overline{g(a)}\big) = 0.$$

2. let f and g be two functions that fit some periodic boundary conditions ($f(a) = f(b)$, $f'(a) = f'(b)$, likewise for g). Then,

$$\begin{aligned} \Big[r(f'\overline{g} - f\overline{g}')\Big]_a^b &= r(b)\big(f'(b)\overline{g(b)} - f(b)\overline{g'(b)}\big) - r(a)\big(f'(a)\overline{g(a)} - f(a)\overline{g'(a)}\big) \\ &= \big(r(b) - r(a)\big)\big(f'(a)\overline{g(a)} - f(a)\overline{g'(a)}\big) \end{aligned}$$

   which means that there is the need of the supplementary condition $r(a) = r(b)$ in order for the periodic boundary conditions to be actually self-adjoint.

3.2 Regular Sturm-Liouville problems

Definition 3.4. *A* ***regular Sturm-Liouville problem*** *on the interval* $[a, b]$ *is defined by the following information:*

1. *L such that $L(f) = (rf')' + pf$, a linear differential formally self-adjoint operator, where $r, r', p \in \mathcal{C}\big([a, b], \mathbb{R}\big)$ and $r > 0$ on $[a, b]$;*

2. *$B_1(f) = 0$ and $B_2(f) = 0$ a set of homogeneous boundary conditions that are self-adjoint with respect to the operator L;*

3. *$w \in \mathcal{C}\big([a, b], (0, +\infty)\big)$.*

The object himself is the boundary value problem

$$\begin{cases} L(f) + \lambda w f = 0 \Longleftrightarrow \big(r(x)f'(x)\big)' + p(x)f(x) + \lambda w(x)f(x) = 0 \\ B_1(f) = B_2(f) = 0, \end{cases}$$

with λ an arbitrary constant.

Remark 3.5. Since earlier we assumed r to be non-vanishing on $[a, b]$, it has to be either $r > 0$ or $r < 0$. If it is the latter, then we replace r, p and λ with $-r, -p$ and $-\lambda$, so the problem remains unchanged.

Remark 3.6. The function $w \in \mathcal{C}\big([a, b], (0, +\infty)\big)$ is called **weight function**, and it is associated to

$$L^2_w[a, b] = \left\{ f : \int_a^b |f(x)|^2 w(x) dx < +\infty \right\},$$

which is the so-called **weighted L^2 space**, a generalization of L^2 spaces. It is a normed inner product space too, thanks to the definitions

$$\langle f, g\rangle_w = \int_a^b f(x)\overline{g(x)}w(x)dx = \langle wf, g\rangle = \langle f, wg\rangle$$

and

$$\|f\|_w = \left(\int_a^b |f(x)|^2 w(x) dx \right)^{1/2}.$$

For most values of λ, the only solution of a given regular Sturm-Liouville problem is $f(x) \equiv 0$. Eventually, there are non-trivial solutions: in that case λ is called **eigenvalue** and its corresponding solution is called **eigenfunction** of the given regular Sturm-Liouville problem. However they are not relative to the operator L, but instead to the operator $M(f) = \frac{1}{w}L(f)$. Of course if f and g satisfy the given regular Sturm-Liouville problem, then so does any of their linear combination, giving sense to the notion of

eigenspace (relative to the given λ of course). The next theorem will sum up the main properties of eigenvalues and eigenfunctions relative to a given regular Sturm-Liouville problem:

Theorem 3.7. *Let a regular Sturm-Liouville problem be given. Then:*

1. *all eigenvalues are real;*

2. *if f and g are eigenfunctions with eigenvalues λ and μ respectively, with $\lambda \neq \mu$, then they are orthogonal in $L^2_w[a,b]$, i.e.*

$$\langle f, g\rangle_w = \int_a^b f(x)\overline{g(x)}w(x)dx = 0;$$

3. *the eigenspace relative to any eigenvalue λ is at most 2-dimensional. If the boundary conditions are separated, the eigenspace is always 1-dimensional.*

Proof. (1): given λ an eigenvalue and f an eigenfunction with eigenvalue λ, we know that

$$\lambda\|f\|_w^2 = \langle \lambda wf, f\rangle = -\langle L(f), f\rangle = -\langle f, L(f)\rangle = \langle f, \lambda wf\rangle = \overline{\lambda}\langle f, wf\rangle = \overline{\lambda}\|f\|_w^2,$$

and since $\|f\|_w^2 > 0$, we obtain $\lambda = \overline{\lambda}$, which means that λ is real.

(2): assuming that $L(f) + \lambda wf = 0$ and that $L(g) + \mu wg = 0$, with f and g non-zero and $\lambda, \mu \in \mathbb{R}$ (thanks to the previous point), we can see that

$$\lambda\langle f, g\rangle_w = \langle \lambda wf, g\rangle = -\langle L(f), g\rangle = -\langle f, L(g)\rangle = \langle f, \mu wg\rangle = \mu\langle f, g\rangle_w.$$

Hence if $\lambda \neq \mu$ then it must be that $\langle f, g\rangle_w = 0$.

(3): the idea is that for any constants c_1, c_2 there exists a unique solution of $L(f) + \lambda wf = 0$ satisfying the initial conditions $f(a) = c_1$ and $f'(a) = c_2$. That is, a solution is determined by two constants, ergo the space of all solutions is 2-dimensional. Hence the space of the solutions that fit the given boundary conditions is at most 2-dimensional. If the given boundary conditions are separated, one of them will be like $\alpha f(a) + \alpha' f'(a) = 0$, which imposes the relation $\alpha c_1 + \alpha' c_2 = 0$ between c_1 and c_2, reducing the dimension of the solution space to one. For most of the eigenvalues λ, the other boundary condition will make the dimension drop to zero. □

The next theorem guarantees instead, that eigenfunctions of regular Sturm-Liouville problems do exist, and that there are enough of them to form orthonormal bases.

