Undergraduate Thesis

Major in Mathematics

Faculty of Mathematics
University of Barcelona

Fundamental Theorems of Functional Analysis and Applications

Author: Joan Carles Bastons Garcia

Advisor: Dr. María Jesús Carro Rossell
Department: Matemàtica Aplicada i Anàlisi

Barcelona, January 18, 2016

Abstract

Among the fundamental theorems of Functional Analysis are the open mapping theorem, the closed graph theorem, the uniform boundedness principle, the Banach-Steinhaus theorem and the Hahn-Banach theorem. We study them in the context of Banach spaces and applications in Analysis like the divergence of Fourier series, the Riesz representation theorem, the existence of nowhere differentiable continuous functions, etc. Apart from Mathematics, we demonstrate that those theorems can play an important role in Physics by examining two more applications: the moment problem and a rocket ascent. The whole thesis outlines that those theorems are applied in many disciplines and can lead to relevant results.

Acknowledgments

I would like to express my gratitude to Dr. María Jesús Carro for her guidance and advice while developing this thesis. I am also grateful for her attention and encouragement to work with the rigor that Mathematics require.

I would also like to thank my parents and sister for their support in all aspects while carrying out this thesis and through majoring years. I am also grateful to the rest of my family and friends.

Contents

Introduction **1**

1 Preliminaries **3**

1.1 A brief review on Banach spaces . . . 3

1.1.1 Banach spaces, operators and dual space . . . 3

1.1.2 Product and quotient space . . . 4

1.2 Density of the continuous functions in L^1 . . . 7

1.3 Riemann-Stieltjes integration . . . 8

1.3.1 Functions of bounded variation . . . 8

1.3.2 The Riemann-Stieltjes integral and properties . . . 10

2 Fundamental theorems of Functional Analysis **15**

2.1 Baire's theorem, the open mapping theorem and the closed graph theorem 15

2.1.1 Baire's theorem . . . 15

2.1.2 The open mapping theorem . . . 16

2.1.3 The closed graph theorem . . . 17

2.2 The uniform boundedness principle and the Banach-Steinhaus theorem . 18

2.3 The Hahn-Banach theorem . . . 20

2.3.1 Analytic version . . . 20

2.3.2 Geometric version . . . 23

3 Applications to areas of Analysis **27**

3.1 Applications to Real Analysis . . . 27

3.1.1 Existence of nowhere differentiable continuous functions . . . 27

3.1.2 The Riesz representation theorem . . . 29

3.2 Application to Functional Analysis: separable Banach spaces . . . 32

3.3 Application to Harmonic Analysis: divergence of Fourier series . . . 33

3.3.1 In $(\mathcal{C}([a,b]), ||\cdot||_\infty)$ . . . 34

3.3.2 In $(L^1([a,b]), ||\cdot||_1)$ . . . 36

3.4 Application to Numerical Analysis: divergence of Lagrange interpolation 39

3.5 Application to Differential Equations 41

4 Applications to other areas **43**

4.1 The moment problem 43

4.2 Minimum norm problems 46

4.2.1 The Chebyshev approximation 50

4.2.2 Optimal control of rockets 50

Conclusions **55**

Bibliography **57**

Index **59**

Introduction

Among the fundamental theorems of Functional Analysis are the open mapping theorem, the closed graph theorem, the uniform boundedness principle, the Banach-Steinhaus theorem and the Hahn-Banach theorem. They date from the first third of the past century, when they were formulated in the context of Banach spaces. As some of their names suggest, they refer to properties of operators like boundedness, continuity, extension and openness. Thus, their appliance goes beyond Functional Analysis and they are present in other branches of Mathematics such as Harmonic Analysis and Differential Equations, among others.

This project aims to study those theorems and their applications from a multidisciplinary approach. For this purpose, applications in different areas of Mathematics are carefully chosen, taking into consideration variety and the necessary background. Other purposes of this thesis are to be self-contained and to put into practice a wide range of skills and knowledge acquired in my majoring years.

The thesis is composed of four chapters. The first one consists of concepts and results which are not central to this thesis, though they are often auxiliary, like the Riemann-Stieltjes integral. In Chapter 2, those fundamental theorems are formulated for Banach spaces, following the versions given in [1, 2]. These two chapters are the starting point to develop the rest of the thesis.

In Chapter 3, the theorems are applied to some areas of Analysis, beginning with the existence of nowhere differentiable continuous functions followed by the Riesz representation theorem. In Harmonic Analysis, from a historical overview of the convergence of Fourier series, the uniform and the L^1-norm convergences are studied. The chapter also includes a result in Numerical Analysis, the Lagrange interpolating polynomial does not converge uniformly. The references for these applications are taken from [3, 6, 9], though in most cases these results are improved by studying more general versions or by complementing the proofs.

Another goal of this project is to analyze and solve physical problems with the theorems in Chapter 2. For this end, Chapter 4 contains two physical applications; the first one is the moment problem, suggested by T.J. Stieltjes as a mechanical problem in [12], which has evolved to different versions in Probability or in systems with infinitely linear equations. The second section is the study of the optimal rocket ascent in terms of fuel expenditure, developed in [7, 9, 14].

Chapter 1

Preliminaries

In this introductory chapter, concepts and results which will be often used in the following chapters are provided to the reader in order to assure full comprehension. In the first section, notions of Banach spaces, continuous linear operators and dual space are introduced. However, many properties are not proved because they were studied in the course *Anàlisi real i funcional*, taught at University of Barcelona, and they can also be found in any introductory manual. The rest of the sections contain results that have not been studied in any previous course. The second section only contains a proof of the density of the continuous functions in L^1, a result that will be necessary in some applications. Finally, the last section consists of two parts, the first one focuses on functions of bounded variation, while the second one on the definition of the Riemann-Stieltjes integral and properties that will be mainly required in Chapter 4. In these sections and in the whole thesis, $\mathbb{K}$ will denote either $\mathbb{R}$ or $\mathbb{C}$.

1.1 A brief review on Banach spaces

1.1.1 Banach spaces, operators and dual space

Definition 1.1.1. A vector space E is Banach if and only if it is normed and complete.

Example 1.1.2. The following normed vector spaces are Banach.

(i) $(\mathbb{R}^n, |\cdot|)$ and $(\mathbb{C}^n, |\cdot|)$.

(ii) $(\mathcal{C}([a,b]), ||\cdot||_\infty)$.

(iii) $L^p([a,b]) := \{f : f \in \mathscr{L} \text{ and } ||f||_p^p := \int_a^b |f(x)|^p dx < \infty\}$, where $\mathscr{L}$ is the set of all Lebesgue-measurable functions and $1 \leq p < \infty$.

(iv) $L^\infty([a,b]) := \{f : f \in \mathscr{L} \text{ and } ||f||_\infty := \sup_{a \leq x \leq b} |f(x)| < \infty\}$.

(v) $l^1 := \{a = \{a_n\}_{n\in\mathbb{N}} \subset \mathbb{K} : ||a||_1 := \sum_{n=1}^\infty |a_n| < \infty\}$.

Proposition 1.1.3. *A normed vector space is Banach if and only if every absolutely convergent series is convergent.*

Proposition 1.1.4. *Let E be a Banach space. If a subspace $A \subseteq E$ is closed in E, then A is a Banach space.*

Definition 1.1.5. Let E, F be two normed vector spaces. A *linear operator* $T : E \to F$ is a linear function between the two normed vector spaces and *the norm of the operator* is defined by

$$||T||:= \sup_{||x||_E \leq 1} ||Tx||_F = \sup_{||x||_E = 1} ||Tx||_F = \sup_{x \neq 0} \frac{||Tx||_F}{||x||_E}.$$

Proposition 1.1.6. *Let E, F be two normed vector spaces and $T : E \to F$ a linear operator. The following conditions are equivalent:*

(i) T is continuous.

(ii) $||T|| < \infty$.

(iii) There exists a constant $C > 0$ so that $||Tx||_F \leq C||x||_E$ for all $x \in E$.

If T is continuous, $||T||$ is the lowest constant that satisfies (iii).

Remark 1.1.7. Continuous linear operators between two normed vector spaces are often called bounded, motivated by Proposition 1.1.6 (iii).

Remark 1.1.8. Let E and F be two normed vector spaces. We denote as $\mathcal{L}(E, F)$ the set of all continuous linear operators between E and F.

Proposition 1.1.9. *Let E and F be two normed vector spaces. Then, $\mathcal{L}(E, F)$ is a normed vector space with the norm of operators. If F is Banach, so is $\mathcal{L}(E, F)$.*

Definition 1.1.10. Let E be a normed vector space over $\mathbb{K}$. The dual space of E is defined as $E^* = \mathcal{L}(E, \mathbb{K})$.

Corollary 1.1.11. *The dual space of a normed vector space is a Banach space.*

1.1.2 Product and quotient space

Proposition 1.1.12. *If E, F are two Banach spaces over $\mathbb{K}$, then $E \times F$ is a Banach space with the norm $||(x, y)|| = \max(||x||_E, ||y||_F)$.*

Proof. First, we show that $E \times F$ is normed.

(i) $0 \leq ||x||_E \leq \max(||x||_E, ||y||_F) = ||(x, y)||$ for all $(x, y) \in E \times F$, and $||(x, y)|| = 0$ if and only if $||x||_E = ||y||_F = 0$ if and only if $x = 0$ and $y = 0$.

(ii) For all $(x, y) \in E \times F$ and all $\lambda \in \mathbb{K}$,

$$||\lambda(x, y)|| = ||(\lambda x, \lambda y)|| = \max(||\lambda x||_E, ||\lambda y||_F) = |\lambda| \max(||x||_E, ||y||_F) = |\lambda| \; ||(x, y)||.$$

(iii) For all $(x, y), (z, t) \in E \times F$,

$$\begin{aligned} ||(x, y) + (z, t)|| &= ||(x + z, y + t)|| = \max(||x + z||_E, ||y + t||_F) \\ &\leq \max(||x||_E + ||z||_E, ||y||_F + ||t||_F) \\ &\leq \max(||x||_E, ||y||_F) + \max(||z||_E, ||t||_F) = ||(x, y)|| + ||(z, t)||. \end{aligned}$$

Next, we prove that $E \times F$ is a complete space. Consider a Cauchy sequence $\{(x_n, y_n)\}_n$ in $E \times F$, then

$$||x_n - x_m||_E \leq ||(x_n, y_n) - (x_m, y_m)|| \to 0 \text{ as } n, m \to \infty, \text{ and}$$
$$||y_n - y_m||_F \leq ||(x_n, y_n) - (x_m, y_m)|| \to 0 \text{ as } n, m \to \infty.$$

Since $\{x_n\}_{n\geq 1}$ is a Cauchy sequence in the Banach space E, there exists $x = \lim_{n\to\infty} x_n$ in E. Similarly, there exists $y = \lim_{n\to\infty} y_n$ in F.

$$||(x, y) - (x_n, y_n)|| = \max(||x - x_n||_E, ||y - y_n||_F) \leq ||x - x_n||_E + ||y - y_n||_F \to 0,$$

as $n \to \infty$. □

Let E be a vector space and F a subspace of E. We say that $u, v \in E$ are related if and only if $u - v \in F$. This is an equivalence relation and the equivalence class of a vector u is $[u] = \{u + v : v \in F\} = u + F$. The quotient set $E/F = \{x + F : x \in E\}$ is a vector space.

Lemma 1.1.13. *Let $(E, ||\cdot||)$ be a normed vector space over $\mathbb{K}$ and F a closed subspace of E. Then, E/F with the functional*

$$\begin{aligned} ||\cdot||_q \colon \ & E/F \to \mathbb{R} \\ & x + F \mapsto ||x + F||_q = \inf_{y\in F} ||x + y|| \end{aligned}$$

is a normed vector space.

Proof. We will show that $||\cdot||_q$ is a norm on E/F. We first notice that $||x + F||_q \geq 0$ for all $x \in E$.
Next, if $[0] = [x] = x + F = F$, then $||x + F||_q = ||0 + F||_q = 0$. Conversely, suppose that $||x + F||_q = 0$ for some $x \in E$. Then, there exists a sequence $\{y_k\}_k$ in F so that $\lim_k ||x + y_k|| = 0$, that is, $x = \lim_k(-y_k)$. Since F is closed, $x \in F$ and, hence, $[x] = [0]$.
For all $x \in E$ and all $\lambda \in \mathbb{K}$, with $\lambda \neq 0$,

$$||\lambda(x + F)||_q = ||\lambda x + F||_q = \inf_{y\in F} ||\lambda x + y|| = \inf_{y'=\frac{y}{\lambda}\in F} ||\lambda x + \lambda y'|| = |\lambda| \ ||x + F||_q.$$

Finally, the triangle inequality is readily shown. Given $x, y \in E$ and $\epsilon > 0$, consider $z_1, z_2 \in F$ so that

$$||x + z_1|| \leq ||x + F||_q + \frac{\epsilon}{2},$$
$$||y + z_2|| \leq ||y + F||_q + \frac{\epsilon}{2}.$$

Then,

$$\begin{aligned} ||(x + F) + (y + F)||_q &= \inf_{z\in F} ||x + y + z|| \leq ||x + z_1 + y + z_2|| \leq ||x + z_1|| + ||y + z_2|| \\ &\leq ||x + F||_q + ||y + F||_q + \epsilon. \end{aligned}$$ □

Proposition 1.1.14. *If $(E, ||\cdot||)$ is a Banach space and F is a closed subspace of E, then $(E/F, ||\cdot||_q)$ is a Banach space.*

Proof. Let $\{x_n + F\}_n$ be a Cauchy sequence in E/F. We will show that there exists a subsequence that converges in E/F and, hence, the sequence is convergent. Consider $\{x_{n_k} + F\}_k$ a subsequence such that

$$||x_{n_{k+1}} - x_{n_k} + F||_q < \tfrac{1}{2^k} \quad \text{for all integers } k \geq 1.$$

Next, we build a sequence $\{y_k\}_k$ in F so that

$$||(x_{n_{k+1}} - y_{k+1}) - (x_{n_k} - y_k)|| < \frac{1}{2^{k-1}} \quad \text{for all integers } k \geq 1.$$

Indeed, consider $y_1 = 0$, then from

$$\inf_{y\in F} ||(x_{n_1} - y_1) - (x_{n_2} - y)|| = \inf_{y\in F} ||x_{n_1} - x_{n_2} + y|| = ||x_{n_1} - x_{n_2} + F||_q < \frac{1}{2},$$

it follows that there exists $y_2 \in F$ so that

$$||(x_{n_1} - y_1) - (x_{n_2} - y_2)|| < 2 \cdot \frac{1}{2} = 1.$$

Similarly, from

$$\inf_{y\in F} ||(x_{n_2} - y_2) - (x_{n_3} - y)|| = \inf_{y\in F} ||x_{n_2} - x_{n_3} - y|| = ||x_{n_2} - x_{n_3} + F||_q < \frac{1}{2^2},$$

it follows that there exists $y_3 \in F$ so that

$$||(x_{n_2} - y_2) - (x_{n_3} - y_3)|| < 2 \cdot \frac{1}{2^2} = \frac{1}{2}.$$

Recursively, we obtain a sequence $\{z_k = x_{n_k} - y_k\}_k$ in E, such that

$$||z_{k+1} - z_k|| < \frac{1}{2^{k-1}} \quad \text{for all integers } k \geq 1.$$

Besides, the sequence $\{z_k\}_k$ is Cauchy in E. Indeed, given $\epsilon > 0$, there exists $k_0 \in \mathbb{N}$ so that $\frac{1}{2^{k_0}} < \epsilon$. For all $n, m > k_0 + 1$, with $n > m$, we have that

$$\begin{aligned} ||z_n - z_m|| &\leq \sum_{i=0}^{n-m-1} ||z_{m+1+i} - z_{m+i}|| \leq \sum_{i=0}^{n-m-1} \frac{1}{2^{m+i-1}} \leq \sum_{i=0}^{\infty} \frac{1}{2^{m+i-1}} = \frac{\frac{1}{2^{m-1}}}{1 - \frac{1}{2}} \\ &= \frac{1}{2^{m-2}} \leq \frac{1}{2^{k_0}} < \epsilon. \end{aligned}$$

Since $\{z_k\}_k \subset E$ is a Cauchy sequence in a Banach space, this sequence converges to a vector $z \in E$. Finally,

$$||(x_{n_k} + F) - (z + F)||_q = ||x_{n_k} - z + F - y_k||_q = ||z_k - z + F||_q \leq ||z_k - z|| \to 0, \text{ as } k \to \infty.$$

$\square$

1.2 Density of the continuous functions in L^1

Proposition 1.2.1. *$\mathcal{C}([-\pi,\pi])$ is dense in $L^1([-\pi,\pi])$.*

Proof. Let $f \in L^1([-\pi,\pi])$, we will prove that for all $\epsilon > 0$ there exists $g \in \mathcal{C}([-\pi,\pi])$ so that $||f-g||_1 \leq \epsilon$.
Let us start with a function $\chi_{(a,b)} \in L^1([-\pi,\pi])$ with $-\pi \leq a < b \leq \pi$. Given $\epsilon > 0$ small enough, we consider a continuous function $g_\epsilon : [-\pi,\pi] \to \mathbb{R}$ defined by

$$g_\epsilon(x) = \begin{cases} 0, & \text{if} \quad -\pi \leq x < a, \\ \frac{x-a}{\epsilon}, & \text{if} \quad a \leq x < a+\epsilon, \\ 1, & \text{if} \quad a+\epsilon \leq x < b-\epsilon, \\ \frac{b-x}{\epsilon}, & \text{if} \quad b-\epsilon \leq x < b, \\ 0, & \text{if} \quad b \leq x \leq \pi. \end{cases}$$

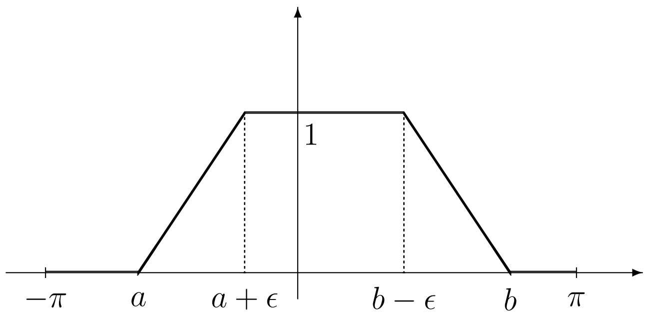

Figure 1.2.1: Representation of the continuous function g_ϵ equal to $\chi_{(a,b)}$ on $(a+\epsilon, b-\epsilon)$.