Theorem 3.8. *Let a regular Sturm-Liouville problem on $[a,b]$ be given. Then there is an orthonormal basis $\{\phi_n\}_{n\in\mathbb{N}}$ of eigenfunctions of $L^2_w[a,b]$. If λ_n is the eigenvalue with respect to ϕ_n, then $\lambda_n \xrightarrow[n\to+\infty]{} +\infty$. Furthermore, if $f \in \mathcal{C}^2[a,b]$ and fits the boundary conditions, then $\sum\langle f,\phi_n\rangle\phi_n$ converges uniformly to f.*

With these tools at hand we can finally prove the following:

Theorem 3.9. *The sets*

$$\left\{e^{i\frac{n\pi x}{L}}\right\}_{n\in\mathbb{Z}} \quad \text{and} \quad \left\{cos\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}_0} \cup \left\{sin\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}}$$

are orthogonal bases for $L^2[-L,L]$, while

$$\left\{cos\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}_0} \quad \text{and} \quad \left\{sin\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}}$$

are orthogonal bases for $L^2[0,L]$.

Proof. To prove that those sets are actual orthogonal bases we just have to show the regular Sturm-Liouville problem they come from, meaning:

1. by solving the problem

$$\begin{cases} u''(x) = \lambda u(x) \\ u(0) = u(L) = 0 \end{cases} \tag{3.3}$$

   we obtain

$$\left\{sin\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}},$$

   an orthogonal basis of $L^2[0,L]$;

2. by solving the problem

$$\begin{cases} u''(x) = \lambda u(x) \\ u'(0) = u'(L) = 0 \end{cases}$$

   we obtain

$$\left\{cos\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}_0},$$

   an orthogonal basis of $L^2[0,L]$;

3. by solving the problem

$$\begin{cases} u''(x) = \lambda u(x) \\ u(-L) = u(L) = 0 \\ u'(-L) = u'(L) = 0 \end{cases}$$

   we obtain

$$\left\{e^{i\frac{n\pi x}{L}}\right\}_{n\in\mathbb{Z}},$$

   an orthogonal basis of $L^2[-L,L]$.

Let us consider the first case. One can easily write the general integral of the equation $u''(x) = \lambda u(x)$ for any λ, however one will end up with null integration constants, most of the time, in the attempt of satisfying the boundary conditions. More practically:

1. if $\lambda = 0$, then $u(x) = c_1 x + c_2$, which satisfies the boundary conditions $u(0) = u(L) = 0$ only if $c_1 = c_2 = 0$;

2. if $\lambda > 0$, then $u(x) = c_1 e^{\lambda x} + c_2 e^{-\lambda x}$, which satisfies the boundary conditions $u(0) = u(L) = 0$ only if $c_1 = c_2 = 0$;

3. if $\lambda < 0$, then $u(x) = c_1 cos(\sqrt{-\lambda} x) + c_2 sin(\sqrt{-\lambda} x)$, which satisfies the boundary conditions $u(0) = u(L) = 0$ when $c_1 = 0$ and for all c_2, provided that $sin(\sqrt{-\lambda} L) = 0$, that is

$$\begin{aligned} \sqrt{-\lambda} L &= n\pi \quad \text{for } n \in \mathbb{N} \\ \lambda &= -\frac{n^2\pi^2}{L^2} \quad \text{for } n \in \mathbb{N}. \end{aligned}$$

Hence the eigenvalues and eigenfunctions of problem (3.3) are, respectively,

$$\lambda_n = -\frac{n^2\pi^2}{L^2} \quad \text{and} \quad u_n(x) = c_n sin\left(\frac{n\pi x}{L}\right) \quad \text{for } n \in \mathbb{N}.$$

The other cases are similar.

In order to obtain orthonormal bases for $L^2[-L, L]$ and $L^2[0, L]$ we need the following normalizing constants:

1. $\sqrt{\frac{1}{2L}}$ for $\left\{e^{i\frac{n\pi x}{L}}\right\}_{n\in\mathbb{Z}}$ on $[-L, L]$;

2. $\sqrt{\frac{1}{L}}$ for $\left\{cos\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}_0} \cup \left\{sin\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}}$ on $[-L, L]$ (for $n = 0$ we need $\sqrt{\frac{1}{2L}}$);

3. $\sqrt{\frac{2}{L}}$ for $\left\{cos\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}_0}$ and $\left\{sin\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}}$ on $[0, L]$ (for $n = 0$ we need $\sqrt{\frac{1}{L}}$).

□

Now that we have orthonormal bases for $L^2[-L, L]$ and $L^2[0, L]$, we can use the following proposition, which we shall state without proof, to transform them into orthonormal bases for every other $L^2[\cdot, \cdot]$ spaces. More precisely:

Proposition 3.10. *Let $\{\phi_n\}_{n\in\mathbb{N}}$ be an orthonormal basis for $L^2[a, b]$, let $c > 0$ and let $d \in \mathbb{R}$. Then if we define $\psi_n(x) = c^{1/2}\phi_n(cx + d)$, the set $\{\psi_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for $L^2[\frac{a-d}{c}, \frac{b-d}{c}]$.*

3.3 Convergence and derivation of Fourier series

Using the bases that we have just extrapolated from regular Sturm-Liouville problems we are going to find the classic Fourier coefficients. Afterwards, with those formulae at our disposition, we can go further in the study of convergence of Fourier series.

More precisely, by using the orthonormal base of $L^2[-L, L]$

$$\left\{\sqrt{\frac{1}{L}}cos\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}_0} \cup \left\{\sqrt{\frac{1}{L}}sin\left(\frac{n\pi x}{L}\right)\right\}_{n\in\mathbb{N}},$$

we obtain, for every $n \in \mathbb{N}$,

$$\begin{aligned}\langle f, \sqrt{\frac{1}{L}}cos\left(\frac{n\pi x}{L}\right)\rangle\sqrt{\frac{1}{L}}cos\left(\frac{n\pi x}{L}\right) &= \frac{1}{L}\langle f, cos\left(\frac{n\pi x}{L}\right)\rangle cos\left(\frac{n\pi x}{L}\right)\\ &= \frac{1}{L}\int_{-L}^{L} f(x)cos\left(\frac{n\pi x}{L}\right) dx\ cos\left(\frac{n\pi x}{L}\right),\end{aligned}$$

and

$$\begin{aligned}\langle f, \sqrt{\frac{1}{L}}sin\left(\frac{n\pi x}{L}\right)\rangle\sqrt{\frac{1}{L}}sin\left(\frac{n\pi x}{L}\right) &= \frac{1}{L}\langle f, sin\left(\frac{n\pi x}{L}\right)\rangle sin\left(\frac{n\pi x}{L}\right)\\ &= \frac{1}{L}\int_{-L}^{L} f(x)sin\left(\frac{n\pi x}{L}\right) dx\ sin\left(\frac{n\pi x}{L}\right),\end{aligned}$$

therefore

$$\begin{aligned}a_n &= \frac{1}{L}\int_{-L}^{L} f(x)cos\left(\frac{n\pi x}{L}\right) dx\\ b_n &= \frac{1}{L}\int_{-L}^{L} f(x)sin\left(\frac{n\pi x}{L}\right) dx,\end{aligned} \tag{3.4}$$

although for $n = 0$ the coefficient relative to the sine function vanishes, while the coefficient relative to the cosine becomes

$$\frac{1}{2L}\int_{-L}^{L} f(x)cos\left(\frac{n\pi x}{L}\right) dx = \frac{1}{2}a_0.$$

In the same way, by using the orthonormal base of $L^2[-L, L]$

$$\left\{\sqrt{\frac{1}{2L}}e^{i\frac{n\pi x}{L}}\right\}_{n\in\mathbb{Z}}$$

we obtain, for every $n \in \mathbb{N}$,

$$\begin{aligned}\langle f, \sqrt{\frac{1}{2L}}e^{i\frac{n\pi x}{L}}\rangle\sqrt{\frac{1}{2L}}e^{i\frac{n\pi x}{L}} &= \frac{1}{2L}\langle f, e^{i\frac{n\pi x}{L}}\rangle e^{i\frac{n\pi x}{L}}\\ &= \frac{1}{2L}\int_{-L}^{L} f(x)e^{i\frac{n\pi x}{L}} dx\ e^{i\frac{n\pi x}{L}},\end{aligned}$$

therefore

$$c_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{i\frac{n\pi x}{L}}dx.$$

Both the definition and the lemma we are about to state will be of good use to us in the next chapter.

Definition 3.11. *Let f be defined on $[0, L]$. Then:*

1. *the extension of f on the interval $[-L, L]$ defined by*

$$f_{even}(-x) = f(x)$$

for $x \in [0, L]$ will be called ***even extension****;*

2. *the extension of f on the interval $[-L, L]$ defined by*

$$f_{even}(-x) = -f(x)$$

for $x \in [0, L]$ will be called ***odd extension****.*

Lemma 3.12. *If we consider Fourier coefficients of the sine and cosine series as in 3.4, then:*

1. *if f is even,*

$$a_n = \frac{2}{L}\int_0^L f(x)cos\left(\frac{n\pi x}{L}\right)dx \quad \text{and} \quad b_n = 0;$$

2. *if f is odd,*

$$a_n = 0 \quad \text{and} \quad b_n = \frac{2}{L}\int_0^L f(x)sin\left(\frac{n\pi x}{L}\right)dx.$$

Proof. We need to observe that

$$\int_{-L}^{L} F(x)dx = \begin{cases} 2\int_0^L F(x)dx & \text{if F is even} \\ 0 & \text{if F is odd} \\ \int_{-L}^{L} F(x)dx & \text{otherwise.} \end{cases}$$

Therefore if f is even, then $f(x)cos(x)$ is even while $f(x)sin(x)$ is odd; but if f is odd, then $f(x)cos(x)$ is odd while $f(x)sin(x)$ is even. Hence, the statement holds. □

Definition 3.13. *$\mathcal{PC}[a, b]$ is the space of* ***piecewise continuous*** *functions on $[a, b]$: f belongs in this set if it is continuous on $[a, b]$ except maybe at finitely many points $x_1, x_2, ..., x_k$, and if, at each of these points, both the left-hand and the right-hand limits exist and are finite, which means that, for $1 \leq j \leq k$,*

$$\exists f(x_j-) = \lim_{h\to 0^+} f(x_j - h) \quad \text{and} \quad \exists f(x_j+) = \lim_{h\to 0^+} f(x_j + h).$$

Definition 3.14. *$\mathcal{PS}[a,b]$ is the space of* ***piecewise smooth*** *functions on $[a,b]$: f belongs in this set if $f \in \mathcal{PC}[a,b]$, if f' exists and is continuous on (a,b) except maybe at finitely many points $x_1, x_2, ..., x_K$ (which of course include all discontinuities of f), and if $f'(a+), f'(b-), f'(x_j-), f'(x_j+)$ exist for all $1 \leq j \leq K$.*

Definition 3.15. *A function f is said to be in $\mathcal{PC}(\mathbb{R})$ (or $\mathcal{PS}(\mathbb{R})$) if $f \in \mathcal{PC}[a,b]$ (or $f \in \mathcal{PS}[a,b]$) for any bounded interval $[a,b]$.*

Remark 3.16. Thanks to Weierstrass theorem, it is obvious that

$$\mathcal{PS}[a,b] \subset \mathcal{PC}[a,b] \subset L^2[a,b],$$

hence everything we have proven until now in $L^2[a,b]$ will hold in $\mathcal{PC}[a,b]$ and in $\mathcal{PS}[a,b]$ too.

Let us now define the partial sums of the Fourier series of f:

$$S_N^f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{N}\left(a_n cos\left(\frac{n\pi x}{L}\right) + b_n sin\left(\frac{n\pi x}{L}\right)\right) = \sum_{n=-N}^{N} c_n e^{i\frac{n\pi x}{L}}. \tag{3.5}$$

By this definition, it is possible to prove the following:

Theorem 3.17. *If $f \in \mathcal{PS}(\mathbb{R})$ is 2L-periodic and S_N^f is defined as in (3.5), then*

$$\lim_{N\to+\infty} S_N^f(x) = \frac{1}{2}\big(f(x-) + f(x+)\big)$$

for every $x \in \mathbb{R}$. In particular,

$$\lim_{N\to+\infty} S_N^f(x) = f(x)$$

for every x at which f is continuous. In other words, if $f \in \mathcal{PS}(\mathbb{R})$ then its Fourier series pointwise converges to it almost everywhere.

This actually tells us that if we take any $f, g \in \mathcal{PS}(\mathbb{R})$, redefine them as $\frac{1}{2}\big(f(x-) + f(x+)\big)$ and $\frac{1}{2}\big(g(x-) + g(x+)\big)$ at their discontinuities, and observe that they have the same Fourier series, then they are the same function. More precisely, this theorem implies that, given the right circumstances, Fourier series are **unique**.