Then,

$$||\chi_{(a,b)} - g_\epsilon||_1 = \int_{-\pi}^{\pi} \left|\chi_{(a,b)}(x) - g_\epsilon(x)\right| dx = \epsilon.$$

Now, let us consider $U_N = \cup_{i=1}^N (a_i, b_i) \subset [-\pi,\pi]$ and the function $\chi_{U_N} \in L^1([-\pi,\pi])$. We can assume that U_N is a finite union of disjoint intervals. Therefore,

$$\chi_{U_N}(x) = \sum_{i=1}^N \chi_{(a_i,b_i)}(x) \quad \text{for all } x \in [-\pi,\pi].$$

For all $1 \leq i \leq N$ there exists $g_i \in \mathcal{C}([-\pi,\pi])$ so that $||g_i - \chi_{(a_i,b_i)}||_1 \leq \frac{\epsilon}{N}$. Besides, $\sum_{i=1}^N g_i \in \mathcal{C}([-\pi,\pi])$. Hence,

$$\left|\left| \sum_{i=1}^N g_i - \chi_{U_N} \right|\right|_1 = \left|\left| \sum_{i=1}^N \left(g_i - \chi_{(a_i,b_i)}\right) \right|\right|_1 \leq \sum_{i=1}^N ||g_i - \chi_{(a_i,b_i)}||_1 \leq \sum_{i=1}^N \frac{\epsilon}{N} = \epsilon.$$

Consider an open subset of $[-\pi,\pi]$, U, and the function $\chi_U \in L^1([-\pi,\pi])$. Notice that U can be expressed as

$$U = \biguplus_{i=1}^{\infty} (a_i, b_i).$$

Since $\{U_N = \biguplus_{i=1}^N (a_i, b_i)\}_N$ is a family of subsets of U with $U_N \subset U_{N+1}$, there exists $N_0 \in \mathbb{N}$ so that

$$|U \setminus U_N| < \tfrac{\epsilon}{2} \quad \text{whenever } N \geq N_0.$$

Besides,

$$|U \setminus U_N| = \int_{-\pi}^{\pi} \chi_{U \setminus U_N}(x)dx = \int_{-\pi}^{\pi} \big(\chi_U(x) - \chi_{U_N}(x)\big)dx = ||\chi_U - \chi_{U_N}||_1.$$

As we have shown before, there exists $g_N \in \mathcal{C}([-\pi, \pi])$ so that $||\chi_{U_N} - g_N||_1 \leq \frac{\epsilon}{2}$. Hence,

$$||\chi_U - g_N||_1 \leq ||\chi_U - \chi_{U_N}||_1 + ||\chi_{U_N} - g_N||_1 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

We continue with a measurable subset, A, of $[-\pi, \pi]$ and the function $\chi_A \in L^1([-\pi, \pi])$. Since the Lebesgue measure is regular, there exists an open set U so that

$$A \subset U \subset [-\pi, \pi] \text{ and } ||\chi_U - \chi_A||_1 = |U \setminus A| < \tfrac{\epsilon}{2}.$$

According to the previous case, there exists $g \in \mathcal{C}([-\pi, \pi])$ so that $||\chi_U - g||_1 < \frac{\epsilon}{2}$. Hence,

$$||\chi_A - g||_1 \leq ||\chi_A - \chi_U||_1 + ||\chi_U - g||_1 < \epsilon.$$

We next consider a simple function, $s = \sum_{i=1}^n a_i \chi_{A_i}$ with $a_i \in \mathbb{R} \setminus \{0\}$ and A_i measurable subsets of $[-\pi, \pi]$. Given $\epsilon > 0$, for all $1 \leq i \leq n$ there exists $g_i \in \mathcal{C}([-\pi, \pi])$ so that $||\chi_{A_i} - g_i||_1 < \frac{\epsilon}{|a_i| n}$. Note that $\sum_{i=1}^n a_i g_i$ is continuous. Then,

$$\left|\left| s - \sum_{i=1}^n a_i g_i \right|\right|_1 = \left|\left| \sum_{i=1}^n a_i \chi_{A_i} - \sum_{i=1}^n a_i g_i \right|\right|_1 \leq \sum_{i=1}^n |a_i| \; ||\chi_{A_i} - g_i||_1 < \epsilon.$$

Given $f \in L^1([-\pi, \pi])$ and $\epsilon > 0$, there exists a simple function s so that $||f - s||_1 < \frac{\epsilon}{2}$ (every measurable function can be approximated by simple functions). Since s is a simple function, there exits $g \in \mathcal{C}([-\pi, \pi])$ so that $||g - s||_1 < \frac{\epsilon}{2}$. Hence, $||g - f||_1 < \epsilon$. □

1.3 Riemann-Stieltjes integration

1.3.1 Functions of bounded variation

Definition 1.3.1. Let $f : [a, b] \to \mathbb{R}$.

(i) If $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ is a partition of $[a, b]$, we define

$$P(f) := \sum_{k=1}^n |f(x_k) - f(x_{k-1})|.$$

(ii) *The total variation of* f on $[a, b]$ is defined by

$$V(f) := \sup_P P(f),$$

where the supremum is taken over all partitions of $[a, b]$.

(iii) For all $x \in [a,b]$, we denote $V_{[a,x]}(f)$ as *the total variation of f on the interval* $[a,x]$, which is a function of x.

(iv) f is said to be a *function of bounded variation on* $[a,b]$, $f \in BV([a,b])$, if and only if $V(f) < \infty$.

Definition 1.3.2. A partition P^* of an interval $[a,b]$ is said to be *thinner* than P if $P \subset P^*$.

Lemma 1.3.3. *Let $f \in BV([a,b])$ and $a \leq x < y \leq b$. Then,*

$$V_{[a,y]}(f) = V_{[a,x]}(f) + V_{[x,y]}(f).$$

Proof. Let P be a partition of $[a,y]$, which may not include x. So, let $P^* = P \cup \{x\}$. Then,

$$P^*(f) - P(f) = |f(x_i) - f(x)| + |f(x) - f(x_{i-1})| - |f(x_i) - f(x_{i-1})| \geq 0.$$

Since P^* is a partition of $[a,y]$, it can be expressed as $P^* = P_1 \cup P_2$, where P_1 contains the points of P^* belonging to $[a,x]$ and P_2 the ones belonging to $[x,y]$. Then,

$$P(f) \leq P^*(f) = P_1(f) + P_2(f) \leq V_{[a,x]}(f) + V_{[x,y]}(f).$$

Taking the supremum over P,

$$V_{[a,y]}(f) \leq V_{[a,x]}(f) + V_{[x,y]}(f).$$

Conversely, let P_1 and P_2 be two fixed partitions of $[a,x]$ and $[x,y]$, respectively. Then, $P = P_1 \cup P_2$ is a partition of $[a,y]$. Therefore,

$$P(f) = P_1(f) + P_2(f) \leq V_{[a,y]}(f).$$

We first take the supremum over P_1,

$$V_{[a,x]}(f) + P_2(f) \leq V_{[a,y]}(f),$$

and now over P_2,

$$V_{[a,x]}(f) + V_{[x,y]}(f) \leq V_{[a,y]}(f).$$

□

Remark 1.3.4. If $a \leq x < y \leq b$, then $V_{[a,y]}(f) = V_{[a,x]}(f) + V_{[x,y]}(f) \geq V_{[a,x]}(f)$.

Proposition 1.3.5. *If $f \in BV([a,b])$, then there exist two increasing functions g, h defined on $[a,b]$ such that $f = g - h$.*

Proof. Let $g(x) = V_{[a,x]}(f)$ be an increasing function such that

$$0 \leq g(x) \leq V(f) < \infty \quad \text{for all } a \leq x \leq b.$$

Since f and g are bounded, $h(x) = g(x) - f(x)$ is well-defined on $[a,b]$. Finally, if $x < y$, then

$$h(y) - h(x) = V_{[a,y]}(f) - V_{[a,x]}(f) - (f(y) - f(x)) = V_{[x,y]}(f) - (f(y) - f(x)) \geq 0.$$

Hence, h is increasing. □

1.3.2 The Riemann-Stieltjes integral and properties

Definition 1.3.6. *The norm of a partition* $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ is defined by

$$||P||= \max_{1\leq i\leq n} |x_i - x_{i-1}|.$$

Definition 1.3.7. Let $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ be a partition of $[a, b]$, $f, g : [a, b] \to \mathbb{R}$ two arbitrary functions and

$$\mu = \{\mu_k \in [x_{k-1}, x_k] : \; 1 \leq k \leq n\}.$$

(i) We define the *Riemann-Stieltjes sum*

$$P(f, g, \mu) := \sum_{k=1}^{n} f(\mu_k)\big(g(x_k) - g(x_{k-1})\big) = \sum_{k=1}^{n} f(\mu_k)\Delta g_k.$$

(ii) f is *Riemann-Stieltjes integrable* with respect to g on $[a, b]$, $f \in RS([a, b], g)$, if and only if there exists $L \in \mathbb{R}$ such that for all $\epsilon > 0$ there exists $\delta > 0$ such that $||P||< \delta$ implies that $|P(f, g, \mu) - L|< \epsilon$. In that case, *the Riemann-Stieltjes integral* is defined by

$$\int_a^b f \; dg := \lim_{||P||\to 0} P(f, g, \mu) = L.$$

Proposition 1.3.8. *Let* $a < b < c$*. If* $\int_a^b f \; dg$*,* $\int_b^c f \; dg$ *and* $\int_a^c f \; dg$ *exist, then*

$$\int_a^c f \; dg = \int_a^b f \; dg + \int_b^c f \; dg.$$

Proof. Let $\epsilon > 0$. There exists $\delta_1 > 0$ such that, for any partition P_1 of $[a, b]$ with $||P_1||< \delta_1$, it holds that

$$\left|P_1(f, g, \mu_1) - \int_a^b f \; dg\right| < \frac{\epsilon}{3}.$$

Similarly, there exists $\delta_2 > 0$ such that, for any partition P_2 of $[b, c]$ with $||P_2||< \delta_2$, it holds that

$$\left|P_2(f, g, \mu_2) - \int_b^c f \; dg\right| < \frac{\epsilon}{3}.$$

Finally, there exists $\delta_3 > 0$ such that, for any partition P_3 of $[a, c]$ with $||P_3||< \delta_3$, it holds that

$$\left|P_3(f, g, \mu_3) - \int_a^c f \; dg\right| < \frac{\epsilon}{3}.$$

Let $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Consider P_1 and P_2 two partitions of $[a, b]$ and $[b, c]$, respectively, so that $||P_i||< \delta$ for $i = 1, 2$. Let $P = P_1 \cup P_2$ be a partition of $[a, c]$ with $||P||< \delta \leq \delta_3$. Then,

$$\begin{aligned}\left| \int_a^c f \; dg - \left(\int_a^b f \; dg + \int_b^c f \; dg\right)\right| &\leq \left|\int_a^c f \; dg - P(f, g, \mu)\right| + \left|P_1(f, g, \mu_1) - \int_a^b f \; dg\right| \\ &\quad + \left|P_2(f, g, \mu_2) - \int_b^c f \; dg\right| < \epsilon.\end{aligned}$$

□

Proposition 1.3.9. *If f is Riemann integrable on $[a,b]$ and $g \in \mathcal{C}^1([a,b])$, then f belongs to $RS([a,b],g)$ and*

$$\int_a^b f \; dg = \int_a^b f(x)g'(x)dx.$$

Proof. By the mean value theorem, $\Delta g_k = g'(z_k)\Delta x_k$ with $z_k \in [x_{k-1}, x_k]$ for all $1 \leq k \leq n$. Then,

$$P(f,g,\mu) = \sum_{k=1}^{n} f(\mu_k)g'(z_k)\Delta x_k = \sum_{k=1}^{n} f(\mu_k)g'(\mu_k)\Delta x_k + \sum_{k=1}^{n} f(\mu_k)[g'(z_k) - g'(\mu_k)]\Delta x_k.$$

Since f and g' are Riemann integrable on $[a,b]$, so is $f \cdot g'$ and

$$\sum_{k=1}^{n} f(\mu_k)g'(\mu_k)\Delta x_k \to \int_a^b f(x)g'(x)dx \quad \text{as } ||P|| \to 0.$$

We next show that the second addend tends to zero as the partition becomes thinner. Given $\epsilon > 0$, we can choose $\delta > 0$ so that if $|x-y| < \delta$, then

$$|g'(x) - g'(y)| < \frac{\epsilon}{||f||_\infty (b-a)} \; ,$$

since g' is uniformly continuous on $[a,b]$. Therefore, for any partition $||P|| < \delta$, we have that $|\mu_k - z_k| < \delta$ for all $1 \leq k \leq n$, and

$$\left| \sum_{k=1}^{n} f(\mu_k)[g'(z_k) - g'(\mu_k)]\Delta x_k \right| < \sum_{k=1}^{n} |f(\mu_k)| \frac{\epsilon}{||f||_\infty (b-a)} \Delta x_k \leq \epsilon.$$

In case $f = 0$ or $a = b$, the proof is immediate. □

In the case that f is bounded and g increasing, the Riemann-Stieltjes theory is very similar to the Riemann integral theory. For this reason, we introduce the upper and lower sums and integrals.

Definition 1.3.10. Let $f, g : [a,b] \to \mathbb{R}$ be two functions such that f is bounded and g is increasing. Let $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ be a partition of $[a,b]$, we denote

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x) \quad \text{and} \quad M_i = \sup_{x \in [x_{i-1}, x_i]} f(x) \quad \text{for all } 1 \leq i \leq n.$$

(i) The *upper sum* is defined by

$$U(P,f,g) := \sum_{i=1}^{n} M_i \Delta g_i.$$

(ii) The *lower sum* is defined by

$$L(P,f,g) := \sum_{i=1}^{n} m_i \Delta g_i.$$

(iii) The *upper integral* is defined by

$$\overline{\int_a^b} f \, dg := \inf_P U(P, f, g),$$

where the infimum is taken over all partitions of $[a, b]$.

(iv) The *lower integral* is defined by

$$\underline{\int_a^b} f \, dg := \sup_P L(P, f, g),$$

where the supremum is taken over all partitions of $[a, b]$.

Remark 1.3.11. Notice that $L(P, f, g) \leq P(f, g, \mu) \leq U(P, f, g)$.

Lemma 1.3.12. *Let $f, g : [a, b] \to \mathbb{R}$ be two functions such that f is bounded and g is increasing. Let $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ be a partition of $[a, b]$. If P^* is a thinner partition of $[a, b]$, then*

$$L(P, f, g) \leq L(P^*, f, g) \quad \text{and} \quad U(P^*, f, g) \leq U(P, f, g).$$

Proof. It is enough to consider $P^* = P \cup \{x^*\}$ with $x_{i-1} < x^* < x_i$ for some $0 \leq i \leq n$. We denote $m' = \inf_{x \in [x_{i-1}, x^*]} f(x)$ and $m'' = \inf_{x \in [x^*, x_i]} f(x)$. Then,

$$\begin{aligned} L(P^*, f, g) - L(P, f, g) &= m'[g(x^*) - g(x_{i-1})] + m''[g(x_i) - g(x^*)] - m_i[g(x_i) - g(x_{i-1})] \\ &= (m' - m_i)[g(x^*) - g(x_{i-1})] + (m'' - m_i)[g(x_i) - g(x^*)] \geq 0. \end{aligned}$$

Similarly, if we denote $M' = \sup_{x \in [x_{i-1}, x^*]} f(x)$ and $M'' = \sup_{x \in [x^*, x_i]} f(x)$, then

$$\begin{aligned} U(P, f, g) - U(P^*, f, g) &= M_i[g(x_i) - g(x_{i-1})] - M'[g(x^*) - g(x_{i-1})] - M''[g(x_i) - g(x^*)] \\ &= (M_i - M')[g(x^*) - g(x_{i-1})] + (M_i - M'')[g(x_i) - g(x^*)] \geq 0. \end{aligned}$$

$\square$

Proposition 1.3.13. *Let $f, g : [a, b] \to \mathbb{R}$ be two functions such that f is bounded and g is increasing. Then,*

$$\underline{\int_a^b} f \, dg \leq \overline{\int_a^b} f \, dg.$$

Proof. Let P_1 and P_2 be two partitions of $[a, b]$ so that $P_1 \subset P_2$. Then, by Lemma 1.3.12,

$$L(P_1, f, g) \leq L(P_2, f, g) \leq U(P_2, f, g) \leq U(P_1, f, g).$$

Therefore, $L(P_1, f, g) \leq U(P_2, f, g)$ and, by taking the infimum over P_2,

$$L(P_1, f, g) \leq \overline{\int_a^b} f \, dg.$$

Finally, by taking the supremum over P_1,

$$\underline{\int_a^b} f \, dg \leq \overline{\int_a^b} f \, dg.$$

$\square$

Proposition 1.3.14. *If $f \in \mathcal{C}([a,b])$ and g is increasing on $[a,b]$, then $f \in RS([a,b],g)$.*

Proof. Given $\epsilon > 0$, there exists $\delta > 0$ such that $|x - x'| < \delta$ implies that

$$|f(x) - f(x')| < \frac{\epsilon}{g(b) - g(a)},$$

since f is uniformly continuous on $[a,b]$. Then, for any partition P of $[a,b]$ with $||P|| < \delta$, we have that

$$\begin{aligned} 0 \leq U(P,f,g) - L(P,f,g) &= \sum_{i=1}^{n} (M_i - m_i)[g(x_i) - g(x_{i-1})] \leq \sum_{i=1}^{n} (M_i - m_i)\Delta g_i \\ &< \frac{\epsilon}{g(b) - g(a)} \sum_{i=1}^{n} \Delta g_i = \epsilon. \end{aligned}$$

If $g(b) = g(a)$, then g is constant on $[a,b]$ and $U = L = 0$.
According to Lemma 1.3.13, we obtain that

$$0 \leq \overline{\int_a^b} f \ dg - \underline{\int_a^b} f \ dg \leq U(P,f,g) - L(P,f,g) < \epsilon.$$

Since the inequality holds for all $\epsilon > 0$,

$$\underline{\int_a^b} f \ dg = A = \overline{\int_a^b} f \ dg.$$

Finally, note that

$$L(P,f,g) \leq P(f,g,\mu) \leq U(P,f,g)$$

implies that

$$|P(f,g,\mu) - A| < \epsilon$$

whenever $||P|| < \delta$ for some $\delta > 0$. Hence, $f \in RS([a,b],g)$. □

Proposition 1.3.15. *Let $c \in \mathbb{R}$ and $f \in RS([a,b],g_i)$ with $i = 1,2$. Then,*

$$\int_a^b f \ d(cg_1 + g_2) = c\int_a^b f \ dg_1 + \int_a^b f \ dg_2.$$

Proof. Let P be a partition of $[a,b]$,

$$P(f, cg_1 + g_2, \mu) = \sum_k f(\mu_k)[c\Delta g_{1,k} + \Delta g_{2,k}] = cP(f,g_1,\mu) + P(f,g_2,\mu).$$

Then,

$$\begin{aligned} \left| P(f, cg_1 + g_2, \mu) - c\int_a^b f \ dg_1 - \int_a^b f \ dg_2 \right| &\leq |c| \left| P(f,g_1,\mu) - \int_a^b f \ dg_1 \right| \\ &+ \left| P(f,g_2,\mu) - \int_a^b f \ dg_2 \right| \to 0 \text{ as } ||P|| \to 0. \end{aligned}$$

□

Corollary 1.3.16. *If $f \in \mathcal{C}([a,b])$ and $g \in BV([a,b])$, then $f \in RS([a,b],g)$.*

Proof. By Proposition 1.3.5, g is the difference of two increasing functions. By Proposition 1.3.14 and Proposition 1.3.15, the statement is proved. □

Example 1.3.17. Consider the Heaviside function H_a on $[a,b]$ defined by

$$H_a(x) = \begin{cases} 0, & x = a, \\ 1, & a < x \leq b. \end{cases}$$

For all $f \in \mathcal{C}([a,b])$,

$$\int_a^b f \ dH_a(x) = f(a).$$

Indeed, given $\epsilon > 0$, consider $\delta > 0$ so that $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \delta$, since f is uniformly continuous on $[a,b]$. Then, for any partition P with $||P|| < \delta$,

$$|f(a) - P(f, H_a, \mu)| = \left| f(a) - \sum_{k=1}^{n} f(\mu_k)[H_a(x_k) - H_a(x_{k-1})] \right| = |f(a) - f(\mu_1)| < \epsilon.$$

Proposition 1.3.18. *If $f \in \mathcal{C}([a,b])$ and $g \in BV([a,b])$, then*

$$\left| \int_a^b f \ dg \right| \leq ||f||_\infty V(g).$$

Proof. Given $\epsilon > 0$, let P be a partition of $[a,b]$ so that

$$\left| \int_a^b f \ dg - P(f, g, \mu) \right| < \epsilon.$$

Then,

$$\left| \int_a^b f \ dg \right| \leq |P(f,g,\mu)| + \epsilon \leq \sum_k \left| f(\mu_k)[g(x_k) - g(x_{k-1})] \right| + \epsilon \leq ||f||_\infty V(g) + \epsilon.$$

Since the inequality holds for all $\epsilon > 0$,

$$\left| \int_a^b f \ dg \right| \leq ||f||_\infty V(g).$$

□

Chapter 2

Fundamental theorems of Functional Analysis

This chapter is central to this thesis, given that it contains some basic theorems of Functional Analysis and Baire's theorem. These theorems are divided into three different sections; in the first one, Baire's theorem, the open mapping theorem and the closed graph theorem are studied together because their proofs are related. The second section includes theorems about sequences of bounded linear operators such as the uniform boundedness principle and the Banach-Steinhaus theorem. Finally, the last section contains different versions of the Hahn-Banach theorem, which are about extension and separation properties.

2.1 Baire's theorem, the open mapping theorem and the closed graph theorem

2.1.1 Baire's theorem

Theorem 2.1.1 (Baire's theorem). *Let X be a complete metric space. If $\{G_n\}_{n\geq 1}$ is a sequence of dense and open subsets of X, then $A = \cap_{n=1}^{\infty} G_n$ is also dense.*

Proof. By definition, A is dense in X if and only if $\bar{A} = X$, that is, for all $x \in X$ and all $r > 0$ $B(x,r) \cap A \neq \emptyset$. This is equivalent to prove that $A \cap G \neq \emptyset$ for every nonempty open set G in X.
Since G_1 is dense and open, $G_1 \cap G$ is a nonempty open set. Therefore, there exist $a_1 \in G_1 \cap G$ and $r_1 > 0$ so that $\overline{B(a_1, r_1)} \subset G_1 \cap G$. Similarly, G_2 is dense and open, so there exist $a_2 \in G_2 \cap B(a_1, r_1)$ and $0 < r_2 < r_1/2$ so that $\overline{B(a_2, r_2)} \subset G_2 \cap B(a_1, r_1)$. By induction, we can build the sequences $\{a_n\}_{n\geq 1} \subset X$ and $\{r_n\}_{n\geq 1} \subset \mathbb{R}^+$ with

$$0 < r_{n+1} < \tfrac{r_n}{2} = \tfrac{r_1}{2^n} \text{ and } \overline{B(a_{n+1}, r_{n+1})} \subset G_{n+1} \cap B(a_n, r_n).$$

Besides, $\{a_n\}_{n\geq 1}$ is a Cauchy sequence in X. Indeed, given $n, m \in \mathbb{N}$ with $m < n$, then $a_n \in \overline{B(a_m, r_m)}$ and $d(a_n, a_m) \leq r_m < \frac{r_1}{2^{m-1}} \to 0$ as $n, m \to \infty$. Since X is complete, there exists $a = \lim_{n\to\infty} a_n$ in X.
Finally, it is readily shown that $a \in A \cap G$, that is, $a \in G \cap G_m$ for all $m \in \mathbb{N}$.