Having said when Fourier series converge pointwise to their respective functions, there is one more important result left, which is the following:

Theorem 3.18. *Let $f \in \mathcal{C}(\mathbb{R}) \cap \mathcal{PS}(\mathbb{R}))$ be 2L-periodic. Then its Fourier series converges to it uniformly and absolutely on $\mathbb{R}$.*

To sum up all convergence results about Fourier series (associated, of course, to a $2L$-periodic function f), we have:

1. norm convergence when $f \in L^2$;
2. pointwise convergence when $f \in \mathcal{PS}$;
3. uniform and absolute convergence when $f \in \mathcal{C} \cap \mathcal{PS}$.

In the next chapter, we will need to know under what conditions we are able to derivate Fourier series term by term, and that's what the next theorems are about.

Theorem 3.19. *Let $f \in \mathcal{C}[-L, L] \cap \mathcal{PS}[-L, L]$ be 2L-periodic. If a_n, b_n, c_n are the Fourier coefficients of f and a'_n, b'_n, c'_n are the Fourier coefficients of f', then*

$$a'_n = \frac{n\pi}{L} b_n, \quad b'_n = -\frac{n\pi}{L} a_n, \quad c'_n = i\frac{n\pi}{L} c_n.$$

Proof. To prove this, we merely need to integrate by parts; that is, taking c'_n as an example

$$\begin{aligned} c'_n &= \frac{1}{2L} \int_{-L}^{L} f'(x) e^{-i\frac{n\pi x}{L}} dx \\ &= \frac{1}{2L} f(x) e^{-i\frac{n\pi x}{L}} \Big|_{-L}^{L} - \frac{1}{2L} \int_{-L}^{L} f(x) \left(-i\frac{n\pi}{L} e^{-i\frac{n\pi x}{L}} \right) dx, \end{aligned}$$

where

$$\frac{1}{2L} f(x) e^{-i\frac{n\pi x}{L}} \Big|_{-L}^{L} = 0$$

because $f(L) = f(-L)$ and $e^{-in\pi} = (-1)^n = e^{in\pi}$. Since the procedure for a'_n and b'_n is completely analogous, the statement holds. $\square$

The theorems 3.17 and 3.19 lead us directly to the next one:

Theorem 3.20. *Let $f \in \mathcal{C}[-L, L] \cap \mathcal{PS}[-L, L]$ be 2L-periodic and let $f' \in \mathcal{PS}[-L, L]$. If*

$$\frac{1}{2} a_0 + \sum_{n \in \mathbb{N}} \left(a_n cos\left(\frac{n\pi x}{L}\right) + b_n sin\left(\frac{n\pi x}{L}\right) \right) = \sum_{n \in \mathbb{Z}} c_n e^{i\frac{n\pi x}{L}}$$

is the Fourier series of f, then the Fourier series of f' is

$$\sum_{n \in \mathbb{N}} \left(\frac{n\pi}{L} b_n cos\left(\frac{n\pi x}{L}\right) - \frac{n\pi}{L} a_n sin\left(\frac{n\pi x}{L}\right) \right) = \sum_{n \in \mathbb{Z}} i\frac{n\pi}{L} c_n e^{i\frac{n\pi x}{L}}$$

for all x where f' exists. In the points x where f' does not exist (left-hand limit different from right-hand limit), the series converges to $\frac{1}{2}\big(f(x-) + f(x+)\big)$.

Proof. Since $f' \in \mathcal{PS}[-L, L]$, for theorem 3.17 we have that

$$\lim_{N\to+\infty} S_N^{f'}(x) = \frac{1}{2}\big(f'(x-) + f'(x+)\big),$$

and since $f \in \mathcal{C}[-L, L] \cap \mathcal{PS}[-L, L]$ is $2L$-periodic we can apply theorem 3.19 and substitute the coefficients, hence the statement holds. □

There is just one item left on the list of all ingredients we shall need for later, which is the following.

Theorem 3.21. *Let f be $2L$-periodic. If $f \in \mathcal{C}^{k-1}[-L, L]$ and $f^{(k-1)} \in \mathcal{PS}[-L, L]$, then the Fourier coefficients are such that*

$$\sum_{n\in\mathbb{N}} |n^k a_n|^2 < +\infty, \quad \sum_{n\in\mathbb{N}} |n^k b_n|^2 < +\infty, \quad \sum_{n\in\mathbb{Z}} |n^k c_n|^2 < +\infty.$$

Conversely, if there exist $M > 0$ and $\alpha > 1$ such that the Fourier coefficients of f satisfy either

$$|a_n| \leq \frac{M}{n^{k+\alpha}} \quad \text{and} \quad |b_n| \leq \frac{M}{n^{k+\alpha}}$$

or

$$|c_n| \leq \frac{M}{|n|^{k+\alpha}},$$

then $f \in \mathcal{C}^k$.

Chapter 4

The 1-dimensional wave equation

In this chapter we will first give an introduction on linear partial differential operators, which will be a fundamental ingredient for the contents of the current chapter. Then we are going to derive the model that approximates the waving of a string, which is fixed ad both ends, by making "reasonable" physical assumptions and translating them into mathematical language. Later, we will solve the partial differential equation through the method of the separation of variables, while paying particular attention to the operators we are using. Finally, we shall consider the global version of the one-dimensional wave equation problem, which shall allow us to derive d'Alembert's formula and investigate its properties.

4.1 Linear partial differential operators

Definition 4.1. *A* ***linear partial differential operator*** *is an operation that transforms a function u of $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ into another function $L(u)$, and its general form is the following:*

$$L(u) = a(\boldsymbol{x})u + \sum_{i=1}^{n} b_i(\boldsymbol{x}) \frac{\partial u}{\partial x_i} + \sum_{i,j=1}^{n} c_{ij}(\boldsymbol{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + ...$$

There can be higher-order terms, but the sum conventionally contains ***finitely many*** *terms. The operator by itself may be written as*

$$L = a(\boldsymbol{x}) + \sum_{i=1}^{n} b_i(\boldsymbol{x}) \frac{\partial}{\partial x_i} + \sum_{i,j=1}^{n} c_{ij}(\boldsymbol{x}) \frac{\partial^2}{\partial x_i \partial x_j} + ...$$

The word "linear" in the definitions refers to the fact that, given $u_1, u_2, ..., u_k$ any adequately regular functions and given $c_1, c_2, ..., c_k$ any constants, then

$$L(c_1 u_1 + c_2 u_2 + ... + c_k u_k) = c_1 L(u_1) + c_2 L(u_2) + ... + c_k L(u_k).$$

Definition 4.2. *A* ***linear partial differential equation*** *is an equation of the form*

$$L(u) = F,$$

where L is any linear partial differential operator, and F is a function of $\boldsymbol{x}$. If $F \equiv 0$ then the equation is called ***homogeneous****, otherwise it is called* ***inhomogeneous****.*

Partial differential equations, such as the one-dimensional wave equation, usually have too many solutions to be able to describe them all explicitly in a reasonable way. Therefore **boundary conditions** come into play: with these one can drastically improve the accuracy of the research of a particular solution. These too present themselves with the form

$$B(u) = f,$$

with B being a linear differential operator and f a function defined on the boundary of the domain of the equation at hand.