Indeed, $a_n \in \overline{B(a_m, r_m)}$ whenever $n \geq m$, together with $a = \lim_{n\to\infty} a_n$, implies that $a \in \overline{B(a_m, r_m)} \subset G_m$ for all $m \in \mathbb{N}$. Besides, $a \in \overline{B(a, r_1)} \subset G$ and, hence, $A \cap G \neq \emptyset$. □

Corollary 2.1.2. *Let $X = \cup_{n=1}^{\infty} F_n$ be a complete metric space and $\{F_n,\ n \in \mathbb{N}\}$ a sequence of closed sets in X. Then, there is one F_n with nonempty interior.*

Proof. Since $X = \cup_{n=1}^{\infty} F_n$, $\emptyset = X^c = \cap_{n=1}^{\infty} F_n^c$, where the sets F_n^c are open. Baire's theorem states that there is at least one F_n^c not dense. Thus, $\overline{F_n^c} \neq X$ and, consequently, $X \setminus \overline{F_n^c} = \operatorname{int}(F_n) \neq \emptyset$. □

2.1.2 The open mapping theorem

Definition 2.1.3. A linear operator $T : E \to F$ is said to be *open* if $T(G)$ is an open set in F for any open set G in E.

Theorem 2.1.4 (open mapping theorem)**.** *Let E, F be two Banach spaces and $T : E \to F$ a surjective continuous linear operator. Then, T is an open mapping.*

Proof. We want to prove that $T(G)$ is an open set in F for any open set G in E.

1. *It is enough to prove that $T(B(0, r))$ is a neighborhood of zero in F for all $r > 0$.* Let $G \subset E$ be an open set. Since T is surjective, we consider $Ta \in T(G)$ with $a \in G$. Since G is open, there is $r > 0$ so that $B(a, r) = a + B(0, r) \subset G$. By linearity, $T(B(a, r)) = Ta + T(B(0, r)) \subset T(G)$. The hypothesis assures that $T(B(0, r))$ is a neighborhood of zero, so $T(B(a, r))$ is a neighborhood of Ta in F. Hence, $T(G)$ is open.

2. *For all $r > 0$, $\overline{T(B(0, r))}$ is a neighborhood of zero in F, that is, there is $\sigma > 0$ so that $B(0, \sigma) \subset \overline{T(B(0, r))}$.*
   Consider the following expressions,
$$E = \bigcup_{n=1}^{\infty} B(0, nr/2) \text{ and } F = T(E) = T\left(\bigcup_{n=1}^{\infty} B(0, nr/2)\right) = \bigcup_{n=1}^{\infty} T\big(B(0, nr/2)\big).$$
   Note that $F \subset \cup_{n=1}^{\infty} \overline{T(B(0, nr/2))} \subseteq \overline{\cup_{n=1}^{\infty} T(B(0, nr/2))} = \bar{F} = F$. Hence,
$$F = \bigcup_{n=1}^{\infty} \overline{T(B(0, nr/2))}.$$
   By Corollary 2.1.2, there is $N \in \mathbb{N}$ such that $\operatorname{int}(\overline{T(B(0, Nr/2))} \neq \emptyset$. We can assume $N = 1$ because $\overline{T(B(0, Nr/2))} = N \cdot \overline{T(B(0, r/2))} \cong \overline{T(B(0, r/2))}$. Hence, there exist $y \in F$ and $\sigma > 0$ so that
$$B(y, \sigma) = y + B(0, \sigma) \subseteq \overline{T(B(0, r/2))}.$$
   Besides, there exists a sequence $\{x_n\}_n \subset B(0, r/2)$ such that $y = \lim_n Tx_n$ and also $-y = \lim_n T(-x_n)$. Therefore, $-y \in \overline{T(B(0, \frac{r}{2}))}$. Finally, we have that
$$B(0, \sigma) \subseteq -y + \overline{T(B(0, r/2))} \subseteq \overline{T(B(0, r/2))} + \overline{T(B(0, r/2))} \subseteq \overline{T(B(0, r))}.$$

3. *Fixed $s > 0$, $T(B(0,s))$ is a neighborhood of zero in F.*
   We write $s = \sum_{n=1}^{\infty} r_n$ with $r_n > 0$ (obviously, $r_n \to 0$ as $n \to \infty$). According to the second step of this proof, for all $n \geq 1$ there exists $\sigma_n > 0$ such that $B(0, \sigma_n) \subset \overline{T(B(0, r_n))}$. We can assume that $\sigma_n \downarrow 0$.
   Let $y \in B(0, \sigma_1) \subseteq \overline{T(B(0, r_1))}$. Since T is surjective, there exists $x_1 \in B(0, r_1)$ so that
   $$||y - Tx_1||_F < \sigma_2.$$
   It follows that $y - Tx_1 \in B(0, \sigma_2) \subset \overline{T(B(0, r_2))}$. Then, there exists $x_2 \in B(0, r_2)$ so that
   $$||y - Tx_1 - Tx_2||_F < \sigma_3.$$
   By induction, if
   $$y - Tx_1 - \ldots - Tx_{n-1} \in B(0, \sigma_n) \subset \overline{T(B(0, r_n))},$$
   then there exists $x_n \in B(0, r_n)$ so that
   $$||y - Tx_1 - \ldots - Tx_{n-1} - Tx_n||_F < \sigma_{n+1}.$$
   Since E is a Banach space and
   $$\sum_{n=1}^{\infty} ||x_n||_E < \sum_{n=1}^{\infty} r_n = s < \infty,$$
   there exists $x = \sum_{n=1}^{\infty} x_n \in E$. Note that $||x||_E \leq \sum_{n=1}^{\infty} ||x_n||_E < s$ implies that $x \in B(0, s)$. Since T is continuous,
   $$y = \lim_{n \to \infty} T\left(\sum_{k=1}^{n} x_k \right) = Tx \in T(B(0, s)).$$
   Hence, $B(0, \sigma_1) \subset T(B(0, s))$ and $T(B(0, s))$ is a neighborhood of zero. □

Corollary 2.1.5 (Banach isomorphism theorem)**.** *Let E, F be two Banach spaces and $T : E \to F$ a bijective continuous linear operator. Then, T^{-1} is also a bijective continuous linear operator. In particular, T is an isomorphism.*

Proof. By the open mapping theorem, T is open. Since T is bijective and open, there exists T^{-1} and it is continuous. □

2.1.3 The closed graph theorem

Theorem 2.1.6 (closed graph theorem)**.** *Let E, F be two Banach spaces over $\mathbb{K}$ and T a linear operator between E and F. Then, $G(T) = \{(x, y) \in E \times F : y = Tx\}$ is a closed set in $E \times F$ if and only if T is continuous.*

Proof. First, assume that $G(T)$ is closed. The linearity of T implies that $G(T) \subseteq E \times F$ is a subspace. Indeed, given $(x, Tx), (z, Tz) \in G(T)$ and $\lambda, \mu \in \mathbb{K}$, we have
$$\lambda(x, Tx) + \mu(z, Tz) = (\lambda x + \mu z, \lambda Tx + \mu Tz) = (\lambda x + \mu z, T(\lambda x + \mu z)).$$

Thus, $\lambda(x,Tx)+\mu(z,Tz)\in G(T)$. By Proposition 1.1.12, $E\times F$ is Banach and, by Proposition 1.1.4, so is $G(T)$. Consider

$$\begin{aligned}\pi_E : G(T) &\longrightarrow E\\ (x,Tx) &\mapsto x\end{aligned}$$

a bijective continuous linear operator. By Corollary 2.1.5, π_E^{-1} is continuous. Then, $T=\pi_F\circ\pi_E^{-1}$ is continuous because it is the composition of two continuous operators. Conversely, assume that T is continuous. We consider a sequence $\{(x_n,Tx_n)\}_n\subset G(T)$ convergent to $(x,y)\in E\times F$. Then,

$$\begin{aligned}||x_n-x||_E &\leq ||(x_n,Tx_n)-(x,y)||_{E\times F}\to 0 \text{ as } n\to\infty, \text{ and}\\ ||Tx_n-y||_F &\leq ||(x_n,Tx_n)-(x,y)||_{E\times F}\to 0 \text{ as } n\to\infty.\end{aligned}$$

We have obtained that $\{x_n\}_n$ converges to x in E and $\{Tx_n\}_n$ converges to y in F. Since T is continuous, $Tx=y$. Hence, $(x,y)\in G(T)$, i.e., $G(T)$ is closed. □

2.2 The uniform boundedness principle and the Banach-Steinhaus theorem

Definition 2.2.1. Let E be a Banach space and A a subset of E. A is *$\mathcal{G}_\delta$-dense* if and only if it is a countable intersection of open sets.

Theorem 2.2.2 (uniform boundedness principle)**.** *Let E,F be two Banach spaces and $\{T_i, i\in I\}$ a family of bounded linear operators between E and F. Then*

- *(a) either $\sup_{i\in I}||T_i||=M<\infty$, or*
- *(b) there is a $\mathcal{G}_\delta$-dense set A in E such that $\sup_{i\in I}||T_i(x)||_F=\infty$ for all $x\in A$.*

Proof. For all $n\in\mathbb{N}$, we define

$$G_n=\big\{x\in E:\ \sup_{i\in I}||T_i(x)||_F>n\big\}=\bigcup_{i\in I}\big\{x\in E:\ ||T_i(x)||_F>n\big\}.$$

The sets G_n, $n\in\mathbb{N}$, are open, since T and the norm are continuous. We consider two cases.

(a) There exists a set G_m not dense in E. In this case, we can find a ball $B_E(a,r)$ such that $\overline{B_E(a,r)}\cap G_m=\emptyset$. From this, it follows that

$$\sup_{i\in I}||T_i(x+a)||_F\leq m \quad \text{whenever } ||x||_E\leq r.$$

Then, for all $i\in I$,

$$||T_i(x)||_F\leq ||T_i(x+a)||_F+||T_i(a)||_F\leq 2m \quad \text{whenever } ||x||_E\leq r.$$

Given $y\in E\setminus\{0\}$, consider $x=\frac{r\cdot y}{||y||_E}$ (note that $||x||_E=r$). Then, for all $i\in I$,

$$||T_i(y)||_F = \frac{||y||_E}{r} ||T_i(x)||_F \leq \frac{2m}{r} ||y||_E.$$

Therefore, T_i is continuous for all $i \in I$ with $||T_i|| \leq \frac{2m}{r}$. Hence,

$$\sup_{i \in I} ||T_i|| \leq \frac{2m}{r} < \infty.$$

(b) Otherwise, G_n is dense for all $n \in \mathbb{N}$. Then, by Baire's theorem, the set $A = \cap_{n \geq 1} G_n$ is dense in E. Consequently, $\sup\limits_{i \in I} ||T_i(x)||_F = \infty$ for all $x \in A$. □

Corollary 2.2.3. *If $\{T_n\}_{n \geq 1}$ is a sequence of bounded linear operators between two Banach spaces E, F so that $T(x) := \lim_n T_n(x)$ for all $x \in E$, then T is a bounded linear operator between E and F so that $||T|| \leq \sup_n ||T_n|| < \infty$.*

Corollary 2.2.4 (Banach-Steinhaus theorem)**.** *Let $\{T_n\}_{n \geq 1}$ be a sequence of bounded linear operators between two Banach spaces E, F. Suppose that*

(1) there exists $T(x) := \lim_n T_n(x)$ for all $x \in D$, D a dense set in E, and

(2) the sequence $\{T_n(x)\}_{n \geq 1}$ is bounded for all $x \in E$.

Then, $T : D \to F$ defined by $T(x) := \lim_n T_n(x)$ extends to a bounded linear operator $T : E \to F$ such that

$$||T|| \leq \liminf_n ||T_n||.$$

Proof. The second hypothesis assures that, for all $x \in E$, there is $M_x > 0$ so that $||T_n(x)||_F \leq M_x$ for all $n \in \mathbb{N}$, that is, $\sup_n ||T_n(x)|| \leq M_x < \infty$. By the uniform boundedness principle, $\sup_n ||T_n|| = M < \infty$.
For all $x \in E$, we will show that $\{T_n(x)\}_{n \geq 1}$ is a Cauchy sequence. Since D is dense in E, for all $\epsilon > 0$ there exists $z \in D$ so that $||x - z|| < \frac{\epsilon}{4M}$. According to hypothesis (1), $\{T_n(z)\}_n$ is a Cauchy sequence and, then, given $\epsilon > 0$ there is $n \in \mathbb{N}$ so that $||T_p(z) - T_q(z)|| < \frac{\epsilon}{2}$ whenever $p, q > n$. Thus,

$$||T_p(x) - T_q(x)||_F \leq ||T_p(x) - T_p(z)||_F + ||T_p(z) - T_q(z)||_F + ||T_q(z) - T_q(x)||_F < \epsilon.$$

Since F is a Banach space, for all $x \in E$ there exists $T(x) := \lim_n T_n(x)$ in F. Therefore,

$$\begin{aligned} T : E &\to F \\ x &\mapsto T(x) := \lim_n T_n(x) \end{aligned}$$

is a well-defined linear operator. Finally,

$$||T(x)||_F = \lim_n ||T_n(x)||_F = \liminf_n ||T_n(x)||_F \leq \liminf_n ||T_n|| \, ||x||_E.$$

Hence, $||T|| \leq \liminf\limits_n ||T_n|| < \infty$. □

2.3 The Hahn-Banach theorem

2.3.1 Analytic version

Definition 2.3.1. A *convex functional* on a real vector space E is a function $p : E \to \mathbb{R}$ such that, for all $x, y \in E$ and all $\lambda \geq 0$,

(i) $p(x + y) \leq p(x) + p(y)$.

(ii) $p(\lambda x) = \lambda p(x)$.

Theorem 2.3.2 (Hahn-Banach theorem). *Let E be a real vector space, F a subspace of E, p a convex functional on E and u a linear form on F dominated by p. Then, there exists a linear form $v : E \to \mathbb{R}$ such that $v(y) = u(y)$ for all $y \in F$ and $v(x) \leq p(x)$ for all $x \in E$.*

Proof. If $F = E$, there is nothing to prove. Let us suppose that $F \neq E$, so there exists $y \in E \setminus F$ and we can consider

$$F \oplus [y] = \{z + \alpha y : z \in F \text{ and } \alpha \in \mathbb{R}\}.$$

Then, u can be extended to this subspace by

$$\begin{aligned} v : F \oplus [y] &\to \mathbb{R} \\ z + ty &\mapsto u(z) + t \cdot s \end{aligned}$$

where $s \in \mathbb{R}$ will be conveniently chosen. Note that v is well-defined (if $ty \in F$, then $t = 0$), extends u and is linear, since u is linear.
We are looking for a real number $s = v(y)$ so that $v(x) \leq p(x)$ for all $x \in F \oplus [y]$. Given $z, z' \in F$,

$$u(z) - u(z') = u(z - z') \leq p(z - z') \leq p(z + y) + p(-y - z'), \text{ and hence}$$
$$-p(-y - z') - u(z') \leq p(z + y) - u(z).$$

There exists $s \in \mathbb{R}$ so that

$$\sup_{z' \in F}\{-p(-y - z') - u(z')\} \leq s \leq \inf_{z \in F}\{p(z + y) - u(z)\}.$$

Then, v is dominated by p. Indeed, let $z + t \cdot y \in F \oplus [y]$.

(i) If $t > 0$, from $s \leq p(\frac{z}{t} + y) - u(\frac{z}{t})$, it follows that

$$v(z + ty) = u(z) + ts \leq u(z) + t[p(\tfrac{z}{t} + y) - u(\tfrac{z}{t})] = p(z + ty).$$

(ii) If $t < 0$, we define $\alpha := -t > 0$ and, according to the selection of s, it follows that

$$-p(-y + \tfrac{z}{\alpha}) - u(\tfrac{-z}{\alpha}) \leq s,$$
$$-p(-\alpha y + z) - u(-z) \leq \alpha s, \text{ and finally}$$
$$v(z + ty) = u(z) - \alpha s \leq p(-\alpha y + z) = p(z + ty).$$

So far, we have proved that a one-dimensional extension is always possible.
Now, (H, h) is said to be an extension of (F, u), $(F, u) \leq (H, h)$, if H is a subspace of E with $F \subset H$ and h is a linear form that extends u on H so that $h(x) \leq p(x)$ for all $x \in H$. Then, the set

$$\mathcal{F} = \{(H, h) : (H, h) \text{ is an extension of } (F, u)\}$$

is partially ordered and nonempty. Let us consider a chain $\mathcal{T} = \{(H_i, h_i)\}_{i \in I}$ in $\mathcal{F}$ (a totally ordered subset of $\mathcal{F}$) and we will show that $\mathcal{T}$ is upper bounded by an element of $\mathcal{F}$. We define $K = \cup_{i \in I} H_i$ and $k(x) = h_i(x)$ if $x \in H_i$.

(1) *K is a vector subspace of E.* Indeed, let $x_1, x_2 \in K$ and $\lambda, \mu \in \mathbb{R}$, there are $i_1, i_2 \in I$ such that $x_1 \in H_{i_1}$, $x_2 \in H_{i_2}$ and, since the order is total in $\mathcal{T}$, we can assume that $H_{i_1} \subseteq H_{i_2}$. Thus, $x_1, x_2 \in H_{i_2}$ and also $\lambda x_1 + \mu x_2 \in H_{i_2} \subseteq K$. Obviously, F is a subspace of K.

(2) *k is a linear form dominated by p that extends u on K.* First of all, k is well-defined: if $x \in H_{i_1}$ and $x \in H_{i_2}$, since the order is total in $\mathcal{T}$, we assume that $(H_{i_1}, h_{i_1}) \leq (H_{i_2}, h_{i_2})$. Then, $h_{i_1}(x) = h_{i_2}(x)$.
Secondly, k is linear: given $x_1, x_2 \in K$ and $\lambda, \mu \in \mathbb{R}$, as we have proceeded before, we assume $x_1, x_2, \lambda x_1 + \mu x_2 \in H_i$. Then,

$$k(\lambda x_1 + \mu x_2) = h_i(\lambda x_1 + \mu x_2) = \lambda h_i(x_1) + \mu h_i(x_2) = \lambda k(x_1) + \mu k(x_2).$$

Thirdly, if $y \in F \subseteq K$, there is H_i that contains y and then $k(y) = h_i(y) = u(y)$. Finally, if $x \in K$, there is H_i that contains x and then $k(x) = h_i(x) \leq p(x)$.

Hence, $(K, k) \in \mathcal{F}$ and, by construction, $(H_i, h_i) \leq (K, k)$ for all $i \in I$. According to Zorn's Lemma, $\mathcal{F}$ has at least one maximal element (V, v). It is clear that $V = E$, otherwise we could find $y \in E \setminus V$ and there would exist a one-dimensional extension of (V, v), which contradicts the fact that (V, v) is maximal. □

Definition 2.3.3. Let E be a real or complex vector space, $p : E \to [0, \infty)$ is a *seminorm* if and only if for all $x, y \in E$ and all $\lambda \in \mathbb{K}$,

(i) $p(x + y) \leq p(x) + p(y)$ and

(ii) $p(\lambda x) = |\lambda| p(x)$.

Remark 2.3.4. A seminorm is a convex functional.

The following version of the Hahn-Banach theorem is referred to real and complex vector spaces.

Theorem 2.3.5 (Hahn-Banach theorem)**.** *Let E be a real or complex vector space, p a seminorm on E, F a subspace of E and $u : F \to \mathbb{K}$ a linear form such that*

$$|u(y)| \leq p(y) \quad \textit{for all } y \in F.$$

Then, there exists a linear form $v : E \to \mathbb{K}$ so that $v(z) = u(z)$ for all $z \in F$ and $|v(x)| \leq p(x)$ for all $x \in E$.