Example 4.3. Let us consider the spatial setting in which we will find ourselves, that is the interval $[0, L]$. In the previous chapter, while studying the Lagrange's identity, and again in this chapter, during both the derivation of the model and the solution of the one-dimensional wave equation, we will find some of the most common yet important kind of boundary conditions: the **separated** ones, which are called like that because they concern one endpoint at a time, namely

$$\alpha f(a) + \alpha' f'(a) = 0 \quad \text{and} \quad \beta f(b) + \beta' f'(b) = 0,$$

with $\alpha, \alpha', \beta, \beta' \in \mathbb{R}$, $(\alpha, \alpha') \neq (0, 0)$ and $(\beta, \beta') \neq (0, 0)$.

Another set of commonly used non-separated boundary conditions consists of the **periodic** ones, namely

$$f(a) = f(b), \quad f'(a) = f'(b).$$

Boundary conditions can be homogeneous ($f \equiv 0$) or inhomogeneous too.

The linearity of L and B can be restated in the following way:

Theorem 4.4 (The Superposition Principle)**.** *If $u_1, u_2, ..., u_k$ satisfy the linear partial differential equation $L(u_j) = F_j$ and the boundary conditions $B(u_j) = f_j$ for $1 \leq j \leq k$ and $c_1, c_2, ..., c_k$ are any given constants, then $u = c_1u_1 + c_2u_2 + ... + c_ku_k$ satisfies*

$$\begin{cases} L(u) = c_1F_1 + c_2F_2 + ... + c_kF_k \\ B(u) = c_1f_1 + c_2f_2 + ... + c_kf_k \end{cases}$$

This principle is of great importance, and we will use it in many different situations. Let us see a couple of examples of what it allows us to do:

Example 4.5. 1. Given a boundary problem like

$$\begin{cases} L(u) = F \\ B(u) = f, \end{cases}$$

then the superposition principle allows us to study the solutions of the homogeneous boundary problem

$$\begin{cases} L(u) = 0 \\ B(u) = 0, \end{cases}$$

which is usually easier to handle. This works because if we have just one solution v of the inhomogeneous problem and we want to find another solution u, then $w = u - v$ solves the homogeneous boundary problem, since

$$\begin{cases} L(w) = L(u - v) = L(u) - L(v) = F - F = 0 \\ B(w) = B(u - v) = B(u) - B(v) = f - f = 0. \end{cases}$$

That means that in order to describe any solution of the inhomogeneous boundary problem, it is enough for one to study the solutions of the homogeneous boundary problem, and then just find one solution of the inhomogeneous boundary problem.

2. Considering the same boundary problem as the one from the previous example, we separately study the problems

$$\begin{cases} L(u) = F \\ B(u) = 0, \end{cases}$$

and

$$\begin{cases} L(u) = 0 \\ B(u) = f. \end{cases}$$

If we name u_1 a solution of the first problem and u_2 a solution of the second problem, then $v = u_1 + u_2$ is such that

$$\begin{cases} L(v) = L(u_1 + u_2) = L(u_1) + L(u_2) = F + 0 = F \\ B(v) = B(u_1 + u_2) = B(u_1) + B(u_2) = 0 + f = f. \end{cases}$$

What this means is that, starting from an inhomogeneous boundary problem with inhomogeneous boundary conditions, one can break it down into many problems, supposedly easier ones, and once one has solved them, one knows that the function, resulting from the sum of each solution found, is a solution of the initial problem.

4.2 Derivation of the model

Let us consider a **perfectly flexible** string of length L and linear mass density ρ_0, with the latter being constant when at rest. Being perfectly flexible means that the string will have no resistance to bending. Let $t \geq 0$ represent time, let $x \in [0, L]$ represent the projection on the horizontal axis of each point of the string and let $u(x,t)$ represent the vertical displacement of each point x of the string at every fixed time t. Since we have to deal with the curve

$$\begin{cases} x = x \\ u = u(x,t) \end{cases}$$

with $x \in [0, L]$ and fixed time t, the line element is

$$ds = \sqrt{1 + u_x^2} dx,$$

and since the mass remains constant throughout the whole movement, we can state that

$$\rho ds = \rho_0 dx \Longrightarrow \rho\sqrt{1 + u_x^2} = \rho_0.$$

Let us assume now that the only force in action here is the tension (i.e. no gravity, no air resistance, etc.) and that the only movement is **vertical** with very **small oscillations** (relatively to the length L of the string). We will call $\mathbf{T}(t,x)$ the tension vector that represents the force applied by the right-hand side of the string, with respect to the point $\big(x, u(x,t)\big)$, to the left-hand side of it. Of course $-\mathbf{T}(t,x)$ represents the opposite vector. Both of them are tangent to the string thanks to the hypothesis of perfect flexibility.