Proof. We distinguish two cases.
The real case: p is a seminorm and, by Theorem 2.3.2, there exists a linear form $v : E \to \mathbb{R}$ that extends u, and such that $v(x) \leq p(x)$ and $-v(x) = v(-x) \leq p(-x) = p(x)$ for all $x \in E$. Hence, $|v(x)| \leq p(x)$ for all $x \in E$.
The complex case: first, we regard E as real vector space and we consider $u_1 = \mathrm{Re}(u)$ a real linear form on F such that $|u_1(x)| \leq |u(x)| \leq p(x)$ for all $x \in F$. From the real case, there exists a real linear form $v_1 : E \to \mathbb{R}$ such that $v_1|_F = u_1$ and $|v_1(z)| \leq p(z)$ for all $z \in E$.
Regarding E as a complex vector space, we define the complex form $v(z) = v_1(z) - iv_1(iz)$, which is linear: for all $x, y \in E$ and all $\lambda \in \mathbb{R}$,

(i) $$\begin{aligned} v(x+y) &= v_1(x+y) - iv_1(i(x+y)) = v_1(x) + v_1(y) - i(v_1(ix) + v_1(iy)) \\ &= v(x) + v(y). \end{aligned}$$

(ii) $v(ix) = v_1(ix) - iv_1(-x) = v_1(ix) + iv_1(x) = i[-iv_1(ix) + v_1(x)] = iv(x)$.

(iii) $v(\lambda x) = v_1(\lambda x) - iv_1(\lambda x) = \lambda v(x)$.

Next, we show that v extends u on E. Since $\mathrm{Re}(v)|_F = v_1|_F = u_1 = \mathrm{Re}(u)$,

$$\mathrm{Im}((v(x)) = -v_1(ix) = -u_1(ix) = -\mathrm{Re}(u(ix)) = -\mathrm{Re}(iu(x)) = \mathrm{Im}(u(x))$$

for all $x \in F$. Hence, $v|_F = u$.
Finally, we want to prove that $|v|$ is dominated by p. Note that for a given $x \in E$, $|v(x)| = \lambda_x v(x)$ with $|\lambda_x| = 1$. Therefore,

$$|v(x)| = \lambda_x v(x) = v(\lambda_x x) = v_1(\lambda_x x) \leq |v_1(\lambda_x x)| \leq p(x).$$

□

Theorem 2.3.6 (Hahn-Banach theorem in normed vector spaces)**.** *Let E be a normed vector space.*

(a) If F is a subspace of E and u a continuous linear form on F, then there exists a continuous linear form v on E which extends u and $||u|| = ||v||$.

(b) For all $a \in E$, there exists a linear form on E such that $v(a) = ||a||_E$ and $||v|| = 1$.

Proof. (a) The function $p : E \to [0, +\infty)$ defined by $p(x) = ||u||\ ||x||_E$ is a seminorm on E. Since u is continuous, $|u(z)| \leq ||u||\ ||z||_F$ for all $z \in F$. By Theorem 2.3.5, there exists a linear form v on E which extends u and such that, for all $x \in E$,

$$|v(x)| \leq ||u||\ ||x||_E.$$

Therefore, v is continuous and $||v|| \leq ||u||$. Besides,

$$||u|| = \sup_{x \in F \setminus \{0\}} \frac{|u(x)|}{||x||_E} \leq \sup_{z \in E \setminus \{0\}} \frac{|v(z)|}{||z||_E} = ||v||.$$

Hence, $||u|| = ||v||$.
(b) Suppose that $a \neq 0$, otherwise there is nothing to prove, and let u be a linear form on $[a]$ defined by $u(\lambda a) = \lambda ||a||_E$, with $\lambda \in \mathbb{K}$. Notice that $|u(\lambda a)| \leq ||\lambda a||_E$ for all $\lambda a \in [a]$. By Theorem 2.3.5, there exists a linear form v on E that extends u and $|v(x)| \leq ||x||_E$ for all $x \in E$. Hence, $||v|| \leq 1$. Since $v(a) = u(a) = ||a||_E$, $||v|| = 1$. □

2.3.2 Geometric version

Definition 2.3.7. Let K be a convex subset of a real vector space E containing the origin. Then, *the Minkowski functional of* K is defined as

$$\begin{aligned} p_K &: E \to [0, +\infty] \\ x &\mapsto p_K(x) = \inf\left\{t > 0 : \ \frac{x}{t} \in K\right\}. \end{aligned}$$

Lemma 2.3.8. *The Minkowski functional is a convex functional.*

Proof. It is clear that $p_K(\lambda x) = \lambda p_K(x)$ for all $\lambda \geq 0$. To prove the subadditivity, we consider $x, y \in E$ and note that if $p_K(x) = \infty$ or $p_K(y) = \infty$, there is nothing to prove. If $\frac{x}{s}, \frac{y}{t} \in K$ with $s, t > 0$, then, since K is convex,

$$\frac{x+y}{t+s} = \frac{s}{t+s}\frac{x}{s} + \frac{t}{t+s}\frac{y}{t} \in K.$$

Hence, $p_K(x+y) \leq t+s$ and, by taking the infimum with respect to t, $p_K(x+y) \leq p_K(x)+s$. Finally, by taking the infimum with respect to s, we conclude that

$$p_K(x+y) \leq p_K(x) + p_K(y).$$

□

Remark 2.3.9. Let K be a convex subset of a real normed vector space E containing the origin. If x is an internal point of K, then $p_K(x) < 1$. Indeed, since x is an internal point, there is $\delta > 0$ such that $(1+\delta)x \in K$ and it follows that $p_K(x) \leq \frac{1}{1+\delta} < 1$.

Remark 2.3.10. Let K be a convex set containing the origin. If $x \notin K$, then $p_K(x) \geq 1$. Indeed, suppose that $p_K(x) < 1$, i.e., there is $0 < \lambda < 1$ such that $\frac{x}{\lambda} \in K$. Since K is convex, $x = \lambda\frac{x}{\lambda} + (1-\lambda)0 \in K$, which contradicts the hypothesis.

Lemma 2.3.11. *Let K be a convex subset of a real normed vector space E with nonempty interior. Then, for all $y \in E \setminus K$ there is a nonzero linear form $f : E \to \mathbb{R}$ such that*

$$f(K) \leq f(y).$$

Besides, if K is open, then f can be chosen so that

$$f(K) < f(y).$$

Proof. We consider the case that the origin is an internal point of K. Let $y \in E \setminus K$, by Remark 2.3.10, $p_K(y) \geq 1$. We define

$$\begin{aligned} f &: [y] \to \mathbb{R} \\ ty &\mapsto t \end{aligned}$$

a nonzero linear form on $[y]$ such that

$$f(ty) = t \leq t \cdot p_K(y) = p_K(ty) \quad \text{for all } t > 0.$$

If $t < 0$, the inequality follows immediately. By Theorem 2.3.2, there is a linear form $\bar{f}$ that extends f on E and $\bar{f}(x) \le p_K(x)$ for all $x \in E$. Then, for all $x \in K$

$$\bar{f}(x) \le p_K(x) \le 1 = f(y) = \bar{f}(y).$$

If K is open, then $p_K(x) < 1$ for all $x \in K$, as we have shown in Remark 2.3.9, and $\bar{f}(x) \le p_K(x) < 1 = \bar{f}(y)$.
Now suppose that 0 is not an internal point of K. We know that there is an internal point $x_0 \neq 0$ in K. Consider

$$C = \{\bar{x} := x - x_0 : x \in K\} \text{ and } \bar{y} = y - x_0 \text{ with } y \notin K.$$

Since $0 \in \operatorname{int}(C)$ and $\bar{y} \notin C$, the argument above assures that there exists a linear form f on E such that

$$f(\bar{x}) \le f(\bar{y}) \quad \text{for all } \bar{x} \in C.$$

Hence, $f(x) \le f(y)$ for all $x \in K$. Similarly, the inequality is strict if K is open. □

Definition 2.3.12. Let E be a real normed vector space and consider two disjoint subsets A and B in E. A and B are said to be *separated* if there is a nonzero linear form f on E such that

$$f(A) < f(B).$$

Furthermore, if

$$\sup f(A) < \inf f(B)$$

we say that A and B are *strictly separated*.

Theorem 2.3.13 (geometric form of the Hahn-Banach theorem)**.** *Let A and B be two nonempty disjoint convex subsets of a real normed vector space E.*

(a) If one of them is open, then A and B are separated.

(b) If A and B are closed and one of them is compact, then A and B are strictly separated.

Proof. Let A be the open set. We define

$$K = A - B = \bigcup_{y \in B} \{x - y : x \in A\},$$

an open set, since it is the union of open sets. Let $z_i = x_i - y_i \in K$ with $x_i \in A$ and $y_i \in B$ for $i = 1, 2$, we show that K is a convex set: for all $0 < \lambda < 1$,

$$\lambda z_1 + (1-\lambda) z_2 = \lambda(x_1 - y_1) + (1-\lambda)(x_2 - y_2) = \lambda x_1 + (1-\lambda) x_2 - (\lambda y_1 + (1-\lambda) y_2) \in K,$$

since $\lambda x_1 + (1-\lambda)x_2 \in A$ and $\lambda y_1 + (1-\lambda) y_2 \in B$. Besides, note that $0 \notin K$ because $A \cap B = \emptyset$. By Lemma 2.3.11, there exists a nonzero linear form $f : E \to \mathbb{R}$ such that

$$f(z) < f(0) = 0 \quad \text{for all } z = x - y \in K \text{ with } x \in A \text{ and } y \in B.$$

Consequently, $f(x) < f(y)$ for all $x \in A$ and all $y \in B$. Besides, there is $\gamma \in \mathbb{R}$ such that

$$\sup_{x\in A} f(x) \leq \gamma \leq \inf_{y\in B} f(y).$$

To prove (b), suppose that B is the compact set and consider for any $r > 0$,

$$A(r) := \bigcup_{x\in A} B(x,r) \text{ and } B(r) := \bigcup_{y\in B} B(y,r),$$

two nonempty open sets. It is readily shown that they are convex: given $z_i \in A(r)$, there exist $x_i \in A$ such that $z_i \in B(x_i, r)$ for $i = 1, 2$. For all $0 < \lambda < 1$, $\lambda x_1 + (1-\lambda)x_2 \in A$, since A is convex. Then, $B(\lambda x_1 + (1-\lambda)x_2, r) \subset A(r)$ and we obtain

$$||\lambda z_1 + (1-\lambda)z_2 - \lambda x_1 - (1-\lambda)x_2|| \leq \lambda||z_1 - x_1|| + (1-\lambda)||z_2 - x_2|| < \lambda r + (1-\lambda)r = r.$$

Hence, $A(r)$ is convex. Similarly, $B(r)$ is also convex.
Furthermore, there exists $r_0 > 0$ such that $A(r) \cap B(r) = \emptyset$ for all $r \leq r_0$. Indeed, assume the opposite case: for all $r > 0$ $A(r) \cap B(r) \neq \emptyset$. Since $A \cap B = \emptyset$, it must exist $\{x_n\}_n \subset A$, $\{y_n\}_n \subset B$ and $\{z_n\}_n \subset E$ such that

$$x_n = y_n + z_n \text{ for all } n \geq 1 \text{ and } z_n \to 0 \text{ as } n \to \infty.$$

Since B is compact, there is a subsequence $\{y_{n_k}\}_k$ that converges in B. Hence, $\{x_{n_k}\}_k$ also converges and $\lim_k x_{n_k} \in A$, since A is closed. Thus, $\lim_k x_{n_k} = \lim_k y_{n_k} \in A \cap B$, which is a contradiction.
Therefore, by (a), the two sets $A(r_0)$ and $B(r_0)$ are separated. Then, there exist a linear form $f : E \to \mathbb{R}$ and $\gamma \in \mathbb{R}$ such that

$$f(x+v) \leq \gamma \leq f(y+w)$$

for all $x \in A$, all $y \in B$ and all $v, w \in E$ with $||v|| = ||w|| = \frac{r_0}{2}$. Hence,

$$f(x) + \tfrac{r_0}{2}||f|| = f(x) + \sup_{||v||=\frac{r_0}{2}} f(v) \leq \gamma \leq f(y) + \inf_{||w||=\frac{r_0}{2}} f(w) = f(y) - \frac{r_0}{2}||f||$$

for all $x \in A$ and all $y \in B$. Since $||f||\frac{r_0}{2} > 0$, it follows that for all $x \in A$ and all $y \in B$,

$$f(x) \leq \gamma - ||f||\tfrac{r_0}{2} < \gamma + ||f||\tfrac{r_0}{2} \leq f(y).$$

□

Remark 2.3.14. For a complex normed vector space E, separation refers to E as a real normed vector space. Note that if f is a linear real form such that $f(A) < f(B)$, then $u(x) = f(x) - if(ix)$ is a complex linear form such that $Re(u) = f$.

Chapter 3

Applications to areas of Analysis

In this chapter, some important results related to different areas of Analysis are examined, showing the great variety of applications that the fundamental theorems in Chapter 2 have. Each section is associated with one area of Analysis, including Real Analysis, Functional Analysis, Harmonic Analysis, Numerical Analysis and Differential Equations.

3.1 Applications to Real Analysis

3.1.1 Existence of nowhere differentiable continuous functions

In 1872, Karl Weierstrass provided an example of a continuous function on $[0,1]$ that was nowhere differentiable. In this subsection, the existence of such a function is proved without giving any specific example. Even more, we prove that these functions are dense among the continuous ones.

Theorem 3.1.1. *There exists $f \in \mathcal{C}([a,b])$ which is nowhere differentiable.*

Proof. First, for all $n \in \mathbb{N}$ we define the following sets,

$$F_n := \left\{ f \in \mathcal{C}([a,b]) : \text{ there exists } c \in [a,b] \text{ so that } \sup_{h \neq 0} \left| \frac{f(c+h)-f(c)}{h} \right| \leq n \right\}.$$

In this proof, we will assume that h is taken so that $c+h \in [a,b]$.

(1) *If $f \in \mathcal{C}([a,b])$ is differentiable at at least one point $c \in [a,b]$, then $f \in \cup_{n\in\mathbb{N}} F_n$.*

Since f is differentiable at $c \in [a,b]$, given $\epsilon > 0$ there exists $h_0 > 0$ so that, for all $0 < |h| < h_0$,

$$\left| \frac{f(c+h)-f(c)}{h} \right| \leq \left| \frac{f(c+h)-f(c)}{h} - f'(c) \right| + |f'(c)| < \epsilon + |f'(c)| < \infty.$$

On the other hand, for all $|h| \geq h_0$,

$$\left| \frac{f(c+h)-f(c)}{h} \right| \leq \frac{2}{h_0} ||f|| < \infty.$$

Consequently, there exists $n_0 \in \mathbb{N}$ so that

$$\sup_{h\neq 0}\left|\frac{f(c+h)-f(c)}{h}\right| \leq n_0.$$

Hence, $f \in F_{n_0} \subset \cup_{n\in\mathbb{N}} F_n$.

(2) *For all $n \in \mathbb{N}$, F_n is a closed set.*

Let $\{f_k\}_k \subset F_n$ be a uniformly convergent sequence to f, we will show that $f \in F_n$. For all $k \in \mathbb{N}$ there exists at least one $c_k \in [a,b]$ so that

$$\sup_{h\neq 0}\left|\frac{f_k(c_k+h)-f_k(c_k)}{h}\right| \leq n.$$

Since $\{c_k\}_k \subset [a,b]$, by the Weierstrass theorem, there exists a subsequence $\{c_{k_j}\}_j$ convergent to $c \in [a,b]$. Given $h \neq 0$ with $c+h \in [a,b]$, let $h_j = h + c - c_{k_j}$ for all $j \in \mathbb{N}$. Then, there exists $j_0 \in \mathbb{N}$, so that $h_j \neq 0$ for all $j \geq j_0$, since $\{h_j\}_j$ is convergent to h. Besides, we have the subsequence $\{f_{k_j}\}_j$ so that

$$|f_{k_j}(c_{k_j}+h_j) - f(c+h)| = |f_{k_j}(c_{k_j}+h_j) - f(c_{k_j}+h_j)| \leq ||f_{k_j} - f|| \to 0, \text{ and}$$
$$|f_{k_j}(c_{k_j}) - f(c)| \leq |f_{k_j}(c_{k_j}) - f(c_{k_j})| + |f(c_{k_j}) - f(c)| \leq ||f_{k_j} - f|| + |f(c_{k_j}) - f(c)| \to 0,$$

as $j \to \infty$, since $\{f_{k_j}\}$ converges uniformly to f, $\{c_{k_j}\}_j$ converges to c and f is uniformly continuous on $[a,b]$. Consequently,

$$\left|\frac{f(c+h)-f(c)}{h}\right| = \lim_{\substack{j\to\infty \\ j\geq j_0}}\left|\frac{f_{k_j}(c_{k_j}+h_j)-f(c_{k_j})}{h_j}\right| \leq n.$$

Hence, $f \in F_n$.

(3) *For all $n \in \mathbb{N}$,* $\mathrm{int}(F_n) = \emptyset$.

Given $f \in F_n$ and $\epsilon > 0$, we will show that there exists a function $g \in B(f,\epsilon)$ such that $g \notin F_n$. By the Stone-Weierstrass theorem, there exists a polynomial p so that $||f - p|| < \frac{\epsilon}{2}$. The fact that $p \in \mathcal{C}^\infty([a,b])$ will be important from now on. Indeed, $|p(x) - p(y)| < \frac{\epsilon}{16}$ whenever $|x - y| < \delta$ for some $\delta > 0$, and there exists a constant $M > 0$ such that $|p'(x)| \leq M$ for all $x \in [a,b]$. Next, we choose $h > 0$ so that

$$h < \min\left\{\delta, \frac{\epsilon}{4(M+n)}\right\}.$$

Let $P = \{a = x_0 < x_1 < \ldots < x_k = b\}$ be a partition with $||P|| \leq h$. We consider $g : [a,b] \to \mathbb{R}$ a piecewise affine function defined on each interval $[x_i, x_{i+1}]$, $0 \leq i \leq k-1$, by

$$g(x_i) = p(x_i) + (-1)^i\frac{\epsilon}{8},$$
$$g(x_{i+1}) = p(x_{i+1}) + (-1)^{i+1}\frac{\epsilon}{8}, \text{ and}$$
$$g(x) = \frac{x_{i+1}-x}{x_{i+1}-x_i}g(x_i) + \frac{x-x_i}{x_{i+1}-x_i}g(x_{i+1}),\ x_i < x < x_{i+1}.$$

Clearly, $g \in \mathcal{C}([a,b])$ and it is differentiable on $[a,b]$ except at $\{x_1,\ldots,x_{k-1}\}$. Besides, given $x \in [a,b]$, assume $x \in [x_i, x_{i+1}]$ for some $0 \le i \le k-1$, we have that

$$\begin{aligned}|g(x)-p(x)| &\le |g(x)-g(x_i)|+|g(x_i)-p(x_i)|+|p(x_i)-p(x)| \\ &< |g(x_{i+1})-g(x_i)|+\frac{\epsilon}{8}+\frac{\epsilon}{16} \le |p(x_{i+1})-p(x_i)|+\frac{\epsilon}{4}+\frac{\epsilon}{8}+\frac{\epsilon}{16} \\ &< \frac{\epsilon}{16}+\frac{\epsilon}{4}+\frac{\epsilon}{8}+\frac{\epsilon}{16}=\frac{\epsilon}{2}.\end{aligned}$$

Consequently, $||g-p||< \frac{\epsilon}{2}$ and $g \in B(f,\epsilon)$. Next, we show that $g \notin F_n$. For all $x \in (x_i, x_{i+1})$, $0 \le i \le k-1$, we have

$$g'(x)=\frac{g(x_{i+1})-g(x_i)}{x_{i+1}-x_i}=\frac{p(x_{i+1})-p(x_i)+(-1)^{i+1}\frac{\epsilon}{4}}{x_{i+1}-x_i}=p'(z_i)+\frac{(-1)^{i+1}\epsilon}{4(x_{i+1}-x_i)},$$

where, by the mean value theorem, $z_i \in (x_i, x_{i+1})$. Finally,

$$|g'(x)|=\left|\frac{(-1)^{i+1}\epsilon}{4(x_{i+1}-x_i)}+p'(z_i)\right| \ge \frac{\epsilon}{4(x_{i+1}-x_i)}-|p'(z_i)|\ge \frac{\epsilon}{4h}-M>n.$$

The inequality also holds for $x=a$ and $x=b$.

Baire's theorem implies that

$$\bigcup_{n\in\mathbb{N}} F_n \subsetneq \mathcal{C}([a,b]).$$

Hence, there exists a continuous function which is nowhere differentiable. □

Corollary 3.1.2. *The nowhere differentiable continuous functions on $[a,b]$ are dense in $\mathcal{C}([a,b])$.*

Proof. For all $n \in \mathbb{N}$, F_n^c is open and $X \setminus \bar{F}_n^c = \text{int}(F_n) = \emptyset$, namely, $\bar{F}_n^c = X$. By Baire's theorem, $\cap_{n\in N} F_n^c$ is dense in $\mathcal{C}([a,b]$. □

3.1.2 The Riesz representation theorem

In this subsection, the dual space of $\mathcal{C}([a,b])$ is characterized by the Riesz representation theorem. This theorem will be crucial in Chapter 4.

Theorem 3.1.3 (Riesz representation theorem)**.** *If $u^* \in \mathcal{C}([a,b])^*$, then there exists a function of bounded variation $g:[a,b]\to\mathbb{R}$ such that*

$$u^*(f)=\int_a^b f dg$$

for all $f \in \mathcal{C}([a,b])$ and $||u^*||=V(g)$.

Proof. First of all, notice that $(\mathcal{C}([a,b]),||\cdot||_\infty)$ is a normed subspace of $(L^\infty([a,b]),||\cdot||_\infty)$. By Theorem 2.3.6, there exists a continuous linear form $v^*: L^\infty([a,b]) \to \mathbb{R}$ that extends u^* and $||v^*||=||u^*||$.

For all $s \in (a,b]$, we define $\chi_s := \chi_{[a,s]} \in L^\infty([a,b])$ and, if $s = a$, $\chi_a = 0$. Besides, consider

$$g : [a,b] \to \mathbb{R}$$
$$s \mapsto v^*(\chi_s)$$

a function of bounded variation in $[a,b]$. Indeed, let $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ be a partition of $[a,b]$ and set

$$s_i = \frac{|g(x_i) - g(x_{i-1})|}{g(x_i) - g(x_{i-1})} \in \{-1,1\} \text{ and } s_i = 0 \text{ if } g(x_i) - g(x_{i-1}) = 0.$$

Then,

$$\begin{aligned}\sum_{i=1}^n |g(x_i) - g(x_{i-1})| &= \sum_{i=1}^n s_i\big(g(x_i) - g(x_{i-1})\big) \\ &= v^*\bigg(\sum_i s_i(\chi_{x_i} - \chi_{x_{i-1}})\bigg) \leq ||v^*||.\end{aligned}$$

Hence, $g \in BV([a,b])$ and $V(g) \leq ||u^*||$. Our next step is to prove that u^* is a Riemann-Stieltjes integral. Let $f \in \mathcal{C}([a,b])$, note that f is uniformly continuous, then for a given $\epsilon > 0$ there is $\delta > 0$ so that

$$|f(x) - f(y)| \leq \epsilon \quad \text{whenever } |x - y| \leq \delta.$$

Consider partitions $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ such that $||P|| \leq \delta$ for any $n \geq 1$, and functions

$$f_P := f(x_1)\chi_{[x_0,x_1]} + f(x_2)\chi_{(x_1,x_2]} + \ldots + f(x_n)\chi_{(x_{n-1},x_n]}.$$

Then, $||f - f_P||_\infty \leq \epsilon$. That is, $f = \lim_{||P|| \to 0} f_P$ in $L^\infty([a,b])$. Consequently,

$$u^*(f) = v^*(f) = \lim_{||P|| \to 0} v^*(f_P) = \lim_{||P|| \to 0} \sum_i f(x_i)[g(x_i) - g(x_{i-1})] = \lim_{||P|| \to 0} P(f,g,P).$$

Hence,

$$u^*(f) = \int_a^b f(t)dg(t) \quad \text{for all } f \in \mathcal{C}([a,b]).$$

Finally, by Proposition 1.3.18, $|u^*(f)| \leq V(g)||f||_\infty$ and, consequently, $||u^*|| \leq V(g)$. Hence, $||u^*|| = V(g)$. □

The Riesz representation theorem assures existence, but not uniqueness. For this purpose, the concept of normalised functions of bounded variation is introduced.

Definition 3.1.4. A function $g \in BV([a,b])$ is said to be *normalised* if $g(a) = 0$ and it is right continuous on (a,b). We denote $g \in NBV([a,b])$.

Lemma 3.1.5. *If $g \in BV([a,b])$, then there exists a unique function $h \in NBV([a,b])$ such that*

$$\int_a^b f \; dg = \int_a^b f dh$$

for all $f \in \mathcal{C}([a,b])$, and $V(h) \leq V(g)$.

Proof. We start with the existence of such a function by defining

$$h(x) = \begin{cases} 0, & \text{if } x = a, \\ g(x^+) - g(a), & \text{if } a < x < b, \\ g(b) - g(a), & \text{if } x = b. \end{cases}$$

Since g is a function of bounded variation, by Proposition 1.3.5, g is the difference of two increasing functions on $[a,b]$. Therefore, the right limit $g(x^+)$ exists for all $x \in (a,b)$ and g is continuous except at at most a countable number of points in $[a,b]$. To sum up, h is well-defined, right continuous and has at most a countable number of discontinuities.
Given $\epsilon > 0$, we consider a partition $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ and a set of points $\{y_1, \ldots, y_n\} \subset (a,b)$ at which g is continuous. Thus, g satisfies

$$x_j < y_j \text{ and } |g(x_j^+) - g(y_j)| < \frac{\epsilon}{2n} \quad \text{for all } 1 \le j \le n-1.$$

We take $y_0 = a$ and $y_n = b$. Then, for all $2 \le j \le n-1$,

$$\begin{aligned} |h(x_1) - h(x_0)| &= |g(x_1^+) - g(a)| \le |g(x_1^+) - g(y_1)| + |g(y_1) - g(y_0)| < |g(y_1) - g(y_0)| + \frac{\epsilon}{2n}, \\ |h(x_j) - h(x_{j-1})| &= |g(x_j^+) - g(x_{j-1}^+)| \le |g(x_j^+) - g(y_j)| + |g(y_j) - g(y_{j-1})| \\ &\quad + |g(y_{j-1}) - g(x_{j-1}^+)| < |g(y_j) - g(y_{j-1})| + \frac{\epsilon}{n}, \\ |h(x_n) - h(x_{n-1})| &= |g(x_n) - g(x_{n-1}^+)| \le |g(y_n) - g(y_{n-1})| + |g(y_{n-1}) - g(x_{n-1}^+)| \\ &\quad < |g(y_n) - g(y_{n-1})| + \frac{\epsilon}{2n}. \end{aligned}$$

It follows that

$$\begin{aligned} \sum_{j=1}^{n} \left| h(x_j) - h(x_{j-1}) \right| &\le \sum_{j=1}^{n} \left| g(y_j) - g(y_{j-1}) \right| + \frac{\epsilon + 2(n-2)\epsilon + \epsilon}{2n} \\ &= \sum_{j=1}^{n} |g(y_j) - g(y_{j-1})| + \frac{(n-1)\epsilon}{n} < V(g) + \epsilon. \end{aligned}$$

Since $V(h) < V(g) + \epsilon$ for all $\epsilon > 0$, $V(h) \le V(g)$. Thus, $h \in BV([a,b])$, that is, $h \in NBV([a,b])$.
Note that $h(x) = g(x) - g(a)$ for all $x \in [a,b]$, except at the points of discontinuities, which are at most countable as we have discussed previously. Let $f \in \mathcal{C}([a,b])$. Given $\epsilon > 0$, there exists $\delta > 0$ and a partition P of $[a,b]$ with $||P|| < \delta$, which does not contain any point of discontinuity of h. Then, $P(f, g, \mu) = P(f, h, \mu)$ and

$$\left| \int_a^b f \, dg - \int_a^b f \, dh \right| \le \left| \int_a^b f \, dg - P(f, g, \mu) \right| + \left| P(f, h, \mu) - \int_a^b f \, dh \right| < \epsilon.$$

Next, we prove the uniqueness of h. Suppose that there exists $h_0 \in NBV([a,b])$ such that

$$\int_a^b f \, dg = \int_a^b f \, dh_0 \quad \text{for all } f \in \mathcal{C}([a,b]).$$

Let $H(x) = h(x) - h_0(x)$. Note that $H(a) = 0 - 0 = 0$ and also

$$H(b) = H(b) - H(a) = \int_a^b dH = \int_a^b dh - \int_a^b dh_0 = 0.$$

Let $c \in (a,b)$ and $\gamma > 0$ be small enough. We define the following continuous function,

$$f(x) = \begin{cases} 1, & \text{if } a \leq x \leq c, \\ 1 - \frac{x-c}{\gamma}, & \text{if } c < x \leq c+\gamma, \\ 0, & \text{if } c+\gamma < x \leq b. \end{cases}$$

Then, from

$$0 = \int_a^b f \ dh - \int_a^b f \ dh_0 = \int_a^b f \ dH = \int_a^c \ dH + \int_c^{c+\gamma} \left(1 - \frac{x-c}{\gamma}\right) dH(x),$$

it follows that

$$H(c) = \int_a^c dH = -\int_c^{c+\gamma} \left(1 - \frac{x-c}{\gamma}\right) dH(x).$$

By Proposition 1.3.18, $|H(c)| \leq V_{[c,c+\gamma]}(H)$. Since H is right continuous on $[a,b]$, so is $v(x) = V_{[a,x]}(H)$, $a \leq x \leq b$. Given $\epsilon > 0$, there exists $\delta > 0$ such that for all $0 < \gamma < \delta$,

$$|H(c)| \leq v(c+\gamma) - v(c) < \epsilon.$$

Hence, $H(c) = 0$ for all $c \in (a,b)$, i.e., $H = 0$. □

Theorem 3.1.6. *There exists a bijection between $NBV([a,b])$ and $\mathcal{C}([a,b])^*$.*

Proof. Consider the following mapping,

$$\begin{aligned} \varphi : NBV([a,b]) &\to \mathcal{C}([a,b])^* \\ g &\mapsto T_g(f) := \int_a^b f \ dg. \end{aligned}$$

By Corollary 1.3.16, φ is well-defined and, by Theorem 3.1.3 and Lemma 3.1.5, the bijection is clear. □

Corollary 3.1.7. *For all $u^* \in \mathcal{C}([a,b])$ there exists a unique $g \in NBV([a,b])$ so that*

$$u^*(f) = \int_a^b f \ dg$$

for all $f \in \mathcal{C}([a,b])$, and $||u^*|| = V(g)$.

3.2 Application to Functional Analysis: separable Banach spaces

In this section, we will prove that every separable Banach space is isomorphic to a quotient of the space l^1.

Definition 3.2.1. A Banach space is said *separable* if it contains a countable dense subset.

Theorem 3.2.2. *Every separable Banach space is isomorphic to a quotient space of l^1.*

Proof. Let $(E, ||\cdot||)$ be a separable Banach space and $A = \{x_n : n \in \mathbb{N}\}$ a countable dense subset of $\overline{B_E(0,1)}$. Given $y \in l^1$, the series $\sum_n y_n x_n$ is absolutely convergent, since

$$\sum_n ||y_n x_n||_E \leq \sum_n |y_n| < \infty.$$

Therefore, the series is convergent and we can define

$$\begin{aligned} T : l^1 &\to E \\ y &\mapsto \sum_n y_n x_n \end{aligned}$$

a well-defined continuous linear operator. We next show that T is surjective. Let $x \in \overline{B_E(0,1)}$, since A is dense in $\overline{B_E(0,1)}$, there exists $x_{n_1} \in A$ so that $||2(x - x_{n_1})||_E < 1$. Then, there exists $x_{n_2} \in A$ so that

$$||2(x - x_{n_1}) - x_{n_2}||_E < \tfrac{1}{2}, \text{ that is, } ||x - x_{n_1} - \tfrac{1}{2}x_{n_2}||_E < \tfrac{1}{2^2}.$$

By induction, we build a subsequence $\{x_{n_k}\}_k$ such that

$$\left|\left| x - \sum_{k=1}^{m} \frac{1}{2^{k-1}} x_{n_k} \right|\right|_E < \frac{1}{2^m} \quad \text{for all integers } m \geq 1.$$

Consequently, the partial sums sequence converges to x in E. Thus, $x = T(y)$ with $y_i = \frac{1}{2^{k-1}}$ if $i = n_k$ and $y_i = 0$ otherwise. In case $||x|| > 1$, consider $x = ||x|| \frac{x}{||x||}$.
By the isomorphism theorem, we can define

$$\begin{aligned} \bar{T} : l^1/ker(T) &\to E \\ y + ker(T) &\mapsto T(y) \end{aligned}$$

a bijective continuous linear operator. Note that $\ker(T) = T^{-1}\{0\}$ is closed and, by Proposition 1.1.14, $l^1/\ker(T)$ is a Banach space. Then, by Corollary 2.1.5, $E \cong l^1/ker(T)$.

□

3.3 Application to Harmonic Analysis: divergence of Fourier series

Consider the Hilbert space $L^2([-\pi, \pi])$ and the orthonormal basis $\{e^{int}, n \in \mathbb{Z}\}$. For all $g \in L^2([-\pi, \pi])$, its Fourier series is defined by

$$g(t) = \sum_{k=-\infty}^{\infty} \hat{g}(k) e^{ikt} \text{ with } \hat{g}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) e^{-ikt} dt.$$

We define the *nth Fourier partial sums of* g by

$$S_n g(t) := \sum_{k=-n}^{n} \hat{g}(k) e^{ikt}.$$

The Fourier partial sums sequence converges to g in the sense $||S_ng - g||_2 \to 0$ as $n \to \infty$ (See [2], pp. 78-84). The question whether the convergence is also pointwise or not immediately arises. Furthermore, given a function $f \in L^p([-\pi, \pi])$, does its Fourier series converge to f in the L^p-norm? And does it pointwise? In this section we will show that the uniform boundedness principle plays an important role in answering some of these questions.

We will give a brief historical perspective of the convergence of the Fourier series. In 1873, Paul du Bois-Reymond gave an example of a continuous function whose Fourier series diverged at a point. Later, in 1921, Andrey Kolmogorov gave an example of a function in L^1 whose Fourier series diverged almost everywhere. It wasn't until 1966 that Lennart Carleson proved that in L^2 the Fourier series converges almost everywhere. A year later, Richard Hunt generalized Carleson's result: he proved that the Fourier series of every function in L^p, $1 < p < \infty$, converges almost everywhere. Finally, the convergence in the L^p-norm, $1 < p < \infty$, is also remarkable.

3.3.1 In $(\mathcal{C}([a, b]), ||\cdot||_\infty)$

Let us start with the case of continuous functions, given $f \in \mathcal{C}([-\pi, \pi])$, we will show that the convergence at a point fails and so the uniform one.

Definition 3.3.1. The *Dirichlet kernels* $D_n : [-\pi, \pi] \to \mathbb{R}$, $n \in \mathbb{N}$, are defined by

$$D_n(t) = \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikt}.$$

Proposition 3.3.2. *The Dirichlet kernels satisfy the following properties:*

(i) $D_n(0) = \frac{2n+1}{2\pi}$, *and*

(ii) $D_n(t) = \frac{\sin(n+\frac{1}{2})t}{2\pi \sin\frac{t}{2}}$ *for all nonzero* $t \in [-\pi, \pi]$.

Proof. (i) is obvious. Consider $t \in [-\pi, \pi]$, $t \neq 0$, and $n \in \mathbb{N}$,

$$\begin{aligned} 2\pi D_n(t) &= \sum_{k=-n}^{n} e^{ikt} = e^{-int} \sum_{k=0}^{2n} e^{ikt} = e^{-int}\, \frac{1 - e^{it(2n+1)}}{1 - e^{it}} = \frac{e^{-int} - e^{it(n+1)}}{1 - e^{it}} \\ &= e^{it/2}\, \frac{e^{-it(n+\frac{1}{2})} - e^{it(n+\frac{1}{2})}}{1 - e^{it}} = \frac{e^{it/2}}{e^{it/2}} \frac{e^{-it(n+\frac{1}{2})} - e^{it(n+\frac{1}{2})}}{e^{-it/2} - e^{it/2}} = \frac{\sin(n+\frac{1}{2})t}{\sin\frac{t}{2}}. \end{aligned}$$

□

Lemma 3.3.3. *Let* $g \in L^1([-\pi, \pi])$, *its Fourier partial sums can be expressed as*

$$S_ng(t) = (g * D_n)(t) := \textstyle\int_{-\pi}^{\pi} g(s) D_n(t-s) ds$$

for all $t \in [-\pi, \pi]$ *and all* $n \in \mathbb{N}$.

Proof. The proof is immediate,

$$\begin{aligned} S_ng(t) &= \sum_{k=-n}^{n} \hat{g}(k) e^{ikt} = \frac{1}{2\pi} \sum_{k=-n}^{n} \int_{-\pi}^{\pi} g(s) e^{-iks} ds\; e^{ikt} = \frac{1}{2\pi} \sum_{k=-n}^{n} \int_{-\pi}^{\pi} g(s) e^{ik(t-s)} ds \\ &= \int_{-\pi}^{\pi} g(s) \left(\frac{1}{2\pi} \sum_{k=-n}^{n} e^{ik(t-s)} \right) ds = \int_{-\pi}^{\pi} g(s) D_n(t-s) ds. \end{aligned}$$

□

Remark 3.3.4. In particular, note that $S_n g(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s) D_n(s) ds$.

Lemma 3.3.5. *The sequence $\{||D_n||_1\}_{n\geq 0}$ is unbounded.*

Proof.

$$2\pi||D_n||_1 = \int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}\right|dx \geq 4\int_0^{\frac{\pi}{2}}\left|\frac{\sin(2n+1)y}{y}\right|dy = 4\int_0^{\frac{(2n+1)\pi}{2}}\frac{|\sin z|}{z}dz$$
$$= 4\sum_{k=0}^{2n}\int_{k\frac{\pi}{2}}^{(k+1)\frac{\pi}{2}}\frac{|\sin z|}{z}dz \geq 4\sum_{k=0}^{2n}\frac{2}{(k+1)\pi}\int_{k\frac{\pi}{2}}^{(k+1)\frac{\pi}{2}}|\sin z|dz = \frac{8}{\pi}\sum_{k=0}^{2n}\frac{1}{k+1}.$$

Hence, $\sup_n ||D_n||_1 = \infty$. □

Theorem 3.3.6. *There exists $g \in \mathcal{C}([-\pi,\pi])$ whose Fourier series diverges at the origin.*

Proof. For all $n \in \mathbb{N}$, consider $u_n : \mathcal{C}([-\pi,\pi]) \to \mathbb{R}$ a linear form defined by

$$u_n(g) = S_n g(0) = \int_{-\pi}^{\pi} g(x) D_n(x) dx.$$

Besides,

$$|u_n(g)| \leq \int_{-\pi}^{\pi}|g(x)|\ |D_n(x)|dx \leq ||g||_\infty \int_{-\pi}^{\pi}|D_n(x)|dx$$

implies that u_n is continuous with

$$||u_n|| \leq \int_{-\pi}^{\pi}|D_n(x)|dx = ||D_n||_1.$$

For all $n \in \mathbb{N}$, we define $g_n(x) = \text{sign}(D_n(x))$ a discontinuous function at the zeros of $D_n(x)$. Note that, $D_n(x)$ has a finite number of zeros. Indeed, $D_n(x) = 0$ if and only if $\sin(n+\frac{1}{2})x = 0$, $x \neq 0$, if and only if $(n+\frac{1}{2})x = \pm k\pi$ with $1 \leq k \leq n$ if and only if $x = \pm k\pi/(n+\frac{1}{2})$ with $1 \leq k \leq n$. Hence, $D_n(x)$ has $2n$ zeros.

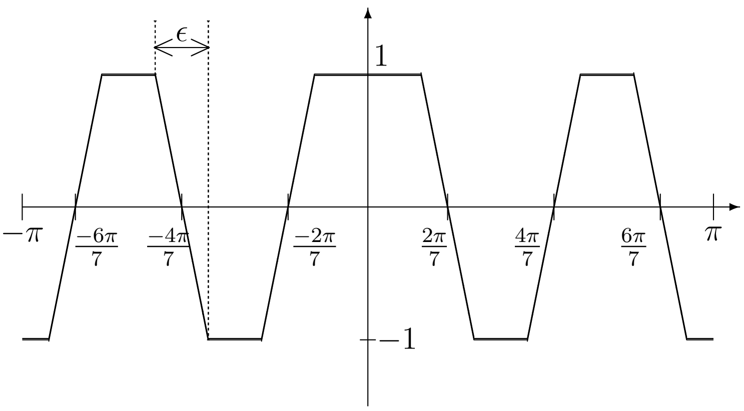

Figure 3.3.1: Representation of the continuous function g_n^ϵ for $n = 3$.

Given $\epsilon > 0$ small enough, let $g_n^\epsilon : [-\pi,\pi] \to \mathbb{R}$ denote the continuous piecewise affine function that is equal to g_n on $[-\pi,\pi] \setminus I_n^\epsilon$, where I_n^ϵ denotes the intersection of $[-\pi,\pi]$

with the union of the open intervals of length ϵ centered at the $2n$ zeros of the Dirichlet kernel D_n that belong to the interval $[-\pi, \pi]$. Note that $||g_n^\epsilon||= 1$.