At this point, we name $T(t,x)$ the intensity of $\mathbf{T}(t,x)$ and $\alpha = \alpha(t,x)$ the slope of the string at the point $\big(x, u(x,t)\big)$ with respect to the resting position. That simply translates to

$$tg\alpha = u_x. \tag{4.1}$$

If we then consider the interval $[x, x + \Delta x]$ of arbitrary length Δx and impose that the horizontal components of $-\mathbf{T}(t,x)$ and $\mathbf{T}(t, x + \Delta x)$ cancel themselves out, then we have

$$\begin{aligned} T(t, x + \Delta x)cos\big(\alpha(t, x + \Delta x)\big) + T(t,x)cos\big(\pi + \alpha(t,x)\big) = 0 \\ T(t, x + \Delta x)cos\big(\alpha(t, x + \Delta x)\big) - T(t,x)cos\big(\alpha(t,x)\big) = 0. \end{aligned}$$

At this point, one can see that if we divide by Δx and let $\Delta x \to 0$

$$\frac{\partial}{\partial x}\Big(T(t,x)cos\big(\alpha(t,x)\big)\Big) = 0$$

and therefore we have that

$$T(t,x)cos\big(\alpha(t,x)\big) = \tau, \tag{4.2}$$

which tells us that the horizontal component of the tension does not depend on the position to which it is applied. Moreover, if we assume that the tension's intensity is proportional to the length of the string's part generating it, then τ is independent of time too, because the length of the string is constantly L'

Let us now consider the tension's vertical component on the string's section relative to the interval $[x, x+\Delta x]$. Thanks to (4.1) and (4.2), the latter meaning that $T = \tau/cos\alpha$, we can see that

$$\begin{aligned} &T(t,x+\Delta x)sin\big(\alpha(t,x+\Delta x)\big) - T(t,x)sin\big(\alpha(t,x)\big) \\ &= \tau\Big(tg\big(\alpha(t,x+\Delta x)\big) - tg\big(\alpha(t,x)\big)\Big) \\ &= \tau\big(u_x(t,x+\Delta x) - u_x(t,x)\big) \\ &= \tau\int_x^{x+\Delta x} u_{xx}(t,y)dy. \end{aligned}$$

If we name $f(t,x)$ a possible external force, we can add it to the last step of (4.2) to obtain

$$\tau\int_x^{x+\Delta x} u_{xx}(t,y)dy + \int_x^{x+\Delta x} f(t,y)\rho(t,y)\sqrt{1+u_x^2(t,y)}dy.$$

Thanks to the fundamental principle of dynamics, the following represents a force too:

$$\int_x^{x+\Delta x} u_{tt}(t,y)\rho(t,y)\sqrt{1+u_x^2(t,y)}dy.$$

Therefore we are allowed to write

$$\begin{aligned} &\int_x^{x+\Delta x} u_{tt}(t,y)\rho(t,y)\sqrt{1+u_x^2(t,y)}dy \\ &= \tau\int_x^{x+\Delta x} u_{xx}(t,y)dy + \int_x^{x+\Delta x} f(t,y)\rho(t,y)\sqrt{1+u_x^2(t,y)}dy, \end{aligned}$$

which is equal to

$$\begin{aligned} \int_x^{x+\Delta x} \rho_0 u_{tt}(t,y)dy - \int_x^{x+\Delta x} \tau u_{xx}(t,y)dy - \int_x^{x+\Delta x} \rho_0 f(t,y)dy = 0 \\ \int_x^{x+\Delta x} \big(\rho_0 u_{tt}(t,y) - \tau u_{xx}(t,y) - \rho_0 f(t,y)\big)dy = 0. \end{aligned}$$

Since the interval of integration $[x, x+\Delta x]$ is of arbitrary length we have

$$\rho_0 u_{tt} - \tau u_{xx} - \rho_0 f = 0,$$

and by defining $c^2 = \tau/\rho_0$ (and dividing the equation by ρ_0), we obtain

$$u_{tt} - c^2 u_{xx} = f.$$

4.3 Separation of variables

In this section we will utilize the technique of **separation of variables** to solve the homogeneous version of the linear partial differential equation we obtained in section 4.2, that is

$$u_{tt}(x,t) = c^2 u_{xx}(x,t).$$

More precisely, our goal is to find a solution u of the form

$$u(x,y) = X(x)Y(y),$$

and if this method is to work, when we substitute this formula into the equation we should be able to reorganize the terms in a way that the left-hand side contains only objects that depend on one variable, and the right-hand side contains only objects that depend on the other variable (we should obtain something like $P(x) = Q(y)$). This method could be used with more than two variables, but that does not concern us.

Let us dive straight into the case of the 1-dimensional wave equation. The Cauchy-Dirichlet problem presents itself as

$$\begin{cases} u_{tt} - c^2 u_{xx} = f & \text{for } 0 < x < L \text{ and } t > 0 \\ u(0,t) = u(L,t) = 0 & \text{for } t > 0 \\ u(x,0) = u_0(x) & \text{for } 0 < x < L \\ u_t(x,0) = v_0(x) & \text{for } 0 < x < L, \end{cases}$$

where the actual partial differential equation presents is $L(u) = f$,which is an inhomogeneous linear partial differential equation, with

$$L = \frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2}, \tag{4.3}$$

as its operator. Furthermore, the boundary of the domain $\Omega = [0, L]$ (we ignore temporarily the time variable t) is $\partial\Omega = \{0, L\}$, and the boundary conditions are of the form $B_1(u) = f_1$ and $B_2(u) = f_2$ with

$$B_1(u(x,t)) = u(0,t) \quad \text{and} \quad B_2(u(x,t)) = u(L,t)$$

as operators. Therefore, since $f_1 = f_2 = 0$, we are in front of homogeneous periodic boundary conditions.

Anyway, we will study the homogeneous version of the equation (which means that

$L(u) = 0$ with L as described in (4.3)) so the problem really is the following:

$$\begin{cases} u_{tt} - c^2 u_{xx} = 0 & \text{for } 0 < x < L \text{ and } t > 0 \\ u(0,t) = u(L,t) = 0 & \text{for } t > 0 \\ u(x,0) = u_0(x) & \text{for } 0 < x < L \\ u_t(x,0) = v_0(x) & \text{for } 0 < x < L. \end{cases}$$

Now let us look for a solution $U(x,t) = X(x)T(t)$; nothing guarantees that a solution like this one exists, however we will be able to justify this assumption a posteriori. By substituting $U(x,t)$ in the homogeneous equation, we obtain

$$X(x)T''(t) = c^2 X''(x)T(t)$$
$$\frac{X''(x)}{X(x)} = \frac{T''(t)}{c^2 T(t)},$$

but for this equation to be verified when $0 < x < L$ and $t > 0$, we need both sides to be constant. Therefore we need a certain $\lambda \in \mathbb{R}$ to have

$$\begin{aligned} \frac{X''(x)}{X(x)} &= \lambda \text{ for } 0 < x < L \\ \frac{T''(t)}{c^2 T(t)} &= \lambda \text{ for } t > 0. \end{aligned} \tag{4.4}$$