Then, by the dominated convergence theorem ($|(g_n^\epsilon(x) - g_n(x))D_n(x)|\leq |D_n(x)|$),

$$\left|u_n(g_n^\epsilon) - \int_{-\pi}^{\pi} |D_n(x)|dx\right| = \left|\int_{-\pi}^{\pi} (g_n^\epsilon(x) - g_n(x))D_n(x)dx\right| \to 0 \text{ as } \epsilon \to 0.$$

Since $||g_n^\epsilon||= 1$,

$$||u_n||\geq |u_n(g_n^\epsilon)|\to ||D_n||_1 \quad \text{as } \epsilon \to 0.$$

By Lemma 3.3.5, $\sup_n||u_n||= \sup||D_n||_1= \infty$. Finally, by the uniform boundedness principle, there exists $g \in \mathcal{C}([-\pi, \pi])$ such that $\sup_n|u_n(g)|= \infty$. □

Theorem 3.3.6 shows that the Fourier series does not converges pointwise in $\mathcal{C}([a, b]$.

3.3.2 In $(L^1([a, b]), ||\cdot||_1)$

Next, we similarly study the convergence in $L^1([-\pi, \pi])$ with the L^1-norm.

Definition 3.3.7. For all $N \in \mathbb{N}$, the Fejer's kernel $F_N : [-\pi, \pi] \to \mathbb{R}$ is defined by

$$F_N(x) = \frac{1}{N+1}\sum_{k=0}^{N} D_k(x).$$

Lemma 3.3.8. *For all* $N \in \mathbb{N}$*,*

(i) $F_N(0) = \frac{N+1}{2\pi}$*, and*

(ii) $F_N(x) = \frac{1}{2\pi(N+1)}\frac{\sin^2[(N+1)\frac{x}{2}]}{sin^2\frac{x}{2}}$ *for all nonzero* $x \in [-\pi, \pi]$*.*

Proof. (i) For all $N \in \mathbb{N}$,

$$\begin{aligned}F_N(0) &= \frac{1}{N+1}\sum_{k=0}^{N} D_k(0) = \frac{1}{N+1}\sum_{k=0}^{N}\frac{2k+1}{2\pi}\\ &= \frac{1}{2\pi(N+1)}[N(N+1) + N + 1] = \frac{N+1}{2\pi}.\end{aligned}$$

(ii) For all $N \in \mathbb{N}$ and all $x \in [-\pi, \pi]$, $x \neq 0$,

$$\begin{aligned}2\pi(N+1)F_N(x) &= 2\pi\sum_{k=0}^{N} D_k(x) = \sum_{k=0}^{N}\frac{\sin(k+\frac{1}{2})x}{\sin\frac{x}{2}} = \frac{1}{\sin\frac{x}{2}}\mathrm{Im}\left(\sum_{k=0}^{N} e^{i(k+\frac{1}{2})x}\right)\\ &= \frac{1}{\sin\frac{x}{2}}\mathrm{Im}\left(\frac{e^{i(N+\frac{3}{2})x} - e^{i\frac{1}{2}x}}{e^{ix}-1}\right)\\ &= \frac{1}{\sin\frac{x}{2}}\mathrm{Im}\left(e^{i\frac{x}{2}}\frac{e^{i(N+1)x}-1}{e^{ix}-1}\right) = \frac{1}{\sin\frac{x}{2}}\mathrm{Im}\left(\frac{e^{i(N+1)x}-1}{e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}}\right)\\ &= \frac{1}{2\sin^2\frac{x}{2}}\mathrm{Im}\left(\frac{\cos[(N+1)x] + i\sin[(N+1)x] - 1}{i}\right)\\ &= \frac{1}{2\sin^2\frac{x}{2}}\big(1-\cos[(N+1)x]\big) = \frac{\sin^2[(N+1)\frac{x}{2}]}{\sin^2\frac{x}{2}}.\end{aligned}$$

□

Lemma 3.3.9. *For all $N \in \mathbb{N}$, the Fejer's kernels satisfy the following properties:*

(i) $F_N(x) \geq 0$ *for all* $x \in [-\pi, \pi]$.

(ii) $\int_{-\pi}^{\pi} F_N(x)dx = 1$.

(iii) Given $0 < \delta < \pi$, $\lim\limits_{N \to \infty} \int_{\delta}^{\pi} F_N(x)dx = 0$.

Proof. (i) By Lemma 3.3.8, it is clear that $F_N(x) \geq 0$ for all $x \in [-\pi, \pi]$.

(ii)

$$\int_{-\pi}^{\pi} F_N(x)dx = \int_{-\pi}^{\pi} \frac{1}{2\pi(N+1)} \sum_{k=0}^{N} \sum_{m=-k}^{k} e^{imx}dx = \frac{1}{2\pi(N+1)} \sum_{k=0}^{N} \sum_{m=-k}^{k} \int_{-\pi}^{\pi} e^{imx}dx.$$

If $m = 0$, then

$$\frac{1}{2\pi(N+1)} \sum_{k=0}^{N} \int_{-\pi}^{\pi} 1 \ dx = 1.$$

If $m \neq 0$, then

$$\int_{-\pi}^{\pi} e^{imx}dx = \frac{1}{im}[e^{im\pi} - e^{-im\pi}] = 0.$$

Hence, $\int_{-\pi}^{\pi} F_N(x)dx = 1$.

(iii) Let $0 < \delta < \pi$, if $\delta \leq x \leq \pi$, then $1 \leq \frac{1}{\sin^2 \frac{x}{2}} \leq \frac{1}{\sin^2 \frac{\delta}{2}}$. For all $\epsilon > 0$ and all $\delta \leq x \leq \pi$ there exists $N_0 \in \mathbb{N}$ so that

$$0 \leq F_N(x) = \frac{1}{2\pi(N+1)} \frac{\sin^2[(N+1)\frac{x}{2}]}{\sin^2 \frac{x}{2}} \leq \frac{1}{2\pi(N+1)} \frac{1}{\sin^2 \frac{\delta}{2}} < \frac{\epsilon}{\pi - \delta}$$

whenever $N \geq N_0$. Finally,

$$\int_{\delta}^{\pi} F_N(x)dx < \int_{\delta}^{\pi} \frac{\epsilon}{\pi - \delta} \ dx = \epsilon.$$

□

Lemma 3.3.10. *If* $f \in L^1([-\pi, \pi])$, *then*

$$\lim_{y \to 0} ||f(\cdot - y) - f||_1 = 0.$$

Proof. Given $f \in L^1([-\pi, \pi])$ and $\epsilon > 0$, by Proposition 1.2.1, there exists $g \in \mathcal{C}([-\pi, \pi])$ so that $||f - g||_1 < \frac{\epsilon}{3}$. Since g is uniformly continuous on $[-\pi, \pi]$, for some $\delta > 0$ we have that

$$||g(\cdot - y) - g||_1 = \int_{-\pi}^{\pi} |g(x-y) - g(x)|dx < \frac{\epsilon}{3} \quad \text{whenver } |y| < \delta.$$

Then, for all $|y| < \delta$,

$$||f(\cdot - y) - f||_1 \leq ||f(\cdot - y) - g(\cdot - y)||_1 + ||g(\cdot - y) - g||_1 + ||g - f||_1 < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

□

Proposition 3.3.11. *Let $f \in L^1([-\pi,\pi])$, $f * F_N \to f$ in $L^1([-\pi,\pi])$ as $N \to \infty$.*

Proof. Let $\epsilon > 0$.

$$\begin{aligned}
||f * F_N - f||_1 &= \int_{-\pi}^{\pi} \big|(f * F_N)(x) - f(x)\big| dx = \int_{-\pi}^{\pi} \left| \int_{-\pi}^{\pi} f(x-y)F_N(y)dy - f(x) \right| dx \\
&= \int_{-\pi}^{\pi} \left| \int_{-\pi}^{\pi} \big(f(x-y) - f(x)\big) F_N(y)dy \right| dx \\
&\leq \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |f(x-y) - f(x)| F_N(y)dydx \\
&= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |f(x-y) - f(x)|dx\ F_N(y)dy \\
&= \int_{-\pi}^{-\delta} \int_{-\pi}^{\pi} |f(x-y) - f(x)|dx\ F_N(y)dy \\
&+ \int_{-\delta}^{\delta} \int_{-\pi}^{\pi} |f(x-y) - f(x)|dx\ F_N(y)dy \\
&+ \int_{\delta}^{\pi} \int_{-\pi}^{\pi} |f(x-y) - f(x)|dx\ F_N(y)dy.
\end{aligned}$$

Let us analyze these three integrals separately. First, by Lemma 3.3.9, there exists $N_0 \in \mathbb{N}$ so that

$$\begin{aligned}
\int_{-\pi}^{-\delta} \int_{-\pi}^{\pi} |f(x-y) - f(x)|dx\ F_N(y)dy &\leq 2||f||_1 \int_{-\pi}^{-\delta} F_N(y)dy < \frac{\epsilon}{3}, \text{ and} \\
\int_{\delta}^{\pi} \int_{-\pi}^{\pi} |f(x-y) - f(x)|dx\ F_N(y)dy &\leq 2||f||_1 \int_{\delta}^{\pi} F_N(y)dy < \frac{\epsilon}{3}
\end{aligned}$$

whenever $N \geq N_0$. The other integral requires Lemma 3.3.10,

$$\begin{aligned}
\int_{-\delta}^{\delta} \int_{-\pi}^{\pi} |f(x-y) - f(x)|dx\ F_N(y)dy &= \int_{-\delta}^{\delta} ||f(\cdot - y) - f||_1 F_N(y)dy \\
&\leq \int_{-\delta}^{\delta} \frac{\epsilon}{3} F_N(y)dy \leq \frac{\epsilon}{3}
\end{aligned}$$

for some $\delta > 0$ small enough. Hence,

$$||f * F_N - f||_1 < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

□

Theorem 3.3.12. *There exists $f \in L^1([-\pi,\pi])$ whose Fourier series does not converge to f in the L^1-norm.*

Proof. According to Young's inequality, $||S_n g||_1 = ||D_n * g||_1 \leq ||D_n||_1\ ||g||_1$. Hence, the linear operator $S_n : L^1([-\pi,\pi]) \to L^1([-\pi,\pi])$ is continuous. By Proposition 3.3.11, $||S_n F_N||_1 = ||D_n * F_N||_1 \to ||D_n||_1$ as $N \to \infty$. Since $||F_N||_1 = 1$, it follows $||S_n|| \geq ||D_n||_1$. By Lemma 3.3.5, $\sup_n ||S_n|| = \infty$ and, by the uniform boundedness principle, there exists $f \in L^1([-\pi,\pi])$ so that $\sup_n ||S_n f||_1 = \infty$. Hence, $S_n f$ does not converge to f with the L^1-norm. □

3.4 Application to Numerical Analysis: divergence of Lagrange interpolation

In this section, we will show that the Lagrange interpolating polynomial does not converge uniformly to the function. We will consider an interval $[a, b] \subset \mathbb{R}$, $a < b$, and $\mathcal{P}_n([a, b])$ will denote the vector space of polynomials of degree $\leq n$ on $[a, b]$.

Definition 3.4.1. For all $n \in \mathbb{N}$, let $a \leq x_0 < x_1 < \ldots < x_n \leq b$ be $(n+1)$ different points. Given a continuous function $f : [a, b] \to \mathbb{R}$, its *Lagrange interpolating polynomial of degree* $\leq n$ *associated with the* $(n+1)$ *nodes* x_i, $0 \leq i \leq n$, is given by

$$L_n f(x) = \sum_{j=0}^{n} f(x_j) p_j(x), \ \ a \leq x \leq b,$$

where the $(n+1)$ polynomials $p_j \in \mathcal{P}_n([a, b])$, $0 \leq j \leq n$, are defined by

$$p_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{x - x_i}{x_j - x_i}.$$

Lemma 3.4.2. *For all $n \in \mathbb{N}$, let $a \leq x_0 < x_1 < \ldots < x_n \leq b$ be $(n+1)$ different points. Given any function $f \in \mathcal{C}([a, b])$, its Lagrange interpolating polynomial is the only polynomial in $\mathcal{P}_n([a, b])$ that satisfies that*

$$L_n f(x_i) = f(x_i) \quad \text{for all } 0 \leq i \leq n.$$

Proof. We first notice that $p_j(x_i) = \delta_{ij}$ for all $0 \leq i, j \leq n$. Indeed,

$$p_j(x_i) = \prod_{\substack{k=0 \\ k \neq j}}^{n} \frac{x_i - x_k}{x_j - x_k} = \delta_{ij}.$$

Then, given a function $f \in \mathcal{C}([a, b])$, we have that

$$L_n f(x_i) = \sum_{j=0}^{n} f(x_j) \delta_{ij} = f(x_i) \quad \text{for all } 0 \leq i \leq n.$$

Let us consider the canonical basis $\{1, x, \ldots, x^n\}$, finding a polynomial

$$q(x) = \sum_{i=0}^{n} a_i x^i \in \mathcal{P}_n([a, b]),$$

such that $q(x_i) = f(x_i)$ for all $0 \leq i \leq n$, implies solving the linear system

$$\begin{pmatrix} f(x_0) \\ \vdots \\ f(x_n) \end{pmatrix} = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}$$

Since all the nodes are different, all columns are linearly independent, which means that the $(n+1) \times (n+1)$ matrix has a nonzero determinant. Besides, we already know that the Lagrange interpolating polynomial is a solution. Hence, the solution is unique and, obviously, it is the Lagrange interpolating polynomial. $\square$

Remark 3.4.3. Notice that for all $p \in \mathcal{P}_n([a,b])$, $L_n p = p$.

Remark 3.4.4. For all $q \in \mathcal{P}_n([a,b])$ and all $x \in [a,b]$,

$$q(x) = \sum_{j=0}^{n} q(x_j) p_j(x).$$

Then, the polynomials $p_j(x)$, $0 \leq j \leq n$ are $(n+1)$ generators of $\mathcal{P}_n([a,b])$ and, hence, they are a basis.

Proposition 3.4.5. *For all $n \in \mathbb{N}$, let $a \leq x_0 < x_1 < \ldots < x_n \leq b$ be $(n+1)$ different points. The operator*

$$\begin{aligned} L_n : \mathcal{C}([a,b]) &\to \mathcal{C}([a,b]) \\ f &\mapsto L_n f \end{aligned}$$

is linear and continuous with $||L_n|| = \sup_{a \leq x \leq b} \left(\sum_{j=0}^{n} |p_j(x)| \right)$.

Proof. The linearity is clear. We note that L_n is continuous,

$$\begin{aligned} ||L_n|| &= \sup_{||f||_\infty = 1} ||L_n f||_\infty = \sup_{||f||_\infty = 1} \left|\left| \sum_{j=0}^{n} f(x_j) p_j(x) \right|\right|_\infty \leq \left|\left| \sum_{j=0}^{n} p_j(x) \right|\right|_\infty = \sup_{a \leq x \leq b} \left| \sum_{j=0}^{n} p_j(x) \right| \\ &\leq \sup_{a \leq x \leq b} \sum_{j=0}^{n} |p_j(x)|. \end{aligned}$$

Next, notice that $\sum_{j=0}^{n} |p_j(x)|$ is continuous on $[a,b]$ and, since $[a,b]$ is compact, there exists $c \in [a,b]$ such that

$$\sup_{a \leq x \leq b} \sum_{j=0}^{n} |p_j(x)| = \sum_{j=0}^{n} |p_j(c)| < \infty.$$

Now, let $g : [a,b] \to \mathbb{R}$ be the continuous piecewise affine function defined by $g(a) = \text{sign}(p_0(c))$, $g(x_j) = \text{sign}(p_j(c))$ for all $0 \leq j \leq n$ and $g(b) = \text{sign}(p_n(c))$. Note that there are at most $(n+1)$ changes of sign, so this continuous piecewise affine function g exists. Besides, g is nonzero. Indeed, if we consider the polynomial $q(x) = 1$, then we have that

$$L_n q(x) = \textstyle\sum_{j=0}^{n} p_j(x) = 1, \; a \leq x \leq b,$$

and, hence, $\sum_{j=0}^{n} |p_j(c)| \geq 1$. This means that there exists at least one $0 \leq j \leq n$ such that $p_j(c) \neq 0$. Therefore $||g||_\infty = 1$. Finally,

$$||L_n|| \geq ||L_n g||_\infty \geq \left| L_n g(c) \right| = \left| \sum_{j=0}^{n} g(x_j) p_j(c) \right| = \sum_{j=0}^{n} |p_j(c)|.$$

We conclude that $||L_n|| = \sup_{a \leq x \leq b} \left(\sum_{j=0}^{n} |p_j(x)| \right)$. $\square$

Theorem 3.4.6. *For all integers $n \geq 1$, given the $(n+1)$ nodes $x_i = a + \frac{(b-a)i}{n}$ with $0 \leq i \leq n$, there exists a function $f \in \mathcal{C}([a,b])$ whose Lagrange interpolating polynomial does not converge uniformly to f.*

Proof. By Theorem 3.4.5, $||L_n||= \sup_{a\leq x\leq b}\left(\sum_{j=0}^{n}|p_j(x)|\right)$. Then,

$$\begin{aligned}
||L_n|| &\geq \sum_{j=0}^{n}\left|p_j\left(a+\frac{(b-a)}{2n}\right)\right| = \sum_{j=0}^{n}\left|\prod_{\substack{i=0\\ i\neq j}}^{n}\frac{\frac{b-a}{2n}-\frac{(b-a)i}{n}}{\frac{(b-a)j}{n}-\frac{(b-a)i}{n}}\right| = \sum_{j=0}^{n}\left|\prod_{\substack{i=0\\ i\neq j}}^{n}\frac{\frac{1}{2}-i}{j-i}\right| \\
&= \sum_{j=0}^{n}\left|\prod_{\substack{i=0\\ i\neq j}}^{n}\frac{2i-1}{2(j-i)}\right| = \sum_{j=0}^{n}\frac{\left|\prod_{i=0}^{n}(2i-1)\right|}{2^n j!(n-j)!(2j-1)} = \sum_{j=0}^{n}\frac{1}{2^n j!(n-j)!(2j-1)}\frac{(2n)!}{n!2^n} \\
&\geq \sum_{j=0}^{n}\frac{1}{2^n j!(n-j)!2n}\frac{(2n)!}{n!2^n} = \sum_{j=0}^{n}\frac{n!}{j!(n-j)!}\frac{(2n)!}{n2^{2n+1}(n!)^2} = \frac{(2n)!}{n2^{2n+1}(n!)^2}\sum_{j=0}^{n}\binom{n}{j} \\
&= \frac{(2n)!}{n2^{2n+1}(n!)^2}2^n = \frac{(2n)!}{n2^{n+1}(n!)^2} = \frac{2^n n!\prod_{j=1}^{n}(2j-1)}{n2^{n+1}(n!)^2} = \frac{\prod_{j=1}^{n}(2j-1)}{2n\cdot n!} \\
&\geq \frac{\prod_{j=2}^{n}(2j-2)}{2n\cdot n!} = \frac{2^{n-1}(n-1)!}{2n\cdot n!} = \frac{2^{n-2}}{n^2}.
\end{aligned}$$

Hence, $\sup_{n\geq 1}||L_n||= \infty$. By the uniform boundedness principle, there exists $f \in \mathcal{C}([a,b])$ so that $\sup_{n\geq 1}||L_n f||_\infty = \infty$. $\square$

Remark 3.4.7. For any election of the nodes (See [5]), there exists a constant $c > 0$ so that

$$||L_n||= \sup_{a\leq x\leq b}\left(\sum_{j=0}^{n}|p_j(x)|\right) \geq \frac{2}{\pi}log(n) - c.$$

Then, $\sup_{n\geq 1}||L_n||= \infty$ and, by the uniform boundedness principle, there exists f in $\mathcal{C}([a,b])$ such that $\sup_{n\geq 1}||L_n f||_\infty = \infty$.

3.5 Application to Differential Equations

Our purpose is to prove that given a Cauchy problem, the solutions vary continuously with the datum and the boundary conditions. We will study the case of a second order ordinary differential equation because it is quite common, especially in many physical problems.

Theorem 3.5.1. *Let functions $a, b, c \in \mathcal{C}([0,1])$ be given such that the two-point boundary value problem*

$$a(x)u''(x) + b(x)u'(x) + c(x)u(x) = f(x),\ 0 \leq x \leq 1,\ u(0) = u_0 \text{ and } u'(0) = v_0$$

has one and only one solution $u \in \mathcal{C}^2([0,1])$ for any $f \in \mathcal{C}([0,1])$. Then, there exists a constant $C > 0$ such that

$$||u''||+||u'||+||u||\leq C(||f||+|u_0|+|v_0|)$$

where $||\cdot||$ *denotes the supremum norm of the space* $\mathcal{C}([0,1])$.

Proof. We consider the Banach space $\mathcal{C}^2([0,1])$ with the norm $||v||_*= ||v''||+||v'||+||v||$. Let us define

$$\begin{aligned} T:\ \mathcal{C}^2([0,1]) &\to \mathcal{C}([0,1]) \times \mathbb{R}^2 \\ u &\mapsto (a(x)u''(x) + b(x)u'(x) + c(x)u(x), u(0), u'(0)) \end{aligned}$$

a linear bijective operator. Besides, T is continuous,

$$\begin{aligned} ||Tu|| &\leq ||a(x)u''(x) + b(x)u'(x) + c(x)u(x)||+|u(0)|+|u'(0)| \\ &\leq \max\{||a(x)||, ||b(x)||, ||c(x)||\}||u||_*+||u||+||u'||+||u''|| \\ &= [\max\{||a(x)||, ||b(x)||, ||c(x)||\} + 1]\ ||u||_*. \end{aligned}$$

Then, T is a bijective continuous linear operator between two Banach spaces. By Corollary 2.1.5, there exists T^{-1} and it is continuous. Hence, there exists a constant $M > 0$ so that

$$||u||+||u'||+||u''||= ||T^{-1}(f, u_0, v_0)||\leq M(||f||+|u_0|+|v_0|).$$ □

Chapter 4

Applications to other areas

In this chapter, we aim to study three applications: the moment problem, the Chebyshev approximation and the optimal control of a rocket. These problems can be formulated in terms of linear forms, thus the Hahn-Banach theorem and the Riesz representation theorem will play an important role. Besides, even though some of these applications are related to Physics, they require developing an interesting mathematical background.

4.1 The moment problem

The term *moment problem* appeared for the first time in *Recherches sur les fractions continues*, a paper published by T. Stieltjes in 1894 (See[12]). At the end of Chapter 4, he proposed the following mechanical problem: "Find a positive mass distribution on $[0, \infty)$, given the moments of order k, $k \in \mathbb{N}$". In current notation, the moment of order k of a mass distribution on $[0, \infty)$ is

$$\mu_k = \int_0^\infty x^k dm(x).$$

We remark that μ_0 is the total mass, M, whereas μ_1/M is the position of the center of mass. The center of mass of an object is a fundamental concept in Physics because it is very useful for solving problems with multiple particles and non punctual objects. It is defined as an average of the masses factored by the distances from a reference point. Furthermore, μ_2 is the moment of inertia, which depends on the distribution of mass around the rotation axis and measures the tendency of an object to rotate.

Our next theorem proves that the finite moment problem has always a solution, i.e., given $\mu_0, \dots, \mu_N$, $(N+1)$ real numbers, there exits a function of bounded variation defined on an interval $[a, b]$ whose moments are those values.

Theorem 4.1.1. *The finite moment problem has always a solution.*

Proof. Let $\mu_0, \dots, \mu_N \in \mathbb{R}$ be the first $(N+1)$ moments, we will show that there exists a function $g \in BV([a, b])$, $-\infty < a < b < \infty$, so that

$$\int_a^b x^k dg(x) = \mu_k \quad \text{for all integers } 0 \leq k \leq N.$$

Let us consider $\mathcal{P}_N([a,b])$, the vector space of polynomials of order $\leq N$. Obviously, $\mathcal{P}_N([a,b])$ is a vector subspace of $\mathcal{C}([a,b])$ and its dimension is $N+1$, since $\{1, x, x^2, \dots, x^N\}$ is the canonical basis. We consider a linear form $u : \mathcal{P}_N[a,b] \to \mathbb{R}$ defined by

$$u\Big(\sum_{i=0}^{N} a_i x^i\Big) = \sum_{i=0}^{N} a_i \mu_i.$$

We want to show that u is continuous. Let $\{p_n(x)\}_n$ be a convergent sequence to $p(x)$ in $\mathcal{P}_N([a,b])$. Since we have the supremum norm, the convergence is uniform and also pointwise. By the Lagrangian interpolation formula, we have

$$p_n(x) = \sum_{j=0}^{N} p_n(x_j) q_j(x), \text{ for all } x \in [a,b] \text{ and all } n \in \mathbb{N},$$

where $a = x_0 < x_1 < ... < x_N = b$ is a fixed partition of the interval $[a,b]$ and $q_0, \dots, q_N$ are polynomials defined by

$$q_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{x - x_i}{x_j - x_i}, \; a \leq x \leq b \text{ and } 0 \leq j \leq N.$$

By Remark 3.4.4, $\{q_0, \dots, q_N\}$ is a basis of $\mathcal{P}_N([a,b])$. We want to prove that the coefficients of $p_n(x)$ in the canonical basis converge to the coefficients of $p(x)$ in the same basis. Indeed, let A be the matrix of change of basis and $(a_{0,n}, \dots, a_{N,n})$ the coordinates of $p_n(x)$ in the canonical basis. Then,

$$\begin{pmatrix} a_{0,n} \\ \vdots \\ a_{N,n} \end{pmatrix}_{\text{canonical}} = A \begin{pmatrix} p_n(x_0) \\ \vdots \\ p_n(x_N) \end{pmatrix}_{q_j} \xrightarrow[n\to\infty]{} A \begin{pmatrix} p(x_0) \\ \vdots \\ p(x_N) \end{pmatrix}_{q_j}$$

As $n \to \infty$, the coordinates of $p_n(x)$ in the canonical basis tend to the ones of $p(x)$. Hence, $u(p_n) \to u(p)$ as $n \to \infty$ and, thus, u is continuous.
By Theorem 2.3.6, there exists a continuous linear form $v : \mathcal{C}([a,b]) \to \mathbb{R}$ that extends u. Then, by the Riesz representation theorem, there exists $g \in BV([a,b])$ so that

$$v(f) = \textstyle\int_a^b f(x) dg(x) \quad \text{for all } f \in \mathcal{C}([a,b]).$$

Therefore, for all integers $0 \leq k \leq N$,

$$\mu_k = u(x^k) = v(x^k) = \int_a^b x^k dg(x).$$

□

Remark 4.1.2. In case of complex values, we can split the problem into real and imaginary parts and apply the previous theorem at each one.

So far, we have considered the case of a finite number of given moments in $\mathbb{K}$. Nevertheless, we could consider the case of a sequence $\{\mu_n\}_n$ in $\mathbb{K}$. Besides, the Riesz representation theorem provides a formulation of the original moment problem in terms of linear forms. For instance, consider a linear system in infinitely unknowns $(x_0, x_1, \ldots)$ and equations

$$\sum_{k=0}^{\infty} a_{j,k} x_k = b_j \quad \text{for } j \in \mathbb{N}.$$

If the sequences $a_j = (a_{j,0}, a_{j,1}, \ldots)$ belong to a normed vector space E, the solution of the linear system is a linear form $v = (x_0, x_1, \ldots)$ such that $v(a_j) = b_j$ for all $j \in \mathbb{N}$. What conditions must satisfy $a_j \in E$ and the constants b_j so as to guarantee the existence of such a linear form v?
From a more general approach, given a Banach space E, an arbitrary index set I, $\{x_i\}_{i\in I} \subset E$ and $\{c_i\}_{i\in I} \subset \mathbb{K}$, is there any linear form v on E so that $v(x_i) = c_i$ for all $i \in I$? Despite the fact that this problem not always has a solution, the Hahn Banach theorem allows us to give the following characterization.

Theorem 4.1.3. *Let $(E, ||\cdot||)$ be a real or complex normed space, $\{x_i\}_{i\in I}$ a sequence in E and $\{c_i\}_{i\in I}$ a sequence in $\mathbb{K}$. It is equivalent:*

(i) There exists $u \in E^*$ such that $u(x_i) = c_i$ for all $i \in I$.

(ii) There exists a constant $M > 0$ so that

$$\left|\sum_{k=1}^{n} \alpha_k c_{i_k}\right| \leq M \left|\left|\sum_{k=1}^{n} \alpha_k x_{i_k}\right|\right|$$

for any linear combination $\alpha_1 x_{i_1} + \ldots + \alpha_n x_{i_n}$ of elements of $\{x_i\}_{i\in I}$.

Proof. If there exists $u \in E^*$ such that $u(x_i) = c_i$ for all $i \in I$, then for any linear combination of elements of $\{x_i\}_{i\in I}$ we have

$$\left|\sum_{k=1}^{n} \alpha_k c_{i_k}\right| = \left|\sum_{k=1}^{n} \alpha_k u(x_{i_k})\right| = \left|u\left(\sum_{k=1}^{n} \alpha_k x_{i_k}\right)\right| \leq ||u|| \left|\left|\sum_{k=1}^{n} \alpha_k x_{i_k}\right|\right|.$$

Conversely, assume (ii) and consider

$$Y := \operatorname{span}\{x_i : i \in I\} = \left\{\sum_{k=1}^{n} \alpha_k x_{i_k} : \alpha_k \in \mathbb{K} \text{ and } n \geq 1\right\}.$$

For all $\sum_{k=1}^{n} \alpha_k x_{i_k} \in Y$ we define

$$u\left(\sum_{k=1}^{n} \alpha_k x_{i_k}\right) = \sum_{k=1}^{n} \alpha_k c_{i_k},$$

a linear form on E. First, we show that u is well-defined: if $\sum_{k=1}^{n} \alpha_k x_{i_k} = \sum_{l=0}^{m} \beta_l x_{i_l}$, then

$$u\left(\sum_{k=1}^{n} \alpha_k x_{i_k}\right) = u\left(\sum_{l=0}^{m} \beta_l x_{i_l}\right).$$

Indeed, it is enough to suppose $\sum_{k=1}^{n} \alpha_k x_{i_k} = 0$, then

$$\left|u\left(\sum_{k=1}^{n} \alpha_k x_{i_k}\right)\right| = \left|\sum_{k=1}^{n} \alpha_k c_{i_k}\right| \leq M \left|\left|\sum_{k=1}^{n} \alpha_k x_{i_k}\right|\right| = 0.$$

Therefore, u is well-defined on Y and, by (ii), continuous with $||u|| \leq M$. By Theorem 2.3.6, there exists $v \in E^*$ that extends u and such that $v(x_i) = c_i$ for all $i \in I$. ☐

4.2 Minimum norm problems

In this first part, we provide the necessary background for studying the Chebyshev approximation and the rocket problem.

Definition 4.2.1. Let $(E, ||\cdot||)$ be a normed vector space over $\mathbb{K}$ and $F \subset E$ a subspace. The orthogonal complement of F is defined by

$$F^{\perp} = \{u^* \in E^* : u^*(u) = 0 \text{ for all } u \in F\}.$$

Lemma 4.2.2. *Let F be a subspace of a normed vector space E over $\mathbb{K}$. Then, for all $u_0 \in E$ such that $d(u_0, F) := \inf_{u \in F} ||u - u_0|| > 0$, there exists a continuous linear form $v^* : E \to \mathbb{K}$ such that $v^* \in F^{\perp}$, $||v^*|| = 1$ and $v^*(u_0) = d(u_0, F)$.*

Proof. Since $d(u_0, F) > 0$, $u_0 \notin F$ and consider $F_0 = F \oplus [u_0]$. Then, each $u \in F_0$ has a unique representation $u = \lambda u_0 + v$ with $\lambda \in \mathbb{K}$ and $v \in F$. Let $u^* : F_0 \longrightarrow \mathbb{K}$ defined by

$$u^*(v + \lambda u_0) = \lambda d(u_0, F)$$

be a well-defined linear form on F_0. Besides, $u^*(u) = 0$ if $u \in F$, otherwise $\lambda \neq 0$ and it follows that

$$|u^*(u)| = |u^*(\lambda u_0 + v)| = |\lambda| d(u_0, F) \leq |\lambda| \left|\left| u_0 - \left(-\frac{1}{\lambda} v \right) \right|\right| = ||\lambda u_0 + v|| = ||u||.$$

Therefore, u^* is continuous with $||u^*|| \leq 1$. By the Hahn Banach theorem, there exists a continuous linear form $v^* : E \to \mathbb{K}$ that extends u^* and $||v^*|| = ||u^*|| \leq 1$. For all $\epsilon > 0$ there exists $v \in F$ so that

$$||u_0 - v|| \leq d(u_0, F) + \epsilon.$$

From $v^*(u_0 - v) = u^*(u_0 - v) = d(u_0, F)$, it follows that

$$\frac{v^*(u_0 - v)}{||u_0 - v||} \geq \frac{d(u_0, F)}{d(u_0, F) + \epsilon} \to 1 \text{ as } \epsilon \to 0.$$

Hence, $||v^*|| = 1$. ☐

Definition 4.2.3. Let E be a real normed vector space, $F \subset E$ a subspace and $u_0 \in E$ a fixed element.

(i) The *primal problem* consists in finding $v \in F$ so that

$$\alpha := \inf_{u \in F} ||u_0 - u|| = ||u_0 - v||. \tag{4.2.1}$$

(ii) The *dual problem* consists in finding $v^* \in F^{\perp}$ with $||v^*|| \leq 1$ so that

$$\beta := \sup_{\substack{u^* \in F^{\perp} \\ ||u^*|| \leq 1}} u^*(u_0) = v^*(u_0). \tag{4.2.2}$$

Proposition 4.2.4. *If* $\dim F < \infty$*, then the primal problem* (4.2.1) *has always a solution.*

Proof. Since $0 \in F$, $\alpha := \inf_{u \in F} ||u_0 - u|| \leq ||u_0||$. Let us define the set

$$F_0 = \{u \in F : ||u - u_0|| \leq ||u_0||\} = \overline{B_F(u_0, ||u_0||)} \subset F.$$

Since F is a finite-dimensional subspace, F_0 is a compact set and, by the Weierstrass theorem, there exists $v \in F_0$ so that

$$||v - u_0|| = \inf_{u \in F_0} ||u - u_0|| = \inf_{u \in F} ||u - u_0|| = \alpha. \qquad \square$$

Theorem 4.2.5. *Let E be a real normed vector space and $F \subset E$ a subspace. Given $u_0 \in E$, the following conditions hold:*

(i) $\alpha = \beta$.

(ii) The dual problem (4.2.2) *has always a solution.*

(iii) Let u^* be a solution of the dual problem (4.2.2)*. Then, $u \in F$ is a solution of the primal problem* (4.2.1) *if and only if*

$$u^*(u_0 - u) = ||u - u_0||.$$

Proof. (i) For all $\epsilon > 0$ there exists $u \in F$ so that

$$||u_0 - u|| \leq \alpha + \epsilon.$$

For all $u^* \in F^\perp$, with $||u^*|| \leq 1$, we have that

$$u^*(u_0) = u^*(u_0 - u) \leq ||u^*|| \, ||u - u_0|| \leq \alpha + \epsilon.$$

Therefore, $\beta \leq \alpha + \epsilon$ for all $\epsilon > 0$ and, thus, $\beta \leq \alpha$.
If $\alpha > 0$, by Lemma 4.2.2, there exists $u^* : E \to \mathbb{R}$ with $u^* \in F^\perp$, $||u^*|| = 1$ and such that $u^*(u_0) = \alpha$. Hence, $\beta \geq \alpha$. In case $\alpha = 0$, consider $u^* = 0$, which satisfies $u^*(u_0) = 0$ and, thus, $\beta \geq 0 = \alpha$. In both cases, we conclude that $\alpha = \beta$.

(ii) In (i) we have already shown the existence of $u^* \in F^\perp$, with $||u^*|| \leq 1$, so that $u^*(u_0) = \beta$.

(iii) Assume that there exists $u^* \in F^\perp$, with $||u^*|| \leq 1$, so that $u^*(u_0) = \alpha = \beta$. Then, $u \in F$ is a solution of problem (4.2.1) if and only if

$$||u_0 - u|| = \alpha = \beta = u^*(u_0) = u^*(u_0 - u). \qquad \square$$

Next, we take into consideration the primal problem and the dual one for a dual space E^*.

Definition 4.2.6. Let E be a real normed vector space, $F \subset E$ a subspace and $u_0^* \in E^*$ a given linear form.

(i) The *modified primal problem* consists in finding $v^* \in F^\perp$ so that

$$\alpha := \inf_{u^* \in F^\perp} ||u_0^* - u^*|| = ||u_0^* - v^*||. \tag{4.2.3}$$

(ii) The *associated dual problem* consists in finding $v \in F$ with $||v|| \leq 1$ so that

$$\beta := \sup_{\substack{u \in F \\ ||u|| \leq 1}} u_0^*(u) = u_0^*(v). \tag{4.2.4}$$

Theorem 4.2.7. *Let E be a real normed vector space and $F \subset E$ a subspace. Given $u_0^* \in E^*$, the following conditions hold:*

(i) $\alpha = \beta$.

(ii) The modified primal problem (4.2.3) *has always a solution.*

(iii) Let $u^* \in F^\perp$ be a solution of the modified primal problem (4.2.3). *Then, $u \in F$ with $||u|| \leq 1$ is a solution of the dual problem* (4.2.4) *if and only if*

$$(u_0^* - u^*)(u) = ||u_0^* - u^*||.$$

Proof. (i) For all $u^* \in F^\perp$,

$$||u_0^* - u^*|| = \sup_{||u|| \leq 1} |(u_0^* - u^*)(u)| \geq \sup_{\substack{||u|| \leq 1 \\ u \in F}} u_0^*(u) = \beta.$$

Hence, $\alpha \geq \beta$. Let $v^* : F \to \mathbb{R}$ be the restriction of $u_0^* : E \to \mathbb{R}$ to F. Then,

$$||v^*|| = \sup_{\substack{||u|| \leq 1 \\ u \in F}} |u_0^*(u)| = \sup_{\substack{||u|| \leq 1 \\ u \in F}} u_0^*(u) = \beta.$$

By the Hahn-Banach theorem, there exists a continuous linear form $v_0^* : E \to \mathbb{R}$ that extends v^* and $||v_0^*|| = ||v^*||$. If $w^* = u_0^* - v_0^*$, then for all $u \in F$ we have that $w^*(u) = u_0^*(u) - v_0^*(u) = v^*(u) - v^*(u) = 0$. That is, $w^* \in F^\perp$. Besides,

$$||w^* - u_0^*|| = ||v_0^*|| = ||v^*|| = \beta.$$

Hence, $\alpha \leq \beta$ and, thus, $\alpha = \beta$.

(ii) Notice that w^* is a solution of the modified primal problem (4.2.3).

(iii) Let $u^* \in F^\perp$ be a solution of the modified primal problem (4.2.3). Then, $u \in F$ with $||u|| \leq 1$ is a solution of the dual problem (4.2.4) if and only if

$$(u_0^* - u^*)(u) = u_0^*(u) = \beta = \alpha = ||u_0^* - u^*||.$$

□

Definition 4.2.8. Let $[a, b]$ be an interval, $-\infty < a < b < \infty$, and $c \in [a, b]$. We define

$$\delta_c : \mathcal{C}([a, b]) \to \mathbb{R}$$
$$u \mapsto \delta_c(u) := u(c)$$

a continuous linear form with $||\delta_c|| = 1$.