The boundary conditions $u(t,0) = u(t,L) = 0$ translate to $X(0) = X(L) = 0$, hence the first equation of (4.4) has its own conditions:

$$\begin{cases} X''(x) = \lambda X(x) \text{ for } 0 < x < L \\ X(0) = X(L) = 0. \end{cases} \tag{4.5}$$

We have already found the eigenvalues and eigenfunctions of this Sturm-Liouville problem in chapter 3, and they are, respectively,

$$\lambda_n = -\frac{n^2\pi^2}{L^2} \quad \text{and} \quad X_n(x) = c_n sin\left(\frac{n\pi x}{L}\right) \quad \text{for } n \in \mathbb{N}.$$

Now we can use them to solve the second equation of (4.4), namely

$$T''(t) = -\frac{n^2\pi^2c^2}{L^2}T(t)$$

from which we obtain the solutions

$$T_n(t) = \alpha_n cos\left(\frac{n\pi ct}{L}\right) + \beta_n sin\left(\frac{n\pi ct}{L}\right) \quad \text{for } n \in \mathbb{N}.$$

Ultimately, the separated-variables solutions of the partial differential equation and of the boundary conditions, are of the following form (for $n \in \mathbb{N}$):

$$u_n(t,x) = T_n(t)X_n(t) = sin\left(\frac{n\pi x}{L}\right)\left(\alpha_n cos\left(\frac{n\pi ct}{L}\right) + \beta_n sin\left(\frac{n\pi ct}{L}\right)\right)$$

Since $u_n(0,x) = \alpha_n sin\left(\frac{n\pi x}{L}\right)$, generally none of these solutions will satisfy the initial condition $u(0,x) = u_0(x)$. However, we can leverage the superposition principle: since the equation at hand is, as stated earlier, linear and homogeneous, we know that any linear combination of u_n will still remain solutions of (4.5). Therefore we look for an infinite series made of these solutions, and choose the right coefficients for it to both converge and satisfy the initial conditions. Thus, we write

$$u(x,t) = \sum_{n=1}^{+\infty} sin\left(\frac{n\pi x}{L}\right)\left(\alpha_n cos\left(\frac{n\pi ct}{L}\right) + \beta_n sin\left(\frac{n\pi ct}{L}\right)\right)$$

and impose

$$\begin{aligned} u(0,x) &= \sum_{n=1}^{+\infty} \alpha_n sin\left(\frac{n\pi x}{L}\right) = u_0(x) \text{ for } 0 < x < L \\ u_t(0,x) &= \sum_{n=1}^{+\infty} \beta_n \frac{n\pi c}{L} sin\left(\frac{n\pi x}{L}\right) = v_0(x) \text{ for } 0 < x < L. \end{aligned} \tag{4.6}$$

For the function $u(x,t)$ to be an actual solution, we have to be able to derivate it twice in both t and x, so our main focus has to be the behaviour of the coefficients α_n and β_n, hence we can do this by working on $u_0(x)$ and $v_0(x)$. We have both the ingredients we need to make all of this happen, namely theorems 3.19 and 3.21, but we must tread carefully.

First, we need to adjust the intervals on which $u_0(x)$ and $v_0(x)$ are defined, precisely from $[0,L]$ to $[-L,L]$ and we can do that by performing an **odd extension** of both functions, this way we can preserve the only-sine series too. Therefore, by following the definition 3.11 and applying lemma 3.12, the Fourier coefficients of the series expansion of $u_{0,odd}$ and $v_{0,odd}$ become, respectively

$$\begin{aligned} \alpha_{n,u_0} &= \frac{2}{L}\int_0^L u_{0,odd}(x) sin\left(\frac{n\pi x}{L}\right) dx = \frac{2}{L}\int_0^L u_0(x) sin\left(\frac{n\pi x}{L}\right) dx \\ \beta_{n,v_0} &= \frac{2}{L}\int_0^L v_{0,odd}(x) sin\left(\frac{n\pi x}{L}\right) dx = \frac{2}{L}\int_0^L v_0(x) sin\left(\frac{n\pi x}{L}\right) dx, \end{aligned}$$

which means that

$$\begin{aligned} \alpha_n &= \frac{2}{L}\int_0^L u_0(x) sin\left(\frac{n\pi x}{L}\right) dx \\ \beta_n &= \frac{2}{n\pi c}\int_0^L v_0(x) sin\left(\frac{n\pi x}{L}\right) dx. \end{aligned}$$

Let us now rewrite $u(x,t)$:

$$u(x,t) = \sum_{n=1}^{+\infty} \left(\alpha_n sin\left(\frac{n\pi x}{L}\right) cos\left(\frac{n\pi ct}{L}\right) + \beta_n sin\left(\frac{n\pi x}{L}\right) sin\left(\frac{n\pi ct}{L}\right)\right).$$

Suppose we can derive twice in t or x; one way or another we are going to end up with n^2 as a coefficient. To control this explosive behaviour we can utilize theorem 3.21 on the Fourier coefficients of u_0 and v_0. More precisely, we need

$$n^2|\alpha_n| \leq \frac{M_\alpha}{n^2} \quad \text{and} \quad n^2|\beta_n| \leq \frac{M_\beta}{n^2},$$

which are conditions equal to

$$|\alpha_n| \leq \frac{M_\alpha}{n^4} \quad \text{and} \quad |\beta_n| \leq \frac{M_\beta}{n^4},$$

for which the requirements $u_0 \in \mathcal{C}^4$ and $v_0 \in \mathcal{C}^3$ are enough (we have to keep in mind that $v_0(x) = u_t(0,x)$ and $M_\alpha, M_\beta > 0$ are constants).

Remark 4.6. Actually, leaning on 3.21, it would be enough to ask $u_0 \in \mathcal{C}^3$ with $u_0^{(3)} \in \mathcal{PS}$ and $v_0 \in \mathcal{C}^2$ with $v_0^{(2)} \in \mathcal{PS}$, but it is a very negligible upgrade from the previous requirements.