Lemma 4.2.9. *Let $u^* \in \mathcal{C}([a, b])^*$ with $||u^*|| \neq 0$. If there exists $u \in \mathcal{C}([a, b])$ so that $u^*(u) = ||u^*|| \, ||u||_\infty$ and $|u(x)|$ achieves its maximum at precisely N points $x_1, \ldots, x_N$ of $[a, b]$, then there exist $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$ such that*

$$u^* = \sum_{i=1}^N \alpha_i \delta_{x_i} \text{ and } \sum_{i=1}^N |\alpha_i| = ||u^*||.$$

Proof. Since $u^* \in \mathcal{C}([a,b])^*$, by the Riesz representation theorem, there exists $g \in BV[a,b]$ such that

$$u^*(v) = \int_a^b v(x)dg(x) \quad \text{for all } v \in \mathcal{C}([a,b]),$$

and $V(g) = ||u^*|| \neq 0$. Besides, we assume $a < x_i < b$ for all $1 \leq i \leq n$. Given $\epsilon > 0$ small enough, we define

$$J = [a,b] \setminus \biguplus_{i=1}^{N} (x_i - \epsilon, x_i + \epsilon) = [a, x_1 - \epsilon] \biguplus_{i=1}^{N-1} [x_i + \epsilon, x_{i+1} - \epsilon] \biguplus [x_N + \epsilon, b],$$

which is a finite union of closed intervals, and $V_J(g)$ will denote the sum of the total variations of g at each interval of the union J. If we denote $I_i = [x_i - \epsilon, x_i + \epsilon]$, note that

$$V_J(g) + \sum_{i=1}^{N} V_{I_i}(g) \leq V(g).$$

We first show that $V_J(g) = 0$. Indeed, we assume the opposite thesis, so $V_J(g) > 0$. Then,

$$\begin{aligned} u^*(u) = |u^*(u)| &= \left| \int_J u(x)dg(x) + \sum_{i=1}^{N} \int_{x_i-\epsilon}^{x_i+\epsilon} u(x)dg(x) \right| \\ &\leq V_J(g) \max_{x \in J} |u(x)| + \sum_{i=1}^{N} V_{I_i}(g) ||u||_\infty < \left(V_J(g) + \sum_{i=1}^{N} V_{I_i}(g) \right) ||u||_\infty \\ &\leq V(g) ||u||_\infty = ||u^*|| \, ||u||_\infty. \end{aligned}$$

Therefore, $u^*(u) < ||u^*|| \, ||u||_\infty$, which contradicts the hypothesis $u^*(u) = ||u^*|| \, ||u||_\infty$. Hence, $V_J(g) = 0$ and it follows that g is constant at each interval of the union J, so g has the following form:

$$g(x) = \begin{cases} \beta_0, & \text{if} \quad a \leq x < x_1, \\ \beta_1, & \text{if} \quad x_1 \leq x < x_2, \\ & \vdots \\ \beta_N, & \text{if} \quad x_N \leq x \leq b. \end{cases}$$

For all $v \in \mathcal{C}([a,b])$,

$$\begin{aligned} u^*(v) &= \int_a^b v(x)dg(x) := \lim_{||P|| \to 0} \sum_j v(\mu_j)[g(y_j) - g(y_{j-1})] = \sum_{i=1}^{N} v(x_i)(\beta_i - \beta_{i-1}) \\ &= \sum_{i=1}^{N} v(x_i)\alpha_i. \end{aligned}$$

Therefore, $u^* = \sum_{i=1}^{N} \alpha_i \delta_{x_i}$ and $||u^*|| \leq \sum_{i=1}^{N} |\alpha_i|$.
Let $w \in \mathcal{C}([a,b])$ be a continuous piecewise affine function such that $w(x_i) = \text{sign}(\alpha_i)$ for all $1 \leq i \leq N$. Besides, $||w||_\infty = 1$ because, in case $w = 0$, this would imply that $||u^*|| \leq \sum_{i=1}^{N} |\alpha_i| = 0$, which is impossible. Finally,

$$||u^*|| \geq |u^*(w)| = \left| \sum_{i=1}^{N} \alpha_i \delta_{x_i}(w) \right| = \left| \sum_{i=1}^{N} \alpha_i w(x_i) \right| = \sum_{i=1}^{N} |\alpha_i|.$$

Hence, $||u^*|| = \sum_{i=1}^{N} |\alpha_i|$. In the case that $|u(x)|$ achieves the maximum at $x = a$ or $x = b$, the proof is similar with the intervals $[a, a + \epsilon]$ and $[b - \epsilon, b]$, respectively. $\square$

4.2.1 The Chebyshev approximation

The approximation of continuous functions by polynomials is often useful in Analysis. Given a continuous function $u : [a, b] \to \mathbb{R}$, $-\infty < a < b < \infty$, we want to determine whether there exists a polynomial $p \in \mathcal{P}_N([a, b])$ such that

$$\alpha := \inf \left\{ ||u - q||_\infty : q \in \mathcal{P}_N([a, b]) \right\} = ||u - p||_\infty. \tag{4.2.5}$$

Even though this problem belongs to Analysis, it has been included in this section as an example of the dual theory developed previously.

Theorem 4.2.10. *Let $u \in \mathcal{C}([a, b])$, problem (4.2.5) has a solution $p \in \mathcal{P}_N([a, b])$. Besides, $|u(x) - p(x)|$ achieves its maximum at at least $(N + 2)$ points of $[a, b]$.*

Proof. Since $\mathcal{P}_N([a, b])$ is a finite-dimensional subspace of $\mathcal{C}([a, b])$, by Proposition 4.2.4, there exists $p \in \mathcal{P}_N([a, b])$ that is a solution of problem (4.2.5).
In order to prove the second part of the statement, we assume that $u \notin \mathcal{P}_N([a, b])$, otherwise the proof is immediate. Therefore, $\alpha = ||u - p||_\infty > 0$ and, by Lemma 4.2.2 and Theorem 4.2.5, there exists $u^* \in \mathcal{P}([a, b])^\perp$ a solution of the dual problem with $||u^*|| = 1$ and such that

$$u^*(u - p) = ||u - p||.$$

Now, suppose that $|u(x) - p(x)|$ achieves its maximum at $x_1, \dots, x_M \in [a, b]$, with $1 \le M < N + 2$. By Lemma 4.2.9, there exist $\alpha_1, \dots, \alpha_M \in \mathbb{R}$ such that

$$u^* = \sum_{i=1}^{M} \alpha_i \delta_{x_i} \text{ and } \sum_{i=1}^{M} |\alpha_i| = ||u^*|| = 1.$$

Note that we can assume that $\alpha_M \neq 0$. Since $M - 1 \le N$, we choose a polynomial $q(x) \in P_N([a, b])$ such that $q(x_i) = 0$, for all $1 \le i \le M - 1$, and $q(x_M) \neq 0$. Then,

$$u^*(q) = \sum_{i=1}^{N} \alpha_i q(x_i) = \alpha_M q(x_M) \neq 0,$$

which contradicts the fact that $u^* \in \mathcal{P}_N([a, b])^\perp$. $\square$

4.2.2 Optimal control of rockets

We want to study the motion of a rocket ascent that reaches an altitude $h > 0$ with the minimum fuel expenditure. If $x = x(t)$ denotes its vertical position, then, by Newton's Second Law,

$$mx''(t) = f(t) - mg \quad 0 < t < T, \tag{4.2.6}$$
$$x(0) = x'(0) = 0 \text{ and } x(T) = h,$$

where m is the mass of the rocket, $f(t)$ the force provided by the engines and T is the time required for achieving the altitude $h > 0$. The variation of mass due to fuel burning is neglected and we will work in physical units, so $m = g = 1$.
The fuel expenditure during $[0, T]$ is given by

$$\int_0^T |f(t)| dt,$$

which depends on the time T. Then, for each time $T > 0$, $\alpha(T)$ denotes the minimum fuel expenditure required for achieving the altitude $h > 0$ in a time T. In terms of minimum norm problems, we have that

$$\alpha(T) := \inf_f \int_0^T |f(t)|dt, \tag{4.2.7}$$

where the infimum is taken over all integrable functions $f : [0,T] \to \mathbb{R}$. Besides, once $\alpha(T)$ is determined, we can adjust $T > 0$ so as to minimize $\alpha(T)$.
We first integrate equation (4.2.6),

$$x'(t) = \int_0^t f(s)ds - t,$$

$$x(t) = \int_0^t \int_0^u f(s)dsdu - \frac{t^2}{2} = \int_0^t f(s) \int_s^t duds - \frac{t^2}{2} = \int_0^t (t-s)f(s)ds - \frac{t^2}{2}.$$

Since $x(T) = h$, it follows the condition

$$h = \int_0^T (T-s)f(s)ds - \frac{T^2}{2}. \tag{4.2.8}$$

To sum up, we are looking for an integrable function $f : [0,T] \to \mathbb{R}$ that is the solution of problem (4.2.7), satisfies (4.2.8) and a time $T > 0$ that minimizes $\alpha(T)$.
We next show that we could have restricted our problem to continuous functions as we take into consideration that $\mathcal{C}([0,T])$ is dense in $L^1([0,T])$. Indeed, if $f \in L^1([0,T])$ is a solution of the problem, for all $\epsilon > 0$ there exists $g \in \mathcal{C}([0,T])$ so that $||f-g||_1 < \frac{\epsilon}{T}$. Then,

$$\left| \int_0^T (T-t)f(t)dt - \int_0^T (T-t)g(t)dt \right| \leq \int_0^T (T-t)|f(t)-g(t)|dt \leq T||f-g||_1 < \epsilon, \text{ and}$$

$$\left| \int_0^T |f(t)|dt - \int_0^T |g(t)|dt \right| \leq \int_0^T |f(t)-g(t)|dt < \frac{\epsilon}{T}.$$

Hence,

$$(h-\epsilon) + \frac{T^2}{2} < \int_0^T (T-t)g(t)dt < (h+\epsilon) + \frac{T^2}{2}, \text{ and}$$

$$\alpha(T) \leq \int_0^T |g(t)|dt < \alpha(T) + \frac{\epsilon}{T}.$$

Therefore, the problem could have formulated in terms of continuous functions instead of integrable ones by taking ϵ smaller than the dimensions of the rocket and so that the extra fuel consumed would be insignificant. Nevertheless, both cases omit the possibility of applying an impulse δ_t at a precise instant t because condition (4.2.8) would not be satisfied. As we will show, impulses play an important role in this optimization problem. For this reason, we will proceed with a more general approach.

(i) For a given altitude $h > 0$ and a final time $T > 0$, we are looking for $g \in NBV([0,T])$, such that

$$\int_0^T (T-t)dg(t) = h + \frac{T^2}{2},$$

and it minimizes $V(g) := \int_0^T |dg(t)| = \alpha(T)$.

(ii) To find $T > 0$ that minimizes $\alpha(T)$.

By Corollary 3.1.7, these problems are equivalent to the following ones.

(i) For a given altitude $h > 0$ and a final time $T > 0$, we are looking for a functional $v^* \in \mathcal{C}([0,T])^*$ such that

$$\alpha(T) = \inf_{\substack{u^* \in \mathcal{C}([0,T])^* \\ u^*(T-t) = h + \frac{T^2}{2}}} ||u^*|| = ||v^*|| \text{ , and} \tag{4.2.9}$$

$$h = v^*(T-t) - \frac{T^2}{2}. \tag{4.2.10}$$

(ii) To find $T > 0$ that minimizes $\alpha(T)$.

We next show that (4.2.9) and (4.2.10) generalize (4.2.7) and (4.2.8), respectively.

Proposition 4.2.11. *Let $g : [0,T] \to \mathbb{R}$ be a continuous function and $v^* \in \mathcal{C}([0,T])^*$ the functional defined by*

$$v^*(u) = \int_0^T u(t)g(t)dt \quad \text{for all } u \in \mathcal{C}([0,T]).$$

Then,

$$||v^*|| = \int_0^T |g(t)|dt.$$

Proof. Let $\rho(t) = \int_0^t g(s)ds$, that is, $\rho'(t) = g(t)$ for all $t \in [0,T]$. By Proposition 1.3.9, we have that

$$v^*(u) = \int_0^T u(t)g(t)dt = \int_0^T u(t)\rho'(t)dt = \int_0^T u(t)d\rho(t) \quad \text{for all } u \in \mathcal{C}([0,T]).$$

Let $0 = t_0 < t_1 < \ldots < t_n = T$ be a partition of $[0,T]$. We define

$$\Delta := \sum_{i=1}^n |\rho(t_j) - \rho(t_{j-1})| \leq \sum_{i=1}^n \int_{t_{j-1}}^{t_j} |g(t)|dt = \int_0^T |g(t)|dt.$$

Hence, $\Delta \leq V(\rho) \leq \int_0^T |g(t)|dt$. Since $V(\rho) \leq T||g||_\infty < \infty$, $\rho \in NBV([0,T])$ and $V(\rho) = ||v^*||$. Besides, by the mean value theorem,

$$\Delta = \sum_{i=1}^n |g(s_j)|(t_j - t_{j-1}) \quad \text{with } s_j \in [t_j, t_{j-1}] \text{ for all } 1 \leq j \leq n.$$

As $n \to \infty$,

$$\Delta \to \int_0^T |g(t)|dt.$$

Since $\Delta \leq V(\rho)$,

$$||v^*|| = V(\rho) = \int_0^T |g(t)|dt.$$

□

It is clear that condition (4.2.10) generalizes (4.2.8).

Theorem 4.2.12. *(a) The solution of the problem (i) is* $u^* = T\delta_0$.

(b) The solution of the problem (ii) is $T = \sqrt{2h}$.

Proof. (a) Let $F = \{\lambda(T-t) : \lambda \in \mathbb{R}\}$ be a subspace of $\mathcal{C}([0,T])$. By Theorem 2.3.6, there exists $u_0^* \in \mathcal{C}([0,T])^*$ such that

$$u_0^*(T-t) = h + \frac{T^2}{2}.$$

Then,

$$\alpha(T) := \inf_{\substack{u^* \in \mathcal{C}([a,b])^* \\ h = u^*(T-t) - \frac{T^2}{2}}} ||u^*|| = \inf_{u^* - u_0^* \in F^\perp} ||(u_0^* - u^*) - u_0^*||.$$

By Theorem 4.2.7, there exists $u^* - u_0^* \in F^\perp$ such that $\alpha(T) = ||u^*|| = ||(u_0^* - u^*) - u_0^*||$. Since $||T-t||_\infty = T$ and u_0^* is linear,

$$\alpha(T) = \sup_{\substack{u \in F \\ ||u||_\infty \le 1}} u_0^*(u) = \sup_{||\lambda(T-t)||_\infty \le 1} \lambda\left(h + \frac{T^2}{2}\right) = \sup_{|\lambda| \le \frac{1}{T}} \lambda\left(h + \frac{T^2}{2}\right) = \frac{1}{T}\left(h + \frac{T^2}{2}\right).$$

Since $u_0^* - u^* \in F^\perp$,

$$||u^*|| = \alpha(T) = u_0^*\big(T^{-1}(T-t)\big) = (u_0^* - u^*)\big(T^{-1}(T-t)\big) + u^*\big(T^{-1}(T-t)\big) = u^*\big(T^{-1}(T-t)\big).$$

Namely,

$$u^*(T^{-1}(T-t)) = ||u^*||\ ||T^{-1}(T-t)||_\infty.$$

Besides, $u(t) = \frac{1}{T}(T-t)$ achieves its maximum at precisely $t = 0$. By Lemma 4.2.9, there exists $\gamma \in \mathbb{R}$ such that $u^* = \gamma\delta_0$ with $|\gamma| = ||u^*|| = \alpha(T)$. Therefore, $u^* = \pm\alpha(T)\delta_0$. Since $u^*(T-t) = h + \frac{T^2}{2} > 0$ and $\delta_0(T-t) = T > 0$, it follows that $u^* = \alpha(T)\delta_0$.

(b) From (a) we have that $\alpha(T) = ||u^*|| = \frac{1}{T}(h + \frac{T^2}{2})$. Then,

$$\alpha'(T) = \frac{-h}{T^2} + \frac{1}{2} = 0 \ \Rightarrow T = \sqrt{2h}.$$

Since $\alpha''(T) > 0$, $T = \sqrt{2h}$ is a minimum. □

Remark 4.2.13. According to Example 1.3.17 and Corollary 3.1.7, we have that

$$u^*(f) = \sqrt{2h} f(0) = \sqrt{2h}\int_0^T f\ dH_0 = \sqrt{2h}\int_0^T f(t) d\delta_0(t).$$

The last equality refers to the Lebesgue-Stieltjes integral with respect to the Dirac delta distribution, which allows us to conclude that the physical solution to the rocket problem is an impulse at the initial instant.

Remark 4.2.14. If we want to express the solutions in (SI) units, we have $u^* = m\sqrt{2hg}\delta_0$, $T = \sqrt{\frac{2h}{g}}$ and the minimum fuel expenditure is $\alpha = m\sqrt{2hg}$.

We have concluded that the most efficient program is an impulse at the initial instant, so the generalization of the rocket problem was necessary. We have made many simplifications in order to adapt the problem to our theory. This problem should be understood as a model to the physical situation and should be contrasted with experiments.

Conclusions

The development of this thesis has provided the assimilation of many cross-curricular concepts and results. The present work could not only be regarded as a continuation of the course *Anàlisi real i functional*, taught at University of Barcelona, but also as an implementation of the skills and knowledge acquired during the Major in Mathematics.

The restriction to Banach spaces has not supposed any limitation for formulating powerful theorems that have lead us to relevant applications such as: the nowhere differentiable continuous functions are dense among the continuous ones, the Riesz representation theorem, the existence of functions whose Fourier series respectively diverge in $(\mathcal{C}([a,b]), ||\cdot||_\infty)$ and in $(L^1([a,b]), ||\cdot||_1)$, the existence of a continuous function whose Lagrange interpolating polynomial does not converge uniformly to the function, etc. Most of them have been fully studied without difficulty, though the rocket problem, in Chapter 4, has required a less formal approach, as Physics does in general, and the solution should be contrasted with experimental evidence.

Another important achievement has been to develop the ability of being critic with the proofs taken from the references and complementing them. One interesting aspect that could have been added to the references is to provide examples to some applications of Chapter 3. For instance, functions whose Fourier series diverge or the Lagrange interpolating polynomial, or an example of a nowhere differentiable continuous function, etc.

Further work should focus on the formulation of these fundamental theorems in Fréchet spaces and examine their applications, for example, the existence of a solution in a partial differential equation. It would also be interesting to seek applications in areas such as Algebra, Economics, Geometry, Probability, etc.

Bibliography

[1] Cerdà, J.:*Linear Functional Analysis*, Graduate Studies in Mathematics, vol. 116, American Mathematical Society, Providence, Rhode Island, 2010.

[2] Cerdà, J.: *Introducció a l'anàlisi funcional*, Edicions de la Universitat de Barcelona, Barcelona, 2004.

[3] Ciarlet, P.G.: *Linear and Nonlinear Functional Analysis with Applications*, Society for Industrial and Applied Mathematics, Philadelphia, 2013.

[4] Eidelman, Y.; Milman, V.; Tsolomitis, A. : *Functional Analysis: An Introduction*, Graduate Studies in Mathematics, vol. 66, American Mathematical Society, Providence, Rhode Island, 2004.

[5] Erdös P.; Turán, P.: *An extremal problem in the theory of interpolation*, Acta Math, Academy of Science of Hungary, Volume 12, Issue 1, March 1964, pp. 221-234.

[6] Miana, P.J.: *Curso de Análisis funcional*, manuals, 2006.
`http://www.unizar.es/matematicas/personales/pjmiana/AnalisisFuncional/Curso%20AF-Miana.pdf`

[7] Luenberger, D.: *Optimization by vector space methods*, John Wiley & Sons, New York, 1969.

[8] Pietsch, A.: *History of Banach Spaces and Linear Operators*, Birkhäuser, Boston, 2007.

[9] Rabindranath, S.: *A First Course in Functional Analysis*, Anthem Press, London, 2013.

[10] Richardson, L.F.: *Advanced calculus: An Introduction to Linear Analysis*, John Wiley & Sons, 2008, Hoboken, New Jersey.

[11] Rudin, W.: *Principios de Análisis Matemático*, McGraw-Hill, 3rd edition, Mexico D.F., 1980.

[12] Stieltjes, T.J. : *Recherches sur les fractions continues*, Annales de la Faculté des sciences de Toulouse 1^{er} série, volume 8, section 4, 1984, pp. 1-122.
`http://archive.numdam.org/ARCHIVE/AFST/AFST_1894_1_8_4/AFST_1894_1_8_4_J1_0/AFST_1894_1_8_4_J1_0.pdf`

[13] Shohat, J.A.; Tamarkin, J.D.: *The Problem of Moments*, Mathematical Surveys and Monographs, American Mathematical Society, fourth edition, Providence, Rhode Island, 1970.

[14] Zeidler, E.: *Applied Functional Analysis: Applications to Mathematical Physics*, Applied Mathematical Sciences, vol. 109, Springer-Verlag, New York, 1995.

Index

$\mathcal{G}_\delta$-dense set, 18

Baire's theorem, 15
Banach isomorphism theorem, 17
Banach space, 3
  product space, 4
  quotient space, 5
  separable, 32
Banach-Steinhaus theorem, 19

Chebyshev approximation, 50
closed graph theorem, 17
convex functional, 20

Dirichlet kernels, 34
Dual space, 4

Fejer's kernel, 36
Fourier series, 33
  divergence
    $\mathcal{C}([a,b])$, 35
    $L^1([a,b])$, 38
function
  nowhere differentiable, 27
  of bounded variation, 8
    normalised, 30
  Riemann-Stieltjes integrable, 10
  total variation of a function, 8

Hahn-Banach theorem
  analytic
    complex, 21
    normed spaces, 22
    real, 20
  geometric, 24

Lagrange interpolating polynomial, 39

Minkowski functional, 23
moment problem, 43
  finite, 43
  infinite, 45

open mapping theorem, 16
operator
  linear, 4
    bounded, 4
    continuous, 4
    open, 16
optimal control of rockets, 50
orthogonal complement, 46

Riemann-Stieltjes
  integral, 10
  sum, 10
Riesz representation theorem, 29

seminorm, 21
separated sets, 24
  strictly, 24

uniform boundedness principle, 18