4.4 D'Alembert's formula

Let us consider the global Cauchy problem ($L = +\infty$):

$$\begin{cases} u_{tt} - c^2 u_{xx} = 0 & \text{for } t > 0,\ x \in \mathbb{R} \\ u(0,x) = g(x) \\ u_t(0,x) = h(x). \end{cases}$$

The equation

$$(\partial_t^2 - c^2\partial_x^2)u = 0$$

can be rewritten as

$$(\partial_t - c\partial_x)(\partial_t + c\partial_x)u = 0.$$

If we let $u(x,t) = v(y,\eta)$ with $y = x - ct$ and $\eta = x + ct$, then

$$\partial_t u = -c\partial_y v + c\partial_\eta cv$$
$$\partial_x u = \partial_y v + \partial_\eta v,$$

which is

$$\partial_t - c\partial_x = -2c\partial_y$$
$$\partial_t + c\partial_x = 2c\partial_\eta.$$

So now we obtained

$$u_{tt} - c^2 u_{xx} = (\partial_t - c\partial_x)(\partial_t + c\partial_x)u = -4c^2\partial_y\partial_\eta v = 0 \quad \Longrightarrow \quad v_{y\eta} = 0,$$

and $v_{y\eta} = 0$ has, as general integral,

$$\begin{aligned} \frac{\partial}{\partial \eta}\left(\frac{\partial}{\partial y} v\right) &= 0 \\ v(y,\eta) &= \int f(\eta)d\eta + G(y) \\ v(y,\eta) &= F(\eta) + G(y), \end{aligned}$$

where F and G are differentiable functions. By turning everything back to the initial variables we obtain

$$u(x,t) = F(x+ct) + G(x-ct),$$

i.e. u is the result of the overlapping of two waves (called **solitons**), traveling in opposite directions at the same speed c.

By imposing the initial conditions

$$\begin{cases} g(x) = u(0,x) = F(x) + G(x) \\ h(x) = u_t(0,x) = c\big(F'(x) - G'(x)\big), \end{cases}$$

we obtain

$$\begin{cases} F + G = g \\ c(F - G) = H \qquad \text{where } H(x) = \int h(x)dx. \end{cases}$$

If we then multiply by c the first one and either add or subtract the second one to it we get either $2cF = cg + H$ or $2cG = cg - H$, hence

$$\begin{aligned} F &= \frac{1}{2}g + \frac{1}{2c}H \\ G &= \frac{1}{2}g - \frac{1}{2c}H, \end{aligned}$$

which finally leads us to

$$\begin{aligned} u(x,t) &= \frac{1}{2}\big(g(x+ct) + g(x-ct)\big) + \frac{1}{2c}\big(H(x+ct) - H(x-ct)\big) \\ &= \frac{1}{2}\big(g(x+ct) + g(x-ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} h(y)dy, \end{aligned}$$

which is known as d'Alembert's formula. Whenever $g \in \mathcal{C}^k$ and $h \in \mathcal{C}^{k-1}$ then $u \in \mathcal{C}^k$ and it is the only one of that class by construction. Furthermore, one can notice the total absence of regularization aspects in the equation, which means that the regularity of u really solely depends on the regularity of the initial data.

Now that we obtained this formula, we want to investigate what is so special about it, apart from what has already been said about its regularity. The formula looks like this:

$$u(x,t) = \frac{1}{2}\big(g(x+ct) + g(x-ct)\big) + \frac{1}{2c}\int_{x-ct}^{x+ct} h(y)dy,$$

and to understand it better we need to think about the space-time plane, which, since we are considering the one-dimensional case, is the set $\mathcal{ST}_1 = \{(x,t) \in \mathbb{R} \times [0,+\infty)\}$.

If we pick any point $(x_0,t_0) \in \mathcal{ST}_1$, then we have

$$u(x_0,t_0) = \frac{1}{2}\big(g(x_0+ct_0) + g(x_0-ct_0)\big) + \frac{1}{2c}\int_{x_0-ct_0}^{x_0+ct} h(y)dy,$$

which shows that in order to compute $u(x_0,t_0)$ it is sufficient to know the value of g at $x_0 \pm ct_0$ and the value of h at every x such that $x \in [x_0 - ct_0, x_0 + ct_0]$. If we take a step back then we can notice that for every point $(x_0,t_0) \in \mathcal{ST}_1$ there are exactly two straight lines that go through it and through either $(x_0 + ct_0, 0)$ or $(x_0 - ct_0, 0)$. What is so special about them? Let us call γ_+ the one that goes through (x_0,t_0) and $(x_0+ct_0,0)$ and γ_- the one that goes through (x_0,t_0) and $(x_0-ct_0,0)$. Respectively, their equations in $\mathcal{ST}_1$ are

$$\begin{aligned} \gamma_+ &: t = -\frac{1}{c}(x-x_0) + t_0 \\ \gamma_- &: t = \frac{1}{c}(x-x_0) + t_0. \end{aligned} \tag{4.7}$$

Hence if we consider the formula

$$u(x,t) = F(x+ct) + G(x-ct),$$

we can easily see that F is constant on γ_+, whereas G is constant on γ_-. More precisely, if we substitute the first equation of (4.7) in $F(x+ct)$, we obtain $F(x_0+ct_0)$, whereas if we substitute the second equation of (4.7) in $G(x+ct)$, we obtain $F(x_0-ct_0)$. What this means physically, is that as the solitons travel, their height remains constant except when they cross one another. The curves γ_+ and γ_- are the vehicles through which information is carried by the equation. That is why these lines are called **characteristics**.

Moreover, if we take any point $(x_0,t_0) \in \mathcal{ST}_1$, then the characteristics that cross it outline four different regions of $\mathcal{ST}_1$, but only two of them are important: the southern one, which is called **domain of dependence** relative to (x_0,t_0), and the northern one, which is called **domain of influence** relative to (x_0,t_0). The first one has that name because it is the one representing the past of the solution, and its points represent all the different states of the solitons before they cross at time t_0, whereas the second one has that name because it is the one representing the future of the solution, and its points represent all the different states of the solitons after they cross at time t_0.

Bibliography

[1] M. Bramanti: *Metodi di Ortogonalità*, notes, University of Milano, 2012.

[2] M. Bramanti: *Metodi di Analisi Matematica per l'Ingegneria*, Esculapio, 2017.

[3] M. Cicognani: *Equazioni Differenziali*, notes, University of Cesena, 2013.

[4] G. Folland: *Fourier Analysis and Its Applications*, Wadsworth & Brooks/Cole, 1992.

[5] M. Fowler: *Physics 152: Oscillation and Waves*, notes, University of Virginia, 2007.

[6] F. John: *Partial Differential Equations*, Springer-Verlag, 1982.

Acknowledgements

Coming soon...